Universidad Nacional Autónoma de México

Facultad de Estudios Superiores Acatlán

Actuaría

Análisis comparativo de pruebas de simetría para cópulas bivariadas

ΤΕSΙS

José Manuel Pedraza Ramírez

Asesor: Dr. Arturo Erdely Ruiz

Noviembre, 2012



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. La única manera de lidiar con este mundo sin libertad es volverte tan absolutamente libre que tu mera existencia sea un acto de rebelión.

Albert Camus.

Algún día, me aferraré con fuerza al mundo. Ahora tengo un dedo apoyado en él, es un principio.

Ray Bradbury, Fahrenheit 451.

Agradecimientos:

A mis padres Bernabe y Guadalupe, fuente constante de valores, amor e inspiración. Gracias a ustedes, hoy soy lo que soy.

A esos grandes amigos quienes me apoyaron en esta etapa tan importante de mi vida. En particular a Emmanuel,

sin tu entrega y dedicación no lo habría logrado.

Al excelente profesor Arturo Erdely, por su disposición y paciencia para dar revisión y seguimiento a este trabajo.

A mi querida Universidad...

IV

Índice general

Introducción

1.	Cóp	ulas: Definición y Propiedades	1
	1.1.	Preliminares	3
	1.2.	Cópulas	7
	1.3.	Teorema de Sklar	12
	1.4.	Cópulas y Variables Aleatorias	15
	1.5.	Cópulas Arquimedianas	21
	1.6.	Dependencia	24
		Medidas de Concordancia	24
		Medidas de Dependencia	28
	1.7.	Cópula Empírica	30
	1.8.	Simulación	32
2.	Sim	etría	37
	2.1.	Conceptos Generales de Simetría	37
	2.2.	Medidas de No-Intercambiabilidad	40
	2.3.	Construcción de cópulas asimétricas	49
	2.4.	Pruebas de Simetría para cópulas bivariadas	52
		Prueba de Tri-simetría (Erdely y González-Barrios)	53
		Prueba de Simetría (Genest, Nešlehová y Quessy)	55
3.	Aná	lisis Comparativo	61
	3.1.	Potencia de las Pruebas	62
	3.2.	Interpretación de Resultados	75
Co	onclu	siones	79

 \mathbf{VII}

ÍNDICE GENERAL

A. Código en R	83
A.1. Instalación del Software R	83
A.2. Algoritmos de Simulación	84
A.3. Pruebas de Simetría	87
B. Diagramas de Dispersión	91
Bibliografía	97

VI

Introducción

La construcción de modelos cuyo objetivo es describir el grado de incertidumbre en un sistema formado por varios componentes es un problema imprescindible en la estadística. En los últimos años, esto ha recibido un creciente interés, en parte por el reciente descubrimiento de la noción del concepto de cópula, por los investigadores que trabajan en las ciencias aplicadas, como finanzas, ciencia actuarial, hidrología, teoría del riesgo, entre otras.

Cuando se considera un sistema formado por varios componentes, usualmente es más sencillo modelar el comportamiento de cada componente de manera individual. En la mayoría de los casos es una tarea más difícil tomar en cuenta la dependencia entre ellos, es por esta razón que muchos practicantes formulan el supuesto poco realista de independencia. Gracias al teorema de Sklar, las cópulas son los objetos matemáticos que permiten dividir la información relacionada con los componentes individuales del sistema de la dependencia entre ellos.

Específicamente, la ventaja principal proporcionada por el enfoque con cópulas consiste en el hecho de que la selección de un modelo apropiado, el cual considere la dependencia entre (X, Y), puede proceder de forma independiente a la elección de las distribuciones marginales. En otros modelos bivariados, la elección de una distribución conjunta que capture la dependencia entre (X, Y), fuerza la elección de las distribuciones marginales. Por ejemplo, si suponemos que la dependencia entre (X, Y) está descrita por una distribución normal con una matriz de correlaciones dada, estamos suponiendo implícitamente que las marginales son distribuciones normal univariadas.

Actualmente no existe un procedimiento estándar para escoger una cópula que mejor se ajuste a un conjunto de datos. Usualmente familias conocidas de cópulas (Clayton, Gumbel, Normal, Plackett, etc.) son consideradas y probadas con el fin de escoger un modelo que se adapte mejor. Regularmente la cópula es seleccionada de la familia Arquimediana, familia bien conocida de cópulas de fácil construcción y simulación.

Gran parte de las cópulas conocidas, entre ellas las Arquimedianas, cumplen la propiedad de ser simétricas, por lo que la simetría es un aspecto fundamental a considerar al momento de seleccionar el modelo más adecuado, surgiendo así la necesidad de validar el supuesto de simetría de cópulas. Con la finalidad de cubrir esta necesidad, aparecen dos pruebas de simetría para cópulas bivariadas: la prueba de Tri-simetría de Erdely y González-Barrios (2010) y la prueba de simetría de Genest, Nešlehová y Quessy (2011). Entonces, ahora surge la interrogante ¿Qué prueba es mejor para validar simetría? El responder a esta pregunta es el objetivo primordial de este trabajo.

Se dedica el primer capítulo a lo siguiente: definir de manera formal a las cópulas; vía el teorema de Sklar, se exhibe el vínculo que existe entre las cópulas y las funciones de distribución conjuntas de variables aleatorias continuas; se estudian varias familias paramétricas, entre ellas las cópulas Arquimedianas; se revisa el concepto de dependencia de variables aleatorias; se define a la cópula empírica como una función de la muestra y una aproximación a la cópula subyacente; por último, se enuncian algunos algoritmos para simular vectores aleatorios.

En el segundo capítulo se revisan algunas definiciones de simetría para variables aleatorias incluyendo el concepto de intercambiabilidad, se analiza la relación que éste tiene con las cópulas simétricas y se mencionan las características deseables para una medida de no-intercambiabilidad o de asimetría. Por último, se exponen los detalles de la prueba de Tri-simetría de Erdely y la prueba de simetría de Genest.

En el tercer capítulo, se comparan la potencias de ambas pruebas mediante la simulación de datos provenientes de algunos modelos conocidos de cópulas simétricas y de algunas versiones asimétricas de éstas, para así determinar bajo qué escenarios cada una de ellas da mejores resultados y determinar cuál es la mejor opción.

Por último, se presentan dos apéndices: en el primero se muestran las rutinas computacionales empleadas; en el segundo, los diagramas de dispersión de las cópulas usadas en el tercer capítulo.

Capítulo 1

Cópulas: Definición y Propiedades básicas

La palabra *cópula* es un sustantivo proveniente del latín que significa enlace, ligamiento o atadura y en gramática es la parte que conecta al sujeto y el predicado. Fue empleada por primera vez por Sklar (1959) en el teorema (que actualmente lleva su nombre) describiendo funciones que "unen" distribuciones univariadas para formar distribuciones multivariadas.

Se podría decir que la historia de las cópulas comienza con Fréchet (1951). Él estudió el siguiente problema: dadas las funciones de distribución F_1 y F_2 de dos variables aleatorias X_1 y X_2 definidas en el mismo espacio de probabilidad $(\Omega, \mathscr{F}, \mathbb{P})$ ¿Qué se puede decir acerca del conjunto $\Gamma(F_1, F_2)$ de funciones de distribución bivariadas cuyas marginales son F_1 y F_2 ? Es inmediato ver que el conjunto $\Gamma(F_1, F_2)$, llamado clases de Fréchet de F_1 y F_2 , no es vacío ya que si X_1 y X_2 son independientes, entonces la función de distribución $(x_1, x_2) \mapsto F(x_1, x_2) = F_1(x_1) F_2(x_2)$ siempre pertenece a $\Gamma(F_1, F_2)$. Pero, no era claro cómo eran los otros elementos de $\Gamma(F_1, F_2)$.

Después, en 1959, Sklar obtuvo el más profundo resultado respecto a esto, introduciendo el concepto, y el nombre de *cópula*, probando así el teorema que actualmente lleva su nombre.

Sin embargo, anterior al trabajo de Sklar, se pueden encontrar muchos de los resultados básicos acerca de las cópulas en el trabajo de Hoeffding (1940), en el cual define "distribuciones estandarizadas", donde la única di-



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. ferencia con las cópulas es el espacio en el que tienen cabida (las cópulas son funciones en $[0, 1] \times [0, 1]$ mientras que las de Hoeffding son funciones en $\left[-\frac{1}{2}, \frac{1}{2}\right] \times \left[-\frac{1}{2}, \frac{1}{2}\right]$).

Hoeffding también obtuvo las mejores cotas posibles para las funciones que propuso, caracterizó las distribuciones correspondientes a estas cotas y estudió medidas de dependencia que son invariantes ante transformaciones estrictamente crecientes. Sin conocer el trabajo de Hoeffding, Fréchet (1951) obtuvo muchos de los mismos resultados tales como "Cotas de Fréchet" y "Clases de Fréchet". En reconocimiento de la compartida responsabilidad de estas importantes ideas, nos referiremos a ellas como "Cotas de Fréchet-Hoeffding" y "Clases de Fréchet-Hoeffding."

Durante aproximadamente 15 años, todos los resultados de cópulas fueron en la teoría de espacios métricos probabilísticos. Sin embargo, actualmente juegan un papel muy importante en la estadística. Así, afirma Fisher (1997), "Las cópulas son de interés para los estadísticos por dos principales razones: Primero, como una forma de estudiar medidas de dependencia libres de escala; y segundo, como un punto de partida para construir familias de distribuciones bivariadas, algunas veces con el propósito de simular."

No fue sino hasta la década de los noventas que el concepto de cópula se volvió más popular. En 1990 Dall'Aglio organizó la primera conferencia dedicada a cópulas llamada "Distribuciones de Probabilidad con marginales dadas", seguida de algunas otras en Seattle (1993), en Praga (1996), en Barcelona (2000), en Quebéc (2004) en Tartu (2007) y en São Paulo (2010); así como la aparición de dos libros: Joe (1997) y Nelsen (2006). Pero, según Durante y Sempi (2010), la principal razón del creciente interés se encuentra en el descubrimiento de la noción de las cópulas en varios campos de aplicación, como finanzas e hidrología. Como Schweizer (2007) escribió:

La "época de i.i.d." se acabó: y cuando la dependencia se toma en serio, las cópulas naturalmente entran en juego. Sólo le resta a la comunidad estadística en su totalidad reconocer este hecho. Y cuando cada texto estadístico contenga una sección o capítulo de cópulas, el tema se convertirá de la época.

Sin embargo varias críticas han surgido recientemente acerca de las cópulas y sus aplicaciones; y varias personas han empezado a hablar acerca de la

"moda de cópulas" (Mikosch (2006) y Embrechts (2009)). Así Durante y Sempi (2010) en sus palabras:

Desde nuestro punto de vista, esas críticas fueron una reacción muy natural debido a la amplia difusión de la aplicación de las cópulas, no siempre en una forma bien motivada. Debe decirse que muchas personas han malinterpretado las cópulas como la solución a "todos los problemas de la dependencia estocástica". ¡Éste definitivamente no es el caso! Las cópulas son una herramienta indispensable para entender muchos problemas acerca de la dependencia estocástica, pero no son la "panacea" para todos los modelos estocásticos.

Sólo queda agregar, como afirman Durante y Sempi (2010) y Nelsen (2006), que a pesar de la amplia gama de interés que existe acerca de las cópulas el concepto sigue aún en su "infancia". Hay muchos problemas abiertos y mucho trabajo por hacer.

Este primer capítulo tiene como objetivos: definir de manera formal a las cópulas como funciones reales que van del cuadrado unitario al intervalo [0, 1], las cuales cumplen con características específicas; vía el teorema de Sklar se exhibe el vínculo que existe entre las cópulas y las funciones de distribución conjuntas de variables aleatorias continuas en teoría de la probabilidad; estudiar familias paramétricas denominadas cópulas Arquimedianas; el concepto de dependencia; la cópula empírica y algoritmos para simular vectores aleatorios.

1.1. Preliminares

Antes de enunciar la definición formal de cópulas, introduciremos la siguiente notación y algunos conceptos básicos de acuerdo con Nelsen (2006) para así extender la idea de una función *no drecreciente* al caso bivariado.

Consideremos al conjunto de los números reales $\mathbf{R} = (-\infty, \infty)$, $\mathbf{\bar{R}}$ denotará a los números reales extendidos $\mathbf{\bar{R}} = [-\infty, \infty] = \mathbf{R} \cup \{-\infty, \infty\}$ y $\mathbf{\bar{R}}^2$ al plano extendido $\mathbf{\bar{R}} \times \mathbf{\bar{R}}$.

Definición 1.1.1. Sean $[x_1, x_2]$, $[y_1, y_2]$ en $\overline{\mathbf{R}}^2$ con $x_1 \leq x_2$, $y_1 \leq y_2$. Un rectángulo en $\overline{\mathbf{R}}^2$ es el producto cartesiano de dos intervalos cerrados $B = [x_1, x_2] \times [y_1, y_2]$. Los vértices del rectángulo B son los puntos (x_1, y_1) , $(x_1, y_2), (x_2, y_1) y (x_2, y_2)$. El cuadrado unitario \mathbf{I}^2 es el producto $\mathbf{I} \times \mathbf{I}$ donde $\mathbf{I} = [0, 1]$.

Definición 1.1.2. Una función real de doble entrada H es una función cuyo dominio, DomH, es un subconjunto de $\overline{\mathbf{R}}^2$ y rango, RanH, es un subconjunto de \mathbf{R} .

Definición 1.1.3. Sean S_1 y S_2 dos subconjuntos no vacíos de $\overline{\mathbf{R}}$, y sea H una función real de doble entrada tal que $DomH = S_1 \times S_2$. Sea $B = [x_1, x_2] \times [y_1, y_2]$ un rectángulo donde todos sus vértices están en DomH. Entonces el H-volumen de B está dado por:

$$V_H(B) = H(x_2, y_2) - H(x_2, y_1) - H(x_1, y_2) + H(x_1, y_1)$$
(1.1)



Figura 1.1.1: Rectángulo donde se calcula el H-volumen

Definición 1.1.4. Una función real de doble entrada H se dice que es 2creciente si $V_H(B) \ge 0$ para todos los rectángulos B cuyos vértices están en DomH.

Hay que notar que la característica "2-creciente" es un concepto diferente a "no decreciente en cada argumento" por lo que no deben confundirse. De hecho como veremos en los siguientes dos ejemplos:

> 2-creciente \Rightarrow no decreciente en cada argumento, y no decreciente en cada argumento \Rightarrow 2-creciente.

Ejemplo 1.1.5. Sea H una función definida en I^2 como $H(x, y) = \max(x, y)$. Sean $a_1 \ y \ a_2$ en I fijos, entonces para todo $x_1 \le x_2, \ y_1 \le y_2$

$$H(x_1, a_2) = \max(x_1, a_2) \\\leq \max(x_2, a_2) \\= H(x_2, a_2) \\H(a_1, y_1) = \max(a_1, y_1) \\\leq \max(a_1, y_2) \\= H(a_1, y_2)$$

Por lo que H(x, y) es no decreciente en cada argumento, sin embargo H no cumple la propiedad 2-creciente:

$$V_H(\mathbf{I}^2) = H(1,1) - H(1,0) - H(0,1) + H(0,0)$$

= máx(1,1) - máx(1,0) - máx(0,1) + máx(0,0)
= 1 - 1 - 1 + 0
= -1
 ≤ 0

Ejemplo 1.1.6. Sea H una función en I^2 tal que H(x, y) = (2x - 1)(2y - 1). Entonces H es 2-creciente ya que para todo $x_1 \le x_2, y_1 \le y_2$ en I:

$$V_{H}([x_{1}, x_{2}] \times [y_{1}, y_{2}]) = H(x_{2}, y_{2}) - H(x_{2}, y_{1}) - H(x_{1}, y_{2}) + H(x_{1}, y_{1})$$

= $(2x_{2} - 1) (2y_{2} - 1) - (2x_{2} - 1) (2y_{1} - 1)$
 $- (2x_{1} - 1) (2y_{2} - 1) + (2x_{1} - 1) (2y_{1} - 1)$
= $2 (2x_{2} - 1) (y_{2} - y_{1}) - 2 (2x_{1} - 1) (y_{2} - y_{1})$
= $4 (x_{2} - x_{1}) (y_{2} - y_{1})$
 ≥ 0

Sin embargo H es decreciente en x para cada y en $(0, \frac{1}{2})$ y es decreciente en y para cada x en $(0, \frac{1}{2})$ ya que si tomamos a_1 , a_2 fijos en $(0, \frac{1}{2})$, $x_1 \le x_2$ y $y_1 \le y_2$

$$H(x_1, a_2) = (2x_1 - 1) (2a_2 - 1)$$

$$\geq (2x_2 - 1) (2a_2 - 1)$$

$$= H(x_2, a_2)$$

$$H (a_1, y_1) = (2a_1 - 1) (2y_1 - 1)$$

$$\geq (2a_1 - 1) (2y_2 - 1)$$

$$= H (a_1, y_2)$$

Definición 1.1.7. Sean $S_1 \ y \ S_2$ dos subconjuntos no vacíos de $\mathbf{\bar{R}}$, y sea Huna función real de doble entrada tal que $DomH = S_1 \times S_2$. Supongamos que $a_1 = \inf(S_1) \ y \ a_2 = \inf(S_2)$, decimos que la función H es fijada si $H(x, a_2) = 0 = H(a_1, y)$ para todo (x, y) en DomH.

Lema 1.1.8. Sean S_1 y S_2 subconjuntos no vacíos de \overline{R} , y sea H una función fijada 2-creciente con dominio $S_1 \times S_2$. Entonces H es no decreciente en cada argumento.

Definición 1.1.9. Suponga que $b_1 = \sup(S_1)$ y $b_2 = \sup(S_2)$. Entonces decimos que una función H de $S_1 \times S_2$ a **R** tiene marginales, y las marginales de H son funciones F y G dadas por:

 $DomF = S_1, y F(x) = H(x, b_2)$ para todo x en S_1 . $DomG = S_2, y G(y) = H(b_1, y)$ para todo y en S_2 .

Ejemplo 1.1.10. Sea H una función con dominio $[-1,1] \times [0,\infty]$ dada por

$$H(x,y) = \frac{(x+1)(e^y - 1)}{x + 2e^y - 1}$$

Entonces H es fijada por que H(x, 0) = 0 = H(-1, y); y H tiene marginales F(x) y G(y) dadas por

$$F(x) = H(x, \infty) = \frac{x+1}{2} \ y \ G(y) = H(1, y) = 1 - e^{-y}$$

Lema 1.1.11. Sean S_1 y S_2 dos subconjuntos no vacíos de \mathbf{R} y sea H una función 2-creciente, fijada, con marginales, cuyo dominio es $S_1 \times S_2$. Sean (x_1, y_1) y (x_2, y_2) puntos cualesquiera en $S_1 \times S_2$. Entonces

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

Definición 1.1.12. Una función de distribución es una función F con dominio \bar{R} tal que

1. F es no decreciente

2. $F(-\infty) = 0 \ y \ F(\infty) = 1$.

Definición 1.1.13. Una función de distribución conjunta es una función H con dominio $\overline{\mathbf{R}}^2$ tal que

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

Entonces H es fijada y, porque $DomH = \overline{\mathbf{R}}^2$, H tiene marginales F y G dadas por $F(x) = H(x, \infty)$ y $G(y) = H(\infty, y)$. Además, por consecuencia del Lema 1.1.8, F y G son funciones de distribución. Citando a Nelsen (2006) con respecto a las definiciones anteriores:

Notemos que no hay nada de "probabilístico" en estas definiciones de funciones de distribución. No se mencionan a las variables aleatorias, ni continuidad por la izquierda o por la derecha. Todas las funciones de distribución de una o dos variables aleatorias que normalmente se encuentran en estadística satisfacen la primera o la segunda de las definiciones anteriores. Entonces cualquiera de los resultados que se obtenga para dichas funciones de distribución se cumplirán cuando hablemos de variables aleatorias, independientemente de alguna otra restricción que se pueda imponer.

1.2. Cópulas

Ya definidos los conceptos a utilizar, podemos continuar con la definición formal de una cópula bivariada, así como sus propiedades básicas.

Definición 1.2.1. Una Cópula bivariada es una función $C : \mathbf{I}^2 \to \mathbf{I}$ con las siguientes propiedades:

1. Para cada u, v en I,

$$C(u,0) = 0 = C(0,v), \qquad (1.2)$$

$$C(u, 1) = u, \ C(1, v) = v$$
 (1.3)

2. Para cada u_1, u_2, v_1, v_2 en **I** tal que $u_1 \leq u_2$ y $v_1 \leq v_2$,

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \ge 0$$
(1.4)

Por consecuencia de la Definición anterior, al ser C una función fijada en 0 y 2-creciente, las cópulas son no decrecientes en sus argumentos y por lo tanto para todo u, v en I,

$$0 \le C\left(u, v\right) \le 1$$

En el siguiente ejemplo mencionaremos una cópula importante que nos encontraremos frecuentemente en los siguientes capítulos.

Ejemplo 1.2.2. Sea Π una función de I^2 a I definida por $\Pi(u, v) = uv$ entonces para todo u, v en I

$$\Pi (u, 0) = u * 0 = 0 = 0 * v = \Pi (0, v)$$
$$\Pi (u, 1) = u, \ \Pi (1, v) = v$$

y para cualesquiera u_1 , u_2 , v_1 , v_2 en I tal que $u_1 \leq u_2$, $v_1 \leq v_2$

$$V_{\Pi} \left([u_1, u_2] \times [v_1, v_2] \right) = \Pi \left(u_2, v_2 \right) - \Pi \left(u_2, v_1 \right) - \Pi \left(u_1, v_2 \right) + \Pi \left(u_1, v_1 \right)$$

$$= u_2 v_2 - u_2 v_1 - u_1 v_2 + u_1 v_1$$

$$= u_2 \left(v_2 - v_1 \right) - u_1 \left(v_2 - v_1 \right)$$

$$= (u_2 - u_1) \left(v_2 - v_1 \right)$$

$$> 0$$

Por lo que Π es una cópula.

Teorema 1.2.3. Sea C una cópula. Entonces para cada u_1 , u_2 , v_1 , v_2 en I:

$$|C(u_2, v_2) - C(u_1, v_1)| \le |u_2 - u_1| + |v_2 - v_1|$$
(1.5)

Por la desigualdad (1.5) se dice que las cópulas cumplen la condición de Lipschitz y ésta misma establece la continuidad de las cópulas.

Definición 1.2.4. Sea C una cópula, y sea a un número en I. La sección horizontal de C en a es la función de $I \mapsto I$ dada por $t \mapsto C(t, a)$; la sección vertical de C en a es la función de $I \mapsto I$ dada por $t \mapsto C(a, t)$; y la sección diagonal de C es la función δ_C de $I \mapsto I$ definida como $\delta_C(t) = C(t, t)$.

1.2. CÓPULAS

Corolario 1.2.5. La sección horizontal, vertical y diagonal de una cópula C son todas no decrecientes y uniformemente continuas en I.

Teorema 1.2.6. Sean M y W functiones de \mathbf{I}^2 a \mathbf{I} definidas como $M(u, v) = \min(u, v)$ y $W(u, v) = \max(u + v - 1, 0)$ entonces M y W son cópulas. Además para toda cópula C y cualesquiera u, v en \mathbf{I} ,

$$W(u,v) \le C(u,v) \le M(u,v) \tag{1.6}$$

La desigualdad (1.6) es la versión en cópulas de la desigualdad de las cotas de Fréchet-Hoeffding. Nos referiremos a M como la cota superior de Fréchet-Hoeffding y a W como la cota inferior de Fréchet-Hoeffding.

Se sigue de la Definición 1.2.1 y del Teorema 1.2.3 que la gráfica de una cópula es una superficie continua dentro del cubo unitario I^3 ; y por el Teorema 1.2.6, esta superficie se encuentra dentro de las gráficas de las cotas de Fréchet-Hoeffding (veáse figura 1.2.1).

Una forma simple pero útil para mostrar la gráfica de una cópula son los diagramas de contorno o conjuntos de nivel, éstos son los conjuntos en I^2 dados por C(u, v) = a, para diferentes valores de a en I. En la figura 1.2.2 se muestran los diagramas de contorno de las cópulas $M, W \ge \Pi$.



Figura 1.2.1: Gráficas de las cópulas $M, W \ge \Pi$



Figura 1.2.2: Diagramas de contorno para las cópulas $M, W \neq \Pi$

Teorema 1.2.7. Sea C una cópula. Para cualquier v en **I**, la derivada parcial $\frac{\partial}{\partial u}C(u, v)$ existe para casi todo u, y además,

$$0 \le \frac{\partial}{\partial u} C\left(u, v\right) \le 1 \tag{1.7}$$

De forma análoga, para cualquier u en I, la derivada parcial $\frac{\partial}{\partial v}C(u, v)$ existe para casi todo v, y además.

$$0 \le \frac{\partial}{\partial v} C\left(u, v\right) \le 1 \tag{1.8}$$

Más aún, las funciones $u \mapsto \frac{\partial}{\partial u} C(u, v) \ y \ v \mapsto \frac{\partial}{\partial v} C(u, v)$ están bien definidas y son no decrecientes en casi todo **I**.

Teorema 1.2.8. Sea C una cópula. Si $\frac{\partial}{\partial v}C(u,v)$ y $\frac{\partial^2}{\partial u\partial v}C(u,v)$ son continuas en \mathbf{I}^2 y $\frac{\partial}{\partial u}C(u,v)$ existe para todo u en (0,1) cuando v = 0, entonces $\frac{\partial}{\partial u}C(u,v)$ y $\frac{\partial^2}{\partial v\partial u}C(u,v)$ existe en $(0,1)^2$ y $\frac{\partial^2}{\partial u\partial v}C(u,v) = \frac{\partial^2}{\partial v\partial u}C(u,v)$.

Proposición 1.2.9. Sean C_i cópulas y a_i en I para i=1,2,...,n tal que $\sum_{i=0}^{n} a_i = 1$. Entonces cualquier combinación lineal convexa de cópulas también es cópula, es decir,

$$\sum_{i=0}^{n} a_i C_i$$

es cópula.

1.2. CÓPULAS

La siguiente proposición nos da una técnica de construcción de cópulas propuesta por Siburg y Stoimenov (2008) denominada "pegado de cópulas". Por ejemplo, dadas dos cópulas bivariadas C_1 , C_2 y un valor fijo $0 < \theta < 1$ podemos escalar C_1 a $[0, \theta] \times [0, 1]$, C_2 a $[\theta, 1] \times [0, 1]$ y pegarlas en una sola cópula tal y como se muestra en la Figura 1.2.3.



Figura 1.2.3: Técnica de pegado de cópulas

Proposición 1.2.10. Sean C_1 y C_2 dos cópulas bivariadas y θ en (0,1) entonces la función $C_{1,2,\theta}$ definida como:

$$C_{1,2,\theta}(u,v) = \begin{cases} \theta C_1\left(\frac{u}{\theta},v\right), & 0 \le u \le \theta, \\ (1-\theta) C_2\left(\frac{u-\theta}{1-\theta},v\right) + \theta v, & \theta \le u \le 1. \end{cases}$$
(1.9)

es una cópula.

A continuación, a manera de ejemplos, veremos algunas cópulas que serán de nuestro interés en los siguientes capítulos.

Ejemplo 1.2.11. Las familias de cópulas de Fréchet y Mardia. Si una familia de cópulas incluye M, Π y W es llamada completa

(a) Sean α, β en I con $\alpha + \beta \leq 1$ y,

$$C_{\alpha,\beta}(u,v) = \alpha M(u,v) + (1 - \alpha - \beta) \Pi(u,v) + \beta W(u,v) \quad (1.10)$$

Entonces $C_{\alpha,\beta}$ es una cópula. Esta familia completa de dos parámetros se debe a Fréchet (1958).

(b) Sea θ en [-1,1] y sea,

$$C_{\theta}(u,v) = \frac{\theta^2 (1+\theta)}{2} M(u,v) + (1-\theta^2) \Pi(u,v) + \frac{\theta^2 (1-\theta)}{2} W(u,v)$$
(1.11)

Entonces C_{θ} es cópula. Esta familia completa de un parámetro se debe a Mardia (1970).

Ejemplo 1.2.12. Las familias de cópulas de Cuadras-Augé y Marshall-Olkin.

(a) Sean θ en I y,

$$C_{\theta}(u,v) = \left[\min(u,v)\right]^{\theta} \left[uv\right]^{1-\theta} = \begin{cases} uv^{1-\theta}, & \text{si } u \le v, \\ u^{1-\theta}v, & \text{si } v \le u, \end{cases}$$
(1.12)

Nótese que $C_0 = \Pi \ y \ C_1 = M$, esta familia se debe a Cuadras y Augé (1981).

(b) Sean α, β en I y,

$$C_{\alpha,\beta}(u,v) = \begin{cases} u^{1-\alpha}v, & \text{si } v^{\beta} \le u^{\alpha}, \\ uv^{1-\beta}, & \text{si } u^{\alpha} \le v^{\beta}, \end{cases}$$
(1.13)

Esta familia se debe a Marshall y Olkin (1967a,b), nótese que cuando $\alpha = \beta = \theta$ (1.13) se reduce a la familia de Cuadras-Augé.

1.3. Teorema de Sklar

El teorema que da nombre a esta sección, es pieza clave y punto de partida en el estudio y construcción de la teoría de cópulas y fundamental para su aplicación en la teoría de la probabilidad y la estadística ya que aclara el papel que juegan las cópulas para explicar el vínculo que existe entre una distribución multivariada y sus respectivas distribuciones marginales. **Teorema 1.3.1** (Teorema de Sklar (1959)). Sea H una función de distribución conjunta con marginales F y G. Entonces existe una cópula C tal que para todo x, y en $\bar{\mathbf{R}}$,

$$H(x,y) = C(F(x), G(y))$$

$$(1.14)$$

Si además $F \ y \ G$ son continuas, entonces C es única, de otro modo sólo es única sobre $RanF \times RanG$. Más aún, si C es cualquier cópula, $y \ F \ y \ G$ cualesquiera funciones de distribución univariadas, entonces la función Hdefinida como (1.14) es una función de distribución conjunta con marginales $F \ y \ G$.

Para detalles de su demostración véase Nelsen (2006).

Definición 1.3.2. Sea F una función de distribución. Una cuasi inversa de F es cualquier función $F^{(-1)}$ con dominio I tal que,

1. Si t está en RanF, entonces $F^{(-1)}(t)$ es cualquier numero x en $\bar{\mathbf{R}}$ tal que F(x) = t, es decir, para todo t en RanF

$$F\left(F^{(-1)}\left(t\right)\right) = t$$

2. Si t no está en RanF, entonces

$$F^{(-1)}(t) = \inf \{ x | F(x) \ge t \} = \sup \{ x | F(x) \le t \}.$$

El Teorema de Sklar nos garantiza la existencia de la cópula subyacente para toda distribución conjunta; el siguiente resultado nos da una forma de obtener dicha cópula, ya que de manera general se puede "extraer" la cópula de la distribución. Para esto, es suficiente el conocimiento de H, porque a partir de ella se pueden obtener sus marginales, $F ext{ y } G$, y posteriormente sus respectivas cuasi inversas.

Corolario 1.3.3. Sean H, F, G y 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 cualquier u, v en **I**

$$C(u,v) = H\left(F^{(-1)}(u), G^{(-1)}(v)\right).$$
(1.15)

Ejemplo 1.3.4. Sea H una función de distribución definida como

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

con marginales F y G dadas por

$$F(x) = \begin{cases} 0, & x < -1, \\ \frac{(x+1)}{2}, & x \in [-1,1], y \ G(y) = \begin{cases} 0, & y < 0, \\ 1 - e^{-y}, & y \ge 0, \end{cases}$$

Las cuasi inversas de F y G están dadas por $F^{(-1)}(u) = 2u-1 y G^{(-1)}(v) = -\ln(1-v)$ para u, v en I. Porque RanF = RanG = I, usando (1.15) la cópula C subyacente es

$$C(u,v) = \frac{uv}{u+v-uv}.$$
(1.16)

Г		٦	
		1	
		1	

Ejemplo 1.3.5. Distribución exponencial bivariada de Gumbel (1960). Sea H_{θ} una función de distribución dada por

$$H_{\theta}(x,y) = \begin{cases} 1 - e^{-x} - e^{-y} + e^{-(x+y+\theta xy)}, & x \ge 0, y \ge 0, \\ 0, & \text{en otro caso.} \end{cases}$$

donde θ está en **I**. Entonces las distribuciones marginales son distribuciones exponenciales $F(x) = 1 - e^{-x} y G(y) = 1 - e^{-y}$, con cuasi inversas $F^{(-1)}(u) = -\ln(1-u) y G^{(-1)}(v) = -\ln(1-v)$ para u, v en **I**. Entonces la correspondiente cópula es

$$C_{\theta}(u,v) = u + v - 1 + (1-u)(1-v)e^{-\theta \ln(1-u)\ln(1-v)}.$$
 (1.17)

Ejemplo 1.3.6. Sean N_{ρ} y Φ las distribuciones normal estándar bivariada y univariada respectivamente, entonces la cópula normal está dada por

$$C(u,v) = N_{\rho} \left(\Phi^{(-1)}(u), \Phi^{(-1)}(v) \right)$$

= $\frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\Phi^{(-1)}(u)} \int_{-\infty}^{\Phi^{(-1)}(v)} exp\left[\frac{-\left(s^2 - 2\rho st + t^2\right)}{2\left(1-\rho^2\right)} \right] dsdt$
(1.18)

Observación 1.3.7. Como ya habíamos mencionado, el dominio de las cópulas es el cuadrado unitario I^2 , pero si lo extendemos a \bar{R}^2 podemos ver a las cópulas como una distribución conjunta con distribuciones marginales uniformes en I. Es decir si definimos la función H_C como

$$H_{C}(x,y) = \begin{cases} 0, & x < 0 \text{ o } y < 0, \\ C(x,y), & (x,y) \in \mathbf{I}^{2}, \\ x, & y > 1, x \in \mathbf{I} \\ y, & x > 1, y \in \mathbf{I}, \\ 1, & x > 1 y y > 1. \end{cases}$$

Entonces H_C es una distribución conjunta con marginales uniformes. En conclusión, podemos ver a las cópulas como restricciones de distribuciones conjuntas al cuadrado unitario \mathbf{I}^2 cuyas marginales son distribuciones uniformes.

1.4. Cópulas y Variables Aleatorias

En las secciones anteriores hemos hablado de funciones de distribución sin mencionar en algún momento a las variables aleatorias, por lo que en esta sección mostraremos la relación que éstas mantienen con las cópulas. Para esto usaremos las letras mayúsculas como X, Y, Z para representar a las variables aleatorias, y letras minúsculas x, y, z para representar sus valores. Decimos que F es una función de distribución de la variable X si para todo x en $\overline{\mathbf{R}}, F(x) = P[X \leq x]$; decimos que una variable aleatoria es continua si su función de distribución es continua.

Al hablar de dos variables adoptaremos la misma convención; dichas variables son elementos de un vector aleatorio cuyos valores están descritos por una distribución conjunta. Decimos que H es una función de distribución conjunta del vector aleatorio (X, Y) si para todo x, y en $\overline{\mathbf{R}}$, $H(x,y) = P[X \leq x, Y \leq y]$. Ahora estamos en posición de enunciar el teorema de Sklar en términos de variables aleatorias y sus funciones de distribución.

Teorema 1.4.1. Sean X, Y variables aleatorias con función de distribución F y G respectivamente y con función de distribución conjunta H. Entonces existe una cópula C tal que (1.14) se cumple. Si F y G son continuas, C es única, de otro modo C está determinada de manera única en Ran $F \times RanG$.

A la cópula C del Teorema 1.4.1 la llamaremos la cópula de (X, Y) y será denotada como C_{XY} .

Ejemplo 1.4.2. Sea X una variable aleatoria con función de distribución F, y sea Y = X, entonces para todo x, y en $\overline{\mathbf{R}}$

$$H(x, y) = P(X \le x, Y \le y)$$

= $P(X \le x, X \le y)$
= $P(X \le \min(x, y))$
= $F(\min(x, y))$
= $\min(F(x), F(y))$

Entonces aplicando el Corolario 1.3.3 tenemos:

$$C_{XY}(u,v) = H\left(F^{(-1)}(u), G^{(-1)}(v)\right)$$

= $H\left(F^{(-1)}(u), F^{(-1)}(v)\right)$
= mín $\left(F\left(F^{(-1)}(u)\right), F\left(F^{(-1)}(v)\right)\right)$
= mín (u,v)
= $M(u,v)$

Por lo que podemos concluir que $C_{XX} = M$.

Ejemplo 1.4.3. Sea X una variable aleatoria con función de distribución F, $y \operatorname{sea} Y = -X \operatorname{con} \operatorname{función} \operatorname{de} \operatorname{distribución} G.$ Entonces $G(y) = P(Y \le y) =$ $P(-X \leq y) = P(X > -y) = 1 - F(-y)$, y para todo x, y en \bar{R}

$$H(x, y) = P(X \le x, Y \le y)$$

= $P(X \le x, -X \le y)$
= $P(-y \le X \le x)$
= $\begin{cases} F(x) - F(-y), & si - y < x \\ 0, & si - y \ge x \end{cases}$
= $\begin{cases} F(x) + G(y) - 1, & si - y < x \\ 0, & si - y \ge x \end{cases}$
= $máx (F(x) + G(y) - 1, 0)$

La última igualdad se da ya que -y < x es equivalente a F(x) + G(y) - 1 > 0

x

1.4. CÓPULAS Y VARIABLES ALEATORIAS

y entonces aplicando el Corolario 1.3.3 tenemos:

$$C_{XY}(u, v) = H\left(F^{(-1)}(u), G^{(-1)}(v)\right)$$

= máx $\left(F\left(F^{(-1)}(u)\right) + G\left(G^{(-1)}(v)\right) - 1, 0\right)$
= máx $(u + v - 1, 0)$
= $W(u, v)$

Por lo tanto $C_{X,-X} = W$.

Recordemos que dos variables aleatorias son independientes si y sólo si su función de distribución conjunta es el producto de las distribuciones marginales, entonces el siguiente Teorema muestra que la cópula producto $\Pi(u, v) = uv$ caracteriza a las variables aleatorias independientes.

Teorema 1.4.4. Sean X, Y variables aleatorias continuas. Entonces X, Y son independientes si y sólo si $C_{XY} = \Pi$.

Una propiedad bastante importante acerca de las cópulas es que ante transformaciones estrictamente monótonas a las variables aleatorias, las cópulas son invariantes o cambian en formas predecibles. Los siguientes teoremas nos muestran el comportamiento de las cópulas ante dichas transformaciones.

Teorema 1.4.5. Sean X, Y variables aleatorias continuas con cópula C_{XY} . Si α y β son funciones estrictamente crecientes en RanX y RanY, respectivamente, entonces $C_{\alpha(X)\beta(Y)} = C_{XY}$. Es decir, C_{XY} es invariante bajo transformaciones estrictamente crecientes de X, Y.

Cuando al menos α o β es estrictamente decreciente, la cópula de $\alpha(X)$ y $\beta(Y)$ es una simple transformación de C_{XY} , de forma más precisa:

Teorema 1.4.6. Sean X, Y variables aleatorias continuas con cópula C_{XY} . Sean α y β funciones estrictamente monótonas en RanX y RanY, respectivamente.

1. Si α es estrictamente creciente y β estrictamente decreciente, entonces

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

2. si α es estrictamente decreciente y β estrictamente creciente, entonces

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

3. Si α y β son ambas estrictamente decrecientes, entonces

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

Nótese que en cada caso la forma de la cópula es independiente de la elección de α y β .

Ejemplo 1.4.7. Retomando los Ejemplos 1.4.2 y 1.4.3 y aplicando el Teorema 1.4.6 tenemos que para toda variable X

- 1. Si $Y = \alpha(X)$ con α una función creciente entonces $C_{XY} = M$.
- 2. Si $Y = \beta(X)$ con β una función decreciente entonces $C_{XY} = W$.

Hasta el momento, se decidió omitir la teoría de la medida en la definición de variables aleatorias, por lo que se mencionarán algunos aspectos relacionados respecto a este tema.

En general, cualquier función de distribución induce una medida de probabilidad al asignar cada elemento de $\overline{\mathbf{R}}$ un valor en I. De igual forma, cada función de distribución conjunta H induce una medida de probabilidad en \mathbf{R}^2 vía $V_H([-\infty, x] \times [-\infty, y]) = H(x, y)$. Como las cópulas son funciones de distribución conjuntas (con marginales uniformes (0, 1)), cada cópula Cinduce una medida de probabilidad en I^2 vía $V_C([0, u] \times [0, v]) = C(u, v)$, es decir, la C-medida de un conjunto es su C-volumen V_C .

Por otra parte, podemos dividir a la cópula de la siguiente manera.

Definición 1.4.8. La parte absolutamente continua de una cópula C es:

$$A_{C}(u,v) = \int_{0}^{u} \int_{0}^{v} \frac{\partial^{2}}{\partial s \partial t} C(s,t) dt ds \qquad (1.19)$$

y la parte singular de una cópula C es

$$S_C(u, v) = C(u, v) - A_C(u, v)$$
(1.20)

 $y \ entonces$

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

Definición 1.4.9. Si $C = A_C$ en \mathbf{I}^2 -esto es, si consideramos a C como una función de distribución conjunta, entonces su densidad conjunta está dada por $\frac{\partial^2}{\partial u \partial v} C(u, v)$ - entonces C es absolutamente continua. Si $C = S_C$ en \mathbf{I}^2 -si $\frac{\partial^2}{\partial u \partial v} C(u, v) = 0$ en casi todo \mathbf{I}^2 - entonces C es singular.

Definición 1.4.10. La C-medida de la parte absolutamente continua se define como $A_C(1,1)$ y la C-medida de la parte singular como $S_C(1,1)$

Definición 1.4.11. El soporte de una cópula es el complemento de la unión de todos los subconjuntos abiertos de \mathbf{I}^2 con C-volumen V_C cero. Cuando el soporte de C es \mathbf{I}^2 , decimos que C tiene "soporte completo".

Observación 1.4.12. Cuando C es singular, su soporte tiene medida de Lebesgue cero (e inversamente). Sin embargo, existen cópulas que tienen soporte completo y tienen tanto parte absolutamente continua como parte singular.

Ejemplo 1.4.13. El soporte de la cota superior de Fréchet-Hoeffding M es la diagonal principal de \mathbf{I}^2 , es decir, la gráfica de v = u para u en \mathbf{I} , por lo que M es singular. Ya que para cualquier rectángulo que se encuentre totalmente por encima o debajo de la diagonal principal, su C-volumen es cero. Similarmente, el soporte de la cota inferior de Fréchet-Hoeffding W es la diagonal secundaria de \mathbf{I}^2 , es decir, la gráfica de v = 1 - u para u en \mathbf{I} , y entonces W es singular también.

Ejemplo 1.4.14. La cópula producto $\Pi(u, v) = uv$ es absolutamente continua, porque para todo (u, v) en I^2 ,

$$A_{\Pi}(u,v) = \int_{0}^{v} \int_{0}^{v} \frac{\partial^{2}}{\partial s \partial t} \Pi(s,t) \, dt ds = \int_{0}^{v} \int_{0}^{v} 1 dt ds = uv = \Pi(u,v) \,.$$

Finalizamos esta sección con algunos ejemplos de cópulas que serán de utilidad en los siguientes capítulos.

Ejemplo 1.4.15. Distribución logística bivariada de Gumbel (Gumbel (1961)). Sean (X, Y) variables aleatorias con función de distribución conjunta dada por

$$H(x,y) = (1 + e^{-x} + e^{-y})^{-1}$$

para todo x,y en \mathbf{R} . Entonces (X, Y) tienen distribución logística estándar (univariada), es decir,

$$F(x) = (1 + e^{-x})^{-1} y G(y) = (1 + e^{-y})^{-1}.$$

Entonces la cópula de (X, Y) es la cópula dada por (1.16) en el Ejemplo 1.3.4.

Ejemplo 1.4.16. Distribuciones bivariadas de valores extremos tipo B (Johnson y Kotz (1970)). Sean (X, Y) variables aleatorias con función de distribución conjunta dada por

$$H_{\theta}(x,y) = exp\left[-\left(e^{-\theta x} + e^{-\theta y}\right)^{1/\theta}\right]$$

para todo x, y en $\overline{\mathbf{R}}$, donde $\theta \geq 1$. Entonces la cópula de (X, Y) está dada por

$$C_{\theta}(u,v) = exp\left(-\left[\left|\ln u\right|^{\theta} + \left|\ln v\right|^{\theta}\right]^{1/\theta}\right).$$
(1.21)

Esta familia paramétrica de cópulas es conocida como la familia de Gumbel-Hougaard (Hutchison y Lai (1990)).

Ejemplo 1.4.17. Nótese que la distribución logística bivariada de Gumbel sufre del defecto de carecer de un parámetro, lo cual limita su uso en aplicaciones. Esto puede corregirse en varias formas, una de ellas es definir H_{θ} como

$$H_{\theta}(x,y) = \left(1 + e^{-x} + e^{-y} + (1-\theta)e^{-x-y}\right)^{-1}$$

para todo x, y en $\bar{\mathbf{R}}$ con θ en [-1, 1]. Entonces:

- (a) Las marginales son distribuciones logísticas estándar;
- (b) cuando $\theta = 1$, tenemos la distribución logística bivariada de Gumbel;
- (c) cuando $\theta = 0$, (X, Y) son independientes; y
- (d) la cópula de (X, Y) está dada por

$$C_{\theta}(u,v) = \frac{uv}{1 - \theta(1 - u)(1 - v)}$$
(1.22)

Ésta es la familia de cópulas Ali-Mikhail-Haq (Hutchison y Lai (1990)).

1.5. Cópulas Arquimedianas

En esta sección analizaremos una clase importante de cópulas conocida como cópulas Arquimedianas. Estas cópulas tienen un amplio campo de aplicaciones por varias razones como: la facilidad con la que son construidas; la gran variedad de modelos paramétricos que pertenecen a esta clase; las buenas propiedades que los miembros de esta clase poseen.

Definición 1.5.1. Sea φ una función continua, estrictamente decreciente de \mathbf{I} a $[0,\infty]$ tal que $\varphi(1) = 0$. La pseudo inversa de φ es la función $\varphi^{[-1]}$ con $Dom\varphi^{[-1]} = [0,\infty]$ y $Ran\varphi^{[-1]} = \mathbf{I}$ dada por

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{(-1)}(t), & 0 \le t \le \varphi(0), \\ 0, & \varphi(0) \le t \le \infty. \end{cases}$$
(1.23)

Observación 1.5.2. Notemos que $\varphi^{[-1]}$ es continua y no creciente en $[0, \infty]$, y estrictamente decreciente en $[0, \varphi(0)]$. Así como, $\varphi^{[-1]}(\varphi(u)) = u$ en I, y

$$\varphi\left(\varphi^{\left[-1\right]}\left(t\right)\right) = \begin{cases} t, & 0 \le t \le \varphi\left(0\right), \\ \varphi\left(0\right), & \varphi\left(0\right) \le t \le \infty, \end{cases}$$
$$= \min\left(t, \varphi\left(0\right)\right).$$

Finalmente, si $\varphi(0) = \infty$, entonces $\varphi^{[-1]} = \varphi^{-1}$, la inversa usual.

Lema 1.5.3. Sea φ una función continua estrictamente decreciente de \mathbf{I} a $[0,\infty]$ tal que $\varphi(1) = 0$, y sea $\varphi^{[-1]}$ la pseudo inversa de φ definida como (1.23). Sea C la función de \mathbf{I}^2 a \mathbf{I} dada por

$$C(u,v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)). \qquad (1.24)$$

Entonces C satisface las ecuaciones (1.2) y (1.3) de la definición de cópula.

Lema 1.5.4. Sean φ , $\varphi^{[-1]}$ y C como en el Lema 1.5.3. Entonces C es 2-creciente si y sólo si para todo v en I y $u_1 \leq u_2$,

$$C(u_2, v) - C(u_1, v) \le u_2 - u_1.$$
(1.25)

Teorema 1.5.5. Sea φ una función continua estrictamente decreciente de $I a [0, \infty]$ tal que $\varphi(1) = 0$, y sea $\varphi^{[-1]}$ la pseudo inversa de φ definida como (1.23). Entonces la función C de I^2 a I dada por (1.24) es una cópula si y sólo si φ es convexa.

Definición 1.5.6. A las cópulas obtenidas como en el Teorema 1.5.5 se les denomina Arquimedianas y a la función φ se le denomina generador de la cópula. Si $\varphi(0) = \infty$ se dice que φ es un generador estricto y C es una cópula Arquimediana estricta.

Proposición 1.5.7. Sea C una cópula Arquimediana, C es estricta si y sólo si C(u, v) > 0 para todo (u, v) en $(0, 1]^2$.

Proposición 1.5.8. La sección diagonal de una cópula Arquimediana C con generador φ está dada por $\delta_C(u) = \varphi^{[-1]}[2\varphi(u)]$. Entonces para todo u en $(0,1), \delta_C(u) < u$.

Ejemplo 1.5.9.

- 1. Sea $\varphi(t) = -\ln t$ para t en [0,1]. Porque $\varphi(0) = \infty$, φ es estricto. Entonces $\varphi^{[-1]}(t) = \varphi^{-1}(t) = exp(-t)$, y generando C con (1.24) tenemos C $(u,v) = exp(-[(-\ln u) + (-\ln v)]) = uv = \Pi(u,v)$. Entonces Π es una cópula Arquimediana estricta.
- 2. Sea $\varphi(t) = 1 t$ para t en [0,1]. Entonces $\varphi^{[-1]} = 1 t$ para t en [0,1] y 0 para t > 1; es decir, $\varphi^{[-1]}(t) = \max(1-t,0)$. Usando (1.24) $C(u,v) = \max(u+v-1,0) = W(u,v)$. Entonces W es también Arquimediana.
- 3. La sección diagonal de la cópula M está dada por $\delta_M(u) = u$, y por Proposición 1.5.8 concluimos que M no es Arquimediana.

Teorema 1.5.10. Sea C una cópula Arquimediana con generador φ . Entonces:

- 1. C es simétrica, es decir, C(u, v) = C(v, u) para todo u, v en I;
- 2. C es asociativa, es decir, C(C(u, v), w) = C(u, C(v, w)) u, v, w en I;
- 3. Si c > 0 es cualquier constante, entonces $c\varphi$ es también un generador de C.

Teorema 1.5.11. Sea C una cópula asociativa tal que $\delta_C(u) < u$ para todo u en (0, 1). Entonces C es Arquimediana.

Una pregunta natural acerca de estas cópulas es: ¿Cuál es el significado del término Arquimediana? Recordemos la propiedad Arquimediana en los

1.5. CÓPULAS ARQUIMEDIANAS

números reales: Si $a \neq b$ son números reales positivos, entonces existe un entero n tal que na > b. Las cópulas Arquimedianas se pueden ver como una operación binaria que asigna un par de elementos u, v en I a C(u, v) en I. Del Teorema 1.5.10 podemos ver que esta "operación" C es conmutativa y asociativa, y en consecuencia del Corolario 1.2.5 preserva el orden, es decir, $u_1 \leq u_2 \neq v_1 \leq v_2$ implica $C(u_1, v_1) \leq C(u_2, v_2)$; por lo que a (I, C) se le denomina semigrupo abeliano ordenado. Para u en I, definimos a u_C^n de forma recursiva como: $u_C^1 = u, \neq u_C^{n+1} = C(u, u_C^n)$.

Teorema 1.5.12. Sea C una cópula Arquimediana generada por φ . Entonces para cualesquiera u, v en I, existe un entero positivo n tal que $u_C^n < v$

El teorema anterior es la versión de la propiedad Arquimediana para (I, C), y es el motivo por el cual las cópulas Arquimedianas reciban dicho nombre.

De acuerdo con el Teorema 1.5.5 podemos generar tantas cópulas Arquimedianas como generadores podamos identificar. A continuación listaremos en el Cuadro 1.5.1 algunas familias de las cópulas Arquimedianas más conocidas. (Para un catálogo más amplio veáse Nelsen (2006)).

Familia	$\varphi_{ heta}\left(t ight)$	$C_{ heta}\left(u,v ight)$	$ heta\in$
Clayton Ali-Mikhail -Haq	$\frac{1}{\theta} \left(t^{-\theta} - 1 \right)$ $\ln \frac{1 - \theta (1 - t)}{t}$	$\left[\max\left(u^{-\theta} + v^{-\theta} - 1, 0\right)\right]^{-1/\theta}$ $\frac{uv}{1 - \theta(1 - u)(1 - v)}$	$[-1,\infty)\setminus\{0\}$ [-1,1)
Gumbel- Hougaard	$(-\ln t)^{\theta}$	$exp\left(-\left[\left \ln u\right ^{\theta}+\left \ln v\right ^{\theta}\right]^{1/\theta} ight)$	$[1,\infty)$
Frank	$-\ln rac{e^{- heta t}-1}{e^{- heta}-1}$	$-\frac{1}{\theta}\ln\left(1+\frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right)$	$(-\infty,\infty)\setminus\{0\}$
Gumbel -Barnett	$\ln\left(1-\theta\ln t\right)$	$uvexp\left(- heta\ln u\ln v ight)$	(0, 1]

Cuadro 1.5.1: Algunas familias Arquimedianas

1.6. Dependencia

La dependencia es un tema bastante importante dentro de la probabilidad y la estadística, siendo un paso primordial en el estudio del comportamiento conjunto de variables aleatorias.

Las cópulas toman una gran importancia dentro de este asunto porque proveen una manera natural de medir la dependencia ya que guardan toda la información del comportamiento conjunto de variables aleatorias y por ende la información necesaria para descrbir su estructura de dependencia.

Para explorar el papel que las cópulas juegan en el estudio de la dependencia, comenzaremos analizando dos medidas de asociación conocidas como la correlación de Kendall y la correlación de Spearman, que son medidas de concordancia.

Medidas de Concordancia

De manera informal diremos que un par de variables aleatorias son concordantes si valores "grandes" de una tienden a estar asociados con valores "grandes" de la otra y valores "pequeños" de una con valores "pequeños" de la otra. De forma más específica, sean (x_i, y_i) y (x_j, y_j) observaciones de un vector aleatorio (X, Y) de variables aleatorias continuas. Decimos que (x_i, y_i) y (x_j, y_j) son concordantes si $x_i < x_j$ y $y_i < y_j$, o si $x_i > x_j$ y $y_i > y_j$. De manera similar, decimos que (x_i, y_i) y (x_j, y_j) son discordantes si $x_i < x_j$ y $y_i > y_j$, o si $x_i > x_j$ y $y_i < y_j$. También hay que tener en cuenta la formulación alternativa: (x_i, y_i) y (x_j, y_j) son concordantes si $(x_i - x_j) (y_i - y_j) > 0$ y discordantes si $(x_i - x_j) (y_i - y_j) < 0$.

Definición 1.6.1. Sea (X, Y) un vector de variables aleatorias continuas con función de distribución conjunta H. Sean (X_1, Y_1) y (X_2, Y_2) vectores aleatorios independientes e idénticamente distribuidos, cada uno con función de distribución conjunta H. Entonces la tau de Kendall está definida como la probabilidad de concordancia menos la probabilidad de discordancia:

$$\tau = \tau_{X,Y} = P\left[(X_1 - X_2) \left(Y_1 - Y_2 \right) > 0 \right] - P\left[(X_1 - X_2) \left(Y_1 - Y_2 \right) < 0 \right]$$
(1.26)

Definición 1.6.2. Sea (X, Y) un vector de variables aleatorias continuas con función de distribución conjunta H. Sean (X_1, Y_1) , (X_2, Y_2) y (X_3, Y_3) tres vectores aleatorios independientes con función de distribución conjunta H. La rho de Spearman se define como la parte proporcional a la probabilidad de concordancia menos la probabilidad de discordancia de los dos vectores (X_1, Y_2) y (X_2, Y_3) , es decir, un par de vectores con las mismas marginales, pero uno tiene función de distribución conjunta H, mientras que los componentes del otro son independientes:

$$\rho_{X,Y} = 3\left(P\left[(X_1 - X_2)\left(Y_1 - Y_3\right) > 0\right] - P\left[(X_1 - X_2)\left(Y_1 - Y_3\right) < 0\right]\right)$$
(1.27)

Teorema 1.6.3. Sean X, Y variables aleatorias continuas cuya cópula es C. Entonces la tau de Kendall (que denotaremos como $\tau_{X,Y}$ o τ_C) y la rho de Spearman (que denotaremos como $\rho_{X,Y}$ o ρ_C) para (X,Y) están dadas por:

$$\tau_{X,Y} = \tau_C = 4 \int \int_{\mathbf{I}^2} C(u,v) \, dC(u,v) - 1 \tag{1.28}$$

y

$$\rho_{X,Y} = \rho_C = 12 \int \int_{I^2} C(u,v) \, du \, dv - 3 \tag{1.29}$$

Observación 1.6.4. Nótese que la integral en (1.28) puede interpretarse como la esperanza de la función C(U, V) de variables aleatorias uniformes (0, 1) U y V cuya función de distribución conjunta es C, es decir,

$$\tau_C = 4E \left[C \left(U, V \right) \right] - 1 \tag{1.30}$$

Observación 1.6.5. Recordando que las cópulas las podemos ver como funciones de distribución conjuntas con marginales uniformes (0,1) U y V, la expresión para ρ_C en (1.29) se puede reescribir en la siguiente forma:

$$\rho_{X,Y} = \rho_C = \frac{E\left(UV\right) - E\left(U\right)E\left(V\right)}{\sqrt{Var\left(U\right)}\sqrt{Var\left(V\right)}}.$$
(1.31)

En consecuencia la rho de Spearman para un par de variables aleatorias continuas X, Y es idéntica al coeficiente de correlación lineal de Pearson para las variables aleatorias $U = F(X) \ y \ V = G(Y)$.

Corolario 1.6.6. Sean X, Y variables aleatorias con C una cópula Arquimediana generada por φ . Entonces la tau de Kendall τ_C para X, Y está dada por

$$\tau_C = 1 + 4 \int_0^1 \frac{\varphi(t)}{\varphi'(t)} dt.$$
(1.32)

Ejemplo 1.6.7.

(a) Sea C_{θ} un miembro de la familia Clayton de cópulas Arquimedianas dada en el Cuadro 1.5.1. Entonces para $\theta \geq -1$,

$$\frac{\varphi_{\theta}\left(t\right)}{\varphi_{\theta}'\left(t\right)} = \frac{t^{\theta+1}-t}{\theta} \text{ cuando } \theta \neq 0, \ y \ \frac{\varphi_{0}\left(t\right)}{\varphi_{0}'\left(t\right)} = t \ln t \text{ para } \theta = 0;$$

por lo que

$$\tau_{\theta} = \frac{\theta}{\theta + 2}.$$

(b) Sea C_{θ} un miembro de la familia Gumbel-Hougaard de cópulas Arquimedianas dada en el Cuadro 1.5.1. Entonces para $\theta \geq 1$,

$$\frac{\varphi_{\theta}\left(t\right)}{\varphi_{\theta}'\left(t\right)} = \frac{t\ln t}{\theta},$$

y por lo tanto

$$\tau_{\theta} = \frac{\theta - 1}{\theta}$$

		_	

Teorema 1.6.8. Sean (X, Y) variables aleatorias continuas cuya cópula es C. Entonces la tau de Kendall y la rho de Spearman para (X, Y) se pueden expresar como:

$$\tau_C = 1 - 4 \int \int_{\mathbf{I}^2} \frac{\partial}{\partial u} C(u, v) \frac{\partial}{\partial v} C(u, v) \, du dv \tag{1.33}$$

y

$$\rho_C = 12 \int \int_{I^2} \left[C(u, v) - uv \right] du dv.$$
 (1.34)

Ejemplo 1.6.9. Sea $C_{\alpha,\beta}$ un miembro de la familia Marshall-Olkin (1.13) para $0 < \alpha, \beta < 1$:

$$C_{\alpha,\beta}(u,v) = \begin{cases} u^{1-\alpha}v, & \text{si } v^{\beta} \leq u^{\alpha}, \\ uv^{1-\beta}, & \text{si } u^{\alpha} \leq v^{\beta}, \end{cases}$$
1.6. DEPENDENCIA

(a) Las parciales de $C_{\alpha,\beta}$ no existen en la curva $u^{\alpha} = v^{\beta}$ por lo que

$$\frac{\partial}{\partial u} C_{\alpha,\beta}(u,v) \frac{\partial}{\partial v} C_{\alpha,\beta}(u,v) = \begin{cases} (1-\alpha) u^{1-2\alpha} v, & u^{\alpha} > v^{\beta}, \\ (1-\beta) u v^{1-2\beta}, & u^{\alpha} < v^{\beta}, \end{cases}$$

entonces

$$\int \int_{\mathbf{I}^2} \frac{\partial}{\partial u} C_{\alpha,\beta}\left(u,v\right) \frac{\partial}{\partial v} C_{\alpha,\beta}\left(u,v\right) du dv = \frac{1}{4} \left(1 - \frac{\alpha\beta}{\alpha - \alpha\beta + \beta}\right),$$

de lo que obtenemos

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

(b) Tenemos que

$$\int \int_{\mathbf{I}^2} C_{\alpha,\beta}\left(u,v\right) du dv = \frac{1}{2} \left(\frac{\alpha+\beta}{2\alpha-\alpha\beta+2\beta}\right),$$

por lo que

$$\rho_{\alpha,\beta} = rac{3lphaeta}{2lpha - lphaeta + 2eta}.$$

Ejemplo 1.6.10. Retomando los Ejemplos 1.6.9 y 1.2.12 sabemos que si en la familia Marshall-Olkin hacemos $\alpha = \beta = \theta$ ésta se convierte en la familia de Cuadras-Augé de cópulas entonces para esta familia tenemos que:

$$\tau_{\theta} = \frac{\theta^2}{\theta (2 - \theta)}$$
(1.35)
$$\rho_{\theta} = \frac{3\theta^2}{\theta (4 - \theta)}$$

En la siguiente definición enunciaremos un conjunto de propiedades deseables para una medida de concordancia

Definición 1.6.11. Una medida de asociación κ entre dos variables aleatorias continuas X, Y cuya cópulas es C es una medida de concordancia si satisface las siguientes propiedades (la denotaremos como $\kappa_{X,Y}$ o κ_C):

- 1. κ está definida para todo par de variables aleatorias continuas;
- 2. $-1 \le \kappa_{X,Y} \le 1$, $\kappa_{X,X} = 1$, $y \kappa_{X,-X} = -1$;
- 3. $\kappa_{X,Y} = \kappa_{Y,X};$
- 4. Si X, Y son independientes, entonces $\kappa_{X,Y} = \kappa_{\Pi} = 0$;
- 5. $\kappa_{-X,Y} = \kappa_{X,-Y} = -\kappa_{X,Y};$
- 6. Si C_1 y C_2 son cópulas tales que $C_1(u, v) \leq C_2(u, v)$ para todo u, v en I, entonces $\kappa_{C_1} \leq \kappa_{C_2}$;
- 7. Si $\{(X_n, Y_n)\}$ es una sucesión de variables aleatorias continuas con cópulas C_n , y si $\{C_n\}$ converge puntualmente a C, entonces $\lim_{n\to\infty} \kappa_{C_n} = \kappa_C$.

Teorema 1.6.12. Sea κ una medida de concordancia para variables aleatorias continuas X, Y:

- 1. Si Y es casi seguramente una función creciente de X, entonces $\kappa_{X,Y} = \kappa_M = 1;$
- 2. Si Y es casi seguramente una función decreciente de X, entonces $\kappa_{X,Y} = \kappa_W = -1;$
- 3. Si α y β son casi seguramente funciones estrictamente monótonas en RanX y RanY, respectivamente, entonces $\kappa_{\alpha(X),\beta(Y)} = \kappa_{X,Y}$.

En el siguiente teorema veremos que la tau de Kendall y la rho de Spearman son medidas de concordancia de acuerdo a la definición anterior

Teorema 1.6.13. Si X, Y son variables aleatorias continuas cuya cópula es C, entonces la tau de Kendall (1.28)y la rho de Spearman (1.29) satisfacen las propiedades de la Definición 1.6.11 y Teorema 1.6.12 de una medida de concordancia.

Medidas de Dependencia

En la Definición 1.6.11 se presentó una lista de siete propiedades que una medida de concordancia debe cumplir, y vimos que la tau de Kendall y la rho de Spearman cumplen con estas propiedades. Sin embargo existe un defecto en estas medidas, ya que por la cuarta propiedad tenemos que si dos variables aleatorias continuas son independientes entonces la medida es igual a cero, pero de forma inversa no se cumple. Existen ejemplos en que la medida de concordancia es cero pero las variables aleatorias son dependientes. A continuación presentaremos un conjunto de propiedades deseables para una medida no paramétrica de dependencia de acuerdo con Nelsen (2006)

Definición 1.6.14. Una medida λ de asociación entre dos variables aleatorias continuas (X, Y) cuya cópula es C es una medida de dependencia si satisface las siguientes propiedades (la denotaremos como $\lambda_{X,Y}$ o λ_C):

- 1. λ está definida para cada par de variables aleatorias continuas X, Y;
- 2. $\lambda_{X,Y} = \lambda_{Y,X};$
- 3. $0 \leq \lambda_{X,Y} \leq 1;$
- 4. $\lambda_{X,Y} = 0$ si y sólo si X, Y son independientes;
- 5. $\lambda_{X,Y} = 1$ si y sólo si alguno de X, Y es casi seguramente una función estrictamente monótona de la otra;
- 6. Si α y β son functiones estrictamente monótonas crecientes en RanX y RanY, respectivamente, entonces $\lambda_{\alpha(X),\beta(Y)} = \lambda_{X,Y}$;
- 7. Si $\{(X_n, Y_n)\}$ es una sucesión de variables aleatorias continuas con cópulas C_n , y si $\{C_n\}$ converge puntualmente a C, entonces $\lim_{n\to\infty} \lambda_{C_n} = \lambda_C$.

Teorema 1.6.15. Para cualquier p en $[1, \infty)$ la distancia L_p entre una cópula $C y \prod$ dada por:

$$\mathcal{O}_{C}(p) = \left(k_{p} \int_{0}^{1} \int_{0}^{1} |C(u,v) = uv|^{p} \, du dv\right)^{1/p}$$
(1.36)

en donde

$$k_p = \frac{\Gamma \left(2p+3\right)}{2 \left(\Gamma \left(p+1\right)\right)^2}$$

es una medida de dependencia.

A continuación se presentarán las medidas de dependencias más estudiadas. **Definición 1.6.16.** 1. $\mathcal{V}_{C}^{2}(2)$ es el índice de dependencia de Hoeffding (1940) donde

$$\mathcal{O}_{C}^{2}(2) = \Phi_{X,Y}^{2} = \Phi_{C}^{2} = 90 \int \int_{I^{2}} |C(u,v) - uv|^{2} \, du dv \qquad (1.37)$$

2. $\mathcal{O}_{C}(1)$ es la medida de Schweizer y Wolff (1981) donde

$$\mho_{C}(1) = \sigma_{X,Y} = \sigma_{C} = 12 \int \int_{\mathbf{I}^{2}} |C(u,v) - uv| \, du dv.$$
(1.38)

3. Cuando $p = \infty$ donde

$$\mathfrak{V}_{C}\left(\infty\right) = \Lambda_{X,Y} = \Lambda_{C} = 4 \sup_{u,v \in I^{2}} \left| C\left(u,v\right) - uv \right|.$$
(1.39)

1.7. Cópula Empírica

La cópula empírica es la versión muestral de la cópula subyacente de variables aleatorias y puede ser usada para la estimación de la misma. Algunas de las aplicaciones de la cópula empírica son: obtener expresiones de la muestra para algunas medidas de asociación o para la construcción de pruebas de hipótesis no paramétricas, como son el caso de pruebas de independencia o bondad de ajuste para cópulas. En nuestro caso es la herramienta esencial para la construcción de Pruebas de Simetría.

Comenzaremos recordando la definición de función de distribución empírica para después definir de manera formal a la cópula empírica de acuerdo con Deheuvels (1979).

Definición 1.7.1. Sea una muestra aleatoria X_1, X_2, \ldots, X_n con función de distribución F. La función de distribución empírica se denota y define como:

$$F_{n}(x) = \frac{1}{n} \sum_{k=1}^{n} \mathbb{I}_{(-\infty,x]}(X_{k})$$
 (1.40)

donde I es la función indicadora definida como

$$\mathbb{I}_{A}(x) = \begin{cases} 1, & x \in A, \\ 0, & \text{en otro caso.} \end{cases}$$

Definición 1.7.2. Sea una muestra aleatoria $(X_1, Y_1), \ldots, (X_n, Y_n)$ con función de distribución conjunta H y marginales F y G. Sea C la cópula subyacente y F_n y G_n las funciones de distribución empíricas de F y G, respectivamente. La cópula empírica es la función \hat{C}_n dada por

$$\hat{C}_n\left(\frac{i}{n}, \frac{j}{n}\right) = \frac{1}{n} \sum_{k=1}^n \mathbb{I}_{\{nF_n(X_k) \le i, \ nG_n(Y_k) \le j\}}.$$
(1.41)

donde *i*, *j* están en $\{0, 1, ..., n\}$.

Y la frecuencia de la cópula empírica \hat{c}_n está dada por

$$\hat{c}_n\left(\frac{i}{n},\frac{j}{n}\right) = \begin{cases} 1/n, & \text{si } \left(X_{(i)},Y_{(j)}\right) \text{ es un elemento de la muestra,} \\ 0, & \text{en otro caso.} \end{cases}$$
(1.42)

Donde $X_{(i)}$ y $Y_{(j)}$, $1 \leq i, j \leq n$ denotan los estadísticos de orden de la muestra.

Observación 1.7.3. Nótese que sin pérdida de generalidad podemos suponer que $X_1 < X_2 < \ldots < X_n$ y la cópula empírica (1.41) se reduce a

$$\hat{C}_n\left(\frac{i}{n}, \frac{j}{n}\right) = \frac{1}{n} \sum_{k=1}^n \mathbb{I}_{\{nG_n(Y_k) \le j\}}.$$
(1.43)

 $y \ que \ \hat{C}_n \ y \ \hat{c}_n \ están \ relacionadas \ vía$

$$\hat{C}_n\left(\frac{i}{n}, \frac{j}{n}\right) = \sum_{p=1}^i \sum_{q=1}^j \hat{c}_n\left(\frac{p}{n}, \frac{q}{n}\right)$$
(1.44)

$$\hat{c}_n\left(\frac{i}{n}, \frac{j}{n}\right) = \hat{C}_n\left(\frac{i}{n}, \frac{j}{n}\right) - \hat{C}_n\left(\frac{i-1}{n}, \frac{j}{n}\right) - \hat{C}_n\left(\frac{i}{n}, \frac{j-1}{n}\right) + \hat{C}_n\left(\frac{i-1}{n}, \frac{j-1}{n}\right).$$
(1.45)

A continuación introduciremos algunos conceptos de acuerdo con Genest y otros (2011).

Definición 1.7.4. Sea una muestra aleatoria $(X_1, Y_1), \ldots, (X_n, Y_n)$ de una distribución conjunta H con marginales F y G. Sea C la cópula subyacente. Cuando F y G son conocidas podemos construir una muestra aleatoria de C haciendo para todo i en $\{1, \ldots, n\}$

$$(U_i, V_i) = (F(X_i), G(Y_i))$$
 (1.46)

a las cuales denominaremos observaciones de la cópula.

Sin embargo, cuando $F \ge G$ son desconocidos no es posible obtener observaciones de C, por lo que introducimos la siguiente definición:

Definición 1.7.5. Sea una muestra aleatoria $(X_1, Y_1), \ldots, (X_n, Y_n)$ de una distribución conjunta H con marginales F, G (desconocidas) y cópula C. Definimos a las pseudo observaciones de la cópula para todo i en $\{1, \ldots, n\}$ como

$$\left(\hat{U}_{i},\hat{V}_{i}\right)=\left(F_{n}\left(X_{i}\right),G_{n}\left(Y_{i}\right)\right).$$
(1.47)

Donde F_n y G_n son las funciones de distribución empíricas de F y G respectivamente.

Obsrvación 1.7.6. En la Definición 1.7.2 definimos a la cópula empírica $C_n\left(\frac{i}{n}, \frac{j}{n}\right)$ sólo para i, j en $\{0, 1, \ldots, n\}$ entonces podemos extender ésta definición para u, v en **I** de la siguiente manera

$$\hat{C}_n(u,v) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}\left(\hat{U}_i \le u, \hat{V}_i \le v\right).$$
 (1.48)

Como habíamos dicho, un uso de la cópula empírica es darnos versiones muestrales de algunas medidas de asociación. El siguiente teorema nos da la expresión de la versión muestral de la tau de Kendall (1.28) y la rho de Spearman (1.29). Para más detalles véase Nelsen (2006).

Teorema 1.7.7. Sean \hat{C}_n y \hat{c}_n la cópula empírica y la frecuencia de la cópula empírica, respectivamente, de una muestra $(X_1, Y_1), \ldots, (X_n, Y_n)$. Si r y t denotan las versiones muestrales de la tau de Kendall y la rho de Spearman entonces:

$$r = \frac{12}{n^2 - 1} \sum_{i=1}^{n} \sum_{j=1}^{n} \left[\hat{C}_n \left(\frac{i}{n}, \frac{j}{n} \right) - \frac{i}{n} \cdot \frac{j}{n} \right],$$
(1.49)

$$t = \frac{2n}{n-1} \sum_{i=2}^{n} \sum_{j=2}^{n} \sum_{p=1}^{i-1} \sum_{q=1}^{j-1} \left[\hat{c}_n\left(\frac{i}{n}, \frac{j}{n}\right) \hat{c}_n\left(\frac{p}{n}, \frac{q}{n}\right) - \hat{c}_n\left(\frac{i}{n}, \frac{q}{n}\right) \hat{c}_n\left(\frac{p}{n}, \frac{j}{n}\right) \right]$$
(1.50)

1.8. Simulación

En esta sección discutiremos el problema de generar una muestra proveniente de una distribución conjunta dada. Estas muestras se pueden usar para estudiar modelos matemáticos de sistemas en el mundo real o para estudios estadísticos. En nuestro caso la simulación de vectores aleatorios nos será útil para la comparación de dos pruebas de simetría, las cuales se definirán en el siguiente capítulo.

Antes de enunciar los procedimientos para generar vectores aleatorios, recordemos un algoritmo para el caso de una variable denominado el método de la "Transformada Inversa".

Alogoritmo 1. Para obtener una observación x de una variable aleatoria X con función de distribución F:

- 1. Generar una variable aleatoria u que sea uniforme (0,1);
- 2. Calcular $x = F^{(-1)}(u)$, donde $F^{(-1)}$ es cualquier cuasi inversa de F (ver Definición 1.3.2).

Hay una gran variedad de procedimientos para generar observaciones (x, y) de un par de variables aleatorias (X, Y) con función de distribución conjunta H, en esta sección ocuparemos a las cópulas como herramienta. Como consecuencia del Teorema de Sklar, sólo tenemos que generar observaciones (u, v) de las variables (U, V) con distribución marginal uniforme (0, 1) cuya función de distribución conjunta es C, la cópula de (X, Y), y transformar dichas observaciones con el Algoritmo 1. Un procedimiento general para generar las observaciones (u, v) es con el método de la distribución condicional. Para dicho método, necesitamos la distribución condicional de V dado U = u.

Definición 1.8.1. Sean $U \ y \ V$ dos variables aleatorias distribuidas uniformemente en $(0,1) \ y$ con función de distribución conjunta C. La distribución condicional de V dado U = u la denotaremos como $c_u(v) \ y$ está dada por:

$$c_{u}(v) = P\left[V \le v | U = u\right] = \lim_{\Delta u \to 0} \frac{C\left(u + \Delta u, v\right) - C\left(u, v\right)}{\Delta u} = \frac{\partial C\left(u, v\right)}{\partial u}.$$
(1.51)

Recordemos, por el Teorema 1.2.7, que la función $c_u(v)$ existe y es no decreciente en casi todo I.

Algoritmo 2. Para generar observaciones del vector aleatorio (U, V) con función de distribución conjunta C y marginales uniformes (0, 1):

- 1. Generar dos variables independientes uniformes (0,1) u y t;
- 2. Hacer $v = c_u^{(-1)}(t)$, donde $c_u^{(-1)}$ denota la cuasi inversa de c_u .

3. El par deseado es (u, v).

Para el caso de la familia de las cópulas Arquimedianas existen los dos siguientes algoritmos:

Algoritmo 3. Para generar observaciones (u, v) cuya función de distribución conjunta es una cópula Arquimediana C con generador φ :

- 1. Generar dos variables independientes uniformes (0,1) u y t;
- 2. Hacer $w = K_C^{(-1)}(t)$ donde K_C está dado como:

$$K_C(t) = t - \frac{\varphi(t)}{\varphi'(t^+)}$$
(1.52)

- 3. Hacer $u = \varphi^{[-1]}(s\varphi(w)) \ y \ v = \varphi^{[-1]}((1-s)\varphi(w));$
- 4. El par deseado es (u, v).

Algoritmo 4. Para generar observaciones (u, v) cuya función de distribución conjunta es una cópula Arquimediana C con generador φ :

- 1. Generar dos variables independientes uniformes (0,1) u y t;
- 2. Hacer $w = \varphi'^{(-1)} (\varphi'(u)/t);$
- 3. Hacer $v = \varphi^{[-1]} (\varphi(w) \varphi(u));$
- 4. El par deseado es (u, v).

Los algoritmos anteriores generan de manera general observaciones provenientes de cualquier cópula Arquimediana; el siguiente algoritmo genera de manera particular observaciones de la cópula Clayton dada en el Cuadro 1.5.1

Algoritmo 5. Para generar observaciones (u, v) cuya función de distribución conjunta es una cópula Clayton con parámetro $\theta > 0$:

- 1. Generar dos variables aletorias exponenciales (media $\mu = 1$) independientes x y y; y una variable z con distribución gamma ($\alpha = \theta, \beta = 1$) independiente de x y y;
- 2. Hacer $u = [1 + (x/z)]^{-\theta}$ y $v = [1 + (y/z)]^{-\theta}$;
- 3. El par deseado es (u, v).

El siguiente algoritmo genera observaciones de la cópula Marshall-Olkin (1.13):

Algoritmo 6. Para generar observaciones (u, v) cuya función de distribución conjunta es una cópula Marshall-Olkin con parámetros $0 < \alpha, \beta, < 1$:

1. Generar tres variables uniformes (0,1) independientes r, s y t;

2. Hacer
$$u = exp\left(-\min\left(\frac{-\ln r}{1-\alpha}, \frac{-\ln t}{\alpha}\right)\right) y$$

 $v = exp\left(-\min\left(\frac{-\ln s}{1-\beta}, \frac{-\ln t}{\beta}\right)\right);$

3. El par deseado es (u, v).

El siguiente algoritmo genera observaciones de la cópula Ali-Mikhail-Haq (1.22):

Algoritmo 7. Para generar observaciones (u, v) cuya función de distribución conjunta es una cópula de la familia Ali-Mikhail-Haq con parámetro θ :

- 1. Generar dos variables uniformes (0,1) independientes u, t;
- 2. Hacer a = 1 u; $b = -\theta (2at + 1) + 2\theta^2 a^2 t + 1$; $y = c = \theta^2 (4a^2t 4at + 1) \theta (4at 4t + 2) + 1$;
- 3. Hacer $v = 2t (a\theta 1)^2 / (b + \sqrt{c});$
- 4. El par deseado es (u, v).

Observación 1.8.2. Sean (U_1, V_1) y (U_2, V_2) vectores aleatorios con función de distribución conjunta C_1 y C_2 , respectivamente y sea X una variable aleatoria Bernoulli con parámetro α independiente de (U_1, V_1) y (U_2, V_2) , definamos al vector aleatorio (U, V) como

$$(U,V) = \begin{cases} (U_1, V_1), & \text{si } X = 1, \\ (U_2, V_2), & \text{si } X = 0 \end{cases}$$
(1.53)

entonces

$$C(u, v) = P[U \le u, V \le v]$$

= $P[U \le u, V \le v| (U, V) = (U_1, V_1)] * P[(U, V) = (U_1, V_1)]$
+ $P[U \le u, V \le v| (U, V) = (U_2, V_2)] * P[(U, V) = (U_2, V_2)]$
= $P[U_1 \le u, V_1 \le v] * P[X = 1] + P[U_2 \le u, V_2 \le v] * P[X = 0]$
= $\alpha C_1 + (1 - \alpha) C_2$

En la Proposición 1.2.9 vimos que la combinación lineal convexa de cópulas también es cópula y, con ayuda de la Observación 1.8.2, si queremos simular un vector aleatorio proveniente de alguna cópula de la forma $C = \alpha C_1 + (1 - \alpha) C_2$ hay que seguir el siguiente procedimiento:

Algoritmo 8. Para generar observaciones (u, v) cuya función de distribución conjunta es una cópula C que es combinación lineal convexa de dos cópulas, es decir, si $C = \alpha C_1 + (1 - \alpha) C_2$ con α en **I**:

- 1. Generar de manera independiente observaciones $(u_1, v_1) y (u_2, v_2)$ provenientes de las cópulas $C_1 y C_2$, respectivamente
- 2. Generar una observación x de una variable aleatoria Bernoulli con parámetro α independiente de (u_1, v_1) y (u_2, v_2) ;
- 3. Hacer $u = xu_1 + (1 x)u_2$ $y v = xv_1 + (1 x)v_2$
- 4. El par deseado es (u, v).

Para hacer simulaciones de una cópula $C_{1,2,\theta}$, que es un pegado a partir de otras dos, es decir que la cópula está dada por (1.9) sólo es necesario hacer uso del Algoritmo 2 y el hecho de que:

$$\frac{\partial}{\partial u}C_{1,2,\theta}\left(u,v\right) = \begin{cases} \frac{\partial}{\partial u}C_1\left(\frac{u}{\theta},v\right), & 0 \le u \le \theta, \\ \frac{\partial}{\partial u}C_2\left(\frac{u-\theta}{1-\theta},v\right), & \theta \le u \le 1. \end{cases}$$

Algoritmo 9. Para generar observaciones (u, v) cuya función de distribución conjunta es una cópula $C_{1,2,\theta}$ que fue construida con la técnica de pegado de las cópulas $C_1 y C_2$:

- 1. Generar una variable aleatoria u uniforme (0,1)
- 2. Generar una observación $\left(\frac{u}{\theta}, v\right)$ proveniente de la cópula C_1 si $0 \le u \le \theta$ ó $\left(\frac{u-\theta}{1-\theta}, v\right)$ proveniente de la cópula C_2 si $\theta \le u \le 1$;
- 3. El par deseado es (u, v).

Capítulo 2

Simetría

En este capítulo revisaremos el concepto de intercambiabilidad de variables aleatorias y su importancia, debido a sus aplicaciones en varios campos de la estadística. También, veremos la relación entre el concepto de intercambiabilidad de variables aleatorias y el de simetría en cópulas; analizaremos algunos conceptos y medidas de simetría propuestos hasta el momento; y para finalizar, se mostrarán algunas propuestas de pruebas de hipótesis para validar simetría en cópulas bivariadas, tales como la de Erdely y González-Barrios (2010) y Genest y otros (2011).

2.1. Conceptos Generales de Simetría

En esta sección revisaremos algunos conceptos de simetría de variables aleatorias, comenzando por algunas definiciones que podemos encontrar en Nelsen (2006).

Definición 2.1.1. Sea X una variable aleatoria y a un número real. Decimos que X es simétrica respecto de a si las funciones de distribución de las variables aleatorias X - a y a - X son la misma. Esto es, si para cualquier x en \mathbf{R} , $P[X - a \le x] = P[a - X \le x]$. Cuando X es continua con función de distribución F, esto es equivalente a

$$F(a+x) = 1 - F(a-x).$$
(2.1)

Cuando F es discontinua, (2.1) se cumple sólo para los puntos de continuidad de F.

Si consideramos ahora el caso bivariado, ¿Cuándo podemos decir que un vector aleatorio (X, Y) es simétrico respecto del punto (a, b)? Hay diferentes

formas de responder a esta pregunta, y cada respuesta nos lleva a diferentes tipos de simetría bivariada.

Definición 2.1.2. Sean X, Y variables aleatorias y sea (a,b) un punto en \mathbb{R}^2 .

- 1. (X, Y) es marginalmente simétrico respecto de (a, b) si X, Y son simétricas respecto de a y b, respectivamente.
- 2. (X, Y) es radialmente simétrico respecto de (a, b) si la función de distribución conjunta de X - a y Y - b es la misma que la función de distribución conjunta de a - X y b - Y.
- 3. (X, Y) es conjuntamente simétrico respecto de (a, b) si los siguientes cuatro pares de variables aleatorias tienen la misma función de distribución conjunta: (X - a, Y - b), (X - a, b - Y), (a - X, Y - b) y(a - X, b - Y),

Teorema 2.1.3. Sean X, Y variables aleatorias continuas con función de distribución conjunta H y marginales F y G respectivamente. Sea (a,b) un punto en \mathbb{R}^2 . Entonces (X,Y) es radialmente simétrico respecto de (a,b) si y sólo si para todo (x,y) en \mathbb{R}^2

$$H(a + x, b + y) = 1 - F(a - x) - G(b - y) + H(a - x, b - y)$$
(2.2)

Observación 2.1.4. Es inmediato que simetría conjunta implica simetría radial y es fácil ver que simetría radial implica simetría marginal (haciendo $x = \infty$ en (2.2) nos queda (2.1), de forma similar para $y = \infty$). Por lo que la simetría conjunta es una condición muy fuerte, sin embargo, es fácil mostrar que si dos variables aleatorias son conjuntamente simétricas, éstas tienen que estar no correlacionadas, es decir, el coeficiente de correlación de pearson es igual a cero (cuando hay existencia de segundos momentos).

Teorema 2.1.5. Sean X, Y variables aleatorias continuas con función de distribución conjunta H, distribuciones marginales F y G respectivamente y cópula C. Supongamos que X, Y son simétricas respecto de a y b, respectivamente. Entonces (X, Y) es radialmente simétrico respecto de (a, b) si y sólo si C satisface la siguiente ecuación:

$$C(u,v) = u + v - 1 + C(1 - u, 1 - v)$$
 para todo (u,v) en I^2 . (2.3)

Otra forma muy importante de simetría es la intercambiabilidad de variables aleatorias **Definición 2.1.6.** Decimos que las variables aleatorias X, Y son intercambiables si los vectores (X, Y) y (Y, X) son idénticamente distribuidos. Es decir, si la función de distribución conjunta de (X, Y) es H, entonces H(x, y) = H(y, x) para todo x, y en $\overline{\mathbf{R}}$.

Citando a Nelsen (2007) con respecto a la definición anterior

Desde su introducción por De Finetti en 1930, las variables aleatorias intercambiables han encontrado importantes aplicaciones en varias áreas de estadística: leyes de límites, teoría de valores extremos, estadística Bayesiana, procesos estocásticos, etc.

Véase Galambos (1982) para un estudio de propiedades y aplicaciones de variables aleatorias intercambiables.

Definición 2.1.7. *Decimos que una cópula* C *es simétrica si para todo* (u, v) *en* I^2 , C(u, v) = C(v, u)

Teorema 2.1.8. Sean X, Y variables aleatorias continuas con función de distribución conjunta H, cópula C y marginales F y G, respectivamente. Entonces X, Y son intercambiables si y sólo si F = G y C es simétrica.

Es decir, para que dos variables aleatorias sean intercambiables, tienen que suceder dos cosas: primero, las variables aleatorias tienen que ser idénticamente distribuidas; segundo, la cópula subyacente tiene que ser simétrica. En ese sentido, sólo necesitamos que una de éstas dos condiciones no se cumpla para poder afirmar que dos variables son no intercambiables.

Ya definido el concepto de simetría para cópulas, veremos un concepto menos general propuesto por Erdely y González-Barrios (2010), denominado tri-simetría.

Definición 2.1.9. Decimos que una cópula C es tri-simétrica si:

$$C(2/3, 1/3) = C(1/3, 2/3)$$
(2.4)

Observación 2.1.10. Hay que notar que si una cópula es simétrica entonces es tri-simétrica, pero, si una cópula es tri-simétrica, no necesariamente es simétrica. Es decir, simetría implica tri-simetría pero tri-simetría no implica simetría.

2.2. Medidas de No-Intercambiabilidad

Comenzaremos esta sección definiendo las propiedades deseables para una medida de no-intercambiabilidad, según Durante, Klement, Sempi y Úbeda-Flores (2010), para después mostrar algunas medidas que cumplen con dichas propiedades.

Sea $\mathcal{H}(F)$ la clase de vectores aleatorios (X, Y) tal que X, Y son idénticamente distribuidos con función de distribución continua F. Observemos que esta definición no depende del espacio de probabilidad en que los vectores aleatorios están definidos.

Definición 2.2.1. Una función $\hat{\mu} : \mathcal{H}(F) \mapsto \mathbf{R}_+$ es una medida de nointercambiabilidad para $\mathcal{H}(F)$ si satisface las siguientes propiedades:

- 1. $\hat{\mu}$ está acotada, es decir, existe K en **R** tal que para todo (X,Y) en $\mathcal{H}(F), \hat{\mu}(X,Y) \leq K;$
- 2. $\hat{\mu}(X,Y) = 0$ si y sólo si (X,Y) es intercambiable;
- 3. $\hat{\mu}$ es simétrica, es decir, para cualquier (X, Y) en $\mathcal{H}(F)$ $\hat{\mu}(X, Y) = \hat{\mu}(Y, X);$
- 4. $\hat{\mu}(X,Y) = \hat{\mu}(f(X), f(Y))$ para toda función estrictamente monótona f y para todo (X,Y) en $\mathcal{H}(F)$;
- 5. Si (X_n, Y_n) y (X, Y) son vectores aleatorios con función de distribución conjunta H_n y H, respectivamente, y si H_n converge débilmente a H, entonces $\hat{\mu}(X_n, Y_n)$ converge a $\hat{\mu}(X, Y)$.

En vista del Teorema de Sklar (Teorema 1.3.1), es posible establecer una correspondencia uno a uno entre el conjunto $\mathcal{H}(F)$ y el conjunto \mathcal{C} de todas las cópulas. De hecho, para cada (X, Y) en $\mathcal{H}(F)$ podemos asociar de manera única la cópula C de (X, Y) dada por $C(u, v) = H\left(F^{(-1)}(u), F^{(-1)}(v)\right)$, donde H es la función de distribución conjunta de (X, Y) y $F^{(-1)}$ es una cuasi inversa de F. Inversamente, para cada cópula C en \mathcal{C} podemos asociar de manera única la función de distribución conjunta H definida en $\mathbf{\bar{R}}^2$ como H(x,y) = C(F(x), F(y)). Esta correspondencia nos permite definir una medida de no-intercambiabilidad directamente en el conjunto \mathcal{C} de todas las cópulas.

Definición 2.2.2. Una función $\mu : C \mapsto \mathbf{R}_+$ es una medida de no-intercambiabilidad para C si satisface las siguientes propiedades:

- 1. Existe un K en \mathbf{R}_+ tal que, para todo C en \mathcal{C} , $\mu(C) \leq K$;
- 2. $\mu(C) = 0$ si y sólo si C es simétrica;
- 3. $\mu(C) = \mu(C^{t})$ para todo C en C, donde $C^{t}(u, v) = C(v, u)$
- 4. $\mu(C) = \mu(\hat{C})$ para todo C en C, donde $\hat{C}(u,v) = u + v 1 + C(1-u, 1-v);$
- 5. Si $\{C_n\}$ y C están en C, y si $\{C_n\}$ converge uniformemente a C, entonces $\{\mu(C_n)\}$ converge a $\mu(C)$.

Ahora consideraremos una amplia clase de medidas de no-intercambiabilidad

Definición 2.2.3. Sea la función $\mu_p(C) : C \mapsto \mathbf{R}_+$ definida como la distancia L_p entre $C \ y \ C^t$ para cada $p \ en \ [1, \infty]$, es decir

$$\mu_p(C) = \left(\int_0^1 \int_0^1 |C(u,v) - C(v,u)|^p \, du dv\right)^{1/p} \tag{2.5}$$

para p en $[1,\infty)$, y para $p = +\infty$

$$\mu_{\infty}(C) = \sup_{(u,v) \in I^{2}} |C(u,v) - C(v,u)|$$
(2.6)

Teorema 2.2.4. La función μ_p es una medida de no-intercambiabilidad para todo p en $[1, \infty]$

Proposición 2.2.5. Sea $\mu : \mathcal{C} \mapsto \mathbf{R}_+$ una medida de no-intercambiabilidad. Entonces existe una constante K_{μ} en \mathbf{R}_+ y una cópula C_{μ} tal que $\mu(C_{\mu}) = K_{\mu}$ y $\mu(C_{\mu}) \ge \mu(C)$ para toda cópula $C \ne C_{\mu}$.

Lo que significa que la cota en el punto uno de la Definición 2.2.2 es siempre alcanzable.

Lema 2.2.6. Para cualquier cópula C y cualesquiera u, v en I,

$$|C(u,v) - C(v,u)| \le \min(u,v,1-u,1-v,|u-v|).$$
(2.7)

Teorema 2.2.7. Para cualquier cópula C,

$$\sup_{(u,v)\in I^2} |C(u,v) - C(v,u)| \le \frac{1}{3},$$
(2.8)

y la desigualdad es lo mejor posible.

Entonces $K_{\mu_{\infty}} = \frac{1}{3}$. Para los valores K_{μ_p} cuando p está en $[1, \infty)$ el siguiente resultado se cumple.

Proposición 2.2.8. Para cada p en $[1, \infty)$ y cada C en C tenemos

$$\mu_p(C) \le \left(\frac{2 \cdot 3^{-p}}{(p+1)(p+2)}\right)^{1/p}.$$
(2.9)

Por el resultado anterior podemos suponer, sin pérdida de generalidad, que cualquier medida de no-intercambiabilidad está normalizada, es decir, que su rango es I, dividiendo μ entre K_{μ} . Por lo que motivamos a la siguiente definición.

Definición 2.2.9. Sea $\mu : C \mapsto I$ una medida de no-intercambiabilidad y C una cópula. Decimos que C es máxima no-intercambiable con respecto a μ si $\mu(C) = 1$.

Observación 2.2.10. Notemos que una medida de no-intercambiabilidad $\mu : \mathcal{C} \mapsto \mathbf{R}_+$ induce una relación \prec_{μ} en \mathcal{C} , definida de la siguiente manera: Para todo A, B en \mathcal{C}

$$A \prec_{\mu} B$$
 si y sólo si $\mu(A) \leq \mu(B)$

Ésta relación \prec_{μ} es reflexiva y transitiva, pero no necesariamente antisimétrica, es decir, es un cuasi-orden. Por otra parte, podemos definir una relación \mathcal{R}_{μ} en el conjunto \mathcal{C} de esta forma: Para todo A, B en \mathcal{C}

$$A \mathcal{R}_{\mu}B$$
 si y sólo si $\mu(A) = \mu(B)$

Es fácil probar que \mathcal{R}_{μ} es un relación de equivalencia y entonces induce una partición en \mathcal{C} .

Ahora daremos otros ejemplos de medidas de no-intercambiabilidad

Ejemplo 2.2.11. Dadas las medidas μ_{p_1} , μ_{p_2} podemos obtener que su producto y sus combinaciones convexas también son medidas de no-intercambiabilidad. Otro ejemplo está dado por

$$\mu(C) = \mu_p(C) + |C(a, 1 - a) - C(1 - a, 1)|,$$

donde a es un punto fijo en (0,1).

A partir de ahora centraremos nuestra atención en la versión normalizada de la medida de no-intercambiabilidad μ_{∞} definida en (2.6), que también podemos encontrar en Nelsen (2007).

Definición 2.2.12. Sea C una cópula. El grado de no-intercambiabilidad de C respecto a μ_{∞} , denotado por $\delta(C)$, está dado por

$$\delta(C) = 3 \cdot \mu_{\infty}(C) = 3 \left(\sup_{(u,v) \in I^2} |C(u,v) - C(v,u)| \right)$$
(2.10)

Cabe mencionar que el grado de no-intercambiabilidad de una cópula C se puede generalizar para toda medida de no-intercambiabilidad μ . Hay que notar que el grado de no-intercambiabilidad de una cópula C también puede verse como una medida de asimetría de la cópula.

A partir de este momento sólo nos referiremos a la medida de no-intercambiabilidad $\delta(C)$, por lo que, haciendo referencia a la Definición 2.2.9, si Ces una cópula máxima no-intercambiable con respecto a $\delta(C)$, sólo diremos que C es máxima no-intercambiable.

Definición 2.2.13. Definimos a \mathcal{M}_1 como el conjunto de cópulas que son máxima no-intercambiables, es decir, $\mathcal{M}_1 = \{C | \delta(C) = 1\}$

Teorema 2.2.14. Para cada θ en I hay una cópula C_{θ} tal que $\delta(C_{\theta}) = \theta$

Ejemplo 2.2.15. Sea C_s una cópula simétrica y C_1 una cópula máxima no-intercambiable, es decir, C_1 está en \mathcal{M}_1 . Sea $C = \alpha C_1 + (1 - \alpha) C_s$ con α en I, entonces

$$\begin{split} \delta\left(C\right) &= 3 \sup_{(u,v) \in I^{2}} \left|C\left(u,v\right) - C\left(v,u\right)\right| \\ &= 3 \sup_{(u,v) \in I^{2}} \left|\alpha C_{1}\left(u,v\right) + (1-\alpha) C_{s}\left(u,v\right) - \alpha C_{1}\left(v,u\right) - (1-\alpha) C_{s}\left(v,u\right)\right| \\ &= 3 \sup_{(u,v) \in I^{2}} \left|\alpha \left(C_{1}\left(u,v\right) - C_{1}\left(v,u\right)\right) + (1-\alpha) \left(C_{s}\left(u,v\right) - C_{s}\left(v,u\right)\right)\right| \\ &= 3 \sup_{(u,v) \in I^{2}} \left|\alpha \left(C_{1}\left(u,v\right) - C_{1}\left(v,u\right)\right)\right| \\ &= \alpha \cdot 3 \sup_{(u,v) \in I^{2}} \left|C_{1}\left(u,v\right) - C_{1}\left(v,u\right)\right| \\ &= \alpha \cdot \delta\left(C_{1}\right) \\ &= \alpha \end{split}$$

Ejemplo 2.2.16. Dado θ en [0, 1/3] definimos las siguientes cópulas

$$M_{\theta}(u,v) = \min\{u, v, (u-1+\theta)_{+} + (v-\theta)_{+}\}, \qquad (2.11)$$

$$W_{\theta}(u,v) = \max\left\{0, u+v-1, \theta - (\theta - u)_{+} - (1 - \theta - v)_{+}\right\}, \qquad (2.12)$$

donde $x_+ = \max\{x, 0\}$. De forma específica, M_{θ} es una cópula cuyo soporte consiste en dos segmentos de línea en \mathbf{I}^2 , uno que conecta $(0, \theta)$ a $(1 - \theta, 1)$ y el otro que conecta $(1 - \theta, 0)$ y $(1, \theta)$; W_{θ} es una cópula cuyo soporte consiste en tres segmentos de línea en \mathbf{I}^2 , uno que conecta $(0, 1 - \theta)$ a (θ, θ) , otro que conecta $(\theta, 1)$ a $(1 - \theta, 1 - \theta)$ y otro que conecta $(1 - \theta, \theta)$ a (1, 0)Véase Figura 2.2.1. Se puede probar que para cada θ en [0, 1/3], $\delta(M_{\theta}) =$ $\delta(W_{\theta}) = 3\theta$.

La familia de cópulas antes mencionada fue propuesta en Durante y otros (2010), pero previamente los casos particulares $M_{1/3}$ y $W_{1/3}$ fueron estudiados y son parte fundamental en el trabajo de Nelsen (2007). La cópula $M_{1/3}$ será retomada en el siguiente capítulo.



Figura 2.2.1: Soportes de $M_{1/3}$ y $W_{1/3}$

Definiremos cuatro conjuntos de cópulas con los cuales describiremos al conjunto de cópulas máxima no-intercambiables

Definición 2.2.17. Sean C_1 , C_2 , C_3 y C_4 los siguientes conjuntos de cópulas:

$$C_{1} = \{C|C(2/3, 1/3) = 0\}, C_{2} = \{C|C(1/3, 2/3) = 1/3\}, C_{3} = \{C|C(1/3, 2/3) = 0\}, C_{4} = \{C|C(2/3, 1/3) = 1/3\},$$

Notemos que C_1 y C_4 son disjuntos, al igual que C_2 y C_3 ; una cópula Cestá en C_1 (C_2) si y sólo si C^t está en C_3 (C_4); además $M_{1/3}$ dada en (2.11) y $W_{1/3}$ dada en (2.12) están en $C_1 \cap C_2$.

Teorema 2.2.18. Sea \mathcal{M}_1 el conjunto de cópulas máxima no-intercambiables y C_1 , C_2 , C_3 y C_4 como en la Definición 2.2.17, entonces:

- (a) $\mathcal{M}_1 = (C_1 \cap C_2) \cup (C_3 \cap C_4)$, es decir, una cópula C está en \mathcal{M}_1 si y sólo si C(1/3, 2/3) = 1/3 y C(2/3, 1/3) = 0 ó C(1/3, 2/3) = 0 y C(2/3, 1/3) = 1/3
- (b) Para cada C en \mathcal{M}_1 , C está en $C_1 \cap C_2$ si y sólo si $W_{1/3}(u,v) \leq C(u,v) \leq M_{1/3}(u,v)$ para todo u, v en \mathbf{I} , $y \in C$ está en $C_3 \cap C_4$ si y sólo si $W_{1/3}^t(u,v) \leq C(u,v) \leq M_{1/3}^t(u,v)$ para todo u, v en \mathbf{I}
- (c) Toda cópula máxima no-intercambiable tiene la misma sección diagonal, es decir, para cada C en \mathcal{M}_1 , $C(u, u) = (u - 1/3)_+ + (u - 2/3)_+$.

Observación 2.2.19.

- (a) En consecuencia del Teorema 2.2.18, un tercio de la masa de probabilidad asociada a una cópula máxima no-intercambiable en C₁ ∩ C₂ está contenido en cada uno de los cuadrados [0,1/3] × [1/3,2/3], [1/3,2/3], × [2/3,1], y [2/3,1] × [0,1/3]; y un tercio de la masa de probabilidad asociada a una cópula máxima no-intercambiable en C₃ ∩ C₄ está contenido en cada uno de los cuadrados [0,1/3] × [2/3,1], [1/3,2/3], [1/3,2/3], y [2/3,1] × [1/3,2/3].
- (b) En consecuencia directa del Teorema 2.2.18 inciso (a) dada una cópula C, el valor |C (2/3, 1/3) - C (1/3, 2/3)| determina si C es máxima nointercambiable o totalmente asimétrica.

Ejemplo 2.2.20. Sea c^{3S} la densidad de una cópula definida como

$$c^{3s}(u,v) = \begin{cases} 0, & (u,v) \in I_1 \cup I_3 \cup I_4 \cup I_5 \cup I_8 \cup I_9, \\ 3, & (u,v) \in I_2 \cup I_6 \cup I_7, \end{cases}$$
(2.13)

donde $I_1 = (0, 1/3] \times (0, 1/3], I_2 = (1/3, 2/3] \times (0, 1/3], I_3 = (2/3, 1] \times (0, 1/3], I_4 = (0, 1/3] \times (1/3, 2/3], I_5 = (1/3, 2/3] \times (1/3, 2/3], I_6 = (2/3, 1] \times (1/3, 2/3], I_7 = (0, 1/3] \times (2/3, 1], I_8 = (1/3, 2/3] \times (2/3, 1] y I_9 = (2/3, 1] \times (2/3, 1], véase figura 2.2.2.$

Definimos para todo (u, v) en I^2 , $C^{3S}(u, v) = \int_0^1 \int_0^1 c^{3S}(s, t) \, ds dt$, entonces integrando y simplificando tenemos que

$$C^{3S}(u,v) = \begin{cases} 0, & si \ (u,v) \in I_1 \cup I_4, \\ (3u-1)v, & si \ (u,v) \in I_2, \\ v, & si \ (u,v) \in I_3, \\ u-1/3, & si \ (u,v) \in I_5, \\ 1/3 + (u-2/3)(3v-1), & si \ (u,v) \in I_6, \\ (3v-2)u, & si \ (u,v) \in I_7, \\ u+v-1, & si \ (u,v) \in I_8 \cup I_9. \end{cases}$$
(2.14)

Ésta cópula máxima no intercambiable fue construida por Erdely y González-Barrios (2010) y es denominada como cópula "3-squared".

1			
2	I ₇	I ₈	I ₉
3	I ₄	I ₅	Ι ₆
3	I ₁	I ₂	I ₃
	1	L 2 3 3	1

Figura 2.2.2: Regiones I_1, \ldots, I_9 de la densidad (2.13)

Proposición 2.2.21. Sea C(u, v) la cópula asociada al vector aleatorio (U, V) con marginales uniformes (0, 1) tal que

$$|C(2/3, 1/3) - C(1/3, 2/3)| = 0$$

Definimos I_j , j = 1, 2, ..., 9 como en el Ejemplo 2.2.20 (véase Figura 2.2.2), y sea $p_j = P(\{(U, V) \in I_j\})$, entonces

$$p_2 = p_4, \ p_3 = p_7, \ p_6 = p_8$$

Proposición 2.2.22. Sea C una cópula tri-simétrica. Entonces

$$\delta(C) = 3 \sup_{(u,v) \in I^2} |C(u,v) - C(v,u)| \le \frac{1}{2}.$$
(2.15)

y la desigualdad es la mejor posible.

Las cópulas $W ext{ y } M$, ambas simétricas, son respectivamente la cota superior e inferior del conjunto de cópulas simétricas, es decir, aquellas cópulas C tal que $\delta(C) = 0$. Ahora veremos las mejores cotas posibles del conjunto de las cópulas máxima no-intercambiables, es decir, las mejores cotas para el conjunto $\mathcal{M}_1 = \{C | \delta(C) = 1\}$. Para esto tenemos que considerar el concepto de cuasi-cópula (el cual es más general que el de cópula).

Definición 2.2.23. Una cuasi-cópula es una función $Q : \mathbf{I}^2 \mapsto \mathbf{I}$ que satisface las condiciones (1.2) y (1.3) de la definición de cópula, pero, en lugar de (1.4) satisface condiciones más débiles:

- 1. Q es no decreciente en cada argumento,
- 2. Q satisface la condición de Lipschitz

$$|Q(u_1, v_1) - Q(u_2, v_2)| \le |u_1 - u_2| + |v_1 - v_2|$$

para todo (u_1, v_1) y (u_2, v_2) en I^2 .

Teorema 2.2.24. Para todo C en \mathcal{M}_1 , tenemos que $B \leq C \leq A$, donde

$$B(u,v) = \begin{cases} \min\{u-1/3, v-1/3\}, & (u,v) \in [1/3, 2/3]^2, \\ W(u,v), & \text{en otro caso.} \end{cases}$$
(2.16)

y

$$A(u,v) = \begin{cases} \max\{0, u - 1/3, v - 1/3, u + v - 1\}, & -1/3 \le v - u \le 1/3, \\ M(u,v), & \text{en otro caso.} \end{cases}$$
(2.17)

Además, A es una cuasi-cópula, B es una cópula y éstas son las mejores cotas posibles.

Por último, hay que comentar que el grado de dependencia de un vector aleatorio afecta el grado de no-intercambiabilidad de la cópula.



Figura 2.2.3: Gráficas de las cotas $A \ge B$

Corolario 2.2.25. Sean X, Y variables aleatorias continuas con cópula C máxima no-intercambiable. Si $\rho_{X,Y}$ y $\tau_{X,Y}$ denotan la rho de Spearman y la tau de Kendall para X y Y, respectivamente. Entonces

1.
$$-5/9 \le \rho_{X,Y} \le -1/3$$
,

2.
$$-5/9 \le \tau_{X,Y} \le 1/9$$
.

Como ya habíamos mencionado, las cópulas son un herramienta útil para medir dependencia de variables aleatorias. Seguramente si al ajustar una cópula a nuestros datos, escogemos una cópula Arquimediana o, de forma más general, una cópula simétrica, esto implicaría el supuesto de intercambiabilidad de nuestras variables (si las marginales son idénticamente distribuidas). Entonces es necesario probar el supuesto de que el vector (X, Y) tiene la misma función de distribución conjunta que el vector (Y, X). Sin embargo, este supuesto no siempre se cumple y entonces hay que proponer modelos más realistas, es decir, ajustar una cópula seleccionada de una familia formada por elementos no simétricos. En este contexto, Durante y otros (2010) se hacen la siguiente pregunta "¿Cómo una medida de no-intercambiabilidad nos ayuda a modelar datos?" y proponen un procedimiento simple.

Supongamos que tenemos una muestra aleatoria $(X_1, Y_1), \ldots, (X_n, Y_n)$ proveniente de un vector aleatorio continuo (X, Y), donde X, Y son idénticamente distribuidas. Consideremos una medida de no-intercambiabilidad $\mu : \mathcal{C} \mapsto \mathbf{R}_+$, por ejemplo $\delta(C)$

- Construir la cópula empírica \hat{C}_n ,
- Calcular δ_n dado por

$$\delta_{n} = 3 \max_{(u,v) \in \mathbf{I}^{2}} \left| \hat{C}_{n}(u,v) - \hat{C}_{n}(v,u) \right|.$$

de forma intuitiva, δ_n describe el grado de no-intercambiabilidad de la muestra aleatoria,

- Construir una familia (paramétrica) de cópulas $\{C_{\theta}\}_{\theta}$ en la clase de equivalencia de todas las cópulas que tienen medida de no-intercambiabilidad igual a δ_n ,
- Usando un procedimiento estándar (por ejemplo, el método de momentos), tratar de describir la dependencia ajustando un valor del parámetro de la familia $\{C_{\theta}\}_{\theta}$.

2.3. Construcción de cópulas asimétricas

En esta sección introduciremos algunos métodos para la construcción de cópulas asimétricas paramétricas. Empezaremos con un método propuesto por Liebscher (2008), que está conectado con productos de cópulas y donde un caso particular fue propuesto por Khoudraji (1996).

Teorema 2.3.1. Supongamos que $C_1, \ldots, C_k : \mathbf{I}^2 \mapsto \mathbf{I}$ son cópulas. Sean $g_{ji} : \mathbf{I} \mapsto \mathbf{I}$ para $j = 1, \ldots, k, i = 1, 2$ funciones con la propiedad de que cada una de ellas es estrictamente creciente o es idénticamente igual a 1. Supongamos que $\prod_{j=1}^{k} g_{ji}(v) = v$ para v en $\mathbf{I}, i = 1, 2, y \lim_{v \to 0^+} g_{ij}(v) = g_{ji}(0)$ para $j = 1, \ldots, k, i = 1, 2$. Entonces para todo u, v en \mathbf{I}

$$\bar{C}(u,v) = \prod_{j=1}^{k} C_j(g_{j1}(u), g_{j2}(v))$$
(2.18)

es también una cópula.

En general es obvio que \overline{C} es una cópula asimétrica. El caso especial de dos factores (k = 2) y funciones potencia (funciones de la forma $g_{ji}(x) = cx^a$, con c, a en **R**) fue estudiado por Khoudraji (1996). También notemos que las funciones $g_{ji} \equiv 1$ también están incluidas en el Teorema 2.3.1.

Observación 2.3.2. Si los supuestos del Teorema 2.3.1 se cumplen, entonces las funciónes g_{ji} tienen las siguientes propiedades

- $g_{ji}(1) = 1 \ y \ g_{ji}(0) = 0$,
- g_{ji} es continua en (0, 1],
- Si hay al menos dos funciones g_{j_1i}, g_{j_2i} con $1 \leq j_1, j_2 \leq k$ que no sean idénticamente igual a 1, entonces $g_{ji}(x) > x$ se cumple para x en $(0, 1), j = 1, \ldots, k$.

Ahora daremos cuatro alternativas para funciones g_{ji} que son aplicables en el Teorema 2.3.1.

Ejemplo 2.3.3. 1. $g_{ji}(v) = v^{\theta_{ji}} \text{ para } j = 1, ..., k \text{ donde } \theta_{ji} \text{ está en } [0, 1]$ $y \sum_{i=1}^{k} \theta_{ji} = 1;$

- 2. $g_{ji}(v) = v^{\theta_{ji}} exp\{(v-1)\alpha_{ji}\}$ para j = 1, ..., k, donde $\sum_{j=1}^{k} \theta_{ji} = 1$, $\sum_{j=1}^{k} \alpha_{ji} = 0$ y θ_{ji} está en (0,1), α_{ji} en $(-\infty,1)$, $\theta_{ji} \ge -\alpha_{ji}$;
- 3. $g_{ji}(v) = v^{\theta_{ji}} (1 e^{-\gamma_i v})^{-\alpha_{ji}} (1 e^{-\gamma_i})^{\alpha_{ji}}$ para j = 1, ..., k, donde $\sum_{j=1}^{k} \theta_{ji} = 1, \sum_{j=1}^{k} \alpha_{ji} = 0 \ y \ \theta_{ji}$ está en $(0, 1), \ \gamma_i$ en $(0, \infty), \ \alpha_{ji}$ en $(-\infty, 1), \ \alpha_{ji} \le \theta_{ji};$

4.
$$g_{1i}(v) = exp\left(\theta_i - \sqrt{|lnv| + \theta_i^2}\right), g_{2i}(v) = v \cdot exp\left(-\theta_i + \sqrt{|lnv| + \theta_i^2}\right)$$
$$para \ \theta_i \ge \frac{1}{2}.$$

El siguiente ejemplo nos muestra como usar la técnica de construcción del Teorema 2.3.1.

Ejemplo 2.3.4. Suponga que $C : \mathbf{I}^2 \mapsto \mathbf{I}$ es una cópula, $C_1 = C \ y \ C_2$ es la cópula independencia.

1. Sea $g_{1i}(v) = v^{1-\theta_i}$, entonces aplicando el Teorema 2.3.1 obtenemos

$$\bar{C}(u,v) = C\left(u^{1-\theta_1}, v^{1-\theta_2}\right) \cdot u^{\theta_1} \cdot v^{\theta_2}$$
(2.19)

la cual es cópula para θ_i en (0,1).

2. Definimos a K_{θ} como la cópula (2.19), haciendo $\theta_1 = \theta \ y \ \theta_2 = 0$ entonces

$$K_{\theta}(u,v) = u^{\theta} C\left(u^{1-\theta}, v\right)$$
(2.20)

diremos que la cópula K_{θ} es la versión asimétrica de C usando la ecuación de Khoudraji.

50

La cópula definida en (2.20) será retomada en el siguiente capítulo.

Lema 2.3.5. Sean $(U_1, V_1), \ldots, (U_k, V_k)$ k vectores independientes con funciones de distribución conjunta H_1, \ldots, H_n en \mathbf{I}^2 , respectivamente. También, sean $g_{ji} : \mathbf{I} \mapsto \mathbf{I}$ para $j = 1, \ldots, k$, i = 1, 2 como en el Teorema 2.3.1. Definimos $g_{ji}^{-1}(v) = 0$ para $v \leq g_{ji}(0)$. Entonces $(u, v) \mapsto H(u, v) = \prod_{j=1}^k H_j(g_{j1}(u), g_{j2}(v))$ es la función de distribución conjunta del vector aleatorio $\left(\max_{j\in J_i} \left\{g_{j1}^{-1}(U_j)\right\}, \max_{j\in J_i} \left\{g_{j1}^{-1}(V_j)\right\}\right)$ donde para $i = 1, 2, J_i = \{j \in \{1, \ldots, k\} : g_{ji} \neq 1\}.$

El Lema 2.3.5 nos da un procedimiento simple para generar observaciones de la cópula \hat{C} , definida en (2.18), cuando tenemos disponibles observaciones de las cópulas C_j . Sean $(U_1, V_1), \ldots, (U_k, V_k)$ k vectores independientes con funciones de distribución conjunta C_1, \ldots, C_k , respectivamente. Entonces \hat{C} es la función de distribución conjunta del vector aleatorio $\left(\max_{j\in J_i} \left\{g_{j1}^{-1}(U_j)\right\}, \max_{j\in J_i} \left\{g_{j1}^{-1}(V_j)\right\}\right)$ donde J_i es el conjunto de índices j con $g_{ji} \neq 1$.

Proposición 2.3.6. Sea τ_0 la tau de Kendall del producto de cópulas (2.18). Denotaremos a la tau de Kendall de C_j como τ_j (j = 1, ..., k) respectivamente. Entonces

$$\tau_0 \le \min_{j=1,\dots,k} \tau_j.$$

A continuación daremos más alternativas para la construcción de cópulas asimétricas.

1. Algunas veces aplicándole a una cópula simétrica alguna de las primeras dos transformaciones presentadas en el Teorema 1.4.1, nos da como resultado una cópula asimétrica, por ejemplo:

Ejemplo 2.3.7. Sea C_p la cópula Clayton definida en el Cuadro 1.5.1, aplicándole la transformación: $\overline{C}_p(u, v) = u - C_p(u, 1 - v)$ tenemos que

$$\bar{C}_{p}(u,v) = u - \left[\max\left(u^{-p} + (1-v)^{-p} - 1, 0\right)\right]^{-1/p}$$

la cual es una cópula asimétrica.

2. Usando la técnica de pegado de cópulas definida en (1.9), por ejemplo:

Ejemplo 2.3.8. Sean las cópulas M y W, entonces usando la técnica de pegado de cópulas con $\theta = 1/2$ tenemos

$$C_{1,2,\frac{1}{2}}(u,v) = \begin{cases} \min(u,v/2), & 0 \le u \le \frac{1}{2}, \\ \max(u+v-1,v/2) & \frac{1}{2} \le u \le 1. \end{cases}$$

3. Haciendo una combinación lineal convexa entre una cópula máxima no-intercambiable y una cópula simétrica, tenemos una cópula con un grado de asimetría determinado por el parámetro de la combinación misma (ver Ejemplo 2.2.15).

2.4. Pruebas de Simetría para cópulas bivariadas

Gran parte de los modelos paramétricos de cópulas cumplen la propiedad de simetría, por lo cual si queremos ajustar una de éstas cópulas a nuestros datos y las marginales son idénticamente distribuidas, estaríamos tomando el supuesto de intercambiabilidad de nuestras variables. Entonces es necesario validar este supuesto mediante una prueba de hipótesis.

Recordando la Definición 2.1.6, sean dos variables aleatorias X, Y con función de distribución conjunta H, sabemos que éstas son intercambiables si la siguiente hipótesis se cumple:

$$\mathcal{H}_0$$
: Para todo $(x, y) \in \mathbf{R}^2$ $H(x, y) = H(y, x)$,

entonces por el Teorema 2.1.8 sabemos que \mathcal{H}_0 se cumple si y sólo si:

- (i) F(x) = G(x) para todo x en \mathbf{R} ,
- (ii) C(u, v) = C(v, u) para todo u, v en \mathbf{I}^2 .

Por lo que \mathcal{H}_0 se puede rechazar si (i), (ii) o ambas no se cumplen. Donde la condición (i) puede ser validada mediante procedimientos gráficos o procedimientos estadísticos formales, pero no es del todo claro cómo probar la condición (ii). El propósito de esta sección es presentar dos pruebas, propuestas por Erdely y González-Barrios (2010) y Genest y otros (2011), para contrastar la hipótesis

$$\mathscr{H}_{0}: \text{ para todo } (u, v) \in \mathbf{I}^{2}, \ C(u, v) = C(v, u), \qquad (2.21)$$

contra la alternativa

 \mathscr{H}_1 : existe $(u, v) \in \mathbf{I}^2$, tal que $C(u, v) \neq C(v, u)$. (2.22)

Cabe mencionar que el problema de validar \mathcal{H}_0 ha sido de interés general desde hace ya varios años. Unos de los primeros contribuyentes a este tema fueron Bell y Haller (1969), quienes abordaron el problema en el modelo Gaussiano bivariado, y Hollander (1971), quien propuso una prueba que es consistente en contra de una amplia clase de alternativas.

Prueba de Tri-simetría (Erdely y González-Barrios)

La prueba de Tri-simetría de Erdely y González-Barrios (2010), como lo dice su nombre, es una prueba para validar la siguiente hipótesis para cualquier cópula C absolutamente continua:

$$\mathscr{H}_0^*$$
: C es una cópula tri-simétrica, (2.23)

contra la alternativa

$$\mathscr{H}_1^*$$
: C no es una cópula tri-simétrica. (2.24)

Aunque esta prueba no contrasta directamente la hipótesis nula de simetría dada en (2.21), nos puede ser de mucha utilidad, ya que, en consecuencia del Teorema 2.2.18 y Proposición 2.2.22, si aceptamos la hipótesis de tri-simetría podemos asegurar que el nivel de asimetría δ no va a ser más grande que 1/2, por lo que nuestra cópula se encuentra más cerca de ser simétrica que de ser máxima no-intercambiable. Si se rechaza la hipótesis de tri-simetría entonces también podemos rechazar simetría. Por lo que podemos utilizar esta prueba para contrastar la hipótesis nula de simetría para una cópula.

Definición 2.4.1. Sea $(X_1, Y_1), \ldots, (X_n, Y_n)$ una muestra aleatoria de tamaño n de un vector aleatorio continuo (X, Y) en \mathbb{R}^2 , con cópula $C_{X,Y}$. Definimos $S_n = |T_n|$, donde

$$T_n = \hat{C}_n \left(\frac{1}{3}, \frac{2}{3} \right) - \hat{C}_n \left(\frac{2}{3}, \frac{1}{3} \right), \qquad (2.25)$$

donde \hat{C}_n es la cópula empírica definida en (1.41).

Recordemos que una cópula es tri-simétrica si y sólo si C(1/3, 2/3) = C(2/3, 1/3), en ese sentido, S_n es un estadístico apropiado para tri-simetría. Por tal razón estudiaremos la distribución exacta de T_n y su distribución asintótica bajo independencia. Para hacer esto comenzaremos con algunos resultados generales. **Teorema 2.4.2.** Sea $(X_1, Y_1), \ldots, (X_n, Y_n)$ una muestra aleatoria de tamaño n de un vector aleatorio continuo (X, Y) de variables aleatorias independientes en \mathbb{R}^2 . Sea i un entero fijo tal que $1 \le i \le n/2$. Definimos las variables aleatorias

$$Z_n = n\hat{C}_n(i/n, i/n), \ V_n = n\hat{C}_n(i/n, 1 - i/n), \ W_n = n\hat{C}_n(1 - i/n, i/n),$$
(2.26)

donde \hat{C}_n es la cópula empírica. Sea

$$T_{i,n} = V_n - W_n, \tag{2.27}$$

entonces

$$P(T_{i,n} = t) = \sum_{v=0}^{i} P(V_n = v, W_n = v - t), \ t \in \{-i, \dots, -1, 0, 1, \dots, i\},$$
(2.28)

donde

$$P(V_{n} = v, W_{n} = w) = \sum_{z=0}^{i} \frac{\binom{i-z}{v-z} \binom{n+z-2i}{v}}{\binom{n-i}{i}} \frac{\binom{i-z}{w-z} \binom{n+z-2i}{w}}{\binom{n-i}{i}} \frac{\binom{i}{z} \binom{n-i}{i-z}}{\binom{n}{i}}$$
(2.29)

para $z \in \{0, 1, \dots, i\} \ y \ v, w \in \{z, z+1, \dots, \min\{i, n-2i+z\}\}.$

Corolario 2.4.3. Sea $(X_1, Y_1), \ldots, (X_n, Y_n)$ una muestra aleatoria de tamaño n de un vector aleatorio continuo (X, Y) de variables aleatorias independientes en \mathbb{R}^2 . Sea i un entero fijo tal que $1 \le i \le n/2$. Definimos las variables V_n , W_n y $T_{i,n}$ como en el Teorema 2.4.2. Sea $S_{i,n} = |T_{i,n}|$, entonces

$$P(S_{i,n} = s) = \begin{cases} P(T_{i,n} = 0), & s = 0, \\ 2P(T_{i,n} = s), & s = 1, 2, \dots, i. \end{cases}$$
(2.30)

Ahora obtendremos la esperanza y la varianza de $T_{i,n}$ dada en la ecuación (2.27).

Proposición 2.4.4. Sea $(X_1, Y_1), \ldots, (X_n, Y_n)$ una muestra aleatoria de tamaño n de un vector aleatorio continuo (X, Y) de variables aleatorias independientes en \mathbb{R}^2 . Sea i un entero fijo tal que $1 \le i \le n/2$. Sean las variables V_n , W_n y $T_{i,n}$ como en el Teorema 2.4.2. Entonces

$$E(T_{i,n}) = 0 \ y \ Var(T_{i,n}) = \frac{2i^2(n-2i)}{n(n-1)}.$$
(2.31)

Ya hemos encontrado la distribución exacta de $T_{i,n}$ bajo independencia. Sin embargo, para valores grandes de n podemos encontrar su distribución asintótica bajo estandarización y condiciones apropiadas de i.

Teorema 2.4.5. Sea $(X_1, Y_1), \ldots, (X_n, Y_n)$ una muestra aleatoria de tamaño n de un vector aleatorio continuo (X, Y) de variables aleatorias independientes en \mathbb{R}^2 . Sea $0 \le K < 1/2$ un valor fijo y sea $i = [n \cdot K]$ donde $[\cdot]$ denota la función máximo entero. Sean las variables aleatorias Z_n, V_n, W_n y $T_{i,n}$ como en las ecuaciones (2.26) y (2.27) del Teorema 2.4.2. Definimos

$$X_n = \frac{T_{i,n}}{\sqrt{\frac{2i^2(n-2i)}{n(n-1)}}},$$
(2.32)

entonces X_n es asintóticamente distribuido N(0,1).

Si $n \geq 3$ un entero fijo, y sea i = [n/3] entonces $1 \leq i \leq n/2$. Supongamos que $(X_1, Y_1), \ldots, (X_n, Y_n)$ es una muestra aleatoria de un vector aleatorio continuo (X, Y) con cópula subyacente $C_{X,Y}$. Con el fin de probar (2.23) usaremos el estadístico $T_{i,n}$ definido en (2.27) y su distribución bajo independecia. Hay que observar que si \mathscr{H}_0^* se cumple, entonces $|T_{i,n}|$ debe ser pequeño. Por lo que, valores grandes de $|T_{i,n}|$ nos dan evidencia en contra de \mathscr{H}_0^* . Por supuesto la distribución de $T_{i,n}$ depende de la cópula $C_{X,Y}$, sin embargo, usamos la distribución de $T_{i,n}$ bajo independencia, ya que las muestras independencia tienen una gran variabilidad para los valores del estadístico $T_{i,n}$, así que tomaremos a la independencia en representación de la clase de cópulas absolutamente continuas y simétricas.

Prueba de Simetría (Genest, Nešlehová y Quessy)

Ahora describiremos de manera muy general el estadístico propuesto por Genest y otros (2011) para contrastar la hipótesis nula dada en (2.21). Para esto, notemos que es natural comparar los valores de \hat{C}_n en (u, v) y (v, u), para todo u, v en **I**. Intuitivamente, $\hat{C}_n(u, v)$ y $\hat{C}_n^t(u, v) = \hat{C}_n(v, u)$ deberían ser valores cercanos bajo \mathscr{H}_0 . **Definición 2.4.1.** Definimos a S_n como una medida de discrepancia entre $\hat{C}_n \ y \ \hat{C}_n^t$ dada por

$$S_n = \int_0^1 \int_0^1 \left\{ \hat{C}_n(u,v) - \hat{C}_n(v,u) \right\}^2 d\hat{C}_n(u,v) \,. \tag{2.33}$$

Notemos que S_n es la versión con cópulas análoga al estadístico Cramérvon Mises dado en Hollander (1971) como

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left\{ H_n\left(x,y\right) - H_n\left(y,x\right) \right\}^2 dH_n\left(x,y\right).$$

El siguiente resultado nos da una expresión alternativa para S_n que es conveniente para propósitos computacionales.

Proposición 2.4.2. El estadístico S_n puede ser expresado como

$$S_n = \frac{1}{n^3} \sum_{k=1}^n \mathbf{1}^t B_k \mathbf{1}.$$
 (2.34)

donde **1** es un vector de unos de dimensión $n \times 1$ y B_1, \ldots, B_n son matrices de $n \times n$ con entrada en la posición (i, j) dada por

$$B_{kij} = \mathbb{I}\left(\hat{U}_i \vee \hat{U}_j \leq \hat{U}_k, \hat{V}_i \vee \hat{V}_j \leq \hat{V}_k\right) - \mathbb{I}\left(\hat{U}_i \vee \hat{V}_j \leq \hat{U}_k, \hat{V}_i \vee \hat{U}_j \leq \hat{V}_k\right) \\ -\mathbb{I}\left(\hat{U}_i \vee \hat{V}_j \leq \hat{V}_k, \hat{V}_i \vee \hat{U}_j \leq \hat{U}_k\right) + \mathbb{I}\left(\hat{U}_i \vee \hat{U}_j \leq \hat{V}_k, \hat{V}_i \vee \hat{V}_j \leq \hat{U}_k\right),$$

donde $(\hat{U}_i, \hat{V}_i), \ldots, (\hat{U}_n, \hat{V}_n)$ son las pseudo observaciones de la cópula definidas en (1.47) y para arbitrarios a, b en $\mathbf{R}, a \lor b = \max(a, b)$.

Se pueden obtener límites débiles del estadístico S_n , dado en (2.33), del comportamiento asintótico del proceso de la cópula empírca, definido para todo (u, v) en \mathbf{I}^2 como

$$\hat{\mathbb{C}}_{n}\left(u,v\right) = n^{1/2} \left\{ \hat{C}_{n}\left(u,v\right) - C\left(u,v\right) \right\}$$

La secuencia $\{\hat{\mathbb{C}}_n\}$ admite un límite débil $\hat{\mathbb{C}}$ denotado como $\hat{\mathbb{C}}_n \rightsquigarrow \hat{\mathbb{C}}$, cuando C es regular en el siguiente sentido:

Definición 2.4.3. Decimos que una cópula C es regular si

(i) las derivadas parciales de C dadas por $\dot{C}_1(u,v) = \partial C(u,v) / \partial u y$ $\dot{C}_2(u,v) = \partial C(u,v) / \partial v$ existen en casi todo \mathbf{I}^2 , con la convención de que las derivadas de un lado son usadas en los puntos frontera; (ii) \dot{C}_1 es continuo en $(0,1) \times [0,1]$ y \dot{C}_2 es continuo en $[0,1] \times (0,1)$.

El límite $\hat{\mathbb{C}}$ es un proceso Gaussiano centrado definido para todo (u,v) en $I\!\!I^2$ como

$$\hat{\mathbb{C}}(u,v) = \mathbb{C}(u,v) - \dot{C}_{1}(u,v)\mathbb{C}(u,1) - \dot{C}_{2}(u,v)\mathbb{C}(1,v),$$

donde \mathbb{C} es un campo aleatorio Gaussiano centrado cuya función de covarianzas está dada para todo u, v, s, t en I como

$$\Gamma_{\mathbb{C}}(u, v, s, t) = C(u \wedge s, v \wedge t) - C(u, v)C(s, t),$$

donde para todo a, b en $\mathbf{R}, a \wedge b = \min(a, b)$.

Ahora consideremos el proceso empírico simetrizado $\hat{\mathbb{D}}_n = n^{1/2} \left(\hat{C}_n - \hat{C}_n^t \right)$, explícitamente definido para todo (u, v) en I^2 como

$$\hat{\mathbb{D}}_n(u,v) = n^{1/2} \left\{ \hat{C}_n(u,v) - \hat{C}_n(v,u) \right\}.$$

Bajo \mathscr{H}_0 , el límite del proceso puede ser deducido fácilmente del comportamiento asintótico de \hat{C}_n , porque tenemos que $\hat{\mathbb{D}}_n(u, v) = \hat{\mathbb{C}}_n(u, v) - \hat{\mathbb{C}}_n(v, u)$ para todo (u, v) en I^2 .

Proposición 2.4.4. Si C es una cópula simétrica y regular, entonces $\hat{\mathbb{D}}_n$ converge débilmente, cuando $n \to \infty$, a un campo aleatorio Gaussiano $\hat{\mathbb{D}}$ definido para todo (u, v) en I^2 como

$$\hat{\mathbb{D}}(u,v) = \mathbb{D}(u,v) - \dot{C}_1(u,v) \mathbb{D}(u,1) - \dot{C}_2(u,v) \mathbb{D}(1,v)$$

en términos de un campo aleatorio Gaussiano centrado \mathbb{D} con función de covarianza dada para todo u, v, s, t en **I** como

$$\Gamma_{\mathbb{D}} = (u, v, s, t) = 2 \left\{ \Gamma_{\mathbb{C}} (u, v, s, t) - \Gamma_{\mathbb{C}} (u, v, t, s) \right\}.$$

Proposición 2.4.5. Si C es una cópula regular y simétrica, entonces cuando $n \to \infty$,

$$nS_n = \int_0^1 \int_0^1 \left\{ \hat{\mathbb{D}}_n(u,v) \right\}^2 d\hat{C}_n(u,v) \rightsquigarrow \quad \mathbb{D}_S = \int_0^1 \int_0^1 \left\{ \hat{\mathbb{D}}(u,v) \right\}^2 dC(u,v) \, .$$

Proposición 2.4.6. Si C es una cópula regular, entonces cuando $n \to \infty$,

$$S_n \xrightarrow{p} S(C) = \int_0^1 \int_0^1 \{C(u, v) - C(v, u)\}^2 dC(u, v).$$
 (2.35)

La distribución del estadístico nS_n bajo \mathscr{H}_0 depende de la forma de la cópula, la cual es generalmente desconocida, por lo que es imposible calcular P-valores mediante los procedimientos tradicionales. El teorema del límite central del multiplicador (Van Der[~]Vaart y Wellner (1996), Sección 2.9) provee una solución a este problema mediante las réplicas bootstrap del proceso de la cópula empírica.

Siguendo a Bücher y Dette (2010), fijando M en N y para cada h en $\{1, \ldots, M\}$, sea $\xi^{(h)} = \left(\xi_1^{(h)}, \ldots, \xi_n^{(h)}\right)$ un vector de variables aleatorias independientes no negativas con media y varianza iguales a uno. Estas variables deberán ser completamente independientes de los datos. Hacemos

$$\bar{\xi}_n^{(h)} = \frac{1}{n} \left(\xi_1^{(h)} + \dots + \xi_n^{(h)} \right),$$

у

$$\Xi_n^{(h)} = \left(\frac{\xi_1^{(h)}}{\bar{\xi}_n^{(h)}} - 1, \dots, \frac{\xi_n^{(h)}}{\bar{\xi}_n^{(h)}} - 1\right).$$

Para (u, v) en I^2 , sea $P_n(u, v)$ un vector de dimensión $n \times 1$ donde el *i*-ésimo componente está dado por

$$P_{in}(u,v) = \mathbb{I}\left(\hat{U}_i \le u, \hat{V}_i \le v\right) - \mathbb{I}\left(\hat{U}_i \le v, \hat{V}_i \le u\right).$$
(2.36)

Para cada h en $\{1, \ldots, M\}$ una versión bootstrap ponderada $\mathbb{D}_n^{(h)}$ de \mathbb{D} puede ser definida haciendo para todo (u, v) en I^2

$$\mathbb{D}_{n}^{(h)}(u,v) = n^{-1/2} \Xi_{n}^{(h)} P_{n}(u,v) \,. \tag{2.37}$$

De la representación asintótica de $\hat{\mathbb{D}}$ bajo \mathscr{H}_0 en la Proposición 2.4.4, una réplica bootstrap de este proceso está dada, para todo (u, v) en I^2 , por

$$\hat{\mathbb{D}}_{n}^{(h)} = \mathbb{D}_{n}^{(h)}(u,v) - \dot{C}_{1n}(u,v) \mathbb{D}_{n}^{(h)}(u,1) - \dot{C}_{2n}(u,v) \mathbb{D}_{n}^{(h)}(1,v)$$
$$= n^{-1/2} \Xi_{n}^{(h)} \left\{ P_{n}(u,v) - \dot{C}_{1n}(u,v) P_{n}(u,1) - \dot{C}_{2n}(u,v) P_{n}(1,v) \right\}.$$

Donde \dot{C}_{1n} y \dot{C}_{2n} son estimaciones de las derivadas parciales \dot{C}_1 y \dot{C}_2 , respectivamente. De forma más específica, sea ℓ_n en (0, 1/2) y para un v arbitrario

en \boldsymbol{I}

$$\dot{C}_{1n}(u,v) = \begin{cases} \frac{\hat{C}_n(2\ell_n,v)}{2\ell_n} & u \in [0,\ell_n), \\\\ \frac{\hat{C}_n(u+\ell_n,v) - \hat{C}_n(u-\ell_n,v)}{2\ell_n} & u \in [\ell_n, 1-\ell_n] \\\\ \frac{\hat{C}_n(1,v) - \hat{C}_n(1-2\ell_n,v)}{2\ell_n} & u \in (1-\ell_n,1]. \end{cases}$$

de manera similar, para un u arbitrario en I

$$\dot{C}_{2n}(u,v) = \begin{cases} \frac{\hat{C}_n(u,2\ell_n)}{2\ell_n} & v \in [0,\ell_n), \\\\ \frac{\hat{C}_n(u,v+\ell_n) - \hat{C}_n(u,v-\ell_n)}{2\ell_n} & v \in [\ell_n, 1-\ell_n], \\\\ \frac{\hat{C}_n(u,1) - \hat{C}_n(u,1-2\ell_n)}{2\ell_n} & v \in (1-\ell_n, 1]. \end{cases}$$

La siguiente proposición implica que asintóticamente, $\hat{\mathbb{D}}_n^{(1)}, \ldots, \hat{\mathbb{D}}_n^{(M)}$ son copias independientes de $\hat{\mathbb{D}}$ bajo \mathscr{H}_0 .

Proposición 2.4.7. Sea C una cópula regular y simétrica. Suponga que

$$\lim_{n \to \infty} \ell_n = 0, \quad \inf_{n \in \mathbb{N}} n^{1/2} \ell_n > 0$$

Entonces para todo M en \mathbf{N} , $(\hat{\mathbb{D}}_n, \hat{\mathbb{D}}_n^{(1)}, \dots, \hat{\mathbb{D}}_n^{(M)}) \rightsquigarrow (\hat{\mathbb{D}}, \hat{\mathbb{D}}^{(1)}, \dots, \hat{\mathbb{D}}^{(M)})$ cuando $n \to \infty$, donde los procesos $\hat{\mathbb{D}}^{(1)}, \dots, \hat{\mathbb{D}}^{(M)}$ son copias independientes de $\hat{\mathbb{D}}$.

Réplicas bootstrap del estadístico S_n ahora pueden ser definidas, para cada n en N y h en $\{1, \ldots, M\}$, como

$$n\hat{S}_{n}^{(h)} = \int_{0}^{1} \int_{0}^{1} \left\{ \hat{\mathbb{D}}_{n}^{(h)}(u,v) \right\}^{2} d\hat{C}_{n}(u,v) \,.$$

El siguiente resultado es una consecuencia inmediata de la Proposición 2.4.7.

Corolario 2.4.8. Suponga que las condiciones de la Proposición 2.4.7 se cumplen. Entonces para todo M en N y cuando $n \to \infty$,

$$\left(n\hat{S}_n, n\hat{S}_n^{(1)}, \dots, n\hat{S}_n^{(M)}\right) \rightsquigarrow \left(\mathbb{D}_S, \mathbb{D}_S^{(1)}, \dots, \mathbb{D}_S^{(M)}\right),$$

donde cada límite consiste en procesos independientes e idénticamente distribuidos. Se sigue del Corolario 2.4.8 que un P-valor aproximado para la prueba de \mathscr{H}_0 basada en S_n está dado por

$$\frac{1}{M}\sum_{h=1}^{M}\mathbb{I}\left(\hat{S}_{n}^{(h)} > S_{n}\right).$$

Por conveniencia, hacemos $\hat{\mathbb{D}}_n^{(h)} = n^{-1/2} \Xi_n^{(h)} Q_n$ para cada h en $\{1, \ldots, M\}$, donde para todo (u, v) en \mathbf{I}^2

$$Q_n(u,v) = P_n(u,v) - \dot{C}_{1n}(u,v) P_n(u,1) - \dot{C}_{2n}(u,v) P_n(1,v),$$

depende sólo de las pseudo observaciones recolectadas de los datos. Mientras que

$$\hat{S}_{n}^{(h)} = \frac{1}{n^{3}} \sum_{i=1}^{n} \left\{ \Xi_{n}^{(h)} Q_{n} \left(\hat{U}_{i}, \hat{V}_{i} \right) \right\}^{2}.$$

Capítulo 3

Análisis Comparativo

En este capítulo analizaremos la potencia de las dos pruebas de simetría presentadas en el capítulo anterior, Erdely y González-Barrios (2010) y Genest y otros (2011), mediante la simulación de datos provenientes tanto de cópulas simétricas como de cópulas asimétricas, para después interpretar dichos resultados y así determinar cual de ellas es la prueba más potente bajo distintos escenarios propuestos. Para llevar a cabo el objetivo se hizo uso del software libre R (R Development Core Team (2011)) y el paquete cópula (Yan (2007), Kojadinovic y Yan (2010), Hofert y Mächler (2011)).

Dada una cópula C y un nivel de significancia α , para contrastar la hipótesis nula \mathscr{H}_0 : C es una cópula simétrica, la potencia empírica de una prueba de simetría será calculada de la siguiente manera:

- 1. Generamos N muestras de tamaño n de nuestra cópula C;
- Para cada una de las muestras generadas, calculamos el P-valor resultante al contrastar *H*₀;
- 3. Determinar para que muestras se rechaza la hipótesis nula con la siguiente regla: rechazamos \mathcal{H}_0 si el P-valor es menor a α ;
- 4. La potencia empírica de la prueba está dada por la siguiente razón:

 $\frac{\text{\# de muestras en las que se rechaza } \mathcal{H}_0}{N}$

3.1. Potencia de las Pruebas

En esta sección compararemos las potencias empíricas de las pruebas antes presentadas, Genest y otros (2011) y Erdely y González-Barrios (2010), mediante la simulación de N = 1000 muestras de tamaño n = 50, 100, 250 para diferentes modelos conocidos de cópulas y sus versiones asimétricas bajo diferentes escenarios de concordancia caracterizados por rho de Spearman (ρ_S), con un nivel de significancia $\alpha = 0.05$. Cabe destacar que las familias de cópulas, los tamaños de muestra y los niveles de concordancia usados en esta sección fueron presentados en los trabajos de Genest y otros (2011) o Erdely y González-Barrios (2010).

En la prueba de Tri-Simetría de Erdely fue usada la distribución exacta del estimador para los tamaños de muestra n = 50, 100 y la distribución asintótica para n = 250; los P-valores aproximados en la prueba de simetría de Genest fueron calculados con base en M = 1000 aproximaciones bootstrap, y las variables $\xi_1^{(h)}, \ldots, \xi_n^{(h)}$ fueron tomadas de variables aleatorias exponenciales con media igual a uno.

Los cuadros 3.1.1 y 3.1.2 nos muestran las potencias calculadas de forma empírica basadas en 1000 muestras aleatorias de tamaño n = 50, 100, 250para diferentes valores de rho de Spearman de los siguientes modelos simétricos de cópulas. Notemos que, por ser cópulas simétricas, esperamos que las pruebas rechacen la hipótesis nula en un número reducido de muestras. Las potencias de las pruebas en cada tabla dentro de esta sección, serán representadas a modo de porcentaje.

1. La cópula independecia (IN), defnida en el Ejemplo 1.2.2, que está dada para todo (u,v) en $I\!\!I^2$ por

$$C^{IN}(u,v) = \Pi(u,v) = uv.$$

2. La cópula Clayton (CL), definida en el Cuadro 1.5.1, que está dada para todo (u, v) en I^2 y θ en $(0, \infty)$ por

$$C_{\theta}^{CL}(u,v) = \left[\max\left(u^{-\theta} + v^{-\theta} - 1, 0\right)\right]^{-1/\theta}.$$

3. La cópula Normal o Gaussiana (GA), definida en (1.18), que está dada para todo (u,v) en $I\!\!I^2$ por

$$C_r^{GA}\left(u,v\right) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi \left(1-r^2\right)^{1/2}} exp\left\{-\frac{s^2+t^2-2rst}{2\left(1-r^2\right)}\right\} dt ds$$
donde
r está en (-1,1) y Φ denota la función de distribución de una variable al
eatoria normal estándar.

4. La cópula Gumbel-Hougaard (GH), definida en (1.21), que está dada para todo (u,v) en $I\!\!I^2$ por

$$C_{\theta}^{GH}(u,v) = exp\left\{-\left[\left|\ln u\right|^{\theta} + \left|\ln v\right|^{\theta}\right]^{1/\theta}\right\}.$$

5. La cópula Cuadras-Augé (CA), definida en (1.12), que está dada para todo $\theta,\,u,\,v$ en \pmb{I} como

$$C_{\theta}^{CA}(u,v) = \min\left(u^{1-\theta}v, uv^{1-\theta}\right).$$

Cuadro 3.1.1: Potencia de las pruebas de Erdely y Genest, estimadas a partir de 1000 muestras de tamaño n = 50 de cinco cópulas simétricas.

Cópula	$ ho_S$	Genest	Erdely
n = 50			
IN	0.0	4.1	2.1
	0.25	4.4	0.5
CL	0.5	3.2	0.5
	0.75	2.7	0.0
	0.25	3.6	1.6
GA	0.5	2.9	0.4
	0.75	4.0	0.0
	0.25	4.5	1.1
GH	0.5	3.9	0.4
	0.75	2.7	0.1
	0.25	4.5	1.1
CA	0.5	16.1	0.6
	0.75	49.4	0.0

Cópula	$ ho_S$	Genest	Erdely
n = 100			
IN	0.0	3.6	3.7
	0.25	2.8	2.7
CL	0.5	2.6	1.2
	0.75	3.2	0.1
	0.25	3.0	3.1
GA	0.5	2.4	0.8
	0.75	2.1	0.1
	0.95	2.0	0.4
<u>AII</u>	0.25	2.9	2.4
GH	0.5	2.3	1.0
	0.75	1.9	0.0
	0.25	4.4	3.2
CA	0.5	12.6	1.7
	0.75	53.8	0.4
	0.00		0
n = 250			
IN	0.0	3.6	5.4
	0.25	3.9	3.6
CL	0.5	2.7	1.1
	0.75	3.1	0.0
	0.25	2.0	n 0
$C\Lambda$	0.25	2.9	2.0
GA	0.5	2.2	1.5
	0.75	2.2	0.1
	0.25	4.9	2.7
GH	0.5	3.6	1.3
	0.75	2.9	0.0
	0.25	5.3	2.8
CA	0.5	20.7	2.5
	0.75	60.0	0.6

Cuadro 3.1.2: Potencia de las pruebas de Erdely y Genest, estimadas a partir de 1000 muestras de tamaño n = 100, 250 de cinco cópulas simétricas.

3.1. POTENCIA DE LAS PRUEBAS

Para generar una versión asimétrica de una cópula paramétrica C se hace uso de la ecuación de Khoudraji, definida en (2.20) como:

$$K_{\theta}(u,v) = u^{\theta} C\left(u^{1-\theta},v\right) \text{ para todo } (u,v) \text{ en } I^{2}$$

dichas cópulas serán la base para determinar la potencia de las pruebas para el caso de cópulas asimétricas. Para θ en (0,1), los símbolos K_{θ}^{CL} , K_{θ}^{GA} , K_{θ}^{GH} y K_{θ}^{CA} denotan las versiones asimétricas de las cópulas Clayton, Gaussiana, Gumbel-Hougaard y Cuadras-Augé respectivamente.

La Figura 3.1.1 muestra los valores de la medida de asimetría definida en (2.10) como $\delta(C) = 3 \sup_{(u,v) \in I^2} |C(u,v) - C(v,u)|$ para las versiones asimétricas de las cópulas antes mencionadas con diferentes valores de ρ_S y θ . Como podemos observar, para valores pequeños de ρ_S tenemos un nivel de asimetría pequeño (digamos $\rho_S \leq 0.5$). En todos los casos la simetría máxima se alcanza alrededor de $\theta = 0.5$.



Figura 3.1.1: Valores de $\delta(K_{\theta})$ en función de θ . Línea continua $\rho_S = 0.9$; línea discontinua $\rho_S = 0.7$; línea punteada $\rho_S = 0.5$; línea discontinua punteada $\rho_S = 0.3$; línea discontinua larga $\rho_S = 0.1$.

Mostramos en los cuadros 3.1.3, 3.1.4 y 3.1.5 las potencias de las pruebas y su grado de asimetría para diferentes valores de θ y ρ_S de las versiones asimétricas de las cópulas ya mencionadas.

Cuadro 3.1.3:					
Cópula	θ	$ ho_S$	$\delta\left(K_{\theta}\right)$	Genest	Erdely
n = 50					
		0.5	0.08	20.1	1.4
	0.25	0.7	0.12	45.9	2.1
		0.9	0.17	87.9	1.5
		0.5	0.12	23.9	5.1
X_{θ}^{CA}	0.5	0.7	0.17	53.2	8.4
		0.9	0.22	88.3	14.6
		0.5	0.10	15.2	3.9
	0.75	0.7	0.14	28.9	5.9
		0.9	0.17	46.8	9.4
		0.5	0.02	3.7	0.6
	0.25	0.7	0.05	7.4	0.3
		0.9	0.11	36.4	0.7
		0.5	0.03	4.2	1.5
$\frac{CL}{\theta}$	0.5	0.7	0.06	7.0	1.2
		0.9	0.14	42.2	4.0
		0.5	0.02	3.9	0.9
	0.75	0.7	0.04	5.5	2.6
		0.9	0.10	14.6	4.2
		0.5	0.02	4.4	0.9
	0.25	0.7	0.04	6.5	0.4
		0.9	0.09	25.5	0.6
		0.5	0.03	4.5	0.5
$\mathcal{F}A$	0.5	0.7	0.06	8.9	1.6
·		0.9	0.13	46.9	5.3
		0.5	0.03	3.7	2.0
	0.75	0.7	0.06	10.1	2.2
		0.9	0.12	26.3	5.5
		0.5	0.02	4.9	0.8
	0.25	0.7	0.05	7.9	0.5
		0.9	0.09	6.7	0.5
		0.5	0.04	8.0	1.5
$_{ heta}^{GH}$	0.5	0.7	0.08	7.4	2.0
-		0.9	0.15	7.5	5.0
		0.5	0.05	7.6	1.8
	0.75	0.7	0.09	8.1	3.6
		0.9	0.13	7.3	6.3

Cópula	θ	$ ho_S$	$\delta\left(K_{ heta} ight)$	Genest	Erdely
n = 100					
		0.5	0.08	36.4	5.6
	0.25	0.7	0.12	75.3	7.2
		0.9	0.17	97.7	8.0
		0.5	0.12	47.7	14.8
K_{θ}^{CA}	0.5	0.7	0.17	88.1	27.2
		0.9	0.22	99.8	47.8
		0.5	0.10	27.9	12.8
	0.75	0.7	0.14	58.2	20.4
		0.9	0.17	85.0	30.3
		0.5	0.02	5.2	0.2
	0.25	0.3	0.02	0.0 10.2	2.3
	0.25	0.7	0.05	10.3 71.5	1.7
		0.9	0.11	11.0	4.5
K^{CL}	0.5	0.3	0.05	13.0	5.0 / 1
\mathbf{n}_{θ}	0.0	0.1	0.00	80.0	4.1 22 5
		0.5	0.14	30.0 4.2	20.0 4-3
	0.75	0.5	0.02	$\frac{4.2}{7.0}$	4.0 5 3
	0.15	0.1	0.10	32.1	15 1
		0.5	0.10	02.1	10.1
		0.5	0.02	3.3	1.9
	0.25	0.7	0.04	7.0	1.9
		0.9	0.09	54.9	5.1
		0.5	0.03	6.5	3.5
K_{θ}^{GA}	0.5	0.7	0.06	17.1	7.1
		0.9	0.13	85.5	26.8
		0.5	0.03	6.4	4.7
	0.75	0.7	0.06	17.4	8.2
		0.9	0.12	56.8	22.4
		0.5	0.02	5.9	2.4
	0.25	0.0	0.02 0.05	14.3	2.1 2.4
	0.20	0.9	0.09	64.3	4.8
		0.5	0.04	10.0	4.9
K^{GH}_{\circ}	0.5	0.7	0.08	31.6	10.0
0	0.0	0.9	0.15	92.7	30.6
		0.5	0.05	12.0	6.2
	0.75	0.7	0.09	28.8	11.2
		0.9	0.13	70.7	26.0

Cuadro 3.1.4:

		Cua	adro 3.1.5:		
Cópula	heta	$ ho_S$	$\delta\left(K_{ heta} ight)$	Genest	Erdely
n = 250					
		0.5	0.08	77.7	11.3
	0.25	0.7	0.12	99.1	20.5
		0.9	0.17	100.0	32.8
		0.5	0.12	92.4	41.5
X_{θ}^{CA}	0.5	0.7	0.17	99.9	72.3
		0.9	0.22	100.0	94.5
		0.5	0.10	69.8	29.3
	0.75	0.7	0.14	95.6	49.3
		0.9	0.17	99.8	72.5
		0.5	0.02	7.2	1.8
	0.25	0.7	0.05	32.2	4.8
		0.9	0.11	99.3	24.7
		0.5	0.03	9.1	5.2
CL	0.5	0.7	0.06	42.2	10.5
0		0.9	0.14	99.7	71.5
		0.5	0.02	5.6	5.7
	0.75	0.7	0.04	17.2	8.4
		0.9	0.10	77.0	41.5
		0.5	0.02	5.4	3.8
	0.25	0.7	0.04	23.7	3.8
		0.9	0.09	97.1	23.9
		0.5	0.03	12.8	6.1
GA	0.5	0.7	0.06	50.7	16.3
0		0.9	0.13	100.0	78.0
		0.5	0.03	9.3	7.5
	0.75	0.7	0.06	45.3	17.5
		0.9	0.12	97.1	54.2
		0.5	0.02	9.6	3.8
	0.25	0.7	0.05	35.4	6.2
		0.9	0.09	98.4	27.5
		0.5	0.04	22.7	9.9
GH	0.5	0.7	0.08	77.4	27.5
σ		0.9	0.15	100.0	85.3
		0.5	0.05	27.0	9.9
	0.75	0.7	0.09	74.0	29.7
		0.9	0.13	99.5	67.4
		0.0	0.10	00.0	01

En las fíguras 3.1.2, 3.1.3 y 3.1.4 graficamos la potencia de las pruebas y el nivel de asimetría de la siguiente cópula:

Combinación lineal convexa entre la cópula Clayton (2) y la cópula $M_{1/3}$ propuesta por Nelsen , es decir, para todo α , u, v en I

$$(1-\alpha) C_2^{CL}(u,v) + \alpha M_{1/3}(u,v),$$

donde C_2^{CL} es la cópula Clayton con parámetro $\theta = 2$ y la cópula $M_{1/3}$, definida en (2.11), es la cópula propuesta en Nelsen (2007) que alcanza el nivel máximo de asimetría y está dada para todo u, v en I como

$$M_{1/3}(u,v) = \min\left\{u, v, (u-2/3)_{+} + (v-1/3)_{+}\right\},\$$

donde para cualquier a en \mathbf{R} , $a_+ = \max(a, 0)$.



COMBINACIÓN LINEAL CONVEXA CLAYTON (2) y M1/3 n=50

Figura 3.1.2:



Figura 3.1.3:



Δ LÍNEA CONTINUA (ERDELY) LÍNEA PUNTEADA (GENEST) BARRAS (NIVEL DE ASIMETRÍA)

Figura 3.1.4:

En las fíguras 3.1.5, 3.1.6 y 3.1.7 graficamos la potencia de las pruebas y el nivel de asimetría de la siguiente cópula:

Combinación lineal convexa de la cópula "3 squared" y la cópula Independencia, es decir, la cópula dada para todo α , u, v en I por

$$(1-\alpha)\Pi(u,v) + \alpha C^{3S},$$

donde $\Pi\left(u,v\right)=uv$ es la cópula que denota la independencia y $C^{3S},$ definida en (2.14), está dada por

$$C^{3S}(u,v) = \begin{cases} 0, & si \ (u,v) \in I_1 \cup I_4, \\ (3u-1)v, & si \ (u,v) \in I_2, \\ v, & si \ (u,v) \in I_3, \\ u-1/3, & si \ (u,v) \in I_5, \\ 1/3 + (u-2/3)(3v-1), & si \ (u,v) \in I_6, \\ (3v-2)u, & si \ (u,v) \in I_7, \\ u+v-1, & si \ (u,v) \in I_8 \cup I_9, \end{cases}$$

donde I_j , $j = 1, \ldots, 9$ están definidos como en el Ejemplo 2.2.20 (véase Figura 2.2.2).



Figura 3.1.5:



COMBINACIÓN LINEAL CONVEXA INDEPENDENCIA Y 3-SQUARED n=100

Figura 3.1.6:



Figura 3.1.7:

En las fíguras 3.1.8, 3.1.9 y 3.1.10 graficamos la potencia de las pruebas y el nivel de asimetría de la siguiente cópula:

El pegado en $\theta=1/2$ de una cópula ${\rm Clayton}(p)$ con $p\geq 0$ y una versión asimétrica de la misma dada por

$$u - C_p^{CL}\left(u, 1 - v\right),$$

con la técnica de pegado de cópulas que está definida como:

$$C_{1,2,\theta}(u,v) = \begin{cases} \theta C_1\left(\frac{u}{\theta},v\right), & 0 \le u \le \theta, \\ (1-\theta) C_2\left(\frac{u-\theta}{1-\theta},v\right) + \theta v, & \theta \le u \le 1. \end{cases}$$



PEGADO CÓPULA CLAYTON Y SU VERSIÓN ASIMÉTRICA n=50

Figura 3.1.8:



PEGADO CÓPULA CLAYTON Y SU VERSIÓN ASIMÉTRICA n=100

Figura 3.1.9:





Figura 3.1.10:

3.2. Interpretación de Resultados

Ya mostradas las potencias de las dos pruebas de simetría bajo diferentes escenarios, lo siguiente será describir y analizar los resultados a fin de determinar condiciones generales bajo las cuales cada una de ellas presenta un mejor comportamiento.

Para comenzar el análisis recordemos que, al contrastar \mathcal{H}_0 , esperamos que la potencia de las pruebas sea menor a nuestro nivel de significancia $\alpha = 0.05$ cuando se trata de una cópula simétrica y lo más cercano posible a 1 cuando se trata de una cópula asimétrica.

Notemos que en los Cuadros 3.1.1 y 3.1.2, donde aplicamos las pruebas a cópulas simétricas, tenemos que las potencias cumplen con lo antes mencionado, es decir, sus potencias son menores a 5 %. También observemos en el Cuadro 3.2.1, donde se presenta un resumen de los resultados obtenidos en los Cuadros 3.1.1 y 3.1.2, que en la mayoría de los casos la prueba de Erdely tiene un mejor desempeño. Es importante destacar que la cópula Cuadras-Augé no cumple con las condiciones de regularidad presentadas en la Definición 2.4.3, lo cuál explica el mal comportamiento que la prueba de Genest tiene en esta cópula.

Cuadro 3.2.1: Número de casos en los que cada unas de las pruebas de hipótesis presenta mejores resultados cuando las cópulas son simétricas.

Tamaño de la muestra	Genest	Erdely	Casos totales
$50 \\ 100 \\ 250$	0	13	13
	2	11	13
	1	12	13

En los Cuadros 3.1.3, 3.1.4 y 3.1.5, donde aplicamos las pruebas a cópulas con niveles bajos de asimetría, se resalta lo siguiente: en la prueba de Genest, para los tamaños de muestra n = 50, 100 y 250 y grados de asimetría mayores a 0.22, 0.17 y 0.11, respectivamente, la potencia presenta valores desde 85 % y hasta 100 % en algunos casos; en la prueba de Erdely hay que notar que la potencia presenta cierta inestabilidad, para distintas cópulas con niveles de asimetría similares ésta se presenta muy distinta, más aún, con las cópulas expuestas, la potencia sólo se acerca a 100% en la cópula Cuadras-Augé cuando el grado de asimetría es 0.22 y el tamaño de muestra es de 250. Con el Cuadro 3.2.2 vemos que en casi todos los casos la prueba de Genest es más potente que la de Erdely.

Tamaño de la muestra	Genest	Erdely	Casos totales
50	36	0	36
100	34	2	36
250	35	1	36

Cuadro 3.2.2: Número de casos en los que cada unas de las pruebas de hipótesis presenta mejores resultados cuando las cópulas son asimétricas.

En las Figuras de la 3.1.2 a la 3.1.10 se observa que la potencia de ambas pruebas de simetría se aproximan a 100% a medida que las cópulas van alcanzando sus niveles máximos de asimetría y también que la potencia mejora conforme el tamaño de muestra se incrementa.

En las Figuras 3.1.2, 3.1.3 y 3.1.4 la prueba de simetría de Genest está por encima de la de Erdely para los tamaños de muestra n = 100 y 250, cuando el tamaño de muestra es 50 la prueba de Erdely es más potente para valores de α mayores a 0.65; en las Figuras 3.1.5, 3.1.6 y 3.1.7 la prueba de Erdely es más potente que la de Genest en todos los tamaños de muestra, esto probablemente se debe a la forma en que la cópula "3-squared" fue construida; por último, en las Figuras 3.1.8, 3.1.9 y 3.1.10 la prueba de simetría de Genest es más potente que la de Erdely para todos los valores de p y todos los tamaños de muestra.

Con base a las observaciones anteriores decimos lo siguiente: la prueba de Erdely demuestra ser ligeramente mejor para rechazar \mathscr{H}_0 cuando las cópulas son simétricas, pero con cópulas asimétricas la prueba de Genest resulta considerablemente más eficiente que la de Erdely.

Otro aspecto fundamental que hay que tener en cuenta respecto a ambas pruebas es el "costo" computacional que estas tienen. La prueba de Erdely es fácil de programar y su tiempo de proceso es bastante corto, incluso con muestras grandes. Mientras que la prueba de Genest requiere de hacer simulaciones y el algoritmo es un poco más complejo, lo cuál representa un tiempo de proceso mucho más largo y mayor dificultad en su programación. Es decir, computacionalmente hablando la prueba de Erdely representa una mejor opción.

En el desarrollo de este capítulo se han presentando distintos escenarios bajo los cuales cada una de las pruebas se desenvuelve mejor. Sin embargo, no se debe perder de vista que al trabajar con datos reales sólo tenemos conocimiento del tamaño de muestra, desconociendo las condiciones de asimetría de la cópula de las variables o si dicha cópula cumple con las condiciones de regularidad. En ese sentido, dado que la prueba de Genest presenta resultados más acertados, ésta resulta ser una mejor opción para contrastar \mathscr{H}_0 .

Dado que la prueba de Genest no es la más potente bajo cualquier escenario, no podemos afirmar que ésta es Uniformemente Más Potente, sólo podemos concluir que la prueba de simetría de Genest es **Más Potente** que la prueba de Tri-simetría de Erdely en la mayoría de los casos analizados.

Sólo basta remarcar que en este trabajo se usaron ciertas familias de cópulas, se pudieron haber elegido otras familias de cópulas y los resultados podrían ser diferentes. Sin embargo, existe cierta tranquilidad ya que las cópulas utilizadas son en general las más utilizadas para modelar un conjunto de datos, ya sea que la cópula se ajuste directamente a los datos o que se ocupe para construir un modelo más complejo.

Conclusiones

En el presente trabajo se presentó la definición de cópula como una función con ciertas características, después se asoció a las funciones cópula con las funciones de distribución conjunta, vía el Teorema de Sklar. Este resultado tan importante nos dio una manera eficiente de medir dependencia de variables aleatorias, debido a que en la cópula recae toda la información acerca de la dependencia de nuestras variables, por lo cual ya no hay pretexto para usar supuestos como la distribución normal multivariada o la independencia para un modelado conjunto, que en ocasiones resultan poco realistas.

Posteriormente se definió a la cópula empírica como una función de la muestra, la cual es una aproximación de la cópula subyacente. Hay que destacar su importancia ya que, gracias a esta función, fue posible la construcción de estimadores para algunas medidas de asociación, así como la creación de pruebas estadísticas no paramétricas tales como las pruebas de simetría. Aunque en este trabajo se presentó sólo la definición para dimensión dos, cabe mencionar que, hoy en día, el computo de la cópula empírica para dimensiones mayores a tres, puede llegar a ser un proceso sumamente tardado o hasta inoperante computacionalmente hablando. El tiempo de proceso y la capacidad de almacenamiento están íntimamente ligados con el tamaño de la muestra (n) y la dimensión de la cópula (d), ya que la cantidad de valores a calcular y a almacenar son n^d , número que puede llegar a ser excesivamente grande.

En el segundo capítulo se estudiaron conceptos generales de simetría para variables y vectores aleatorios, incluido entre ellos el de intercambiabiliad. Es importante enfatizar que aunque este concepto está ampliamente ligado al de simetría de cópulas, estos no son equivalentes, es decir, a menos que las variables sean idénticamente distribuidas, la simetría no implica intercambiabilidad. Aún así, ambos conceptos son un aspecto fundamental a considerar cuando se lleva acabo un modelado con cópulas.

También se definieron propiedades deseables para una medida de nointercambiabilidad de un par de variables aleatorias en términos de la cópula subyacente y se mostraron algunas medidas que cumplen con dichas propiedades, considerando como mejor opción a la medida propuesta por Nelsen, denominándola simplemente como el grado de no-intercambiabilidad o de asimetría de la cópula.

Gran parte de los familias conocidas de cópulas son simétricas y al igual que la dependencia, la simetría no se puede tomar a la ligera, al momento de ajustar nuestros datos hay que proponer modelos más realistas que contemplen el grado tanto de dependencia como de asimetría.

Hasta hace unos años el validar la simetría era algo complicado, durante las últimas décadas se han dado diversas propuestas para atacar el problema, pero ninguna lo solucionó por completo. No fue sino hasta hace dos años que el tema se retomó con mayor éxito: primero, Erdely y González-Barrios (2010) con su prueba de Tri-Simetría, siendo ésta la mejor opción hasta ese momento, pese a que ésta no es una prueba propiamente de simetría; un año más tarde, se publicó en Genest y otros (2011) la primer prueba de simetría para cópulas bivariadas hasta el momento.

Expuesto lo anterior, sólo queda responder a la pregunta: ¿Qué prueba es más eficiente para validar simetría? Responder dicha interrogante es algo muy complicado con procedimientos formales de la teoría de la probabilidad (debido a lo complejo que es deducir la distribución de los estadísticos de prueba), por lo que se recurrió a una herramienta estadística bastante útil: la simulación, haciendo lo anterior una tarea mucho más sencilla. Cumpliendo los objetivos de este trabajo, se concluye que la prueba más eficiente para validar simetría es la de Genest y otros (2011) (al menos en la mayoría de los casos analizados en este trabajo).

Se ha concluido que la prueba de Genest y otros (2011) es la mejor opción para contrastar la hipótesis nula de simetría de una cópula, sin embargo hay que tener en cuenta dos detalles respecto a dicha prueba: se hace el supuesto de que la cópula subyacente cumple las condiciones de regularidad en la cópula, y en el caso de no cumplirse este supuesto la prueba no promete tener un buen comportamiento; por otra parte, la prueba de Genest (computacionalmente hablando) es más "costosa" que la de Erdely y González-Barrios (2010).

Actualmente existen algunos trabajos considerados guías rápidas para ajuste paramétrico con cópulas (véase De[~]Matteis (2001) y Trivedi y Zimmer (2007)), pero las cópulas han tenido un crecimiento impresionante en los últimos años, con la aparición de nuevas técnicas estadísticas y metodologías de construcción de cópulas tales como: pruebas de simetría, pruebas de independencia (Genest y Rémillard (2004)), pruebas de Arquimeneidad (Bücher, Dette y Volgushev (2011)), "gluing copulas" (Siburg y Stoimenov (2008)), "pair copulas" (Czado (2010) o la metodología "vine-copulas" (Kurowicka y Joe (2010)). Es por esto que los procedimientos para ajustes paramétricos con cópulas deben ser actualizados y, si es necesario, estandarizados incorporando las técnicas hasta el momento desarrolladas.

CONCLUSIONES

82

Apéndice A

Código en R

Todos los resultados mostrados en este trabajo fueron obtenidos usando el software libre R (R Development Core Team (2011)).

A.1. Instalación del Software R

- 1. Acudir a la siguiente dirección de internet www.r-project.org;
- 2. En el menú vertical de la izquierda, hacer click en CRAN;
- 3. En la lista de países, localizar USA y hacer click en el enlace de la University of California, Los Angeles;
- 4. Hacer click en Windows;
- 5. Hacer click en base;
- 6. Hacer click derecho con el mouse en Downoload R 2.15.0 for Windows, luego "Guardar vínculo como..." y descargar el archivo en algún directorio que recuerdes.
- 7. Hacer doble click en el archivo descargado y seguir la instrucciones del Ayudante de Instalación (Wizard), permitiendo que continúe con las opciones que propone por default, hasta concluir la instalación.

Al finalizar lo anterior, en el escritorio deberá estar un icono azul con una letra R mayúscula. En él hay que hacer doble click para ejecutar el programa R. Puede ocurrir durante la instalación que al tratar de instalar R en el directorio "Archivos de Programa" el acceso sea denegado (especialmente si se está instalando en una computadora de una empresa, que por lo general tienen restricciones), en tal caso mediante el botón "Examinar" elija otro directorio de instalación, por ejemplo en "Mis Documentos" o algún otro que le sea permitido al usuario.

Para nuestros fines, es necesario instalar el paquete **copula** mediante el siguiente procedimiento:

- 1. En la página de R www.r-project.org hacer click en CRAN
- 2. Escoger un servidor-espejo cercano, por ejemplo el de México
- 3. En el menú vertical del lado izquierdo, hacer click en Packages
- 4. Hacer click en el paquete de interés copula
- 5. Descargar el archivo PDF que se indica en Reference manual (para consultar dicho manual)
- 6. La instalación del paquete se hace directamente desde R, haciendo click en el menú Paquetes y luego Instalar paquete(s)...
- 7. Hacer click en un servidor-espejo, por ejemplo México, y Aceptar
- 8. Hacer click en el paquete deseado, y Aceptar
- 9. R instalará automáticamente el paquete. La instalación sólo es necesario hacerlo una vez.
- 10. Para utilizar el paquete en cualquier sesión, es neceario cargarlo mediante el menú Paquetes y luego Cargar paquete...

A.2. Algoritmos de Simulación

Simulación Arquimedianas

```
SimulArq<-function(gen,n)
{
#Simulación Arquimedianas
#Dar la función generadora en términos de u
#en forma de expresión... expression()
#y el número de simulaciónes deseadas n
u<-runif(n,0,1)
t<-runif(n,0,1)</pre>
```

```
phi<-function(u) eval(gen)</pre>
phiq<-function(u,q) phi(u)-q</pre>
dphi<-function(u) eval(D(gen,"u"))</pre>
dphiw<-function(u,k) dphi(u)-k
invphi<-function(q)</pre>
{
if (0<=q&q<=phi(0))
{
inv<-uniroot(phiq,c(.000000001,.999999999),tol=.0001,q=q)$root</pre>
}
else inv<-0
inv
}
dginv<-function(k)</pre>
ſ
if (k>dphi(0)&k<=0)
{
ginv<-uniroot(dphiw,c(.00000001,.999999999),tol=.0001,k=k)$root
}
else ginv<-0
ginv
}
w<-sapply(dphi(u)/t,dginv)
v<-sapply(phi(w)-phi(u),invphi)</pre>
cbind(u,v)
}
```

Simulación cópula Marshall-Olkin

```
marshall<-function(n,a,b){
#Simulaciones cópula Marshall-Olkin
#n=número de observación de la cópula
#a,b= párametros de la cópula donde 0<=a,b<=1
r<-runif(n)
s<-runif(n)
t<-runif(n)
u=exp(-1*mapply(min,-log(r)/(1-a),-log(t)/a))
v=exp(-1*mapply(min,-log(s)/(1-b),-log(t)/b))
cbind(u,v)
}</pre>
```

```
Simulación Ecuación de Khoudraji
```

```
khoudraji<-function(tabla,d)
{
#Simulación ecuación khoudraji
#tabla=observaciones de una cópula C a la cual
#le aplicaremos la ecuación de khodraji
#Kd(u,v)= u^d*C(u^(1-d),v)
n<-length(tabla[,1])
k<-mapply(max,(tabla[,1])^(1/(1-d)),runif(n)^(1/d))
cbind(k,tabla[,2])
}</pre>
```

Simulación Combinación Lineal Convexa

```
clc<-function(a,cop1,cop2)
{
#Simulación combinación lineal
#cop1,cop2=simulaciones de dos cópulas
#con el mismo número de observaciones
#a=parámetro de la combinación lineal
#a*cop1+(1-a)*cop2
n=length(cop1[,1])
b<-rbinom(n,1,a)
b*cop1+(1-b)*cop2
}</pre>
```

Simulación Pegado de Cópulas

```
gluing<-function(theta,cop1,cop2){
#Simulación de un pegado de cópulas
#cop1,cop2=simulaciones de dos cópulas
#con el mismo número de observaciones
#theta=parámetro del pegado de cópulas
n=length(cop1[,1])
n1=floor(n*theta)
n2=ceiling(n*(1-theta))
cop1[,1]=cop1[,1]*theta
cop2[,1]=cop2[,1]*(1-theta)+theta
rbind(cop1[1:n1,],cop2[1:n2,])
}</pre>
```

86

A.3. Pruebas de Simetría

Cópula Empírica

```
copemp<-function(tabla) #Cn(i/n,j/n) copula empírica
{
    #tabla=muestra aleatoria (X1,Y2),...,(Xn,Yn)
    #copemp es la matriz de la cópula empírica
    #copemp[i+1,j+1] equivale a Cn(i/n,j/n)
f<-function(i,r) (r<=i)
    tabla<-tabla[order(tabla[,1]),]
n<-length(tabla[,1])
y<-c(n+1,rank(tabla[,2]))
nn<-sapply(0:n,f,r=y)
    copula<-apply(nn,2,cumsum)
    copula/n
}</pre>
```

Prueba de Simetría de Erdely

```
TS.test<-function(tabla)
{
#Prueba de Trisimetria
#usando distribución exacta del estimador
C<-copemp(tabla)
n<-length(tabla[,1])</pre>
i<-round((n/3),0)
Tin<-abs(n*(C[i+1,n-i+1]-C[n-i+1,i+1]))
f<-function(i,n,z,t,v)</pre>
{
choose(i-z,v-t-z)*(choose(n+z-2*i,v-t)/choose(n-i,i))
}
g<-function(i,n,z,t,v)
ł
f(i,n,z,0,v)*f(i,n,z,t,v)*dhyper(z,i,n-i,i)
}
m<-function(t,i,n) 2*sum(outer(0:i,0:i,g,i=i,n=n,t=t))</pre>
if(Tin>0) p<-sum(sapply(Tin:i,m,i=i,n=n))</pre>
else p<-1
a<-noquote("Prueba de Tri-Simetría")
b<-noquote(paste("|Tn|=",round(Tin,0)))</pre>
```

```
c<-noquote(paste("p-value =",p))</pre>
list(a,b,c)
}
TSa.test<-function(tabla)
ſ
#Prueba de tri-simetría usando distribución asintótica
C<-copemp(tabla)
n<-length(tabla[,1])</pre>
i<-floor(n/3)
Tin<-abs(n*(C[i+1,n-i+1]-C[n-i+1,i+1]))
Xn<-Tin/sqrt(2*i^2*(n-2*i)/(n*(n-1)))</pre>
a<-noquote("Prueba de Tri-Simetría")
p<-2*pnorm(-Xn,0,1)</pre>
b<-noquote(paste("Xn=",Xn))</pre>
c<-noquote(paste("p-value =",p))</pre>
list(a,b,c)
}
Prueba de Simetría de Genest
GenestS.test<-function(tabla,M)</pre>
{
#Prueba de Simetría de Genest
n<-length(tabla[,1])</pre>
ln<-n/(sqrt(n))</pre>
```

```
tabla<-tabla[order(tabla[,1]),]</pre>
v<-rank(tabla[,2])
C<-copemp(tabla)
f<-function(i,j,k)</pre>
ſ
(i<=k&v[i]<=v[k])*((j<=k&v[j]<=v[k])-(v[j]<=k&j<=v[k]))+
(i<=v[k]&v[i]<=k)*((j<=v[k]&v[j]<=k)-(v[j]<=v[k]&j<=k))
}
g<-function(k) sum(outer(1:n,1:n,f,k=k))
Pn<-function(uu,vv) (1:n<=uu&v<=vv)-(1:n<=vv&v<=uu)
C1n<-function(uu,vv)
ſ
if (0<=uu&uu<ln) x<-C[floor(2*ln)+1,floor(vv)+1]</pre>
else if (ln<=uu&uu<=(n-ln))
```

88

```
{
x<-C[floor(uu+ln)+1,floor(vv)+1]-C[floor(uu-ln)+1,floor(vv)+1]</pre>
}
else if ((n-ln)<uu&uu<=n)</pre>
ſ
x < -C[n+1,floor(vv)+1] - C[floor(n-2*ln)+1,floor(vv)+1]
}
x*n/(2*ln)
}
C2n<-function(uu,vv)
{
if (0<=vv&vv<ln) x<-C[floor(uu)+1,floor(2*ln)+1]</pre>
else if (ln<=vv&vv<=(n-ln))</pre>
{
x<-C[floor(uu)+1,floor(vv+ln)+1]-C[floor(uu)+1,floor(vv-ln)+1]</pre>
}
else if ((n-ln)<vv&vv<=n)</pre>
-
x < -C[floor(uu)+1,n+1] - C[floor(uu)+1,floor(n-2*ln)+1]
}
x*n/(2*ln)
}
Qn<-function(uu,vv)
{
Pn(uu,vv)-C1n(uu,vv)*Pn(uu,n)-C2n(uu,vv)*Pn(n,vv)
}
Qnv<-function(i) Qn(i,v[i])</pre>
Sn<-sum(sapply(1:n,g))/n^3</pre>
Snh<-function(h)
{
xi<-rexp(n,1)</pre>
phi<-xi/mean(xi)-1
sum((t(phi)%*%sapply(1:n,Qnv))^2)/n^3
}
pvalue<-mean((sapply(1:M,Snh)>Sn))
a<-noquote("Prueba de Simetría de Genest")
b<-noquote(paste("Sn=",Sn))</pre>
c<-noquote(paste("p-value =",pvalue))</pre>
list(a,b,c)
}
```

Apéndice B

Diagramas de Dispersión



Figura B.0.1: 500 Simulaciones de las cópulas $\Pi,\,M$ yW



Figura B.0.2: 500 Simulaciones de la cópula Clayton, $\theta=-0.8$ (izquierda), $\theta=4$ (derecha)



Figura B.0.3: 500 Simulaciones de la cópula Normal, $\rho=-0.75$ (izquierda), $\rho=0.75$ (derecha)



Figura B.0.4: 500 Simulaciones de la cópula Gumbel-Hougaard, $\theta=2$ (izquierda), $\theta=5$ (derecha)



Figura B.0.5: 500 Simulaciones de la cópula Cuadras-Augé, $\theta=0.25$ (izquierda), $\theta=0.75$ (derecha)



Figura B.0.6: 500 Simulaciones de la versión Asimétrica de la Cópula Clayton, K_{δ}^{CL} con $\delta = 0.5$, $\theta = -0.8$ (izquierda), $\theta = 5$ (derecha)



Figura B.0.7: 500 Simulaciones de la versión Asimétrica de la Cópula Normal, K^{GA}_{δ} con $\delta=0.5,\,\rho=-0.75$ (izquierda), $\rho=0.75$ (derecha)



Figura B.0.8: 500 Simulaciones de la versión Asimétrica de la Cópula Gumbel-Hougaard, K_{δ}^{GH} con $\delta = 0.5$, $\theta = 25$ (izquierda), $\theta = 5$ (derecha)



Figura B.0.9: 500 Simulaciones de la versión Asimétrica de la Cópula Cuadras-Augé, K_{δ}^{CA} con $\delta=0.25,\,\theta=25$ (izquierda), $\theta=0.75$ (derecha)



Figura B.0.10: 500 Simulaciones de las cópulas 3-squared (izquierda) y la cópula dada por Nelsen ${\cal M}_{1/3}$ (derecha)



Figura B.0.11: 500 Simulaciones del pegado de las cópulas Clayton y su versión asimétrica $(u-C_p^{CL}(u,1-v)), p = -0.8$ (izquierda), p = 5 (derecha)



Figura B.0.12: 500 Simulaciones de la combinación lineal convexa de la cópula Clayton (2) y la cópula dada por Nelsen $M_{1/3}$, $\alpha = 0.25$ (izquierda), $\alpha = 0.=75$ (derecha)



Figura B.0.13: 500 Simulaciones de la combinación lineal convexa de la cópula 3-squared y la cópula Independencia, $\alpha = 0.25$ (izquierda), $\alpha = 0.75$ (derecha)

Bibliografía

- BELL, C.B. y HALLER, H.S. (1969). «Bivariate symmetry tests: parametric and nonparametric». *The Annals of Mathematical Statistics*, pp. 259–269.
- BÜCHER, A. y DETTE, H. (2010). «A note on bootstrap approximations for the empirical copula process». *Statistics & probability letters*, 80(23-24), pp. 1925–1932.
- BÜCHER, A.; DETTE, H. y VOLGUSHEV, S. (2011). «A test for Archimedeanity in bivariate copula models». Arxiv preprint arXiv:1109.6501.
- CUADRAS, C.M. y AUGÉ, J. (1981). «A continuous general multivariate distribution and its properties». *Communications in Statistics-Theory and Methods*, **10(4)**, pp. 339–353.
- CZADO, C. (2010). «Pair-copula constructions of multivariate copulas». Copula Theory and Its Applications, pp. 93–109.
- DE MATTEIS, R. (2001). «Fitting copulas to data». Institute of Mathematics/University of Zurich.
- DEHEUVELS, P. (1979). «La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance». Acad. Roy. Belg. Bull. Cl. Sci.(5), 65(6), pp. 274–292.
- DURANTE, F.; KLEMENT, E.P.; SEMPI, C. y ÚBEDA-FLORES, M. (2010). «Measures of non-exchangeability for bivariate random vectors». *Statistical Papers*, **51(3)**, pp. 687–699.
- DURANTE, F. y SEMPI, C. (2010). «Copula theory: an introduction». Copula Theory and its Applications, pp. 3–31.
- EMBRECHTS, P. (2009). «Copulas: A personal view». Journal of Risk and Insurance, **76(3)**, pp. 639–650.

- ERDELY, A. y GONZÁLEZ-BARRIOS, J.M. (2010). «A nonparametric symmetry test for absolutely continuous bivariate copulas». Statistical Methods & Applications, pp. 1–25.
- FISHER, N.I. (1997). «Copulas». Encyclopedia of Statistical Sciences, Update, 1, pp. 159–163.
- FRÉCHET, M. (1951). «Sur les tableaux de corrélation dont les marges sont données». Ann. Univ. Lyon, III. Sér., Sect. A, 14, pp. 53–77.
- FRÉCHET, M. (1958). «Remarques au sujet de la note précédente». CR Acad. Sci. Paris Sér. I Math, 246, pp. 2719–2720.
- GALAMBOS, J. (1982). «Exchangeability». Encyclopedia of statistical sciences, 2, pp. 573–577.
- GENEST, C.; NEŠLEHOVÁ, J. y QUESSY, J.F. (2011). «Tests of symmetry for bivariate copulas». Annals of the Institute of Statistical Mathematics, pp. 1–24.
- GENEST, C. y RÉMILLARD, B. (2004). «Test of independence and randomness based on the empirical copula process». *Test*, **13(2)**, pp. 335–369.
- GUMBEL, E.J. (1960). «Bivariate exponential distributions». Journal of the American Statistical Association, pp. 698–707.
- GUMBEL, E.J. (1961). «Bivariate logistic distributions». Journal of the American Statistical Association, pp. 335–349.
- HOEFFDING, W. (1940). Masstabinvariante Korrelationstheorie. Schriften des Matematischen Instituts und des Instituts für Angewandte Mathematik der Universität Berlin 5 Heft 3:179-233.
- HOFERT, MARIUS y MÄCHLER, MARTIN (2011). «Nested Archimedean Copulas Meet R: The nacopula Package». Journal of Statistical Software, **39(9)**, pp. 1–20. http://www.jstatsoft.org/v39/i09/
- HOLLANDER, M. (1971). «A nonparametric test for bivariate symmetry». Biometrika, 58(1), pp. 203–212.
- HUTCHISON, TP y LAI, CD (1990). «Continuous Bivariate Distributions, Emphasizing Applications». Adelaide: Rumsby Scientific Publishing.
- JOE, H. (1997). Multivariate models and dependence concepts. volumen 73. Chapman & Hall/CRC.
- JOHNSON, N.L. y KOTZ, S. (1970). *Distributions in statistics*. Houghton Mifflin New York.
- KHOUDRAJI, A. (1996). Contributions à l'étude des copules et à la modélisation de valeurs extrêmes bivariées.
- KOJADINOVIC, IVAN y YAN, JUN (2010). «Modeling Multivariate Distributions with Continuous Margins Using the copula R Package». Journal of Statistical Software, 34(9), pp. 1–20. http://www.jstatsoft.org/v34/i09/
- KUROWICKA, D. y JOE, H. (2010). Dependence Modeling: Vine Copula Handbook. World Scientific Publishing Co. Pte. Ltd..
- LIEBSCHER, E. (2008). «Construction of asymmetric multivariate copulas». Journal of Multivariate analysis, **99(10)**, pp. 2234–2250.
- MARDIA, K.V. (1970). Families of bivariate distributions. Griffin London.
- MARSHALL, A.W. y OLKIN, I. (1967a). «A generalized bivariate exponential distribution». *Journal of Applied Probability*, pp. 291–302.
- MARSHALL, A.W. y OLKIN, I. (1967b). «A multivariate exponential distribution». Journal of the American Statistical Association, pp. 30–44.
- MIKOSCH, T. (2006). «Copulas: Tales and facts». Extremes, 9(1), pp. 3–20.
- NELSEN, R. B. (2006). An introduction to copulas. Springer Verlag.
- NELSEN, R. B. (2007). «Extremes of nonexchangeability». Statistical Papers, 48, pp. 329–336.
- R DEVELOPMENT CORE TEAM (2011). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. http://www.R-project.org/
- SCHWEIZER, B. (2007). «Introduction to copulas». Journal of Hydrologic Engineering, 12, p. 346.
- SCHWEIZER, B. y WOLFF, EF (1981). «On nonparametric measures of dependence for random variables». *The Annals of Statistics*, pp. 879–885.

- SIBURG, K.F. y STOIMENOV, P.A. (2008). «Gluing copulas». Communications in Statistics Theory and Methods, 37(19), pp. 3124–3134.
- SKLAR, A. (1959). «Fonctions de répartition à n dimensions et leurs marges». Publ. Inst. Statist. Univ. Paris, 8(1), p. 11.
- TRIVEDI, P.K. y ZIMMER, D.M. (2007). Copula modeling: an introduction for practitioners. volumen 1. Now Pub.
- VAN DER VAART, A.W. y WELLNER, J.A. (1996). Weak convergence and empirical processes. Springer Verlag.
- YAN, JUN (2007). «Enjoy the Joy of Copulas: With a Package copula». Journal of Statistical Software, 21(4), pp. 1-21. http://www.jstatsoft.org/v21/i04/