



**UNIVERSIDAD NACIONAL  
AUTÓNOMA DE MÉXICO**

**FACULTAD DE ESTUDIOS SUPERIORES  
ACATLÁN**

**DESARROLLO HUMANO E INTENSIDAD  
MIGRATORIA: MODELADO CONJUNTO VÍA CÓPULAS**

**T E S I S**

**QUE PARA OBTENER EL TÍTULO DE  
ACTUARIO**

**PRESENTA:  
ISRAEL EMMANUEL AMBRIZ LOBATO**

**ASESOR: DR. ARTURO ERDELY RUIZ**

**MARZO, 2013**



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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

*Muere lentamente  
quien no gira el volante cuando está infeliz  
con su trabajo, o su amor,  
quien no arriesga lo cierto ni lo incierto para ir  
detrás de un sueño,  
quien no se permite, ni siquiera una vez en su vida,  
huir de los consejos sensatos...*

*Fragmento del Poema Muere Lentamente, Pablo Neruda.*

*“Porque somos mortales no somos libres.  
La única libertad que tenemos es la de  
hacer lo que nos gusta.”*

*Alejandro Jodorowsky.*

## Gracias:

- A mi Madre, Angeles** La persona que más quiero, admiro y respeto en este mundo, has sido mi mayor apoyo, me encuentro infinitamente agradecido contigo y con la vida misma por haberte elegido como mi madre. Me enseñaste a ser fuerte, a luchar por lo que deseo de la vida, siempre siguiendo mis ideales y principios. Por ti aprendí a disfrutar de los pequeños detalles de la vida. Eres la persona que más cree en mí, quien más se preocupa por que yo siga adelante. Todo lo anterior me forjó como el adulto y profesionista comprometido que soy ahora. Mamá, te amo.
- A mi Tía, Piri** Gran ejemplo de lucha, de triunfo, una persona comprometida y respetable. Gracias por confiar siempre en mí. Sin tu apoyo no hubiera sido lo mismo. Eres un pilar fundamental en mi vida. Gracias por tus consejos, enseñanzas y el gran amor que me brindas. Te quiero mucho tía.
- A mi Padre, Jose Luis** Soy quien soy gracias a ti. Sin tu apoyo no hubiera llegado a donde estoy, sin tu educación no tendría el carácter y la tenacidad que me han dado lo que tengo, sin tu amor no sería el hombre que soy.
- A mis Abuelitos** Por su infinito amor, por su memorable paciencia e incondicional apoyo. Los adoro Mamá Luisa y Papá Pepe
- A Toda mi Familia** Fuente inagotable de amor. En especial, gracias a mi hermanita Fer por quererme tanto y soportarme. Los quiero un montón a todos.
- A mis Amigos** Por su apoyo, por sus palabras de aliento y por su respetuosa y paciente espera al presente trabajo. En especial a Gaby Miranda y a Manu Pedraza. Gaby, te estaré agradecido siempre por todo lo que me enseñaste, el apoyo que me brindaste y el cariño que me diste. Gracias Manu, siempre me has extendido la mano cuando más lo he necesitado, gracias por la paciente y entregada revisión al presente documento.
- A mi Universidad** Mi queridísima UNAM que me ha dado múltiples satisfacciones en la vida.
- A mis Sinodales y Asesor** Por su amable atención y paciencia en la revisión del trabajo. En especial al Dr. Arturo Erdely, por ser quien me inspiró y motivó para la realización de la Tesis, por compartirme su conocimiento y por sus atenciones para la correcta realización del trabajo.

# Índice general

<b>Introducción</b>	<b>v</b>
<b>1. Cópulas</b>	<b>1</b>
1.1. Conceptos Preliminares . . . . .	7
1.2. Definición y Propiedades Básicas . . . . .	10
1.3. Teorema de Sklar . . . . .	13
1.4. Cópulas y Variables Aleatorias . . . . .	17
1.5. Las cotas de Fréchet-Hoeffding . . . . .	19
1.6. Cópula de Supervivencia . . . . .	22
1.7. Simetría . . . . .	22
1.8. Orden . . . . .	24
1.9. Simulación Bivariada con Cópulas . . . . .	25
1.10. Cópulas Arquimedianas . . . . .	27
1.11. Aproximaciones de la Cópula . . . . .	29
Cópula Empírica . . . . .	30
Suavizamiento vía polinomios de Bernstein . . . . .	31
Observaciones de la cópula . . . . .	31
1.12. Regresión con Cópulas . . . . .	31
<b>2. Dependencia</b>	<b>33</b>
2.1. Correlación Lineal de Pearson . . . . .	34
2.2. Concordancia . . . . .	36
2.3. Rho de Spearman . . . . .	38
2.4. Medidas de Dependencia . . . . .	40
2.5. Dependencias Estrictamente Monótonas . . . . .	42
<b>3. Migración y Desarrollo Humano</b>	<b>47</b>
3.1. Motivación . . . . .	48
3.2. Índice de Desarrollo Humano (IDH) . . . . .	49
3.3. Índice de Intensidad Migratoria (IIM) . . . . .	55
3.4. Ajuste Paramétrico Univariado . . . . .	62
IDH . . . . .	63
IIM . . . . .	64

<b>4. Segmentación y Ajuste Paramétrico</b>	<b>67</b>
4.1. Comportamiento Conjunto	68
Dependencia entre IDH e IIM	71
Comportamiento de la Cópula Empírica entre IDH e IIM	73
4.2. Algoritmo de Segmentación Bivariada	75
Segmentación en Dependencias Estrictamente Monótonas	76
Comportamiento por Submuestra	78
Comportamiento de la Cópula Empírica por Submuestra	87
4.3. Ajuste Paramétrico y Simulación	90
4.4. Regresión Lineal y con Cópulas	93
<b>Conclusiones</b>	<b>99</b>
<b>A. Algoritmos en R</b>	<b>105</b>
A.1. Análisis Descriptivo Conjunto	105
Cálculo de la cópula empírica: <code>cee</code>	105
Diagonal de la cópula empírica: <code>diagonal</code>	106
Contornos de la cópula empírica: <code>contour</code>	106
Pseudo-observaciones de la cópula: <code>pseudo</code>	107
Suavizamiento univariado vía polinomios de Bernstein: <code>Q</code>	107
Simulación bivariada conjunta: <code>rcopN</code>	108
Regresiones $\alpha$ -cuantil	109
Prueba de Simetría: <code>exchTest</code>	111
Cópulas incluidas en el paquete “cópula”	111
A.2. Análisis de la Dependencia	112
Prueba de Independencia: <code>indepTestSim</code> y <code>indepTest</code>	112
Medidas de Asociación empíricas	113
Concordancia y Dependencia de Cópulas Paramétricas	114
A.3. Modelado Conjunto	114
Segmentación: (1) inicio, (2) final	114
Bondad de Ajuste para Cópulas: <code>gofCopula</code>	115
Simulación con la cópula “Gluing”: <code>Gluing.Function</code>	116
Regresión $\alpha$ -cuantil semiparamétrica: <code>Regresion.C.IDH.IIM</code>	118
Regresión $\alpha$ -cuantil paramétrica: <code>Regresion.C.IDH.IIM.exp</code>	119
Regresión lineal: (1) <code>lm</code> , (2) <code>predict</code>	120
A.4. Tópicos Adicionales	121
Gráficos Boxplot: <code>boxplot</code>	121
<b>Bibliografía</b>	<b>123</b>

# Introducción

Las exigencias del mundo actual demandan modelos estadísticos cada vez más precisos. Exigencias que van desde normativas más estrictas para instituciones financieras y de seguros hasta el representar de manera más asertiva fenómenos macroeconómicos o sociodemográficos cada vez más complejos.

En la última década, se ha tomado más en serio el desarrollar modelos estadísticos en donde no sólo se explique por separado el comportamiento aislado de cada variable, sino modelos más flexibles que representen la interacción entre éstas, es decir modelos conjuntos. El hacer modelos conjuntos requiere tomar muy en consideración el concepto de *dependencia*. Analizar la dependencia entre variables aleatorias es un tema que suele tomarse muy a la ligera, o que simplemente, por facilidad, se opta por suponer independencia.

En la gran mayoría de las áreas en donde se ocupa a la estadística como una herramienta para la toma de decisiones, se conoce sólo al coeficiente de correlación lineal de Pearson como la fuente única que proporciona información sobre la relación entre variables continuas, sin saber que éste no puede decirnos todo lo que quisiéramos saber en relación a la estructura de dependencia de las mismas, además de que, cuando el coeficiente de correlación lineal llega a ser cercano a cero no existe una implicación directa respecto a la interacción entre las variables, pese a que esto se llega a asociar erróneamente con independencia; muchas otras son las limitantes que éste tiene, las cuales serán mencionadas en el desarrollo de este trabajo. Un método muy socorrido cuando se pretende describir la relación entre variables es el modelo de regresión lineal, mismo que tiene alcances sumamente limitados, además de la serie de supuestos que trae consigo previos y posteriores a su implementación. Cuando se desea hacer modelos en donde se involucre la función de distribución conjunta de variables aleatorias, por facilidad, es muy común que se suponga que las variables se ajusten a una distribución normal multivariada, implicando así normalidad en las distribuciones marginales univariadas, lo cual es bien sabido que es muy difícil que ocurra con fenómenos cotidianos.

Las *cóputas* son instrumentos matemáticos, que utilizados en la teoría de la probabilidad se pueden entender como la relación funcional entre la distribución conjunta de variable aleatorias y sus respectivas marginales univariadas. En ese sentido la *cóputa* guarda toda la información necesaria para lograr describir la interacción o dependencia entre las variables, por lo tanto proporciona la información suficientes para conocer de manera más acertada a la función de distribución conjunta. A lo largo de este trabajo se hará evidente el hecho de que con el uso de la *cóputa* subyacente a las variables de estudio se logran modelos conjuntos mucho más robustos respecto a los obtenidos con técnicas estadísticas convencionales, además de prescindir de supuestos poco alcanzables como los de normalidad, linealidad, entre otros.

En algunas ocasiones, la dependencia entre un par de variables puede llegar a ser un tanto complicada y difícil de representar con *cóputas* conocidas; en tales casos, el hacer ajustes con *cóputas* puede convertirse en un problema. Un camino que se puede tomar, es el de segmentar la muestra de estudio en dependencias más estables, con la finalidad de lograr ajustes paramétricos por separado a cada muestra resultante de la segmentación.

Para la presente investigación se buscó un par de variables que guardarán una dependencia poco clara y que además el coeficiente de correlación de Pearson para éstas fuera muy cercano a cero y en ese sentido no pudiera darnos información alguna. Fue así como se decidió modelar la relación conjunta que guardan el Desarrollo Humano y la Intensidad Migratoria en México; utilizando las variables Índice de Desarrollo Humano (IDH) e Índice de Intensidad Migratoria (IIM) del año 2000 a nivel municipal. Como se verá más adelante, estas variables guardan una relación un tanto caótica y difícil de describir en una primera instancia. A lo largo del trabajo se explicará el comportamiento del IIM con respecto al cambio del IDH.

La relación que guarda este par de variables fue analizado previamente en [PNUD \(2006\)](#) y se describió de la siguiente manera:

[...] los municipios más rezagados, los más pobres, tienen una intensidad migratoria menor, al igual que los municipios de mayor desarrollo humano. La mayor intensidad migratoria tiene lugar en municipios rezagados en términos relativos, pero no en los de mayor rezago. [...] esta relación tendría forma de U invertida, donde la mayor intensidad migratoria corresponde a un IDH de 0.716.

La descripción de la relación que se da en [PNUD \(2006\)](#) es un acercamiento básicamente visual sobre el comportamiento conjunto, y llega a quedar con términos ambiguos. Al analizar la descripción surgen las preguntas siguientes: ¿la relación es la misma a niveles bajos que a niveles altos

de desarrollo humano? ¿a qué se refiere con municipios con rezago en términos relativos? ¿qué significa que un par de variables guarden una relación en forma de «U» invertida? ¿qué ocurre con la relación de las variables en la parte superior de la «U» invertida, justo en los puntos donde la relación cambia de sentido? ¿en qué puntos del IDH la relación conjunta cambia de sentidos? ¿la gráfica de una «U» invertida es suficiente para representar la interacción entre las variables?

Con la finalidad de explicar de manera más clara y precisa la relación y dependencia que guardan el par de variables IDH e IIM a nivel municipal, además de proponer una vía de análisis en contraste a algunos de los puntos mencionados al principio, los objetivos generales, en orden de importancia, del presente documento son los siguientes:

- Aun cuando el coeficiente de correlación de Pearson es muy cercano a cero, determinar y modelar la estructura de dependencia que guardan el IDH y el IIM, logrando así explicar su relación conjunta. Se espera determinar un grado de dependencia en las variables que sea considerable y pueda llegar a ser modelado de forma paramétrica.
- Tomando en consideración que la dependencia entre el IDH y el IIM es caótica y difícil de modelar con una sola cópula conocida, se propone una vía de segmentación en dependencias más fáciles de explicar y representar: Algoritmo de Segmentación en Dependencias Estrictamente Monótonas (ASDEM). Se espera determinar los puntos para el IDH en los que la relación cambia de sentido.
- Una vez propuesto el algoritmo de segmentación y por ende el modelo conjunto, se comparan los modelos y alcances al explicar la relación conjunta con aquellos obtenidos al explicar el fenómeno con métodos de regresión lineal. Se espera que un modelo con cópulas sea más robusto y preciso, llevando a una explicación más clara y minuciosa sobre la relación conjunta.

Para lograr los fines antes mencionados se tienen los siguientes objetivos particulares para cada uno de los cuatro capítulos que componen esta tesis.

El primer capítulo tiene como finalidad general explicar a las funciones cópula desde el punto de vista de la teoría de la probabilidad para luego exponer algunos alcances de éstas en el campo de la estadística, teniendo como objetivos particulares: definir a las funciones cópula y presentar sus propiedades más relevantes, además de mostrar algunas familias de cópulas paramétricas; entender, vía el Teorema de Sklar, la relación entre las cópulas y las funciones de distribución conjunta y marginales de variables aleatorias; presentar a la clase de cópulas arquimedianas, cópulas paramétricas de fácil

construcción y propiedades deseables; explicar los métodos de simulación conjunta a partir de la cópula subyacente de un par de variables aleatorias; explicar la construcción de la cópula empírica y de la cópula Bernstein, siendo esta última una función suavizada vía polinomios de Bernstein de la cópula empírica; finalmente, se presenta un método de regresión con cópulas con base en un cuantil dado.

El capítulo dos tiene como objetivo central el analizar vías idóneas cuando se pretende describir la dependencia de variables aleatorias, con los siguientes objetivos particulares: explicar las limitaciones y alcances tanto conceptuales como de aplicación que se llegan a presentar con medidas de correlación lineal, concordancia y dependencia; exhibir la relevancia que tienen las cópulas en el estudio de estadísticos no paramétricos para analizar la “asociación” o “dependencia” que pueden llegar a presentar pares de variables aleatorias continuas; se define el concepto de dependencia estrictamente monótona, exhibiendo las ventajas de que un par de variables guarden una dependencia de este tipo.

El tercer capítulo pretende explicar de manera marginal el comportamiento de las variables IDH e IIM, con base en los siguientes objetivos particulares: presentar una motivación al estudio de la relación conjunta entre el desarrollo humano y los fenómenos migratorios; explicar las metodologías que llevan a la construcción del IDH y del IIM, describir a las variables IDH e IIM y a las variables con las que estos índices se construyen; hacer ajustes paramétricos univariados al IDH y al IIM.

Finalmente, el cuarto capítulo se centra en modelar el comportamiento conjunto que presentan las variables IDH, IIM a fin de llegar a un ajuste paramétrico con cópulas que explique dicha relación, este capítulo tiene los siguientes objetivos particulares: describir el comportamiento conjunto entre el IDH y el IIM con base en las medidas de asociación presentadas en el capítulo dos y el comportamiento de la cópula empírica subyacente a las variables; proponer y aplicar un algoritmo de segmentación con base en la variable IDH, a fin de lograr submuestras de observaciones en donde la relación entre éstas se logre explicar de una manera más sencilla para cada submuestra (ASDEM); describir el comportamiento de las variables por cada submuestra para lograr un análisis más minucioso sobre la interacción entre el IDH y el IIM; buscar un ajuste paramétrico con cópulas a las variables en cada submuestra, para luego pegar las cópulas en una sola con el método “Gluing Copulas”; hacer simulaciones conjuntas con base en la cópula “Gluing” y compararlas con simulaciones conjuntas con base en la cópula Bernstein, para luego contrastar las simulaciones paramétricas y las no paramétricas con los datos originales; llegar a modelos de regresión con cópulas que expliquen la relación entre las variables IDH e IIM, mismos que

serán comparados con modelos de regresión lineal.

Cabe destacar que la presente investigación se limita a trabajar con cópulas bivariadas. Todos los conceptos dentro del desarrollo teórico de cópulas y sus aplicaciones pueden ser extendidos a dimensiones mayores, pero vale la pena mencionar que los avances teóricos aún no son los suficientes para lograr que la implementación de modelos con cópulas sea eficiente o en algunas ocasiones posible, para algunas aplicaciones, a partir de dimensiones superiores a cuatro.

# Capítulo 1

## Cóputas

Este primer capítulo tiene como finalidad dar a conocer de manera formal las características y propiedades de una cópula, primero, como concepto en la teoría de la probabilidad y luego entenderla como herramienta para explicar el comportamiento conjunto entre variables aleatorias continuas en el campo de la estadística. Antes de comenzar con los conceptos que llevan al entendimiento de las cópulas haremos una introducción histórica sobre las mismas.

¿Qué es una cópula?

Desde un punto de vista, las cópulas son funciones que “unen” o “juntan” distribuciones multivariadas a sus funciones de distribución marginales de una dimensión. En ese sentido, una cópula se puede entender como la *relación funcional* que existe entre una función de distribución conjunta y sus respectivas marginales univariadas. Alternativamente, las cópulas son funciones de distribución multivariada tales que sus marginales de una dimensión son uniformes en el intervalo  $(0, 1)$ .

Según la RAE<sup>1</sup>, entre otros, tiene estos dos significados: (1) Atadura, ligamento de algo con otra cosa, (2) Término que une el predicado con el sujeto. Más adelante, se hace notar cómo el significado (2) tuvo cierta relevancia en el descubrimiento de las cópulas.

La noción de cópula fue introducida por Abe Sklar en 1959, contestando a la problemática planteada por Maurice Fréchet acerca de la relación entre una función de probabilidad multidimensional y sus marginales de menor dimensión. El problema es muy general, pero puntualizando en la dimensión dos es el siguiente: 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

---

<sup>1</sup>Real Academia Española

$(\Omega, \mathfrak{F}, \mathbb{P})$ , ¿Qué se puede decir del conjunto  $\Gamma(F_1, F_2)$  de funciones de distribución bivariadas con marginales  $F_1$  y  $F_2$ ? Es inmediato notar que el conjunto  $\Gamma(F_1, F_2)$ , conocido como las *Clases de Fréchet de  $F_1$  y  $F_2$* , es no vacío, ya que si  $X_1$  y  $X_2$  son independientes entonces  $H$ , la función de distribución conjunta de  $X_1$  y  $X_2$ , es igual al producto de las distribuciones marginales, es decir,  $(x_1, x_2) \mapsto H(x_1, x_2) = F_1(x_1)F_2(x_2)$ , así definida  $H$  siempre pertenece a  $\Gamma(F_1, F_2)$ . Pero, no era claro cómo eran los otros elementos de  $\Gamma(F_1, F_2)$ .

Estudios preliminares abordaron este problema, pero [Sklar \(1959\)](#) es quien obtiene el resultado más profundo al respecto, introduciendo la noción y el nombre de una *cópula*, con el teorema que ahora lleva su nombre. Citando a [Sklar \(1996\)](#):

[...] Bert (Schweizer) y yo estábamos progresando en nuestro trabajo en espacios métricos estadísticos, extendiendo la sugerencia de Menger consideramos conveniente comunicarle nuestros resultados a Fréchet. ... Esto dio inicio a un intercambio de cartas con Fréchet, durante ese tiempo él me envió varios paquetes de reimpresiones, principalmente su trabajo y el de sus colegas sobre distribuciones con marginales dadas. [...] yo particularmente distingo el trabajo de [Dall'Aglio \(1959\)](#), fue importante para mucho de nuestro trabajo subsecuente. Aunque en ese momento, la reimpresión mas significativa para mi fue [Féron \(1956\)](#).

Féron, estudiando las distribuciones de tres dimensiones introdujo funciones auxiliares, definidas en el cubo unitario, que conectaban dichas distribuciones con sus marginales de una dimensión. Yo observé que funciones similares podrían ser definidas en el n-cubo unitario para todo  $n \geq 2$  y similarmente podrían servir para ligar distribuciones n-dimensionales con sus marginales de una dimensión. Al haber trabajado con las propiedades básicas de estas funciones, escribí sobre esto a Fréchet, en inglés. Él me pidió que lo hiciera en francés. Mientras escribía esto, decidí que necesitaba un nombre para estas funciones. Conociendo la palabra *cópula* como un término gramatical para una palabra o expresión que enlaza al sujeto con el predicado, consideré que éste podría ser un nombre apropiado para una función que enlaza una distribución multidimensional a sus marginales de una dimensión, y la usé como tal. Fréchet recibió mis notas, corrigió una declaración matemática, hizo algunas correcciones a mi francés, y publicó las notas en el "Statistical Institute of the University of Paris"<sup>2</sup> como [Sklar \(1959\)](#).

---

<sup>2</sup>Instituto de Estadística de la Universidad de París

Previo al teorema de Sklar y a todo este intercambio de resultados con Fréchet y sus colegas, existió un personaje llamado Wassily Hoeffding quien en su tesis doctoral ([Hoeffding \(1940\)](#)) presenta muchos de los resultados básicos de cópulas, él encontró funciones de distribución bivariadas “estandarizadas” tal que su soporte está contenido en el cuadrado  $[-1/2, 1/2]^2$  y que tienen marginales uniformes en el intervalo  $[-1/2, 1/2]$ . Como opina [Schweizer \(1991\)](#):

[...]si Hoeffding hubiera elegido el cuadrado unitario  $[0, 1]^2$  para su normalización en lugar de  $[-1/2, 1/2]^2$ , él hubiera descubierto a las cópulas.

Hoeffding también obtuvo las mejores cotas posibles para estas funciones, caracterizó las distribuciones correspondientes a éstas (“dependencia funcional”) y estudió medidas de dependencia que son “invariantes bajo cambios de escala”, *i.e.* invariantes bajo transformaciones estrictamente crecientes. Hoeffding, quien era ruso, estudió su doctorado en la universidad de Berlín, y desde que comenzó le preocupaba el ambiente difícil con algunos académicos ganando altos cargos debido a su fuerte apoyo a los nazis y otros que simplemente habían dejado la universidad como Richard von Mises. El único curso de estadística matemática lo impartía Alfred Klose, él había sido alumno de von Mises y llevaba el curso con un libro escrito por él. En general, algunos años posteriores a sus estudios de doctorado no fueron los mejores para que Hoeffding pudiera seguir con lo trabajado en su tesis, pero, pese a lo que se comenta en [Gani \(1982\)](#) aquella tesis y otros trabajos de la carrera de Hoeffding fueron traducidos al inglés y publicados por [Fisher y Sen \(1994\)](#) tres años después de la muerte de Hoeffding. De [Gani \(1982\)](#):

Mi Doktorvater<sup>3</sup> o supervisor de Ph.D.<sup>4</sup> fue Klose. Yo escogí el tema de tesis y trabajé en ella gran parte por mi cuenta, con algunas sugerencias y el aliento de él. Él era un alemán báltico y tenía sus propias ideas acerca de los rusos. Él me advirtió que me abstuviera de hacer afirmaciones exageradas en mi tesis que no pudiera justificar, como él pensaba, los rusos eran propensos a hacerlo.

Sin el conocimiento del trabajo de Hoeffding, [Fréchet \(1951\)](#) obtuvo muchos de los mismos resultados, que llevaron a terminos como “Cotas de Fréchet” o “Clases de Fréchet”. En reconocimiento al trabajo de Hoeffding, actualmente nos referimos a ellas como “Cotas de Fréchet-Hoeffding” y “Clases de Fréchet-Hoeffding”.

---

<sup>3</sup>Del alemán, supervisor de doctorado.

<sup>4</sup>Doctorado en investigación (Ph.D. o D.Phil., siglas en latín de Philosophiae Doctor o Doctor Philosophiae, respectivamente)

¿Por qué las cópulas son importantes en el estudio de la probabilidad y la estadística?

Entre 1959 y 1976, las cópulas fueron usadas en el desarrollo de la teoría de espacios métricos de probabilidad, principalmente en el estudio de operaciones binarias en el espacio de funciones de distribución de probabilidad. Subsecuentemente, se descubrió que con las cópulas se podían definir medidas no paramétricas de dependencia entre variables aleatorias y desde entonces comenzaron a jugar un papel importante en la probabilidad y la estadística matemática.

Como comenta [Fisher \(1997\)](#) en el primer volumen de actualización de la “Encyclopedia of Statistical Sciences”<sup>5</sup>:

Las *cópulas* son de interés para los estadísticos por dos grandes razones: Primero, como una vía de estudiar medidas de dependencia invariantes bajo cambios de escala; y segundo, como un punto de partida para construir familias de distribuciones bivariadas, algunas veces con la finalidad de simulación.

El evento que propició el interés de la comunidad estadística en las cópulas ocurrió a mediados de los setentas, cuando [Schweizer \(2007\)](#), en sus propias palabras:

[...] por accidente, releyendo un artículo de Rényi, titulado “Sobre medidas de dependencia” me di cuenta que [él] podía fácilmente construir medidas de dependencia usando cópulas.

El primer trabajo donde se relaciona de manera explícita a las cópulas con el estudio de dependencia entre variables aleatorias es [Schweizer y Wolff \(1981\)](#). En este artículo, Schweizer y Wolff discuten y modifican los criterios para una medida de dependencia entre un par de variables aleatorias expuestos en [Rényi \(1959\)](#), presentan las propiedades básicas de invarianza de cópulas bajo transformaciones estrictamente monótonas de variables aleatorias (véase Teoremas 1.4.5 y 1.4.6), e introducen la medida de dependencia conocida como  $\sigma$  de [Schweizer & Wolff](#) (véase sección 2.4).

Sin embargo, durante algunos años, el capítulo 6 del libro fundamental [Schweizer y Sklar \(1983\)](#), dedicado a la teoría de espacios métricos probabilísticos, fue la principal fuente de información básica sobre cópulas. De nuevo en las palabras de [Schweizer \(2007\)](#):

Después de la publicación de este artículo y del libro ... el ritmo se aceleró a medida que más ... estudiantes y colegas se vieron

---

<sup>5</sup> “Enciclopedia en Ciencias Estadísticas”

involucrados. Por otra parte, ya que el interés en las cuestiones de dependencia estadística va en aumento, otros vinieron al tema desde diferentes direcciones. En 1986 el artículo titulado tentadoramente “Enjoy de joy of copulas”<sup>6</sup> por [Genest y MacKay \(1986\)](#) capto más atención.

¿Por qué y cómo las cópulas han tenido un crecimiento sumamente significativo en los últimos años?

En 1990, Dall’Aglío organizó la primera conferencia dedicada a cópulas, acertadamente llamada “Probability distributions with given marginals”<sup>7</sup>. Ésta resultó ser la primera de una serie de conferencias que ayudó mucho al desarrollo del campo, ya que cada una de ellas ofreció la oportunidad de presentar algunos resultados y de aprender de otros investigadores, estas conferencias se celebraron en Seattle en 1993, en Praga en 1996, en Barcelona en el año 2000, en Québec en 2004, en Tartu 2007, en Alemania 2007, en São Paulo en 2010, Montréal en 2011 y también la conferencia “Conference for the 50th anniversary of copulas”<sup>8</sup> en Lecce (Italia), el 11 de Junio de 2009.

A finales de los noventas, la popularidad de la noción de cópula creció. Dos libros acerca de cópulas aparecieron y se convirtieron en las referencias fundamentales durante la siguiente década: (1) “Multivariate Models and Dependence Concepts. Monographs on Statistics and Applied Probability”<sup>9</sup>, un libro dedicado a modelos multivariados con una gran parte asignada a cópulas y familias de cópulas, [Joe \(1997\)](#). (2) “An Introduction to Copulas”<sup>10</sup>, la primera y la única introducción a cópulas hasta el momento escrita, estudia familias de cópulas, construcción de cópulas, cópulas arquimedianas, medidas de dependencia entre otros tópicos importantes, [Nelsen \(2006\)](#).

Pero la razón principal de este creciente interés es el descubrimiento que han tenido algunos investigadores de la noción de cópula aplicada a varios campos, como finanzas. A continuación se explica brevemente este fenómeno citando a [Embrechts \(2009\)](#):

Como hemos visto hasta ahora, la noción de cópula es a la vez natural, así como fácil para buscar funciones de distribución

---

<sup>6</sup> “Disfruta el placer de las cópulas”

<sup>7</sup> “Distribuciones de probabilidad con marginales dadas”

<sup>8</sup> “Conferencia por el 50 aniversario de las Cópulas.”

<sup>9</sup> “Modelos Multivariados y Conceptos de Dependencia. Monografía sobre Estadística y Probabilidad Aplicada.”

<sup>10</sup> “Una Introducción a Cópulas”

multivariadas. Pero ¿por qué somos testigos de un crecimiento tan increíble en artículos publicados a partir del final de la década de los noventa (recordar que el concepto se remonta a los años cincuenta e incluso antes, pero no con ese nombre)? Yo puedo dar tres razones: finanzas, finanzas, finanzas. En los ochentas y noventas experimentamos un desarrollo explosivo de metodologías cuantitativas en la administración de riesgos en finanzas y seguros, muchas de las cuales fueron impulsadas por las nuevas pautas regulatorias o el desarrollo de nuevos productos.

La llegada de las cópulas en las finanzas originó una gran cantidad de investigaciones sobre las mismas y, especialmente, sobre las aplicaciones. Al mismo tiempo, diferentes campos como la hidrología han descubierto la importancia de este concepto para la construcción de modelos multivariados más flexibles. Hoy en día, es casi imposible dar una explicación completa de todas las aplicaciones de cópulas a los muchos campos en donde pueden ser utilizadas. Como se comenta en [Schweizer \(2007\)](#):

La “era de i.i.d.” ha terminado: y cuando la dependencia se toma en serio, las cópulas de forma natural entran en juego. Queda para la comunidad estadística en general el reconocer este hecho. Y cuando todos los textos de estadística contengan una sección o capítulo de cópulas, el tema sera de actualidad.

Sin embargo, varias críticas han sido recientemente levantadas sobre las cópulas y sus aplicaciones, y varias personas comenzaron a hablar de la “Locura de Cópulas”, por ejemplo, la discusión muy interesante relacionada con el artículo [Mikosch \(2006\)](#) (ver también [Joe \(2006\)](#), [Genest y Rémillard \(2006\)](#), [Vries y Zhou \(2006\)](#), entre otros artículos en repuesta al del Doctor Mikosch). A continuación una cita al respecto de [Durante y Sempi \(2010\)](#):

Desde nuestro punto de vista, estas críticas fueron una reacción muy natural a tan amplia difusión de las aplicaciones de las cópulas, no siempre de una manera bien motivada. Hay que decir que varias personas han interpretado erróneamente a las cópulas como la solución a “todos los problemas de dependencia estocástica”. ¡Esto definitivamente no es el caso! Las cópulas son una herramienta indispensable para la comprensión de varios problemas sobre la dependencia estocástica, pero no son la “panacea” para todos los modelos estocásticos.

En su implementación, las cópulas pueden llegar a implicar procesos computacionales arduos, más aún, cuando las dimensiones de éstas aumentan. Actualmente, ya son varios los programas matemáticos y estadísticos que cuentan con implementaciones con cópulas, por ejemplo: Mathworks Matlab, Wolfram Mathematica, Microsoft Excel (con un complemento adicional),

RExcel, S-plus, SAS, XploRe y R. Sin embargo, la gran mayoría de éstos cuentan con un catálogo muy limitado de cópulas, en general sólo cuentan con las 5 familias más comunes: normal, t, Clayton, Gumbel y Frank. Además, los métodos y herramientas de ajuste, en general, también se encuentran limitados. El único paquete que, en cierto sentido, se encuentra más actualizado es R, esto por su naturaleza de “software libre” (véase Bacigál (2012) para mayor detalle sobre el contenido con respecto a cópulas de cada herramienta.). Aún así, esto nos lleva a muchos retos que están latentes actualmente en la teoría y aplicación de las cópulas tanto como un concepto probabilístico hasta como una herramienta estadística. Concluiremos esta introducción histórica citando a Durante y Sempi (2010):

A pesar de esta amplia gama de interés acerca de las cópulas, todavía creemos que este concepto se encuentra en “su infancia” y otras tantas investigaciones pueden (y deben) ser conducidas con el fin de subrayar si las cópulas, o conceptos relacionados con cópulas, pueden ser realmente considerados como un “fuerte” concepto matemático de interés en varias aplicaciones.

Con esta gran motivación histórica ya estamos en posición de entrar en materia y comprender el concepto de cópula.

## 1.1. Conceptos Preliminares

Esta sección tiene por finalidad exhibir conceptos básicos que nos serán de mucha ayuda para definir a la función cópula, sus características y propiedades.

Si  $\mathfrak{R}$  denota el conjunto de los número reales  $(-\infty, \infty)$ ,  $\bar{\mathfrak{R}}$  denota al conjunto de los reales extendidos  $\mathfrak{R} \cup \{-\infty, \infty\} = [-\infty, \infty]$ ;  $\mathfrak{R}^2$  y  $\bar{\mathfrak{R}}^2$  representan al plano real  $\mathfrak{R} \times \mathfrak{R}$  y el plano real extendido  $\bar{\mathfrak{R}} \times \bar{\mathfrak{R}}$  respectivamente.

Un rectángulo en  $\bar{\mathfrak{R}}^2$  es el producto cartesiano  $\mathcal{B}$  de dos intervalos cerrados:  $\mathcal{B} = [x_1, x_2] \times [y_1, y_2]$ , donde los vértices del rectángulo son los puntos  $(x_1, y_1)$ ,  $(x_1, y_2)$ ,  $(x_2, y_1)$ ,  $(x_2, y_2)$ . El cuadrado unitario  $\mathcal{I}^2$  es el producto  $\mathcal{I} \times \mathcal{I}$ , donde  $\mathcal{I} = [0, 1]$ .

Una función real  $H$  de doble entrada es aquella con dominio,  $DomH$ , subconjunto de  $\bar{\mathfrak{R}}^2$  y rango,  $RanH$ , subconjunto de  $\mathfrak{R}$ .

**Definición 1.1.1.** Sean  $\mathcal{S}_1$  y  $\mathcal{S}_2$  subconjuntos no vacíos de  $\bar{\mathfrak{R}}$ , y  $H$  una función real de doble entrada tal que el  $DomH = \mathcal{S}_1 \times \mathcal{S}_2$ . Sea  $\mathcal{B} = [x_1, x_2] \times [y_1, y_2]$  un rectángulo cuyos vértices están contenidos en el  $DomH$ . Entonces, el  $H$ -volumen de  $\mathcal{B}$  esta dado por:

$$V_H(\mathcal{B}) = H(x_2, y_2) - H(x_1, y_2) - H(x_2, y_1) + H(x_1, y_1). \quad (1.1.1)$$

**Definición 1.1.2.** Una función real  $H$  de doble entrada es 2-creciente si  $V_H(\mathcal{B}) \geq 0$ , para todo rectángulo  $\mathcal{B}$  cuyos vértices están contenidos en

*DomH.*

Cuando una función es *2-creciente*, en ocasiones se hace referencia al *H-volumen* ó a la *H-medida* de algún rectángulo  $\mathcal{B}$ . Algunos autores se refieren a las funciones *2-crecientes* como *cuasi-monótonas*.

**Ejemplo 1.1.1.** Sea  $\Pi$  una función definida en  $\mathcal{I}^2$  por  $\Pi(x, y) = x \cdot y$ . Entonces  $\Pi$  es *2-creciente*:

Si  $u_1, u_2, v_1, v_2 \in \mathcal{I}$  y además  $u_1 \leq u_2$  y  $v_1 \leq v_2$  entonces:

$$\begin{aligned} V_{\Pi}([u_1, u_2] \times [v_1, v_2]) &= u_2 v_2 - u_2 v_1 - u_1 v_2 + u_1 v_1 \\ &= u_2(v_2 - v_1) - u_1(v_2 - v_1) \\ &= (u_2 - u_1)(v_2 - v_1) \\ &\geq 0, \end{aligned}$$

ya que si  $u_1 \leq u_2$ , entonces  $u_2 - u_1 \geq 0$ , de manera análoga para  $v_1 \leq v_2$ .  $\square$

Cabe mencionar que ninguno de estos enunciados es consecuencia uno del otro: «H es *2-creciente*»  $\not\Rightarrow$  «H es creciente en cada uno de sus argumentos», y «H es creciente en cada uno de sus argumentos»  $\not\Rightarrow$  «H es *2-creciente*», como se observa en los siguientes ejemplos:

**Ejemplo 1.1.2.** Sea  $H$  una función definida en  $\mathcal{I}^2$  por  $H(x, y) = (2x - 1)(2y - 1)$ .

Si  $u_1, u_2, v_1, v_2 \in \mathcal{I}$  y además  $u_1 \leq u_2$  y  $v_1 \leq v_2$  entonces:

$$\begin{aligned} V_H([u_1, u_2] \times [v_1, v_2]) &= (2u_2 - 1)(2v_2 - 1) - (2u_1 - 1)(2v_2 - 1) \\ &\quad - (2v_1 - 1)(2u_2 - 1) + (2u_1 - 1)(2v_1 - 1) \\ &= (2u_2 - 1)[(2v_2 - 1) - (2v_1 - 1)] \\ &\quad + (2u_1 - 1)[(2v_1 - 1) - (2v_2 - 1)] \\ &= (2u_2 - 1)[2(v_2 - v_1)] + (2u_1 - 1)[2(v_1 - v_2)] \\ &= 2(v_2 - v_1)[(2u_2 - 1) - (2u_1 - 1)] \\ &= 2(v_2 - v_1)[2(u_2 - u_1)] \\ &= 4(v_2 - v_1)(u_2 - u_1) \\ &\geq 0, \end{aligned}$$

ya que  $(u_2 - u_1), (v_2 - v_1) \geq 0$ , entonces  $H$  es *2-creciente*, pero si tomamos a  $y$  un número fijo en el intervalo  $(0, \frac{1}{2})$ ,  $H$  es una función decreciente para  $x$ , ya que  $-1 < 2y - 1 < 0$  y  $\forall x \leq x', 2x - 1 \leq 2x' - 1$ , por lo tanto  $(2x' - 1)(2y - 1) \leq (2x - 1)(2y - 1)$ , de manera análoga si  $x \in (0, \frac{1}{2})$ ,  $H$  es decreciente  $\forall y$ .  $\square$

**Ejemplo 1.1.3.** Sea  $H$  una función definida en  $\mathcal{I}^2$  por  $H(x, y) = \text{máx}(x, y)$ . Entonces  $H$  es una función monótona creciente de  $x$  y  $y$ , ya que si fijamos  $y = y_0$  entonces para todo  $x \leq x'$ ,  $\text{máx}(x, y_0) \leq \text{máx}(x', y_0)$ , de forma análoga si fijamos  $x = x_0$ . Por otro lado

$$\begin{aligned} V_H(\mathcal{I}^2) &= H(1, 1) - H(1, 0) - H(0, 1) + H(0, 0) \\ &= \text{máx}(1, 1) - 2 \text{máx}(0, 1) + \text{máx}(0, 0) \\ &= 1 - 2 + 0 \\ &= -1, \end{aligned}$$

por lo tanto  $H$  no es una función 2-creciente.  $\square$

**Lema 1.1.1.** Sean  $S_1$  y  $S_2$  dos subconjuntos no vacíos de  $\mathfrak{R}$ , y sea  $H$  una función 2-creciente con dominio  $S_1 \times S_2$ . Sean  $x_1, x_2$  en  $S_1$  con  $x_1 \leq x_2$ , y sean  $y_1, y_2$  en  $S_2$  con  $y_1 \leq y_2$ . Entonces la función  $t \mapsto H(t, y_2) - H(t, y_1)$  es una función no decreciente en  $S_1$ , y la función  $t \mapsto H(x_2, t) - H(x_1, t)$  es no decreciente en  $S_2$ .

**Definición 1.1.3.** Sean  $S_1, S_2 \subseteq \mathfrak{R}$  no vacíos, t.q.  $a_1 := \inf(S_1)$  y  $a_2 := \inf(S_2)$  y sea  $H$  una función real de doble entrada con  $\text{Dom}H = S_1 \times S_2$ .  $H$  está fijada si:  $H(x, a_2) = 0 = H(a_1, y) \forall (x, y) \in S_1 \times S_2$ .

Una aplicación inmediata del Lema 1.1.1 es la siguiente.

**Lema 1.1.2.** Sean  $S_1, S_2 \subseteq \mathfrak{R}$  no vacíos, y sea  $H$  una función 2-creciente fijada con dominio  $S_1 \times S_2$ . Entonces  $H$  es no decreciente en cada argumento.

**Definición 1.1.4.** Sean  $b_1 = \sup(S_1)$ ,  $b_2 = \sup(S_2)$ , una función  $H$  de  $S_1 \times S_2$  a  $\mathfrak{R}$ , tiene marginales  $F$  y  $G$  dadas de la siguiente manera:

$$\begin{aligned} \text{Dom}F &= S_1, \text{ y } F(x) = H(x, b_2) \forall x \in S_1; \\ \text{Dom}G &= S_2, \text{ y } G(y) = H(b_1, y) \forall y \in S_2. \end{aligned}$$

**Ejemplo 1.1.4.** Sea  $H$  una función con dominio  $[-1, 1] \times [0, \infty]$  dada por:

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

$H$  esta fijada ya que  $H(x, 0) = [(x+1)(1-1)]/(x+2-1) = 0$ , y  $H(-1, y) = [(-1+1)(e^y - 1)]/(-1+2e^y - 1) = 0$ ; y  $H$  tiene marginales  $F$  y  $G$  dadas por:

$$\begin{aligned} F(x) &= H(x, \infty) \\ &= \lim_{y \rightarrow \infty} \frac{(x+1)(e^y - 1) \frac{1}{e^y}}{x + 2e^y - 1 \frac{1}{e^y}} \\ &= (x+1) \lim_{y \rightarrow \infty} \frac{1 - \frac{1}{e^y}}{\frac{(x-1)}{e^y} + 2} \\ &= (x+1)/2, \end{aligned}$$

$$\begin{aligned}
G(y) &= H(1, y) \\
&= \frac{(1+1)(e^y - 1)}{1 + 2e^y - 1} \\
&= \frac{2(e^y - 1)}{2e^y} \\
&= 1 - e^{-y}.
\end{aligned}$$

A continuación un lema importante sobre funciones 2-crecientes fijadas con marginales.

**Lema 1.1.3.** Sea  $S_1, S_2 \subseteq \bar{\mathfrak{R}}$ , y sea  $H$  una función 2-creciente fijada con marginales, t.q.  $DomH = S_1 \times S_2$ .

Sean  $(x_1, y_1)$  y  $(x_2, y_2)$ , puntos arbitrarios en  $S_1 \times S_2$ . Entonces:

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

## 1.2. Definición y Propiedades Básicas

La finalidad de esta sección es definir a las funciones «Cópula» y enunciar sus características y propiedades. Lo primero que se hará será definir a las subcópulas como una familia de funciones fijadas, 2-crecientes y con marginales; para luego definir a las funciones «Cópula» como subcópulas con dominio  $\mathcal{I}^2$ .

**Definición 1.2.1.** Una subcópula bivariada es una función  $C'$  con las siguientes propiedades:

1. El  $DomC' = S_1 \times S_2$ , donde  $S_1, S_2 \subseteq \mathcal{I}$  y  $\{0\}, \{1\} \in S_1, S_2$ ;
2.  $C'$  está fijada y es 2-creciente;
3.  $\forall u \in S_1$  y  $\forall v \in S_2$ ,  $C'(u, 1) = u$  y  $C'(1, v) = v$ .

Podemos notar que  $\forall (u, v) \in DomC'$ ,  $0 \leq C'(u, v) \leq 1$ .

**Definición 1.2.2.** Una cópula bivariada es una función  $\mathcal{C} : \mathcal{I}^2 \rightarrow \mathcal{I}$  con las siguientes propiedades:

1. Está fijada. Para cualquier  $u, v \in \mathcal{I}$ ,  $\mathcal{C}(u, 0) = 0 = \mathcal{C}(0, v)$ ;
2. Tiene marginales  $u$  y  $v$  respectivamente para cada argumento.  $\mathcal{C}(u, 1) = u$ ,  $\mathcal{C}(1, v) = v$ ;
3. Es 2-creciente. Para cualquier  $u_1, u_2, v_1, v_2 \in \mathcal{I}$ , tales que  $u_1 \leq u_2$  y  $v_1 \leq v_2$ , entonces  $\mathcal{C}(u_2, v_2) - \mathcal{C}(u_2, v_1) - \mathcal{C}(u_1, v_2) + \mathcal{C}(u_1, v_1) \geq 0$ .

La distinción respecto al dominio entre una cópula y una subcópula puede parecer poco significativa, pero será de gran interés en la siguiente sección al presentar el teorema de Sklar.

**Ejemplo 1.2.1.** Las siguientes funciones definidas en  $\mathcal{I}^2$  y que van a  $\mathcal{I}$ , son cópulas

1.  $\mathcal{M}(u, v) = \min(u, v)$ ;
2.  $\mathcal{W}(u, v) = \max(u + v - 1, 0)$ ;
3.  $\Pi(u, v) = u \cdot v$ , donde por el Ejemplo 1.1.1 es 2-creciente, también  $\Pi(0, v) = 0 \cdot v = 0 = u \cdot 0 = \Pi(u, 0)$  por lo tanto está fijada y además  $\Pi(u, 1) = u \cdot 1 = u$  y  $\Pi(1, v) = 1 \cdot v = v$ , entonces  $\Pi$  es cópula.

Las pruebas para las primeras dos cópulas se encuentran en [Pérez \(2007\)](#).

**Ejemplo 1.2.2.** Sea  $\alpha$  cualquier número en el intervalo  $(0, 1)$ , entonces la combinación lineal convexa entre las cópulas  $\mathcal{M}$  y  $\mathcal{W}$  a partir de  $\alpha$ ,  $\mathcal{C} = \alpha\mathcal{M} + (1 - \alpha)\mathcal{W}$ , también es una cópula.

Durante el desarrollo de todo el trabajo se emplearan algunas cópulas paramétricas, las cuales en su totalidad son estudiadas en [Nelsen \(2006\)](#).

El siguiente teorema explica la importancia de las cópulas en los puntos 1 y 2 del Ejemplo 1.2.1.

**Teorema 1.2.1.** Sea  $\mathcal{C}'$  una subcópula, entonces  $\forall(u, v) \in \mathcal{I}$ ,

$$\max(u + v - 1, 0) \leq \mathcal{C}'(u, v) \leq \min(u, v). \quad (1.2.1)$$

Las cotas para una subcópula enunciadas en el Teorema 1.2.1 son conocidas como las cotas de *Fréchet-Hoeffding*.

Dado que toda cópula es una subcópula las cotas dadas en (1.2.1) aplican para toda cópula, de hecho estas mismas son también cópulas y son comúnmente denotadas como  $\mathcal{M}(u, v) = \min(u, v)$  y  $\mathcal{W}(u, v) = \max(u + v - 1, 0)$ . Por lo tanto la versión para cópulas de las cotas de *Fréchet-Hoeffding* es la siguiente:  $\forall(u, v) \in \mathcal{I}^2$ ,

$$\mathcal{W}(u, v) \leq \mathcal{C}(u, v) \leq \mathcal{M}(u, v). \quad (1.2.2)$$

Existe otra cópula importante que será utilizada con frecuencia en el desarrollo del trabajo, la *cópula producto*,  $\Pi(u, v) = uv$ .

A partir del Lema 1.1.3 se da el siguiente teorema, el cual establece la continuidad de subcópulas y por tanto de cópulas, vía una condición de Lipschitz en  $\mathcal{I}^2$ .

**Teorema 1.2.2.** Sea  $\mathcal{C}'$  una subcópula. Entonces  $\forall (u_1, u_2), (v_1, v_2) \in \text{Dom}\mathcal{C}'$ ,

$$|\mathcal{C}'(u_2, v_2) - \mathcal{C}'(u_1, v_1)| \leq |u_2 - u_1| + |v_2 - v_1|. \quad (1.2.3)$$

*Demostración.* La prueba es consecuencia directa del Lema 1.1.3.  $\square$

Por lo tanto  $\mathcal{C}'$  es *uniformemente continua* en su dominio.

**Definición 1.2.3.** Sea  $\mathcal{C}$  una cópula y sea  $a$  un número en  $\mathcal{I}$ . La sección horizontal de  $\mathcal{C}$  en  $a$ , es la función que va de  $\mathcal{I}$  a  $\mathcal{I}$  dada por  $t \mapsto \mathcal{C}(t, a)$ ; la sección vertical de  $\mathcal{C}$  en  $a$  es la función que va de  $\mathcal{I}$  a  $\mathcal{I}$  dada por  $t \mapsto \mathcal{C}(a, t)$ ; y la sección diagonal de  $\mathcal{C}$  es la función que va de  $\mathcal{I}$  a  $\mathcal{I}$  dada por  $\delta_{\mathcal{C}}(t) = \mathcal{C}(t, t)$ .

El siguiente corolario es consecuencia inmediata del Lema 1.1.2 y del Teorema 1.2.2.

**Corolario 1.2.1.** Las secciones horizontal, vertical y diagonal son no decrecientes y uniformemente continuas.

Una manera simple pero útil de presentar la gráfica de una cópula es con un diagrama de contornos (Conway (1979)), que son gráficas de sus *conjuntos de nivel*, dados por los conjuntos en  $\mathcal{I}^2$  tal que  $\mathcal{C}(u, v) = a$ , con  $a$  una constante, para constantes específicas en  $\mathcal{I}$ . La Figura 1.1 nos muestra la gráfica de los contornos de las cópulas  $\mathcal{M}$ ,  $\Pi$  y  $\mathcal{W}$ .

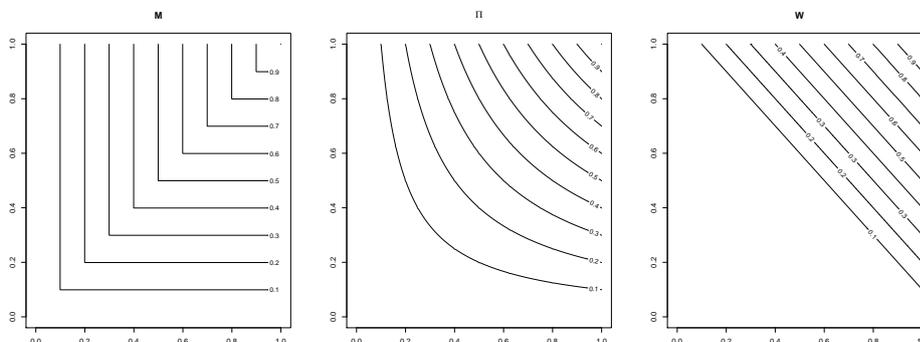


Figura 1.1: Contornos de  $\mathcal{M}$ ,  $\Pi$  y  $\mathcal{W}$ .

En la Figura 1.2 se presenta una gráfica con las secciones diagonales de  $\mathcal{M}$ ,  $\Pi$  y  $\mathcal{W}$ .

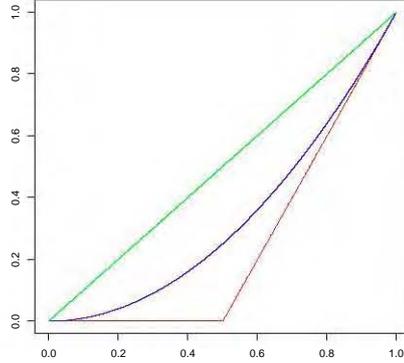


Figura 1.2: Sección diagonal de  $\mathcal{M}$  (verde),  $\Pi$  (azul) y  $\mathcal{W}$  (rojo).

Las gráficas antes expuestas son de mucha utilidad al explorar el comportamiento de la cópula de un par de variables aleatorias. Dicha aplicación se exhibirá de manera mas explícita en el capítulo cuatro, al desarrollar el modelo que explique a las variables de estudio.

**Teorema 1.2.3.** Sea  $\mathcal{C}$  una cópula.  $\forall v \in \mathcal{I}$ , la derivada parcial  $\partial\mathcal{C}(u, v)/\partial u$  existe para casi todo  $u$ , además para dichas  $u$  y  $v$ ,

$$0 \leq \frac{\partial\mathcal{C}(u, v)}{\partial u} \leq 1. \quad (1.2.4)$$

De manera análoga  $\forall u \in \mathcal{I}$ , la derivada parcial  $\partial\mathcal{C}(u, v)/\partial v$  existe para casi todo  $v$ , además para dichas  $u$  y  $v$ ,

$$0 \leq \frac{\partial\mathcal{C}(u, v)}{\partial v} \leq 1. \quad (1.2.5)$$

Además, las funciones  $u \mapsto \partial\mathcal{C}(u, v)/\partial v$  y  $v \mapsto \partial\mathcal{C}(u, v)/\partial u$  están bien definidas como no decrecientes en todo  $\mathcal{I}$ .

### 1.3. Teorema de Sklar

El teorema al que se refiere el título de esta sección es central para la teoría de cópulas. El teorema de Sklar exhibe el papel que desempeñan las cópulas en la relación entre funciones de distribución conjuntas y sus marginales univariadas. Por lo tanto iniciaremos esta sección definiendo a las funciones de distribución.

**Definición 1.3.1.** Una *función de distribución* es una función  $F$  con dominio  $\mathfrak{R}$  tal que:

1.  $F$  es no decreciente.

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

**Definición 1.3.2.** Una *función de distribución conjunta* es una función con dominio  $\bar{\mathfrak{R}}^2$  tal que:

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

Por lo tanto  $H$  esta fijada, y como el  $\text{Dom}H = \bar{\mathfrak{R}}^2$ ,  $H$  tiene marginales  $F$  y  $G$  dadas por  $F(x) = H(x, \infty)$  y  $G(y) = H(\infty, y)$ . A partir del Corolario 1.2.1,  $F$  y  $G$  son funciones de distribución.

Cabe mencionar que hasta el momento no se ha hecho alusión a ningún concepto “probabilístico”. Las funciones de distribución de una o dos variables correspondientes a variables aleatorias cumplen con las características antes mencionadas. Por lo tanto cualquier resultado que se derive de las funciones de distribución aplicara también para variables aleatorias sin ninguna consideración adicional.

**Teorema 1.3.1. Teorema de Sklar.** Sea  $H$  una función de distribución conjunta con marginales  $F$  y  $G$ . Entonces existe una cópula  $\mathcal{C}$  tal que  $\forall x, y \in \bar{\mathfrak{R}}$ ,

$$H(x, y) = \mathcal{C}(F(x), G(y)). \quad (1.3.1)$$

Si  $F$  y  $G$  son continuas, entonces  $\mathcal{C}$  es única; en otro caso  $\mathcal{C}$  esta únicamente determinada en el  $\text{Ran}F \times \text{Ran}G$ . De manera inversa, si  $\mathcal{C}$  es una cópula y  $F$  y  $G$  son funciones de distribución, entonces la función  $H$  definida por (1.3.1) es una función de distribución conjunta con marginales  $F$  y  $G$ .

**Lema 1.3.1.** Sea  $H$  una función de distribución conjunta con marginales  $F$  y  $G$ . Entonces existe una única subcópula  $\mathcal{C}'$  tal que:

1.  $\text{Dom}\mathcal{C}' = \text{Ran}F \times \text{Ran}G$ .
2.  $\forall x, y \in \bar{\mathfrak{R}}, H(x, y) = \mathcal{C}'(F(x), G(y))$ .

**Lema 1.3.2.** Sea  $\mathcal{C}'$  una subcópula. Entonces existe una cópula  $\mathcal{C}$  t.q.  $\mathcal{C}(u, v) = \mathcal{C}'(u, v) \forall (u, v) \in \text{Dom}\mathcal{C}'$ ; i.e., cualquier subcópula puede ser extendida a una cópula. Dicha extensión no siempre es única.

Las pruebas de los Lemas 1.3.1 y 1.3.2 y del Teorema de Sklar 1.3.1 se detallan en [Nelsen \(2006\)](#).

La ecuación (1.3.1) da una expresión para funciones de distribución conjuntas en términos de la cópula y dos funciones de distribución univariadas. Pero (1.3.1) puede ser invertida a manera de expresar a la cópula en términos de una función de distribución conjunta y la “inversa” de dos marginales. Pero, si una de las marginales no es estrictamente creciente, entonces esta no posee una inversa en los términos convencionales. Por lo tanto, primero se definirá el término de *cuasi-inversa* para funciones de distribución.

**Definición 1.3.3.** Sea  $F$  una función de distribución. Entonces una *cuasi-inversa* de  $F$  es una función  $F^{(-1)}$  con dominio  $\mathcal{I}$  tal que:

1. Si  $t$  está en el  $\text{Ran}F$ , entonces  $F^{(-1)}(t)$  es cualquier número  $x \in \bar{\mathfrak{R}}$  tal que  $F(x) = t$ , i.e.  $\forall t \in \text{Ran}F$ ,

$$F(F^{(-1)}(t)) = t.$$

2. Si  $t$  no está en  $\text{Ran}F$ , entonces

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

Cabe mencionar que si  $F$  es estrictamente creciente, esta tiene una única *cuasi-inversa*, la inversa ordinaria, en tal caso se usará la notación acostumbrada  $F^{-1}$ .

**Corolario 1.3.1.** Sean  $H, F, G$  y  $\mathcal{C}'$  como en el Lema 1.3.1, y sea  $F^{(-1)}$  y  $G^{(-1)}$  las *cuasi-inversas* de  $F$  y  $G$ , respectivamente. Entonces  $\forall (u, v) \in \text{Dom}\mathcal{C}'$ ,

$$\mathcal{C}'(u, v) = H(F^{(-1)}(u), G^{(-1)}(v)) \quad (1.3.2)$$

Cuando  $F$  y  $G$  son continuas este resultado aplica para cópulas.

La relevancia de los resultados expuestos en esta sección será más clara cuando se explique en las secciones siguientes la importancia que tienen las cópulas en el estudio de variables aleatorias.

Se finaliza esta sección mencionando algunas familias de cópulas, las cuales serán utilizadas en el capítulo cuatro al momento de hacer un ajuste paramétrico a las variables de estudio. También se menciona un método para “pegar” dos o más cópulas en una sola, éste será empleado en caso de que la cópula de las variables tenga que ser representada por dos o más cópulas.

### Cópulas Elípticas

- *Familia Gaussiana (Normal)*, si  $\Phi$  denota la función de distribución (univariada) de una normal estándar y  $\mathcal{N}_p$  denota la función de distribución normal bivariada estándar con coeficiente de correlación  $\hat{r}$ , entonces para  $\hat{r} \in (-1, 1) \setminus \{0\}$ :

$$\begin{aligned} \mathcal{C}(u, v) &= \mathcal{N}_p(\Phi^{-1}(u), \Phi^{-1}(v)) \\ &= \frac{1}{2\pi\sqrt{1-\hat{r}^2}} \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \exp\left[\frac{-(s^2 - 2\hat{r}st + t^2)}{2(1-\hat{r}^2)}\right] ds dt. \end{aligned} \quad (1.3.3)$$

- *Familia T*, véase [Demarta y McNeil \(2005\)](#).

### Otras Familias de Cópulas

- *Familia Plackett*, para  $\theta > 0$ ,  $\theta \neq 1$ :

$$C_\theta(u, v) = \frac{[1 + (\theta - 1)(u + v)] - \sqrt{[1 + (\theta - 1)(u + v)]^2 - 4uv\theta(\theta - 1)}}{2(\theta - 1)}, \quad (1.3.4)$$

para  $\theta = 1$ ,

$$C_1(u, v) = uv.$$

- *Familia Farlie-Gumbel-Morgenstern (FGM)*,  $\forall \theta \in [-1, 1]$ :

$$C_\theta(u, v) = uv + \theta uv(1 - u)(1 - v). \quad (1.3.5)$$

### “Gluing Copulas” (Pegando Cópulas)

El método “Gluing Copulas” fue propuesto en [Siburg y Stoimenov \(2008\)](#) y es una técnica general para construir cópulas  $n$ -variadas a partir de dos o más cópulas de la misma dimensión, escalando y juntando en una sola, abordaremos el caso de cópulas bivariadas e iniciaremos con la forma para pegar 2 cópulas:

**Proposición 1.3.1.** Dadas dos copulas bivariadas  $C_1$  y  $C_2$  y un valor fijo  $0 < \theta < 1$ , podemos escalar  $C_1$  a  $[0, \theta] \times [0, 1]$  y  $C_2$  a  $[\theta, 1] \times [0, 1]$  y juntarlas en una sola cópula con la siguiente fórmula:

$$C_{1,2,\theta}(u, v) = \begin{cases} \theta C_1\left(\frac{u}{\theta}, v\right), & 0 \leq u \leq \theta, \\ (1 - \theta)C_2\left(\frac{u - \theta}{1 - \theta}, v\right) + \theta v, & \theta \leq u \leq 1. \end{cases} \quad (1.3.6)$$

De igual forma se puede hacer si se desea pegar 3 cópulas en una, de la siguiente manera:

**Proposición 1.3.2.** Dadas tres copulas bivariadas  $C_1$ ,  $C_2$  y  $C_3$  y dos valores fijos  $0 < \theta_1$ ,  $\theta_2 < 1$ , podemos escalar  $C_1$  a  $[0, \theta_1] \times [0, 1]$ ,  $C_2$  a  $[\theta_1, \theta_2] \times [0, 1]$  y  $C_3$  a  $[\theta_2, 1] \times [0, 1]$  y juntarlas en una sola cópula con la siguiente fórmula:

$$C_{1,2,3,\theta_1,\theta_2}(u, v) = \begin{cases} \theta_1 C_1\left(\frac{u}{\theta_1}, v\right), & 0 \leq u \leq \theta_1, \\ (\theta_2 - \theta_1)C_2\left(\frac{u - \theta_1}{\theta_2 - \theta_1}, v\right) + \theta_1 v, & \theta_1 \leq u \leq \theta_2, \\ (1 - \theta_2)C_3\left(\frac{u - \theta_2}{1 - \theta_2}, v\right) + \theta_2 v, & \theta_2 \leq u \leq 1. \end{cases} \quad (1.3.7)$$

De igual forma se puede extender el concepto para cualquier número de cópulas que se deseen pegar. Los pegados de cópulas presentados en las Proposiciones 1.3.1 y 1.3.2 son horizontales, es decir se segmenta con respecto a la variable  $u$  (el primer argumento), pero, en [Siburg y Stoimenov \(2008\)](#) se exhibe también como pegar de manera vertical, con respecto a la variable  $v$ , o hacer combinaciones, pegando tanto de manera horizontal como vertical para una misma cópula. Para efectos del presente trabajo con las Proposiciones anteriores bastará para realizar el ajuste que se desea.

## 1.4. Cópulas y Variables Aleatorias

Se tomará el término “variable aleatoria” (v.a.) desde el punto de vista de la estadística como: una cantidad para la cual su valor esta determinado por una función de distribución de probabilidad (conocida o desconocida). Todos los resultados siguientes son aplicables también desde el punto de vista de la teoría de la probabilidad donde una variable aleatoria es definida con base en la teoría de la medida, *i.e.*, como una función medible en un espacio de probabilidad determinado. Ambos términos son análogos al extenderlos a vectores aleatorios. Por lo tanto, recordaremos las características de una función de distribución de probabilidad (FDP).

**Definición 1.4.1.** Sean  $X$  una v.a. y  $F$  una función t.q.  $F : \bar{\mathfrak{R}} \rightarrow [0, 1]$ , entonces  $F$  es una FDP de  $X$  si para todo  $x \in \bar{\mathfrak{R}}$

$$F(x) = \mathcal{P}(X \leq x). \quad (1.4.1)$$

**Teorema 1.4.1.** Sean  $X$  una v.a. y  $F$  una FDP de  $X$ , entonces  $F$  cumple con las siguientes propiedades:

1.  $F(\infty) = 1$ ;
2.  $F(-\infty) = 0$ ;
3. Si  $x \leq y \Rightarrow F(x) \leq F(y)$ , lo que significa que  $F$  es una función monótona creciente ó no decreciente en  $\bar{\mathfrak{R}}$ .

**Definición 1.4.2.** Sean  $X, Y$  v.a. 's y  $H$  una función t.q.  $H : \bar{\mathfrak{R}}^2 \rightarrow [0, 1]$ , entonces  $H$  es una FDP conjunta de  $X, Y$  si para todo  $(x, y) \in \bar{\mathfrak{R}}^2$

$$H(x, y) = \mathcal{P}(X \leq x, Y \leq y). \quad (1.4.2)$$

**Teorema 1.4.2.** Sean  $X, Y$  v.a. 's y  $H$  una FDP conjunta, entonces  $H$  cumple con lo siguiente:

1.  $H$  está fijada;

2.  $H$  es 2-creciente;
3.  $H$  tiene por marginales  $F$  y  $G$  las FDP de  $X$ ,  $Y$  respectivamente.

Estamos definiendo FDP de variables aleatorias como continuas por la derecha, pero esto es sólo por convención, FDP continuas por la izquierda pueden funcionar de igual manera. Cabe mencionar que una variable aleatoria es continua cuando su función de distribución es continua.

A continuación reafirmaremos el teorema de Sklar en términos de variables aleatorias y sus funciones de distribución:

**Teorema 1.4.3.** Sean  $X$ ,  $Y$  variables aleatorias con funciones de distribución  $F$  y  $G$ , respectivamente, y función de distribución conjunta  $H$ . Entonces existe una cópula  $\mathcal{C}$  tal que (1.3.1) aplica. Si  $F$  y  $G$  son continuas,  $\mathcal{C}$  es única. En otro caso  $\mathcal{C}$  es única en  $\text{Ran}F \times \text{Ran}G$ .

La cópula  $\mathcal{C}$  definida en el Teorema 1.4.3 será llamada la cópula de  $X$ ,  $Y$  o la cópula subyacente de  $X$ ,  $Y$  y simplemente será denotada por  $\mathcal{C}_{XY}$ .

**Teorema 1.4.4.** Sean  $X$ ,  $Y$  variables aleatorias continuas.  $X$ ,  $Y$  son independientes si y sólo si  $\mathcal{C}_{XY} = \Pi$ .

La demostración del Teorema anterior es inmediata por el hecho de que un par de variables aleatorias continuas  $X$ ,  $Y$  con marginales  $F$  y  $G$  respectivamente, son independientes si y sólo si  $H(x, y) = F(x)G(y)$ . Con el Teorema 1.4.3 se sigue la demostración.

Mucha de la utilidad de las cópulas en el estudio de la estadística no paramétrica deriva del hecho de que para transformaciones estrictamente monótonas de variables aleatorias, la cópula entre las mismas es invariante o cambia de manera predeterminada. Recordaremos que si la función de distribución de una variable aleatoria  $X$  es continua, y si  $\alpha$  es una función estrictamente monótona tal que su dominio contiene  $\text{Ran}X$ , entonces la función de distribución de  $\alpha(X)$  es también continua. Dichas consideraciones serán necesarias para entender los siguientes dos teoremas.

**Teorema 1.4.5.** Sean  $X$ ,  $Y$  variables aleatorias continuas, con cópula  $\mathcal{C}_{XY}$ . Si  $\alpha$  y  $\beta$  son funciones estrictamente crecientes en el  $\text{Ran}X$  y  $\text{Ran}Y$ , respectivamente, entonces  $\mathcal{C}_{\alpha(X)\beta(Y)} = \mathcal{C}_{XY}$ . Por lo tanto  $\mathcal{C}_{XY}$  es invariante bajo transformaciones estrictamente crecientes de  $X$ ,  $Y$ .

Cuando al menos una de las funciones  $\alpha$  o  $\beta$  es estrictamente decreciente, se obtienen resultados para los cuales la cópula para las variables  $\alpha(X)$ ,  $\beta(Y)$  es una simple transformación de  $\mathcal{C}_{XY}$  como se muestra en el Teorema 1.4.6.

**Teorema 1.4.6.** Sean  $X$ ,  $Y$  variables aleatorias continuas con cópula  $\mathcal{C}_{XY}$ . Sean  $\alpha$  y  $\beta$  funciones estrictamente monótonas en el  $\text{Ran}X$  y  $\text{Ran}Y$ , respectivamente:

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

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

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

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

3. Si ambas,  $\alpha$  y  $\beta$  son estrictamente decrecientes, entonces:

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

Como las cópulas son funciones de distribución (con marginales uniformes  $(0, 1)$ ), cada cópula  $\mathcal{C}$  induce una medida de probabilidad en  $\mathcal{I}^2$  vía  $V_{\mathcal{C}}([0, u] \times [0, v]) = \mathcal{C}(u, v)$ , de manera similar, la  $\mathcal{C}$ -medida de un conjunto es su  $\mathcal{C}$ -volumen  $V_{\mathcal{C}}$ . Por lo tanto, de manera intuitiva, la  $\mathcal{C}$ -medida de un subconjunto de  $\mathcal{I}^2$  es la probabilidad de que dos variables aleatorias uniformes  $(0, 1)$   $U$  y  $V$ , con función de distribución conjunta  $\mathcal{C}$  tomen valores en ese subconjunto.

El *soporte de una cópula* es el complemento de la unión de todos los subconjuntos abiertos de  $\mathcal{I}^2$  con  $\mathcal{C}$ -medida igual a cero.

**Ejemplo 1.4.1.** El soporte de la cota superior de Fréchet-Hoeffding  $\mathcal{M}$  es la diagonal principal de  $\mathcal{I}^2$ , *i.e.*, la gráfica de  $v = u$  para  $u \in I$ . Lo anterior se sigue de que la  $\mathcal{M}$ -medida de cualquier rectángulo por debajo o por encima de la diagonal principal es cero. De manera similar, el soporte de la cota inferior de Fréchet-Hoeffding  $\mathcal{W}$  es la diagonal secundaria de  $\mathcal{I}^2$ , *i.e.*, la gráfica de  $v = 1 - u$  para  $u \in I$ .

## 1.5. Las cotas de Fréchet-Hoeffding como funciones de distribución

En el Teorema 1.2.1 se encontraron las cotas de Fréchet-Hoeffding como cotas universales para cópulas, *i.e.*, para cualquier cópula  $\mathcal{C}$  y  $\forall u, v \in \mathcal{I}$ .

$$\mathcal{W}(u, v) = \max(u + v - 1, 0) \leq \mathcal{C}(u, v) \leq \min(u, v) = \mathcal{M}(u, v).$$

Como consecuencia del teorema de Sklar, si  $X, Y$  son variables aleatorias continuas con función de distribución conjunta  $H$  y marginales  $F$  y  $G$ , respectivamente, entonces  $\forall x, y \in \mathfrak{R}$ :

$$\max(F(x), G(y) - 1, 0) \leq H(x, y) \leq \min(F(x), G(y)). \quad (1.5.1)$$

Ya que  $\mathcal{M}$  y  $\mathcal{W}$  son cópulas, las cotas antes enunciadas son funciones de distribución conjuntas y son conocidas como las cotas de Fréchet-Hoeffding

para funciones de distribución conjuntas  $H$  con marginales  $F$  y  $G$ . El objetivo de esta sección es conocer qué se puede decir sobre el comportamiento conjunto de un par de variables aleatorias cuando su función de distribución conjunta es una de las cotas de Fréchet-Hoeffding.

Para este fin se introducirá la noción de conjuntos no decrecientes y no crecientes en  $\bar{\mathfrak{R}}^2$ .

**Definición 1.5.1.** Sea  $\mathcal{S} \subseteq \bar{\mathfrak{R}}^2$ ,  $\mathcal{S}$  es no decreciente si  $\forall(x, y), (u, v) \in \mathcal{S}$ ,  $x < u \Rightarrow y \leq v$ . De manera similar,  $\mathcal{S}$  es no creciente si  $\forall(x, y), (u, v) \in \mathcal{S}$ ,  $x < u \Rightarrow y \geq v$ .

**Lema 1.5.1.** Sea  $\mathcal{S} \subseteq \bar{\mathfrak{R}}^2$ ,  $\mathcal{S}$  es no decreciente si y sólo si  $\forall(x, y) \in \bar{\mathfrak{R}}^2$ :

1. 
$$\forall(u, v) \in \mathcal{S}, u \leq x \Rightarrow v \leq y; \quad (1.5.2)$$

ó

2. 
$$\forall(u, v) \in \mathcal{S}, v \leq y \Rightarrow u \leq x. \quad (1.5.3)$$

**Lema 1.5.2.** Sean  $X, Y$  variables aleatorias con función de distribución conjunta  $H$ . Entonces  $H$  es igual a la cota superior de Fréchet-Hoeffding si y sólo si  $\forall(x, y) \in \bar{\mathfrak{R}}^2$ ,  $\mathcal{P}[X > x, Y \leq y] = 0$  ó  $\mathcal{P}[X \leq x, Y > y] = 0$ .

**Teorema 1.5.1.** Sean  $X, Y$  variables aleatorias con función de distribución conjunta  $H$ . Entonces  $H$  es igual a la cota superior de Fréchet-Hoeffding si y sólo si el soporte de  $H$  es un subconjunto no decreciente de  $\bar{\mathfrak{R}}^2$ .

**Ejemplo 1.5.1.** Sean  $X, Y$  variables aleatorias, con  $Y = \alpha(X)$  y  $\alpha$  una función monótona creciente (no decreciente), entonces  $\mathcal{C}_{XY} = \mathcal{M}$ .

*Demostración.* Si  $u = F(x)$  y  $v = G(y) = F(\alpha^{-1}(y))$

$$\begin{aligned} \mathcal{C}_{XY}(u, v) &= \mathcal{C}_{X\alpha(X)}(u, v) \\ &= \mathcal{P}[X \leq x, \alpha(X) \leq y] \\ &= \mathcal{P}[X \leq x, X \leq \alpha^{-1}(y)] \\ &= \mathcal{P}[X \leq \min(x, \alpha^{-1}(y))] \\ &= \min(F(x), F(\alpha^{-1}(y))) \\ &= \min(F(x), G(y)) \\ &= \min(u, v) = \mathcal{M}(u, v). \end{aligned}$$

□

**Lema 1.5.3.** Sea  $\mathcal{S} \subseteq \bar{\mathfrak{R}}^2$ . Entonces  $\mathcal{S}$  es no creciente si y sólo si  $\forall(x, y) \in \bar{\mathfrak{R}}^2$ :

1.

$$\forall(u, v) \in \mathcal{S}, u \leq x \Rightarrow v > y; \quad (1.5.4)$$

ó

2.

$$\forall(u, v) \in \mathcal{S}, v > y \Rightarrow u \leq x. \quad (1.5.5)$$

**Lema 1.5.4.** Sean  $X, Y$  variables aleatorias con función de distribución conjunta  $H$ . Entonces  $H$  es igual a la cota inferior de Fréchet-Hoeffding si y sólo si  $\forall(x, y) \in \bar{\mathfrak{R}}^2, \mathcal{P}[X > x, Y > y] = 0$  ó  $\mathcal{P}[X \leq x, Y \leq y] = 0$ .

**Teorema 1.5.2.** Sean  $X, Y$  variables aleatorias con función de distribución conjunta  $H$ . Entonces  $H$  es igual a la cota inferior de Fréchet-Hoeffding si y sólo si el soporte de  $H$  es un subconjunto no creciente de  $\bar{\mathfrak{R}}^2$

**Ejemplo 1.5.2.** Sean  $X, Y$  variables aleatorias, con  $Y = \beta(X)$  y  $\beta$  una función monótona decreciente (no creciente), entonces  $\mathcal{C}_{XY} = \mathcal{W}$ .

*Demostración.* Si  $u = F(x)$  y  $v = G(y) = 1 - F(\beta^{-1}(y))$   
 $\Rightarrow$

$$\begin{aligned} \mathcal{C}_{XY}(u, v) &= \mathcal{C}_{X\beta(X)}(u, v) \\ &= \mathcal{P}[X \leq x, \beta(X) \leq y] \\ &= \mathcal{P}[X \leq x, X > \beta^{-1}(y)] \\ &= \mathcal{P}[\beta^{-1}(y) < X \leq x], \end{aligned}$$

si  $x \leq \beta^{-1}(y)$  entonces  $\mathcal{P}[\beta^{-1}(y) < X \leq x] = 0$ , y además:

$$\begin{aligned} F(x) &\leq F(\beta^{-1}(y)) \\ 1 - F(\beta^{-1}(y)) &\leq 1 - F(x) \\ G(y) &\leq 1 - F(x) \\ G(y) + F(x) - 1 &\leq 0, \end{aligned}$$

ahora si  $x > \beta^{-1}(y)$  entonces:

$$\begin{aligned} \mathcal{P}[\beta^{-1}(y) < X \leq x] &= F(x) - F(\beta^{-1}(y)) \\ &= F(x) - (1 - G(y)) \\ &= F(x) + G(y) - 1, \end{aligned}$$

y además de manera análoga al caso anterior  $G(y) + F(x) - 1 \geq 0$ . Por lo tanto si  $H$  es la función de distribución conjunta de  $X, Y$  entonces  $H(x, y) = \max(F(x) + G(y) - 1, 0)$  de donde se sigue que  $\mathcal{C}_{XY} = \mathcal{W}$ .  $\square$

## 1.6. Cópula de Supervivencia

En muchas aplicaciones se estudia el tiempo de vida de individuos u objetos en alguna población. La probabilidad de que un individuo sobreviva más de un tiempo  $x$  está dado por una *función de supervivencia*  $\bar{F}(x) = \mathcal{P}[X > x] = 1 - F(x)$ , donde como antes  $F$  denota la función de distribución de una variable aleatoria  $X$ . Por la naturaleza de la aplicación, se espera que el rango de la variable aleatoria sean los reales positivos, sin embargo se trabajará con la variable cuando su rango es  $\mathfrak{R}$ .

Para un par de variables aleatorias  $X, Y$  con función de distribución conjunta  $H$ , la función de supervivencia conjunta se define como  $\bar{H}(x, y) = \mathcal{P}[X > x, Y > y]$ . Las marginales de  $\bar{H}$  son las funciones  $\bar{F}(x) = \bar{H}(x, -\infty)$  y  $\bar{G}(y) = \bar{H}(-\infty, y)$  que son las funciones de supervivencia univariadas. A continuación, se expondrá la relación que existe entre la función de supervivencia conjunta y sus marginales, suponemos que  $X, Y$  son variables aleatorias con cópula  $\mathcal{C}$ . Entonces tenemos:

$$\begin{aligned}\bar{H}(x, y) &= 1 - F(x) - G(y) + H(x, y) \\ &= \bar{F}(x) + \bar{G}(y) - 1 + \mathcal{C}(F(x), G(y)) \\ &= \bar{F}(x) + \bar{G}(y) - 1 + \mathcal{C}(1 - \bar{F}(x), 1 - \bar{G}(y))\end{aligned}$$

Por lo tanto definimos una función  $\hat{\mathcal{C}}$  que va  $\mathcal{I}^2$  a  $\mathcal{I}$  por:

$$\hat{\mathcal{C}}(u, v) = u + v - 1 + \mathcal{C}(1 - u, 1 - v). \quad (1.6.1)$$

De donde:

$$\bar{H}(x, y) = \hat{\mathcal{C}}(\bar{F}(x), \bar{G}(y)). \quad (1.6.2)$$

Nos referiremos a  $\hat{\mathcal{C}}$  como la *cópula de supervivencia*, cabe destacar que ésta cumple con las propiedades de una función cópula. Se debe tener cuidado de no confundir a la cópula de supervivencia  $\hat{\mathcal{C}}$  con la función de supervivencia conjunta  $\bar{C}$  de dos variables aleatorias uniformes  $(0,1)$ , donde su función de distribución conjunta es la cópula  $\mathcal{C}$ . Ya que:

$$\bar{C}(u, v) = \mathcal{P}[U > u, V > v] = 1 - u - v + \mathcal{C}(u, v) = \hat{\mathcal{C}}(1 - u, 1 - v). \quad (1.6.3)$$

## 1.7. Simetría

La finalidad de este capítulo es presentar conceptos básicos de simetría de cópulas ya que la gran mayoría de las cópulas paramétricas son simétricas.

**Definición 1.7.1.** Un par de variables aleatorias  $X, Y$  son intercambiables si los vectores  $(X, Y)$  y  $(Y, X)$  son idénticamente distribuidos. Si la función de distribución conjunta del vector  $(X, Y)$  es  $H$ , entonces  $H(x, y) = H(y, x) \forall x, y \in \mathfrak{R}^2$ . Claramente variables aleatorias intercambiables deben ser idénticamente distribuidas *i.e.*, tienen la misma función de distribución univariada.

**Teorema 1.7.1.** Sean  $X, Y$  variables aleatorias continuas con función de distribución  $H$ , marginales  $F$  y  $G$  respectivamente y cópula  $\mathcal{C}$ . Entonces  $X, Y$  son intercambiables si y sólo si  $F = G$  y  $\mathcal{C}(u, v) = \mathcal{C}(v, u) \forall (u, v) \in \mathcal{I}^2$ .

Cuando  $\mathcal{C}(u, v) = \mathcal{C}(v, u) \forall (u, v) \in \mathcal{I}^2$ , simplemente se dice que  $\mathcal{C}$  es simétrica.

**Lema 1.7.1.** Sea  $\mathcal{C}$  una cópula, entonces:

$$\sup_{(u,v) \in \mathcal{I}^2} |\mathcal{C}(u, v) - \mathcal{C}(v, u)| \leq \frac{1}{3}; \quad (1.7.1)$$

y la desigualdad es la mejor posible.

Cuando una cópula alcanza la cota superior se le llama *máxima no-intercambiable* o *extremadamente asimétrica*.

En Durante, Klement, Sempi y Úbeda-Flores (2010), se analizan algunas medidas de no-intercambiabilidad como por ejemplo; si  $\mathcal{C}$  es una cópula, una medida de no-intercambiabilidad o asimetría es:

$$\delta(\mathcal{C}) = 3 \sup_{(u,v) \in \mathcal{I}^2} |\mathcal{C}(u, v) - \mathcal{C}(v, u)|. \quad (1.7.2)$$

La cual tiene las siguientes propiedades:

1.  $\delta(\mathcal{C}) = 0$  si y sólo si  $\mathcal{C}$  es intercambiable o simétrica.
2. Para cualquier cópula  $\mathcal{C}$ ,  $0 \leq \delta(\mathcal{C}) \leq 1$ .
3.  $\mathcal{C}$  es máxima no-intercambiable o extremadamente asimétrica si y sólo si  $\delta(\mathcal{C}) = 1$ .
4. El valor  $\delta(\mathcal{C}) = 1$  para cópulas máximas no-intercambiables o extremadamente asimétricas es alcanzado en  $(u, v) = (1/3, 2/3)$  o en  $(u, v) = (2/3, 1/3)$ .

**Definición 1.7.2.** Una cópula  $\mathcal{C}$  es *tri-simétrica* si:

$$\mathcal{C}\left(\frac{1}{3}, \frac{2}{3}\right) - \mathcal{C}\left(\frac{2}{3}, \frac{1}{3}\right) = 0. \quad (1.7.3)$$

Observamos que toda cópula simétrica es tri-simétrica, el caso contrario no precisamente se cumple. Si bien el concepto de tri-simetría es más débil que el de simetría puede ayudar al momento de hacer ajustes con cópulas paramétricas. El concepto de tri-simetría se aborda de manera detallada en Erdely y González-Barrios (2010), en donde también se propone una prueba de tri-simetría. Otra prueba, ésta si de simetría, es la que se propone en Genest, Nešlehová y Quessy (2011), la cual será utilizada en el capítulo cuatro al ajustar una cópula paramétrica a las variables de estudio.

## 1.8. Orden

La cotas de Fréchet-Hoeffding,  $\mathcal{W}(u, v) \leq \mathcal{C}(u, v) \leq \mathcal{M}(u, v)$  para toda cópula  $\mathcal{C}$  y para todo  $u, v$  en  $\mathcal{I}$ , sugiere un orden parcial en el conjunto de cópulas.

**Definición 1.8.1.** Si  $\mathcal{C}_1$  y  $\mathcal{C}_2$  son cópulas, diremos que  $\mathcal{C}_1$  es menor que  $\mathcal{C}_2$  (o  $\mathcal{C}_2$  es mayor que  $\mathcal{C}_1$ ), y se escribirá  $\mathcal{C}_1 \prec \mathcal{C}_2$  (o  $\mathcal{C}_2 \succ \mathcal{C}_1$ ) si  $\mathcal{C}_1(u, v) \leq \mathcal{C}_2(u, v) \forall u, v \in \mathcal{I}$ .

En otras palabras, la cota inferior de Fréchet-Hoeffding  $\mathcal{W}$  siempre es menor a toda cópula, de igual manera la cota superior  $\mathcal{M}$  siempre es mayor a toda cópula. Este orden parcial de cópulas es llamado el *orden de concordancia* y tiene relevancia en el siguiente capítulo en donde se discute la relación entre cópulas y propiedades de dependencia para variables aleatorias. Notemos que lo que se propone en la Definición 1.8.1 es un orden parcial ya que no todo par de cópulas es comparable.

**Ejemplo 1.8.1.** La cópula producto  $\Pi$  y la cópula que se obtiene de hacer un promedio de las cotas de Fréchet-Hoeffding (caso particular cuando  $\alpha = 0.5$  en el Ejemplo 1.2.2) no son comparables.

*Demostración.*

Si  $\mathcal{C}(u, v) = [\mathcal{W}(u, v) + \mathcal{M}(u, v)]/2$ , entonces  $\mathcal{C}(1/4, 1/4) = [1/2 + 1/4]/2 = 3/8 > 1/16 = \Pi(1/4, 1/4)$  y  $\mathcal{C}(1/4, 3/4) = [0 + 1/4]/2 = 1/8 < 3/16 = \Pi(1/4, 3/4)$ , por lo tanto no pasa que  $\Pi \prec \mathcal{C}$  y tampoco  $\Pi \succ \mathcal{C}$ .  $\square$

Por otro lado, existen familias de cópulas que están completamente ordenadas. A estas familias paramétricas  $\{\mathcal{C}_\theta\}$  se les llama *ordenadas positivamente* si  $\mathcal{C}_\alpha \prec \mathcal{C}_\beta$  cuando  $\alpha \leq \beta$ ; y *ordenadas negativamente* si  $\mathcal{C}_\alpha \succ \mathcal{C}_\beta$  cuando  $\alpha \leq \beta$ .

**Ejemplo 1.8.2.** La familia Cuadras-Augé de cópulas (para detalle sobre esta familia de véase [Nelsen \(2006\)](#)), es ordenada positivamente. Si  $\mathcal{C}_\alpha$  pertenece a la dicha familia entonces para  $\alpha \in \mathcal{I}$ :

$$\mathcal{C}_\alpha = [\mathcal{M}(u, v)]^\alpha [\Pi(u, v)]^{1-\alpha}.$$

*Demostración.* Para  $0 \leq \alpha \leq \beta \leq 1$  y  $u, v$  en  $(0, 1)$ ,

$$\frac{\mathcal{C}_\alpha(u, v)}{\mathcal{C}_\beta(u, v)} = \left( \frac{uv}{\mathcal{M}(u, v)} \right)^{\beta-\alpha} \leq 1,$$

y por tanto  $\mathcal{C}_\alpha \prec \mathcal{C}_\beta$ .  $\square$

## 1.9. Simulación Bivariada con Cópulas

Una de las aplicaciones fundamentales de las cópulas es la simulación, en el presente trabajo será de mucha utilidad comparar simulaciones conjuntas de los modelos propuestos y los datos reales. Iniciaremos con un teorema que será la base del desarrollo de esta sección.

**Teorema 1.9.1.** (De la Integral de Probabilidad Transformada)

Si  $X$  es una variable aleatoria con función de distribución continua  $F$ , entonces  $F(X) \sim \mathcal{U}(0, 1)$ . Debido a esto, para una variable  $\mathcal{U}$  uniforme  $(0, 1)$ ,  $F^{(-1)}(\mathcal{U}) \sim X$ , donde  $F^{(-1)}$  es alguna cuasi-inversa de  $F$ .

*Demostración.* Véase [Mood \(1963\)](#). □

Con base en el Teorema 1.9.1 es que se propone el método de la función inversa. Para simular una observación  $x$  de la variable aleatoria  $X$  con función de distribución  $F$ .

**Algoritmo 1.9.1.**

1. Generar un valor aleatorio  $u$ , como una observación de una variable aleatoria uniforme  $(0, 1)$ .
2. Hacer  $x = F^{(-1)}(u)$ , donde  $F^{(-1)}$  es alguna cuasi-inversa de  $F$ .

A continuación se expondrá un método para generar observaciones  $(x, y)$  de un vector aleatorio  $(X, Y)$  usando la cópula subyacente de  $\mathcal{C}_{XY}$  como herramienta. En virtud del teorema de Sklar, sólo necesitamos generar un par  $(u, v)$  de observaciones de variables aleatorias  $(\mathcal{U}, \mathcal{V})$  uniformes  $(0, 1)$ , con función de distribución conjunta  $\mathcal{C}_{XY}$ , para después transformar dichas uniformes vía el algoritmo en el párrafo anterior. Un procedimiento para generar el par  $(u, v)$  es el método de la distribución condicional. Para este método necesitamos la función de distribución condicional de  $\mathcal{V}$  dado  $\mathcal{U} = u$ , que denotaremos como  $c_u(v)$ :

$$\begin{aligned} c_u(v) = \mathcal{P}[\mathcal{V} \leq v | \mathcal{U} = u] &= \lim_{\Delta u \rightarrow 0} \frac{\mathcal{C}(u + \Delta u, v) - \mathcal{C}(u, v)}{\Delta u} \\ &= \frac{\partial \mathcal{C}(u, v)}{\partial u}. \end{aligned} \tag{1.9.1}$$

Ahora, estamos en posición de enunciar el algoritmo:

**Algoritmo 1.9.2.**

1. Generar 2 observaciones independientes  $u$  y  $t$  de variables aleatorias uniformes  $(0,1)$ .
2. Hacer  $v = c_u^{(-1)}(t)$ , donde  $c_u^{(-1)}$  denota una cuasi-inversa de  $c_u$ .
3. El par deseado es  $(u, v)$ , aplicar la transformación del Algoritmo 1.9.1, a  $u$  y a  $v$  con respectivas cuasi-inversas de las funciones de distribución de las variables de interés.

Más adelante, en este mismo capítulo se presenta una versión suavizada para funciones de distribución, la cual parte de la función de distribución empírica, la inversa de esta función puede ser sustituida de manera única por la cuasi-inversa en los algoritmos expuestos en esta sección. Esto en el caso de que no se tenga información alguna de la distribución de las variables y sólo se cuente con una muestra de las mismas.

Se concluye esta sección con un último algoritmo para simular cópulas que provengan de un “pegado de cópulas” como en (1.3.6) o (1.3.7), éste será relevante en caso de que la cópula de las variables de estudio tenga que ser construida bajo este método. El algoritmo se presenta cuando la cópula es producto de pegar 2 cópulas, este puede ser generalizado de manera análoga para cuando se tienen más cópulas en el pegado.

Sea  $\mathcal{C}_{1,2,\theta}$  que satisface (1.3.6), entonces:

$$\begin{aligned} \frac{\partial \mathcal{C}_{1,2,\theta}(u, v)}{\partial u} &= \begin{cases} \theta \frac{\partial \mathcal{C}_1(\frac{u}{\theta}, v)}{\partial u} \frac{d(\frac{u}{\theta})}{du}, & 0 \leq u \leq \theta, \\ (1 - \theta) \frac{\partial \mathcal{C}_2(\frac{u-\theta}{1-\theta}, v)}{\partial u} \frac{d(\frac{u-\theta}{1-\theta})}{du}, & \theta \leq u \leq 1. \end{cases} \\ &= \begin{cases} \frac{\partial \mathcal{C}_1(\frac{u}{\theta}, v)}{\partial u}, & 0 \leq u \leq \theta, \\ \frac{\partial \mathcal{C}_2(\frac{u-\theta}{1-\theta}, v)}{\partial u}, & \theta \leq u \leq 1. \end{cases} \end{aligned}$$

Por lo tanto si se desea simular una muestra de tamaño  $N$  de la cópula  $\mathcal{C}_{1,2,\theta}$  se propone el siguiente algoritmo:

**Algoritmo 1.9.3.**

1. Calcular  $N_1 = \lfloor N * \theta \rfloor$  y  $N_2 = N - N_1$ , donde, si  $a$  es un número real,  $\lfloor a \rfloor$  es el máximo entero de  $a$ .

2. Simular de manera independiente una muestra  $\{(u_{1,k}, v_{1,k})\}_{k=1}^{N_1}$  de la cópula  $\mathcal{C}_1$  y otra  $\{(u_{2,k}, v_{2,k})\}_{k=1}^{N_2}$  de la cópula  $\mathcal{C}_2$ .
3. Aplicar las transformaciones: (1)  $\hat{u}_{1,k} = (u_{1,k} * \theta) \forall k = 1, 2, \dots, N_1$ ;  
(2)  $\hat{u}_{2,k} = (u_{2,k}(1 - \theta) + \theta) \forall k = 1, 2, \dots, N_2$ .  
La muestra  $\Gamma := \{(\hat{u}_{1,k}, v_{1,k})\}_{k=1}^{N_1} \cup \{(\hat{u}_{2,k}, v_{2,k})\}_{k=1}^{N_2}$  es, en efecto, una simulación de tamaño  $N$  de la cópula  $\mathcal{C}_{1,2,\theta}$ .
4. Si la cópula proviene de una par variables aleatorias y se desea simular a las mismas, aplicar la transformación del Algoritmo 1.9.1, a todo  $\hat{u}$  y  $v$  en  $\Gamma$  con respectivas cuasi-inversas de las funciones de distribución de las variables de interés.

## 1.10. Cópulas Arquimedianas

Esta sección tiene por finalidad mostrar una clase de cópulas conocida como las cópulas arquimedianas. Estas cópulas han tenido muchas aplicaciones por varias razones, entre ellas son las siguientes: (1) la facilidad con la que estas se construyen, (2) la gran variedad de cópulas que pertenecen a esta clase y (3) la cantidad de propiedades que poseen. Primero daremos algunas definiciones que nos ayudarán a caracterizar a esta clase de cópulas.

**Definición 1.10.1.** Sea  $\phi$  una función continua, estrictamente decreciente que va de  $\mathcal{I}$  a  $[0, \infty]$  tal que  $\phi(1) = 0$ . La *pseudo-inversa* de  $\phi$  es la función  $\phi^{[-1]}$  con  $\text{Dom}\phi^{[-1]} = [0, \infty]$  y  $\text{Ran}\phi^{[-1]} = \mathcal{I}$  dada por:

$$\phi^{[-1]} = \begin{cases} \phi^{-1}(t), & 0 \leq t \leq \phi(0), \\ 0, & \phi(0) \leq t \leq \infty. \end{cases} \quad (1.10.1)$$

donde  $\phi^{-1}$  es la inversa ordinaria de la función.

Observamos que si  $\phi(0) \rightarrow \infty$ , entonces  $\phi^{[-1]} \rightarrow \phi^{-1}$ .

**Lema 1.10.1.** Sea  $\phi$  una función continua, estrictamente decreciente que va de  $\mathcal{I}$  a  $[0, \infty]$  tal que  $\phi(1) = 0$ , y sea  $\phi^{[-1]}$  la pseudo-inversa de  $\phi$ . Sea  $\mathcal{C}$  es la función que va de  $\mathcal{I}^2$  a  $\mathcal{I}$  dada por:

$$\mathcal{C}(u, v) = \phi^{[-1]}(\phi(u) + \phi(v)). \quad (1.10.2)$$

Entonces  $\mathcal{C}$  satisface las condiciones de frontera del punto 1 y 2 de la Definición 1.2.2 para una cópula.

**Teorema 1.10.1.** Sea  $\phi$  una función continua, estrictamente decreciente que va de  $\mathcal{I}$  a  $[0, \infty]$  tal que  $\phi(1) = 0$ , y sea  $\phi^{[-1]}$  la pseudo-inversa de  $\phi$ . Entonces la función  $\mathcal{C}$  que va de  $\mathcal{I}^2$  a  $\mathcal{I}$  dada por (1.10.2) es una cópula si y sólo si  $\phi$  es convexa.

Para detalles de las demostraciones del lema y teorema anteriores, véase [Nelsen \(2006\)](#).

Las cópulas como en (1.10.2) son llamadas arquimedianas, es por eso que los resultados anteriores sirven para entender la estructura de esta clase de cópulas. La función  $\phi$  es llamado un *generador* de la cópula. Si  $\phi(0) = \infty$ , diremos que  $\phi$  es un *generador estricto*. En tal caso,  $\phi^{[-1]} = \phi^{-1}$  y  $\mathcal{C}(u, v) = \phi^{-1}(\phi(u) + \phi(v))$ , diremos que  $\mathcal{C}$  es *estrictamente una cópula arquimediana*.

El siguiente teorema presenta algunas propiedades de las cópulas arquimedianas que serán de utilidad en capítulos siguientes.

**Teorema 1.10.2.** Sea  $\mathcal{C}$  una cópula arquimediana con generador  $\phi$ . Entonces:

1.  $\mathcal{C}$  es simétrica; i.e.,  $\mathcal{C}(u, v) = \mathcal{C}(v, u)$  para todo  $u, v$  en  $\mathcal{I}$ ;
2.  $\mathcal{C}$  es asociativa, i.e.,  $\mathcal{C}(\mathcal{C}(u, v), w) = \mathcal{C}(u, \mathcal{C}(v, w))$  para todo  $u, v, w$  en  $\mathcal{I}$ ;
3. si  $c > 0$  cualquier constante, entonces  $c\phi$  es también un generador de  $\mathcal{C}$ .

**Ejemplo 1.10.1.** (Ejemplos de cópulas arquimedianas)

- (a) Sea  $\phi(t) = -\ln(t)$  para  $t \in [0, 1]$ . Como  $\phi(0) = \infty$ ,  $\phi$  es un generador estricto. Entonces  $\phi^{[-1]} = \phi^{-1} = \exp(-t)$ , y generando  $\mathcal{C}$  vía (1.10.2),  $\mathcal{C}(u, v) = \exp(-[(-\ln(u)) + (-\ln(v))]) = uv = \Pi(u, v)$  Por lo tanto  $\Pi$  es estrictamente una cópula Arquimediana.
- (b) Sea  $\phi(t) = 1 - t$  para  $t \in [0, 1]$ . Entonces  $\phi^{[-1]}(t) = 1 - t$  para  $t \in [0, 1]$  y 0 para  $t > 1$ ; i.e.,  $\phi^{[-1]}(t) = \max(1 - t, 0)$ . Nuevamente usando (1.10.2),  $\mathcal{C}(u, v) = \max(1 - [(1 - u) + (1 - v)], 0) = \max(u + v - 1, 0) = \mathcal{W}(u, v)$ . Entonces  $\mathcal{W}$  es también Arquimediana.
- (c)  $\mathcal{M}$  no es Arquimediana. Véase [Nelsen \(2006\)](#).

A continuación se presentan algunas familias de cópulas arquimedianas, éstas serán utilizadas en el capítulo cuatro como candidatas para el ajuste a las variables de estudio:

- *Familia Clayton (1978)*, para  $\theta \in [-1, \infty) \setminus \{0\}$ :

$$\mathcal{C}_\theta(u, v) = \left[ \max(u^{-\theta} + v^{-\theta} - 1, 0) \right]^{-1/\theta}, \quad (1.10.3)$$

donde  $\phi_\theta(t) = \frac{1}{\theta}(t^{-\theta} - 1)$  es la función generadora de  $\mathcal{C}_\theta$ . Además, cabe destacar a ciertos miembros de esta familia:

$$\mathcal{C}_{-1} = \mathcal{W}, \mathcal{C}_0 = \Pi, \mathcal{C}_1 = \frac{\Pi}{u+v-\Pi} \text{ y } \mathcal{C}_\infty = \mathcal{M}.$$

- *Familia Ali-Mikhail-Haq (AMH)*, para  $\theta \in [-1, 1)$ :

$$\mathcal{C}_\theta(u, v) = \frac{uv}{1 - \theta(1-u)(1-v)}, \quad (1.10.4)$$

donde  $\phi_\theta(t) = \ln\left(\frac{1-\theta(1-t)}{t}\right)$  es la función generadora de  $\mathcal{C}_\theta$ . Además, cabe destacar a ciertos miembros de esta familia:

$$\mathcal{C}_0 = \Pi \text{ y } \mathcal{C}_1 = \frac{\Pi}{u+v-\Pi}.$$

- *Familia Gumbel*, para  $\theta \in [1, \infty)$ :

$$\mathcal{C}_\theta(u, v) = \exp\left(-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{1/\theta}\right) \quad (1.10.5)$$

donde  $\phi_\theta(t) = (-\ln t)^\theta$  es la función generadora de  $\mathcal{C}_\theta$ . Además, cabe destacar a ciertos miembros de esta familia:

$$\mathcal{C}_1 = \Pi \text{ y } \mathcal{C}_\infty = \mathcal{M}.$$

- *Familia Frank*, para  $\theta \in (-\infty, \infty) \setminus \{0\}$ :

$$\mathcal{C}_\theta(u, v) = -\frac{1}{\theta} \ln\left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}\right), \quad (1.10.6)$$

donde  $\phi_\theta(t) = -\ln\frac{e^{-\theta t}-1}{e^{-\theta}-1}$  es la función generadora de  $\mathcal{C}_\theta$ . Además, cabe destacar a ciertos miembros de esta familia:

$$\mathcal{C}_{-\infty} = \mathcal{W}, \mathcal{C}_0 = \Pi \text{ y } \mathcal{C}_\infty = \mathcal{M}.$$

Cabe mencionar que a las familias que contienen a las cópulas  $\mathcal{M}$ ,  $\Pi$  y  $\mathcal{W}$  se les llama *completas*. En ese sentido, las familias *Clayton* (1.10.3) y *Frank* (1.10.6) cumplen con ser completas.

## 1.11. Aproximaciones de la Cópula

En esta sección se revisarán: la expresión empírica de una cópula bivariada, un suavizamiento de la cópula con base en la cópula empírica y un acercamiento a las observaciones correspondientes a la cópula. Estos conceptos son muy importantes como un primer acercamiento al comportamiento de la cópula subyacente de un par de variables aleatorias y por ende para explicar el comportamiento conjunto de las variables. También, como se verá en el siguiente capítulo, la cópula empírica sirve para obtener algunas

versiones muestrales de medidas de asociación.

Iniciaremos con algunos conceptos ya conocidos para una variable. Sea  $\{(x_k, y_k)\}_{k=1}^n$  que denota una muestra de tamaño  $n$  de observaciones independientes e idénticamente distribuidas del vector aleatorio  $(X, Y)$ . Podemos obtener estimadores empíricos de las funciones de distribución univariada de la siguiente manera:

$$F_n(x) = \frac{1}{n} \sum_{k=1}^n \mathbb{I}\{x_{(k)} \leq x\}, \quad G_n(y) = \frac{1}{n} \sum_{k=1}^n \mathbb{I}\{y_{(k)} \leq y\}, \quad (1.11.1)$$

donde  $\mathbb{I}$  es la función indicadora que toma el valor de 1 si el argumento es cierto y 0 si es falso y además  $x_{(k)}$  y  $y_{(k)}$  son los estadísticos de orden de  $X, Y$  respectivamente. Es bien conocido que la distribución empírica  $F_n$  es un estimador consistente de  $F$ , lo que significa que  $F_n(t)$  converge casi seguramente a  $F(t)$  cuando  $n \rightarrow \infty$ , para toda  $t$ , como se muestra en [Billingsley \(1995\)](#). Para simulaciones de variables aleatorias continuas el uso de  $F_n$  no es apropiado, ya que ésta no es una función continua, en ese sentido se necesita suavizar la función de cuantiles  $Q(u) = F_n^{(-1)}(u)$ , donde  $F_n^{(-1)}$  es una cuasi-inversa de  $F_n$ , como en la Definición 1.3.3. Lo anterior es posible con un suavizamiento vía polinomios de Bernstein-Kantorovic como en [Muñoz~Perez y Fernández~Palacín \(1987\)](#), para toda  $u \in \mathcal{I}$ :

$$\tilde{Q}_n(u) = \sum_{k=0}^n \frac{x_{(k)} + x_{k+1}}{2} \binom{n}{k} u^k (1-u)^{n-k}, \quad (1.11.2)$$

donde  $x_{(k)}$  son los estadísticos de orden de  $X$ . Para que la expresión tenga sentido se considera a  $x_{(0)} = x_{(1)}$  y a  $x_{(n+1)} = x_{(n)}$ .

## Cópula Empírica

De manera similar a los conceptos anteriores:

**Definición 1.11.1.** La *cópula empírica* es la función  $\mathcal{C}_n$  dada por:

$$\mathcal{C}_n \left( \frac{i}{n}, \frac{j}{n} \right) = \frac{\# \text{ de pares } (x, y) \text{ en la muestra con } x \leq x_{(i)}, y \leq y_{(j)}}{n},$$

donde  $x_{(i)}$  y  $y_{(j)}$ ,  $1 \leq i, j \leq n$ , denotan los estadísticos de orden de la muestra.

Ésta converge a la cópula verdadera como se prueba en [Ruschendorf \(1976\)](#). La cópula empírica no es una cópula ya que se define en un conjunto finito y no en todo el cuadrado unitario  $[0, 1]^2$ , pero por el Teorema de Sklar  $c_n$  puede ser extendida a una cópula.

### Suavizamiento vía polinomios de Bernstein

Por otra parte una extensión suavizada de la cópula es posible con la cópula Bernstein como en [Sancetta y Satchell \(2004\)](#):

$$\tilde{\mathcal{C}}(u, v) = \sum_{i=0}^n \sum_{j=0}^n \mathcal{C}_n \left( \frac{i}{n}, \frac{j}{n} \right) \binom{n}{i} u^i (1-u)^{n-i} \binom{n}{j} v^j (1-v)^{n-j}, \quad (1.11.3)$$

para todo  $(u, v)$  en el cuadrado unitario  $[0, 1]^2$ , y donde  $\mathcal{C}_n$  es la cópula empírica como en la Definición 1.11.1.

### Observaciones de la cópula

Otra manera muy sencilla de obtener un acercamiento a la cópula subyacente de una par de variables aleatorias se da en las siguientes definiciones.

**Definición 1.11.2.** Sea  $\{(x_k, y_k)\}_{k=1}^n$  una muestra de tamaño  $n$  del vector aleatorio  $(X, Y)$  y sean  $F$  y  $G$  conocidas, las funciones de distribución de las variables  $X, Y$ . Se definen como observaciones de la cópula a los elementos de la muestra  $\{(u_k, v_k)\}_{k=1}^n$  del vector  $(\mathcal{U}, \mathcal{V})$ , donde:

$$(u_k, v_k) = (F(x_k), G(y_k)). \quad (1.11.4)$$

**Definición 1.11.3.** Sea  $\{(x_k, y_k)\}_{k=1}^n$  una muestra de tamaño  $n$  del vector aleatorio  $(X, Y)$  y sean  $F_n$  y  $G_n$  las funciones de distribución empíricas de las variables  $X, Y$ . Se definen como pseudo-observaciones de la cópula a los elementos de la muestra  $\{(\hat{u}_k, \hat{v}_k)\}_{k=1}^n$  del vector  $(\hat{\mathcal{U}}, \hat{\mathcal{V}})$ , donde:

$$(\hat{u}_k, \hat{v}_k) = (F_n(x_k), G_n(y_k)). \quad (1.11.5)$$

## 1.12. Regresión con Cópulas

La regresión es un método que pretende describir a una variable en términos de otra (u otras).

Una propuesta de curva de regresión es la que se da en [Conway \(1986\)](#), la cual especifica valores en la mediana de  $Y$  por cada valor dado de  $X$ , como se describe a continuación.

**Definición 1.12.1.** Sean  $X, Y$  variables aleatorias. Para  $x$  en  $\text{Ran}X$ , sea  $y = \tilde{y}(x)$  que denota una solución a la ecuación  $\mathcal{P}[Y \leq y | X = x] = 1/2$ . Entonces la gráfica de  $y = \tilde{y}(x)$  es la *curva de regresión mediana* de  $Y$  en  $X$ .

De manera más general se extiende el concepto como sigue.

Para un valor  $x$  en  $\text{Ran}X$  y  $0 < \alpha < 1$ , sea  $y = \Phi_\alpha(x)$  que denota una solución a la ecuación  $\mathcal{P}[Y \leq y|X = x] = \alpha$ . La gráfica  $y = \Phi_\alpha(x)$  es la *curva de regresión  $\alpha$ -cuantil* de  $Y$  condicional a que  $X = x$ . Por (1.9.1) tenemos que:

$$\mathcal{P}[Y \leq y|X = x] = c_u(v) \Big|_{u=F(x), v=G(y)}, \quad (1.12.1)$$

y con este resultado se sigue el algoritmo a continuación para la *curva de regresión  $\alpha$ -cuantil* (incluyendo la regresión mediana).

**Algoritmo 1.12.1.**

1. Hacer  $c_u(v) = \alpha$ , donde  $v$  es desconocida.
2. Resolver la curva de regresión  $v = g_\alpha(u)$ .
3. Se remplazan  $u$  por  $\tilde{Q}^{-1}(x)$  y  $v$  por  $\tilde{R}^{-1}(y)$ , donde  $\tilde{Q}$  y  $\tilde{R}$  son las funciones de distribución suavizadas para  $X$ ,  $Y$  respectivamente como en (1.11.2).
4. Resolver la curva de regresión  $y = \Psi_\alpha(x)$ .

Cabe mencionar que en el algoritmo anterior se puede sustituir la inversa de la función en (1.11.2) por alguna cuasi-inversa, en caso de que se tenga algún conocimiento previo sobre la distribución de las variables de estudio y se desee utilizar.

## Capítulo 2

# Dependencia

La dependencia entre variables aleatorias es un concepto fundamental en la teoría de la probabilidad y en la aplicación de la estadística.

Este capítulo se centra en dos objetivos: (1) hacer una breve crítica de como se ha medido la “dependencia” desde mucho tiempo atrás y como se sigue haciendo actualmente, con base en el artículo de [Erdely \(2009\)](#), (2) explicar cómo las cópulas pueden ser usadas para estudiar la dependencia o asociación de variables aleatorias.

Para dar inicio y motivación a este capítulo se hace una cita al prefacio en [Mari y Kotz \(2001\)](#):

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

dedicado a conceptos de dependencia apareció bajo la autoría de [Joe \(1997\)](#). Más aún, pareciera ser que no hay departamento de matemáticas o estadística en Estados Unidos o Europa que ofrezca cursos especialmente dedicados a estudiar conceptos y medidas de dependencia.

## 2.1. Correlación Lineal de Pearson

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

$$r(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\mathbb{V}(X)\mathbb{V}(Y)}}, \quad (2.1.1)$$

suponiendo la existencia de segundos momentos no nulos y finitos.

Con base en el Teorema 1.4.5 y en la definición de la correlación lineal es que se da el siguiente comentario en [Embrechts, Lindskog y McNeil \(2003\)](#).

Las cópulas proveen una manera natural de estudiar y medir dependencia entre variables aleatorias. Como consecuencia directa del Teorema 1.4.5, las propiedades de las cópulas son invariantes bajo transformaciones estrictamente crecientes de las variables aleatorias involucradas. La correlación lineal (o de Pearson) es frecuentemente utilizada en la práctica como medida de dependencia. Sin embargo, como la correlación lineal no es una medida basada en la cópula subyacente, en ocasiones conduce a resultados aberrantes y por tanto no debiera tomarse como la medida canónica de dependencia [...] La popularidad de la correlación lineal se explica por la facilidad con la que puede ser calculada y por ser una medida escalar de dependencia que de forma natural surge en el contexto de las distribuciones elípticas (con conocidos miembros como lo son las distribuciones normal y  $t$  multivariadas). El problema es que a muchas variables aleatorias no les corresponde una distribución elíptica, y utilizar en tal caso la correlación lineal usualmente conduce a conclusiones verdaderamente alejadas de la realidad. Y aún cuando resulte razonable utilizar como distribución conjunta a un miembro de la familia elíptica, hay situaciones en las que utilizar la correlación lineal [...] no tiene sentido; por ejemplo, en el caso de distribuciones de colas pesadas: en tales casos ni siquiera está definida la correlación lineal ante la presencia de segundos momentos infinitos.

Más aún, encontramos en [Embrechts, McNeil y Straumann \(1999\)](#) una lista de observaciones a tomar en cuenta al utilizar la correlación lineal.

La correlación lineal es un campo minado para el desprevenido. No hace falta esforzarse mucho para encontrar en la literatura sobre administración de riesgos financieros enorme confusión y malinterpretación. Esto es preocupante ya que la correlación lineal es un concepto técnico fundamental en Finanzas [...]:

1. La correlación lineal es simplemente una medida escalar de dependencia; no puede decirnos todo lo que quisiéramos saber en relación a la estructura de dependencia de los riesgos.
2. Los valores posibles de la correlación lineal dependen de las distribuciones marginales de los riesgos. Todos los valores entre  $-1$  y  $+1$  no son necesariamente alcanzables.
3. Riesgos con una dependencia positiva perfecta no tienen necesariamente una correlación lineal igual a  $+1$ ; riesgos con una dependencia negativa perfecta no tienen necesariamente una correlación lineal igual a  $-1$ .
4. Correlación lineal igual a cero no implica independencia de los riesgos.
5. La correlación lineal no es invariante bajo transformaciones crecientes de los riesgos. Por ejemplo,  $\log(X)$  y  $\log(Y)$  por lo general no tienen la misma correlación lineal que  $X$ ,  $Y$ .
6. La correlación sólo está definida cuando las varianzas de los riesgos son finitas, así que no es una medida de dependencia adecuada para riesgos de colas pesadas en donde las varianzas son infinitas.

Al expresar la correlación de Pearson con base en la cópula subyacente de las variables aleatorias se llega a lo siguiente:

$$r(X, Y) = \frac{1}{\sqrt{\mathbb{V}(X)\mathbb{V}(Y)}} \int \int_{\mathcal{I}^2} [\mathcal{C}(u, v) - uv] dF^{-1}(u) dG^{-1}(v), \quad (2.1.2)$$

donde  $F$  y  $G$  son las distribuciones marginales de  $X$ ,  $Y$  respectivamente.

De la ecuación (2.1.2) observamos que la correlación lineal no es sólo función de la cópula subyacente sino también del comportamiento marginal de las variables involucradas, entonces, con una misma estructura de relación conjunta (o dependencia), es posible tener distintos valores de correlación lineal tan sólo modificando una o ambas marginales. Más adelante, en la sección 2.4 se verá más a fondo la importancia de esta situación.

## 2.2. Concordancia

De manera informal se dice que un par de variables son concordantes si: el crecimiento de una variable se asocia con el crecimiento de la otra, o el decrecimiento de una variables se asocia con el decrecimiento de la otra. De manera más precisa, sean  $(x_i, y_i)$  y  $(x_j, y_j)$  dos observaciones del vector aleatorio  $(X, Y)$  de variables aleatorias continuas. Diremos que  $(x_i, y_i)$  y  $(x_j, y_j)$  son *concordantes* si ocurren las desigualdades  $x_i < x_j, y_i < y_j$ ; o si  $x_i > x_j, y_i > y_j$ . De manera similar  $(x_i, y_i)$  y  $(x_j, y_j)$  son *discordantes* si se dan las desigualdades  $x_i < x_j, y_i > y_j$ ; o si  $x_i > x_j, y_i < y_j$ . Como formulaciones alternativas para estas definiciones se dan las siguientes:  $(x_i, y_i)$  y  $(x_j, y_j)$  son concordantes si  $(x_i - x_j)(y_i - y_j) > 0$  y discordantes si  $(x_i - x_j)(y_i - y_j) < 0$ .

**Teorema 2.2.1.** Sean  $(X_1, Y_1)$  y  $(X_2, Y_2)$  vectores independientes de variables aleatorias continuas con distribuciones conjuntas  $H_1$  y  $H_2$  respectivamente, con marginales en común  $F$ (de  $X_1$  y  $X_2$ ) y  $G$ (de  $Y_1$  y  $Y_2$ ). Sean  $\mathcal{C}_1$  y  $\mathcal{C}_2$  las cópulas de  $(X_1, Y_1)$  y  $(X_2, Y_2)$  respectivamente, por lo tanto  $H_1(x, y) = \mathcal{C}_1(F(x), G(y))$  y  $H_2(x, y) = \mathcal{C}_2(F(x), G(y))$ . Sea  $\mathcal{Q}$  que denota la diferencia entre las probabilidades de concordancia y discordancia de  $(X_1, Y_1)$  y  $(X_2, Y_2)$ , i.e.

$$\mathcal{Q} = \mathcal{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] - \mathcal{P}[(X_1 - X_2)(Y_1 - Y_2) < 0]. \quad (2.2.1)$$

Entonces

$$\mathcal{Q} = \mathcal{Q}(\mathcal{C}_1, \mathcal{C}_2) = 4 \int \int_{\mathcal{I}^2} \mathcal{C}_2(u, v) d\mathcal{C}_1(u, v) - 1. \quad (2.2.2)$$

Como la función de concordancia del Teorema 2.2.1 juega un papel importante en esta sección, a continuación se enuncian algunas propiedades de la misma en el siguiente corolario.

**Corolario 2.2.1.** Sean  $\mathcal{C}_1, \mathcal{C}_2$  y  $\mathcal{Q}$  como se presento en el Teorema 2.2.1. Entonces:

1.  $\mathcal{Q}$  es simétrica en sus argumentos:  $\mathcal{Q}(\mathcal{C}_1, \mathcal{C}_2) = \mathcal{Q}(\mathcal{C}_2, \mathcal{C}_1)$ .
2.  $\mathcal{Q}$  es no decreciente en cada uno de sus argumentos: si  $\mathcal{C}_1 \prec \mathcal{C}'_1$  y  $\mathcal{C}_2 \prec \mathcal{C}'_2 \forall (u, v) \in \mathcal{I}^2$ , entonces  $\mathcal{Q}(\mathcal{C}_1, \mathcal{C}_2) \leq \mathcal{Q}(\mathcal{C}'_1, \mathcal{C}'_2)$ .
3. Las cópulas pueden ser remplazadas por cópulas de supervivencia en  $\mathcal{Q}$ , i.e.  $\mathcal{Q}(\mathcal{C}_1, \mathcal{C}_2) = \mathcal{Q}(\hat{\mathcal{C}}_1, \hat{\mathcal{C}}_2)$ .

**Ejemplo 2.2.1.** La función  $\mathcal{Q}$  es fácil de evaluar para pares de cópulas básicas  $\mathcal{M}, \mathcal{W}$  y  $\Pi$ .

- (a) Primero, recordar que el soporte de  $\mathcal{M}$  es la diagonal  $v = u$  en  $\mathcal{I}^2$  (Ejemplo 1.4.1). Como  $\mathcal{M}$  tiene marginales uniformes  $(0,1)$ , se sigue que si  $g$  es una función integrable con dominio  $\mathcal{I}^2$ , entonces:

$$\int \int_{\mathcal{I}^2} g(u, v) d\mathcal{M}(u, v) = \int_0^1 g(u, u) du.$$

Entonces tenemos:

$$\mathcal{Q}(\mathcal{M}, \Pi) = 4 \int \int_{\mathcal{I}^2} uv d\mathcal{M}(u, v) - 1 = 4 \int_0^1 u^2 du - 1 = 4\left(\frac{1}{3}\right) - 1 = \frac{1}{3}.$$

- (b) De manera similar, como el soporte de  $\mathcal{W}$  es la diagonal secundaria  $v = 1 - u$  (nuevamente por Ejemplo 1.4.1), tenemos:

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

Entonces:

$$\mathcal{Q}(\mathcal{W}, \Pi) = 4 \int \int_{\mathcal{I}^2} uv d\mathcal{W}(u, v) - 1 = 4 \int_0^1 u(1 - u) du - 1 = \frac{4}{6} - 1 = -\frac{1}{3}.$$

- (c) Finalmente, como  $d\Pi(u, v) = \frac{\partial^2(w)}{\partial u \partial v} du dv = du dv$ ,

$$\mathcal{Q}(\Pi, \Pi) = 4 \int \int_{\mathcal{I}^2} uv d\Pi(u, v) - 1 = 4 \int \int_{\mathcal{I}^2} uv du dv - 1 = 4\left(\frac{1}{4}\right) - 1 = 0.$$

Como  $\mathcal{Q}$  es la diferencia de dos probabilidades, para una cópula arbitraria  $\mathcal{C}$ ,  $\mathcal{Q}(\mathcal{C}, \mathcal{C}) \in [-1, 1]$ ; como consecuencia de la parte dos del Corolario 2.2.1 y con base en el Ejemplo 2.2.1 se sigue que:

$$\mathcal{Q}(\mathcal{C}, \Pi) \in [-1/3, 1/3].$$

este resultado será muy relevante en la siguiente sección.

A continuación se presentan las condiciones deseables de una medida de concordancia.

**Definición 2.2.1.** Una medida numérica  $\kappa$  de asociación entre dos variables aleatorias continuas  $X, Y$  con cópula  $\mathcal{C}$  es una medida de *concordancia* si satisface las siguientes propiedades (se puede denotar como  $\kappa_{X,Y}$  o  $\kappa_{\mathcal{C}}$  según convenga):

1.  $\kappa$  esta definida para todo par de variables aleatorias continuas  $X, Y$ ;

2.  $-1 \leq \kappa \leq 1$ ,  $\kappa_{X,X} = 1$ , y  $\kappa_{X,-X} = -1$ ;
3.  $\kappa_{X,Y} = \kappa_{Y,X}$ ;
4. Si  $X, Y$  son independientes, entonces  $\kappa_{X,Y} = \kappa_{\mathcal{C}} = 0$ ;
5.  $\kappa_{-X,Y} = \kappa_{X,-Y} = -\kappa_{X,Y}$ ;
6. Si  $\mathcal{C}_1$  y  $\mathcal{C}_2$  son cópulas tales que  $\mathcal{C}_1 \prec \mathcal{C}_2$ , entonces  $\kappa_{\mathcal{C}_1} \leq \kappa_{\mathcal{C}_2}$ ;
7. Si  $\{(X_n, Y_n)\}$  es una sucesión de variables aleatorias continuas con cópulas  $\mathcal{C}_n$ , respectivamente, y si  $\mathcal{C}_n$  converge puntualmente a  $\mathcal{C}$ , entonces  $\lim_{n \rightarrow \infty} \kappa_{\mathcal{C}_n} = \kappa_{\mathcal{C}}$ .

Como consecuencia de la Definición 2.2.1, tenemos el siguiente teorema

**Teorema 2.2.2.** Sea  $\kappa$  una medida de concordancia para variables aleatorias continuas  $X, Y$ :

1. Si  $Y = \alpha(X)$  con  $\alpha$  una función creciente casi seguramente de  $X$ , entonces  $\kappa_{X,Y} = \kappa_{\mathcal{M}} = 1$ ;
2. Si  $Y = \beta(X)$  con  $\beta$  una función decreciente casi seguramente de  $X$ , entonces  $\kappa_{X,Y} = \kappa_{\mathcal{W}} = -1$ ;
3. Si  $\alpha$  y  $\beta$  son casi seguramente funciones estrictamente monótonas, ambas crecientes o ambas decrecientes en el  $\text{Ran}X$  y  $\text{Ran}Y$ , respectivamente, entonces  $\kappa_{\alpha(X),\beta(Y)} = \kappa_{X,Y}$ .

Existen varias medidas de concordancia que cumplen con las propiedades antes mencionadas, dentro de las más relevantes por su mayor uso en la aplicación estadística se encuentran la  $\rho$  de Spearman (1904) y la  $\tau$  de Kendall (1938). Por tener una interpretación más natural en términos de la cópula subyacente, el presente trabajo sólo estudiara a la medida  $\rho$  de Spearman como medida de concordancia y su relación con medidas de dependencia.

### 2.3. Rho de Spearman

Sean  $(X_1, Y_1)$ ,  $(X_2, Y_2)$  y  $(X_3, Y_3)$  tres vectores aleatorios independientes con función de distribución conjunta en común  $H$ , marginales  $F$  y  $G$  y cópula  $\mathcal{C}$ . La versión muestral  $\rho_{X,Y}$  de Rho de Spearman se define como la probabilidad de concordancia menos la de discordancia de los dos vectores  $(X_1, Y_1)$  y  $(X_2, Y_3)$ , *i.e.*, un par de vectores con mismas marginales, pero uno de ellos tiene función de distribución  $H$ , mientras los componentes del otro son independientes:

$$\rho_{X,Y} = 3(\mathcal{P}[(X_1 - X_2)(Y_1 - Y_3) > 0] - \mathcal{P}[(X_1 - X_2)(Y_1 - Y_3) < 0]). \quad (2.3.1)$$

(de igual forma se puede hacer uso del vector  $(X_3, Y_2)$ ). Como ya se mencionó el vector  $(X_1, Y_1)$  tiene por función de distribución conjunta a  $H$  mientras que el vector  $(X_2, Y_3)$  tiene por función de distribución conjunta al producto de sus marginales (por que  $X_2, Y_3$  son independientes). Por lo tanto la cópula entre  $X_2, Y_3$  es  $\Pi$ , usando el Teorema 2.2.1 y la parte 1 Corolario 2.2.1, tenemos:

**Teorema 2.3.1.** Sea  $X, Y$  variables aleatorias continuas con cópula  $\mathcal{C}$ . Entonces la versión muestral de Rho de Spearman de  $X, Y$  (la cual se denotará con  $\rho_{X,Y}$  o  $\rho_{\mathcal{C}}$ ) está dada por:

$$\begin{aligned} \rho_{X,Y} = \rho_{\mathcal{C}} &= 3\mathcal{Q}(\mathcal{C}, \Pi) \\ &= 12 \int \int_{\mathcal{I}^2} \mathcal{C}(u, v) d\Pi(u, v) - 3 \\ &= 12 \int \int_{\mathcal{I}^2} \mathcal{C}(u, v) dudv - 3 \\ &= 12 \int \int_{\mathcal{I}^2} [\mathcal{C}(u, v) - uv] dudv. \end{aligned} \quad (2.3.2)$$

**Teorema 2.3.2.** Si  $X, Y$  son variables aleatorias continuas con cópula  $\mathcal{C}$ , entonces la versión muestral de *Rho de Spearman* satisface las propiedades en la Definición 2.2.1 de una *medida de concordancia*.

El hecho de que  $\rho$  cumpla con la propiedad seis en la Definición 2.2.1 es una de las razones para que “ $\prec$ ” sea llamado el “orden de concordancia”.

El siguiente teorema presenta el estadístico muestral de  $\rho$  de Spearman con base en la cópula empírica, éste será empleado para representar a  $\rho$  de Spearman en el capítulo cuatro.

**Teorema 2.3.3.** Sea  $\mathcal{C}_n$  la cópula empírica (como en la Definición 1.11.1), de una muestra  $\{(x_k, y_k)\}_{k=1}^n$ . Si  $\hat{\rho}$  denota el estadístico muestral de  $\rho$  de Spearman, entonces:

$$\hat{\rho} = \frac{12}{n^2 - 1} \sum_{i=1}^n \sum_{j=1}^n \left[ \mathcal{C}_n\left(\frac{i}{n}, \frac{j}{n}\right) - \frac{i}{n} \cdot \frac{j}{n} \right]. \quad (2.3.3)$$

*Demostración.* Véase [Kruskal \(1958\)](#), [Lehmann \(1975\)](#) y [Nelsen \(2006\)](#).  $\square$

## 2.4. Medidas de Dependencia

En el debate que se ha tenido respecto a cómo cuantificar de manera adecuada la dependencia entre variables aleatorias se han llegado a los siguientes acuerdos:

El primero tiene que ver con una propuesta en [Mari y Kotz \(2001\)](#) : una medida de dependencia debiera ser independiente de las distribuciones marginales, es decir, sólo basada en la cópula subyacente. De hecho, esta propuesta coincide con lo ya citado en la sección 2.1, proveniente de [Embrechts y otros \(1999\)](#), [Embrechts y otros \(2003\)](#). Esta propuesta surge como consecuencia del Teorema de Sklar ya que en el caso de variables aleatorias continuas, la estructura de dependencia queda determinada de manera única por la cópula subyacente. De aceptar esta propuesta entonces la correlación lineal (o de Pearson) quedaría excluida como medida de dependencia por dos razones: primero, depende de la existencia de segundos momentos finitos de las distribuciones marginales; segundo, porque aún bajo la existencia de segundos momentos finitos, como consecuencia de (2.1.2), la correlación lineal no sólo depende de la estructura de dependencia que guarden las variables sino también del comportamiento marginal de las mismas, como ya se había comentado en la sección 2.1.

La segunda propuesta de propiedad deseable de una medida de dependencia para variables aleatorias continuas tiene que ver con pedir que dicha medida reporte el valor cero si y sólo si las variables involucradas son independientes, tal propuesta aparece en [Schweizer y Wolff \(1981\)](#) y [Rényi \(1959\)](#). Esta condición no se cumple en el caso de la correlación lineal y de las medidas de concordancia, mismas que, a pesar de que pueden calcularse solamente en términos de la cópula subyacente, no caracterizan la independencia, como se observa en el punto cuatro de la Definición 2.2.1.

Con base en el Teorema 1.4.4, [Schweizer y Sklar \(1983\)](#) sugieren medir de algún modo que tan lejos/cerca esta la cópula subyacente de la cópula  $\Pi$  que representa la independencia:

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

Y fue así como finalmente [Schweizer y Wolff \(1981\)](#) proponen en concreto la siguiente definición de medida de dependencia.

**Definición 2.4.1.** Sean  $X, Y$  cualesquiera variables aleatorias continuas con cópula subyacente  $\mathcal{C}$ . Una medida numérica de asociación para este par de variables aleatorias es *medida de dependencia*, denotada por  $\mu_{X,Y}$  o  $\mu_{\mathcal{C}}$ ; si satisface las siguiente propiedades:

1.  $\mu_{X,Y}$  está definida para cualquier par de variables aleatorias continuas;
2.  $0 \leq \mu_{X,Y} \leq 1$ ;
3.  $\mu_{X,Y} = \mu_{Y,X}$ ;
4.  $X, Y$  son independientes si y sólo si  $\mu_{X,Y} = \mu_{\Pi} = 0$ ;
5.  $\mu_{X,Y} = 1$  si y sólo si  $X, Y$  son cada una casi seguramente función estrictamente monótona de la otra; es decir, si  $\mathcal{C}_{X,Y}$  es igual a  $\mathcal{M}$  o  $\mathcal{W}$ ;
6. Si  $\alpha$  y  $\beta$  son casi seguramente funciones estrictamente monótonas en  $\text{Ran}X$  y  $\text{Ran}Y$ , respectivamente, entonces  $\mu_{\alpha(X),\beta(Y)} = \mu_{X,Y}$ ;
7. Si  $\{(X_n, Y_n)\}$  es una sucesión de variables aleatorias continuas con cópula  $\mathcal{C}_n$ , respectivamente, y  $\{\mathcal{C}_n\}$  converge puntualmente a  $\mathcal{C}$ , entonces  $\lim_{n \rightarrow \infty} \mu_{\mathcal{C}_n} = \mu_{\mathcal{C}}$ .

También en [Schweizer y Wolff \(1981\)](#) comentan:

Ya que  $X, Y$  son independientes si y sólo si  $\mathcal{C}(u, v) = uv$ , esta observación sugiere que cualquier medida de distancia, adecuadamente estandarizada, entre las superficies  $z = \mathcal{C}(u, v)$  y  $z = uv$ , por ejemplo, cualquier distancia  $L_p$ , debiera proporcionar una medida simétrica y no paramétrica de dependencia.

De donde proviene la siguiente definición.

**Definición 2.4.2.** Para cualquier  $p \in [1, \infty)$  la distancia  $L_p$  entre una cópula  $\mathcal{C}$  y  $\Pi$  esta dada por:

$$\mathcal{V}_p(\mathcal{C}) = \left( k_p \int \int_{\mathcal{I}^2} |\mathcal{C}(u, v) - uv|^p du dv \right)^{\frac{1}{p}}, \quad (2.4.1)$$

donde  $k_p$  es una constante elegida de tal modo que (2.4.1) sea igual a 1 cuando  $\mathcal{C} = \mathcal{M}$  o  $\mathcal{W}$ .

De [Nelsen \(2006\)](#), tenemos la siguiente expresión para  $k_p$

**Teorema 2.4.1.** Para cada  $p \in [1, \infty)$  la distancia  $L_p$  calculada mediante  $\mathcal{V}_p(\mathcal{C})$  es una medida de dependencia, en donde:

$$k_p = \frac{\Gamma(2p + 3)}{2[\Gamma(p + 1)]^2}. \quad (2.4.2)$$

Casos particulares estudiados son:

1.  $\mathcal{V}_1(\mathcal{C}) = \sigma(X, Y)$ , la medida de dependencia propuesta en [Schweizer y Wolff \(1981\)](#), donde

$$\mathcal{V}_1(C) = \sigma(X, Y) = \sigma(\mathcal{C}) = 12 \int \int_{\mathcal{I}^2} |\mathcal{C}(u, v) - uv| du dv. \quad (2.4.3)$$

2.  $(\mathcal{V}_2(C))^2 = \phi^2(X, Y)$ , mejor conocido como el índice de dependencia de Hoeffding (1940), donde

$$\mathcal{V}_2(C) = \phi(X, Y) = \phi(\mathcal{C}) = \left( 90 \int \int_{\mathcal{I}^2} |\mathcal{C}(u, v) - uv|^2 du dv \right)^{\frac{1}{2}} \quad (2.4.4)$$

Para concluir esta sección se presentan las versiones muestrales de estas dos medidas de dependencia con base en la cópula empírica, ya que serán empleadas de esa manera en el capítulo cuatro.

**Proposición 2.4.1.** Sea  $\mathcal{C}_n$  la cópula empírica (como en la Definición 1.11.1), de una muestra  $\{(x_k, y_k)\}_{k=1}^n$ . Se proponen  $\hat{\sigma}$  y  $\hat{\phi}$  para denotar a los estadísticos muestrales de  $\sigma$  de Schweizer & Wolff y  $\phi$  en el índice de dependencia de Hoeffding, de la siguiente manera:

$$\hat{\sigma} = \frac{12}{n^2 - 1} \sum_{i=1}^n \sum_{j=1}^n \left| \mathcal{C}_n\left(\frac{i}{n}, \frac{j}{n}\right) - \frac{i}{n} \cdot \frac{j}{n} \right|, \quad (2.4.5)$$

y

$$\hat{\phi} = \left[ \frac{90}{n^2 - 1} \sum_{i=1}^n \sum_{j=1}^n \left| \mathcal{C}_n\left(\frac{i}{n}, \frac{j}{n}\right) - \frac{i}{n} \cdot \frac{j}{n} \right|^2 \right]^{\frac{1}{2}}. \quad (2.4.6)$$

## 2.5. Dependencias Estrictamente Monótonas

En esta última sección se exponen las ventajas de que un par de variables aleatorias tengan una *dependencia estrictamente monótona*. Comenzaremos definiendo el concepto:

**Definición 2.5.1.** Sean  $X, Y$  variables aleatorias continuas con cópula subyacente  $\mathcal{C}_{XY}$ , entonces  $X, Y$  guardan una dependencia estrictamente monótona si  $\mathcal{C}_{XY}(u, v) - \Pi(u, v) > 0$  ó  $\mathcal{C}_{XY}(u, v) - \Pi(u, v) < 0$ ,  $\forall u, v \in \mathcal{I}$ . Notemos que los extremos de la dependencia estrictamente monótona son las cópulas  $\mathcal{M}$  y  $\mathcal{W}$ .

**Proposición 2.5.1.** Sean  $X, Y$  variables aleatorias con dependencia estrictamente monótona, dadas las medidas de asociación  $\rho_{XY}$  de Spearman (2.3.2) y  $\sigma_{XY}$  de Schweizer y Wolff (2.4.3), entonces  $|\rho_{XY}| = \sigma_{XY}$ .

Esta última proposición se da ya que, estas dos medidas son muy parecidas en su versión analítica, con la salvedad del valor absoluto para  $\sigma$ . Entonces, si las diferencias entre una cópula dada y la cópula producto siempre son positivas, se sigue que  $\rho = \sigma$  y en el caso de que las diferencias entre una cópula dada y la cópula producto siempre sean negativas, entonces  $\rho = -\sigma$ .

La importancia de que un par de variables aleatorias guarden una dependencia estrictamente monótona se da por el hecho de que la gran mayoría de las cópulas conocidas cumplen con esta condición. Entonces, al buscar un ajuste paramétrico conjunto de un par de variables aleatorias, tendremos que hallar dependencias estrictamente monótonas en la muestra que puedan llegar a ser modeladas por cópulas conocidas.

Para hacer más evidente este hecho, se concluye esta sección y el presente capítulo con gráficas comparativas del cálculo de  $\rho$  de Spearman y  $\sigma$  de Schweizer & Wolff para distintas cópulas, cambiando sus parámetros a manera de que estos tomen los valores que expliquen mejor las estructuras de dependencia que las cópulas asociadas representan. Estas gráficas nos ayudarán para dos cosas: primero, para exhibir el hecho de que las cópulas paramétricas presentadas cumplen con representar dependencias estrictamente monótonas; y segundo, para familiarizarse con el comportamiento que tienen algunas cópulas al cambio de sus parámetros.

Los intervalos de los parámetros de las cópulas fueron fraccionados para hacer la representación gráfica, se decidió no graficar con líneas rectas para hacer más evidente la interacción entre el cambio de los parámetros y los valores que toman las medidas de asociación.

Cabe mencionar que los cálculos de  $\rho$  de Spearman y  $\sigma$  de Schweizer & Wolff se obtuvieron con la ayuda del paquete “copula” de R<sup>1</sup> (véase Apéndice A.2, [R Development Core Team \(2012\)](#), [Yan \(2007\)](#), [Kojadinovic y Yan \(2010\)](#) y [Hofert y Mächler \(2011\)](#)).

Las gráficas siguientes muestran de lado derecho el cálculo de  $\rho$  y de lado izquierdo el de  $\sigma$ .

La Figura 2.1 muestra los valores que toman las medidas para la cópula Plackett (1.3.4). Esta cópula es completa, pero sólo se muestra con el parámetro  $\theta \in (0, 20]$ ,  $\rho$  de Spearman en  $[-1, 0.773]$ , haciendo así más visible su comportamiento ya que las medidas llegan a valores cercanos a uno para valores de  $\theta > 500$ .

---

<sup>1</sup>Software libre para procesos estadísticos y gráficas.

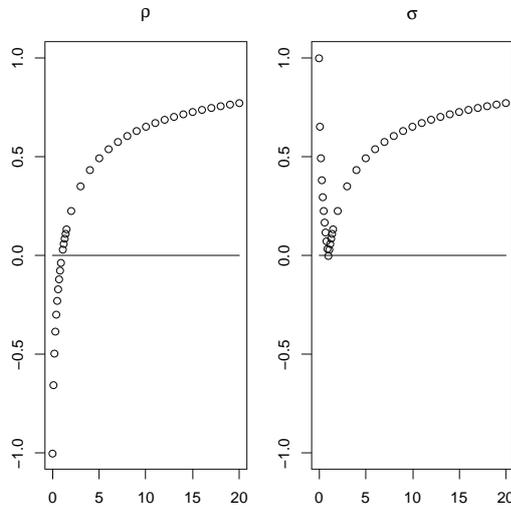


Figura 2.1: Dependencia en cópula Plackett, con  $\theta \in (0, 20]$ .

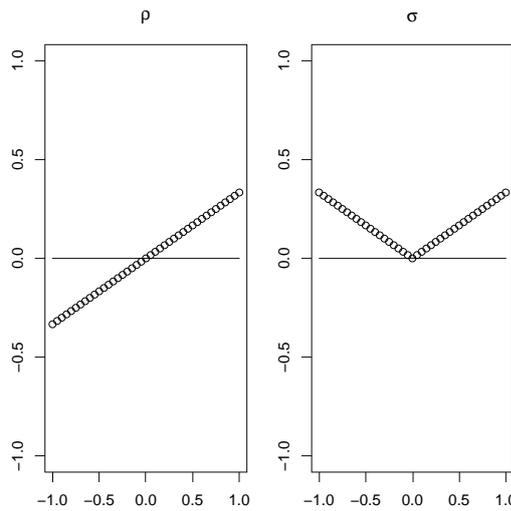


Figura 2.2: Dependencia en cópula FGM, con  $\theta \in [-1, 1]$ .

Las medidas para la cópula Farlie-Gumbel-Morgenstern (1.3.5), evaluada en todos los posibles valores de su parámetro, se muestran en la Figura 2.2, donde se observa que  $|\rho| \leq 1/3$ .

Las dos siguientes son cópulas arquimedianas presentadas en la sección 1.10.

La Figura 2.3 muestra a las medidas para la cópula Frank (1.10.6) con  $\theta \in [-40, 40]$ , es claro que ésta es completa, representando primordialmente

niveles altos de dependencia.

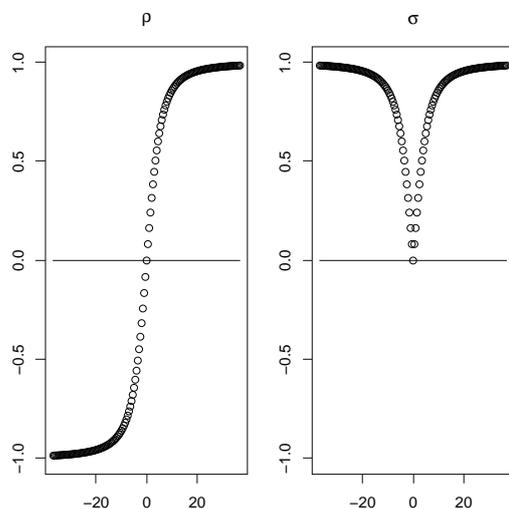


Figura 2.3: Dependencia en cópula Frank, con  $\theta \in [-37, 37]$ .

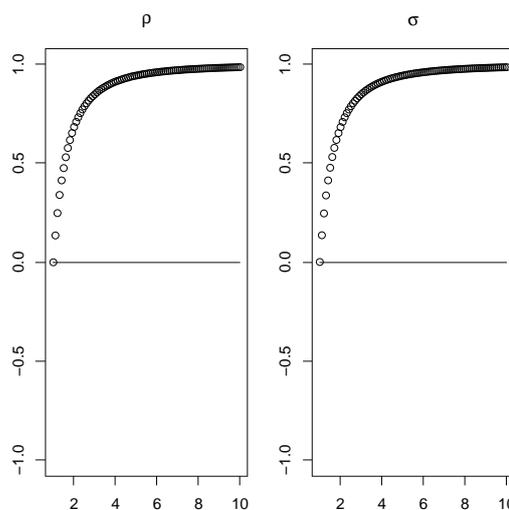


Figura 2.4: Dependencia en cópula Gumbel, con  $\theta \in [1, 10]$ .

En la Figura 2.4 se muestran las medidas para la cópula Gumbel (1.10.5), esta familia se encuentra entre la cópula  $\Pi$  y la cópula  $\mathcal{M}$ .

La Figura 2.5 muestra el cálculo de  $\rho$  y  $\sigma$  para la cópula Gaussiana (1.3.3), ésta es completa al evaluarla en todos los posibles valores para el coeficiente de correlación de Pearson (2.1.1).

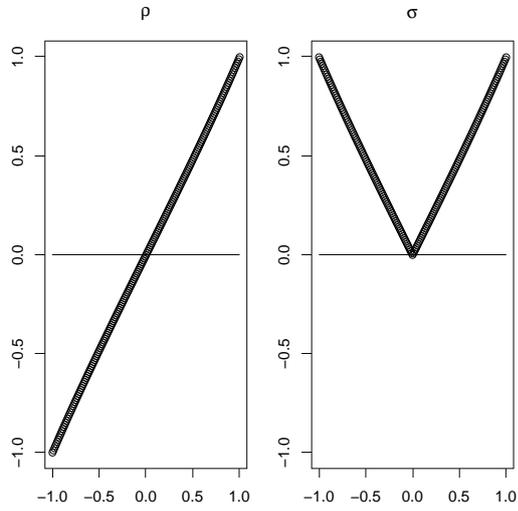


Figura 2.5: Dependencia en cópula Gaussiana (Normal), con  $r \in [-1, 1]$ .

## Capítulo 3

# Intensidad Migratoria y Desarrollo Humano

“Todos somos migrantes”: es una idea que muchos filósofos y pensadores tienen. La vida de migrante es un deseo de muchas personas en el mundo, las razones pueden ser diversas: algunas buscan definirse a sí mismas; otras, conocer lugares nuevos; hay quienes sólo buscan aventura o cualquier otro tipo de experiencia; pero un gran número de ellas, va en busca de una mejor calidad de vida con respecto a la que tienen en su tierra natal, es gente que simplemente huye al quedarse sin posibilidades mejores en su lugar de origen.

Algo que difícilmente se llega a dimensionar, es el impacto que tiene el fenómeno de la migración en la vida de aquellos que ven partir o reciben a los migrantes. En México, hoy en día, los vínculos familiares, afectivos, económicos o simplemente simbólicos que los desplazamientos geográficos de las personas han ido formando (ya sea entre regiones internas o con otros países) constituyen un elemento central en la vida del país. Es por eso que su estudio se ha vuelto imprescindible.

Podemos entender al desarrollo humano como la expansión de la libertad de las personas, donde la libertad es el conjunto de oportunidades para ser y actuar y la posibilidad de elegir con autonomía. En ese sentido, surge la siguiente pregunta: ¿Migrar es una decisión en libertad o es la única opción disponible? Las respuestas son tan variadas como complejos son los problemas de desarrollo del país, donde otro fenómeno aparece: la desigualdad, problema latente de nuestra nación y al cual se hace referencia a lo largo de este trabajo. Es muy intuitivo pensar que el desarrollo humano está íntimamente ligado con la migración, pero ¿de que manera se da esta relación? Esta pregunta es el eje de esta investigación, cuya respuesta es el propósito del capítulo cuatro.

### 3.1. Motivación

A continuación se enumeran una serie de observaciones que hace el *Programa de las Naciones Unidas para el Desarrollo* y que se enuncian en [PNUD \(2006\)](#), las cuales sirven de motivación al estudio del desarrollo humano y la migración:

1. El desarrollo humano es la expansión de la libertad de las personas. La libertad es el conjunto de oportunidades para ser y actuar y la posibilidad de elegir con autonomía. Desde la perspectiva del desarrollo humano, los movimientos migratorios voluntarios son indicativos del grado de libertad del que gozan las personas.
2. El desarrollo humano comienza por dar un lugar a las personas para que no sean sujetos dominados por el azar, la necesidad o la voluntad de otros. Pero la libertad no es sólo un asunto de autonomía individual. También consiste en tener oportunidades accesibles de las cuales escoger. Por esta razón, el desarrollo humano es el potencial que tienen las personas para ser o hacer; es la posibilidad de vivir como se desea.
3. Quien decide migrar ejerce su capacidad de elección, aunque en ocasiones lo hace sobre la base de opciones limitadas y frecuentemente desiguales. La variable más robusta para explicar los flujos migratorios a lo largo de la historia es la existencia de brechas de desarrollo entre dos regiones o países.
4. Los movimientos migratorios cambian la geografía del desarrollo humano. No sólo puede ocurrir el desplazamiento de una población con determinadas características sociodemográficas, como el analfabetismo, sino también la transformación del conjunto de oportunidades de aquellos que ven partir a los migrantes, como de aquellos que los reciben.
5. En la migración está presente un fenómeno de desigualdad: la distribución regional de las posibilidades de desarrollo implica la presencia de zonas que ofrecen mejores condiciones de vida que las que se tienen en el lugar de origen.
6. A su vez, la redistribución geográfica de las personas está acompañada de una modificación en el conjunto de oportunidades accesibles a los individuos. Tanto las zonas emisoras

como las receptoras transforman su potencial económico y social con la migración.

Para analizar la relación que guardan el desarrollo humano y la migración en México, se estudiará el comportamiento conjunto de las variables: Índice de Desarrollo Humano (IDH)<sup>1</sup> e Índice de Intensidad Migratoria México-Estados Unidos (IIM)<sup>2</sup>. Se cuenta con información del año 2000 de 2442 municipios, sólo uno menos de en los que se dividía el país en ese momento, ya que para Nicolás Ruíz (Chiapas) no se disponía de información para ninguna de las 2 variables. En las siguientes dos secciones se describe cada una de las variables de estudio.

### 3.2. Índice de Desarrollo Humano (IDH)

El desarrollo de un país no puede ser entendido desde la perspectiva única del crecimiento económico. El propósito final del desarrollo se encuentra en cada uno de sus habitantes y en las posibilidades que ellos tienen para elegir una vida en la que puedan realizar a plenitud su potencial como seres humanos.

El desarrollo humano consiste en la libertad que gozan los individuos para elegir entre distintas opciones y formas de vida. Los factores fundamentales que permiten a las personas ser libres en ese sentido, son la posibilidad de alcanzar una vida larga y saludable, poder adquirir conocimientos individual y socialmente valiosos, y tener la oportunidad de obtener los recursos necesarios para disfrutar un nivel de vida decoroso.

En el núcleo del concepto de desarrollo humano se encuentran las personas y sus oportunidades, no la riqueza que poseen, el ingreso que devengan, o las mercancías y servicios que consumen.

Es así como el PNUD<sup>3</sup> define al desarrollo humano. Bajo estas ideas es que el IDH mide los adelantos medios de un país en tres aspectos básicos del desarrollo humano:

- Una vida larga y saludable, medida por la esperanza de vida al nacer (*índice de salud*).

---

<sup>1</sup>Fuente: <http://www.undp.org.mx/>

<sup>2</sup>Fuente: estimaciones de CONAPO con base en la muestra del diez por ciento del XII Censo General de Población y Vivienda 2000 (véase <http://www.conapo.gob.mx/>).

<sup>3</sup>Programa de la Naciones Unidas para el Desarrollo.

- Conocimientos, medidos por la tasa de alfabetización de adultos y la tasa bruta de matriculación combinada en educación primaria, secundaria y terciaria (*índice de educación*).
- Un nivel de vida decoroso, medido por el PIB<sup>4</sup> per cápita<sup>5</sup> (en dólares PPC<sup>6</sup>, *índice de ingreso*).

En el cálculo del IDH es necesario crear un índice para cada uno de estos componentes: esperanza de vida, educación, y PIB per cápita. Para el caso del índice de educación se requiere, en primer término, crear un índice de alfabetización de adultos y un índice de matriculación.

Para la construcción de los índices antes mencionados, se seleccionan valores mínimos y máximos (valores de referencia), con los que se compara el logro en cada componente del país o estado en cuestión. Los valores máximos y mínimos son establecidos por el PNUD en el ámbito internacional, éstos se muestran a continuación en el Cuadro 3.1.

Indicador	Valor máximo	Valor mínimo
Esperanza de vida al nacer (años)	85	25
Tasa de alfabetización de adultos (%)	100	0
Tasa bruta de matriculación combinada (%)	100	0
PIB per cápita (dólares PPC)	40000	100

Cuadro 3.1: Valores máximos y mínimos establecidos por el PNUD.

El desempeño en cada componente se expresa como un valor entre 0 y 1, aplicando la siguiente fórmula para los índices de salud, alfabetización de adultos y matriculación:

$$\text{índice del componente} = \frac{\text{valor efectivo} - \text{valor mínimo}}{\text{valor máximo} - \text{valor mínimo}}$$

<sup>4</sup>Producto Interno Bruto: Es el valor total de los bienes y servicios producidos en el territorio de un país en un periodo determinado, libre de duplicaciones. Se puede obtener mediante la diferencia entre el valor bruto de producción y los bienes y servicios consumidos durante el propio proceso productivo, a precios comprador (consumo intermedio). Esta variable se puede obtener también en términos netos al deducirle al PIB el valor agregado y el consumo de capital fijo de los bienes de capital utilizados en la producción (para mayor información consulte <http://www.banxico.org.mx/>).

<sup>5</sup>Locución latina de uso actual que significa literalmente por cada cabeza. El PIB per capita, entonces, indica el PIB dividido por la población al 1 de julio de ese mismo año.

<sup>6</sup>Paridad de Poder de Compra: Es el ajuste económico para comparar de una manera realista el nivel de vida entre distintos países, atendiendo al producto interno bruto per cápita en términos del coste de vida en cada país.

En el cálculo del índice de ingreso se usa el logaritmo del PIB per cápita.

$$\text{índice del ingreso} = \frac{\log(\text{valor efectivo}) - \log(\text{valor mínimo})}{\log(\text{valor máximo}) - \log(\text{valor mínimo})}$$

Para calcular el índice de educación se pondera con dos tercios el logro en la tasa de alfabetización y con un tercio el logro en la tasa bruta de matriculación combinada.

$$\begin{aligned} \text{índice de educación} &= \frac{2}{3}(\text{índice de alfabetización de adultos}) \\ &+ \frac{1}{3}(\text{índice de matriculación}). \end{aligned}$$

Después de obtener el índice de cada componente, se calcula el IDH como un simple promedio entre dichos índices:

$$IDH = \frac{1}{3}\text{índice de salud} + \frac{1}{3}\text{índice de educación} + \frac{1}{3}\text{índice de ingreso}.$$

Existen 4 niveles de desarrollo humano con base en el IDH, establecidos internacionalmente por el PNUD. A continuación, en el Cuadro 3.2 se presentan los valores máximos y mínimos para cada nivel, junto con las frecuencias relativas encontradas para el año 2000 en dichos niveles, de un total de 2442 municipios.

Nivel de Desarrollo Humano	Número de Municipios	Frecuencia Relativa	Mínimo ( $\geq$ )	Máximo ( $<$ )
Bajo	31	1.27 %	0	.5
Medio Bajo	625	25.59 %	.5	.65
Medio Alto	1584	64.86 %	.65	.8
Alto	202	8.27 %	.8	1 (=)

Cuadro 3.2: Niveles de Desarrollo Humano.

El cuadro anterior reitera el desequilibrio a nivel de desarrollo humano que existía en el país, donde el mayor nivel se presentó en una proporción reducida de la población, la mayoría de ésta se encontró en una escala media de desarrollo, pero existía un grupo importante de más de la cuarta parte de la población con rezago en su desarrollo; lo anterior da una idea de los diversos problemas en materia de desarrollo humano a los que se enfrentaba el país.

Para estudiar los índices que componen al IDH se propone la siguiente notación:

- ISI: índice de salud,
- INE: índice de educación,
- IPIB: Índice de ingreso.

La Figura 3.1 muestra una comparativa entre los comportamientos de las variables ISI, INE, IPIB e IDH a nivel de sus respectivos cuartiles y rangos intercuartiles, se hace uso de gráficos boxplot para dichos fines, los cuales también nos ayudarán a detectar valores atípicos en los índices. Los umbrales inferior y superior de estos gráficos (conocidos comúnmente como “whiskers”) para la detección de valores atípicos se emplean bajo la configuración habitual<sup>7</sup> (para mayor detalle de las características de un gráfico boxplot véase Apéndice A.4 y Murrell (2006), todas las gráficas que se presentan en el resto del presente capítulo fueron obtenidas con R).

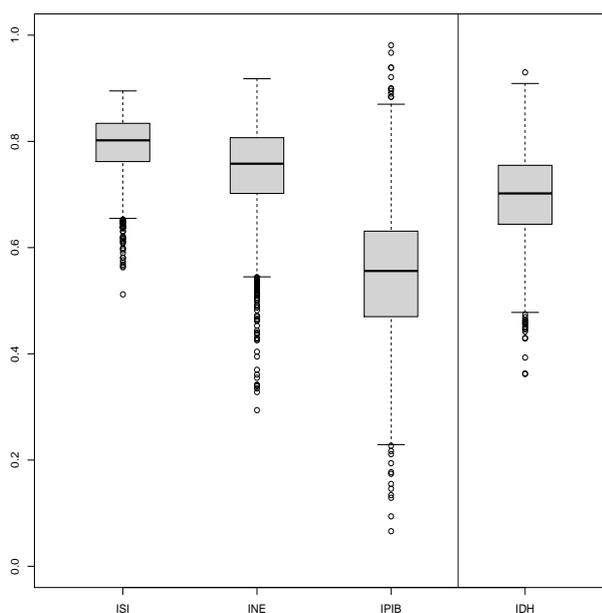


Figura 3.1: ISI, INE, IPIB e IDH.

A partir de la Figura 3.1 y calculando la curtosis y sesgo de las variables, se encuentra que todas éstas sesgadas hacia la izquierda y sus distribuciones

<sup>7</sup>Se considera como la configuración habitual el utilizar los umbrales a 1.5 por el rango intercuartil más el tercer cuartil (umbral superior) y menos el primer cuartil (umbral inferior). Esta consideración puede ser modificada a conveniencia del análisis.

son leptocurticas<sup>8</sup>.

Valor	ISI	INE	IPIB
<b>Valor mínimo</b>	0.512	0.293	0.066
<b>Representante</b>	Metlatónoc (Guerrero)	Coicoyán de las Flores (Oaxaca)	Santos Reyes Yucuná (Oaxaca)
<b>Posición de IDH</b>	2441	2442	2439
<b>Umbral inferior</b>	0.655	0.545	0.229
<b>Atípicos inferiores</b>	38	79	12
<b>Primer cuartil</b>	0.762	0.702	0.470
<b>Mediana</b>	0.802	0.758	0.559
<b>Tercer cuartil</b>	0.834	0.807	0.631
<b>Umbral superior</b>	0.895	0.918	0.870
<b>Atípicos superiores</b>	0	0	11
<b>Valor máximo</b>	0.895	0.918	0.980
<b>Representante</b>	Coacalco de Berriozábal (México)	San Sebastián Tutla (Oaxaca)	Benito Juárez (D.F.)
<b>Posición de IDH</b>	73	20	1

Cuadro 3.3: Cuartiles y umbrales (ISI, INE e IPIB).

El Cuadro 3.3 exhibe los cuartiles y umbrales (inferior y superior) junto con el número de valores atípicos que los superan, muestra los valores máximos y mínimos de las variables, presentando al municipio representante de estos valores y su posición en el IDH<sup>9</sup>.

Con base en la Figura 3.1 y el Cuadro 3.3 se analizará particularmente cada variable que compone al IDH:

1. Si explicamos a la variable ISI en términos de la esperanza de vida, entonces, para el año 2000 la esperanza de vida a nivel municipal iba desde los 55.7 y hasta los 78.7 años, siendo la cota inferior muy alarmante. El primer cuartil en ISI tiene asociada una esperanza de vida de 70.72, por lo tanto se puede decir que en alrededor del 75 % de los municipios hay una esperanza de vida mayor a los 70.5 años. Se consideran a 38 municipios como atípicos ya que se encuentran por debajo del umbral inferior,  $ISI < 0.655$  (esperanza de vida menor a 64.3), donde la mayoría de estos casos se presentan en Oaxaca con 13 municipios, seguido por Veracruz con 7, se observa la presencia de

<sup>8</sup>Alejada de la media.

<sup>9</sup>La posición en el IDH es el puesto que se le asigna a cada municipio con base en el valor de IDH que obtuvieron, en ese sentido, el municipio con el menor valor presentado tiene el puesto 2442 y aquel con el mayor valor tiene el puesto 1.

Chiapas, Guerrero y Chihuahua con 4 municipios cada uno, Puebla con dos y por último los estados de Durango, Jalisco, Nayarit y San Luis Potosí con un sólo caso. El IDH en este conjunto de valores atípicos es Bajo o Medio Bajo, la mayor posición de IDH para el conjunto es 2197.

2. Con base en el INE se observa una notable desigualdad en los municipios a nivel educación. El nivel educativo en promedio es regular. La mediana nos indica que alrededor del 50 % de los municipios presenta un valor para  $INE > 0.758$ . Se consideran a 79 municipios como atípicos ya que se encuentran por debajo del umbral inferior,  $INE < 0.545$ , donde la mayoría de éstos se presenta también en Oaxaca con 29 municipios, seguido por Chiapas y Veracruz con 18 y 12 casos, respectivamente, le sigue Puebla y Guerrero con 9 y 8 municipios en cada uno de éstos y por último Nayarit presenta un sólo caso en esta situación. El IDH en este conjunto de valores atípicos es Bajo o Medio Bajo, la mayor posición de IDH para el conjunto es 2191.
3. IPIB presenta valores casi en todo su rango  $[0,1]$ . Para el año 2000, el PIB per cápita anual ajustado va de 35,594 a 149 dólares estadounidenses PPC. En ese sentido, es rotunda la desigualdad que se vive en el país a nivel ingreso. La mediana se asocia con un PIB per cápita anual ajustado de 2,852 dólares, entonces la probabilidad de que un municipio mexicano tenga un PIB per cápita entre 149 y 2,852 es de 0.5. Analizando los cuartiles del índice se tiene que: el 25 % de los municipios tiene un PIB per cápita menor a 1,671 dólares; también, el 50 % tiene un PIB per cápita entre 1,671 y 4,384 dólares. Se consideran a 12 municipios como atípicos por el umbral inferior ya que se encuentran con un valor  $IPIB < 0.229$  (PIB per cápita menor a 394 dólares), donde 11 de ellos son de Oaxaca y uno de Chiapas. Este grupo de municipios atípicos se encuentra en niveles de Bajo a Medio bajo de IDH, la posición máxima alcanzada en el mismo es de 2134. También se consideran a 11 municipios como atípicos por el umbral superior, estos tienen un valor de  $IPIB > 0.87$  (PIB per cápita mayor a 18,357 dólares), donde 9 de ellos provienen del Distrito Federal, 2 de Nuevo León y en Campeche y Coahuila hay un municipio en cada uno de ellos con esta particularidad. Este grupo es tan solo el 0.47 % de todos los municipios, sus miembros se encuentran en el nivel Alto de IDH, la posición mínima es la 48.

La variable IDH ya se ha descrito con anterioridad, así que ya sólo se hablará sobre sus valores atípicos con base en la Figura 3.1 y el Cuadro 3.4. Se considera un conjunto de 19 municipios atípicos por el umbral inferior para IDH, ya que presentan valores para  $IDH < 0.478$ , donde en su mayoría son municipios de Oaxaca, 10 en total, seguido de Chiapas, Guerrero y

<b>Valor</b>	<b>IDH</b>
<b>Valor mínimo</b>	0.362
<b>Representante</b>	Coicoyán de las Flores (Oaxaca)
<b>Umbral inferior</b>	0.478
<b>Atípicos inferiores</b>	19
<b>Primer cuartil</b>	0.644
<b>Mediana</b>	0.702
<b>Tercer cuartil</b>	0.755
<b>Umbral superior</b>	0.909
<b>Atípicos superiores</b>	1
<b>Valor máximo</b>	0.930
<b>Representante</b>	Benito Juárez (D.F.)
<b>Posición de IDH</b>	1

Cuadro 3.4: Cuartiles y umbrales (IDH).

Veracruz con 4, 3 y 2 casos, respectivamente. Otro valor atípico es el antes mencionado municipio Benito Juárez (Distrito Federal), ya que rebasa el umbral superior, para el año 2000 éste fue el municipio con mayor desarrollo humano, según el IDH.

### 3.3. Índice de Intensidad Migratoria (IIM)

Los objetivos de esta sección son los siguientes: exhibir la construcción del Índice de Intensidad Migratoria México-Estados Unidos (IIM); hacer un análisis descriptivo de las variables involucradas en la construcción de éste; y por último, interpretar y describir a la variable IIM.

El índice de intensidad migratoria es una medida que permite diferenciar entidades federativas y municipios según la intensidad global del fenómeno migratorio a estas escalas.

La unidad de análisis para la construcción del IIM es el hogar; ya que, los hogares son el espacio donde los individuos inician, desarrollan y amplían el espacio de sus capacidades y, en muchos sentidos, en ellos subyacen las razones o motivaciones de muchos de sus comportamientos. En particular, la migración internacional es una estrategia que es vislumbrada por un cada vez mayor número de hogares como opción viable para mejorar las condiciones de vida de sus integrantes. Además, la partida o el regreso de sus miembros implica cambios en la estructura y composición de los hogares, así como en las relaciones intrafamiliares.

En la estimación del índice de intensidad migratoria se consideró como fuente de información la muestra censal del año 2000. El tamaño de esta muestra (alrededor de 2.2 millones de viviendas), permite contar con la información necesaria para la construcción de un índice multivariado a escala estatal y municipal, que de cuenta de las principales manifestaciones de la migración de mexicanos al país vecino. Para los fines de este trabajo sólo se hace uso del índice a nivel municipal.

El índice de intensidad migratoria es el resultado de una técnica de componentes principales, destacando que en virtud de la estructura de los datos, la proporción de la varianza explicada por la primera componente y los resultados de las pruebas estadísticas se justifica la aplicación de esta técnica con el fin de resumir la evidencia captada por la muestra censal (para detalles y resultados sobre la técnica de componentes principales véase [CONAPO \(2000\)](#)).

Las cuatro variables involucradas en la construcción del IIM se presentan a continuación, se denota a las variables con  $I_i$ , para  $i = 1, 2, 3, 4$ :

1. *Hogares que reciben remesas* ( $I_1$ ). Unidades domésticas donde al menos uno de sus miembros declaró recibir transferencias de familiares desde otro país.

El cálculo de este indicador consistió en dividir el número de hogares donde alguno de los miembros reciben remesas entre el monto total de hogares de cada municipio:

$$I_1 = \frac{H^{\text{rem}}}{H^{\text{tot}}} \times 100,$$

donde:

$H^{\text{rem}}$ : número de hogares que reciben remesas,

$H^{\text{tot}}$ : total de hogares.

2. *Hogares con emigrantes en Estados Unidos del quinquenio anterior* ( $I_2$ ). Refiere aquellos hogares donde alguno o algunos de sus miembros dejó el país, en el quinquenio anterior, para establecer su residencia habitual en la Unión Americana.

Este indicador resulta de dividir los hogares que cuentan con emigrantes en Estados Unidos del quinquenio anterior (1995-2000) entre el total de hogares:

$$I_2 = \frac{H^{\text{migEU}}}{H^{\text{tot}}} \times 100,$$

donde:

$H^{\text{migEU}}$ : número de hogares que cuentan con emigrantes en Estados

Unidos del quinquenio anterior,  
 $H^{\text{tot}}$ : total de hogares.

3. *Hogares con migrantes circulares del quinquenio anterior ( $I_3$ )*. Hogares en cuyo seno se ubica algún miembro que después de haber emigrado al vecino país entre 1995 y 2000, regresó a vivir a México.

Para el cálculo de este índice se identificó el número de hogares que cuentan con miembros del hogar que se fueron a Estados Unidos en el periodo indicado, pero que al momento del levantamiento censal ya habían regresado a residir al país y se dividió entre el total de hogares:

$$I_3 = \frac{H^{\text{circ}}}{H^{\text{tot}}} \times 100,$$

donde:

$H^{\text{circ}}$ : número de hogares que cuentan con miembros que se trasladaron a Estados Unidos durante 1995-2000, pero que regresaron en el mismo quinquenio,

$H^{\text{tot}}$ : total de hogares.

4. *Hogares con migrantes de retorno ( $I_4$ )*. Unidades con algún miembro, nacido en México, que en 1995 vivía en Estados Unidos y que regresó a residir al país, de tal manera que al momento del levantamiento censal (2000) se ubicaba, de nueva cuenta, como un habitante del territorio nacional.

En este caso, se identificó el número de hogares con algún miembro que en 1995 vivía en Estados Unidos pero que en 2000 ya había regresado a vivir a México, y se dividió entre el total de hogares:

$$I_4 = \frac{H^{\text{ret}}}{H^{\text{tot}}} \times 100,$$

donde:

$H^{\text{ret}}$ : número de hogares con personas nacidas en nuestro país que en 1995 vivían en Estados Unidos y que para 2000 ya habían regresado a México,

$H^{\text{tot}}$ : total de hogares.

Al igual que el IDH, el IIM implica una clasificación en niveles de intensidad migratoria (para detalles de como se asignaron dichos niveles, véase [CONAPO \(2000\)](#)).

En el Cuadro 3.5 se exhiben los valores máximos y mínimos de IIM que se presentan en cada nivel, así como la distribución de municipios para el año 2000 por nivel de intensidad migratoria. El Cuadro 3.5 muestra que la mayoría de los municipios cuenta con niveles bajos de intensidad migratoria. Nótese que disminuye la frecuencia por grupo a medida que aumenta el nivel

Nivel de Intensidad Migratoria	Frecuencia Relativa	Número de Municipios	Mínimo ( $\geq$ )	Máximo ( $<$ )
Nulo	92	3.77 %	0	0
Muy Bajo	873	35.75 %	-0.87874	-0.58777
Bajo	593	24.28 %	-0.58777	-0.00585
Medio	392	16.05 %	-0.00585	0.72156
Alto	330	13.51 %	0.72156	1.88452
Muy Alto	162	6.63 %	1.88452	6.39536 (=)

Cuadro 3.5: Niveles de Intensidad Migratoria.

de intensidad migratoria.

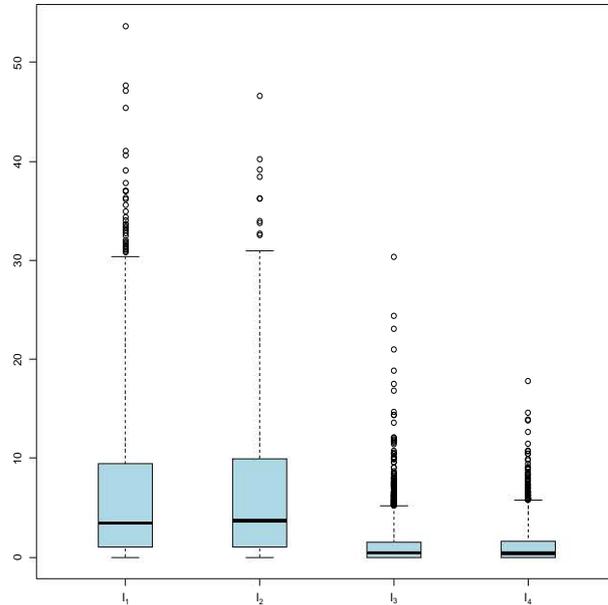
La variable IIM no tiene una escala bien definida ya que, como se observa en el Cuadro 3.5, aparentemente la intensidad migratoria crece a medida que el índice crece, pero si así fuera, la intensidad migratoria nula, representada por el valor de cero, sería el punto de partida para la variable, lo cual no ocurre. Más aún, podría parecer que se traslapa con la intensidad migratoria media, pero no es así, simplemente la intensidad migratoria nula se da cuando las cuatro variables que interactúan en el índice reportan el valor de cero. Por lo antes mencionado, es que se decidió analizar a la intensidad migratoria nula por separado a nivel estatal, como se muestra en el Cuadro 3.6.

Estado	Frecuencia	Frecuencia Relativa
Oaxaca	41	44.57 %
Chiapas	19	20.65 %
Yucatán	15	16.30 %
Puebla	10	10.87 %
Veracruz	5	5.43 %
Sonora	1	1.09 %
Tabasco	1	1.09 %
Total general	92	100 %

Cuadro 3.6: Intensidad Migratoria Nula.

Como se verá más adelante, de manera consistente, los estados que encabezan esta lista (Oaxaca, Chiapas y Yucatán) son también los que presentaron mayor número de municipios con el valor de cero en las variables que componen al IIM.

La Figura 3.2 muestra una comparativa entre los comportamientos de

Figura 3.2:  $I_1$ ,  $I_2$ ,  $I_3$  e  $I_4$ .

las variables  $I_1$ ,  $I_2$ ,  $I_3$  e  $I_4$  a nivel de sus respectivos cuartiles y rangos intercuartiles, cabe mencionar que nuevamente se hace uso de gráficos boxplot para dicho análisis, pero esta vez los umbrales serán localizados a 2.5 por el rango intercuartil, ésto se decidió debido a que, como se comenta más adelante, las variables se encuentran sumamente sesgadas hacia la izquierda y presentan valores primordialmente bajos, con esta decisión se reduce el número de valores atípicos con respecto a tomar la consideración habitual de 1.5. No se incluye en la comparación al IIM ya que esta variable, a pesar de ser el resultado de  $I_1$ ,  $I_2$ ,  $I_3$  e  $I_4$ , cuenta con otra escala como se observa en el Cuadro 3.5.

Es evidente que el valor mínimo para las variables  $I_i$ , con  $i = 1, 2, 3, 4$ , es cero, ésto por la existencia de un nivel de intensidad migratoria nulo. En las gráficas es claro el sesgo positivo tan alto de las cuatro variables, es por eso que los umbrales inferiores se encuentran muy cerca o son iguales a los primeros cuartiles. Calculando la curtosis se llega a que todas las variables tienen distribución leptocúrtica.

El Cuadro 3.7 muestra los cuartiles de las variables  $I_i$  junto con los umbrales superiores para la detección de valores atípicos, el número de municipios atípicos con base en cada variable y por último al municipio que

representa el valor máximo observado en cada una de éstas y su respectivo valor en IIM.

Valor	$I_1$	$I_2$	$I_3$	$I_4$
Primer cuartil	1	0.996	0	0
Mediana	3.423	3.700	0.445	0.396
Tercer cuartil	9.460	9.907	1.492	1.647
Umbral superior	30.408	30.990	5.215	5.751
Atípicos superiores	36	10	117	68
Valor máximo	53.714	46.666	30.389	17.834
Representante	Axutla (Puebla)	Santa Ana Del Valle (Oaxaca)	Teococuilco De Marcos Pérez (Oaxaca)	Santos Reyes Tepejillo (Oaxaca)
IIM	3.237	6.395	3.722	4.263

Cuadro 3.7: Cuartiles y límites intercuartiles ( $I_1$ ,  $I_2$ ,  $I_3$  e  $I_4$ ).

En la tabla anterior, a comparación de las que se presentaron para el IDH, no se considera el umbral inferior ya que, debido al gran sesgo positivo de las variables, el umbral inferior en todos los casos era menor al valor de cero como se observa en la Figura 3.2. A continuación presentamos una descripción general de las variables que constituyen al IIM con base en la Figura 3.2 y el Cuadro 3.7.

1. *Hogares que reciben remesas ( $I_1$ ):* Se observaron 166 municipios sin hogares en donde se recibían remesas de otro país, siendo Oaxaca el estado con mayor número de éstos con esta característica, 74 en total, seguido por Chiapas con 30. En aproximadamente el 50% de los municipios, menos del 3.5% de sus hogares reciben remesas de otra nación y sólo en alrededor del 25%, más del 9.5% de sus hogares las reciben. Se consideran a 36 municipios como atípicos por superar el umbral superior,  $I_1 > 30.408$ , es decir, municipios con más del 30.4% de hogares donde al menos uno de sus miembros declaró recibir transferencias de familiares desde otro país, en este grupo se encuentran 7 municipios de Oaxaca, 6 de Jalisco y Michoacán, 4 de Puebla y Zacatecas, 3 de Guanajuato, 2 de Durango y San Luis Potosí y 1 de Guerrero y Veracruz.
2. *Hogares con emigrantes en Estados Unidos del quinquenio anterior ( $I_2$ ):* Se observaron 215 municipios sin hogares donde alguno o algunos de sus miembros dejó el país, en el quinquenio anterior, para establecer su residencia habitual en la Unión Americana, donde el estado con mayor número de municipios con esta característica es Oaxaca, con 94 de éstos, seguido por Yucatán y Chiapas con 38 y 34, respectivamente. En aproximadamente el 50% de los municipios, en menos del 3.7% de

sus hogares se declaró tener a un familiar emigrante del quinquenio anterior y sólo en alrededor del 25 %, en más del 10 % de los hogares hubo al menos un integrante que haya establecido su residencia en el país vecino durante el tiempo indicado. Se consideran a 10 municipios como atípicos con  $I_2 > 30.99$ , es decir, municipios con más del 30.99 % de hogares donde al menos uno de sus miembros dejó el país entre 1995 y 2000 para establecer su residencia habitual en la Unión Americana, de donde 8 municipios son de Oaxaca, y de Durango e Hidalgo se observó un municipio en cada estado.

3. *Hogares con migrantes circulares del quinquenio anterior ( $I_3$ ):* Se observaron 707 municipios sin hogares con algún miembro que después de haber emigrado al vecino país entre 1995 y 2000, regresó a vivir a México. A nivel estado, con esta característica imperó Oaxaca con 312 municipios, seguido por Yucatán, Chiapas y Puebla con 80, 76 y 66, respectivamente. Alrededor del 75 % de los municipios, tenían menos del 1.5 % de hogares con migrantes circulares del quinquenio anterior. Se consideran a 117 municipios como atípicos ya que presentan valores para  $I_3 > 5.215$ , es decir, municipios con más del 5.2 % de hogares con algún miembro que después de haber emigrado a Estados Unidos, entre 1995 y 2000, regresó a vivir a México. Este grupo de municipios atípicos esta conformado por: 23 de Michoacán, 18 de Oaxaca, 15 de Zacatecas, 14 de Jalisco, 8 de Guanajuato, 7 de Hidalgo, 6 de Chihuahua, 4 en cada uno de los estados de Aguascalientes, Durango y Puebla, 2 en los estados de Guerrero, Morelos, Querétaro y Sonora y finalmente 1 municipio proveniente de los estados de Nuevo Leon, México, San Luis Potosí, Tamaulipas, Tlaxcala y Veracruz.
4. *Hogares con migrantes de retorno ( $I_4$ )* Se observaron 757 municipios sin hogares con migrantes de retorno, nuevamente Oaxaca presentó más casos con esta particularidad, 286 en total, seguido por Veracruz, Chiapas y Puebla con 106, 86 y 85, respectivamente. Tan sólo el 25 % de los municipios mexicanos presenta más del 1.65 % de hogares con algún miembro nacido en México, que en 1995 vivía en Estados Unidos y que regresó a residir al país antes del año 2000. Municipios atípicos fueron 68, por rebasar el umbral superior,  $I_4 > 17.834$ , grupo conformado por: 16 municipios de Oaxaca, 11 de Jalisco, 10 de Puebla y Zacatecas, 9 de Michoacán, 4 de Guerrero, 3 de Guanajuato y San Luis Potosí y por último uno de cada uno de los estados de Hidalgo y Nuevo León.

La variable IIM también se encuentra sumamente sesgada hacia la izquierda. Por lo tanto, el gráfico boxplot correspondiente será empleado bajo la misma configuración que para las variables  $I_i$ , como se muestra en la Figura 3.3. Cabe remarcar que los casos con valor cero fueron omitidos para evitar interpretaciones ambiguas de los estadísticos.

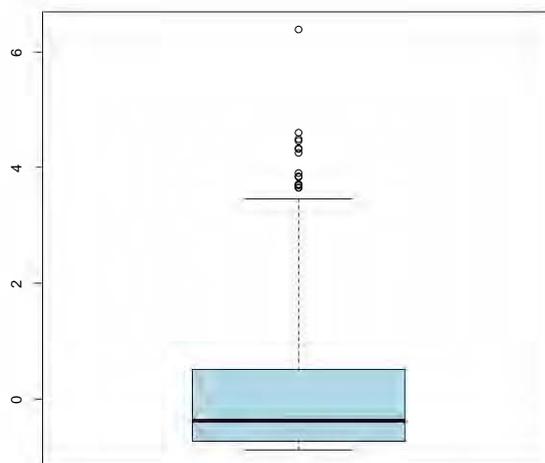


Figura 3.3: IIM.

Valor	IIM
Primer cuartil	-0.716
Mediana	-0.374
Tercer cuartil	0.518
Umbral superior	3.467
Atípicos superiores	15
Valor máximo	6.395
Representante	Santa Ana Del Valle (Oaxaca)

Cuadro 3.8: Cuartiles y límites intercuartiles (IIM).

Con base en los Cuadros 3.5 y 3.8 se observa que más del 50% de los municipios presenta niveles de Muy Bajo a Bajo de Intensidad Migratoria, menos del 25% presenta niveles de Alto a Muy Alto. Se consideran a 15 valores como atípicos, todos pertenecientes al nivel Muy Alto, por superar el umbral superior con  $IIM > 3.467$ , de donde se encontraron a 7 municipios de Oaxaca, 2 en los estados de Guanajuato y Zacatecas y por último uno en cada uno de los estados de Michoacán, Guerrero, Hidalgo y Puebla.

### 3.4. Ajuste Paramétrico Univariado

Recordemos que una de las finalidades primordiales del presente trabajo es proponer un modelo paramétrico que replique la interacción entre las

variables de estudio, en ese sentido, un modelo robusto requiere, además de la cópula subyacente de los datos, información sobre el comportamiento marginal de cada variable. Es por eso que se buscó un ajuste univariado paramétrico para las variables IDH e IIM, haciendo pruebas de bondad de ajuste para un amplio catálogo de distribuciones de entre las que destacan las siguientes: Normal, Weibull, Log-Normal, Exponencial, Gamma y Beta.

## IDH

De lo anterior se obtuvo que con un nivel de significancia  $\alpha = 0.05$   $IDH \sim Weibull(\sigma = 0.427913, S = 5.793495, \theta = 0.3)$ , como se muestra en la Figura 3.4, donde  $\sigma$  es la escala,  $S$  es la forma y  $\theta$  es un umbral inferior para la distribución, este último se tomó de 0.3 ya que para el año 2000 el IDH reportó un valor mínimo de 0.362.

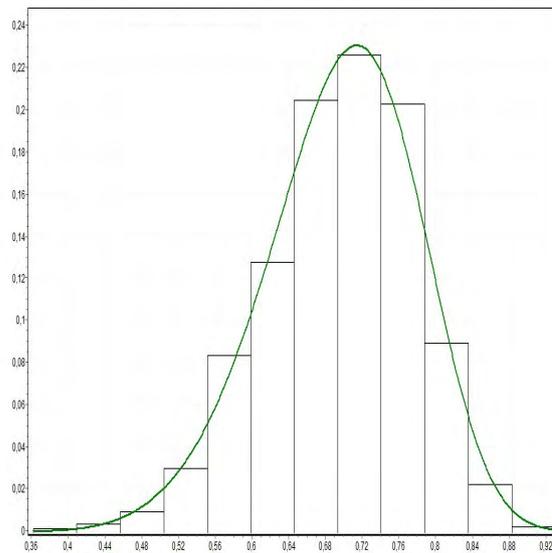


Figura 3.4: IDH.

La evidencia estadística del ajuste se muestra en el Cuadro 3.9.

Prueba	Estadístico	p-valores
Cramer-von Mises	W-Sq 0.06763841	0.250
Anderson-Darling	A-Sq 0.45364999	0.250

Cuadro 3.9: Pruebas de bondad de ajuste para distribución *Weibull*.

El Cuadro 3.10 muestra una comparación con percentiles observados y aquellos provenientes de la distribución ajustada.

Percentil	Observado	Estimado
1	0.485	0.493
5	0.557	0.556
10	0.587	0.590
25	0.644	0.645
50	0.702	0.701
75	0.755	0.752
90	0.794	0.794
95	0.816	0.817
99	0.849	0.856

Cuadro 3.10: Percentil observado contra distribución *Weibull*.

Notemos que si la variable IDH se distribuye como  $Weibull(\sigma, S, \theta)$  con los parámetros antes mencionados, ésta tiene su dominio en  $(\theta, \infty)$ , pero la variable IDH a lo más puede tomar el valor de uno. Se decidió tomar esta distribución por que fue la única para la cual no se rechazó la prueba de bondad de ajuste. Aun así, calculando la probabilidad de que IDH, con esta distribución, sea mayor a uno tenemos que  $\mathcal{P}[IDH > 1] \approx 0$ , justificando así lo adecuado de su uso.

## IIM

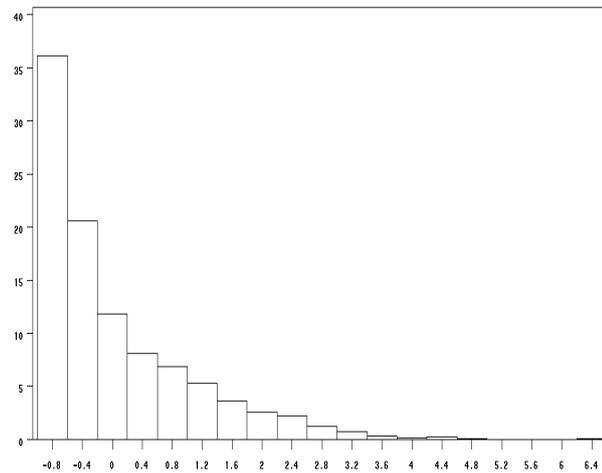


Figura 3.5: IIM.

El comportamiento marginal de la variable  $IIM|IIM \neq 0$  se muestra en la Figura 3.5. Por las razones que se han expuesto previamente, se anali-

zó al IIM omitiendo el valor de cero, además de que ésta se considera como la parte discreta de la variable IIM y la densidad completa de ésta tendría que estar dada por la combinación lineal convexa correspondiente a la densidad de la parte discreta y la de la parte continua. Cabe destacar que, el hecho de que  $IIM|IIM \neq 0$  sea una variable continua implica que la densidad en un punto es nula, por lo tanto  $IIM|IIM \neq 0 \equiv IIM$ . Así que, aun cuando se omitió de la muestra a los elementos que reportaron el valor de cero, se denotará con IIM a la variable a ajustar.

Para la variable en cuestión no se logró un ajuste univariado paramétrico directo. Dos características subyacen de la variable para complicar un ajuste: la primera, la variable se encuentra sumamente sesgada hacia la izquierda, presenta un sesgo de 1.5; segundo, el rango de la variable se encuentra distribuido entre números negativos y positivos,  $-0.878 \leq IIM \leq 6.395$ , lo cual afectó mucho los ajustes, aun modificando los umbrales de las distribuciones.

En ese sentido, se propone la transformación  $Y = \exp(IIM)$ , así el rango de la variable  $Y$  se encuentra en los reales positivos, finalmente, con esta modificación se logró un ajuste. Se realizaron las pruebas de bondad de ajuste para las mismas distribuciones que para IDH y fue así como, con un nivel de significancia  $\alpha = 0.05$ , se llegó a que  $Y \sim \text{Log-Normal}(\sigma = 2.0323, \mu = -1.2316, \theta = 0.41507)$ , como se muestra en la Figura 3.6, donde  $\sigma$  es la forma,  $\mu$  la escala y  $\theta$  un umbral inferior asociado con el mínimo valor reportado por  $Y$ .

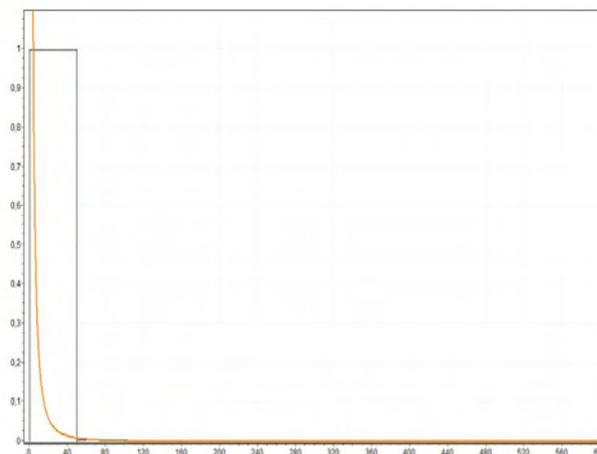


Figura 3.6:  $Y = \exp(IIM)$ .

La evidencia estadística del ajuste se muestra en el Cuadro 3.11.

Prueba	Estadístico	p-valores
Kolmogorov-Smirnov	D 0.01585935	0.150
Anderson-Darling	A-Sq 0.82276	0.4649

Cuadro 3.11: Pruebas de de bondad de ajuste para distribución *Log-Normal*.

El Cuadro 3.12 muestra una comparación con percentiles observados y aquellos provenientes de la distribución ajustada.

Percentil	Observado	Estimado
1	0.41786	0.41765
5	0.42560	0.42538
10	0.43692	0.43664
25	0.48859	0.48916
50	0.68806	0.70692
75	1.67908	1.56473
90	4.58777	4.36380
95	9.01734	8.67859
99	25.46732	33.43369

Cuadro 3.12: Percentil observado contra distribución *Log-Normal*.

Los últimos percentiles teóricos ya no parecen ser tan cercanos a los observados, pero en general el ajuste ofrece buenos resultados, el hecho de que la variable *Log-Normal* tome valores muy grandes con una probabilidad muy baja es consistente con los valores atípos que presenta la variable.

Si  $IIM = \ln(Y)$  y  $\mathcal{A} = [-0.878, \infty)$ , entonces:

$$\begin{aligned}
 f_{IIM}(x) &= \left| \frac{d}{dx} \exp(x) \right| * f_Y(\exp(x)) \mathbb{I}_{\mathcal{A}}(x) \\
 &= \frac{\exp(x) * \exp \left[ -(\ln(\exp(x) - \theta) - \mu)^2 / 2\sigma^2 \right]}{(\exp(x) - \theta) \sqrt{2\pi\sigma^2}} \mathbb{I}_{\mathcal{A}}(x)
 \end{aligned}$$

Donde  $f_{IIM}(\cdot)$ , es la función de densidad de la variable IIM, obteniendo así un modelo explícito para ésta.

Cabe mencionar que los estadísticos y p-valores de las pruebas de bondad de ajuste para el IDH y el IIM fueron obtenidos con R.

## Capítulo 4

# Segmentación y Ajuste Paramétrico Bivariado

En el presente capítulo se analiza la interacción entre el Desarrollo Humano y la Intensidad Migratoria, teniendo como objetivo primordial construir un modelo que explique y replique el comportamiento conjunto de estos fenómenos.

Recordemos que la introducción del capítulo anterior finalizó con la pregunta fundamental de esta investigación: «¿De qué manera se da esta relación?» Para darle respuesta a ésta, se formulan los siguientes cuestionamientos básicos, los cuales se irán resolviendo a lo largo de este capítulo:

- (a) **¿Cuál es el grado de dependencia que existe entre el desarrollo humano y la migración?**
- (b) **¿Qué tipo de dependencia guardan las variables?**
- (c) **¿Cuál es el comportamiento conjunto que se presenta entre el desarrollo humano y la migración?**

Como se analizará más adelante, la interacción entre las variables IDH e IIM es un tanto caótica y difícil de explicar. Es por esto que en la segunda sección de este capítulo se propone un algoritmo de segmentación a la muestra que logre capturar la mayor evidencia de dependencia estrictamente monótona en cada submuestra generada, sin dejar de lado que exista evidencia estadística de que la cópula subyacente de las variables sea simétrica en cada submuestra. Como ya se comentó en los primeros dos capítulos, la dependencia estrictamente monótona y la simetría son características que la gran mayoría de las cópulas paramétricas conocidas tienen.

Para concluir el capítulo y toda la investigación, en la última sección se hace una comparativa entre los resultados obtenidos y los que resultan

al hacer el análisis con métodos estadísticos más convencionales como la regresión lineal.

## 4.1. Comportamiento Conjunto

Esta sección tiene como finalidad presentar un análisis descriptivo sobre la relación conjunta que guardan las variables que componen al vector aleatorio  $(IDH, IIM)$ , para este fin se tienen los siguientes objetivos particulares: calcular y comparar las medidas de asociación presentadas en el capítulo dos; analizar el comportamiento de la cópula empírica de toda la muestra; y por último, evidenciar la complejidad de la estructura de dependencia que guardan las variables, a fin de motivar una partición a la muestra que permita explicar la dependencia total como un conjunto de dependencias más sencillas. Durante el desarrollo del trabajo se denotará con  $X$  a la variable aleatoria  $IDH$  y con  $Y$  a la variable aleatoria  $IIM$ ; es decir,  $X := IDH$  y  $Y := IIM$ , de igual manera,  $(X, Y) := (IDH, IIM)$ .

Cabe destacar que todos los procedimientos empleados y gráficas obtenidas fueron programadas en R (véase [R Development Core Team \(2012\)](#)). Las simulaciones de cópulas paramétricas, ajustes de cópulas paramétricas, prueba de independencia y prueba de simetría ([Genest y otros \(2011\)](#)) fueron empleados con el paquete “copula” de R (para mayor detalle sobre el uso del software, la programación de los procesos y el funcionamiento del paquete “copula”, véase: Apéndice A, [Yan \(2007\)](#), [Kojadinovic y Yan \(2010\)](#) y [Hofert y Mächler \(2011\)](#)).

Primero, recordemos que la muestra del vector aleatorio  $(X, Y)$  esta compuesta por 2442 registros, la Figura 4.1 muestra de manera gráfica el comportamiento conjunto de las variables, vía un gráfico de dispersión. Observamos 4 cosas relevantes: (1) no existe la mínima evidencia de que las variables puedan guardar una relación lineal; (2) no es claro cual es la relación imperante en la muestra, ¿La concordancia o la discordancia?; (3) de manera intuitiva es claro que existe un grado de dependencia entre las variables, pero, ¿Qué tan fuerte es el vínculo? y más aún, ¿Cómo podemos explicarlo?; (4) existen registros repetidos en la muestra, es decir, a un mismo nivel de intensidad migratoria se le asocian más de un nivel de desarrollo humano o viceversa, ésto complica aún más la explicación del fenómeno, además de que un ajuste bivariado con cópulas supone que las variables son absolutamente continuas, es decir, que no tienen puntos con densidad.

Un ejemplo claro de lo que se comenta en (4) es la Intensidad Migratoria Nula, este evento se da para distintos niveles de desarrollo humano y lleva a la construcción de una nueva medida de probabilidad que denotaremos por

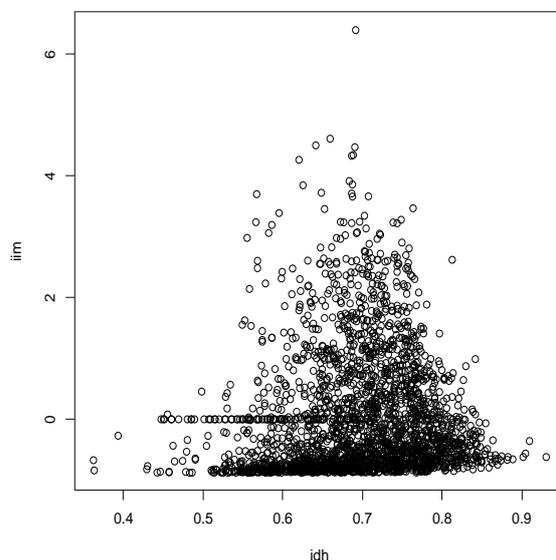


Figura 4.1: Comportamiento conjunto.

$X_N$ , donde abusando un poco de la notación  $X_N := X\{Y = 0\}$ . Son 92 casos en la muestra que cumplen con esta característica, una cantidad poco despreciable, además, la variable IIM no cuenta con una escala aparentemente establecida, como ya se comentó en el capítulo anterior. Es por eso que a la variable  $X_N$  se le estudiará por separado para no caer en interpretaciones ambiguas. A partir de este momento denotaremos con NDH al Nivel de Desarrollo Humano y con NIM al Nivel de Intensidad Migratoria. El Cuadro 4.1 muestra las frecuencias, absolutas y relativas, de intensidad migratoria nula por NDH.

NDH	Frecuencia Absoluta	Frecuencia Relativa
Bajo	10	32.25 %
MedioBajo	61	9.76 %
MedioAlto	20	1.26 %
Alto	1	0.49 %
Total	92	100 %

Cuadro 4.1: Intensidad Migratoria Nula por niveles de IDH.

La variable  $X_N$  explica al desarrollo humano a nivel municipal condicional a que el municipio tiene una intensidad migratoria nula. Cabe destacar que la variable toma valores desde 0.447 y hasta 0.803, es decir, en todos los

niveles de desarrollo humano tiene presencia la intensidad migratoria nula.

El Cuadro 4.2 muestra una tabla de contingencia con las frecuencias relativas entre el NDH y el NIM.

NDH/NIM	Muy Bajo	Bajo	Medio	Alto	Muy Alto	Total
Bajo	0.64 %	0.17 %	0.09 %	0.00 %	0.00 %	0.89 %
Medio Bajo	13.32 %	4.38 %	2.64 %	2.34 %	1.32 %	24.00 %
Medio Alto	19.91 %	16.68 %	12.85 %	11.57 %	5.53 %	66.55 %
Alto	3.28 %	4.00 %	1.11 %	0.13 %	0.04 %	8.55 %
Total	37.15 %	25.23 %	16.68 %	14.04 %	6.89 %	100.00 %

Cuadro 4.2: NDH vs NIM.

De los cuadros 4.1 y 4.2 se observa lo siguiente: para el desarrollo humano “Bajo”, la intensidad migratoria primordialmente se encontró en “Nula” y “Muy Baja”, no hubo un sólo municipio en niveles “Altos”; el nivel “Medio Bajo” es el que presentó mayor frecuencia en los municipios con intensidad migratoria “Nula”, en este nivel imperó la intensidad migratoria “Muy Baja”, aun cuando se observan todos los niveles bajando en frecuencia a medida que el NIM sube; el nivel de desarrollo humano “Medio Alto” es el que presentó mayor frecuencia en toda la muestra, éste interactuó con cada NIM reportando frecuencias considerables, también bajando a medida que la intensidad migratoria sube; el desarrollo humano “Alto” tuvo sólo un caso en los niveles de intensidad migratoria “Nula” y “Muy Alta”, respectivamente, imperó en los niveles “Bajos”.

Retomando el punto (4) respecto a las observaciones repetidas, esto es algo que se debe considerar antes de iniciar cualquier ajuste con cópulas. La justificación es que, por el teorema de Sklar, las cópulas se definen para variables aleatorias continuas y este detalle se puede entender como la parte discreta de una o ambas variables. Es por lo anterior que se propone sustituir estos casos por un sólo elemento constituido del valor repetido y la mediana de las observaciones de la otra variable condicional al valor repetido. En caso de que el número de elementos con esta característica sea considerable, se debe modelar la parte discreta y la parte continua del vector aleatorio.

En el IDH no se encontraron registros repetidos. Para IIM, sólo se presentaron los 92 registros de la intensidad migratoria nula y 2 más con el valor

-0.850725756. Quedando así 2350 registros como se muestran en la Figura 4.2.

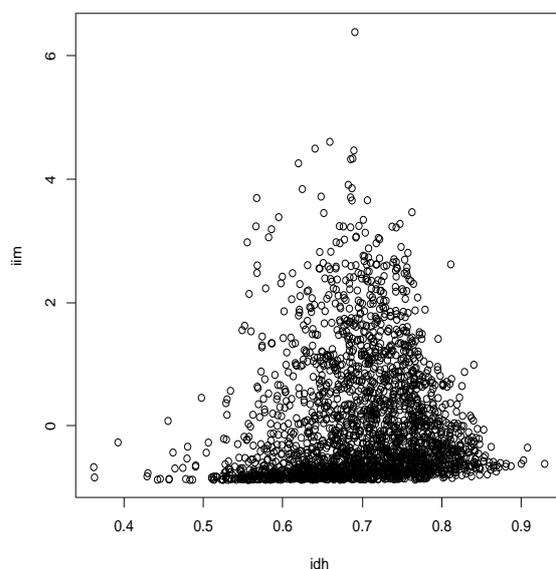


Figura 4.2: Parte continua de los datos.

Los datos presentados en la Figura 4.2 son los que se utilizarán para el desarrollo del modelo conjunto.

### Dependencia entre IDH e IIM

Como un primer paso para describir la relación conjunta entre las variables y hacer un ajuste con cópulas, es necesario tener evidencia estadística de que éstas no son independientes. Para determinar la independencia entre dos variables aleatorias  $X$ ,  $Y$  con función de distribución conjunta  $H$  y funciones de distribución marginal  $F$  y  $G$ , respectivamente, se requiere de una prueba que contraste la hipótesis nula  $\mathcal{H}_0 : H(x, y) = F(x)G(y)$ . En caso de que la prueba no sea rechazada, la relación entre las variables simplemente será determinada por la independencia y por ende no amerita un ajuste con cópulas. Para dichos fines se hace uso de la prueba de independencia contenida en el paquete “copula” de R (véase [Yan \(2007\)](#), [Kojadinovic y Yan \(2010\)](#) y [Hofert y Mächler \(2011\)](#)). Aplicando la antes mencionada prueba a toda la muestra se obtuvo un p-valor=0.0004, por lo tanto no existe evidencia para determinar que las variables aleatorias IDH e IIM son independientes.

Ya se ha determinado que las variables no son independientes, pero ¿Cuál

es el grado de dependencia que éstas tienen? Para dar respuesta a esta pregunta, primero recordaremos a las medidas de asociación del capítulo dos. El Cuadro 4.3 presenta la manera en la que se denotarán durante el resto del presente trabajo.

Notación	Estadístico
$\hat{r}$	Correlación lineal de Pearson
$\hat{\rho}$	Correlación de Spearman (Medida de concordancia)
$\hat{\sigma}$	Medida de dependencia no paramétrica de Schweizer & Wolff
$\hat{\phi}$	Medida de dependencia no paramétrica $\mathcal{V}_2$ ( $\mathcal{V}_2^2$ también conocido como el índice de dependencia de Hoeffding)

Cuadro 4.3: Notación de las medidas de asociación.

Donde para los cálculos de las medidas se usan las equivalencias muestrales de la siguiente manera:  $\hat{r}$  como en (2.1.1),  $\hat{\rho}$  como en (2.3.3),  $\hat{\sigma}$  como en (2.4.5) y  $\hat{\phi}$  como en (2.4.6).

A continuación se comparan las medidas de asociación correspondientes a la muestra.

Medida	Valor	Descripción
$\hat{r}$	0.046	No hay evidencia de relación lineal.
$\hat{\rho}$	0.162	La concordancia impera en la relación conjunta.
$\hat{\sigma}$	0.237	Evidencia significativa de dependencia.
$\hat{\phi}$	0.236	Evidencia significativa de dependencia.

Cuadro 4.4: Medidas de asociación: IDH e IIM.

A partir del Cuadro 4.4 podemos decir lo siguiente sobre la relación conjunta entre las variables de estudio:

- La concordancia es la relación que impera en la estructura de dependencia de las variables, lo que implica que para alguna parte de la muestra, a medida que el IDH crece, el IIM también crece en alguna proporción.
- Existe una diferencia entre  $\hat{\sigma}$  y  $\hat{\rho}$  de casi 0.1, lo que implica que en la muestra existe una parte importante donde impera la discordancia y que está anulando parte de concordancia en  $\hat{\rho}$ .
- Por el punto anterior, existe cierta parte de la muestra donde a medida que crece el IDH, el IIM decrece en alguna proporción.

- En contraste con el tamaño de muestra, la proporción de la muestra donde impera la concordancia es mayor a aquella donde impera la discordancia.
- Existe un grado de dependencia considerable de alrededor del 25 %.

Con base en la información del Cuadro 4.4 podemos conocer de forma muy general la estructura de dependencia, más adelante se determinará en que parte, o partes, de la muestra y en que grado es donde impera la concordancia, así como aquella donde impera la discordancia. En caso de que aplique, también se explicará que ocurre con el resto de la muestra.

### Comportamiento de la Cópula Empírica entre IDH e IIM

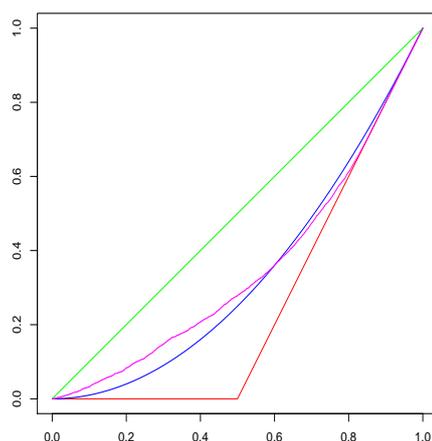


Figura 4.3: Diagonal de la cópula empírica.

La Figura 4.3 muestra una comparativa de la diagonal de la cópula empírica (rosa) y las diagonales de  $\mathcal{M}$  (verde),  $\mathcal{W}$  (rojo) y  $\Pi$  (azul), donde se observa lo siguiente: en alrededor del 55 % de la muestra impera la concordancia; existe una parte de la muestra (alrededor del 8 %) que se encuentra muy cerca de la cópula  $\Pi$  (al menos a nivel de la diagonal); en el resto de la muestra impera la discordancia. Las ocasiones siguientes que se muestre una comparativa de diagonales se mantendrá el mismo formato y colores.

La Figura 4.4 muestra la gráfica de las pseudo-observaciones de la cópula (como en la Definición 1.11.3), ésta da un acercamiento al comportamiento conjunto del vector  $(\hat{U}, \hat{V})$ . Dicha gráfica ofrece un panorama amplio sobre

la simetría de la cópula de los datos<sup>1</sup> y eventualmente puede ayudar a saber que camino tomar para hacer ajustes paramétricos de cópulas. Aplicando la prueba de simetría de [Genest y otros \(2011\)](#) se obtuvo un  $p$ -valor=0.022, por lo tanto, se rechaza la simetría con un nivel de significancia  $\alpha = 0.05$ . La prueba antes mencionada nos da evidencia suficiente para afirmar que la cópula subyacente no es simétrica.

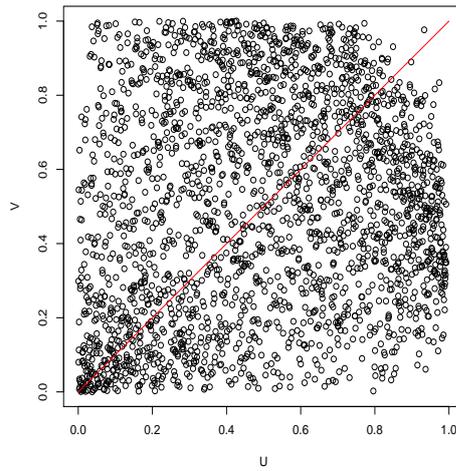


Figura 4.4: Pseudo-observaciones de la cópula.

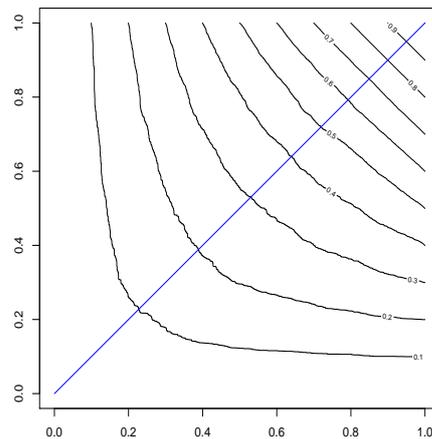


Figura 4.5: Gráfica de contornos.

<sup>1</sup>La gráfica se acompaña de una línea transversal la cual se pretende sirva como “espejo”.

Otra manera de apreciar la simetría de la cópula es mediante la gráfica de contornos (véase sección 1.2) que se presenta en la Figura 4.5.

Finalmente, con base en la diagonal de la cópula, en la comparativa de las medidas de asociación y la prueba de simetría, podemos afirmar que la cópula subyacente no es simétrica y no cumple con guardar una dependencia estrictamente monótona. Como se comentó en los dos primeros capítulos, la gran mayoría de las cópulas paramétricas son simétricas y representan dependencias estrictamente monótonas. Sólo para obtener la evidencia estadística, se le hicieron a los datos pruebas de bondad de ajuste para un catálogo amplio de cópulas<sup>2</sup> y todas fueron rechazadas. Por lo tanto, en lo que sigue de este capítulo, se propondrán **cortes** a la muestra, con la finalidad de buscar que, de manera individual, las variables en cada submuestra guarden una dependencia estrictamente monótona que provenga de una cópula simétrica, ya que es más fácil ajustar e interpretar este tipo de dependencias.

## 4.2. Algoritmo de Segmentación Bivariada

Esta sección es una de las partes más importantes de toda la tesis, ya que en ésta se propone una alternativa para llegar a un modelo conjunto con cópulas que sea capaz de representar, de la manera mas acertada posible, la estructura de dependencia que guardan las variables IDH e IIM. Una vez obtenido este modelo, nos servirá para hacer simulaciones de las variables de estudio y obtener funciones de regresión  $\alpha$ -cuantil con cópulas paramétricas. Para los fines de este trabajo, el modelo, las interpretaciones y conclusiones sobre la interacción entre las variables, se realizarán tomando como variable explicatoria al IDH, en ese sentido, todos los resultados serán enfocados con respecto al impacto que tiene el desarrollo humano sobre la intensidad migratoria.

Tras haber realizado la limpieza de los datos, el IDH y el IIM son variables aleatorias absolutamente continuas, por el Teorema de Sklar (Teorema 1.3.1) existe una cópula,  $\mathcal{C}_{IDH,IIM}$ , subyacente de las variables aleatorias IDH e IIM. Como se determinó anteriormente, dada la estructura de dependencia que guardan las variables y el comportamiento de la cópula empírica, no se encontró una cópula paramétrica que pudiera representar a toda la muestra. Por lo anterior, esta sección presenta una propuesta de segmentación a la muestra original, tomando en consideración muchas de las características de una cópula paramétrica mencionadas previamente.

---

<sup>2</sup>Las cópulas evaluadas fueron las mismas que serán probadas con cada submuestra resultante, más adelante se presenta el catálogo explícito de éstas.

Una vez obtenidas las submuestras de la muestra original, se procederá a aplicar, a las variables en cada submuestra, las pruebas de bondad de ajuste para las siguientes familias de cópulas (véase [Genest, Rémillard y Beaudoin \(2009\)](#)):

- Arquimedianas: Clayton (1.10.3), Frank (1.10.6), Gumbel (1.10.5) y Ali-Mikhail-Haq (AMH, (1.10.4)).
- Elípticas: Gaussiana (Normal, (1.3.3)) y T (véase [Demarta y McNeil \(2005\)](#)).
- Cópulas de valores extremos (véase [Yan \(2007\)](#) y [Kojadinovic y Yan \(2010\)](#)): Husler-Reiss y Galambos.
- Otras cópulas: Plackett (1.3.4) y Farlie-Gumbel-Morgenstern (FGM, (1.3.5)).

Una vez ajustados los datos en cada submuestra, se pegarán las cópulas con el método “Gluing Cópulas” (Proposiciones 1.3.1 y 1.3.2).

Tras conocer el comportamiento tanto marginal como conjunto de las variables de estudio y todos los elementos necesarios para un ajuste con cópulas, estamos en toda la posición de proponer un algoritmo de segmentación y comenzar con los cortes a la muestra.

### Algoritmo de Segmentación en Dependencias Estrictamente Monótonas (ASDEM)

Con la finalidad de buscar submuestras con la mayor evidencia de dependencia estrictamente monótona (véase: Definición 2.5.1 y Proposición 2.5.1), se propone un algoritmo para segmentar una muestra de tamaño  $M$ , éste a su vez incluye dos subprocesos iterativos (**INICIO** y **FINAL**) que interactúan para llevar a la mejor decisión de corte.

#### Algoritmo 4.2.1. ASDEM

Dada una matriz  $\Psi$  de dimensión  $M \times 2$  con las observaciones de un vector aleatorio  $(X, Y)$ , para segmentar los datos respecto a la variable explicatoria  $X$  en submuestras donde cada una de éstas guarde una dependencia estrictamente monótona que pueda ser modelada con una cópula paramétrica conocida, se efectúa lo siguiente:

1. Se ordena de forma ascendente a la matriz  $\Psi$  respecto a la variable  $X$ .
2. Se determina a  $\mathbf{n}$  como el tamaño mínimo que una submuestra puede tener. La primera justificación para el uso de un tamaño de submuestra mínimo es que para muestras muy pequeñas las medidas de asociación

tienden a ser inestables y la segunda es por que se busca que éstas sean representativas. Para lo que concierne al presente trabajo, se utilizó el 5% del tamaño de la muestra que se deseaba segmentar como el tamaño de muestra mínima.

3. El primer subproceso será llamado **INICIO**, éste se compone de  $M-n+1$  iteraciones. En la iteración  $k$  se toma una submuestra de  $\Psi$  de tamaño  $n-1+k$  con las observaciones de la **uno** a la  $n-1+k$  y se calcula  $(dif)_{k,i} = \hat{\sigma}_{k,i} - |\hat{\rho}_{k,i}|$ .

El segundo subproceso será llamado **FINAL**, éste se compone de  $M-n+1$  iteraciones. En la iteración  $k$  se toma una submuestra de  $\Psi$  de tamaño  $n-1+k$  con las observaciones de la  $M-n+2-k$  a la  $M$  y se calcula  $(dif)_{k,f} = \hat{\sigma}_{k,f} - |\hat{\rho}_{k,f}|$ .

De donde  $(dif)_{k,i}$  y  $(dif)_{k,f}$  representan las diferencias entre las medidas empíricas de asociación  $\hat{\sigma}$  y el valor absoluto de  $\hat{\rho}$  correspondientes a la submuestra de la iteración  $k$  de los procesos **INICIO** y **FINAL**, respectivamente. Véase Proposición 2.5.1.

4. Se busca la mínima diferencia en ambos subprocesos,

$$\begin{aligned} min_i &= \min_{\{k \in 1, \dots, M-n+1\}} ((dif)_{k,i}), \\ min_f &= \min_{\{k \in 1, \dots, M-n+1\}} ((dif)_{k,f}). \end{aligned}$$

Donde  $min_i$  corresponde a **INICIO** y  $min_f$  a **FINAL**.

5. Se calcula la mínima de las diferencias mínimas,  $\min(min_i, min_f)$ , y se obtiene el índice asociado a dicho valor (**IN**).
6. Si el subproceso correspondiente a **IN** es **INICIO** se determinan las siguientes submuestras de  $\Psi$ , con las observaciones:
  - de la **1** a la  $n-1+IN$ ;
  - de la  $n+IN$  a la **M**.

Si el subproceso correspondiente a **IN** es **FINAL** se determinan las siguientes submuestras de  $\Psi$ , con las observaciones:

- de la **1** a la  $M-n+1-IN$ ;
- de la  $M-n+2-IN$  a la **M**.

7. Para este momento se tiene un puntos de corte que genera dos submuestras distintas. Se calcula la diferencia entre  $\hat{\sigma}$  y  $|\hat{\rho}|$  para cada una de las submuestras obtenidas en el **punto 6**, se verifica que dichas

diferencias seas cercanas a cero. Las submuestras en las cuales se cumpla la condición pasan **punto 8** del algoritmo, para aquellas que no, puede ocurrir lo siguiente:

- a) en una de las submuestras no se cumple que la diferencia es cercana a cero, en tal caso se regresa al **punto 1** del **ASDEM**, pero esta vez sólo con la submuestra que no cumplió la condición;
  - b) ambas submuestras no cumplieron la condición, en tal caso surgen dos alternativas de acción: primero, buscar otro algoritmo de partición y/o de ajuste; y segundo, regresar al **punto 1** del **ASDEM** pero esta vez en paralelo para cada submuestra, aceptando de antemano que existirá alguna submuestra en la partición última que no representará una dependencia estrictamente monótona y puede que sea complicado su ajuste.
8. Se verifica que en la(s) submuestra(s) exista evidencia de simetría en la cópula, es decir, que no se rechace la prueba de simetría ([Genest y otros \(2011\)](#)) para cada submuestra. La(s) submuestra(s) que cumpla(n) la condición pasa(n) al **punto 9**, para aquellas que no se cumpla, se transforman los datos buscando que la simetría mejore. Si no se rechaza la simetría en la cópula de los datos transformados, se pasa al **punto 9** la(s) submuestra(s) transformada(s), en casos contrario se regresa al **punto 1** del **ASDEM** con aquella(s) submuestra(s) donde no se cumplió la condición.
  9. Se procede al ajuste de la cópulas para  $(X, Y)$  sobre la(s) submuestra(s) con base en [Genest y otros \(2009\)](#).

En el momento en el que al **punto 9** lleguen ambas submuestras, el **ASDEM** estará terminado y se tendrán dos o más submuestras, o transformaciones de las mismas, de la muestra  $\Psi$  con una estructura de dependencia estrictamente monótona. La intención del algoritmo es que se vayan intercalando los mejores cortes entre los subprocesos **INICIO** y **FINAL**. En caso de que se tenga una transformación de los datos,  $(\alpha(X), \beta(Y))$ , en una submuestra y su ajuste con cópulas, simplemente se procede a obtener la cópula de  $(X, Y)$  a partir de la cópula de  $(\alpha(X), \beta(Y))$ .

Las Figuras 4.7 y 4.8 muestran un diagrama de flujo sobre el funcionamiento del Algoritmo de Segmentación en Dependencias Estrictamente Monótonas (ASDEM).

### Comportamiento por Submuestra

A partir de este momento denotaremos por  $\Psi$  a la muestra de tamaño  $M=2350$  (recordar que fue el número de observaciones resultado de la limpieza

que se realizó) del vector aleatorio  $(X, Y)$ , que recordemos, representa al vector aleatorio compuesto por IDH, IIM. Si  $x_i, y_i$  denotan a la observación  $i$  de las variables aleatorias  $X, Y$  respectivamente, bajo la consideración de ordenar la muestra con base en  $X$  (IDH), entonces  $\Psi := \{(x_i, y_i) | i \in [1, 2350]\}$ . Recordemos que se ordena con base en  $X$  por que se pretende que el IDH sea la variable que explique a la IIM, por lo tanto será esta la variable por la cual se generarán los cortes.

Al emplear el Algoritmo 4.2.1 (ASDEM) a las variables de estudio, en una primera iteración, el subproceso **INICIO** presentó la mínima de las diferencias mínimas. Las dos submuestras resultantes presentaron evidencia de dependencia estrictamente monótona, pero para la submuestra a la derecha del corte, se rechazó la prueba de simetría de la cópula (p-value = 0.002498). Teniendo así un primer corte en  $IDH=0.6923088$ , como se muestra en la Figura 4.6.

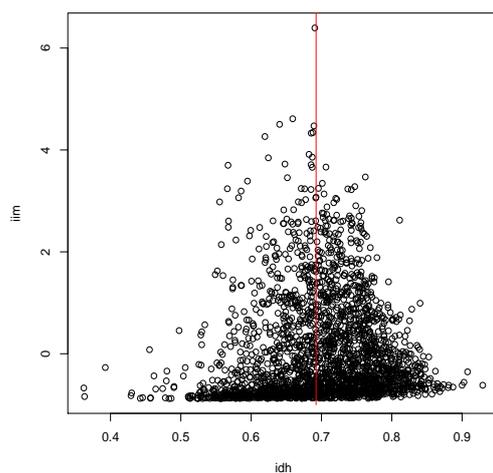


Figura 4.6: Primer corte:  $IDH=0.6923088$ .

A partir de esta primera iteración se espera que se pueda lograr un buen ajuste a las variables en el siguiente conjunto.

- $\Psi_{con} := \{(x_i, y_i) \in \Psi | i \in [1, 1018]\}$ : siendo éste el 30.7% de la muestra total, tomando valores desde 0.3619992 y hasta 0.6923088 para el IDH y el IIM toma todos sus posibles valores. Se denotará con  $X_{con}$  y  $Y_{con}$  a las representantes de las variables  $X, Y$  respectivamente, restringidas a este conjunto.

El subíndice “con” se le da a esta submuestra ya que, como veremos más adelante, en ésta impera la concordancia.

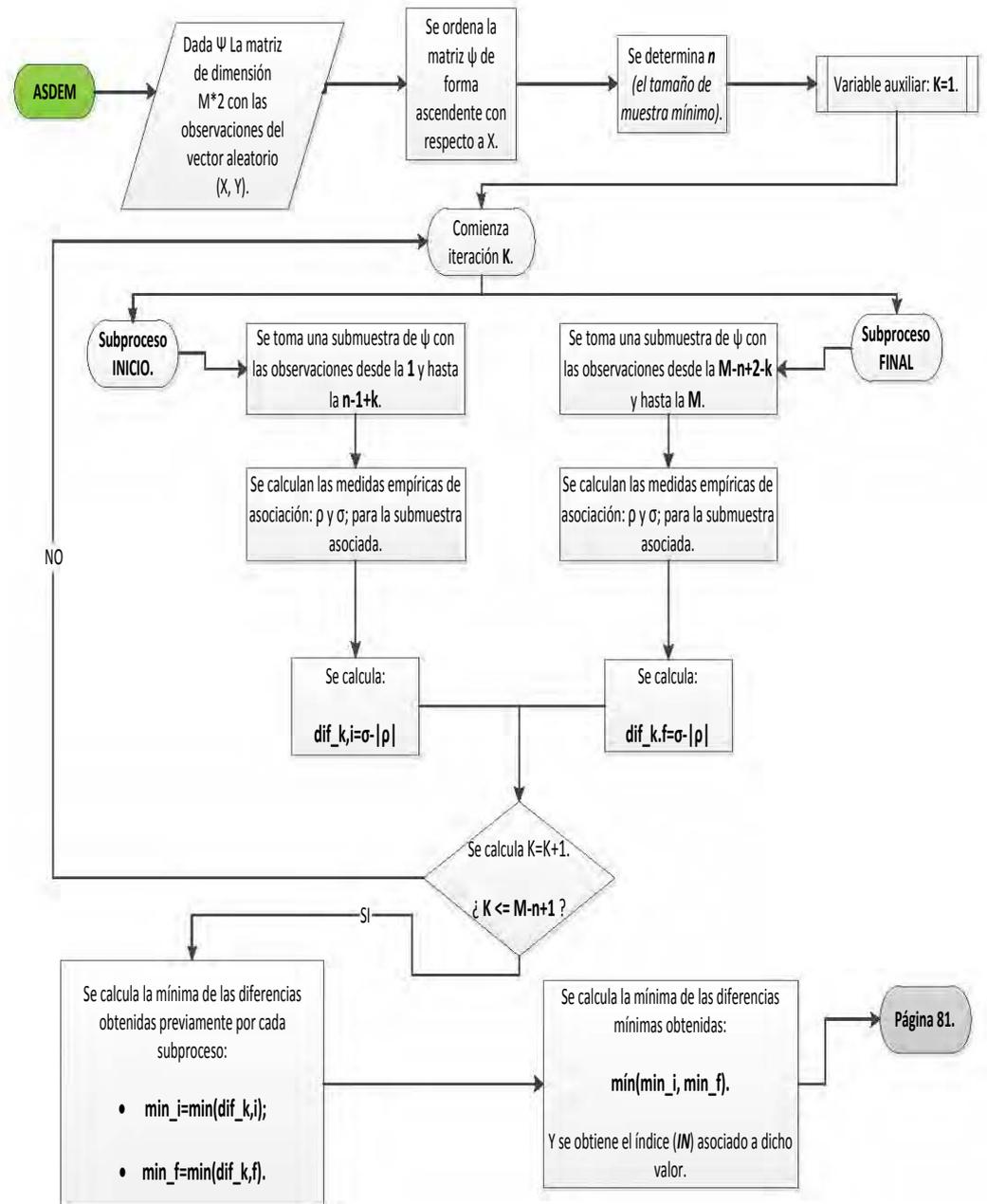


Figura 4.7: Algoritmo de Segmentación en Dependencias Estrictamente Monótonas (1).

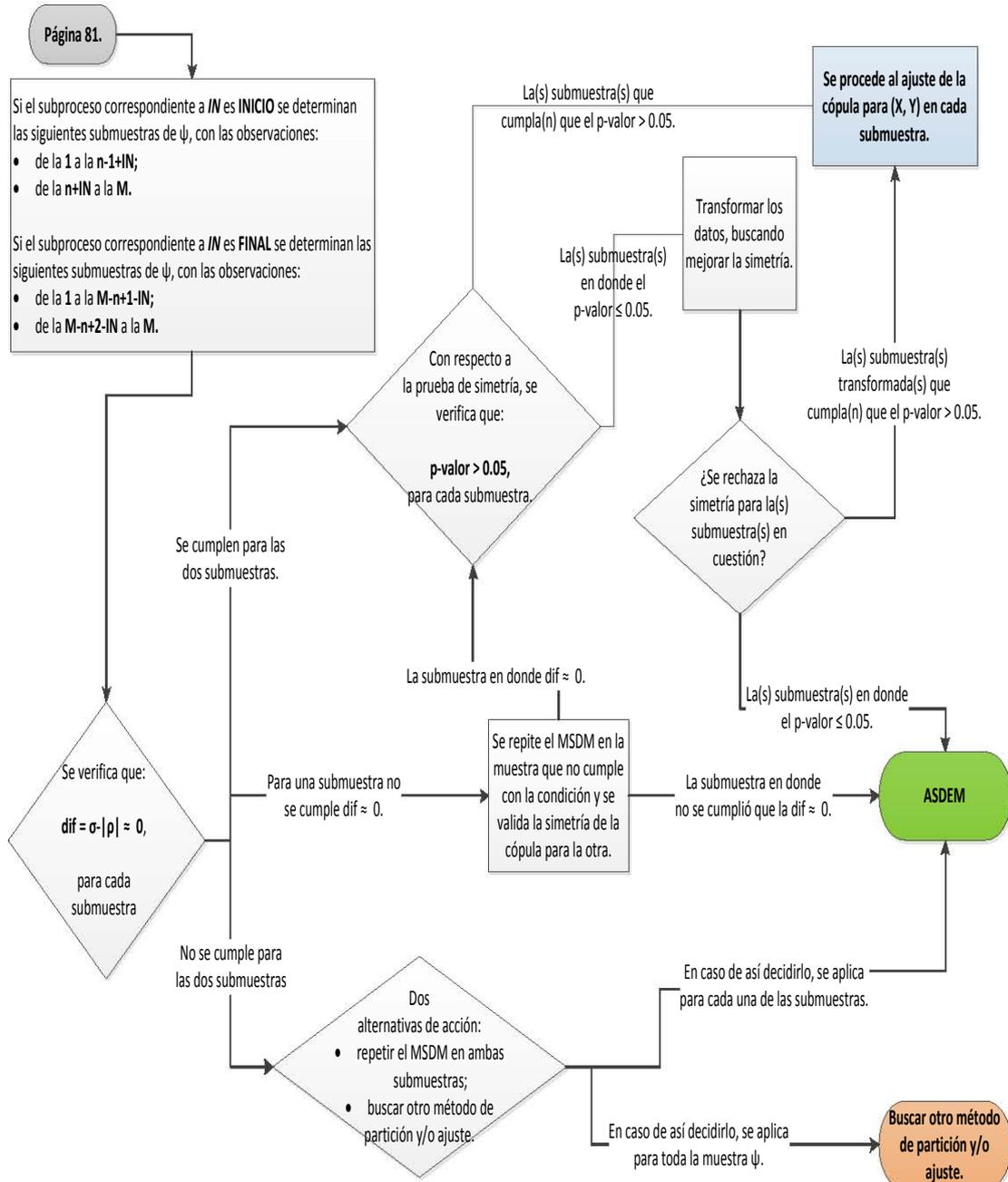


Figura 4.8: Algoritmo de Segmentación en Dependencias Estrictamente Monótonas (2).

Con base en el ASDEM, se buscó transformar los datos en la submuestra  $\Psi^* := \{(x_i, y_i) \in \Psi | i \in [1019, 2350]\}$ , sin haber obtenido éxito sobre la simetría de la cópula de las variables transformadas.

La siguiente iteración del algoritmo se realizó únicamente sobre la submuestra antes mencionada. Esta iteración fue muy similar a la primera, ya que se obtuvieron submuestras con dependencias estrictamente monótonas con la mínima de las diferencias mínimas proveniente del subproceso **INICIO**, pero en la submuestra a la derecha del corte se rechazó la simetría de la cópula correspondiente (p-value = 0.002498). Se generó un siguiente corte en IDH=0.7619343, como se muestra en la Figura 4.9.

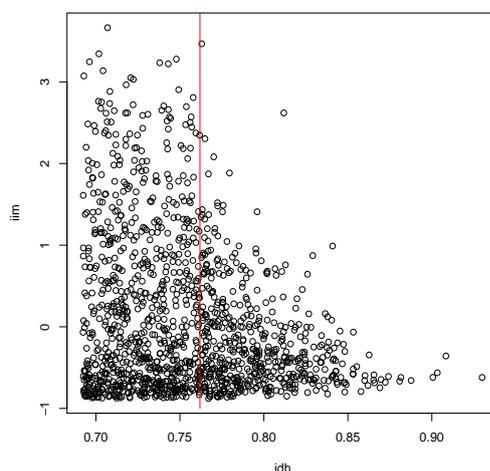


Figura 4.9: Segundo corte corte: IDH=0.7619343.

A partir de esta segunda iteración se obtuvo la siguiente submuestra con un comportamiento conjunto muy peculiar, la “*independencia*” denotada con el subíndice “*ind*”.

- $\Psi_{ind} := \{(x_i, y_i) \in \Psi | i \in [1019, 1817]\}$ : siendo éste el 34% de la muestra total, tomando valores desde 0.6923286 y hasta 0.7619343 para el IDH y desde -0.8734804 y hasta 3.663301 para IIM. Se denotará con  $X_{ind}$  y  $Y_{ind}$  a las representantes de las variables  $X$ ,  $Y$  respectivamente, restringidas a este conjunto.

Cabe mencionar que en dicha submuestra no se presenta una dependencia estrictamente monótona, simplemente las variables son independientes, pero por la naturaleza del algoritmo fue posible detectar dicha relación. Esta submuestra es muy importante, ya que es la parte caótica de la relación

conjunta, se da después de una relación imperante de concordancia y como se analiza más adelante le sigue una parte imperante de discordancia.

Nuevamente, con base en el ASDEM se buscó transformar a la submuestra a la derecha del corte, con lo que no se logró que se mejorara la simetría de la cópula respectiva.

La última iteración se realizó sobre el conjunto  $\Psi^{**} := \{(x_i, y_i) \in \Psi | i \in [1818, 2350]\}$ . Para ambos subprocesos (INICIO y FINAL) se obtuvieron dependencias estrictamente monótonas significativas, esta vez sí se obtuvieron dos submuestras con evidencia de provenir de cópulas simétricas. Se decidió utilizar nuevamente el corte proveniente del subproceso **INICIO**, por presentar mayor evidencia de dependencia estrictamente monótona (una mínima diferencia más cercana a cero). Se generó este último corte en  $IDH=0.7852387$ , como se muestra en la Figura 4.10.

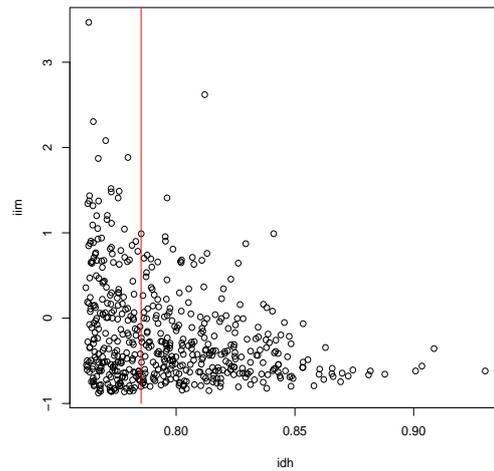


Figura 4.10: Tercer corte corte:  $IDH=0.7852387$ .

Con este corte surgen las siguientes submuestras, denotadas con los subíndices “*dis1*” y “*dis2*” debido a la imperante relación de discordancia en éstas.

- $\Psi_{dis1} := \{(x_i, y_i) \in \Psi | i \in [1818, 2051]\}$ : siendo éste el 9.9% de la muestra total, tomando valores desde 0.7619477 y hasta 0.7852387 para el IDH y desde -0.8770457 y hasta 3.467135 para IIM. Se denotará con  $X_{dis1}$  y  $Y_{dis1}$  a las representantes de las variables  $X$ ,  $Y$  respectivamente, restringidas a este conjunto.

- $\Psi_{dis2} := \{(x_i, y_i) \in \Psi | i \in [2052, 2350]\}$ : siendo éste el 12.7% de la muestra total, tomando valores desde 0.7853664 y hasta 0.9301854 para el IDH y desde -0.8498019 y hasta 2.619973 para IIM. Se denotará con  $X_{dis2}$  y  $Y_{dis2}$  a las representantes de las variables  $X, Y$  respectivamente, restringidas a este conjunto.

Quedando finalmente la partición como se presenta en la Figura 4.11.

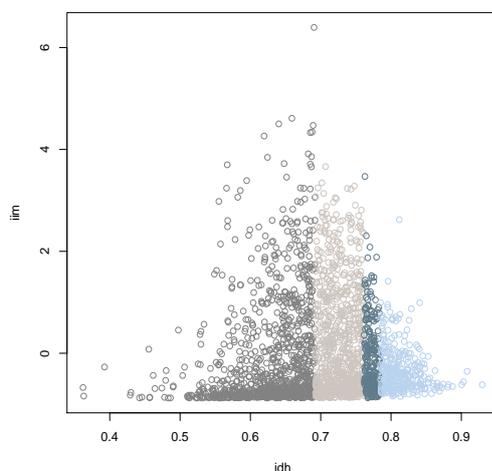


Figura 4.11: Cortes a la muestra.

Para que sea más claro el comportamiento de las variables en cada submuestra, en la Figura 4.12 se muestra el gráfico de dispersión de cada una de éstas. Además, el Cuadro 4.5 presenta las medidas de asociación del capítulo dos para cada submuestra.

<b>Muestra</b>	$\hat{r}$	$\hat{\rho}$	$\hat{\sigma}$	$\hat{\phi}$
$\Psi_{con}$	0.233	0.300	0.300	0.278
$\Psi_{ind}$	-0.015	0.008	0.052	0.052
$\Psi_{dis1}$	-0.225	-0.223	0.225	0.212
$\Psi_{dis2}$	-0.237	-0.227	0.238	0.235

Cuadro 4.5: Medidas de asociación por submuestra.

Las variables en las submuestras  $\Psi_{con}$ ,  $\Psi_{dis1}$  y  $\Psi_{dis2}$  presentan grados de dependencia relevantes, de alrededor del 25%. Como ya se había comentado es evidente que en  $\Psi_{con}$  impera la concordancia y que en  $\Psi_{dis1}$  y  $\Psi_{dis2}$  la discordancia. En estos casos la correlación lineal es un poco parecida a las medidas de concordancia y dependencia, pero se debe recordar que en general esta medida puede ser engañosa y explicar muy poco sobre la relación

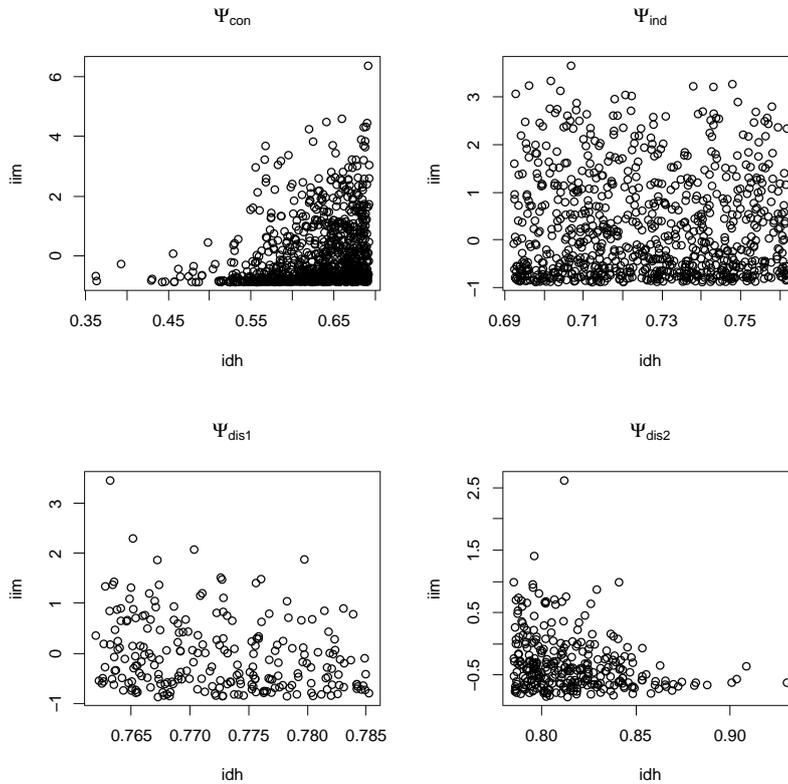


Figura 4.12: Comportamiento en cada submuestra.

conjunta. No cabe duda de que las variables en la submuestra  $\Psi_{ind}$  son independientes.

En la Figura 4.13 se presentan los gráficos de dispersión junto con las regresiones  $\alpha$ -cuantil con base en la cópula Bernstein como en (1.11.3). La línea gruesa representa la regresión mediana y las líneas delgadas son regresiones a los percentiles 75 y 25, respectivamente (véase sección 1.12), estas últimas forman una banda de probabilidad al 50%. Con las gráficas antes mencionadas es más evidente la relación conjunta de las variables en cada submuestra.

El Cuadro 4.6 muestra, por submuestra, los p-valores de la prueba de independencia, siendo ésta la evidencia estadística de que en las submuestras  $\Psi_{con}$ ,  $\Psi_{dis1}$  y  $\Psi_{dis2}$  existe un grado de dependencia que puede ser modelado y que para  $\Psi_{ind}$  es adecuado utilizar a la cópula  $\Pi$  para representar la independencia de las variables.

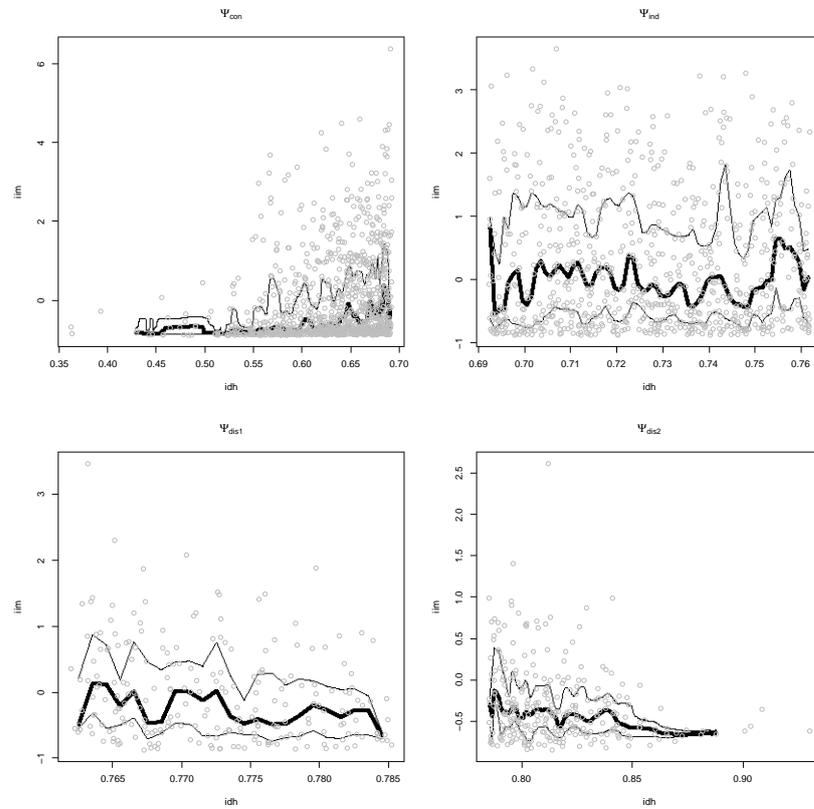


Figura 4.13: Regresiones  $\alpha$ -cuantil de cada submuestra.

Muestra	Independencia p-valor
$\Psi_{con}$	0.0004
$\Psi_{ind}$	0.4750
$\Psi_{dis1}$	0.0024
$\Psi_{dis2}$	0.0004

Cuadro 4.6: Prueba de independencia por submuestra.

### Comportamiento de la Cópula Empírica por Submuestra

La Figura 4.14 muestra la gráfica comparativa por submuestra de las diagonales de cada cópula empírica y las diagonales de  $\mathcal{M}$ ,  $\Pi$  y  $\mathcal{W}$ , donde se reitera el comportamiento en cada submuestra.

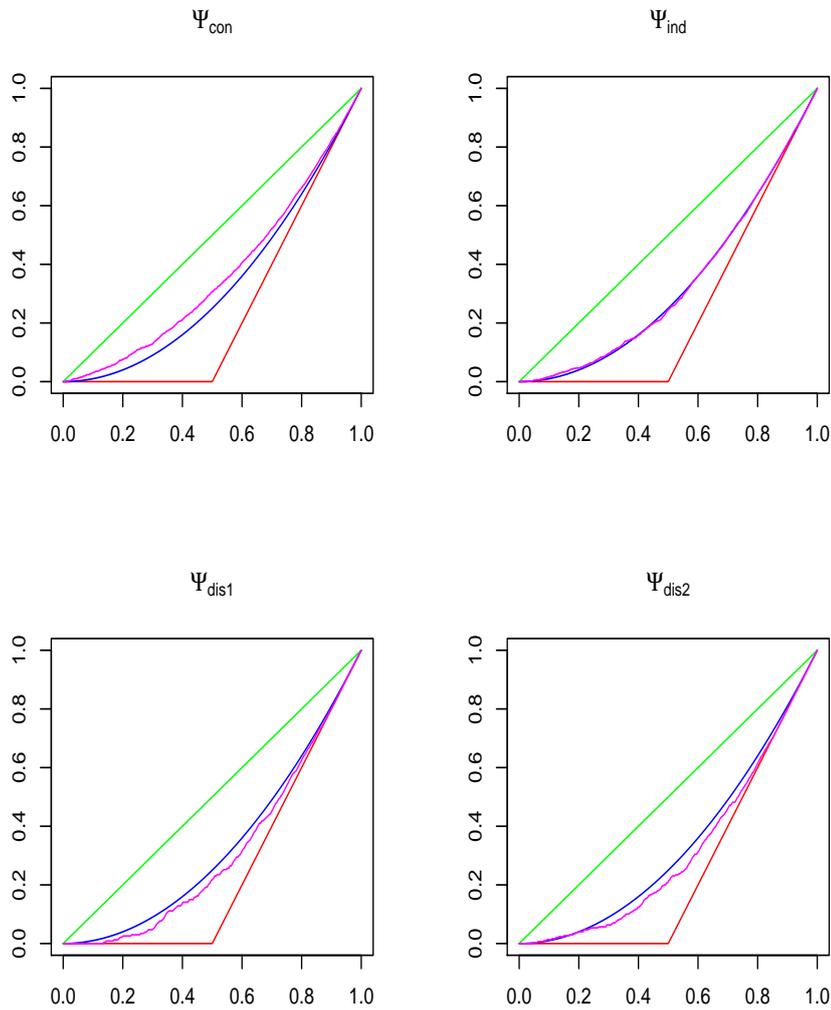


Figura 4.14: Diagonales de cada submuestra.

La Figura 4.15 muestra la gráfica de las pseudo-observaciones de la cópula subyacente de las variables en cada submuestra, donde se reitera la simetría de cada cópula.

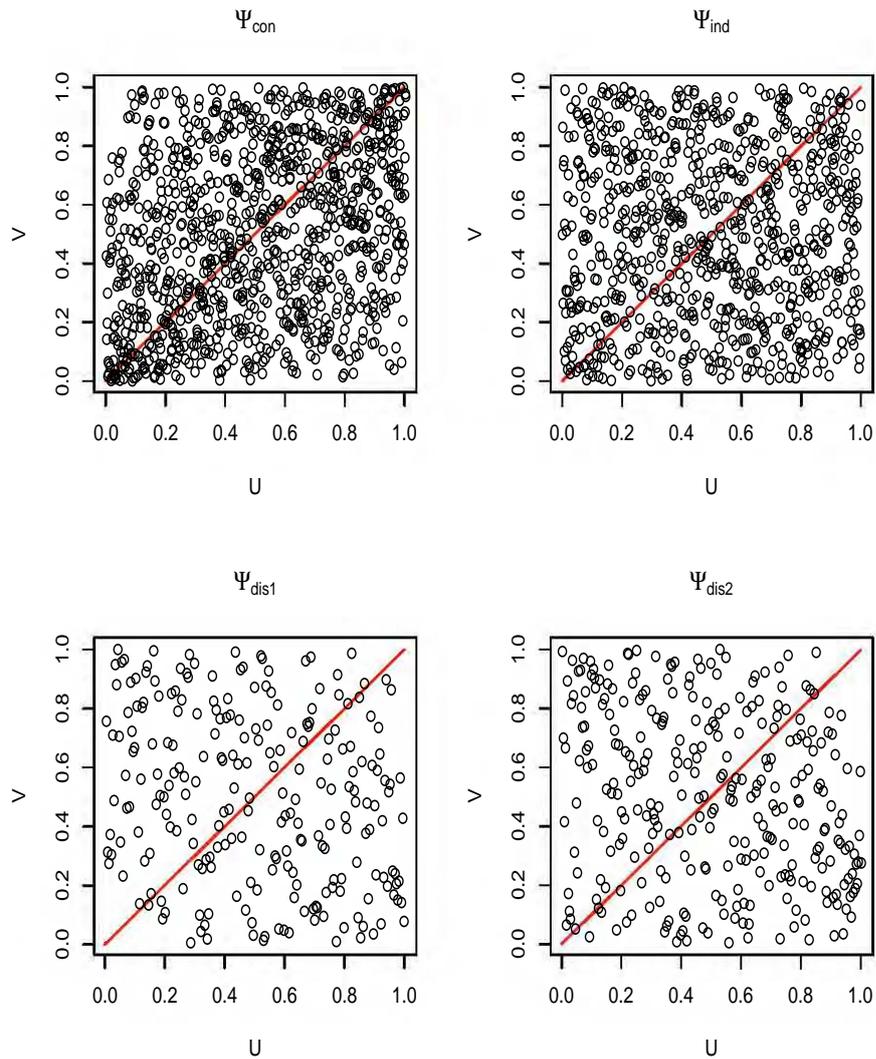


Figura 4.15: Pseudo-observaciones de cada submuestra.

La Figura 4.16 muestra la gráfica de contornos de la cópula subyacente de las variables en cada submuestra, recordar que la gráfica de contornos nos da una idea del comportamiento de las curvas de nivel de la cópula y también podemos apreciar la simetría.

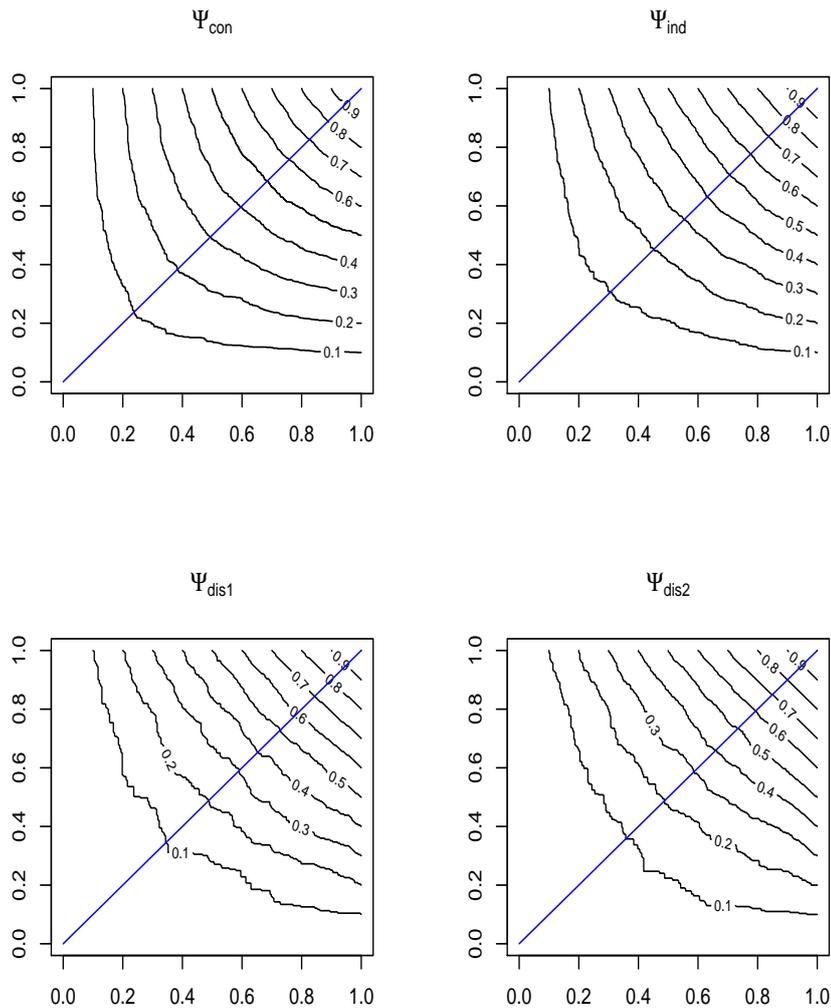


Figura 4.16: Gráfica de contornos de cada submuestra.

Finalmente, para tener evidencia estadística de que la cópula subyacente de las variables en las submuestras  $\Psi_{con}$ ,  $\Psi_{ind}$ ,  $\Psi_{dis1}$  y  $\Psi_{dis2}$  es simétrica, el Cuadro 4.7 presenta los p-valores de la prueba de simetría realizadas a cada una de las submuestras.

Muestra	Simetría p-valor
$\Psi_{con}$	0.919
$\Psi_{ind}$	0.257
$\Psi_{dis1}$	0.889
$\Psi_{dis2}$	0.244

Cuadro 4.7: Prueba de simetría por submuestra.

Con todos los puntos antes comentados, estamos en toda la posición de iniciar con el ajuste a las variables IDH e IIM restringidas a los conjuntos:  $\Psi_{con}$ ,  $\Psi_{dis1}$  y  $\Psi_{dis2}$ .

### 4.3. Ajuste Paramétrico Bivariado y Simulación Conjunta

El Cuadro 4.8 contiene, por cada submuestra, los p-valores de las pruebas de bondad de ajuste para las variables de estudio, junto con los parámetros de éstas. Sólo se incluyen las cópulas para las cuales no se haya rechazado la prueba de hipótesis.

Submuestra	Cópula	Parámetro (s)	p-valor
$\Psi_{con}$	Gaussiana (Normal)	$r=0.3234996$	0.418
	T	$\rho = 0.3223847$ $df=35.61604$	0.380
	Plackett	$\theta = 2.524974$	0.132
$\Psi_{ind}$	II	-	-
$\Psi_{dis1}$	Plackett	$\theta = 0.5351376$	0.998
	T	$\rho = -0.2343981$ $df=233.5822$	0.982
	Frank	$\theta = -1.336393$	0.977
	Gaussiana (Normal)	$r=-0.2350956$	0.973
$\Psi_{dis2}$	Gaussiana (Normal)	$r=-0.2248809$	0.201
	T	$\rho = -0.2254649$ $df=101.4803$	0.192
	Plackett	$\theta = 0.5069886$	0.161
	Frank	$\theta = -1.40362$	0.129

Cuadro 4.8: Ajuste a cada submuestra.

Recordemos lo que denotan los parámetros mencionados en el Cuadro anterior:  $r$ , representa al coeficiente de correlación lineal como en (2.1.1);  $\rho$ , representa a la medida de concordancia Rho de Spearman como en (2.3.2);

$df$ , se utiliza para representar a los grados de libertad en la cópula  $T$ , véase Demarta y McNeil (2005); y por último  $\theta$ , es el parámetro particular de cada cópula.

Finalmente, la cópula seleccionada para cada submuestra es aquella con el mayor p-valor. Entonces, con base en la definición de cada submuestra se tiene lo siguiente:

- $\mathcal{C}_{X_{con}, Y_{con}} := Normal(0.3234996)$ ,
- $\mathcal{C}_{X_{ind}, Y_{ind}} := \Pi$ ,
- $\mathcal{C}_{X_{dis1}, Y_{dis1}} := Plackett(0.5351376)$ ,
- $\mathcal{C}_{X_{dis2}, Y_{dis2}} := Normal(-0.2254649)$ .

Una vez realizada la selección en cada submuestra se pegan las cópulas con el método “Gluing Cópulas” (Proposiciones 1.3.1 y 1.3.2), resultando así la cópula paramétrica subyacente de las variables IDH e IIM.

Dadas las cópulas:  $\mathcal{C}_{X_{con}, Y_{con}}$ ,  $\mathcal{C}_{X_{ind}, Y_{ind}}$ ,  $\mathcal{C}_{X_{dis1}, Y_{dis1}}$ ,  $\mathcal{C}_{X_{dis2}, Y_{dis2}}$ , y tres valores fijos que representen la partición en  $u$ :  $\theta_1 = 0.4331915$ ,  $\theta_2 = 0.7731915$  y  $\theta_3 = 0.872766$ , podemos escalar  $\mathcal{C}_{X_{con}, Y_{con}}$  a  $[0, \theta_1] \times [0, 1]$ ,  $\mathcal{C}_{X_{ind}, Y_{ind}}$  a  $[\theta_1, \theta_2]$ ,  $\mathcal{C}_{X_{dis1}, Y_{dis1}}$  a  $[\theta_2, \theta_3]$  y  $\mathcal{C}_{X_{dis2}, Y_{dis2}}$  a  $[\theta_3, 1]$  y juntarlas en la cópula  $\mathcal{C}_{IDH, IIM}$  como sigue:

$$\mathcal{C}_{IDH, IIM}(u, v) = \begin{cases} \theta_1 \mathcal{C}_{X_{con}, Y_{con}}\left(\frac{u}{\theta_1}, v\right), & 0 \leq u \leq \theta_1, \\ (\theta_2 - \theta_1) \mathcal{C}_{X_{ind}, Y_{ind}}\left(\frac{u - \theta_1}{\theta_2 - \theta_1}, v\right) + \theta_1 v, & \theta_1 \leq u \leq \theta_2, \\ (\theta_3 - \theta_2) \mathcal{C}_{X_{dis1}, Y_{dis1}}\left(\frac{u - \theta_2}{\theta_3 - \theta_2}, v\right) + \theta_2 v, & \theta_2 \leq u \leq \theta_3, \\ (1 - \theta_3) \mathcal{C}_{X_{dis2}, Y_{dis2}}\left(\frac{u - \theta_3}{1 - \theta_3}, v\right) + \theta_3 v, & \theta_3 \leq u \leq 1. \end{cases}$$

En la Figura 4.17 se presentan simulaciones conjuntas del vector aleatorio  $(IDH, IIM)$  y los datos originales: en la parte superior, simulación con la cópula “Gluing”  $\mathcal{C}_{IDH, IIM}$  (Normal + Independencia + Plackett + Normal), de lado izquierdo con marginales Bernstein como en (1.11.2) y de lado derecho con las marginales obtenidas en la sección 3.4 ( $IDH \sim Weibull$  y  $\exp(IIM) \sim Log-Normal$ ); de lado inferior izquierdo, se presenta una simulación totalmente no paramétrica, con base en la cópula y marginales Bernstein (Ecuaciones (1.11.3) y (1.11.2)); de lado inferior derecho se presentan los datos originales. El tamaño de las simulaciones es de 2350, el tamaño de la muestra de estudio. Las simulaciones se realizan con base en los Algoritmos 1.9.2 y 1.9.3.

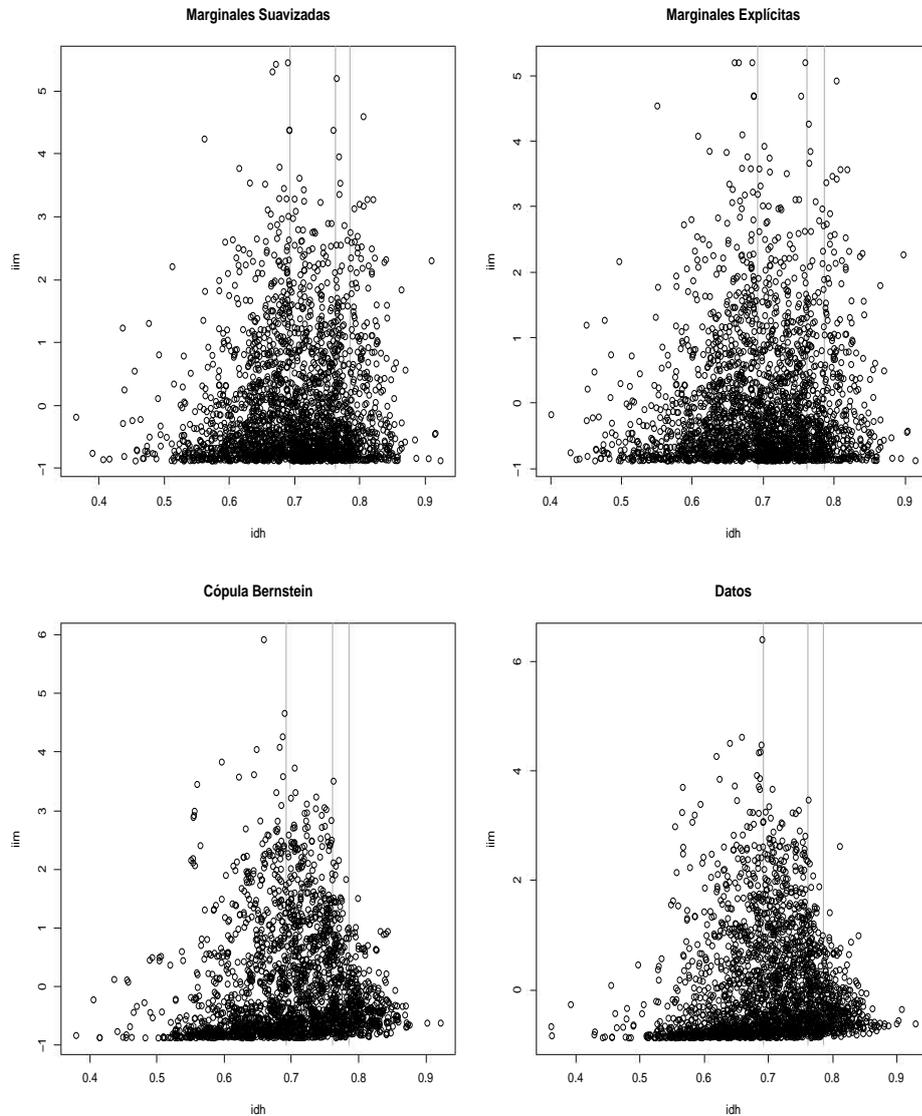


Figura 4.17: Simulaciones conjuntas del vector  $(IDH, IIM)$ .

Cabe destacar que una simulación conjunta con base en la cópula Bernstein (no paramétrica) es, generalmente, más efectiva que una paramétrica, básicamente porque ésta acaba por ser una réplica del fenómeno original. Fundamentalmente, se busca un modelo paramétrico por ser más práctico. En lo que concierne a una simulación no paramétrica con cópulas, ésta requiere mucho más tiempo de cómputo que lo que necesita una simulación paramétrica (varias horas contra un par de minutos).

Con base en la Figura 4.17, se observa que el modelo con la cópula

“Gluing” es adecuado. En la parte que respecta a las submuestras  $\Psi_{con}$ ,  $\Psi_{ind}$  y  $\Psi_{dis1}$  se replicó satisfactoriamente al fenómeno, mientras que en  $\Psi_{dis2}$  el modelo es bueno pero mejorable.

#### 4.4. Regresión con Base en la Cópulas Subyacente y Regresión Lineal Simple

La presente investigación concluye haciendo una breve comparativa de los resultados obtenidos a partir de regresiones con cópulas y los obtenidos con el ya tan conocido modelo de regresión lineal.

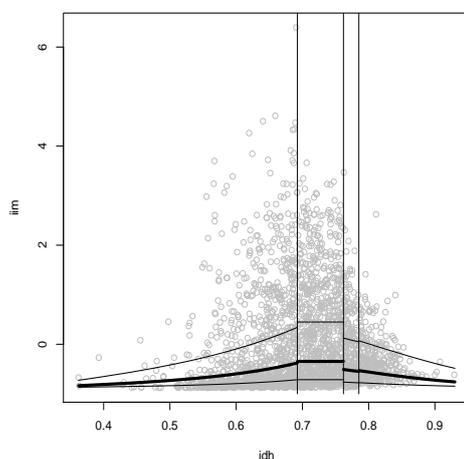


Figura 4.18: Regresiones  $\alpha$ -cuantil paramétrica.

En las Figuras 4.18 y 4.19 se muestran, con líneas gruesas, regresiones medianas y, con líneas delgadas, regresiones al primer y tercer cuartil (percentiles 25 y 75), éstas últimas representando una banda de probabilidad al 50%. Como ha sido a lo largo del presente trabajo, se pretende explicar el cambio en la intensidad migratoria con respecto al desarrollo humano.

En la Figura 4.18 se muestran regresiones  $\alpha$ -cuantil totalmente paramétricas con base en la cópula “Gluing”  $\mathcal{C}_{IDH,IIM}$  (Normal + Independencia + Plackett + Normal), obtenida en la sección 4.3, y las marginales explícitas, obtenidas en la sección 3.4.

La Figura 4.19 muestra: de lado izquierdo, regresión  $\alpha$ -cuantil con base en la cópula y marginales Bernstein (regresión no paramétrica); de lado derecho, regresión  $\alpha$ -cuantil con base en la cópula “Gluing”  $\mathcal{C}_{IDH,IIM}$  (Normal + Independencia + Plackett + Normal) y marginales Bernstein (regresión

semiparamétrica).

A partir de la cópula “Gluing” se obtienen curvas de regresión suaves y discontinuas que logran describir la interacción entre las variables con respecto a un cuantil, las cuales evidentemente llegan a ser más claras y prácticas que las obtenidas con la cópula Bernstein.

Con la cópula “Gluing”, pese a que con las marginales explícitas se logran curvas rectas, la diferencia entre utilizar las marginales Bernstein y las explícitas es mínima.

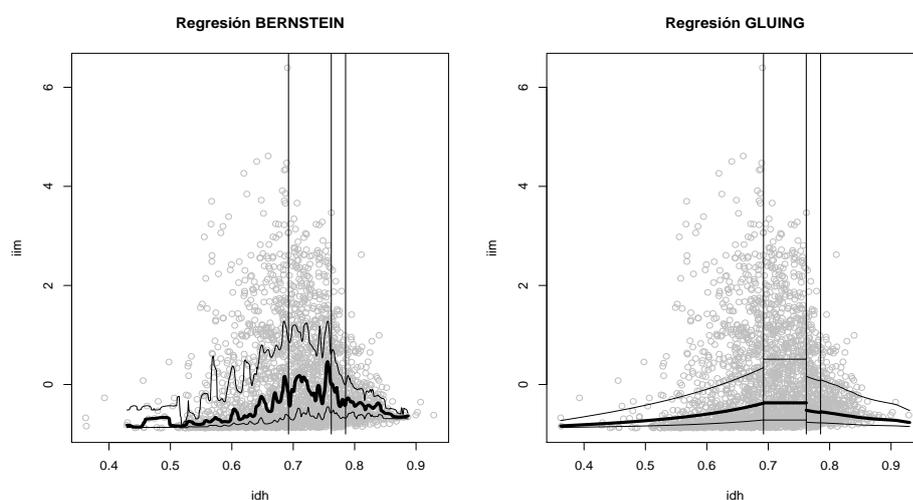


Figura 4.19: Regresiones  $\alpha$ -cuantil: no paramétrica y semiparamétrica

La Figura 4.20 presenta distintos modelos de regresión lineal (regresión lineal simple, polinomios y transformaciones logarítmicas) aplicados a toda la muestra. Se presenta la curva estimada con base en la media y las correspondientes al intervalo de confianza al 95 %.

Se buscó entre varios modelos aquellos que presentaran los mejores resultados. No se encontró algún modelo para el cual no se rechazara la prueba de normalidad a los residuales. No hubo algún modelo con  $R^2 \geq 0.1$ . Básicamente se eligieron los modelos con el mayor valor de  $R^2$  para los cuales se rechazaran todas las pruebas  $t^3$  correspondientes y la prueba  $F^4$ , ambas con un nivel de significancia  $\alpha = 0.05$ .

<sup>3</sup>Dado un modelo de regresión lineal  $Y = \beta_0 + \beta_1 f_1(X) + \beta_2 f_2(X) + \dots + \beta_n f_n(X)$ , la prueba  $t$  contrasta la hipótesis nula  $\mathcal{H}_0 : \beta_i = 0$ , para todo  $i = 0, 1, 2, \dots, n$ .

<sup>4</sup>Dado un modelo de regresión lineal  $Y = \beta_0 + \beta_1 f_1(X) + \beta_2 f_2(X) + \dots + \beta_n f_n(X)$ , la prueba  $F$  contrasta la hipótesis nula  $\mathcal{H}_0 : \beta_1 = \beta_2 = \dots = \beta_n = 0$ .

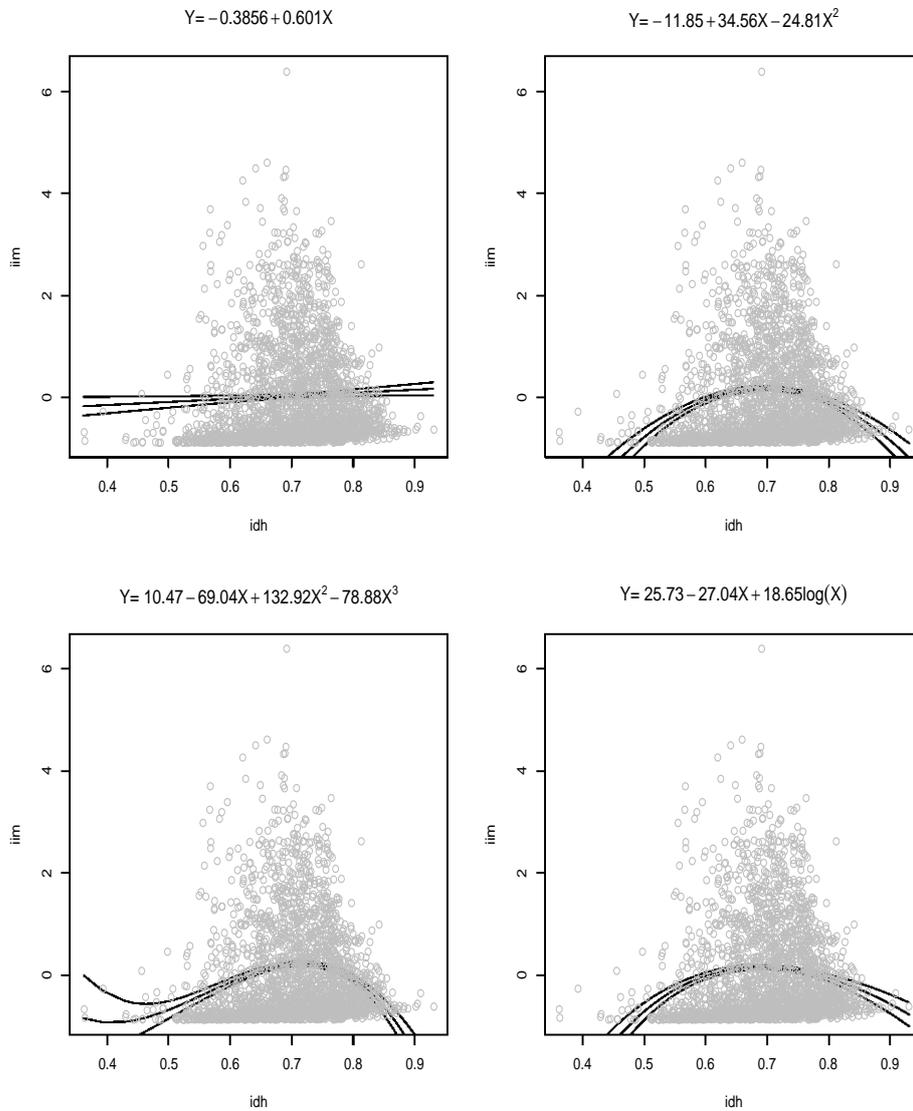


Figura 4.20: Regresiones Lineales.

El Cuadro 4.9 muestra los valores para  $R^2$  ajustada de las regresiones presentadas en la Figura 4.20.

Los mismos modelos de regresión que fueron probados para toda la muestra también se realizaron para cada submuestra. En el Cuadro 4.10 se muestran los mejores modelos para cada submuestra junto con los valores de  $R^2$  correspondientes.

Modelo	$R^2$
$Y = -0.3856 + 0.6010X$	0.001728
$Y = -11.85 + 34.56X - 24.81X^2$	0.04626
$Y = 10.47 - 69.04X + 132.92X^2 - 78.88X^3$	0.05565
$Y = 25.73 - 27.04X + 18.65 \log(X)$	0.03803

Cuadro 4.9:  $R^2$  por modelo de regresión.

Submuestra	Modelo	$R^2$
$\Psi_{con}$	$Y = -11.85 + 34.56X - 24.81X^2$	0.05676
$\Psi_{ind}$	$Y = 0.3775X$	0.06268
$\Psi_{dis1}$	$Y = 19.15 - 24.86X$	0.0469
$\Psi_{dis2}$	$Y = -0.3851X$	0.3225

Cuadro 4.10:  $R^2$  por modelo regresión en cada submuestra.

Finalmente, la Figura 4.21 muestra una regresión discontinua con los mejores modelos para cada submuestra.

A partir de los valores de  $R^2$  presentados en los Cuadros 4.9 y 4.10, es evidente que los modelos de regresión lineal cuentan con una efectividad sumamente baja, además de que los residuales no cumplen con el supuesto de normalidad.

La naturaleza de las regresiones  $\alpha$ -cuantil las hace más específicas desde un punto de vista probabilístico, ya que éstas se basan en la probabilidad condicional del valor de la variable dependiente con respecto de la independiente, esto permite tener un panorama más amplio del cambio que presenta la variable a explicar. Además, las regresiones representan de manera adecuada al fenómeno siempre que se cuente con la cópula subyacente de las variables de estudio, sin la necesidad de que se cumpla con algún supuesto previo o posterior a la implementación del modelo de regresión.

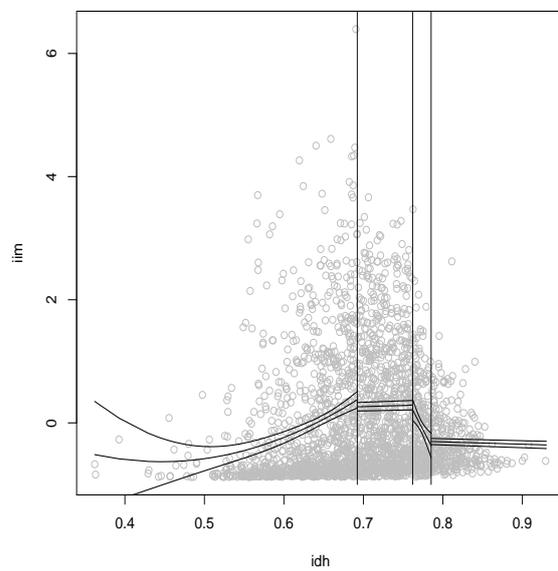


Figura 4.21: Regresión lineal discontinua.



# Conclusiones

La presente tesis comenzó por definir a las funciones cópula bivariadas a fin de exhibir su importancia en la teoría de la probabilidad. Se explicó la relevancia de las cópulas al analizar el comportamiento conjunto entre un par de variables aleatorias, ya que a partir del Teorema de Sklar, las cópulas son funciones que enlazan distribuciones bivariadas con sus marginales univariadas, las cuales por sí mismas son funciones de distribución continuas en el cuadrado unitario, con marginales uniformes (estos conceptos se pueden extender a cualquier dimensión mayor a dos). Se mostró que la cópula de un par de variables aleatorias  $X, Y$  es invariante bajo transformaciones estrictamente crecientes de  $X, Y$ ; y que la cópula subyacente de transformaciones monótonas (crecientes o decrecientes) de  $X$  y/o  $Y$  se puede explicar en función de la cópula de  $X, Y$ .

Al explicar la asociación entre un par de variables aleatorias  $X, Y$ , la cópula subyacente de las mismas  $\mathcal{C}_{X,Y}$  representa la estructura de dependencia que éstas guardan, lo que pone a la cópula  $\mathcal{C}_{X,Y}$  como la fuente única y fundamental para explicar el comportamiento conjunto de estas variables. Es por lo anterior, que para cuantificar el grado de dependencia basta con medir, de algún modo, qué tan lejos/cerca está la cópula subyacente de la cópula producto, siendo esta afirmación la justificación y naturaleza de la construcción de las medidas de dependencia presentadas en el capítulo dos. Es también, que con sólo el conocimiento de la cópula subyacente se logra determinar el grado de concordancia (discordancia) a partir de la medida Rho de Spearman. En contraste, el coeficiente de correlación lineal de Pearson se ve afectado, no sólo por la cópula sino también por las marginales correspondientes, además generalmente, éste es variante bajo transformaciones estrictamente monótonas de las variables, siendo esto último inconsistente con el hecho de que la estructura de dependencia no cambia bajo estas circunstancias. Cabe mencionar que si las variables guardan una relación lineal, el coeficiente de correlación llega a ser útil, pero esta es una relación que difícilmente se da con fenómenos cotidianos.

A lo largo del trabajo se presentaron algunas familias de cópulas paramétricas (Normal, T, Plackett, Farlie-Gumbel-Morgenstern, Cuadras-Augé, entre

otras) de entre las que destacan la clase de cópulas Arquimedianas (Clayton, Ali-Mikhail-Haq, Gumbel, Frank, entre otras) por su fácil construcción y propiedades. También durante el trabajo, se hizo énfasis en el hecho de que la gran mayoría de las cópulas paramétricas son simétricas y representan dependencias estrictamente monótonas. Siendo primordialmente, estas dos las características que se buscaron en la relación conjunta de las variables de estudio, el Índice de Desarrollo Humano (IDH) y el Índice de Intensidad Migratoria (IIM).

Conforme a los objetivos planteados al inicio del presente documento, se buscó hacer una descripción robusta de la relación que guardan las variables de estudio. Para lograr una descripción conjunta sólida se debe entender plenamente el comportamiento marginal, para eso se presenta a continuación una breve reseña del comportamiento marginal de las variables.

Si entendemos al desarrollo humano como el conjunto de oportunidades a las cuales un individuo tiene acceso, con base en el IDH y las variables que lo componen, se observó un desequilibrio en la población mexicana para el año 2000, principalmente en materia de ingreso, lo que significa que las oportunidades en México eran muy distintas dependiendo del municipio donde se localizará el individuo. Pese a que en los municipios imperó el desarrollo humano Medio Alto, existen casos alarmantes con niveles muy bajos, principalmente en los estados de Oaxaca, Guerrero y Veracruz. Los estados en donde se encontraron municipios con mayor desarrollo humano fueron el Distrito Federal y Estado de México.

El IIM presentó valores primordialmente Bajos y Muy Bajos. De entre las variables que componen al IIM, las que presentaron los valores más bajos fueron la proporción de *Hogares con migrantes circulares del quinquenio anterior* y la de *Hogares con migrantes de retorno*. Oaxaca fue el estado con mayor frecuencia de municipios con intensidad migratoria nula, pero también fue el estado con los municipios que reportaron mayor valor para el IIM.

Cabe destacar que Oaxaca es un estado muy plural y con municipios muy distintos, además para el año 2000, era el estado con mayor número de municipios, 570 en total.

Una vez entendido el comportamiento marginal y previo a hacer el modelo y el análisis con cópulas se logró dar una explicación de la relación conjunta con base en los niveles de desarrollo humano y de intensidad migratoria donde observamos que, municipios con un nivel Bajo de desarrollo humano presentan niveles Nulos o Bajos de intensidad migratoria, lo que significa que estos tienen opciones tan limitadas de vida que el migrar no es una alternativa disponible o se recurre poco a ésta; los municipios con un nivel

Medio Bajo o Medio Alto de desarrollo humano presentan más diversidad en sus fenómenos migratorios, imperan los niveles Muy Bajo y Bajo de intensidad migratoria, pero se observa a más municipios en donde la migración se convierte en la opción de vida para gran parte de su población; para los municipios con desarrollo humano Alto la migración deja de ser una opción de vida para la mayoría de sus habitantes, aun cuando se tenga acceso a ésta, al tener una diversidad más amplia de alternativas en su lugar de origen, deciden desarrollarse en éste.

Nótese que la anterior descripción acaba por ser simplemente más rebuscada que la que se propone en PNUD (2006) y que se presentó en la introducción de la presente investigación, pero siguen quedando ambiguos algunos elementos de la relación conjunta. Es por lo anterior, que buscar describir la relación a partir la distribución conjunta de las variables y más aun a partir la cópula subyacente de las variables será la vía para llevarnos a una descripción más precisa y acertada.

Se rechazó la prueba de independencia entre las variables IDH, IIM y con base en la medida  $\sigma$  de Schweizer & Wolff se observó que las variables tienen un grado de dependencia del 23.7%, pero las variables no guardan una dependencia estrictamente monótona, ya que la medida  $\rho$  de Spearman presentó un valor menor al de  $\sigma$ , con una diferencia de casi 0.1, este hecho se reiteró con la gráfica de la diagonal de la cópula empírica donde se observó que la primer parte de la muestra tiende hacia la concordancia, una parte considerable pasa muy cerca de la independencia y el resto de muestra tiende hacia la discordancia. Además se rechazó la simetría en la cópula subyacente de toda la muestra. A toda la muestra se le realizaron pruebas de bondad de ajuste para un catálogo amplio de cópulas, las cuales fueron rechazadas.

Como ya se comentó, las variables IDH, IIM no cumplen con guardar una dependencia estrictamente monótona y la cópula subyacente de éstas no es simétrica. Entonces, con la finalidad de obtener un ajuste paramétrico conjunto se propuso un algoritmo de segmentación denominado “Algoritmo de Segmentación en Dependencias Estrictamente Monótonas (ASDEM)”, el cual tiene como objetivo segmentar a la muestra con base en una variable explicatoria a fin de lograr submuestras con dependencias estrictamente monótonas y que además provengan de cópulas simétricas. Para los fines de la presente investigación se tomó al IDH como la variable que explica al IIM. Una vez obtenidas las submuestras y un ajuste paramétrico conjunto para cada una, se juntaron las cópulas provenientes de cada submuestra en una sola que representará a toda la muestra con el método “Gluing Copulas”.

Finalmente con el algoritmo propuesto, se logró segmentar la muestra en

cuatro submuestras y además se logró describir la relación conjunta a partir de las submuestras como sigue:

- $\Psi_{con}$ , submuestra donde impera la concordancia con un valor  $\hat{\rho} = 0.3$ , lo que significa que el IIM crece en proporción de uno a tres con respecto al crecimiento del IDH cuando este último presenta valores menores a 0.69230, en este caso el IIM puede tomar cualquiera de sus valores, esta submuestra incluye a todos los municipios con niveles de desarrollo humano Bajo y Medio Bajo y unos cuantos del nivel Medio Alto, esta submuestra es el 30.7 % de la muestra total;
- $\Psi_{ind}$ , submuestra en donde las variables IDH, IIM son independientes, todos los municipios en ésta tienen un nivel de desarrollo humano Medio Alto, básicamente lo que ocurre es que a medida que el IDH incrementa en el intervalo que va desde 0.69232 y hasta 0.76193 el IIM puede tomar cualquier valor desde -0.87348 y hasta 3.6633, esta submuestra es el 34 % de la muestra total;
- $\Psi_{dis1}$ , submuestra donde impera la discordancia con un valor  $\hat{\rho} = -0.225$ , lo que implica que en alrededor del 22.5 % de los casos en los que el IDH toma valores desde 0.76194 y hasta 0.78523 el IIM disminuye a medida que el IDH aumenta, nuevamente todos los municipios en esta submuestra presentan el nivel de desarrollo humano Medio Alto y presentan valores de IIM entre -0.87704 y 3.4671, esta submuestra es el 9.9 % de la muestra total;
- $\Psi_{dis2}$ , submuestra donde impera la discordancia con un valor  $\hat{\rho} = -0.238$ , lo que implica que en alrededor del 23.8 % de los casos en los que el IDH es mayor a 0.78536 el IIM disminuye a medida que el IDH aumenta, en esta submuestra se incluyen unos cuantos casos con nivel de desarrollo humano Medio Alto y el resto son todos aquellos con nivel Alto, para este caso el IIM reportó valores entre -0.84980 y 2.6199, esta submuestra es el 12.7 % de la muestra total.

Nótese que la descripción anterior es mucho más robusta que las antes enunciadas y da mucha más claridad e información detallada sobre la interacción de las variables de análisis, además de que determina los puntos críticos en los que la dependencia cambia de sentido.

La submuestra  $\Psi_{ind}$ , por ser la parte inestable de la relación conjunta, toma mucha importancia, ya que viene a ser la manera en la que la relación cambia de sentido (de la concordancia a la discordancia), lo cual puede llegar a ser una necesidad en la relación entre algunos fenómenos cotidianos.

Cabe destacar, que la propuesta que se hace de segmentación es únicamente basada en el comportamiento general de los índices y que tal vez existan factores que segmenten la muestra de una manera más natural, como por ejemplo la geografía de los municipios. Pero la propuesta que se da, cumple con los objetivos planteados al principio del documento, ya que a partir de la segmentación empleada se logra una descripción precisa y detallada de la relación conjunta, además de que se determinan submuestras con dependencias estables que finalmente pudieron ser modeladas con cópulas conocidas.

De igual forma, el “Algoritmo de Segmentación en Dependencias Estrictamente Monótonas” es simplemente una propuesta que se da con la finalidad particular de cumplir con los objetivos del trabajo, incluso cuando puede llegar a ser útil como auxiliar para hacer ajustes bivariados cuando las variables no guarden una dependencia estrictamente monótona, aún se requieren validar muchos factores, como por ejemplo la convergencia de los estadísticos empleados, así como hacer más pruebas sobre los escenarios sobre las cuales tiene un mejor desempeño y de esa forma delimitar su correcta aplicación.

Para validar la efectividad de la segmentación y el posterior ajuste, se generaron simulaciones conjuntas del vector  $(IDH, IIM)$  con base en la cópula “Gluing”, las cuales representaron de manera adecuada al fenómeno.

Para tener un panorama más práctico del modelo conjunto se presentaron regresiones a los percentiles 25, 50 y 75 con base en la cópula “Gluing”, mismas que fueron comparadas con regresiones lineales. Cabe destacar que ni la estimación puntual ni por intervalos de una regresión lineal son totalmente comparables con la estimación por alguno de los cuantiles propuestos, más bien se comparó el modelo completo de regresión lineal contra el de regresión con cópulas. Se lograron modelos de regresión lineal que asemejaban al comportamiento de las variables, pero estos presentaban niveles muy bajos de  $R^2$ , ya que las distribuciones reales para IIM, que se presentaban para un valor fijo de IDH, generalmente se encontraban alejadas de la media, además de que no se cumplían con los supuestos propios del modelo. En cambio, el modelo de regresión con cópulas presentó un panorama mucho más amplio del comportamiento del IIM condicional a un valor fijo de IDH con respecto a un cuantil específico, en ese sentido la interacción de regresiones con cópulas a distintos cuantiles generan bandas de probabilidad para el valor de IIM con respecto al cambio en el IDH, que son conceptualmente más informativas que una banda de confianza en una regresión lineal.

La teoría de cópulas, como rama de la probabilidad y la estadística, ha estado en crecimiento en los últimos años. Como se evidenció a lo largo de la

presente tesis, el hacer ajustes con cópulas paramétricas, puede no ser una tarea trivial. Los métodos actuales de construcción de cópulas son cada vez más complejos y logran cópulas paramétricas que no precisamente guardan dependencias estrictamente monótonas o simetría, en ese sentido el homologar técnicas para ajustes paramétricos que logren representar dependencias complejas sigue siendo un tema abierto.

## Apéndice A

# Algoritmos en R

### A.1. Análisis Descriptivo Conjunto

#### Cálculo de la cópula empírica: `cee`

De la función `cee` resulta una matriz de  $(n + 1) \times (n + 1)$ , la cópula empírica correspondiente a los datos. Véase Definición 1.11.1.

*datos*: una matriz de  $n \times 2$  (más adelante será llamada simplemente matriz de datos), la información correspondiente de la muestra.

```
cee<-function(datos)
{
  F<-function(i,k)
  {
    n=length(k)
    datos=rep(0,(n+1))
    datos[k[1:i]]=rep(1/n,i)
    datos=cumsum(datos)
  }
  n=length(datos[,1])
  copula<-matrix(nrow=(n+1),ncol=(n+1))
  ii<-order(datos[,2])
  h<-rbind(datos[,1],datos[,2])[,ii]
  k<-order(h[1,])+1
  r=k
  copula[1,]=rep(0,(n+1))
  copula[2:(n+1),]=t(sapply(c(1:n),F,k=r))
  copula
}
```

### Diagonal de la cópula empírica: diagonal

Gráfica comparativa de la diagonal de la cópula empírica y las diagonales de  $\mathcal{M}$ ,  $\Pi$  y  $\mathcal{W}$  (véase Figura 1.2), y cálculo de las diferencias entre la diagonal de la cópula empírica y la diagonal de la cópula  $\Pi$ . Al igualar la función **diagonal** a una variable, además de la gráfica, resulta un vector de  $n \times 1$  con las diferencias antes mencionadas. Función auxiliar: **cee**.

*datos*: matriz de datos.

*nombre*: nombre del gráfico (en algunas gráficas se empleará esta variable, ésta siempre requiere un cadena entre comillas "" o una expresión matemática).

```
diagonal<-function(datos,nombre)
{
  copula=cee(datos)
  d=dim(copula)
  diag=vector(mode="numeric",length=d[1])
  for(i in 1:d[1])
  {
    diag[i]=copula[i,i]
  }
  Md<-function(u) u
  Pd<-function(u) u^2
  Wd<-function(u) max(2*u-1,0)
  w=sapply(seq(0,1,.01),Wd)
  p=sapply(seq(0,1,1/(d[1]-1)),Pd)
  m=sapply(seq(0,1,1/(d[1]-1)),Md)
  plot(seq(0,1,.01),w,type="l",col="red",main=nombre)
  lines(seq(0,1,1/(d[1]-1)),p,col="blue")
  lines(seq(0,1,1/(d[1]-1)),m,col="green")
  lines(seq(0,1,1/(d[1]-1)),diag,col=30)
  df=diag[-c(1,d[1])]-p[-c(1,d[1])]
}
```

### Contornos de la cópula empírica: contour

La función **contour** es una función propia de R, para visualizar los contornos de una cópula empírica (véase Figura 1.1). Durante el trabajo, en general, esta gráfica se acompaña de una línea transversal para dar una idea de la simetría de la cópula. Función auxiliar: **cee**

*cee(datos)*: basta con ingresar a ésta la matriz de la cópula empírica calculada con la función **cee**.

```
contour(cee(datos))
lines(c(0,1),c(0,1),col="red")
```

### Pseudo-observaciones de la cópula: pseudo

Produce la gráfica de las pseudo-observaciones de la cópula con base en la muestra y cuando ésta se iguala a una variable se obtiene también la muestra de las mismas. Véase Definición 1.11.3.

*datos*: matriz de datos.

*nombre*: nombre del gráfico.

```
pseudo=function(datos,nombre)
{
  APs1<-ecdf(datos[,1])
  APs2<-ecdf(datos[,2])
  APs=cbind(APs1(datos[,1]),APs2(datos[,2]))
  plot(APs,xlab="u",ylab="v",main=nombre)
  lines(c(0,1),c(0,1),col="red")
  AP=APs
}
```

### Suavizamiento univariado vía polinomios de Bernstein: Q

La función **Q** es el suavizamiento correspondiente a la función de cuantiles como en (1.11.2). Resulta el cuantil  $u$  con base en el suavizamiento.

$u$ : un número  $u$  que pertenece a  $\mathcal{I}$ , el cuantil buscado.

*datos*: una matriz de datos.

```
Q<-function(u,datos)
{
  n=length(datos)
  q=sort(datos)
  eo<-c(q[1],q,q[n])
  sum=0
  for(i in 0:n)
  {
    sum=sum+(eo[i+1]+eo[i+2])/2*dbinom(i,n,u)
  }
  sum
}
```

### Simulación bivariada conjunta: `rcopN`

La función `rcopN` obtiene simulaciones conjuntas de un vector aleatorio con base en la cópula Bernstein de las variables, resulta una matriz de  $N \times 2$  de datos simulados. Véase Algoritmo 1.9.2.

Las función `pui` es auxiliar para ésta y otras funciones más adelante expuestas, aproxima la parcial en  $u$  de la cópula Bernstein. Otra función auxiliar: `cee`.

```

pui<-function(t,uu,ve,cop)
{
fauxu<-function(i,u,n)
{
  a=0
  a=dbinom(i,n,u)*(i-u*n)/(u*(1-u))
  a
}
h<-function(a,b)
{
c=a*b
c
}
  m=length(cop[,1])-1
  sv<- sapply(c(0:m),dbinom,size=m,prob=ve)
  su<- sapply(c(0:m),fauxu,u=uu,n=m)
  pui=sapply(sv,h,b=su)*cop
  sum(pui)-t
}

```

La función `rcopN` tiene por argumentos los siguientes:

*datos*: una matriz de datos.

*N*: el número de simulaciones que se desean.

```

rcopN<-function(datos,N)
{
rcop<-function(i,vec)
{
cp=cee(vec)
simf<-rep(0,2)
uu=runif(1)
y=uniroot(pui,c(0,1),tol=.0001,t=runif(1),u=uu,cop=cp)$root
simf=c(Q(uu,vec[,1]),Q(y,vec[,2]))
}
}

```

```

simf
}
t(sapply(c(1:N),rcop,vec=datos))
}

```

### Regresiones $\alpha$ -cuantil: (1) `rcop_alfa`, (2) `rcop_reg`, (3) `reg_bandas`

Las siguientes funciones obtienen distintas regresiones  $\alpha$ -cuantil con base en la cópula Bernstein, véase (1.11.3).

La función **Qi** es auxiliar para ésta y otra funciones, aproxima la inversa de la función **Q** (Suavizamiento univariado vía polinomios de Bernstein-Kantorovic). Otras funciones auxiliares: **cee** y **pui**.

```

Qi<-function(u,x,vec)
{
  n=length(vec)
  q=sort(vec)
  eo<-c(q[1],q,q[n])
  sum=0
  for(i in 0:n)
  {
    sum=sum+(eo[i+1]+eo[i+2])/2*dbinom(i,n,u)
  }
  sum-x
}

```

(1) Dada una muestra de un par de variables aleatorias continuas  $X$ ,  $Y$ , la función `rcop_alfa` aproxima el valor para la variable  $Y$  con base en la regresión  $\alpha$ -cuantil, para un valor de  $X$  y un  $\alpha$  dados. Véase Algoritmo 1.12.1.

*pred*: el valor  $x$  de la variable  $X$ .

*datos*: una matriz de datos.

*alpha*: el valor de  $\alpha$ .

```

rcop_alfa<-function(pred,datos,alpha)
{
  cp=cee(datos)
  simf<-rep(0,2)
  uu=uniroot(Qi,c(0,1),tol=.01,x=pred,vec=datos[,1])$root
  v=uniroot(pui,c(0,1),tol=.0001,t=alpha,u=uu,cop=cp)$root
  simf=Q(v,datos[,2])
  simf
}

```

}

(2) La función **rcop\_reg** funciona exactamente igual que la anterior, con la salvedad de que sólo calcula la regresión mediana, en ese sentido, sólo requiere de un valor para  $X$  y de la matriz de datos. Véase Definición 1.12.1.

*pred*: el valor  $x$  de la variable  $X$ .

*datos*: una matriz de datos.

```
rcop_reg<-function(pred,datos)
{
  cp=cee(datos)
  simf<-rep(0,2)
  uu=uniroot(Qi,c(0,1),tol=.0001,x=pred,vec=datos[,1])$root
  v=uniroot(pui,c(0,1),tol=.001,t=.5,u=uu,cop=cp)$root
  simf=Q(v,datos[,2])
  simf
}
```

(3) De la función **reg\_bandas** resulta un gráfico de dispersión de un par de variables junto con sus respectivas regresiones a los tres primeros cuartiles (percentiles 25, 50 y 75). Funciones auxiliares: **rcop\_reg** y **rcop\_alfa**.

*datos*: matriz de datos, con la variable predictora en la primer columna..

*nombre*: nombre del gráfico.

```
reg_bandas<-function(datos,nombre)
{
  copula=cee(datos)
  a=sort(datos[,1])
  n=length(datos[,1])
  x=seq(a[3],a[n-3],.001)
  reg=sapply(x,rcop_reg,data=datos,cp=copula)
  reg_25=sapply(x,rcop_alfa,data=datos,cp=copula,alfa=.25)
  reg_75=sapply(x,rcop_alfa,data=datos,cp=copula,alfa=.75)
  plot(datos,main=nombre,col="gray")
  lines(x,reg,lwd=3)
  lines(x,reg_25)
  lines(x,reg_75)
}
```

### Prueba de Simetría: `exchTest`

La función `exchTest` está contenida en el paquete “copula”, calcula el estadístico de prueba y el p-valor asociado para la prueba de simetría de [Genest y otros \(2011\)](#).

*datos*: matriz de datos.

```
exchTest(datos)
```

### Cóputas incluidas en el paquete “cópula”

Estas funciones generan objetos cópula, básicamente sirven como insumo para otras funciones que se presentaran más adelante. Los nombres de las funciones generalmente inician con el nombre de la cópula, en caso contrario se menciona a cual hace referencia. Los argumentos de estas funciones son los siguientes:

*theta*: parámetro particular de la cópula.

*dim*: dimensión de la cópula (dos por default).

*r*: el coeficiente de correlación de Pearson como en (2.1.1), empleado en la cópula Normal.

*rho*:  $\rho$  de Spearman como en (2.3.2), empleado en la cópula T.

*df*: son los grados de libertad de la cópula T (cuatro por default).

A continuación se enuncian las funciones que se emplearon en el presente trabajo:

- `normalCopula(r, dim=2)`

Véase (1.3.3).

- `tCopula(rho, dim = 2, df = 4)`

Véase [Demarta y McNeil \(2005\)](#).

- `plackettCopula(theta)`

Véase (1.3.4).

- `fgmCopula(theta, dim = 2)`

Cópula Farlie-Gumbel-Morgenstern. Véase (1.3.5).

- `claytonCopula(theta, dim = 2)`  
Véase (1.10.3).
- `frankCopula(theta, dim = 2)`  
Véase (1.10.6).
- `gumbelCopula(theta, dim = 2)`  
Véase (1.10.5).
- `amhCopula(theta, dim = 2)`  
Cópula Ali-Mikhail-Haq. Véase (1.10.4).
- `galambosCopula(theta)`
- `huslerReissCopula(param)`

## A.2. Análisis de la Dependencia

### Prueba de Independencia: `indepTestSim` y `indepTest`

Las funciones `indepTestSim` y `indepTest` están contenidas en el paquete “copula”. La prueba de independencia se da en dos pasos:

Paso 1: Simulaciones para el estadístico de prueba:

*n*: tamaño de la muestra.

*p*: dimensión del vector aleatorio (dos para efectos de este trabajo).

*N*: número de simulaciones del estadístico de prueba (1000 por default).

*print.every*: va reportando el número de simulaciones (100 por default).

```
paso1 <-indepTestSim(n, p, N = 1000, print.every = 100)
```

Paso 2: Prueba de hipótesis  $\mathcal{H}_0 : H(x, y) = F(x)G(y)$ :

*x*: matriz de datos.

*paso1*: el resultado del Paso 1.

*alpha*: nivel de la prueba (0.05 por default).

```
paso2 <-indepTest(x, paso1, alpha=0.05)
```

### Medidas de Asociación empíricas: (1) **cor**, (2) **s.w.e**, (3) **hoeffding.e**

(1) La función **cor** es propia de R y se presentan de la siguiente manera:

*datos*: matriz de datos.

*method*: hace referencia a las medidas empíricas  $\hat{\rho}$  de Spearman como en (2.3.3), y el coeficiente de correlación de Pearson como en (2.1.1).

```
cor(datos,method=c('spearman','pearson'))
```

(2) La función **s.w.e** da como resultado el cómputo de la medida empírica  $\hat{\sigma}$  de Schweizer & Wolff como en (2.4.5).

Tanto para **s.w.e** como para **hoeffding.e**, se utiliza a la función **cargam3** como auxiliar para el cómputo de una muestra de la cópula II. Otra función auxiliar: **cee**.

```
cargam3<-function(i,n)
{
r=floor(i/n)+1
c=i-(r-1)*n+1
r*c/n^2
}
```

El argumento de la función **s.w.e** es el siguiente:

*datos*: matriz de datos.

```
s.w.e<-function(datos)
{
cp=cee(datos)
n=length(cp[,1])-1
12*sum(abs(cp[2:(n+1),2:(n+1)]-
matrix(sapply(0:(n*n-1),cargam3,n),ncol=n)))/(n^2-1)
}
```

(3) La función **hoeffding.e** tiene como resultado el cómputo del cuadrado del índice de dependencia de Hoeffding como en (2.4.6).

*datos*: matriz de datos.

```
hoeffding.e<-function(datos)
{
cp=cee(datos)
```

```
n=length(cp[,1])-1
(90*sum((cp[2:(n+1),2:(n+1)]-
matrix(sapply(0:(n*n-1),cargam3,n),ncol=n))^2))^(1/2)/(n-1)
}
```

### Concordancia y Dependencia de Cópulas Paramétricas: (1) s.w.copula, (2) spearmansRho

(1) Cálculo de  $\sigma$  de Schweizer & Wolff como en (2.4.3), para una cópula paramétrica. Nótese que se hace uso de la función **pcopula**, ésta está incluida en el paquete “cópula” y sirve para el cálculo de la distribución de una cópula dada.

*copula*: objeto cópula.

```
s.w.copula=function(copula)
{
CuV=function(v,u) abs(pcopula(copula,c(u,v)) - u*v)
intv=function(x) integrate(Vectorize(CuV),u=x,lower=0.00001,
upper=1)$value
intu=function() integrate(Vectorize(intv),0.0000001,1)$value
12*intu()
}
```

(2) **spearmansRho** es una función incluida en el paquete “copula” de R, calcula la medida de concordancia  $\rho$  de Spearman como en (2.3.2).

*copula(theta)*: objeto cópula.

*copula*: es una función de la lista en **Cópulas incluidas en el paquete “cópula”**, sección A.1.

```
spearmansRho(copula(theta))
```

## A.3. Modelado Conjunto

### Segmentación: (1) inicio, (2) final

A partir del procedimiento de la Subsección 4.2 surgen las funciones **inicio** y **final**, de las cuales resultan vectores con las diferencias entre las medidas empíricas  $\hat{\sigma}$  de Schweizer & Wolff y  $\hat{\rho}$  de Spearman, Ecuaciones (2.4.5) y (2.3.3), con base en el Algoritmo 4.2.1. Funciones Auxiliares: **s.w.e** y **cor**. Estas funciones tienen por argumentos:

*datos*: matriz de datos.

*posicion*: posición de la columna en la matriz en la que se encuentra la variable explicatoria, por la que se pretende hacer la segmentación.

(1) **inicio**:

```

inicio=function(datos,posicion)
{
  evalua_inicio<-function(i,datos,posicion)
  {
    datos=datos[order(datos[,posicion]),]
    muestra=datos[1:i,]
    diff=s.w.e(muestra)-abs(cor(muestra,method='spearman'))
    diff
  }

  n=dim(datos)[1]
  n.min=round(n*.05)
  diferencias=sapply(n.min:n,evalua_inicio,datos,pos)
  diferencias
}

```

(2) **final**:

```

final=function(datos,posicion)
{
  evalua_final<-function(i,datos,posicion)
  {
    M=dim(datos)[1]
    datos=datos[order(datos[,posicion]),]
    muestra=datos[i:M,]
    diff=s.w.e(muestra)-abs(cor(muestra,method='spearman'))
    diff
  }

  n=dim(datos)[1]
  n.min=n-round(n*.05)
  diferencias=sapply(n.min:1,evalua_final,datos,pos)
  diferencias
}

```

### Bondad de Ajuste para Cópulas: `gofCopula`

Esta función viene contenida dentro del paquete “cópula” de R. De ésta resulta la estimación del parámetro ( $s$ ) de la cópula evaluada, el estadístico

de prueba y el p-valor asociado.

*copula(parametro)*: objeto cópula, con un parámetro arbitrario.

*datos*: la matriz de datos a la que se le pretende hacer el ajuste.

`gofCopula(copula(parametro),datos)`

### Simulación con la cópula “Gluing”: `Gluing.Function`

De la función **Gluing.Function** resulta un objeto tipo lista con tres simulaciones del comportamiento del vector aleatorio (IDH,IIM) con base en la cópula “Gluing” obtenida en la sección 4.3, véase Algoritmo 1.9.3: la primera, es una simulación de las pseudo-observaciones de la cópula; la segunda, una simulación semiparamétrica del vector (IDH,IIM), con la cópula “Gluing” y marginales Bernstein; la tercera, es una simulación totalmente paramétrica, con base en la cópula “Gluing” y las marginales obtenidas en la sección 3.4.

Esta función está totalmente encaminada a hacer simulaciones del vector aleatorio (IDH,IIM) con base en la segmentación obtenida en la sección 4.2, los parámetros de ésta son los siguientes:

*cop\_con*: objeto cópula para representar la parte de la submuestra  $\Psi_{con}$ .

*cop\_dis1*: objeto cópula para representar la parte de la submuestra  $\Psi_{dis1}$ .

*cop\_dis2*: objeto cópula para representar la parte de la submuestra  $\Psi_{dis2}$ .

*NM*: tamaño de muestra para la simulación.

*datos*: matriz de datos con las observaciones del vector (IDH,IIM).

Los parámetros *cop\_con*, *cop\_dis1* y *cop\_dis2* pueden ser tomados de entre las cópulas que no hayan rechazado la bondad de ajuste por submuestra, Cuadro 4.8.

Nótese que se hace uso de **rcopula**, función incluida en el paquete “copula”, sirve para hacer simulaciones de cópulas específicas. Función auxiliar: **Q**.

```
Gluing.Function=function(cop_con,cop_dis1,cop_dis2,NM,datos)
{
  salida_final=datos
  N=2350
```



```

F.IIM.u=function(y,u) integrate(d.IIM,theta=0.41507,mu= -1.231,
                                sigma=2.0323,lower=-0.87874,upper=y)$value-u

F.IDH.u=function(q,u) pweibull(q-0.3,shape=5.793,scale=0.427)-u
F.IDH.inv=function(u) uniroot(F.IDH.u,c(0,1),tol=1e-04,u=u)$root
F.IIM.inv=function(u)
{
  if(u>0.999) u=.999
  uniroot(F.IIM.u, c(-0.87874,100), tol = 1e-04, u=u)$root
}
simulacion.exp=cbind(sapply(Rcop[,1],F.IDH.inv),
                    sapply(Rcop[,2],F.IIM.inv))

list(Rcop,simulacion.emp,simulacion.exp)
}

```

### Regresión $\alpha$ -cuantil semiparamétrica: **Regresion.C.IDH.IIM**

La función **Regresion.C.IDH.IIM** obtiene una estimación del IIM dado un valor de IDH respecto de un cuantil  $\alpha$  y con base en la cópula “Gluing” (Normal + Independencia + Plackett + Normal) y las marginales Bernstein, véase 1.12. Funciones auxiliares: **Q** y **Qi**.

*x*: valor para IDH.

*cuan*: el nivel  $\alpha$ , un valor en el intervalo (0, 1).

*datos*: matriz de datos con las observaciones del vector (IDH,IIM).

```

Regresion.C.IDH.IIM=function(x,cuan,datos)
{
  regresion.parametrica=function(obs,copula,cuantil)
  {
    u=uniroot(Qi,c(0,1),tol=.0001,x=obs,vec=datos[,1])$root
    buscav=function(v) dcopula(copula,c(u,v))
    intbusca=function(bcopula,v,alfa) integrate(Vectorize(bcopula),
                                                0.00000001,v)$value-alfa
    v=uniroot(intbusca,c(0,1),tol=.0001,bcopula=buscav,
              alfa=cuantil)$root
    Q(v,datos[,2])
  }

  cop.con=normalCopula(0.3234996)
  cop.dis1=plackettCopula(0.5351376)

```

```

cop.dis2=normalCopula(-0.2248809)

if(x>=0 & x < 0.3619992+.0001)
y=regresion.parametrica(0.3619992+.0001,cop.con,cuan)

if(x >= 0.3619992+.0001 & x <= 0.6923088)
y=regresion.parametrica(x,cop.con,cuan)

if(x > 0.6923088 & x <= 0.7619343) y=Q(cuan,datos[,2])

if(x > 0.7619343 & x <= 0.7852387)
y=regresion.parametrica(x,cop.dis1,cuan)

if(x > 0.7852387) & x <= 0.9301854 -.001)
y=regresion.parametrica(x,cop.dis2,cuan)

if(x>0.9301854-.001 & x<=1)
y=regresion.parametrica(max(dis2[,1])-.001,cop.dis2,cuan)

y
}

```

### Regresión $\alpha$ -cuantil paramétrica: **Regresion.C.IDH.IIM.exp**

La función **Regresion.C.IDH.IIM.exp** obtiene una estimación del IIM dado un valor de IDH respecto de un cuantil  $\alpha$  y con base en la cópula “Gluing” (Normal + Independencia + Plackett + Normal) y las marginales obtenidas en la sección 3.4, véase 1.12.

$x$ : valor para IDH.

$cuan$ : el nivel  $\alpha$ , un valor en el intervalo  $(0, 1)$ .

```

Regresion.C.IDH.IIM.exp=function(x,cuan)
{
d.IIM=function(x,theta,mu,sigma) exp(x)*dlnorm(exp(x)-theta,
meanlog = mu, sdlog = sigma)

F.IIM.u=function(y,u) integrate(d.IIM,theta=0.415,mu=-1.231,
sigma=2.0323,lower=-0.87874,upper=y)$value-u

F.IDH.u=function(q) pweibull(q-0.3,shape=5.793495,
scale = 0.427913)

F.IIM.inv=function(u)

```

```

{
if(u>0.999) u=.999
uniroot(F.IIM.u, c(-0.87874,100), tol = 1e-04, u=u)$root
}
regresion.parametrica=function(obs,copula,cuantil)
{
u=F.IDH.u(obs)
buscav=function(v) dcopula(copula,c(u,v))
intbusca=function(bcopula,v,alfa) integrate(Vectorize(bcopula),
0.00000001,v)$value-alfa
v=uniroot(intbusca,c(0,1),tol=.0001,bcopula=buscav,
alfa=cuantil)$root
F.IIM.inv(v)
}
cop.con=normalCopula(0.3234996)
cop.dis1=plackettCopula(0.5351376)
cop.dis2=normalCopula(-0.2248809)

if(x>=0 & x < 0.3619992+.0001)
y=regresion.parametrica(0.3619992+.0001,cop.con,cuan)

if(x >= 0.3619992+.0001 & x <= 0.6923088)
y=regresion.parametrica(x,cop.con,cuan)

if(x > 0.6923088 & x <= 0.7619343) y=F.IIM.inv(cuan)

if(x > 0.7619343 & x <= 0.7852387)
y=regresion.parametrica(x,cop.dis1,cuan)

if(x > 0.7852387) & x <= 0.9301854 -.001)
y=regresion.parametrica(x,cop.dis2,cuan)

if(x>0.9301854-.001 & x<=1)
y=regresion.parametrica(max(dis2[,1])-.001,cop.dis2,cuan)

y
}

```

### Regresión lineal: (1) `lm`, (2) `predict`

(1) La función `lm` es propia de R. para los fines del presente trabajo se hizo uso de los siguientes resultados de la función `lm`: coeficientes de la curva de regresión, estadísticos de resumen de los residuales, pruebas  $t$  por coeficiente, prueba  $F$ ,  $R^2$  y  $R^2$  ajustada. Ésta tiene el siguiente parámetro:

*formula*: un objeto del tipo formula. Ejemplos:

- $Y \sim X$   
para el modelo  $Y = \beta_0 + \beta_1 * X$ ;
- $Y \sim -1+X$   
para el modelo  $Y = \beta_1 * X$ ;
- $Y \sim X+I(X^2)$   
para el modelo  $Y = \beta_0 + \beta_1 * X + \beta_2 * X^2$ .

donde  $X$  y  $Y$  son vectores con los datos correspondientes.

`lm(formlula)`

(2) Con la función **predict** se obtienen la estimación puntual y por intervalos aplicados a todos los valores de la variable predictora en un modelo de regresión lineal. Esta función tiene los siguientes parámetros:

*lm(formlula)*: un modelo de regresión lineal, obtenido con la función anterior.

*level*: el nivel del intervalo de confianza deseado, para los fines de este trabajo se utilizaron intervalos de confianza al 95 %. Se espera un número en el intervalo (0, 1).

`predict(lm(formlula),level = 0.95)`

## A.4. Tópicos Adicionales

### Gráficos Boxplot: boxplot

Recordemos que un gráfico boxplot tiene las siguientes características: los límites en la “caja” principal son el primer cuartil (percentil 25) y el tercer cuartil (percentil 75); la recta dentro de la “caja” representa la posición de la mediana (percentil 50); el rango intercuartil es la diferencia entre el tercer y el primer cuartil; el umbral superior se encuentran a 1.5 por el rango intercuartil mas el tercer cuartil y el umbral inferior es el primer cuartil menos 1.5 por el rango intercuartil (configuración habitual), el valor de 1.5 puede ser modificado en caso de ser necesario; dichos umbrales nos ayudan a definir una regla de decisión para la detección de valores atípicos, ya que en el gráfico se representan con “puntos” a los valores que sobrepasan dichos umbrales para ser considerados como atípicos. Cuando un umbral aparenta

verse más pequeño que el otro significa que, en el caso del inferior, el umbral es menor que el valor mínimo o que, en el caso del superior, el umbral es mayor al valor máximo observado. Adicional a lo antes mencionado, un gráfico `boxplot` ayuda a detectar el grado dispersión y sesgo de la variable.

La función **`boxplot`** es propia de R y con esta se produce el gráfico `boxplot`. Esta función requiere los siguientes parámetros:

*datos*: un vectos con los datos correspondientes a la variable que se desea graficar.

*range*: valor deseado por el cual será multiplicado el rango intercuartil para la definición de los umbrales, 1.5 por default.

*boxwex*: es el ancho de la caja principal, 0.8 por default, éste sólo afecta la apariencia del gráfico. Para los fines del trabajo se utilizó 0.4.

```
boxplot(datos, range=1.5, boxwex)
```

# Bibliografía

- ANGUS, J.E. (1994). «The probability integral transform and related results». *SIAM review*, pp. 652–654.
- BACIGÁL, T. (2012). «Súasné nástroje modelovania závislosti pomocou kopúl a R Recent tools for modelling dependence with copulas and R». En: *FORUM STATISTICUM SLOVACUM*, p. 62.
- BILLINGSLEY, P. (1995). «Probability and Measure». *Inc. New York*.
- CONAPO (2000). «Metodología del índice de intensidad migratoria México-Estados Unidos».
- CONWAY, D.A. (1979). *Multivariate distributions with specified marginals*. Tesis doctoral, Dept. of Statistics, Stanford University..
- CONWAY, D.A. (1986). «Plackett Family of Distributions». *Encyclopedia of Statistical Sciences*, **7**, pp. 1–5.
- DALLÁGLIO, G (1959). «Sulla compatibilitá delle funzioni di ripartizione doppia». *Rend. Mat. e Appl.*, **18(5)**, pp. 385–413.
- DEMARTA, S. y MCNEIL, A.J. (2005). «The t copula and related copulas». *International Statistical Review*, **73(1)**, pp. 111–129.
- 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*, **198(September)**, pp. 3–31.
- EMBRECHTS, P. (2009). «Copulas: A personal view». *Journal of Risk and Insurance*, **76(3)**, pp. 639–650.
- EMBRECHTS, P.; LINDSKOG, F. y MCNEIL, A. (2003). «Modelling dependence with copulas and applications to risk management». *Handbook of heavy tailed distributions in finance*, **8(329-384)**.

- EMBRECHTS, P.; MCNEIL, A. y STRAUMANN, D. (1999). «Correlation: pitfalls and alternatives». *RISK-LONDON-RISK MAGAZINE LIMITED*-, **12**, pp. 69–71.
- ERDELY, A. (2009). «Cópulas y dependencia de variables aleatorias: Una introducción.» *Miscelánea Matemática*, **48**, pp. 7–28.
- ERDELY, A. y DIAZ-VIERA, M. (2010). «Nonparametric and Semiparametric Bivariate Modeling of Petrophysical Porosity-Permeability Dependence from Well Log Data». *Copula Theory and Its Applications*, pp. 267–278.
- ERDELY, A. y GONZÁLEZ-BARRIOS, J.M. (2010). «A nonparametric symmetry test for absolutely continuous bivariate copulas». *Statistical Methods & Applications*, pp. 1–25.
- FÉRON, R. (1956). «Sur les tableaux de corrélation dont les marges sont données, cas de l'espace à trois dimensions». *Publ. Inst. Statist. Univ. Paris*, **5**, pp. 3–12.
- FISHER, N.I. (1997). «Copulas». *Encyclopedia of Statistical Sciences, Update*, **1**, pp. 159–163.
- FISHER, N.I. y SEN, P.K. (1994). *The collected works of Wassily Hoeffding*. Springer.
- FRÉCHET, M. (1951). «Sur les tableaux de corrélation dont les marges sont données». *Ann. Univ. Lyon Sect. A*, **9**, pp. 53–77.
- GANI, J.M. (1982). *The making of statisticians*. Springer-Verlag.
- GENEST, C. y MACKAY, J. (1986). «The joy of copulas: bivariate distributions with uniform marginals». *American Statistician*, pp. 280–283.
- 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. (2006). «Discussion of “Copulas: Tales and facts,” by Thomas Mikosch». *Extremes*, **9(1)**, pp. 27–36.
- GENEST, C.; RÉMILLARD, B. y BEAUDOIN, D. (2009). «Goodness-of-fit tests for copulas: A review and a power study». *Insurance: Mathematics and Economics*, **44(2)**, pp. 199–213.
- HOEFFDING, W. (1940). *Maszstabinvariante korrelationstheorie*. Tesis doctoral, Universität Berlin.
- HOEFFDING, W. (1948). «A non-parametric test of independence». *The Annals of Mathematical Statistics*, **19(4)**, pp. 546–557.

- 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/>
- JOE, H. (1997). *Multivariate models and dependence concepts*. volumen 73. Chapman & Hall/CRC.
- JOE, H. (2006). «Discussion of “Copulas: Tales and facts,” by Thomas Mikosch». *Extremes*, **9(1)**, pp. 37–41.
- 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/>
- KRUSKAL, W.H. (1958). «Ordinal measures of association». *Journal of the American Statistical Association*, pp. 814–861.
- LEHMANN, E. (1975). «Nonparametrics: Statistical Methods Based on Ranks».
- LEHMANN, E.L. (1966). «Some concepts of dependence». *The Annals of Mathematical Statistics*, pp. 1137–1153.
- MARI, D.D. y KOTZ, S. (2001). *Correlation and dependence*. World Scientific Publishing Company.
- MIKOSCH, T. (2006). «Copulas: Tales and facts». *Extremes*, **9(1)**, pp. 3–20.
- MOOD, A.M. (1963). *Introduction to the theory of statistics*. McGraw-Hill.
- MUÑOZ PEREZ, J. y FERNÁNDEZ PALACÍN, A. (1987). «Estimating the quantile function by Bernstein polynomials». *Computational Statistics & Data Analysis*, **5(4)**, pp. 391–397.
- MURRELL, P. (2006). *R graphics*. CRC Press.
- NELSEN, R.B. (2006). *An introduction to copulas*. Springer Verlag.
- PÉREZ, CID (2007). *Cópulas: modelado y medición de dependencia estocástica bivariada y su aplicación en el cálculo del Valor en Riesgo*. Tesina o Proyecto, Universidad Nacional Autónoma de México.
- PNUD (2006). «Informe sobre Desarrollo Humano México 2006-2007: migración y desarrollo humano».
- QUESADA-MOLINA, J.; RODRÍGUEZ-LALLENA, J. y ÚBEDA-FLORES, M. (2003). «What are copulas». *Monografías del Semin. Matem. Gracia de Galdeano*, **27**, pp. 499–506.

- R DEVELOPMENT CORE TEAM (2012). *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/>
- RÉNYI, A. (1959). «On measures of dependence». *Acta Mathematica Hungarica*, **10(3)**, pp. 441–451.
- RUSCHENDORF, L. (1976). «Asymptotic distributions of multivariate rank order statistics». *The Annals of Statistics*, pp. 912–923.
- SANCETTA, A. y SATCHELL, S. (2004). «The Bernstein copula and its applications to modeling and approximations of multivariate distributions». *Econometric Theory*, **20(3)**, pp. 535–562.
- SCHWEIZER, B. (1991). «Thirty years of copulas». *Advances in probability distributions with given marginals*, pp. 13–50.
- SCHWEIZER, B (2007). «Introduction to copulas.» *J. Hydrol. Eng.*, **12(4)**, pp. 346–346.
- SCHWEIZER, B. y SKLAR, A. (1983). *Probabilistic metric spaces*. North-Holland New York.
- 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)**, pp. 229–231.
- SKLAR, A. (1996). «Random Variables, Distribution Functions, and Copulas: A Personal Look Backward and Forward». *Lecture Notes-Monograph Series*, pp. 1–14.
- VRIES, C.G. y ZHOU, C. (2006). «Discussion of “Copulas: Tales and facts,” by T. Mikosch». *Extremes: statistical theory and applications in science, engineering and economics*, **9(1)**, pp. 23–25.
- 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/>