

UNA FALACIA SOBRE DEPENDENCIA ENTRE VARIABLES
ALEATORIAS ILUSTRADA CON TEORÍA DE CÓPULAS

JARLES ANDRÉS MARIMON HERNÁNDEZ



UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS
FACULTAD DE CIENCIAS Y EDUCACIÓN
PROYECTO CURRICULAR DE MATEMÁTICAS
BOGOTÁ

2016

UNA FALACIA SOBRE DEPENDENCIA ENTRE VARIABLES
ALEATORIAS ILUSTRADA CON TEORÍA DE CÓPULAS

JARLES ANDRÉS MARIMON HERNÁNDEZ

Trabajo de grado para optar al título de Matemático

Dirigida por:

M.Sc. Luis Alejandro Másmela Caíta

UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS
FACULTAD DE CIENCIAS Y EDUCACIÓN
PROYECTO CURRICULAR DE MATEMÁTICAS
BOGOTÁ

2016

A mi hermosa madre Candelaria Hernández

Agradecimientos

Considero que lo que hoy somos refleja la sumatoria de experiencias e historias con las personas que nos hemos atravesado a lo largo de toda nuestra vida, por eso me resulta complejo (en un sentido no matemático) ordenar según el nivel de agradecimiento el conjunto de todas las personas que han hecho posible culminar esta etapa más en mi vida. Sin embargo, hay una persona que resulta ser un elemento máximo y es mi madre, a quien debo y espero compensar no solo por darme la vida sino también por ser mi motivación, ejemplo de perseverancia y esfuerzo, ella que cambió el sueño de muchas noches para trabajar mientras yo dormía o trasnochaba haciendo trabajos, para ella todo mi amor y agradecimiento.

Estar en una ciudad con aproximadamente seis veces la población del lugar donde naciste y a más de mil kilómetros de tu familia además de enseñarte cosas nuevas puede generar cierta melancolía, sin embargo contar con el apoyo del señor Omar Acosta y su familia nos permitió disimularlo y sobre todo para mí en estos años de vida universitaria, sé que la mejor forma de agradecerles es cumpliendo los objetivos que les prometí y en eso estoy.

A mi familia en general le agradezco el apoyo y la motivación recibida, principalmente Ana, Michelle y Yoly.

Ningún navegante puede obviar los faros y un estudiante no puede obviar sus guías en el camino de aprender, con esto busco resaltar y agradecer los conocimientos adquiridos por mis profesores y sus aportes a la construcción de mi proyecto de vida, especialmente al profesor Fernando Villarraga y a mi director Luis Alejandro Másmela quien decidió apoyarme desde el semillero IPREA en la realización de este trabajo y cuyos conocimientos fueron decisivos para su

culminación, agradezco también su paciencia en la preparación de las ponencias que he presentado.

A los amigos que me entendieron cuando no pude estar con ellos por razones académicas, a los que compartieron conmigo muchos días en la universidad, los que hicieron parte de mi equipo de trabajo y todos los demás les agradezco por hacer este camino más ameno.

Incluso en la distancia agradezco a los profesores Alexander Mcneil, Paul Embrechts y Marius Hofert por sugerirme trabajar en este tema tan apasionante y aclararme con la mejor disposición las dudas que se me presentaron.

A todos, ¡gracias!

Índice general

Introducción	III
Objetivos	v
1. Preliminares	1
1.1. Función Inversa Generalizada	1
1.2. Espacios de Probabilidad	2
1.3. Variables Aleatorias	4
1.4. Vectores Aleatorios	7
1.5. Relación Entre Variables Aleatorias	11
2. Teoría de Cópulas y Dependencia	14
2.1. Transformaciones Cuantil y de Probabilidad	15
2.2. Cópulas	17
2.2.1. Ejemplos de Cópulas	20
2.2.2. Cópula Gaussiana	23
2.2.3. Métodos de Construcción de Cópulas	24
2.2.4. Invariancia	26
2.3. Conceptos alternativos de dependencia	27
2.3.1. Comonotonidad y Contramonotonidad	28
2.3.2. Propiedades deseadas de las medidas de dependencia	31
3. La Falacia	33

4. Conclusiones	41
A. Códigos en R	44

Introducción

Si bien hablar de una falacia implica hablar de algo que es falso aun cuando se cree cierto, en el presente trabajo se ilustra por medio de un ejemplo que las distribuciones marginales y el coeficiente de correlación lineal no determinan de manera única la función de distribución conjunta, dado que muchas veces se asume lo contrario, es decir, que la función de distribución conjunta sí está determinada de manera única dadas las distribuciones marginales y el coeficiente de correlación lineal. La herramienta que se usará principalmente es la moderna y grandiosa teoría de cópulas, la cual se ha enriquecido con la necesidad de encontrar una relación (dependencia) entre variables aleatorias y también entre las funciones de distribución multivariadas y sus distribuciones marginales; según [7] la palabra cópula fue empleada por primera vez en un sentido matemático o estadístico por el matemático americano Abe Sklar en el célebre teorema que lleva su nombre y que más adelante estudiaremos. Obviamente esta no es la única medida de dependencia, pues es bien conocido y útil el coeficiente de correlación lineal de Pearson, pero veremos que en general no resulta conveniente considerarlo para medir dependencia.

La idea de estudiar lo que se presenta en este trabajo nace como sugerencia vía e-mail de uno de los autores del artículo guía [3], Alexander Mc Neil, quien es profesor del Departamento de Matemática Actuarial y Estadística de Heriot-Watt University, él también escribió junto a Rüdiger Frey y Paul Embrechts el libro *Quantitative Risk Management*, del cual se complementan las demostraciones que se presentarán. Cabe recalcar que gran parte de la estructura del segundo capítulo se regirá por el contenido de [3] y se complementará con [6], es decir, de ahí se harán las interpretaciones que no se citen en el cuerpo del trabajo.

La organización de este trabajo está compuesta por tres capítulos, el prime-

ro contiene aspectos preliminares de la teoría de la probabilidad tales como los espacios de probabilidad, variables aleatorias, funciones de distribución, vectores aleatorios y se finaliza con la definición, algunas ventajas y desventajas del coeficiente de correlación lineal. En el segundo capítulo entramos a definir cópulas, se da el teorema de Sklar, ejemplos, métodos de construcción y algunas propiedades de las cópulas y se concluye con aspectos propios de las cópulas como medida de dependencia. Finalmente en el tercer capítulo ahondamos en la falacia que nos compete.

Objetivos

Objetivo General

Ilustrar una falacia relacionada con variables aleatorias usando la teoría de cópulas.

Objetivos Específicos

- Reconstruir parcialmente el contenido teórico de los capítulos 2 y 4 del artículo [3], con el fin de dar los resultados necesarios para la ilustración de la falacia.
- Dar algunas propiedades sobre dependencia entre variables aleatorias desde el punto de vista de la teoría de cópulas.
- Visualizar por medio del software R ciertos resultados obtenidos.

Capítulo 1

Preliminares

En este capítulo introduciremos algunos conceptos relacionados con el desarrollo axiomático de la teoría de probabilidad, principalmente espacio de probabilidad, propiedades de las variables aleatorias y vectores aleatorios. No obstante presentaremos antes una proposición cuyos resultados encontraremos a lo largo de todo este trabajo. Los conceptos que se presentan en este capítulo son tomados principalmente de [2], sin embargo también se han tomado algunos de [6] y [9].

1.1. Función Inversa Generalizada

La inversa generalizada de una función T creciente, definida como $T^{\leftarrow}(\alpha) = \inf \mathcal{D}_\alpha$ donde $\mathcal{D}_\alpha = \{x : T(x) \geq \alpha\}$ y por convención $\inf \emptyset = \infty$ nos resultará de gran utilidad, por eso se darán a continuación algunas de sus propiedades. Cabe resaltar que según las características de la función T que se tome, su función inversa generalizada puede estar definida en el conjunto de los reales extendidos $\overline{\mathbb{R}} = [-\infty, \infty]$.

Proposición 1.1.1. *Sea $T : A \subset \mathbb{R} \rightarrow \mathbb{R}$ una función creciente, entonces*

- a) T^{\leftarrow} es una función creciente y continua por derecha.
- b) T es continua $\iff T^{\leftarrow}$ es estrictamente creciente.
- c) T es estrictamente creciente $\iff T^{\leftarrow}$ es continua.

Para las siguientes propiedades asúmase adicionalmente que $T^{\leftarrow}(\alpha) < \infty$.

d) Si T es continua a derecha, $T(x) \geq \alpha \iff T^{\leftarrow}(\alpha) \leq x$

e) $T^{\leftarrow} \circ T(x) \leq x$.

f) $T \circ T^{\leftarrow}(\alpha) \geq \alpha$.

g) T es estrictamente creciente $\implies T^{\leftarrow} \circ T(x) = x$.

h) T es continua $\implies T \circ T^{\leftarrow}(\alpha) = \alpha$.

Observación 1.1.1. Cuando T es estrictamente creciente y continua su inversa generalizada T^{\leftarrow} coincide con su función inversa, cuando este sea el caso usaremos la notación $T^{\leftarrow} = T^{-1}$.

1.2. Espacios de Probabilidad

La probabilidad nos sitúa en una clase particular de experimentos, los **Experimentos Aleatorios**, es decir, experimentos cuyo resultado no puede ser determinado con anticipación, el ejemplo clásico es el lanzamiento al aire de una moneda. Siempre que podamos conocer el conjunto de resultados de un experimento aleatorio lo llamaremos el espacio muestral y se simbolizará con la letra griega Ω , para nuestro ejemplo de la moneda tendremos dos posibles resultados: cara o sello, así $\Omega = \{C, S\}$.

De la teoría de conjuntos tenemos que dado un conjunto A podemos hablar de $\mathcal{P}(A)$, donde $\mathcal{P}(A) := \{a : a \subset A\}$, esto es la esencia de la siguiente definición.

Definición 1.2.1. (σ - **Álgebra**) Sea $\Omega \neq \emptyset$. Una colección \mathfrak{J} de subconjuntos de Ω , es decir, $\mathfrak{J} \subset \mathcal{P}(\Omega)$. Se dice que es una σ -álgebra sobre Ω siempre que se satisfagan las siguientes 3 propiedades:

1. $\Omega \in \mathfrak{J}$

2. Si $A \in \mathfrak{J}$ entonces $A^c \in \mathfrak{J}$. Donde A^c es el complemento de A .

3. Si $A_1, A_2, \dots \in \mathfrak{J}$, entonces $\bigcup_{i=1}^{\infty} A_i \in \mathfrak{J}$.

Los elementos de \mathfrak{J} se llaman eventos y a la pareja (Ω, \mathfrak{J}) se le llama un **espacio medible**.

Al considerar el ejemplo de lanzar la moneda donde $\Omega = \{C, S\}$, se puede ver que los conjuntos $\mathcal{F} = \{\emptyset, \Omega\}$ y $\mathcal{G} = \{\emptyset, \{C\}, \{S\}, \{C, S\}\}$ son σ -álgebras sobre Ω , mientras que $\mathcal{H} = \{\emptyset, \{C\}\}$ no lo es.

Una característica que tienen las σ -álgebras es que si $\Omega \neq \emptyset$ y $\mathfrak{J}_1, \mathfrak{J}_2, \dots$ son σ -álgebras sobre Ω , entonces $\bigcap_{i=1}^{\infty} \mathfrak{J}_i$ resulta ser una σ -álgebra sobre Ω , por otro lado, en general la unión de σ -álgebras sobre Ω no resulta ser una σ -álgebra sobre Ω . Lo anterior nos facilita dar la siguiente definición.

Definición 1.2.2. (σ -Álgebra Generada) Sea $\Omega \neq \emptyset$ y sea \mathcal{A} una colección de subconjuntos de Ω . Sea $\mathcal{M} := \{\mathfrak{J} : \mathfrak{J} \text{ es una } \sigma\text{-álgebra sobre } \Omega \text{ que contiene a } \mathcal{A}\}$. Entonces $\sigma(\mathcal{A}) := \bigcap_{\mathfrak{J} \in \mathcal{M}} \mathfrak{J}$ es la más pequeña σ -álgebra sobre Ω que contiene a \mathcal{A} . $\sigma(\mathcal{A})$ es llamada la σ -álgebra generada por \mathcal{A} .

Puede resultar natural el hecho de pensar que el espacio muestral de cierto experimento sea el conjunto de los números reales, esto quiere decir, $\Omega = \mathbb{R}$. Lo anterior nos motiva a definir la σ -álgebra de Borel, la cual resulta ser la más pequeña σ -álgebra sobre \mathbb{R} que contiene todos los intervalos de la forma $(-\infty, a]$ donde $a \in \mathbb{R}$. Simbolizaremos la σ -álgebra de Borel por \mathcal{B} y diremos que A es un subconjunto de Borel si $A \in \mathcal{B}$.

Una vez dado el concepto de espacio medible podemos a continuación definir una medida de probabilidad y posteriormente espacio de probabilidad.

Definición 1.2.3. (Espacio de Probabilidad) Sea (Ω, \mathfrak{J}) un espacio medible. Una función P de valor real definida sobre \mathfrak{J} es una medida de probabilidad sobre (Ω, \mathfrak{J}) siempre que satisfaga las siguientes condiciones:

1. $P(A) \geq 0$ para todo $A \in \mathfrak{J}$ (Propiedad de nonegatividad).
2. $P(\Omega) = 1$ (Propiedad de normalizado)
3. Si A_1, A_2, \dots son eventos mutuamente excluyentes en \mathfrak{J} , esto quiere decir que

$$A_i \cap A_j = \emptyset \text{ para todo } i \neq j$$

Entonces

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) \quad (\text{Aditividad contable})$$

A la tripleta $(\Omega, \mathfrak{J}, P)$ se le llama un espacio de probabilidad.

1.3. Variables Aleatorias

Los esfuerzos de dar una mirada superficial y ligera a los conceptos que hasta el momento hemos planteado se han hecho con el fin de llegar al concepto de variable aleatoria y vector aleatorio junto con algunas de sus propiedades, el primero de estos conceptos que resulta ser de gran interés nos permite asignarle una característica numérica a los eventos de nuestros experimentos y el segundo concepto nos facilita manipular simultáneamente la relación entre más de una variable aleatoria. Empezamos entonces con el concepto de variable aleatoria en esta sección.

Definición 1.3.1. (Variable Aleatoria) Sea $(\Omega, \mathfrak{J}, P)$ un espacio de probabilidad. Una variable aleatoria (real) es una función $X : \Omega \rightarrow \mathbb{R}$ tal que, para todo $A \in \mathcal{B}$, $X^{-1}(A) \in \mathfrak{J}$, donde \mathcal{B} es la σ -álgebra de Borel sobre \mathbb{R} como se vio anteriormente. Equivalentemente $X : \Omega \rightarrow \mathbb{R}$ es una variable aleatoria si y solo si $X^{-1}((-\infty, x]) \in \mathfrak{J}$ para todo $x \in \mathbb{R}$ lo cual se tiene por el hecho que la σ -álgebra de Borel \mathcal{B} en \mathbb{R} está generada por la colección de todos los intervalos de la forma $(-\infty, x], x \in \mathbb{R}$.

Cuando se tiene una variable aleatoria $X : \Omega \rightarrow \mathbb{R}$ sobre un espacio de probabilidad $(\Omega, \mathfrak{J}, P)$ podemos dar una nueva medida de probabilidad a la que llamaremos P_X definida sobre \mathcal{B} , tal que para todo $B \in \mathcal{B}$

$$P_X(B) = P(X^{-1}(B)) := P(\{X \in B\})^1$$

A esta nueva medida de probabilidad sobre $(\mathbb{R}, \mathcal{B})$ se le llama la distribución de X y está bien definida ya que, $X^{-1}(B) \in \mathfrak{J}$. Se puede verificar que P_X satisface cada una de las propiedades de una medida de probabilidad y formalizamos la definición de esta medida de probabilidad a continuación.

Definición 1.3.2. (Función de Distribución) Sea X una variable aleatoria real. A la función F_X definida sobre \mathbb{R} mediante

$$\begin{aligned} F_X(x) &:= P_X((-\infty, x]) \\ &= P(X \leq x) \end{aligned}$$

¹Por notación se define $\{X \in B\} := \{\omega \in \Omega : X(\omega) \in B\}$ para todo $B \in \mathcal{B}$. En particular $\{X \in (-\infty, x]\} = \{X \leq x\} := \{\omega \in \Omega : X(\omega) \leq x\}$.

se le llama la función de distribución (o función de distribución acumulada **fda**) de la variable aleatoria X . Usaremos la notación $X \sim F$ para referirnos a que la variable aleatoria X tiene como función de distribución a F .

Algunas propiedades que caracterizan las funciones de distribución de una variable aleatoria se muestran a continuación.

Teorema 1.3.1. (Propiedades de la Función de Distribución) Sea X una variable aleatoria real definida sobre $(\Omega, \mathfrak{F}, P)$. La función de distribución F_X satisface que

- a) Si $x < y$ entonces $F_X(x) \leq F_X(y)$. Es decir, F_X es una función no decreciente.
- b) $F_X(x^+) := \lim_{h \rightarrow 0^+} F_X(x+h) = F_X(x)$ para todo $x \in \mathbb{R}$. Esto quiere decir que F_X es continua a derecha en todo punto de \mathbb{R} .
- c) $\lim_{x \rightarrow \infty} F_X(x) = 1$.
- d) $\lim_{x \rightarrow -\infty} F_X(x) = 0$.

Si situamos la inversa generalizada en funciones de distribución $F : \mathbb{R} \rightarrow [0, 1]$, la función inversa generalizada se conoce como la función cuantil, a la que también notaremos F^{-1} en adelante aun cuando F no sea continua y estrictamente creciente. Además, si tomamos $\alpha \in (0, 1)$, se puede asegurar lo siguiente.

Teorema 1.3.2. El ínf de \mathcal{D}_α existe.

Demostración. Para que el ínf exista debemos probar que el conjunto es no vacío y que además está acotado inferiormente.

Veamos primero que \mathcal{D}_α es un conjunto no vacío. Por Teorema 1.3.1.c tendremos que

$$\lim_{x \rightarrow \infty} F(x) = 1.$$

Entonces para cualquier $\epsilon > 0$ existe $N_0 \in \mathbb{N}$ tal que $|F(x) - 1| < \epsilon$ siempre que $x \geq N_0$, en particular $|F(N_0) - 1| \leq \epsilon$ y por estar F acotada superiormente por 1 tendremos que para $0 < \alpha < 1$

$$\begin{aligned} 1 - F(N_0) &\leq \epsilon \\ \alpha = 1 - \epsilon &\leq F(N_0) \end{aligned}$$

es decir, $N_0 \in \mathcal{D}_\alpha$.

\mathcal{D}_α es un conjunto acotado, pues por el Teorema 1.3.1.d tenemos que

$$\begin{aligned}\lim_{x \rightarrow -\infty} F(x) &= 0 \\ \lim_{x \rightarrow \infty} F(-x) &= 0\end{aligned}$$

por tanto para todo $\epsilon > 0$ existe $N_0 \in \mathbb{N}$ tal que si $x \geq N_0$ entonces $|F(-N_0)| < \epsilon$, en particular por ser $\alpha > 0$ y estar F acotada inferiormente por 0 se tendrá que

$$F(-N_0) < \alpha. \quad (1.1)$$

Por otro lado, si $x \in \mathcal{D}_\alpha$ y $x < -N_0$ se tendría que $F(-N_0) \geq F(x) \geq \alpha$ por Teorema 1.3.1.c, hecho que contradice (1.1) y por tanto para cualquier $x \in \mathcal{D}_\alpha$ se debe cumplir que $-N_0 \leq x$. \square

La definición de variable aleatoria que hemos dado coincide con la de una función medible, en ese sentido, si usamos el hecho que toda función $T : \mathbb{R} \rightarrow \mathbb{R}$ monótona es medible, obtenemos el siguiente teorema que caracteriza variables aleatorias bajo transformaciones monótonas.

Teorema 1.3.3. *Si $T : \mathbb{R} \rightarrow \mathbb{R}$ es monótona y $X : \Omega \rightarrow \mathbb{R}$ es una variable aleatoria, entonces $T(X) : \Omega \rightarrow \mathbb{R}$ es también una variable aleatoria.*

Demostración. $T(X)$ es una variable aleatoria, pues si tomamos $B \in \mathcal{B}$ se puede ver que

$$[T(X)^{-1}](B) = X^{-1}[T^{-1}(B)]$$

y dado que $B_1 = T^{-1}(B) \in \mathcal{B}$, por ser X variable aleatoria se tendrá también que $X^{-1}(T^{-1}(B)) \in \mathcal{B}$. \square

La anterior proposición abre paso al siguiente resultado, el cual será de gran utilidad cuando se le apliquen transformaciones crecientes a las variables aleatorias.

Proposición 1.3.4. *Si X es una variable aleatoria con función de distribución F y T es una transformación creciente, entonces*

$$1. \{X \leq x\} \subset \{T(X) \leq T(x)\} \text{ y}$$

$$P[T(X) \leq T(x)] = P[X \leq x] + P[T(X) = T(x), X > x] \quad (1.2)$$

$$2. P[F(X) \leq F(x)] = P[X \leq x].$$

Demostración. Para mostrar la contención consideremos los conjuntos $A = \{X \leq x\} := \{\omega \in \Omega : X(\omega) \leq x\}$ y $B = \{T(X) \leq T(x)\} := \{\omega \in \Omega : T(X(\omega)) \leq T(x)\}$. Sea $\omega_1 \in A$ entonces $X(\omega_1) \leq x$, por ser T creciente entonces se tendrá que $T(X(\omega_1)) \leq T(x)$, es decir, $\omega_1 \in B$. Lo anterior nos permite asegurar que $B = A \cup (B \cap A^c)$, donde $A \cap (B \cap A^c) = \emptyset$, por lo tanto,

$$\begin{aligned} P[T(X) \leq T(x)] &= P(B) = P(A \cup (B \cap A^c)) \\ &= P(A) + P(A \cap A^c) \\ &= P(X \leq x) + P(T(X) \leq T(x), X > x) \\ &= P(X \leq x) + P(T(X) = T(x), X > x) \\ &+ P(T(X) < T(x), X > x) \\ &= P(X \leq x) + P(T(X) = T(x), X > x) \text{ por ser } T \text{ creciente} \end{aligned}$$

queda probada así la primera parte. Nótese que si T fuera estrictamente creciente, entonces $P(T(X) = T(x), X > x) = 0$ y así $P[T(X) \leq T(x)] = P[X \leq x]$.

La segunda proposición resulta de tomar $T = F$ en (1.2) y teniendo en cuenta que para cualquier x , el evento $\{F(X) = F(x), X > x\}$ corresponde a una parte constante de la función de distribución, es decir, su probabilidad es cero. \square

Una vez dada la definición de variable aleatoria junto con algunas propiedades de las funciones de distribución estamos listos para introducir el concepto de vector aleatorio.

1.4. Vectores Aleatorios

Hasta el momento hemos centrado nuestra atención en estudiar una sola característica de un experimento dado, sin embargo junto con la necesidad de generalizar y de manipular simultáneamente el comportamiento de dos o más variables aleatoria surgen los vectores aleatorios.

Debido a que más adelante los vectores aleatorios jugarán un rol importante para modelos en dos dimensiones, restringiremos los resultados de esta sección al caso bidimensional, sin embargo el lector interesado en conocer el comportamiento en mayores dimensiones puede consultar en [2], pues como lo hemos

hecho en las anteriores secciones, los resultados sobre vectores aleatorios serán tomados también de ahí.

Definición 1.4.1. (Vector aleatorio) Sean X_1, X_2 variables aleatorias definidas sobre el mismo espacio de probabilidad $(\Omega, \mathfrak{F}, P)$. A la función $\mathbf{X} : \Omega \rightarrow \mathbb{R}^2$ definida por

$$\mathbf{X}(\omega) := (X_1(\omega), X_2(\omega))^t{}^2$$

se le llama un vector aleatorio bidimensional y lo llamaremos simplemente vector aleatorio en adelante.

Análogo a lo visto en variables aleatorias, a la medida de probabilidad definida por

$$P_{\mathbf{X}}(B) := P(\mathbf{X} \in B); B \in \mathcal{B}_2$$

se le llama la distribución del vector \mathbf{X} .

Observación 1.4.1. Al definir $P_{\mathbf{X}}$ hemos tomado $B \in \mathcal{B}_2$, es decir, en la σ -álgebra de Borel sobre \mathbb{R}^2 . Esta σ -álgebra es la generada por todos los intervalos de la forma

$$(a, b] := \{x = (x_1, x_2) \in \mathbb{R}^2 : a_i \leq x_i \leq b_i, i = 1, 2\}$$

donde $a = (a_1, a_2)$ y $b = (b_1, b_2)$ son elementos de \mathbb{R}^2 y $a \leq b$, es decir, $a_i \leq b_i, i = 1, 2$.

Definición 1.4.2 (Función de Distribución Conjunta). La función de distribución bivariada de un vector aleatorio $\mathbf{X} = (X_1, X_2)^t : \Omega \rightarrow \mathbb{R}^2$ sobre el espacio de probabilidad $(\Omega, \mathfrak{F}, P)$ está definida para todo x_1, x_2 en \mathbb{R} por

$$F(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2).$$

Además, las respectivas funciones de distribución de las variables aleatorias X_1 y X_2 están dadas por

$$F_{X_1}(x_1) = \lim_{x_2 \rightarrow \infty} F(x_1, x_2)$$

$$F_{X_2}(x_2) = \lim_{x_1 \rightarrow \infty} F(x_1, x_2)$$

a cada función F_{X_j} se le llama la **función de distribución marginal** de la variable $X_j, j = 1, 2$.

²El superíndice t indica que $(X_1(\omega), X_2(\omega))^t$ es el vector transpuesto de $(X_1(\omega), X_2(\omega))$.

De lo anterior se tiene que dada la función de distribución conjunta se pueden encontrar cada una de las distribuciones marginales, sin embargo el recíproco no siempre se tiene y será casualmente objeto de estudio en posteriores capítulos. Veamos a continuación algunas caracterizaciones de la función de distribución conjunta.

Teorema 1.4.1. (*Propiedades de la Función de Distribución Conjunta*) Sea $X = (X_1, X_2)^t$ un vector aleatorio. La función de distribución conjunta F de las variables X_1, X_2 tiene las siguientes propiedades.

a)

$$\Delta_a^b F := F(b_1, b_2) + F(a_1, a_2) - F(a_1, b_2) - F(b_1, a_2) \geq 0$$

donde $a = (a_1, a_2), b = (b_1, b_2) \in \mathbb{R}^2$ con $a_1 \leq b_1$ y $a_2 \leq b_2$.

b) F es continua a derecha en cada componente.

c)

$$\lim_{x_1 \rightarrow -\infty} F(x_1, x_2) = 0 \text{ y } \lim_{x_2 \rightarrow -\infty} F(x_1, x_2) = 0.$$

d)

$$\lim_{(x_1, x_2) \rightarrow (\infty, \infty)} F(x_1, x_2) = 1.$$

Siempre que exista una función $f : \mathbb{R}^2 \rightarrow [0, +\infty)$ tal que

$$F(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f(t_1, t_2) dt_1 dt_2,$$

diremos que el vector aleatorio $(X_1, X_2)^t$ es **absolutamente continuo** y llamaremos a f la **función de densidad** de las variables aleatorias X_1, X_2 . Una característica de f que surge del Teorema 1.4.1.d es que

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(t_1, t_2) dt_1 dt_2 = 1.$$

También, como consecuencia directa del teorema fundamental del cálculo podemos obtener el siguiente teorema.

Teorema 1.4.2. *Sea X_1 y X_2 variables aleatorias continuas con función de distribución conjunta F . Entonces la función de densidad de probabilidad f está dada por*

$$f(x_1, x_2) = \frac{\partial^2 F(x_1, x_2)}{\partial x_1 \partial x_2} = \frac{\partial^2 F(x_1, x_2)}{\partial x_2 \partial x_1} \quad (1.3)$$

para todos los puntos (x_1, x_2) donde $f(x_1, x_2)$ sea continua.

Terminaremos esta sección aseverando que tanto para la variable X_1 como para X_2 podemos determinar sus respectivas **funciones de densidad marginal**, las cuales serán las funciones que se obtengan de la siguiente manera:

Supongamos que X_1 y X_2 son variables aleatorias continuas con función de densidad de probabilidad f y definamos la función g dada por

$$g(x_1) := \int_{-\infty}^{\infty} f(x_1, x_2) dx_2$$

Así

$$\begin{aligned} \int_{-\infty}^{x_1} g(u) du &= \int_{-\infty}^{x_1} \int_{-\infty}^{\infty} f(u, x_2) dx_2 du \\ &= \lim_{t \rightarrow \infty} \int_{-\infty}^{x_1} \int_{-\infty}^t f(u, x_2) dx_2 du \\ &= \lim_{t \rightarrow \infty} F(x_1, t) = F_{X_1}(x_1). \end{aligned}$$

Además, dado que

$$\int_{-\infty}^{\infty} g(x_1) dx_1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2) dx_2 dx_1 = 1$$

g será la función de densidad marginal de la variable aleatoria X_1 y de manera análoga podemos hallar la función de densidad marginal de la variable aleatoria X_2 . En adelante notaremos estas funciones por $f_{X_i}(x_i)$, $i = 1, 2$.

Para terminar esta sección mencionaremos el siguiente teorema que se tiene para el caso de variables aleatorias absolutamente continuas.

Teorema 1.4.3. (Teorema de Transformación) *Sea $\mathbf{X} = (X_1, X_2)$ un vector aleatorio con función de densidad $f_{\mathbf{X}}$ y sea $g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ una función inyectiva. Supóngase que tanto g como su inversa $h : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ son continuas. Si las derivadas parciales de h existen y son continuas y si su Jacobiano J es diferente de cero, entonces el vector aleatorio $\mathbf{Y} = (Y_1, Y_2) := g(\mathbf{X})$ tiene función de densidad $f_{\mathbf{Y}}$*

$$f_{\mathbf{Y}}(y_1, y_2) = \begin{cases} |J(y_1, y_2)| f_{\mathbf{X}}(h(y_1, y_2)), & \text{si } (y_1, y_2) \text{ está en el rango de } g \\ 0, & \text{en cualquier otro caso.} \end{cases} \quad (1.4)$$

1.5. Relación Entre Variables Aleatorias

En la literatura nos encontramos con que dos eventos dados A y B se dicen independientes si y solo si $P(A \cap B) = P(A)P(B)$, de lo contrario serán eventos dependientes. Extenderemos un poco el concepto de dependencia e independencia a las variables aleatorias en esta sección.

Definición 1.5.1. (*Variables Aleatorias Independientes*) Sean X_1 y X_2 dos variables aleatorias reales definidas sobre el mismo espacio de probabilidad. Si para cualquier par de conjuntos de Borel A y B de \mathbb{R} tenemos que

$$P(X_1 \in A, X_2 \in B) = P(X_1 \in A)P(X_2 \in B),$$

entonces decimos que X_1 y X_2 son independientes.

Como consecuencia de la anterior definición podemos afirmar que para todo $x_1, x_2 \in \mathbb{R}$ tendremos que

$$F(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2) = P(X_1 \leq x_1)P(X_2 \leq x_2)$$

es decir,

$$F(x_1, x_2) = F_{X_1}(x_1)F_{X_2}(x_2) \text{ para todo } x_1, x_2 \in \mathbb{R}. \quad (1.5)$$

Recíprocamente si se tiene la condición de la ecuación (1.5) entonces las variables serán independientes.

Ya que parece normal encontrarse en la realidad con fenómenos dependientes, un tipo de dependencia que se podría tener es la lineal. Definiremos a continuación dos conceptos importantes para determinar si existe dependencia lineal entre dos variables aleatorias, estos son el de **covarianza** y **coeficiente de correlación**. Como se podrá ver mencionaremos la esperanza matemática y la varianza de una variable aleatoria X notadas por $E(X)$ y $\sigma^2(X)$ respectivamente, más detalles sobre las definiciones o propiedades de estas cantidades se pueden consultar en [2].

Definición 1.5.2. (*Covarianza*) Sean X_1 y X_2 variables aleatorias definidas sobre el mismo espacio de probabilidad y tales que $E(X_1^2) < \infty$ y $E(X_2^2) < \infty$. La covarianza entre X_1 y X_2 está definida por:

$$Cov(X_1, X_2) = E\{[X_1 - E(X_1)][X_2 - E(X_2)]\} \quad (1.6)$$

Algunas propiedades de la covarianza son:

- $Cov(X_1, X_2) = E(X_1X_2) - E(X_1)E(X_2)$
- $Cov(X_1, X_2) = Cov(X_2, X_1)$
- $\sigma^2(X_1) = Cov(X_1, X_1)$
- $Cov(aX_1 + b, X_2) = aCov(X_1, X_2)$ para cualquier $a, b \in \mathbb{R}$

Una resultado adicional surge de la primera propiedad antes mencionada y es que si X_1 y X_2 son independientes, entonces $Cov(X_1, X_2) = 0$, esto debido a que $E(X_1X_2) = E(X_1)E(X_2)$ dada la independencia entre las variables aleatorias X_1 y X_2 . Sin embargo, el recíproco de esta afirmación no se tiene en general. Hemos llegado a la parte final de este capítulo donde presentaremos una medida de dependencia lineal que nos resultará útil y con la cual compararemos algunas bondades de medir dependencia más adelante.

Definición 1.5.3. (Coeficiente de Correlación) Si X_1 y X_2 son variables aleatorias con $0 < \sigma^2(X_1) < \infty$ y $0 < \sigma^2(X_2) < \infty$, el coeficiente de correlación entre ellas está dado por

$$\rho(X_1, X_2) = \frac{Cov(X_1, X_2)}{\sqrt{\sigma^2(X_1)\sigma^2(X_2)}} \quad (1.7)$$

Este coeficiente toma valores en el intervalo $[-1, 1]$ y cumple que si las variables aleatorias X_1 y X_2 son independientes $\rho(X_1, X_2) = 0$ como consecuencia de que $Cov(X_1, X_2) = 0$. Si hay dependencia lineal perfecta entre las variables, es decir, $X_2 = aX_1 + b$ o $P[X_2 = aX_1 + b] = 1$ para $a \in \mathbb{R} \setminus \{0\}$, $b \in \mathbb{R}$ entonces $\rho(X_1, X_2) = \pm 1$, por último cuando hay dependencia lineal imperfecta, $-1 < \rho(X_1, X_2) < 1$.

A pesar de ser el coeficiente de correlación uno de los conceptos más asociados con dependencia, es solo una medida estocástica particular que tiene comportamientos deseados cuando de distribuciones normales multivariadas o más generalmente distribuciones elípticas se trata. Sin embargo, resulta desventajoso por los siguientes aspectos:

- Las varianzas de X_1 y X_2 deben ser finitas o la correlación lineal no esta definida.
- Si dos variables aleatorias son independientes su coeficiente de correlación lineal es cero, el recíproco en general no es cierto, es decir, si el coeficiente de correlación lineal entre dos variables aleatorias es cero, no siempre se puede asegurar independencia entre las variables.

- La correlación lineal no es invariante bajo transformaciones crecientes no lineales $T : \mathbb{R} \rightarrow \mathbb{R}$. Es decir, para dos variables aleatorias X_1, X_2 se tiene en general que

$$\rho(T(X_1), T(X_2)) \neq \rho(X_1, X_2).$$

Capítulo 2

Teoría de Cópulas y Dependencia

Ante la necesidad de encontrar alguna relación funcional que determinara la dependencia entre variables aleatorias y entre las funciones de distribución multivariadas y sus marginales de menor dimensión, el matemático Abe Sklar da la definición de la función cópula junto con el teorema que lleva su nombre, esto sumado a los trabajos realizados por Höfding, Maurice Fréchet, Ferón, Berthold Schweizer y otros, se convierte en el punto de partida del desarrollo de lo que ahora se conoce como la teoría de cópulas, así se afirma en [4]. Mostraremos entonces en este capítulo algunos aspectos y propiedades relevantes de la teoría de cópulas.

Ya que jugará un papel importante durante el desarrollo del presente capítulo, presentamos el hecho que una variable aleatoria X se distribuye uniforme estándar y lo notamos $X \sim U(0, 1)$, si su función de distribución está dada por:

$$F(x) = \begin{cases} 0, & \text{si } x < 0 \\ x, & \text{si } 0 \leq x \leq 1 \\ 1, & \text{si } x > 1 \end{cases} \quad (2.1)$$

2.1. Transformaciones Cuantil y de Probabilidad

Dado lo anterior, la siguiente proposición sobre las **transformaciones cuantil y de probabilidad** nos proporciona un par de relaciones entre la función cuantil y la distribución uniforme estándar.

Proposición 2.1.1. *Sea X una variable aleatoria con función de distribución F y sea F^{-1} la función cuantil de F , es decir,*

$$F^{-1}(\alpha) = \inf \mathcal{D}_\alpha$$

$\alpha \in (0, 1)$. Entonces

1. *Para cualquier variable aleatoria U que se distribuya uniforme estándar, tenemos que $F^{-1}(U) \sim F$.
Esto nos proporciona un método para simular variables aleatorias con función de distribución F .*
2. *Si F es continua entonces la variable aleatoria $F(X)$ tiene distribución uniforme estándar, esto es, $F(X) \sim U(0, 1)$.*

Demostración. Para la primera parte sea la variable aleatoria $Y = F^{-1}(U)$, como consecuencia de la Proposición 1.1.1.d tendremos que,

$$P[Y \leq y] = P[F^{-1}(U) \leq y] = P[U \leq F(y)] = F(y), \quad y \in \mathbb{R}$$

De la última igualdad podemos concluir que $F^{-1}(U) \sim F$.

Para la segunda parte, si $U = F(X)$ tendremos también como consecuencia de la Proposición 1.1.1.d que,

$$P[U \leq u] = P[F(X) \leq u] = P[X \leq F^{-1}(u)], \quad u \in (0, 1)$$

Además, ya que $X \sim F$ y de la Proposición 1.1.1.h

$$P[X \leq F^{-1}(u)] = F \circ F^{-1}(u) = u, \quad u \in (0, 1)$$

Es decir, $F(X) \sim U(0, 1)$ como se quería mostrar. \square

Ilustraremos lo anterior con una variable aleatoria $X \sim \exp(1)$, es decir, X se distribuye exponencial con parámetro 1 y dado que la función de distribución de una variable que se distribuye exponencial con parámetro $\lambda > 0$ es

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & \text{si } x \geq 0 \\ 0, & \text{si } x < 0 \end{cases} \quad (2.2)$$

podremos obtener sin mayores complicaciones su función cuantil, la cual coincide con su función inversa y está dada por:

$$F^{-1}(\alpha) = \begin{cases} \frac{-\ln(1-\alpha)}{\lambda}, & \text{si } 0 \leq \alpha < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases} \quad (2.3)$$

Para ilustrar el primer resultado de la proposición supongamos que queremos extraer una muestra de mil datos que sigan una distribución dada, digamos que sea de una variable $X \sim \exp(1)$. Lo primero que debemos hacer es generar 1000 datos que se distribuyan $U(0, 1)$.

A continuación evaluamos estos 1000 datos en la función cuantil de la función de distribución exponencial y construimos un histograma que permita ilustrar el comportamiento de los datos. Lo anterior se puede apreciar en la Figura 2.1, donde notamos un comportamiento similar entre el histograma obtenido y una curva $\exp(1)$.

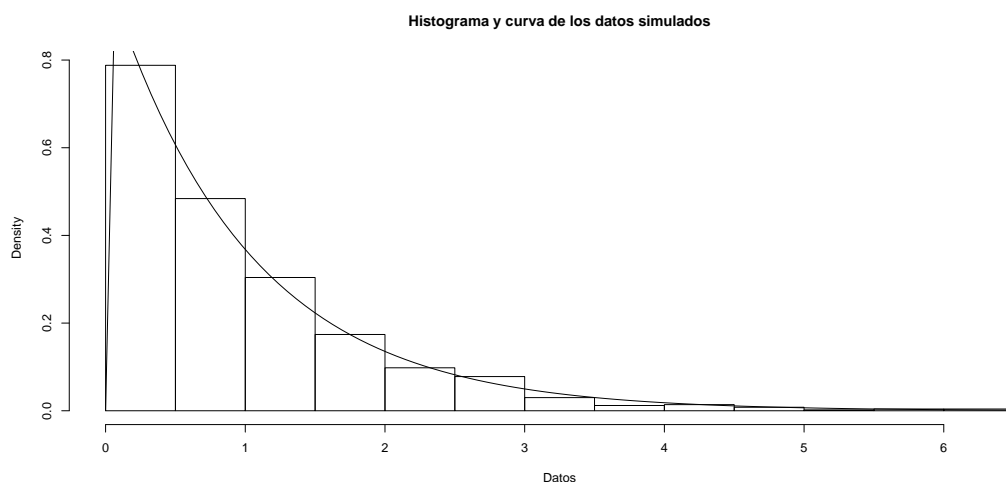


Figura 2.1: Histograma de la función cuantil evaluada en los datos generados y una curva exponencial de parámetro 1 superpuesta.

Para el segundo resultado de la proposición generamos ahora 1000 datos de una distribución exponencial con parámetro 1, evaluamos estos en su función de distribución y una vez más construimos un histograma que nos permita observar el comportamiento de los datos. Como se puede apreciar en la Figura 2.2 obtenemos barras a la misma altura que nos permiten inferir una distribución de tipo uniforme en los datos, es decir, $F(X) \sim U(0, 1)$.

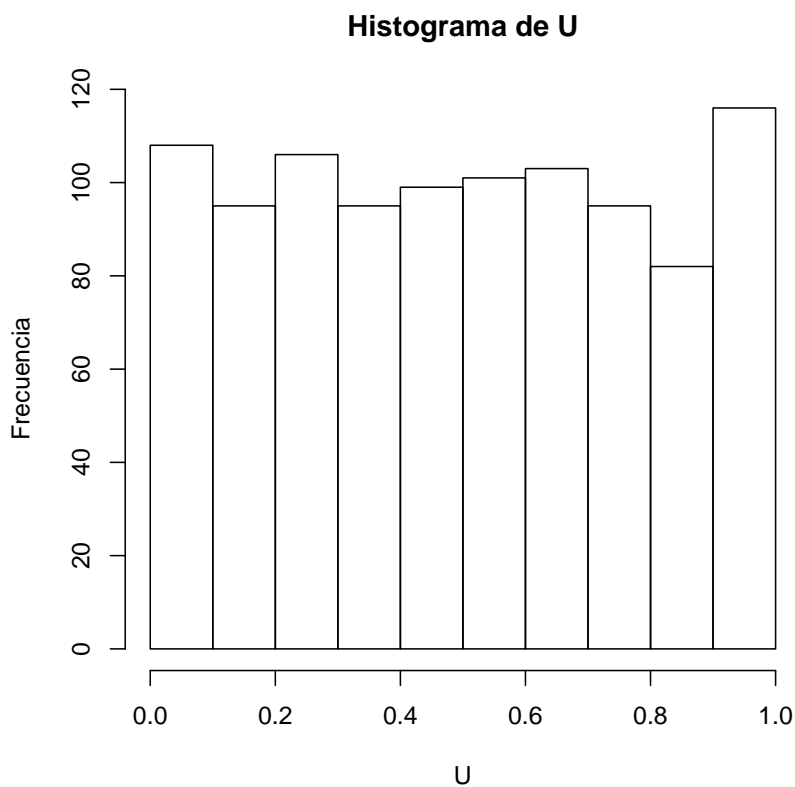


Figura 2.2: Histograma de las probabilidades de los datos generados.

2.2. Cópulas

Si bien la estructura de dependencia entre dos variables aleatorias dadas X_1, X_2 está completamente descrita por su función de distribución conjunta $F(x_1, x_2) = P[X_1 \leq x_1, X_2 \leq x_2]$, como respuesta a la idea de obtener información de dependencia dado el comportamiento marginal de las variables surge el concepto de cópula.

Si consideramos la transformación

$$\begin{aligned} T : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x_1, x_2)^t &\longmapsto (F_{X_1}(x_1), F_{X_2}(x_2))^t \end{aligned}$$

conseguimos de esta manera transformar el vector $X = (X_1, X_2)$ componente a componente en nuevas variables aleatorias que, asumiendo su continuidad se distribuyen $U(0, 1)$, es decir, son variables con distribución uniforme estándar según la Proposición (2.1.1). La función de distribución de este nuevo vector aleatorio es lo que llamaremos la cópula del vector aleatorio $(X_1, X_2)^t$ o de la función de distribución F . Lo anterior sumado a la Proposición 1.3.4.2 resulta del hecho que

$$\begin{aligned} F(x_1, x_2) &= P[X_1 \leq x_1, X_2 \leq x_2] \\ &= P[F_{X_1}(X_1) \leq F_{X_1}(x_1), F_{X_2}(X_2) \leq F_{X_2}(x_2)] \\ &= C(F_{X_1}(x_1), F_{X_2}(x_2)). \end{aligned} \tag{2.4}$$

Definición 2.2.1. (Cópula) Una cópula C es la función de distribución de un vector aleatorio sobre $[0, 1]^2$ con marginales uniformes- $(0, 1)$. Alternativamente si definimos $u := F_1(x_1)$, $v := F_2(x_2)$, una cópula es una función $C : \mathbb{I}^2 \rightarrow \mathbb{I}$ con las siguientes propiedades:

1. $C(u, v)$ es no decreciente en cada componente.
2. Para cualquier $u, v \in \mathbb{I} := [0, 1]$

$$C(u, 0) = 0 = C(0, v) \tag{2.5}$$

$$C(u, 1) = u, C(1, v) = v \tag{2.6}$$

3. Para cualesquiera $u_1, u_2, v_1, v_2 \in \mathbb{I}$ tales que $u_1 \leq u_2$ y $v_1 \leq v_2$

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0 \tag{2.7}$$

Se puede comprobar la equivalencia de las anteriores definiciones teniendo en cuenta las propiedades de las funciones de distribución por un lado y usando el hecho que las marginales de la cópula son uniformes $(0, 1)$.

El teorema que daremos a continuación puede considerarse el manantial de la teoría de cópulas y contiene la esencia del estudio de funciones de distribución multivariadas y cópulas, con su publicación en 1959 por el matemático Abe Sklar aparece por primera vez en la literatura matemática la palabra cópula. Su demostración se puede encontrar en [7, pág 21].

Teorema 2.2.1. (Teorema de Sklar) *Sea F una función de distribución conjunta bivariada con marginales F_{X_1} y F_{X_2} . Entonces existe una copula $C : [0, 1]^2 \rightarrow [0, 1]$ tal que para cualesquiera $x_1, x_2 \in \overline{\mathbb{R}}$*

$$F(x_1, x_2) = C(F_{X_1}(x_1), F_{X_2}(x_2)) \quad (2.8)$$

Si F_{X_1} y F_{X_2} son continuas, entonces C es única; en cualquier otro caso, C sólo está determinada de forma única sobre el conjunto $\text{Ran}F_{X_1} \times \text{Ran}F_{X_2}$. También, si C es una copula y F_{X_1} y F_{X_2} son funciones de distribución univariadas, entonces la función F definida mediante (2.8) es una función de distribución conjunta bivariada con marginales F_{X_1} y F_{X_2} .

Si evaluamos en (2.8) $x_i = F_{X_i}^{-1}(u_i)$, $i = \{1, 2\}$, y suponemos que las distribuciones marginales F_{X_i} , $i = \{1, 2\}$ son continuas, entonces obtenemos que

$$C(F_{X_1}(F_{X_1}^{-1}(u_1)), F_{X_2}(F_{X_2}^{-1}(u_2))) = F(F_{X_1}^{-1}(u_1), F_{X_2}^{-1}(u_2))$$

y por la Proposición 1.1.1.h podremos asegurar que

$$C(u_1, u_2) = F(F_{X_1}^{-1}(u_1), F_{X_2}^{-1}(u_2)). \quad (2.9)$$

La anterior expresión proporciona una representación explícita de la cópula C en términos de F y sus marginales, la cual será de gran utilidad más adelante cuando hablemos de métodos de construcción de cópulas.

Otra característica importante es que cualquier cópula está acotada inferior y superiormente como se verá a continuación.

Teorema 2.2.2. *Si C es una copula, entonces, para todo $(x_1, x_2) \in \mathbb{I}^2$,*

$$\max\{x_1 + x_2 - 1, 0\} \leq C(x_1, x_2) \leq \min\{x_1, x_2\} \quad (2.10)$$

Además, $C_l(x_1, x_2) = \max\{x_1 + x_2 - 1, 0\}$ y $C_u(x_1, x_2) = \min\{x_1, x_2\}$ son cópulas y se conocen como las cotas de Fréchet-Hoeffding para cópulas.

Demostración. Veamos primero que C está acotada superior e inferiormente, sea $(x_1, x_2) \in \mathbb{I}^2$, por ser C función de distribución y por tanto no decreciente, se tiene que $C(x_1, x_2) \leq C(x_1, 1) = x_1$ y $C(x_1, x_2) \leq C(1, x_2) = x_2$, se puede concluir así que $C(x_1, x_2) \leq \min\{x_1, x_2\}$. Por otro lado dado que $x_1 \leq 1$ y $x_2 \leq 1$ por las condiciones (2.6) y (2.7) tendremos que

$$\begin{aligned} C(1, 1) - C(x_1, 1) - C(1, x_2) + C(x_1, x_2) &\geq 0 \\ C(x_1, x_2) &\geq C(x_1, 1) + C(1, x_2) - C(1, 1) \\ C(x_1, x_2) &\geq x_1 + x_2 - 1 \end{aligned}$$

Además, ya que en general $C(x_1, x_2) \geq 0$, podemos asegurar que $C(x_1, x_2) \geq \max\{x_1 + x_2 - 1, 0\}$ y tendremos la primera parte del teorema.

Para demostrar que C_l y C_u son cópulas se debe tener en cuenta que si tenemos la variable $U \sim U(0, 1)$ y si llamamos a $V = 1 - U$, también $V \sim U(0, 1)$ y por tanto se puede ver que

$$C_l(x_1, x_2) = P[U \leq x_1, 1 - U \leq x_2] \quad (2.11)$$

$$C_u(x_1, x_2) = P[U \leq x_1, U \leq x_2] \quad (2.12)$$

Es decir, C_l y C_u son las funciones de distribución de los vectores $(U, 1 - U)^t$ y $(U, U)^t$ respectivamente, en la Figura (2.5) se puede apreciar las gráficas correspondientes a C_l y C_u . \square

La distribución del vector $(U, 1 - U)^t$ se concentra sobre diagonal entre los puntos $(0, 1)$ y $(1, 0)$, mientras que la de $(U, U)^t$ se concentra sobre la diagonal entre los puntos $(0, 0)$ y $(1, 1)$ como se puede ver en la Figura (2.3).

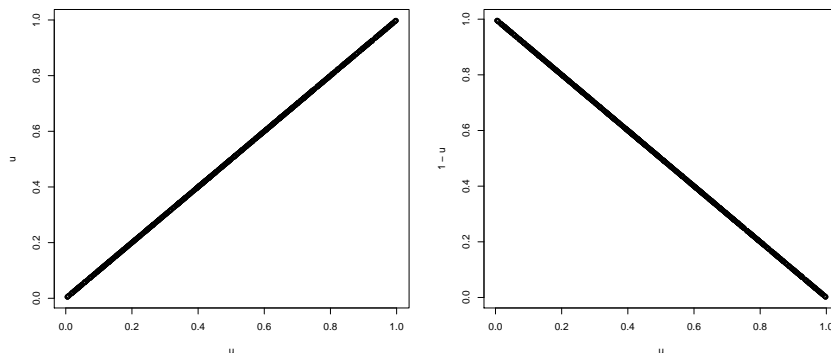


Figura 2.3: Izq: Concentración de la distribución de $(U, U)^t$. Der: Concentración de la distribución de $(U, 1 - U)^t$.

2.2.1. Ejemplos de Cópulas

A continuación se dan algunos ejemplos de cópulas, las cuales pueden ser clasificadas en fundamentales, implícitas y explícitas. Las cópulas fundamentales juegan un papel muy importante a la hora de referirse a medidas de dependencia como se verá más adelante; las cópulas implícitas se extraen de funciones de distribución conocidas usando el teorema de Sklar pero no necesariamente

están expresadas de forma sencilla; las expresiones de las cópulas explícitas son más sencillas y se pueden construir usando las características matemáticas de las cópulas.

Cópulas fundamentales.

■ Cópula de independencia:

$$\Pi(x_1, x_2) = x_1 \cdot x_2$$

Esta cópula caracteriza variables aleatorias independientes, es decir, si por un lado tenemos en cuenta el hecho que dos variables aleatorias X_1 y X_2 con marginales F_{X_1} , F_{X_2} respectivamente y función de distribución conjunta F , son independientes si y solo si $F(x_1, x_2) = F_{X_1}(x_1)F_{X_2}(x_2)$, y por otro lado consideramos la expresión (2.8), podemos afirmar entonces que X_1 y X_2 son independientes si y solo si su cópula es la cópula de independencia.

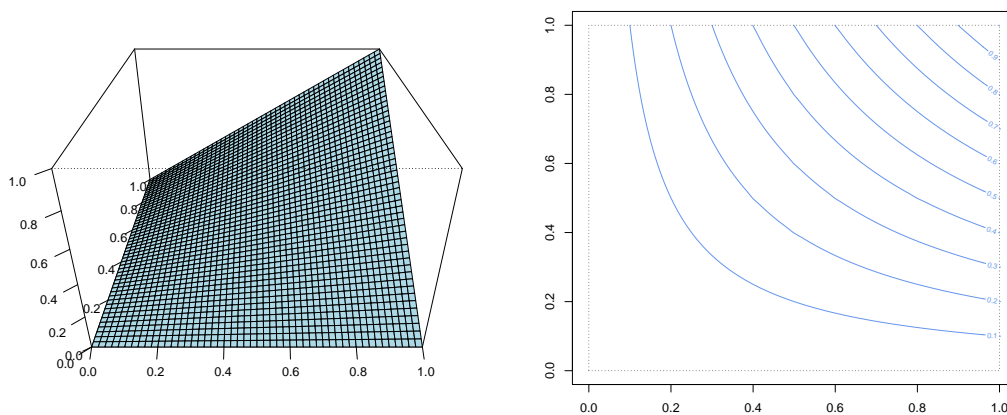


Figura 2.4: Izq: Gráfica de la cópula Π . Der:Contorno de la cópula Π .

■ Cotas Fréchet-Hoeffding

$$C_l(x_1, x_2) = \max\{x_1 + x_2 - 1, 0\}$$

$$C_u(x_1, x_2) = \min\{x_1, x_2\}$$

Este par de cópulas de las que hablamos antes describen dependencia negativa perfecta y positiva perfecta respectivamente, conceptos que examinaremos con mayor detalles más adelante.

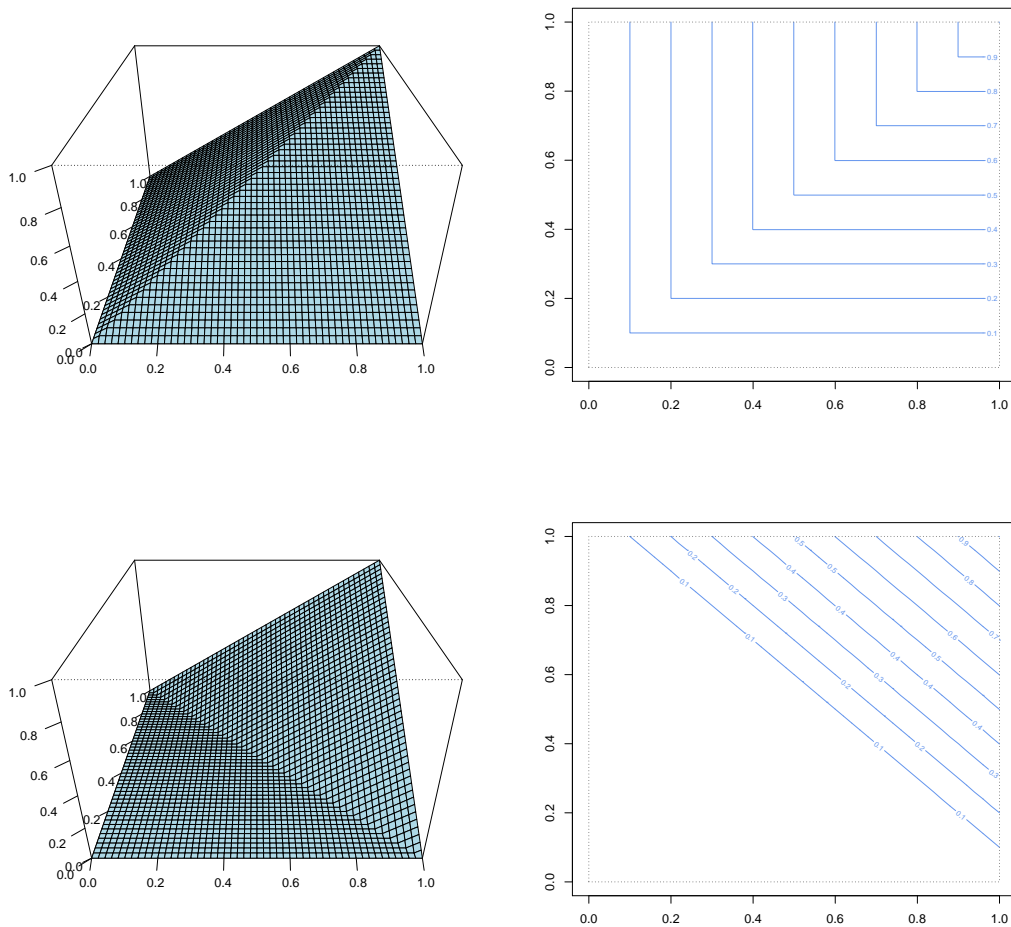


Figura 2.5: Arriba: Gráfica de la cópula C_u y su contorno. Abajo: Gráfica de la cópula C_l y su contorno

Cópulas explícitas.

- Familia de cópulas de Frank con parámetro $\theta \in \mathbb{R}, \theta \neq 0$

$$C_{\theta}^{Fr}(x_1, x_2) = -\frac{1}{\theta} \ln \left(1 + \frac{(\exp(-\theta x_1) - 1)(\exp(-\theta x_2) - 1)}{\exp(-\theta) - 1} \right) \quad (2.13)$$

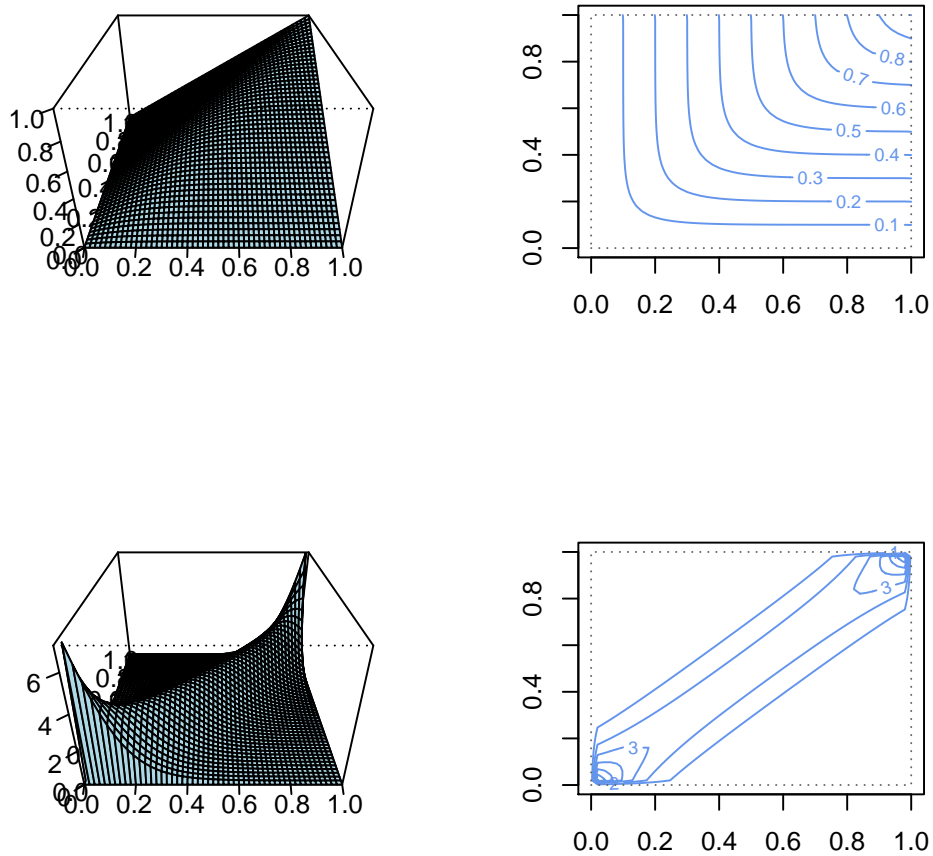


Figura 2.6: Arriba: Gráfica y contorno de la cópula de Frank de parámetro $\theta = 10$. Abajo: Gráfica y contorno de la función de densidad de la cópula de Frank de parámetro $\theta = 10$.

2.2.2. Cópula Gaussiana

La cópula Gaussiana hace parte de las cópulas implícitas y le dedicaremos especial atención porque jugará un papel fundamental en el desarrollo del objetivo del presente trabajo. Recordemos que si una variable aleatoria X tiene distribución normal estándar, entonces su función de distribución Φ está dada por:

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt.$$

Por otro lado, si el vector aleatorio $(X_1, X_2)^t$ tiene función de distribución normal estándar bivariada entonces su función de distribución conjunta es

$$\Phi_2^\rho(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{\frac{-(y_1^2 - 2\rho y_1 y_2 + y_2^2)}{2(1-\rho^2)}\right\} dy_2 dy_1,$$

donde $-1 < \rho < 1$ es el coeficiente de correlación entre las variables X_1 y X_2 . Lo anterior nos permite definir la cópula Gaussiana C_ρ^{Ga} como $C_\rho^{Ga}(u, v) = \Phi_2^\rho(\Phi^{-1}(u), \Phi^{-1}(v))$ la cual se obtiene como consecuencia del Teorema de Sklar y la expresión (2.9), es decir,

$$C_\rho^{Ga}(u, v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi(1-\rho)^{1/2}} \exp\left\{\frac{-(s^2 - 2\rho st + t^2)}{2(1-\rho^2)}\right\} ds dt, \quad (2.14)$$

por tanto, variables con función de distribución $C_\rho^{Ga}(\Phi(x), \Phi(v))$ son variables normal estándar bivariada con coeficiente de correlación ρ .

2.2.3. Métodos de Construcción de Cópulas

A continuación ilustraremos un par de métodos que permiten construir cópulas bivariadas bajo ciertas condiciones. El primero será el método de inversión, el cual es una consecuencia de (2.9) y el segundo que se describe en [3] será de gran importancia más adelante. Una larga lista de métodos de construcción geométricos y algebraicos de cópulas se puede encontrar en el Capítulo 3 de [7].

- **Método de Inversión:** este método ya se usó para definir la cópula Gaussiana y como se ha dicho consiste en aplicar la expresión (2.9), es decir, dada una función de distribución bivariada F con marginales continuas F_{X_1} y F_{X_2} siempre se tendrá que

$$C(x_1, x_2) = F(F_{X_1}^{-1}(x_1), F_{X_2}^{-1}(x_2))$$

es una cópula. El siguiente ejemplo que se encuentra en [7, pág 22] ilustra el método de inversión.

Ejemplo 2.2.1. Sea

$$F(x_1, x_2) = \begin{cases} \frac{(x_1+1)(e^{x_2}-1)}{x_1+2e^{x_2}-1}, & (x_1, x_2) \in [-1, 1] \times [0, \infty) \\ 1 - e^{-x_2}, & (x_1, x_2) \in [-1, \infty) \times [0, \infty) \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

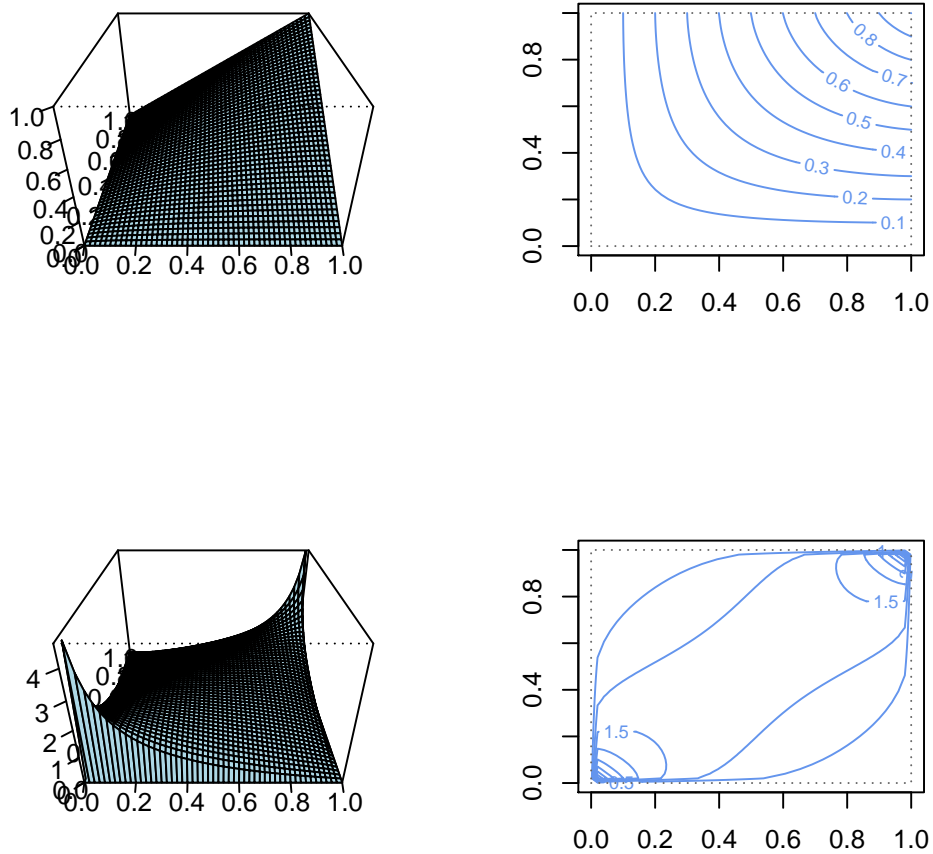


Figura 2.7: Arriba: Gráfica y contorno de la cópula Gaussiana de parámetro $\rho = 0,5$. Abajo: Gráfica y contorno de la función de densidad de la cópula Gaussiana de parámetro $\rho = 0,5$.

Sus marginales están dadas por

$$F_{X_1}(x_1) = \begin{cases} 0, & \text{si } x_1 < -1 \\ \frac{(x_1+1)}{2}, & \text{si } x_1 \in [-1, 1] \\ 1, & \text{si } x_1 > 1 \end{cases}$$

y

$$F_{X_2}(x_2) = \begin{cases} 0, & \text{si } x_2 < 0 \\ 1 - e^{-x_2}, & \text{si } x_2 \geq 0. \end{cases}$$

Por tanto las funciones cuantiles serán $F_{X_1}^{-1}(u_1) = 2u_1 - 1$ y $F_{X_2}^{-1}(u_2) =$

$-\ln(1-u_2)$ para todo $u_1, u_2 \in [0, 1]$, entonces según el método de inversión

$$\begin{aligned} F(F_{X_1}^{-1}(u_1), F_{X_2}^{-1}(u_2)) &= \frac{u_1 u_2}{u_1 + u_2 - u_1 u_2} \\ &= C(u_1, u_2) \end{aligned}$$

es una cópula.

- Por otro lado, si tenemos dos funciones $f, g : [0, 1] \rightarrow \mathbb{R}$ donde $\int_0^1 f(x_1) dx_1 = \int_0^1 g(x_2) dx_2 = 0$ y $f(x_1)g(x_2) \geq -1$ para todo $x_1, x_2 \in [0, 1]$. Podemos generar una variedad de cópulas teniendo en cuenta que $h(x_1, x_2) = 1 + f(x_1)g(x_2)$ es una función de densidad bivariada sobre $[0, 1]^2$. Por lo cual,

$$\begin{aligned} C(x_1, x_2) &= \int_0^{x_1} \int_0^{x_2} h(u_1, u_2) du_2 du_1 \\ &= \int_0^{x_1} \int_0^{x_2} 1 + f(u_1)g(u_2) du_2 du_1 \\ &= x_1 x_2 + \left(\int_0^{x_1} f(u_1) du_1 \right) \left(\int_0^{x_2} g(u_2) du_2 \right) \quad (2.15) \end{aligned}$$

es una cópula. Si por ejemplo escogemos $f(x_1) = \alpha(1 - 2x_1), g(x_2) = (1 - 2x_2), |\alpha| \leq 1$ obtenemos la cópula de Farlie-Gumbel-Morgenstern de parámetro α , $C(x_1, x_2) = x_1 x_2 [1 + \alpha(1 - x_1)(1 - x_2)]$.

2.2.4. Invariancia

Resulta conveniente en muchas situaciones ver el comportamiento de nuestros modelos probabilísticos cuando les aplicamos cierto tipo de transformaciones; veremos en la siguiente proposición cómo las cópulas son invariantes bajo transformaciones estrictamente crecientes y continuas de las marginales. En consecuencia esta será una gran ventaja frente a la propiedad de invariancia del coeficiente de correlación lineal que se cumple únicamente bajo transformaciones crecientes lineales. Además se puede demostrar que en general bajo transformaciones estrictamente monótonas y marginales continuas, los cambios en la cópula son predecibles, Ver [7, pág 26].

Proposición 2.2.3. *Si $(X_1, X_2)^t$ tiene cópula C y T_1, T_2 son funciones continuas estrictamente crecientes, entonces $(T_1(X_1), T_2(X_2))^t$ también tiene cópula C .*

Demostración. Sea C la función de distribución del vector $(U_1, U_2)^t$ (para el caso de marginales F_{X_i}), $i = 1, 2$ continuas se puede tomar $U_i = F_{X_i}(X_i)$, $i = 1, 2$). Entonces tendremos que

$$\begin{aligned}
& C(F_{T_1(X_1)}(x_1), F_{T_2(X_2)}(x_2)) \\
&= P[U_1 \leq F_{T_1(X_1)}(x_1), U_2 \leq F_{T_2(X_2)}(x_2)] \\
&= P[F_{T_1(X_1)}^{-1}(U_1) \leq x_1, F_{T_2(X_2)}^{-1}(U_2) \leq x_2] \\
&= P[(F_{X_1} \circ T_1^{-1})^{-1}(U_1) \leq x_1, (F_{X_2} \circ T_2^{-1})^{-1}(U_2) \leq x_2] \\
&= P[T_1 \circ F_{X_1}^{-1}(U_1) \leq x_1, T_2 \circ F_{X_2}^{-1}(U_2) \leq x_2] \\
&= P[T_1(X_1) \leq x_1, T_2(X_2) \leq x_2]
\end{aligned}$$

□

Observación 2.2.1. Nótese que la segunda igualdad se puede asegurar gracias a la Proposición 1.1.1.d, además la continuidad de las transformaciones T_i es necesaria para el caso general de variables aleatorias X_1 y X_2 , ya que así por la Proposición 1.1.1.b, T_i^{-1} resulta siendo estrictamente creciente, lo que sumado a las Proposiciones 1.1.1.g y 1.3.4 nos permite asegurar que

$$\begin{aligned}
F_{T_i(X_i)}(u) &= P[T_i(X_i) \leq u] \\
&= P[T_i^{-1} \circ T(X_i) \leq T_i^{-1}(u)] \\
&= P[X_i \leq T_i^{-1}(u)] \\
&= F_{X_i}(T_i^{-1}(u)) \\
&= (F_{X_i} \circ T^{-1})(u)
\end{aligned}$$

por tanto $(F_{T_i(X_i)})^{-1} = (F_{X_i} \circ T^{-1})^{-1} = T \circ F_{X_i}^{-1}$, $i = 1, 2$.

2.3. Conceptos alternativos de dependencia

En esta sección se introducen algunas nociones de conceptos sobre dependencia entre variables aleatorias con un enfoque desde la teoría de cópulas, entre ellos comonotonidad y contramonotonidad, principalmente veremos la relación de estos conceptos con la existencia de dependencia positiva perfecta o negativa perfecta entre variables aleatorias.

2.3.1. Comonotonicidad y Contramonotonicidad

Definición 2.3.1. *Se dice que las variables aleatorias X_1, X_2 son comonotónicas si $(X_1, X_2)^t$ tiene por cópula a C_u ; y se dicen contramonotónicas si $(X_1, X_2)^t$ tiene por cópula a C_l .*

Dado que no se dan condiciones sobre las distribuciones marginales de X_1 y X_2 , se hace la aclaración nuevamente en virtud del teorema de Sklar que cuando las marginales tengan discontinuidades, C_u o C_l harán parte de la familia de posibles cópulas de $(X_1, X_2)^t$, mientras que para marginales continuas estas estarán determinadas de manera única. Por otro lado siempre que las variables aleatorias X_1 y X_2 tengan distribuciones marginales continuas podremos hablar de dependencia positiva perfecta o negativa perfecta entre ellas, es decir, para el primer caso que exista una transformación creciente T tal que $X_2 = T(X_1)$ y para el segundo caso que $X_2 = T(X_1)$ cuando T sea una función decreciente.

La condición de dependencia positiva perfecta junto con la definición de cópula nos permite asegurar que la cópula C_u o cópula de comonotonicidad representa dependencia perfecta positiva entre las variables X_1 y X_2 . Lo anterior se tiene dado que si X_1 tiene distribución F_{X_1} , entonces la distribución de X_2 estará dada por $F_{X_2} = F_{T(X_1)} = F_{X_1} \circ T^{\leftarrow}$, así la cópula del vector $(X_1, X_2)^t = (X_1, T(X_1))^t$ es, por la continuidad de F_{X_1} y F_{X_2} , la función de distribución del vector

$$(F_{X_1}(X_1), F_{X_1} \circ T^{\leftarrow} \circ T(X_1)) = (U, U)^t$$

donde $U = F_{X_1}(X_1)$ y $T^{\leftarrow} \circ T(X_1) = X_1$ por la Proposición 1.1.1.g. Es decir, la cópula resulta ser la función de distribución del vector (U, U) y por tanto igual a C_u .

De manera análoga la cópula C_l o cópula de contramonotonicidad representa dependencia perfecta negativa entre las variables X_1 y X_2 como se afirma en [6, pág 190].

El siguiente teorema nos da condiciones necesarias y suficientes para establecer comonotonicidad o contramonotonicidad entre variables aleatorias, además nos permite interpretar dichas variables aleatorias como transformaciones crecientes o decrecientes de alguna otra variable aleatoria.

Teorema 2.3.1. *El vector aleatorio $(X_1, X_2)^t$ tiene una de las cópulas C_u o C_l , es decir, que para el primer caso*

$$F(x_1, x_2) = \min \{F_{X_1}(x_1), F_{X_2}(x_2)\}$$

o para el otro caso

$$F(x_1, x_2) = \text{máx} \{F_{X_1}(x_1) + F_{X_2}(x_2) - 1, 0\};$$

si y solo si existen dos funciones monótonas $u, v : \mathbb{R} \rightarrow \mathbb{R}$ y una variable aleatoria Z de valor real tal que

$$(X_1, X_2)^t =_d (u(Z), v(Z))^t, \quad (2.16)$$

con ambas crecientes en el primer caso y con u creciente y v decreciente para el otro, aquí $=_d$ denota la igualdad en las distribuciones.

Demostración. Demostraremos primero la doble equivalencia asumiendo que $(X_1, X_2)^t$ tiene cópula C_u , es decir, X_1 y X_2 son variable aleatorias comonotónicas y luego mostraremos el caso cuando sean contramonotónicas.

Sea U una variable aleatoria tal que $U \sim U(0, 1)$, si suponemos que $(X_1, X_2)^t$ tiene cópula C_u entonces tendremos de (2.12) que

$$\begin{aligned} F(x_1, x_2) &= \text{mín} \{F_{X_1}(x_1), F_{X_2}(x_2)\} \\ &= P[U \leq F_{X_1}(x_1), U \leq F_{X_2}(x_2)] \\ &= P[F_{X_1}^{-1}(U) \leq F_{X_1}^{-1} \circ F_{X_1}(x_1), F_{X_2}^{-1}(U) \leq F_{X_2}^{-1} \circ F_{X_2}(x_2)] \\ &= P[F_{X_1}^{-1}(U) \leq x_1, F_{X_2}^{-1}(U) \leq x_2] \end{aligned}$$

la última igualdad que se tiene de la Proposición 1.1.1.d nos asegura si llamamos $u = F_{X_1}^{-1}$ y $v = F_{X_2}^{-1}$ (las cuales son funciones crecientes) que

$$(X_1, X_2) =_d (u(Z), v(Z))^t.$$

Ahora supongamos que (2.16) se tiene para u y v crecientes, es decir,

$$F(x_1, x_2) = P(u(Z) \leq x_1, v(Z) \leq x_2) = P(Z \in A_1, Z \in A_2), \quad (2.17)$$

donde A_1 y A_2 son intervalos de la forma $(-\infty, \alpha)$ o $(-\infty, \alpha]$ y son elementos de \mathcal{B} , es decir, de la σ -álgebra de Borel de la cual hablamos en el anterior capítulo; se debe cumplir entonces que $A_1 \subset A_2$ o $A_2 \subset A_1$, lo cual nos lleva a concluir que,

$$F(x_1, x_2) = \text{mín}\{P(Z \in A_1), P(Z \in A_2)\} = \text{mín}\{F_1(x_1), F_2(x_2)\}$$

como se quería probar.

De manera similar se demostrará la doble equivalencia asumiendo que $(X_1, X_2)^t$

tiene cópula C_l . Sea U una variable aleatoria tal que $U \sim U(0, 1)$, así $1 - U = V \sim U(0, 1)$ y tendremos de (2.11) que

$$\begin{aligned}
F(x_1, x_2) &= \text{máx} \{F_{X_1}(x_1) + F_{X_2}(x_2) - 1, 0\} \\
&= P[U \leq F_{X_1}(x_1), 1 - U \leq F_{X_2}(x_2)] \\
&= P[F_{X_1}^{-1}(U) \leq F_{X_1}^{-1} \circ F_{X_1}(x_1), F_{X_2}^{-1}(1 - U) \leq F_{X_2}^{-1} \circ F_{X_2}(x_2)] \\
&= P[F_{X_1}^{-1}(U) \leq x_1, F_{X_2}^{-1}(1 - U) \leq x_2] \\
&= P[F_{X_1}^{-1}(U) \leq x_1, F_{X_2}^{-1} \circ g(U) \leq x_2], \quad g(x) = 1 - x.
\end{aligned}$$

también la penúltima igualdad se tiene por la Proposición 1.1.1.d y se cumple así la primera implicación tomando como función creciente a $u := F_{X_1}^{-1}$ y como función decreciente $v := F_{X_2}^{-1} \circ g$.

Recíprocamente si (2.16) se tiene para u creciente y v decreciente entonces también tendremos la expresión (2.17), donde $A_1 := \{Z \in u^{-1}((-\infty, x_1])\}$ y $A_2 := \{Z \in v^{-1}((-\infty, x_2])\}$. Si $A_1 \cap A_2 \neq \emptyset$ entonces

$$P[A_1 \cup A_2] = P[\Omega] = 1 = P[A_1] + P[A_2] - P[A_1 \cap A_2]$$

debido a la monotonicidad de u y v . Por tanto se tendrá que

$$P[A_1 \cap A_2] = P[u(Z) \leq x_1, v(Z) \leq x_2] = F_{X_1}(x_1) + F_{X_2}(x_2) - 1.$$

Por otro lado si $A_1 \cap A_2 = \emptyset$, entonces

$$F_{X_1}(x_1) + F_{X_2}(x_2) - 1 \leq 0.$$

En cualquiera de los casos siempre se cumplirá que

$$P[u(Z) \leq x_1, v(Z) \leq x_2] = \text{máx}\{F_{X_1}(x_1) + F_{X_2}(x_2) - 1, 0\}$$

y se completa así la demostración. □

Una consecuencia inmediata del anterior teorema cuando se tiene la continuidad de F_{X_1} y F_{X_2} es la siguiente:

Corolario 1. Sean X_1, X_2 variables aleatorias con funciones de distribución continuas F_{X_1} y F_{X_2} respectivamente, entonces

$$\begin{aligned}
C = C_l &\iff X_2 = T(X_1), \text{ donde } T = F_{X_2}^{-1} \circ (1 - F_{X_1}) \text{ es una función decreciente,} \\
C = C_u &\iff X_2 = T(X_1), \text{ donde } T = F_{X_2}^{-1} \circ F_{X_1} \text{ es una función creciente.}
\end{aligned}$$

En el Capítulo 1 se definió la covarianza entre dos variables aleatorias con varianzas finitas y se dieron algunas de sus propiedades, otra forma de hallar la covarianza se da en el siguiente teorema que establece la que se conoce como identidad de Höfding.

Teorema 2.3.2. *Si $(X_1, X_2)^t$ tiene función de distribución conjunta F y marginales F_1 y F_2 , entonces la covarianza de X_1 y X_2 siempre que sea finita está dada por*

$$\text{cov}\{(X_1, X_2)^t\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (F(x_1, x_2) - F_1(x_1)F_2(x_2)) dx_1 dx_2 \quad (2.18)$$

Demostración. Ver [6, pág 203]. □

2.3.2. Propiedades deseadas de las medidas de dependencia

Las medidas de dependencia como correlación lineal, por ejemplo, expresan la dependencia entre dos variables aleatorias en un número. A continuación se darán algunas propiedades con las que deben cumplir ciertas medidas. Consideremos una medida de dependencia $\delta(\cdot, \cdot)$ que asigna un número real a cualquier par de variables aleatoria de valor real X_1 y X_2 . Se espera que esta medida cumpla las siguientes propiedades:

P.1 $\delta(X_1, X_2) = \delta(X_2, X_1)$ (Simetría).

P.2 $-1 \leq \delta(X_1, X_2) \leq 1$ (Normalización).

P.3 $\delta(X_1, X_2) = 1 \Leftrightarrow X_1, X_2$ son comonotónicas y $\delta(X_1, X_2) = -1 \Leftrightarrow X_1, X_2$ son contramonotónicas.

P.4 Si $T : \mathbb{R} \rightarrow \mathbb{R}$ es una función estrictamente monótona sobre el rango de X_1 :

$$\delta(T(X_1), X_2) = \begin{cases} \delta(X_1, X_2), & \text{si } T \text{ es creciente} \\ -\delta(X_1, X_2), & \text{si } T \text{ es decreciente} \end{cases} \quad (2.19)$$

Si nos fijamos el coeficiente de correlación satisface únicamente las condiciones 1 y 2. Las anteriores propiedades pueden ser restringidas o extendidas, por ejemplo una propiedad que sería de gran utilidad agregar es

P.5 $\delta(X_1, X_2) = 0 \Leftrightarrow X_1, X_2$ son independientes.

Sin embargo, las propiedades P.4 y P.5 son incompatibles tal como lo veremos en la siguiente proposición.

Proposición 2.3.3. *Ninguna medida de dependencia satisface las propiedades P.4 y P.5 simultáneamente.*

Demostración. Supongamos que una medida δ satisface P.4 y veamos que no puede satisfacer P.5. Sea $\mathbf{X} = (X_1, X_2)^t$ un vector aleatorio uniformemente distribuido sobre el círculo unitario en \mathbb{R}^2 , es decir,

$$f_{\mathbf{X}}(x_1, x_2) = \begin{cases} \frac{1}{2\pi}, & \text{si } x_1^2 + x_2^2 = 1 \\ 0, & \text{en cualquier otro caso} \end{cases} \quad (2.20)$$

por tanto podemos asegurar por el Teorema 1.4.3 que $(X_1, X_2)^t = (Y_1, Y_2)^t = \mathbf{Y}$, donde $Y_1 = \cos \varphi$, $Y_2 = \sin \varphi$ y $\varphi \sim U(0, 2\pi)$. Además, $(-X_1, X_2)^t =_d (X_1, X_2)^t$, por tanto por P.4 tendremos que

$$\delta(-X_1, X_2) = \delta(X_1, X_2) = -\delta(X_1, X_2).$$

Lo anterior implica que $\delta(X_1, X_2) = 0$ aun cuando X_1 y X_2 son dependientes, dada la identidad $\cos \varphi = \sin(\varphi + \frac{\pi}{2})$. \square

Si se deseara incluir la propiedad 5, tendríamos entonces que considerar únicamente aquellas medidas que tomen valores positivos para cualquier par de variables aleatorias.

Capítulo 3

La Falacia

La distribución normal bivariada es una distribución con comportamiento ideal en el sentido que, sus marginales resultan ser normales al igual que sus condicionales. Un problema que surge con bastante frecuencia en cursos básicos de probabilidad es pensar en un resultado recíproco a la afirmación anterior, esto es, proponer una distribución con marginales normales cuya distribución bivariada no sea normal. Una forma más general de dar respuesta a este problema, planteado en [7, pág 23], acude a la teoría de cópulas y se plantea a manera de falacia¹ de una manera más general como sigue.

Falacia 1. *Distribuciones marginales y correlación determinan de manera única la distribución conjunta.*

En el Capítulo 1 nos encontramos con una forma de hallar las marginales de una función de distribución bivariada por medio de un límite, esto a su vez nos permite encontrar el coeficiente de correlación. Sin embargo, si conocemos las distribuciones marginales de dos variables aleatorias X_1 y X_2 y su correspondiente coeficiente de correlación, entonces existen infinitas distribuciones posibles para el vector $(X_1, X_2)^t$ y lo ilustraremos con el siguiente ejemplo tomado de [3].

Sean X_1, X_2 variables aleatorias con distribuciones normal estándar $\Phi(x_1)$ y $\Phi(x_2)$ respectivamente, sea también $\rho(X_1, X_2) = \rho$. Entonces si $(X_1, X_2)^t$ se distribuye normal estándar su función de distribución F está dada según la expresión (2.14) por

$$F(x_1, x_2) = C_\rho^{Ga}(\Phi(x_1), \Phi(x_2)).$$

¹Entiéndase como una proposición cuyo argumento parece válido sin serlo.

Cualquier otra cópula $C \neq C_\rho^{Ga}$ proporciona una distribución bivariada con marginales que tienen distribución normal estándar la cual no es normal bivariada con correlación ρ . En efecto, construiremos una cópula C como en (2.15) definiendo f y g del siguiente modo:

$$f(x_1) = \mathbf{1}_{\{(\gamma, 1-\gamma)\}}(x_1) + \frac{2\gamma - 1}{2\gamma} \mathbf{1}_{\{(\gamma, 1-\gamma)^c\}}(x_1)$$

$$g(x_2) = -\mathbf{1}_{\{(\gamma, 1-\gamma)\}}(x_2) - \frac{2\gamma - 1}{2\gamma} \mathbf{1}_{\{(\gamma, 1-\gamma)^c\}}(x_2)$$

donde $\frac{1}{4} \leq \gamma \leq \frac{1}{2}$ y

$$\mathbf{1}_{\{(a,b)\}}(\alpha) = \begin{cases} 1, & \text{si } \alpha \in (a, b) \\ 0, & \text{si } \alpha \notin (a, b). \end{cases}$$

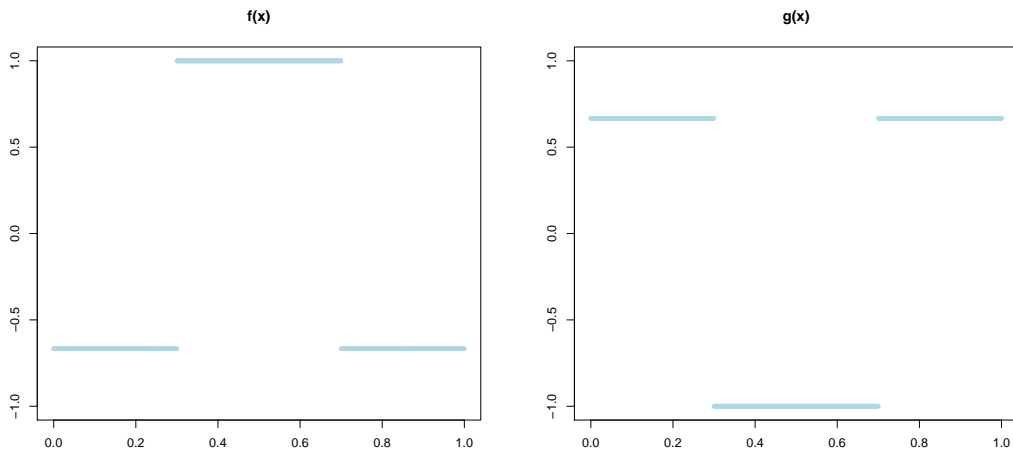


Figura 3.1: Gráfica de las funciones f y g con parámetro $\gamma = 0,3$

Las gráficas de f y g para $\gamma = 0,3$ se ilustran en la Figura (3.1) y en este caso $h(x_1, x_2) = 1 + f(x_1)g(x_2)$ es una función de densidad, pues f y g satisfacen las condiciones requeridas, en efecto,

- $f, g : [0, 1] \longrightarrow \mathbb{R}$.

- $\int_0^1 f(x_1)dx_1 = 0$, pues

$$\begin{aligned}
\int_0^1 f(x_1)dx_1 &= \int_0^1 \left[\mathbf{1}_{\{(\gamma, 1-\gamma)\}}(x_1) + \frac{2\gamma-1}{2\gamma} \mathbf{1}_{\{(\gamma, 1-\gamma)^c\}}(x_1) \right] dx_1 \\
&= \int_0^\gamma \left[\mathbf{1}_{\{(\gamma, 1-\gamma)\}}(x_1) + \frac{2\gamma-1}{2\gamma} \mathbf{1}_{\{(\gamma, 1-\gamma)^c\}}(x_1) \right] dx_1 \\
&\quad + \int_\gamma^{1-\gamma} \left[\mathbf{1}_{\{(\gamma, 1-\gamma)\}}(x_1) + \frac{2\gamma-1}{2\gamma} \mathbf{1}_{\{(\gamma, 1-\gamma)^c\}}(x_1) \right] dx_1 \\
&\quad + \int_{1-\gamma}^1 \left[\mathbf{1}_{\{(\gamma, 1-\gamma)\}}(x_1) + \frac{2\gamma-1}{2\gamma} \mathbf{1}_{\{(\gamma, 1-\gamma)^c\}}(x_1) \right] dx_1 \\
&= \left(\frac{2\gamma-1}{2\gamma} \right) x_1 \Big|_0^\gamma + x_1 \Big|_\gamma^{1-\gamma} + \left(\frac{2\gamma-1}{2\gamma} \right) x_1 \Big|_{1-\gamma}^1 \\
&= \left(\frac{2\gamma-1}{2\gamma} \right) \gamma + (1-\gamma-\gamma) + \frac{2\gamma-1}{2\gamma} - \frac{2\gamma-1}{2\gamma}(1-\gamma) \\
&= 2\gamma \left(\frac{2\gamma-1}{2\gamma} \right) + 1 - 2\gamma \\
&= 0.
\end{aligned}$$

- Análogamente $\int_0^1 g(x_2)dx_2 = 0$, pues

$$\begin{aligned}
\int_0^1 g(x_2)dx_2 &= \int_0^1 \left[-\mathbf{1}_{\{(\gamma, 1-\gamma)\}}(x_2) - \frac{2\gamma-1}{2\gamma} \mathbf{1}_{\{(\gamma, 1-\gamma)^c\}}(x_2) \right] dx_2 \\
&= \int_0^\gamma \left[-\mathbf{1}_{\{(\gamma, 1-\gamma)\}}(x_2) - \frac{2\gamma-1}{2\gamma} \mathbf{1}_{\{(\gamma, 1-\gamma)^c\}}(x_2) \right] dx_2 \\
&\quad + \int_\gamma^{1-\gamma} \left[-\mathbf{1}_{\{(\gamma, 1-\gamma)\}}(x_2) - \frac{2\gamma-1}{2\gamma} \mathbf{1}_{\{(\gamma, 1-\gamma)^c\}}(x_2) \right] dx_2 \\
&\quad + \int_{1-\gamma}^1 \left[-\mathbf{1}_{\{(\gamma, 1-\gamma)\}}(x_2) - \frac{2\gamma-1}{2\gamma} \mathbf{1}_{\{(\gamma, 1-\gamma)^c\}}(x_2) \right] dx_2 \\
&= - \left(\frac{2\gamma-1}{2\gamma} \right) x_2 \Big|_0^\gamma - x_2 \Big|_\gamma^{1-\gamma} - \left(\frac{2\gamma-1}{2\gamma} \right) x_2 \Big|_{1-\gamma}^1 \\
&= - \left(\frac{2\gamma-1}{2\gamma} \right) \gamma - (1-\gamma-\gamma) - \frac{2\gamma-1}{2\gamma} + \frac{2\gamma-1}{2\gamma}(1-\gamma) \\
&= -2\gamma \left(\frac{2\gamma-1}{2\gamma} \right) - 1 + 2\gamma \\
&= 0.
\end{aligned}$$

- Además $f(x_1)g(x_2) \geq -1$ para todo $x_1, x_2 \in [0, 1]$, para verlo consideremos los siguientes casos:

1. $x_1 \in (\gamma, 1-\gamma)$ y $x_2 \in (\gamma, 1-\gamma)^c$

2. $x_1 \in (\gamma, 1 - \gamma)$ y $x_2 \in (\gamma, 1 - \gamma)$
3. $x_1 \in (\gamma, 1 - \gamma)^c$ y $x_2 \in (\gamma, 1 - \gamma)$
4. $x_1 \in (\gamma, 1 - \gamma)^c$ y $x_2 \in (\gamma, 1 - \gamma)^c$

Para cada uno de los casos tendremos que $f(x_1)g(x_2)$ es igual a las funciones $-\frac{2\gamma-1}{2\gamma}$, -1 , $-\frac{2\gamma-1}{2\gamma}$ y $-\frac{(2\gamma-1)^2}{4\gamma^2}$ respectivamente, las cuales son mayores o iguales a -1 , en efecto si suponemos que $-\frac{2\gamma-1}{2\gamma} < -1$ entonces,

$$\begin{aligned}\frac{2\gamma-1}{2\gamma} &> 1 \\ 2\gamma-1 &> 2\gamma\end{aligned}$$

lo cual es una contradicción.

También si suponemos que $-\frac{(2\gamma-1)^2}{4\gamma^2} < -1$ entonces,

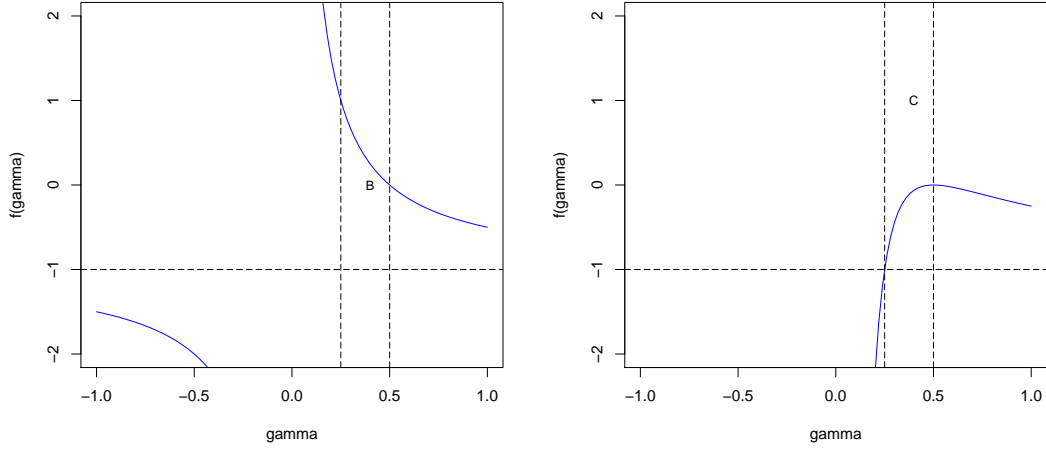
$$\begin{aligned}\frac{(2\gamma-1)^2}{4\gamma^2} &> 1 \\ (2\gamma-1)^2 &> 4\gamma^2 \\ 4\gamma^2 - 4\gamma + 1 &> 4\gamma^2 \\ 4\gamma &< 1 \\ \gamma &< \frac{1}{4}\end{aligned}$$

esto contradice el hecho que $1/4 \leq \gamma$ y por tanto las funciones $-\frac{2\gamma-1}{2\gamma}$ y $-\frac{(2\gamma-1)^2}{4\gamma^2}$ son mayores o iguales a -1 . Lo anterior también se puede ver en la Figura (3.2).

La cópula definida por $h(x_1, x_2)$ con dominio $\Phi(x_1)$ y $\Phi(x_2)$ es

$$\begin{aligned}C(\Phi(x_1), \Phi(x_2)) &= \int_0^{\Phi(x_1)} \int_0^{\Phi(x_2)} h(u_1, u_2) du_2 du_1 \\ &= \Phi(x_1)\Phi(x_2) + \left(\int_0^{\Phi(x_1)} f(u_1) du_1 \right) \left(\int_0^{\Phi(x_2)} g(u_2) du_2 \right).\end{aligned}$$

Del teorema de Sklar sabemos que $C(\Phi(x_1), \Phi(x_2)) = F(x_1, x_2)$, donde F es una función de distribución con marginales $\Phi(x_1)$ y $\Phi(x_2)$, es decir que su función


 (a) Gráfica de $-\frac{2\gamma-1}{2\gamma}$

 (b) Gráfica de $-\frac{(2\gamma-1)^2}{4\gamma^2}$

Figura 3.2: Las regiones B y C nos permiten comprobar que para $\frac{1}{4} \leq \gamma \leq \frac{1}{2}$ las funciones dadas son mayores o iguales a -1 .

de densidad conjunta $f(x_1, x_2)$ está dada por:

$$\begin{aligned}
 f(x_1, x_2) &= \frac{\partial^2}{\partial x_1 \partial x_2} F(x_1, x_2) \\
 &= \frac{\partial^2}{\partial x_1 \partial x_2} \left[\Phi(x_1)\Phi(x_2) + \left(\int_0^{\Phi(x_1)} f(u_1) du_1 \right) \left(\int_0^{\Phi(x_2)} g(u_2) du_2 \right) \right] \\
 &= \Phi'(x_1)\Phi'(x_2) + f(\Phi(x_1))\Phi'(x_1)g(\Phi(x_2))\Phi'(x_2) \\
 &= \Phi'(x_1)\Phi'(x_2)[1 + f(\Phi(x_1))g(\Phi(x_2))] \\
 &= \Phi'(x_1)\Phi'(x_2) + h(\Phi(x_1), \Phi(x_2))
 \end{aligned}$$

donde

$$\Phi'(x_1)\Phi'(x_2) = \frac{1}{2\pi} \exp \frac{-(x_1^2 + x_2^2)}{2}.$$

La función $h(x, y)$ desaparece sobre el cuadrado $[\gamma, 1 - \gamma]^2$, esto hace que $f(x_1, x_2)$ también lo haga en el intervalo $[\Phi^{-1}(\gamma), \Phi^{-1}(1 - \gamma)]^2$, es decir, en el centro como se ve en la Figura (3.3); por tanto C para $\gamma < \frac{1}{2}$ y $F(x, y) = C(\Phi(x), \Phi(y))$ no será una distribución normal bivariada. Por otro lado para cada $1/4 \leq \gamma \leq 1/2$ obtendremos una función de distribución distinta, es decir, las marginales dadas determinan infinitas funciones de distribución conjunta. Partiendo de las consideraciones de simetría ($C(u, v) = C(1 - u, v)$), $0 \leq u, v \leq 1$; la correlación

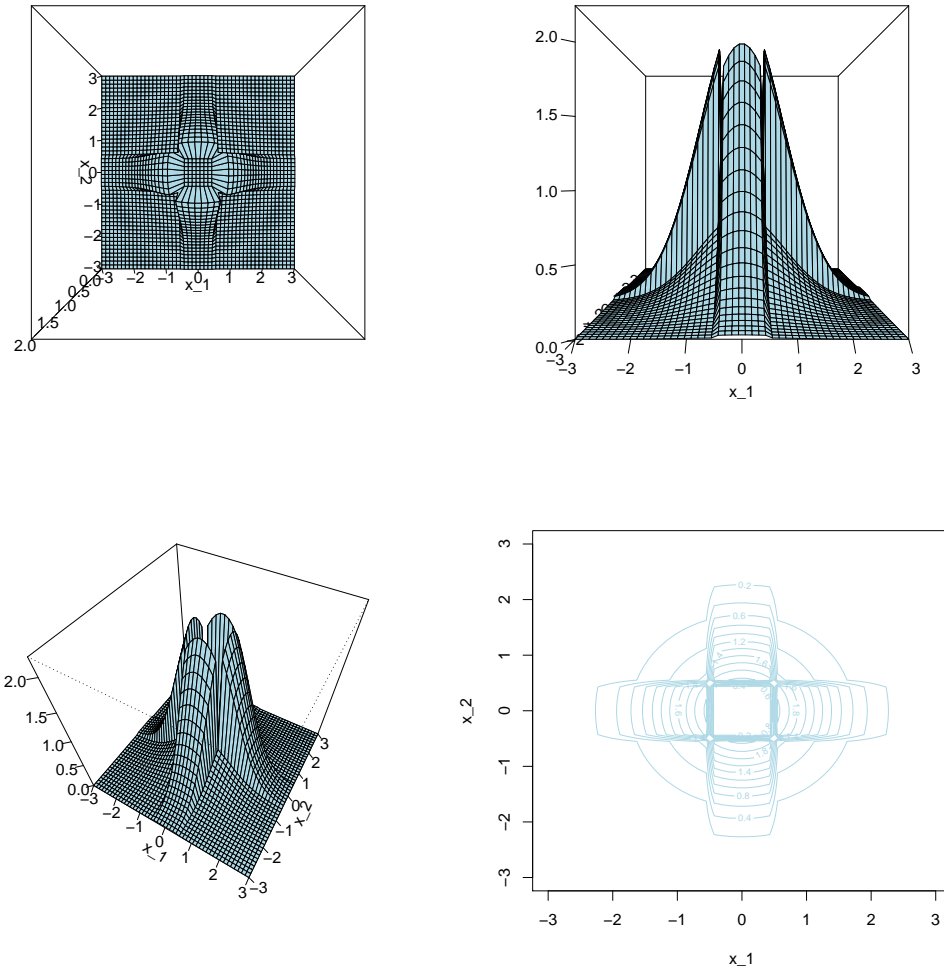


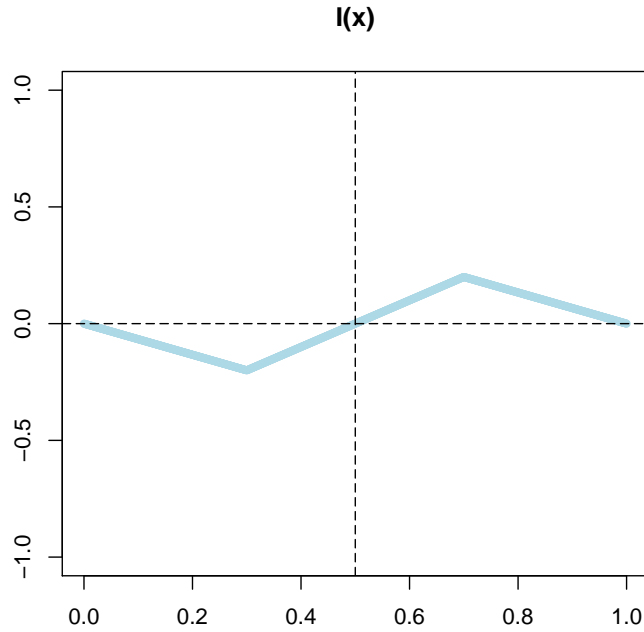
Figura 3.3: Densidad y contorno de una distribución que no es normal bivariada pero que tiene marginales normal estándar.

independientemente de γ es cero, en efecto

$$C(x, y) = xy - \left(\int_0^x f(u) du \right) \left(\int_0^y f(v) dv \right)$$

dado que $g(x) = -f(x)$. Considerando la función I definida sobre $[0, 1]$ por

$$\begin{aligned} I(x) &:= \int_0^x f(u) du \\ &= \int_0^x \left[\mathbf{1}_{\{(\gamma, 1-\gamma)\}}(u) + \frac{2\gamma - 1}{2\gamma} \mathbf{1}_{\{(\gamma, 1-\gamma)^c\}}(u) \right] du, \end{aligned}$$

Figura 3.4: Gráfica de la función I con $\gamma = 0,3$.

explícitamente estará dada por

$$I(x) = \begin{cases} \left(\frac{2\gamma-1}{2\gamma}\right) x, & \text{si } 0 \leq x < \gamma \\ \frac{2\gamma-1}{2} + x - \gamma, & \text{si } \gamma \leq x < 1 - \gamma \\ \left(\frac{2\gamma-1}{2\gamma}\right) (x - 1), & \text{si } 1 - \gamma \leq x \leq 1, \end{cases} \quad (3.1)$$

esta función posee una simetría alrededor de $1/2$ por la forma en que hemos definido f como se puede ver en la Figura 3.4 y por tanto satisface que

$$I(1-x) = \int_0^{1-x} f(u)du = \int_0^1 f(u)du - \int_0^x f(u)du = -\int_0^x f(u)du = -I(x)$$

para todo $x \in (0, 1/2)$.

Lo anterior implica que el coeficiente de correlación es 0 dado que según la

expresión (2.18)

$$\begin{aligned}
Cov\{(X_1, X_2)^t\} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (C(\Phi(x_1), \Phi(x_2)) - \Phi(x_1)\Phi(x_2)) dx_1 dx_2 \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[- \left(\int_0^{\Phi(x_1)} f(u_1) du_1 \right) \left(\int_0^{\Phi(x_2)} f(u_2) du_2 \right) \right] dx_1 dx_2 \\
&= \left[- \int_{-\infty}^{\infty} \int_0^{\Phi(x_1)} f(u_1) du_1 dx_1 \right] \left[\int_{-\infty}^{\infty} \int_0^{\Phi(x_2)} f(u_2) du_2 dx_2 \right] \\
&= \left[- \int_{-\infty}^{\infty} (I(\Phi(x_1))) dx_1 \right] \left[\int_{-\infty}^{\infty} (I(\Phi(x_2))) dx_2 \right]
\end{aligned}$$

Estas últimas integrales se anulan debido a la simetría de la función I alrededor de $1/2$, por tanto se concluye que $Cov\{(X_1, X_2)^t\} = 0$ y de ahí que $\rho = 0$ independientemente de γ .

Ejemplos que complementan esta falacia se pueden encontrar en [5].

Capítulo 4

Conclusiones

- Gracias a la teoría de cópulas se logra ilustrar teórica y gráficamente que las distribuciones marginales de un par de variables aleatorias y su coeficiente de correlación lineal no determinan de manera única la función de distribución conjunta, este hecho evidencia la necesidad de ser cuidadosos a la hora de construir modelos multivariados dados los comportamientos marginales.
- En el escenario de las medidas de dependencia se plantea la necesidad de desarrollar medidas más sofisticadas que el coeficiente de correlación lineal de Pearson, ya que en ocasiones el escalar obtenido no nos proporciona toda la información que nos interesa de un vector aleatorio, también según las propiedades que se esperan tener de una medida de dependencia, este presenta desventajas, pues por un lado variables no correlacionadas en general no son independientes y, por otro lado el coeficiente de correlación no es invariante bajo transformaciones crecientes no lineales.
- El teorema de Sklar constituye una herramienta fundamental en la teoría de Cópulas, dentro de sus bondades resaltamos la facilidad que brinda para construir modelos multivariados dado el comportamiento marginal y su utilidad para obtener información sobre la estructura de dependencia de un vector aleatorio, la cual se encuentra implícita en su función de distribución conjunta.
- El software R permitió la visualización de muchos resultados presentados durante todo el trabajo, principalmente el paquete “copula” facilitó manipular gráficamente los ejemplos de los tres tipos de cópulas presentados:

fundamentales, explícitas e implícitas. Por otro lado la ilustración de la falacia se hizo más notable gracias a que se pudo mostrar gráficamente que la función de densidad obtenida no era una normal estándar bivariada considerando la región en la que se anulaba.

Bibliografía

- [1] BARTLE, R. *Introduction to Real Analysis*. John Wiley & Sons, Inc, USA, 2011.
- [2] BLANCO, L., ARUNACHALAM, V., AND DHARMARAJA, D. *Introduction to Probability and Stochastic Processes with Applications*. John Wiley & Sons, Inc., USA, 2012.
- [3] EMBRECHTS, P., MCNEIL, A., AND STRAUMANN, D. Correlation and dependence in risk management: properties and pitfalls. *Risk management: value at risk and beyond* (2002), 176–223.
- [4] ERDELY, A. Cópulas y dependencia de variables aleatorias: Una introducción. *Miscelánea Matemática 48* (2009), 7–28.
- [5] KOWALSKI, C. Non-normal bivariate distributions with normal marginals. *Taylor & Francis 27* (1973), 103–106.
- [6] MCNEIL, A., FREY, R., AND EMBRECHTS, P. *Quantitative Risk Management: Concepts, techniques and tools*. Princeton university press, New Jersey, 2005.
- [7] NELSEN, R. *An introduction to copulas*. Springer, USA, 2006.
- [8] WANG, S., AND DHAENE, J. Comonotonicity, correlation order and premium principles. *Insurance: Mathematics and Economics 22* (1998), 235–242.
- [9] YOHAJ, V. *Notas de Probabilidades y Estadística*. 2008.

Apéndice A

Códigos en R

Figura (2.1):

```
> u<-runif(1000)
> x<-qexp(u, rate=1) # el comando "qexp" simula la función
cuantil que antes hallamos
> hist(x, freq=F, xlab="Datos", main="Histograma y curva de
los datos simulados")
> curve(dexp(x), add=T)
```

Figura (2.2) en R:

```
> de<-rexp(1000, rate=1)
> U<-pexp(de,rate=1)
> hist(U, main="Histograma de U", ylab="Frecuencia")
```

Figura (2.4):

```
> library(copula)
> prodcop <- normalCopula(0)
> par(mfrow=c(1, 2))
> persp(prodcop, pCopula, main= , theta = 0, phi = 30, col=
"lightblue", xlab=, ylab=, zlab=)
> contour(prodcop,pCopula,theta = 0, phi = 30, col=
"cornflowerblue", xlab=, ylab=, zlab= )
```

Figura (2.5):

```
> library(copula)
> mincop <- normalCopula(1)
> par(mfrow=c(1, 2))
> persp(mincop, pCopula, main= , theta = 0, phi = 30, col=
"lightblue", xlab=, ylab=, zlab=)
> contour(mincop,pCopula,theta = 0, phi = 30, col=
"cornflowerblue", xlab=, ylab=, zlab= )
> maxcop <- normalCopula(-1)
> par(mfrow=c(1, 2))
> persp(maxcop, pCopula, main= , theta = 0, phi = 30, col=
"lightblue", xlab=, ylab=, zlab=)
> contour(maxcop,pCopula,theta = 0, phi = 30, col=
"cornflowerblue", xlab=, ylab=, zlab= )
```

Figura (2.6):

```
> library(copula)
> cfrank <- frankCopula(10)
> par(mfrow=c(2, 2))
> persp(cfrank, pCopula, theta = 0, phi = 30, col=
"lightblue", xlab=, ylab=, zlab=)
> contour(cfrank,pCopula,theta = 0, phi = 30, col=
"cornflowerblue", xlab=, ylab=, zlab= )
> persp(cfrank, dCopula, theta = 0, phi = 30, col=
"lightblue", xlab=, ylab=, zlab=)
> contour(cfrank,dCopula,theta = 0, phi = 30, col=
"cornflowerblue", xlab=, ylab=, zlab= )
```

Figura (2.7):

```
> library(copula)
> norm.cop <- normalCopula(0.5)
> par(mfrow=c(2, 2))
> persp(norm.cop, pCopula, theta = 0, phi = 30, col=
"lightblue", xlab=, ylab=, zlab=)
> contour(norm.cop,pCopula,theta = 0, phi = 30, col=
"cornflowerblue", xlab=, ylab=, zlab= )
> persp(norm.cop, dCopula, theta = 0, phi = 30, col=
"lightblue", xlab=, ylab=, zlab=)
> contour(norm.cop,dCopula,theta = 0, phi = 30, col=
"cornflowerblue", xlab=, ylab=, zlab= )
```