



UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE ESTUDIOS SUPERIORES
ACATLÁN

MODELO DE DISTRIBUCIÓN CONJUNTA
DEL ÍNDICE DE PRECIOS Y
COTIZACIONES Y EL TIPO DE CAMBIO
DÓLAR ESTADOUNIDENSE-PESO
MEXICANO PARA LA CREACIÓN DE
INSTRUMENTOS DERIVADOS BASADOS
EN SU DEPENDENCIA.

T E S I S

QUE PARA OBTENER EL TÍTULO DE
A C T U A R I O

PRESENTA:

ENRIQUE LECUONA SIERRA

TUTOR DE TESIS:

DR. ARTURO ERDELY RUÍZ



SANTA CRUZ ACATLÁN, NAUCALPAN,
ESTADO DE MÉXICO, 2023



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.

Agradecimientos

De su versión en inglés, “If I have seen further, it is by standing on the shoulders of giants”, o de su versión en español “Si he visto más lejos, es poniéndome sobre los hombros de gigantes” es una frase que comúnmente se utiliza para describir como el descubrimiento de nuevo conocimiento se alcanza en gran medida gracias a la aportación de conocimiento ya existente. Dicha frase se atribuye comúnmente a Issac Newton que en 1675, en una carta entregada a Robert Hooke, acepta una invitación para iniciar una conversación privada entre ambas partes para compartir, discutir y criticar los estudios y análisis que ambos científicos estaban haciendo en la época.

Para mí el conocimiento científico es como un círculo que está en constante expansión, crece de forma irregular y este crecimiento solo se alcanza mediante lo que ya se conoce. La expansión, tan pequeña como sea, requiere de un esfuerzo titánico por parte del que investiga, requiere de tiempo, esfuerzo, y dedicación completa al proyecto en cuestión.

La frase de Issac Newton es cierta, el conocimiento que se crea depende de forma crucial del punto de partida que se toma a la hora de investigar, de la fuente bibliográfica, del artículo de investigación, de la comparativa y contraste entre distintos autores. Pero considero que los “gigantes” no terminan ahí, ahí terminan los gigantes que son mencionados en el anexo final de este escrito, pero espero este espacio permita denotar de forma clara a mis “gigantes personales”, los cuales fueron tan fundamentales en la creación de este escrito como cualquier otro.

Primeramente agradezco a mis padres Amalia y Enrique y a mi hermana Mary, por ser el pilar más importante en mi vida. También agradezco a mis padrinos Luis e Hilda, por ser un apoyo fundamental en el transcurso de mi vida y en mi desarrollo personal y profesional. Asimismo, agradezco a mis amigos Dani, Rafa, May e Ivette por su apoyo incondicional. Para concluir, agradezco en gran medida al Dr. Arturo Erdely Ruíz por su apoyo para la creación de este escrito, a mis sínodos por su

II

paciencia y dedicación a la hora de revisar este escrito y a la Universidad Nacional Autónoma de México por darme una educación profesional de la mas alta calidad.

¡Muchas gracias!

Dedicatoria

A mi abuelita Raquel
y a mi primo Miguel.
Se les recuerda siempre
con cariño.

Índice general

Resumen	VI
Introducción	VIII
1. Cópulas	1
1.1. Antecedentes históricos de la teoría de cópulas	1
1.2. Introducción probabilística al concepto de cópula	3
1.3. El Teorema de Sklar y sus consecuencias	35
1.4. Simulación y regresión mediante cópulas	46
1.4.1. Simulación mediante cópulas	46
1.4.2. Regresión mediante cópulas	50
1.5. Cópulas Arquimedianas	53
1.6. Medidas de concordancia y dependencia	62
1.6.1. Ejemplo de medida de concordancia: Rho de Spearman	62
1.6.2. Ejemplo de medida de dependencia: Sigma de Schweizer	67
1.7. Tail Dependence	70
1.8. Gluing Copulas	74
2. Instrumentos Financieros Derivados	82
2.1. Antecedentes históricos de la teoría de Derivados	82
2.2. Procesos estocásticos y esperanza condicional	86

2.3. El Modelo Binomial	96
2.3.1. El modelo binomial a un periodo	96
2.3.2. El modelo binomial multiperiodo	103
2.4. El modelo a tiempo continuo	114
2.5. Cálculo estocástico	127
2.6. Sobre una extensión al Cálculo de Itô.	142
3. Modelos discretos de valuación y resultados	159
3.1. Enfoque no paramétrico	159
3.1.1. Distribución empírica univariada	160
3.1.2. Distribución y Cópula empírica bivariada	164
3.2. Activos financieras a tiempo discreto	175
3.2.1. Distribución discreta del subyacente	175
3.2.2. Distribución discreta del derivado	177
3.2.3. Distribución discreta del portafolio de replicación	178
3.3. Análisis de datos y resultados	187
3.3.1. Ajuste de la información	187
3.3.2. Derivados propuestos y resultados	196
Conclusiones	213
Bibliografía	218

Resumen

Los derivados financieros son productos financieros cuyo valor depende del valor de otro activo financiero (ya sea una acción o una tasa de interés, por ejemplo), dicho activo financiero del cual depende el derivado es denominado subyacente. Las implementaciones que tienen los derivados financieros en el mercado financiero se pueden englobar en las siguientes dos grandes ramas:

1. Operaciones de cobertura, también denominadas como operaciones de *hedging*: Protección ante los cambios de otros activos financieros (principalmente cambios en el subyacente).
2. Operaciones de especulación: Se invierte en derivados financieros con el objetivo de generar ganancias en base a su incertidumbre.

Este trabajo de investigación se enfocará en la creación de un modelo de distribución conjunta semiparamétrico del Índice de Precios y Cotizaciones y el tipo de cambio dólar estadounidense-peso mexicano con el objetivo de crear y valorar un derivado financiero que cuente como subyacentes dichos indicadores financieros.

Para la creación del modelo, se utiliza el concepto de cópulas perteneciente al área de la estadística. Una vez creado el modelo de distribución conjunta, se utilizan los resultados clásicos de la teoría de derivados financieros como base para la creación de instrumentos financieros los cuales se benefician del tipo de dependencia existente entre los subyacentes.

Se mostrará a detalle el modelo de distribución conjunta escogido, así como el proceso para valorar y replicar los derivados financieros propuestos, todo esto fundamentado por el marco teórico.

Por último, se establecerá que tanto se benefician los derivados financieros propuestos con la dependencia que existe entre el Índice de Precios y Cotizaciones y el tipo de cambio dólar estadounidense-peso mexicano.

La implementación computacional de este proyecto, se realizó en su totalidad en el Lenguaje de Programación Julia, para más información al respecto de dicho lenguaje véase <https://julialang.org/> (método de instalación, funcionamiento general, etc.) así como <https://docs.julialang.org/en/v1/> para la documentación general del lenguaje y Bezanson, Edelman, Karpinski, y Shah (2017) para el artículo de investigación inicial que dio la creación de este lenguaje. Se realizó una paquetería la cual contiene todo el código original del proyecto, así como las funciones implementadas para los ajustes paramétricos y semiparamétricos de los datos, dicha paquetería está disponible en la página <https://github.com/EnriqueLecuona1998/TesisEnriqueLecuona.jl>.

Introducción

El Capítulo 1 está conformado por 8 secciones las cuales establecen los resultados básicos de la teoría de cópulas. Se inicia con antecedentes históricos en la materia y con la estructura teórica fundamental de dichas funciones. Luego, se desarrollan resultados relacionados con regresión y simulación mediante cópulas y se introduce la familia arquimediana de cópulas, la cual será esencial a la hora de la construcción del modelo semiparamétrico propuesto. A continuación se establecen medidas de concordancia y dependencia, así como el concepto de *tail dependence*. Se termina el capítulo estableciendo el concepto de *gluing copulas*, y como estas ayudan a la construcción de cópulas asimétricas mediante cópulas simétricos.

Posteriormente, el Capítulo 2 está conformado por 6 secciones las cuales establecen la teoría clásica de valuación de derivados financieros. Se inicia con antecedentes históricos en la materia y con la construcción del modelo binomial multiperiodo, estableciendo las propiedades de martingala que existen como consecuencia de la valuación neutral al riesgo. Luego, se establece la relación del modelo discreto con el modelo continuo como consecuencia inmediata del Teorema Central del Límite. Después, se establece el resultado del modelo Black-Sholes-Merton para la valuación de opciones europeas clásicas. Se termina el capítulo con una propuesta de extensión del modelo Black-Sholes-Merton a uno más general en el cual los log-rendimientos de las acciones no siguen una distribución normal, y se plantean consecuencias inmediatas de esta generalización en términos de valuación libre de riesgo.

Por último, el Capítulo 3 está conformado por 3 secciones y en este se desarrolla el modelo semiparamétrico propuesto y se ajusta a los datos analizados. Dicho modelo está compuesto por una cópula paramétrica ajustada mediante metodologías basadas en las pruebas de bondad y ajuste clásicas, y con distribuciones empíricas univariadas. El capítulo inicia estableciendo resultados clásicos de estadística no paramétrica para el caso univariado y bivariado. Luego, se establece el modelo de valuación a tiempo discreto para activos financieros (acciones y derivados de acciones), esto se hace en

el caso univariado y bivariado. Finalmente, se hace un ajuste y valuación práctico de la cópula paramétrica a los datos, así como la valuación y comparación del derivado propuesto con la valuación tradicional dada por el modelo Black-Scholes-Merton. Se realiza una comparación de resultados y se concluye en que momentos es adecuado el modelo propuesto por encima del modelo tradicional y viceversa.

Capítulo 1

Cópulas

En la Sección 1, se dará una introducción histórica al concepto de cópulas, así como una mención rápida de los primeros resultados de dicha área, principalmente utilizando información de Dall’Aglío (1991), Nelsen (2006) y Schweizer (1991).

De la Sección 2 en adelante, se escribirán y desarrollarán los conceptos básicos de la teoría de cópulas, describiendo sus características y las propiedades que sean de interés para este escrito. Esta sección estará basada en Embrechts *et. al.* (2001), Erdely (2009), Nelsen (2006) y el apéndice de Salvadori *et al.* (2007).

1.1. Antecedentes históricos de la teoría de cópulas

El concepto de cópula en la teoría de la estadística tiene sus orígenes en artículos científicos relacionados con el término en inglés “*Distributions with given marginals*”. Dichos artículos tenían como enfoque el crear modelos de distribución multivariados con distribuciones marginales fijas y que además cumplieran condiciones específicas, ya sea que se cumpliera una condición específica en los momentos de la distribución como en Dall’Aglío (1956), en el aspecto de las tablas de correlación establecidas en Fréchet (1951, 1957) o en el aspecto de distribuciones multivariadas con cotas superiores como en Dall’Aglío (1961) y Kellerer (1961).

El desarrollo de esta teoría inicio con el artículo de Fréchet (1951) publicado en *Annales de l’ Universite de Lyon* con los desarrollos citados anteriormente como

ejemplos, y esta etapa preliminar termino con Sklar (1959) en donde se estableció formalmente el concepto de cópula, en el ámbito de la teoría estadística así como las propiedades básicas de dichas funciones.

Tal y como dice Nelsen (2006), la palabra cópula es un sustantivo que proviene del latín cuyo significa es “un enlace, unión”, mientras que en un contexto gramatical de acuerdo a Sklar (1996), la palabra cópula es “un término o expresión que enlaza un sujeto y predicado”. Por lo cual, el término de cópula es muy adecuado para su función en la teoría de la estadística, ya que su objetivo principal, en este contexto, es el de unir distribuciones de probabilidad n -dimensionales con sus distribuciones marginales respectivas, es decir, la cópula es el *enlace* que existe entre dichas funciones.

El término cópula fue utilizado por primera vez en un sentido matemático en un trabajo escrito por Abe Sklar en 1959, donde en respuesta a una consulta por parte de M. Fréchet, Sklar introdujo el concepto de cópula. De acuerdo a Sklar (1996):

Ferón [1956], in studying three-dimensional distributions had introduced auxiliary functions, defined on the unit cube, that connected such distributions with their one-dimensional margins. I saw that similar functions could be defined on the n -cube for all $n \geq 2$ and would similarly serve to link n -dimensional distributions to their one-dimensional margins. Having worked out the basic properties of these functions, I wrote about them to Fréchet, in English. He asked me to write a note about them in French[. . .] Fréchet received my note, corrected one mathematical statement, made some minor corrections to my French, and had the note published by the Statistical Institute of the University of Paris as Sklar (1959). Subsequent developments are summarized in Schweizer (1991).

y durante el periodo de 1959 a 1974, los mayores resultados relacionados con cópulas fueron obtenidos en el desarrollo de la teoría de espacios métricos de probabilidad, ejemplos de esto se puede encontrar en Moynihan y Schweizer (1979); Schweizer y Sklar (1960) y de forma estructurada se pueden encontrar en el libro de Schweizer y Sklar (1983).

A partir de 1974, se puede ver un gran avance en la utilización de las cópulas bivariadas como herramientas para crear medidas de dependencia no paramétricas para pares de variables aleatorias, ejemplos de dichos desarrollos se pueden ver en Schweizer y Wolff (1981) donde se describe lo que hoy se conoce como la σ de Schweizer, o en Vitale (1990) donde se utiliza el concepto de “fixing marginals” en relación

a convergencias en distribución para vectores finitos de variables aleatorias. Todo esto nos hace ver como las áreas de uso que tiene las cópulas en la teoría estadística y en la teoría de la probabilidad no son pocas, y son de amplia utilidad cuando se manejan vectores aleatorios y se desea cuantificar de alguna forma su dependencia sin tomar en cuenta que modelos univariados se están empleando en su construcción.

Por último, es de extremo interés para este trabajo mencionar textos que implementen a las cópulas como alternativas a modelos multivariados clásicos, y entre muchos otros más, podemos mencionar el trabajo de Trivedi y Zimmer (2005) en donde se da una introducción al concepto de cópulas así como su estimación mediante métodos de máxima verosimilitud; en el ámbito de finanzas nos podemos referir a Cherubini *et al.* (2012; 2004) donde se utiliza a las cópulas en distintos ámbitos como lo son el riesgo de crédito, aplicaciones econométricas o en valuación de opciones financieras, y ciertos capítulos de Franke *et al.* (2008) que se enfocan en análisis de portafolios de inversión y su valor en riesgo; concluimos citando a Deniuit *et al.* (2005) y su trabajo en el ámbito de la teoría del riesgo que es de especial interés para el ámbito actuarial.

1.2. Introducción probabilística al concepto de cópula

Esta sección se enfocará en los siguientes dos puntos:

- 1) Plantear los conceptos básicos de la teoría de la probabilidad, para esta parte se utilizará principalmente Erdely (2017).
- 2) Plantear los conceptos y resultados básicos de la teoría de cópulas. Para esta parte se empleará en gran medida el capítulo 2 de Nelsen (2006) y Embrechts *et al.* (2001).

Para iniciar esta sección se necesita plantear resultados básicos de la teoría de la probabilidad. Para esto, es necesario contar con un fenómeno o experimento aleatorio que sea de interés analizar, y la estrategia que proporciona la teoría de la probabilidad para poder realizar análisis a estos fenómenos o experimentos aleatorios es mediante el concepto de un *espacio de probabilidad*.

Definición 1.1. *Un espacio de probabilidad es una tripleta $(\Omega, \mathcal{F}, \mathbb{P})$ donde,*

- (i) Ω es un conjunto no vacío cualquiera.
- (ii) \mathcal{F} es una σ -álgebra de subconjuntos de Ω .
- (iii) \mathbb{P} es una función $\mathbb{P} : \mathcal{F} \mapsto [0, 1]$ que cumple con las siguientes dos propiedades.
- a) $\mathbb{P}(\Omega) = 1$.
- b) La función \mathbb{P} es σ -aditiva, es decir, para una colección $\{A_n\}_{n=1}^{\infty}$ de conjuntos distintos a pares ($A_i \cap A_j = \emptyset$; $i \neq j$), se cumple que

$$\mathbb{P}\left(\biguplus_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n) \quad (1.1)$$

La definición anterior establece las propiedades matemáticas fundamentales que debe cumplir un espacio de probabilidad, pero es de gran importancia detallar el significado que estos conceptos tienen en un contexto más simple. Para un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ se tiene que Ω representa el conjunto de resultados posibles del fenómeno o experimento aleatorio que se desea cuantificar y se le denomina *espacio muestral*; \mathcal{F} representa el conjunto de eventos o resultados que son de interés a considerar en relación a dicho fenómeno o experimento aleatorio, y se le denomina *espacio de eventos*; para cada evento de interés (es decir, $\forall A \in \mathcal{F}$) se asigna un valor en el intervalo $[0, 1]$ que denota que tan posible o probable el evento puede ocurrir, la función que hace esta correspondencia es \mathbb{P} y la denominamos *medida de probabilidad*.

Para una descripción más profunda de las implicaciones que la terna $(\Omega, \mathcal{F}, \mathbb{P})$ tiene en la teoría de la probabilidad, véase a Klenke (2008). Cabe mencionar, que la creación de la terna $(\Omega, \mathcal{F}, \mathbb{P})$ es esencial a la hora que planteamos cualquier modelo probabilístico, tal y como dice Rincón (2007):

El objetivo es asociar un espacio de probabilidad al experimento aleatorio de interés. No existen reglas establecidas para ello y además la posible asignación no es única, pues dependiendo del interés del observador, se puede asociar un espacio de probabilidad u otro.

Con lo cual, es importante siempre tener bien establecida la terna $(\Omega, \mathcal{F}, \mathbb{P})$ antes de realizar cualquier análisis probabilístico.

Una vez creado la tripleta $(\Omega, \mathcal{F}, \mathbb{P})$, se puede cuantificar los resultados de los eventos de interés ($A \in \mathcal{F}$) de la forma deseada. Una de las primeras dificultades

que se tienen a la hora de realizar cálculo de probabilidades, es que el espacio de eventos Ω es un conjunto cualquiera, por lo que puede resultar difícil de manejar o analizar. Para poder resolver este inconveniente, se introduce el concepto de *variable aleatoria*.

Definición 1.2. Una variable aleatoria X (por simpleza se denota como v.a.) es una función $X : \Omega \mapsto \mathbb{R}$ la cual para todo conjunto abierto A en los reales, cumple que

$$X^{-1}(A) \in \mathcal{F} \quad (1.2)$$

Si se cumple esto, se dice que X es \mathcal{F} -medible.

Sin entrar a mucho detalle, se desea que la función X cumpla estas propiedades ya que así es posible *trasladar* la información disponible en $(\Omega, \mathcal{F}, \mathbb{P})$ a una nueva tripleta $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_X)$ la cual también se busca sea espacio de probabilidad pero que además cuente con todas las bondades que están definidas y establecidas en los números reales, como por ejemplo, los resultados existentes en el Cálculo Integral y Diferencial¹.

A continuación, es necesario crear una conexión entre las funciones \mathbb{P} y \mathbb{P}_X , sabiendo que \mathbb{P}_X debe ser también medida de probabilidad, la forma tradicional de hacer esto es la siguiente:

Definición 1.3. Decimos que \mathbb{P}_X es la medida de probabilidad inducida por la v.a. X . Y para todo $B \in \mathcal{B}(\mathbb{R})$, la función \mathbb{P}_X queda definida como sigue:

$$\mathbb{P}_X(B) := \mathbb{P}(X^{-1}(B)) \equiv \mathbb{P}(X \in B) \quad (1.3)$$

gracias a la fórmula 1.3, notamos que la propiedad de medibilidad pedida a X es fundamental².

Consecuentemente, es de interés crear una nueva función que no dependa directamente de $\mathcal{B}(\mathbb{R})$ y que nos permita realizar cálculo de probabilidades de forma mas sencilla, para esto, definimos la *función de distribución* de la variable aleatoria X como sigue:

Definición 1.4. Sea X una v.a. cualquiera, se denota como F_X a la función de distribución de la v.a. X , y queda definida para todo $x \in \mathbb{R}$ como sigue:

$$F_X(x) := \mathbb{P}_X\left(]-\infty, x]\right) \equiv \mathbb{P}(X \leq x) \quad (1.4)$$

¹Para resultados importantes en el área del cálculo diferencial e integral véase, entre mucho otros, a Spivak (1996).

²Cierta literatura se refiere a \mathbb{P}_X como la *ley de X* , por ejemplo, Klenke *op. cit.*.

el objetivo de esta función es el de facilitar el análisis de \mathbb{P}_X , ya que se puede demostrar³ que todo elemento $A \in \mathcal{B}(\mathbb{R})$ se puede descomponer en una colección, a lo más numerable, de conjuntos de la forma $] - \infty, t]$; $t \in \mathbb{R}$, por lo cual, se crea la siguiente relación

$$\mathbb{P} \longleftrightarrow \mathbb{P}_X \longleftrightarrow F_X \quad (1.5)$$

y gracias a 1.1 y 1.3, se puede desarrollar cualquier cálculo de probabilidades en $(\Omega, \mathcal{F}, \mathbb{P})$ mediante la función F_X . Para cualquier conjunto $]a, b[\subseteq \mathbb{R}$; se crea por conveniencia la notación siguiente:

$$\mathbb{P}_X(]a, b[) \equiv \mathbb{P}\left(X \in]a, b[\right) \equiv \mathbb{P}(a < X < b) \quad (1.6)$$

y esta notación se puede extender a cualquier conjunto $]a, b[\subseteq \mathbb{R}$ (donde $|$ puede denotar abierto o cerrado) de forma equivalente.

Las variables aleatorias las podemos clasificar en 3 tipos distintos dependiendo de las características que tiene su función de distribución F_X

Definición 1.5. *Sea X una v.a. cualquiera y sea \mathcal{D}_X el conjunto de discontinuidades de su función de distribución F_X . Dependiendo de las características del conjunto \mathcal{D}_X denotamos a la v.a. como sigue:*

- (i) *Si $\mathcal{D}_X = \emptyset$ entonces decimos que X es una variable aleatoria continua (denotado como v.a. continua).*
- (ii) *Si $\mathcal{D}_X \neq \emptyset$ entonces existen dos posibilidades:*
 - a) *Si $\sum_{x \in \mathcal{D}_X} \mathbb{P}(X = x) = 1$ entonces decimos que X es una variable aleatoria discreta (denotado como v.a. discreta).*
 - b) *Si $\sum_{x \in \mathcal{D}_X} \mathbb{P}(X = x) < 1$ entonces decimos que X es una variable aleatoria mixta (denotado como v.a. mixta).*

si la v.a. es discreta, entonces \mathcal{D}_X es un conjunto a lo mas numerable⁴, y por características de la función F_X ⁵, se cumple para todo $x \in \mathbb{R}$ lo siguiente,

$$\mathbb{P}(X = x) = \lim_{\delta \rightarrow 0^-} \mathbb{P}(x + \delta < X \leq x) = F_X(x) - \lim_{\delta \rightarrow 0^-} F_X(x + \delta)$$

³Para la demostración, véase Rincón (2007).

⁴Rincón, *op. cit.*

⁵Rincón, *op. cit.*

es decir, la probabilidad en el punto x es equivalente al salto que hace la función de distribución en dicho punto. Para este tipo de variables aleatorias, se define la función de masa de probabilidades como sigue,

Definición 1.6. *Sea X una v.a. discreta cualquiera, se denota la función de masa de probabilidades de la v.a. X como p_X y la definimos como sigue para todo $x \in D_X$*

$$p_X(x) := \mathbb{P}(X = x)$$

para este caso, por las propiedades de una función de distribución, el conjunto D_X es a lo más numerable, por lo que toda v.a. discreta tiene a lo más un conjunto numerable de puntos de masa.

Para el caso de variables aleatorias continuas, se tiene el siguiente caso particular muy importante,

Definición 1.7. *Sea X una v.a. continua, si se cumple que la función de distribución de la v.a. X es diferenciable en todo \mathbb{R} y además existe una función $f_X : \mathbb{R} \mapsto \mathbb{R}$ continua, tal que*

$$(i) \ f_X \geq 0.$$

(ii) *Se cumple que para todo $x \in \mathbb{R}$,*

$$f_X(x) = \frac{d}{dx} F_X(x)$$

entonces se dice que f_X es la función de densidad de probabilidades para la v.a. X , y además, se dice que X es una variable aleatoria absolutamente continua (denotado como v.a. abs. continua).

Gracias a la definición 1.7 y al Teorema Fundamental del Cálculo Diferencial e Integral⁶, tenemos que F_X y f_X cumplen la siguiente condición para toda $x \in \mathbb{R}$,

$$F_X(x) = \int_{-\infty}^x f_X(t) dt$$

La definición 1.7 es muy importante, ya que existen muchas v.a. absolutamente continuas que están definidas solamente mediante su función de densidad, ya que su

⁶Existe mucha literatura en relación a este resultado pero, por ejemplo, véase Spivak, *op. cit.*

función de distribución no tiene fórmula explícita (ejemplo de esto es la distribución normal), por lo cual conocer la relación entre F_X y f_X es esencial.

Por último, en el caso de una v.a. mixta, se puede implementar el siguiente resultado,

Teorema 1.8 (Teorema de descomposición para variables aleatorias). *Sea X una v.a. cualquiera definida sobre un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, entonces existe una v.a. continua X_C y una v.a. discreta X_D definidas en el mismo espacio de probabilidad, y un número $\alpha \in [0, 1]$, que cumplen la siguiente igualdad,*

$$F_X(x) = \alpha F_{X_C}(x) + (1 - \alpha) F_{X_D}(x) \quad (1.7)$$

en donde F_{X_C} y F_{X_D} denotan las funciones de distribución de las variables aleatorias X_C y X_D , respectivamente⁷.

esto es de especial interés para variables aleatorias mixtas, ya que nos permite *descomponerlas* en v.a. que son mas sencillas de trabajar.

Una vez se tiene la v.a. X y su respectiva función de distribución F_X , resulta de interés conocer como sería la función de distribución de la variable aleatoria Y que representa una *transformación* de interés para la información contenida en X , esto se denota por $Y := g(X)$ y se tiene la siguiente relación,

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}\left(\{\omega \in \Omega : g(X(\omega)) \leq y\}\right) \quad (1.8)$$

en los casos que la función g sea estrictamente creciente o estrictamente decreciente, se tienen las siguientes equivalencias para 1.8,

(i) Si g es estrictamente creciente:

$$F_Y(y) = \mathbb{P}(X \leq g^{-1}(y)) = F_X(g^{-1}(y))$$

(ii) Si g es estrictamente decreciente:

$$F_Y(y) = \mathbb{P}(X \geq g^{-1}(y)) = 1 - \lim_{\delta \rightarrow 0^-} F_X(g^{-1}(y) + \delta)$$

⁷Para mas información al respecto, véase Domínguez (2001).

para otros casos existen fórmulas establecidas⁸, pero el objetivo de este escrito no es profundizar en estos desarrollos.

Uno de los objetivos que tiene la estadística, es la *predicción* de los fenómenos o experimentos aleatorios que se están considerando, esto se puede hacer mediante *estimaciones puntuales* que normalmente buscan reflejar un punto medio o central en nuestra incertidumbre. Esto tradicionalmente se realiza con el concepto de *esperanza*. Dicho concepto se define a continuación,

Definición 1.9. *Sea X una v.a. cualquiera, se define la esperanza de la variable aleatoria X como la siguiente integral*

$$\mathbb{E}(X) := \int_0^{\infty} (1 - F_X(x)) dx - \int_{-\infty}^0 F_X(x) dx \quad (1.9)$$

si la integral anterior existe, se dice que la v.a. X tiene esperanza finita. Si la integral anterior es infinita se dice que X tiene esperanza no finita. Por último, si la anterior integral no existe se dice que X no tiene esperanza definida.

la definición anterior de esperanza no es la dada en libros avanzados de teoría de la probabilidad, pero para el objetivo y alcance de este escrito es más que suficiente, si se desea profundizar en la construcción rigurosa de la esperanza, véase Klenke *op. cit.*

Existen fórmulas equivalentes para 1.9 en los casos que X sea una v.a. absolutamente continua o discreta, las fórmulas equivalentes son las siguientes, respectivamente,

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx \quad (\text{v.a. abs. continua}) \quad (1.10)$$

$$E(X) = \sum_{x \in \mathcal{D}_X} x \mathbb{P}(X = x) \quad (\text{v.a. discreta}) \quad (1.11)$$

Es importante mencionar que para la variable aleatoria $Y = g(X)$, podemos obtener su esperanza (siempre y cuando exista) de dos formas distintas; una es mediante la ecuación 1.9 (obteniendo primeramente F_Y), y la otra es mediante la siguiente ecuación,

$$\mathbb{E}(Y) = \mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(u) dF_X(u) \quad (1.12)$$

⁸Por ejemplo en Rincón, *op. cit.*

donde la ecuación 1.12 es una *integral de Riemann-Stieltjes*⁹. Si para la fórmula 1.12 tenemos que F_X es diferenciable (tal como es el caso en una v.a. abs. continua) se cumple que,

$$\mathbb{E}(Y) = \int_{-\infty}^{\infty} g(u) f_X(u) du \quad (1.13)$$

lo cual es equivalente a 1.10 cuando g es la función identidad.

Aunque existen resultados probabilísticos muy importantes en relación a la esperanza¹⁰, no es la única estimación puntual que podemos hacer en relación a una v.a. X , existe otra denominada *mediana* y se define a continuación,

Definición 1.10. *Sea X una v.a. cualquiera, se define la mediana de la variable aleatoria X como el número (si existe) $\mathbb{M}(X) \in \mathbb{R}$ tal que*

$$\mathbb{P}(X \leq \mathbb{M}(X)) = \frac{1}{2} = \mathbb{P}(X > \mathbb{M}(X)) \quad (1.14)$$

la mediana (como medida de tendencia central), tiene ciertas ventajas con respecto a la esperanza, como son las siguientes

- (i) Se puede notar por la definición 1.14, que la mediana representa claramente un punto medio en la incertidumbre de X .
- (ii) La mediana existe para toda v.a. abs. continua, en comparación a la esperanza que deja de existir para v.a. abs. continuas con colas muy pesadas (ejemplo de esto es la distribución t de Student con ν grados de libertad para $0 < \nu < 1$).
- (iii) Existen casos en los que aunque la v.a. tiene esperanza finita, dicho valor no representa un valor central de ningún tipo y se puede hacer tan grande o pequeña como la deseemos.

Tenemos casos particulares de funciones g de especial interés; ejemplos de esto es la función $g_X(x) = (x - \mathbb{E}(X))^2$, con esto definimos la *varianza* de la v.a. X ,

Definición 1.11. *Sea una variable aleatoria X cualquiera, se define la varianza de la variable aleatoria X (cuando X tiene esperanza finita) de la siguiente forma,*

$$\mathbb{V}(X) := \mathbb{E}((X - \mathbb{E}(X))^2) \equiv \mathbb{E}(X^2) - \mathbb{E}^2(X) \quad (1.15)$$

⁹Para un desarrollo rápido de dicha integral, véase Rincón *op. cit.*

¹⁰Ejemplo de esto es el **Ley de los grandes números**. Véase, entre otros, Klenke *op. cit.*

si la integral anterior existe, se dice que la v.a. X tiene varianza finita. Si la integral anterior es infinita se dice que X tiene varianza no finita. Por último, si la anterior integral no existe se dice que X no tiene varianza definida.

con esto se termina la introducción básica a la teoría de probabilidad, a continuación se plantea el concepto de dependencia y como este está ligado estrechamente al concepto de cópula.

Cuando se está analizando un fenómeno o experimento aleatorio cualquiera, se plantea un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ para poder realizar un análisis del mismo. Una relación de interés a considerar es la ocurrencia (o supuesta ocurrencia) de un evento $B \in \mathcal{F}$ y como este puede afectar la probabilidad de ocurrencia de otro evento $A \in \mathcal{F}$, esto se denomina probabilidad condicional y queda definida como sigue,

Definición 1.12. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad cualquiera y sea $B \in \mathcal{F}$ un evento tal que $\mathbb{P}(B) > 0$; para cualquier $A \in \mathcal{F}$ se define la probabilidad condicional de A dado B mediante la siguiente ecuación,

$$\mathbb{P}_B(A) \equiv \mathbb{P}(A | B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \quad (1.16)$$

Notesé que para $B \in \mathcal{F}$ tal que $\mathbb{P}(B) > 0$, la función \mathbb{P}_B es también una medida de probabilidad admisible para el par (Ω, \mathcal{F}) ; por lo que la función \mathbb{P}_B genera una *actualización* a la incertidumbre inicial y esto genera un nuevo espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P}_B)$ donde la incertidumbre a cambiado por la ocurrencia (o supuesta ocurrencia) del evento $B \in \mathcal{F}$.

$$\mathbb{P} \xrightarrow{B} \mathbb{P}_B \quad (1.17)$$

Una vez se tiene \mathbb{P}_B , se puede ver para que elementos $A \in \mathcal{F}$ se cumple que la ocurrencia (o supuesta ocurrencia) del evento B no afecta de ninguna forma la incertidumbre que se tenía respecto al evento A , en otras palabras, no se obtiene *nueva información* relacionada con la incertidumbre del evento A a causa de la ocurrencia (o supuesta ocurrencia) del evento B ; esto se define formalmente, a continuación, como *independencia de eventos*,

Definición 1.13 (Independencia de eventos). Sean $A, B \in \mathcal{F}$ eventos cualesquiera y $\mathbb{P}(B) > 0$, decimos que A y B son eventos independientes entre sí, si se cumple que

$$\mathbb{P}_B(A) = \mathbb{P}(A) \quad (1.18)$$

esto es equivalente a la condición siguiente

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}(B) \quad (1.19)$$

si no se cumple la condición 1.18, decimos que A y B son dependientes entre sí.

Se puede extender el concepto de independencia de eventos a variables aleatorias, esto se hace del siguiente modo; primeramente se da la siguiente definición,

Definición 1.14. Sean X y Y dos variables aleatorias definidas en un mismo espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$. Se denota la medida inducida por el vector aleatorio (por simpleza también lo escribe como v.a.) (X, Y) como \mathbb{P}_{XY} ; dicha función está definida de la siguiente forma,

$$\mathbb{P}_{XY} : \mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}) \mapsto [0, 1]$$

y su valor para cualquier evento $A, B \in \mathcal{B}(\mathbb{R})$ es el siguiente:

$$\mathbb{P}_{XY}(A, B) := \mathbb{P}\left(\{\omega \in \Omega : X(\omega) \in A\} \cap \{\omega \in \Omega : Y(\omega) \in B\}\right) \quad (1.20)$$

También se define la función de distribución conjunta de probabilidades del vector aleatorio (X, Y) como la función $F_{XY} : \mathbb{R}^2 \mapsto \mathbb{R}$, con la siguiente regla de correspondencia,

$$F_{XY}(x, y) = \mathbb{P}\left(\{\omega \in \Omega : X(\omega) \leq x\} \cap \{\omega \in \Omega : Y(\omega) \leq y\}\right) \quad (1.21)$$

Y de forma equivalente como en el caso univariado (el caso de \mathbb{P}_X), se tiene una notación equivalente para \mathbb{P}_{XY} , la cual es la siguiente

$$\mathbb{P}_{XY}(A, B) \equiv \mathbb{P}(X \in A, Y \in B) \equiv \mathbb{P}\left((X, Y) \in A \times B\right) \quad (1.22)$$

Se puede utilizar 1.6 en conjunto a 1.22 en los casos particulares de que A o B sean conjuntos de la forma $|a, b| \subseteq \mathbb{R}$ (donde $|$ puede denotar abierto o cerrado).

Una vez hecho esto, se toman las definiciones 1.13 y 1.14 y se construye la siguiente definición,

Definición 1.15 (Independencia de variables aleatorias). *Sea (X, Y) un vector aleatorio cualquiera, se dice que X y Y son variables aleatorias independientes (por simpleza se denota como v.a. ind.) si se cumple que*

$$\mathbb{P}_{XY}(A, B) = \mathbb{P}_X(A) \mathbb{P}_Y(B) \quad (1.23)$$

para cualesquiera eventos $A, B \in \mathcal{B}(\mathbb{R})$.

Y esto en conjunto a lo mencionado en 1.5 nos permite escribir el siguiente resultado,

Teorema 1.16. *Sea (X, Y) un vector aleatorio cualquiera. Las variables aleatorias X y Y son independientes si y sólo si,*

$$F_{XY}(x, y) = F_X(x) F_Y(y) \quad (1.24)$$

Demostración. Para el resultado anterior, véase Klenke *op. cit.*

□

Se puede extender el concepto de probabilidad condicional y relacionarlo con el concepto de función distribución conjunta de probabilidades F_{XY} perteneciente a un v.a. (X, Y) cualquiera,

Definición 1.17. *Sea (X, Y) un vector aleatorio cualquiera y sea F_{XY} su respectiva función de distribución. La función de distribución condicional de Y dado el evento $\{X = x\}$ denotada como $F_{Y|X}(\cdot | x)$, se define del siguiente modo*

(i) *Si $\mathbb{P}(X = x) > 0$; para toda $y \in \mathbb{R}$ se define $F_{Y|X}(\cdot | x)$ como sigue:*

$$F_{Y|X}(y | x) := \mathbb{P}(Y \leq y | X = x) \quad (1.25)$$

(ii) *Si $\mathbb{P}(X = x) = 0$ y $x \in \text{Ran } X$; para toda $y \in \mathbb{R}$ se define $F_{Y|X}(\cdot | x)$ como sigue:*

$$F_{Y|X}(y | x) := \lim_{\delta \rightarrow 0^+} \mathbb{P}(Y \leq y | x \leq X \leq x + \delta) \quad (1.26)$$

para el caso 1.25 se tiene que

$$F_{Y|X}(y | x) = \frac{\mathbb{P}(Y \leq y, X = x)}{\mathbb{P}(X = x)} = \frac{F_{XY}(x, y) - F_{XY}(x-, y)}{F_{XY}(x, +\infty) - F_{XY}(x-, +\infty)}$$

y para el caso 1.26

$$\begin{aligned} F_{Y|X}(y|x) &= \lim_{\delta \rightarrow 0^+} \frac{\mathbb{P}(x \leq X \leq x + \delta, Y \leq y)}{\mathbb{P}(x \leq X \leq x + \delta)} = \lim_{\delta \rightarrow 0^+} \frac{F_{XY}(x + \delta, y) - F_{XY}(x-, y)}{F_X(x + \delta) - F_X(x-)} \\ &= \lim_{\delta \rightarrow 0^+} \frac{F_{XY}(x + \delta, y) - F_{XY}(x-, y)}{F_{XY}(x + \delta, +\infty) - F_{XY}(x-, +\infty)} \end{aligned} \quad (1.27)$$

es decir, teniendo la función de distribución conjunta F_{XY} para un v.a. (X, Y) , siempre se puede obtener la distribución condicional $F_{Y|X}$ para cualquier punto $x \in \text{Ran } X$ y $y \in \mathbb{R}$, y de forma equivalente para $F_{X|Y}$.

En el caso particular de que la función F_{XY} y la función F_X tengan derivada parcial en su primera entrada y derivada en un punto en común $x \in \mathbb{R}$, respectivamente; la ecuación 1.27 es equivalente a lo siguiente,

$$F_{Y|X}(y|x) = \lim_{\delta \rightarrow 0^+} \frac{F_{XY}(x + \delta, y) - F_{XY}(x, y)}{F_X(x + \delta) - F_X(x)} = \frac{\frac{\partial}{\partial x} F_{XY}(x, y)}{\frac{d}{dx} F_X(x)} \quad (1.28)$$

y esto se hace multiplicando en la ecuación 1.27 por un $\delta > 0$ disfrazado de 1, y luego tomando el límite cuando $\delta \rightarrow 0^+$.

Una vez se tiene la distribución condicional, se puede realizar estimación puntual de la v.a. Y dado la ocurrencia (o supuesta ocurrencia) del evento $\{X = x\}$; esto se hace con el objetivo de ver como la dependencia entre las variables aleatorias afecta dicha estimación puntual, en comparación a la estimación puntual univariada de Y . Esto se define mediante el siguiente concepto,

Definición 1.18 (Curva de regresión). *Sea (X, Y) un vector aleatorio cualquiera, y sea F_{XY} la función de distribución conjunta de probabilidades del vector anterior, se define las siguientes dos curvas de regresión de Y dado el evento $\{X = x\}$ para cualquier $x \in \mathbb{R}$ (siempre y cuando las funciones estén bien definidas en dicho punto).¹¹*

(i) *Curva de regresión con la media, denotada como $\mathcal{RC}_{\mathbb{E}}$ (donde exista la esperanza siguiente):*

$$\begin{aligned} \mathcal{RC}_{\mathbb{E}}(Y | X = x) &:= \mathbb{E}(Y | X = x) \\ &:= \int_0^{\infty} (1 - F_{Y|X}(y|x)) dx - \int_{-\infty}^0 F_{Y|X}(y|x) dx \end{aligned} \quad (1.29)$$

¹¹El nombre \mathcal{RC} proviene de las siglas en inglés, *regression curve*.

(ii) *Curva de regresión con la mediana, denotada como $\mathcal{RC}_{\mathbb{M}}$ (donde exista la mediana siguiente):*

$$\mathcal{RC}_{\mathbb{M}}(Y | X = x) := \mathbb{M}(Y | X = x) \quad (1.30)$$

donde la mediana $\mathbb{M}(Y | X = x)$ queda definida como el punto que cumple la siguiente condición,

$$\mathbb{P}(Y \leq \mathbb{M}(Y | X = x) | X = x) = \frac{1}{2} = \mathbb{P}(Y > \mathbb{M}(Y | X = x) | X = x) \quad (1.31)$$

en el caso particular de que el v.a. (X, Y) sea independiente, se tiene por el teorema 1.16 que para cualquier $x \in \mathbb{R}$,

$$F_{Y|X}(y | x) = F_Y(y), \quad y \in \mathbb{R} \quad (1.32)$$

y como consecuencia de la ecuación 1.32, se cumple que (donde estén bien definidas),

$$\begin{aligned} \mathcal{RC}_{\mathbb{E}}(Y | X = x) &= \mathbb{E}(Y), \quad \forall x \in \mathbb{R} \\ \mathcal{RC}_{\mathbb{M}}(Y | X = x) &= \mathbb{M}(Y), \quad \forall x \in \mathbb{R} \end{aligned}$$

es decir, ambas curvas de regresión son constantes y la ocurrencia (o supuesta ocurrencia) de cualquier evento $\{X = x\}$ no afecta de ninguna forma la incertidumbre que se tiene con respecto a la v.a. Y .

Se puede realizar una extensión de los conceptos de función de masa y de densidad de probabilidades al caso bivariado; esto se hace con las siguientes dos definiciones,

Definición 1.19. *Sea (X, Y) un vector aleatorio compuesto por dos variables aleatorias discretas. Se denota la función de masa de probabilidades del vector anterior como p_{XY} y la definimos como sigue,*

$$p_{XY}(x, y) := \mathbb{P}(\{X = x\} \cap \{Y = y\}) \equiv \mathbb{P}(X = x, Y = y) \quad (1.33)$$

esto se hace para todo $x \in \text{Ran } X$, $y \in \text{Ran } Y$. Las funciones p_1 y p_2 definidas como sigue

$$p_1(x) = \sum_{y \in \text{Ran } Y} p_{XY}(x, y); \quad p_2(y) = \sum_{x \in \text{Ran } X} p_{XY}(x, y);$$

se denominan las marginales de p_{XY} .

Definición 1.20. Sea (X, Y) un vector aleatorio con función de distribución conjunta F_{XY} ; dicho vector se denomina absolutamente continuo (v.a. abs. continuo) si existe una función $f_{XY} : \mathbb{R}^2 \mapsto \mathbb{R}$ que cumple las siguientes propiedades,

(i) $f_{XY}(x, y) \geq 0$.

(ii) f_{XY} es integrable en el sentido de Riemann.

(iii) f_{XY} cumple con la siguiente condición,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y) \, dxy = 1$$

(iv) Para todo $x, y \in \mathbb{R}$ se cumple que,

$$F_{XY}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{XY}(t, s) \, dt ds \quad (1.34)$$

en este caso, se dice que f_{XY} es la función de densidad conjunta de probabilidades para el vector (X, Y) . Si definimos las funciones f_1 y f_2 como,

$$f_1(x) := \int_{-\infty}^{\infty} f_{XY}(x, y) \, dy; \quad f_2(y) = \int_{-\infty}^{\infty} f_{XY}(x, y) \, dx \quad (1.35)$$

decimos que f_1 y f_2 son las marginales de f_{XY} .

Gracias al Teorema del Cálculo Diferencial e Integral de varias variables¹² se cumple la siguiente relación, para el caso de un v.a. abs. continuo (X, Y) , entre F_{XY} y f_{XY} :

$$\frac{\partial^2}{\partial x \partial y} F_{XY}(x, y) = f_{XY}(x, y), \quad x, y \in \mathbb{R} \quad (1.36)$$

Además, gracias a las ecuaciones en 1.35, si se escoge un v.a. absolutamente continuo con función de densidad de probabilidades f_{XY} , y marginales f_1, f_2 ; entonces se cumple que

$$f_1(x) = f_X(x); \quad f_2(y) = f_Y(y); \quad (1.37)$$

es decir, las variables aleatorias X y Y , de forma marginal, son absolutamente continuas siempre que el vector (X, Y) sea absolutamente continuo.

¹²Véase, por ejemplo, Spivak (1988).

La ecuación 1.37 se puede también usar en el caso de que (X, Y) sea un vector discreto (es decir, X y Y son v.a. discretas), nada más utilizando el concepto de función de masa de probabilidades en vez del de función de densidad conjunta; es decir, si (X, Y) es un vector discreto con función de masa de probabilidades conjunta p_{XY} y marginales p_1 y p_2 entonces también se cumple que,

$$p_1(x) = p_X(x); \quad p_1(y) = p_Y(y); \quad (1.38)$$

por lo cual también se puede obtener las funciones de masa marginales utilizando la función conjunta.

Uno pensaría que si X y Y son v.a. abs. continuas, entonces el vector (X, Y) es también absolutamente continuo. Esto en general no es cierto y el siguiente ejemplo lo demuestra,

Ejemplo 1.21. Sea (U, V) un vector aleatorio y sean U y V v.a. igualmente distribuidas (es decir, tienen la misma función de distribución), con la siguiente función de densidad de probabilidades,

$$f_U(z) = \mathbf{1}_{\{0 < z < 1\}} = f_V(z)$$

Y sea $H : \mathbb{R}^2 \mapsto \mathbb{R}$; la siguiente función,

$$H(u, v) = \begin{cases} 0 & \text{Si } u, v \leq 0 \\ \min\{u, v\} & \text{Si } u, v \in]0, 1] \\ u & \text{Si } v > 1 \text{ y } u \in]0, 1] \\ v & \text{Si } u > 1 \text{ y } v \in]0, 1] \\ 1 & \text{Si } u, v \geq 1 \end{cases}$$

Nótese como H es una función admisible como distribución conjunta de probabilidades y además notemos que,

$$H(u, +\infty) = F_U(u); \quad H(+\infty, v) = F_V(v)$$

es decir, H es una función admisible como distribución conjunta para el v.a. (X, Y) . La derivada parcial mixta para la función H queda como sigue,

$$\frac{\partial^2}{\partial u \partial v} H(u, v) = 0, \quad u \neq v$$

y en el caso que $u = v$, la anterior derivada parcial de segundo orden no existe.

El caso anterior es un ejemplo claro de que no toda distribución conjunta de probabilidades cuenta con una función de densidad conjunta de probabilidades asociada, y que además esta derivada parcial mixta se puede anular (o no existir) para ciertos valores, esto sucede por el tipo de función de distribución conjunta que fue escogida, a este tipo de distribuciones se les denomina *singulares*.

El concepto de probabilidad condicional se puede también extender utilizando a las funciones de masa conjunta y las funciones de densidad conjunta, esto se hace mediante las siguientes dos definiciones,

Definición 1.22. Sea (X, Y) un vector aleatorio discreto. Para cualquier $x \in \mathcal{D}_X$, se define y denota la función de masa de probabilidades de la variable aleatoria Y condicional al evento $\{X = x\}$ como sigue,

$$p_{Y|X}(y|x) := \mathbb{P}(Y = y | X = x), \quad y \in \text{Ran } Y \quad (1.39)$$

Definición 1.23. Sea (X, Y) un vector aleatorio absolutamente continuo. Para cualquier $x \in \text{Ran } X$, se define y denota la función de densidad de probabilidades de la variable aleatoria Y condicional a X como sigue,

$$f_{Y|X}(y|x) := \frac{f_{XY}(x, y)}{f_X(x)}, \quad y \in \text{Ran } Y \quad (1.40)$$

Recordando la ecuación 1.28; en el caso de que se cuente con un v.a. abs. continuo (X, Y) , se cumple que F_{XY} tiene derivadas parciales y F_X es también derivable, por lo que utilizando el resultado en la ecuación 1.36 e integrando la ecuación 1.40 se tiene que,

$$F_{Y|X}(y|x) = \int_{-\infty}^y f_{Y|X}(s|x) ds = \int_{-\infty}^y \frac{f_{XY}(x, s)}{f_X(x)} ds = \frac{\frac{\partial}{\partial x} F_{XY}(x, y)}{\frac{d}{dx} F_X(x)} \quad (1.41)$$

lo cual relaciona el concepto de función de densidad condicional y función de distribución condicional en el caso particular de que (X, Y) sea un v.a. abs. continuo.

Por último, si se tiene una función $g : A \times B \mapsto \mathbb{R}$ cualquiera, también resulta de interés analizar la variable aleatoria $g(X, Y)$ donde el vector aleatorio (X, Y) tiene como función de distribución F_{XY} ; en este caso la dependencia entre las variables aleatorias X y Y y las distribuciones univariadas F_X y F_Y juegan un papel fundamental en la transformación del vector.

Ejemplo 1.24. Sean (X_1, Y_1) y (X_2, Y_2) dos v.a. abs. continuos con funciones de densidad conjunta de probabilidades f_1 y f_2 , respectivamente. Se define f_1 y f_2 como sigue,

$$\begin{aligned} f_1(x, y) &:= \mathbf{1}_{\{0 < 2y < x < 2\}} \\ f_2(x, y) &:= x(1-y)\mathbf{1}_{\{0 < x < 2\}}\mathbf{1}_{\{0 < y < 1\}} \end{aligned}$$

Ahora bien, utilizando la función $g(x, y) := \frac{y}{x}$; se desea obtener la distribución de las variables aleatorias,

$$Z_1 := g(X_1, Y_1) = \frac{Y_1}{X_1}; \quad Z_2 := g(X_2, Y_2) = \frac{Y_2}{X_2}$$

Para el caso de Z_1 , se tiene que $\text{Ran } Z_1 =]0, \frac{1}{2}[$ y F_{Z_1} queda como sigue

$$\begin{aligned} F_{Z_1}(z) &= \mathbb{P}(Z_1 \leq z) = \mathbb{P}\left(\frac{Y_1}{X_1} \leq z\right) = \mathbb{P}(Y_1 \leq zX_1) \\ &= \int_0^2 \int_0^{zx} f_1(x, y) dy dx = \dots = 2z\mathbf{1}_{\{0 < z < \frac{1}{2}\}} + \mathbf{1}_{\{z \geq \frac{1}{2}\}} \end{aligned}$$

Mientras que para el caso de Z_2 , se tiene que $\text{Ran } Z_2 =]0, +\infty[$ y F_{Z_2} queda como sigue

1) Si $z \in]0, \frac{1}{2}[$

$$\begin{aligned} F_{Z_2}(z) &= \mathbb{P}(Z_2 \leq z) = \mathbb{P}\left(\frac{Y_2}{X_2} \leq z\right) = \mathbb{P}(Y_2 \leq zX_2) \\ &= \int_0^2 \int_0^{zx} f_2(x, y) dy dx = \dots = z\left(\frac{8}{3} - 2z\right) \end{aligned}$$

2) Si $z \in [\frac{1}{2}, +\infty[$

$$\begin{aligned} F_{Z_2}(z) &= \mathbb{P}(Z_2 \leq z) = \mathbb{P}\left(\frac{Y_2}{X_2} \leq z\right) = \mathbb{P}(Y_2 \leq zX_2) \\ &= \int_0^1 \int_{y/z}^2 f_2(x, y) dx dy = \dots = \left(1 - \frac{1}{24z^2}\right) \end{aligned}$$

Quedando F_{Z_1} y F_{Z_2} como siguen,

$$F_{Z_1}(z) = 2z\mathbf{1}_{\{0 < z < \frac{1}{2}\}} + \mathbf{1}_{\{z \geq \frac{1}{2}\}}$$

$$F_{Z_2}(z) = z\left(\frac{8}{3} - 2z\right)\mathbf{1}_{\{0 < z < \frac{1}{2}\}} + \left(1 - \frac{1}{24z^2}\right)\mathbf{1}_{\{z \geq \frac{1}{2}\}}$$

Notemos además una característica especial de f_1 y f_2 en las siguientes ecuaciones,

$$\int_{-\infty}^{+\infty} f_1(x, y) dy = \int_{-\infty}^{+\infty} f_2(x, y) dy$$

$$\int_{-\infty}^{+\infty} f_1(x, y) dx = \int_{-\infty}^{+\infty} f_2(x, y) dx$$

Es decir, se tienen dos v.a. abs. continuos (X_1, Y_1) y (X_2, Y_2) ; tales que $X_1 = X_2$ y $Y_1 = Y_2$ pero $f_{X_1Y_1} \neq f_{X_2Y_2}$ (marginamente se distribuyen igual pero conjuntamente no), y esto es suficiente para que $Z_1 \neq Z_2$. Por lo cual a la hora de realizar una transformación para cualquier v.a. (X, Y) , esto es solamente posible hacerlo contando con la función de distribución conjunta F_{XY} .

Con esto se ve como la información univariada (distribución de probabilidad univariada de X y Y) **no es información suficiente** para conocer la distribución conjunta del vector aleatorio (X, Y) , entonces se requiere crear una estrategia para medir, cuantificar o modelar la relación existente entre X y Y .

La estrategia que se va a seguir en este escrito para medir la dependencia de un v.a. (X, Y) es mediante el concepto de cópulas, y como estas pueden ser utilizadas para crear modelos de distribución multivariados variados. Se denota como $\overline{\mathbb{R}}$ a la recta real extendida, es decir, $\overline{\mathbb{R}} := [-\infty, \infty]$.

Definición 1.25. Sean $A, B \subseteq \overline{\mathbb{R}}$ dos subconjuntos no vacíos en los reales cualesquiera y sea $H : A \times B \mapsto \mathbb{R}$ una función cualquiera. Para cualquier rectángulo $B = [x_1, x_2] \times [y_1, y_2]$ tal que los vértices del rectángulo estén contenidos en el Dominio de H (denotado como $\text{Dom } H$). Se denota el H -volumen del rectángulo B como $V_H(B)$ y se define como sigue,

$$V_H(B) := H(x_2, y_2) - H(x_2, y_1) - H(x_1, y_2) + H(x_1, y_1) \quad (1.42)$$

Nótese que si la función H descrita anteriormente es una función de distribución conjunta de probabilidades asociada a un v.a. (X, Y) ; entonces para cualquier $B = [x_1, x_2] \times [y_1, y_2] \subseteq \mathbb{R}$ se cumple que,

$$\begin{aligned}
V_H(B) &= \mathbb{P}(X \leq x_2, Y \leq y_2) - \mathbb{P}(X \leq x_2, Y \leq y_1) - \mathbb{P}(X \leq x_1, Y \leq y_2) \\
&\quad + \mathbb{P}(X \leq x_1, Y \leq y_1) \\
&= \mathbb{P}(X \leq x_2, y_1 < Y \leq y_2) - \mathbb{P}(X \leq x_1, y_1 < Y \leq y_2) \\
&= \mathbb{P}(x_1 < X \leq x_2, y_1 < Y \leq y_2) \\
&= \mathbb{P}\left((X, Y) \in B\right) - \mathbb{P}(X = x_1, y_1 < Y \leq y_2) - \mathbb{P}(x_1 < X \leq x_2, Y = y_1) \\
&\quad - \mathbb{P}(X = x_1, Y = y_1)
\end{aligned}$$

Por lo mencionado anteriormente, se obtiene el siguiente resultado.

Corolario 1.25.1. *Sea (X, Y) un vector aleatorio cualquiera y F_{XY} su respectiva función de distribución conjunta de probabilidades. Entonces para cualquier rectángulo $B = [x_1, x_2] \times [y_1, y_2]$ se cumple que,*

$$V_{F_{XY}}(B) = \mathbb{P}(x_1 < X \leq x_2, y_1 < Y \leq y_2) \geq 0 \quad (1.43)$$

Si además F_{XY} es una función continua, se cumple que,

$$V_{F_{XY}}(B) = \mathbb{P}\left((X, Y) \in B\right) \quad (1.44)$$

El resultado anterior inspira la siguiente definición,

Definición 1.26. *Una función $H : A \times B \subseteq \mathbb{R}^2 \mapsto \mathbb{R}$ es 2-creciente si $V_H(B) \geq 0$ para cualquier rectángulo B cuyos vértices se encuentren en $\text{Dom } H$.*

Por la definición 1.26 y el corolario 1.25.1 se tiene que toda función de distribución conjunta de probabilidades es 2-creciente por construcción.

Uno pensaría que al decir que una función H es 2-creciente, esto es consecuencia de que H es una función no decreciente en cada argumento de H . Esto en general no es cierto, y el siguiente ejemplo es prueba de eso

Ejemplo 1.27. Sea H una función definida en $\mathbf{I}^2 := [0, 1]^2$ como

$$H(x, y) := (2x - 1)(2y - 1)$$

Entonces se ve que H es 2-creciente, sea $B = [x_1, x_2] \times [y_1, y_2] \subseteq \mathbf{I}^2$, un rectángulo cualquiera, entonces veamos a que equivale $V_H(B)$,

$$\begin{aligned} V_H(B) &= (2x_2 - 1)(2y_2 - 1) - (2x_2 - 1)(2y_1 - 1) - (2x_1 - 1)(2y_2 - 1) \\ &\quad + (2x_1 - 1)(2y_1 - 1) \\ &= 4(x_2y_2 - x_1y_2 - x_2y_1 + x_1y_1) = 4V_F(B) \end{aligned}$$

donde $V_F(B)$ representa la probabilidad de que un v.a. (X, Y) independiente, se encuentre en B , donde $X, Y \sim U(0, 1)$; por lo cual la función H es 2-creciente. Pero si se fija un $y \in (0, \frac{1}{2})$ cualquiera, la función H es decreciente en x y también es una función decreciente en y si fijamos un $x \in (0, \frac{1}{2})$ cualquiera.

Los siguientes lemas serán muy útiles para establecer posteriormente la continuidad de subcópulas y cópulas.

Lema 1.28. Sean A y B dos subconjuntos no vacíos cualesquiera de $\overline{\mathbb{R}}$, y sea H una función 2-creciente con $\text{Dom } H = A \times B$. Sean $x_1, x_2 \in A$ con $x_1 \leq x_2$ y sean $y_1, y_2 \in B$ con $y_1 \leq y_2$. Si definimos las funciones siguientes:

$$\begin{aligned} H_1(t) &:= H(t, y_2) - H(t, y_1) \\ H_2(t) &:= H(x_2, t) - H(x_1, t) \end{aligned}$$

entonces la función H_1 es no-decreciente en A , mientras que H_2 es no-decreciente en B .

Demostración. Esto es consecuencia inmediata de que H es 2-creciente y de la ecuación 1.42. \square

Para el siguiente lema primero se establece la siguiente definición,

Definición 1.29. Sean $A, B \subseteq \overline{\mathbb{R}}$ dos subconjuntos no vacíos cada uno con elementos mínimos denotados, respectivamente, como

$$a := \text{mín } A; \quad b := \text{mín } B$$

y sea $H : A \times B \mapsto \mathbb{R}$ una función. Decimos que H es anulable si se cumple la siguiente condición

$$H(x, b) = 0 = H(a, y) \quad (1.45)$$

Con la definición anterior se establece el siguiente lema,

Lema 1.30. Sean $A, B \subseteq \overline{\mathbb{R}}$ dos subconjuntos no vacíos cualesquiera y sea H una función anulable 2-creciente con $\text{Dom } H = A \times B$. Entonces H es no-decreciente para cada argumento.

Demostración. Sean a y b los elementos mínimos de A y B , respectivamente, entonces tomamos $x_1 = a$ y $y_1 = b$. Utilizando el lema 1.28 se llega al resultado. \square

Volviendo al ejemplo 1.27, nótese que H no es una función anulable, ya que si lo fuera sería también no decreciente en cada entrada, lo cual se sabe no es cierto.

Si tomamos cualquier v.a. (X, Y) y su respectiva función de distribución conjunta de probabilidades F_{XY} , y se cumple que para $\text{Ran } (X, Y) \subseteq \text{Ran } X \times \text{Ran } Y$, tales que $\text{Ran } X, \text{Ran } Y \subset \mathbb{R}$ son subconjuntos propios de los reales; entonces se cumple que F_{XY} es anulable (esto se consigue con los puntos $a := \inf \text{Ran } X$ y $b := \inf \text{Ran } Y$). En el caso de que $\text{Ran } (X, Y) = \mathbb{R}^2$, la igualdad en 1.45 se consigue como sigue para cualquier $(x, y) \in \text{Ran } (X, Y)$,

$$\lim_{b \rightarrow -\infty} F_{XY}(x, b) = 0 = \lim_{a \rightarrow -\infty} F_{XY}(a, y) \quad (1.46)$$

A continuación, se da la siguiente definición.

Definición 1.31. Sean $A, B \subseteq \overline{\mathbb{R}}$ subconjuntos no vacíos tales que cada uno tenga elementos máximos a y b , respectivamente. Entonces se dice que una función $H : A \times B \mapsto \mathbb{R}$ tiene marginales, y las marginales de H son las funciones F y G definidas como sigue:

$$\text{Dom } F = A; \quad F(x) := H(x, b) \quad (1.47)$$

$$\text{Dom } G = B; \quad G(y) := H(a, y) \quad (1.48)$$

Si F_{XY} es una función de distribución conjunta de probabilidades, y se cumple que para $\text{Ran}(X, Y) \subseteq \text{Ran} X \times \text{Ran} Y$, tales que $\text{Ran} X, \text{Ran} Y \subset \mathbb{R}$ son subconjuntos propios de los reales; entonces se pueden obtener las marginales de F_{XY} del siguiente modo,

$$\begin{aligned} F_X(x) &= F_{XY}(x, b); & b &:= \sup \text{Ran} Y \\ F_Y(y) &= F_{XY}(a, y); & a &:= \sup \text{Ran} X \end{aligned}$$

en el caso de que $\text{Ran}(X, Y) = \mathbb{R}^2$, se tiene que las marginales F_X, F_Y se pueden conseguir con la siguiente fórmula,

$$\begin{aligned} F_X(x) &= \lim_{y \rightarrow \infty} F_{XY}(x, y) \\ F_Y(y) &= \lim_{x \rightarrow \infty} F_{XY}(x, y) \end{aligned}$$

Se completan los anteriores dos lemas con el siguiente resultado que involucra funciones anulables 2-crecientes.

Lema 1.32. *Sea $A, B \subseteq \overline{\mathbb{R}}$ dos subconjuntos vacíos cualesquiera y sea H una función anulable 2-creciente, con marginales F y G , respectivamente, y cuyo dominio es $A \times B$. Sean $(x_1, y_1), (x_2, y_2) \in A \times B$. Entonces,*

$$|H(x_2, y_2) - H(x_1, y_1)| \leq |F(x_2) - F(x_1)| + |G(y_2) - G(y_1)| \quad (1.49)$$

Demostración. Por la desigualdad del triángulo para números reales,

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

Supóngase que $x_1 \leq x_2$. El lema 1.28 implica que

$$0 \leq H(x_2, y_2) - H(x_1, y_2) \leq F(x_2) - F(x_1)$$

Una igualdad equivalente se cumple cuando $x_2 \leq x_1$, de esto se sigue los siguientes dos puntos:

(1) Para cualesquiera $x_1, x_2 \in A$

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

(2) Para cualesquiera $y_1, y_2 \in B$

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

completando la prueba. □

Con la información anterior, se puede establecer los conceptos de subcópula y cópula

Definición 1.33. *Una subcópula bivariada (abreviado como subcópula) es una función C' con las siguientes propiedades:*

- (i) $Dom C' = A \times B$, donde $A, B \subseteq \mathbf{I}$ tales que $\{0, 1\} \subseteq A, B$.
- (ii) C' es una función anulable y 2-creciente (entonces es no-decreciente en cada entrada gracias al lema 1.30).
- (iii) Para cualquier $u \in A$ y $v \in B$,

$$C'(u, 1) = u; \quad C'(1, v) = v \quad (1.50)$$

Notemos que para cualquier $(u, v) \in Dom C'$, se cumple que

$$0 = C'(u, 0) \leq C'(u, v) \leq C'(u, 1) = u \leq 1$$

es decir, $Ran C' \subseteq \mathbf{I}$.

Definición 1.34. *Una cópula bivariada (abreviado como cópula) es una subcópula bivariada C cuyo dominio es \mathbf{I}^2 .*

Dadas las definiciones 1.33 y 1.34, podemos establecer las propiedades que debe cumplir una función cualquiera $C : \mathbf{I}^2 \mapsto \mathbf{I}$ para ser una cópula.

Proposición 1.35. *Sea $C : \mathbf{I}^2 \mapsto \mathbf{I}$ una cópula, entonces C cumple las siguientes propiedades:*

$$(i) \text{ Para cualquier } u, v \in \mathbf{I}, \quad C(u, 0) = 0 = C(0, v) \quad (1.51)$$

y

$$C(u, 1) = u; \quad C(1, v) = v \quad (1.52)$$

- (ii) Para cualquier $u_1, u_2, v_1, v_2 \in \mathbf{I}$ tales que $u_1 \leq u_2$ y $v_1 \leq v_2$,

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0 \quad (1.53)$$

La ecuación 1.53 es la propiedad 2-creciente que cumple cualquier cópula C ; y que si se recuerda, asocia a cualquier rectángulo $B \in \mathbf{I}^2$ un número no negativo que representa el C -volumen del rectángulo B y lo denotamos como $V_C(B)$.

Antes de dar una extensión de interés al concepto de cópula, es de especial importancia notar las propiedades que cumple una función de distribución de probabilidades F_X asociada a una variable aleatoria X cualquiera.

Proposición 1.36. *Sea X una variable aleatoria cualquiera y sea F_X su respectiva función de distribución de probabilidades. Entonces F_X cumple las siguientes propiedades:*

- (i) Si $x_1 < x_2$ entonces $F_X(x_1) \leq F_X(x_2)$ (monótona creciente).
- (ii) $\lim_{h \rightarrow 0^+} F_X(x+h) = F_X(x)$ (continuidad por la derecha).
- (iii) $\lim_{x \rightarrow -\infty} F_X(x) = 0$; $\lim_{x \rightarrow +\infty} F_X(x) = 1$.

y además, se puede definir para cualquier función F que cumpla las propiedades (i) hasta (iii) una variable aleatoria X tal que $F_X = F$.

Este resultado se puede extender en relación a una función de distribución conjunta de probabilidades F_{XY} asociada a un v.a. (X, Y) cualquiera, solamente cambiando las propiedades que debe cumplir la función H para ser igual a F_{XY} (es decir, $H = F_{XY}$), las propiedades se dan en la siguiente proposición.

Proposición 1.37. *Sea (X, Y) un vector aleatorio cualquiera y sea F_{XY} su respectiva función de distribución conjunta de probabilidades. Entonces F_{XY} cumple las siguientes propiedades:*

- (i) $F_{XY}(x, -\infty) = 0 = F_{XY}(-\infty, y)$.
- (ii) $F_{XY}(x, +\infty) = F_X(x)$; $F_{XY}(+\infty, y) = F_Y(y)$.
- (iii) $F_{XY}(+\infty, +\infty) = 1$.
- (iv) F_{XY} es 2-creciente.
- (v) F_{XY} es monótona creciente en cada variable.
- (vi) F_{XY} es continua por la derecha en cada variable.

Con la proposición 1.37 se puede considerar una extensión de dominio para cualquier cópula C . Esto se hace con el objetivo de relacionar cualquier cópula C con una función de distribución conjunta de probabilidades H en específico,

Definición 1.38 (Extensión de cópulas a \mathbb{R}^2). *Sea C una cópula cualquiera; se extiende el dominio de C a \mathbb{R}^2 de la siguiente forma,*

$$(u, v) \mapsto \begin{cases} 0; & \text{Si } \min\{u, v\} < 0 \\ C(u, v); & \text{Si } (u, v) \in \mathbf{I}^2 \\ u; & \text{Si } v > 1, u \in \mathbf{I} \\ v; & \text{Si } u > 1, v \in \mathbf{I} \\ 1; & \text{Si } \max\{u, v\} > 1 \end{cases} \quad (1.54)$$

dicha función se denota como $C_{\mathbb{R}^2}$.

La anterior definición permite la creación del siguiente resultado,

Teorema 1.39. *Sea C una cópula cualquiera y sea $C_{\mathbb{R}^2}$ su extensión para \mathbb{R}^2 . Entonces $C_{\mathbb{R}^2}$ es válida como una función de distribución conjunta de probabilidades asociadas a un par de variables aleatorias $U(0,1)$.*

Demostración. Primero se demostrará las propiedades establecidas en la proposición 1.37, y para esto se toma $H = C_{\mathbb{R}^2}$, donde C es la cópula de interés:

- a) Para la propiedad (i), como $C(u, v) = 0$ siempre que $u, v \leq 0$, entonces los límites en (i) quedan como sigue,

$$\lim_{v \rightarrow -\infty} H(u, v) = 0 = \lim_{u \rightarrow -\infty} H(u, v)$$

- b) Para la propiedad (ii), notemos lo siguiente,

$$\begin{aligned} \lim_{v \rightarrow +\infty} H(u, v) &= \mathbf{1}_{\{0 < u < 1\}} + \mathbf{1}_{\{u \geq 1\}} \\ \lim_{u \rightarrow +\infty} H(u, v) &= \mathbf{1}_{\{0 < v < 1\}} + \mathbf{1}_{\{v \geq 1\}} \end{aligned}$$

y en ambas casos, estas funciones son admisibles como función de distribución de probabilidades asociadas a variables aleatorias $U(0,1)$.

- c) Para el caso (iii), por definición de $C_{\mathbb{R}}$,

$$\lim_{(u,v) \rightarrow (+\infty, +\infty)} H(u, v) = 1$$

- d) Para el caso (iv), es claro que H es 2-creciente en \mathbf{I}^2 , y para el caso de $B = [x_1, x_2] \times [y_1, y_2] \subseteq \mathbb{R}^2$ cualquiera se tiene que la extensión establecida en la definición 1.38 cumple con la condición siguiente,

$$H(u, v) = \begin{cases} H(0, v); & \text{Si } u < 0 \\ H(u, 0); & \text{Si } v < 0 \\ H(u, 1); & \text{Si } v > 1 \\ H(1, v); & \text{Si } u > 1 \end{cases} \quad (1.55)$$

por lo que,

$$V_H(B) = V_H(B^*) \geq 0$$

donde $B^* = [\text{máx}\{x_1, 0\}, \text{mín}\{x_2, 1\}] \times [\text{máx}\{y_1, 0\}, \text{mín}\{y_2, 1\}]$ (esto significa que, por entrada, no se acumula volumen por debajo de 0 y todo el volumen ya fue acumulado, por entrada, por encima de 1).

- e) Para el caso (v), esto es consecuencia inmediata del lema 1.30 con ciertos ajustes utilizados en el inciso anterior.
- f) Para el caso (vi), se sabe por el teorema 1.43 que cualquier cópula es uniformemente continua, y eso incluye la propiedad (vi) como caso especial, del mismo modo que los incisos d) y e), esta extensión se puede hacer fácilmente para todo $x, y \in \mathbb{R}$.

Ya encontramos las distribuciones univariadas que corresponden a la función $C_{\mathbb{R}}$, por lo que el teorema 1.39 queda demostrado. □

Las propiedades que la función $C_{\mathbb{R}^2}$ tiene como función de distribución conjunta de probabilidades están dadas en la siguiente proposición,

Proposición 1.40. *Sea (U, V) un vector aleatorio con función de distribución conjunta de probabilidades $C_{\mathbb{R}^2}$ para una cópula cualquiera C . Entonces se cumplen las siguientes propiedades.*

- (i) $\text{Ran}(U, V) \subseteq \mathbf{I}^2$.
- (ii) $\text{Ran } U = \text{Ran } V =]0, 1[^2$ con funciones de distribución de probabilidades marginales iguales y con la forma siguiente,

$$F_U(z) = \mathbf{1}_{\{0 < z < 1\}} + \mathbf{1}_{\{z \geq 1\}} = F_V(z) \quad (1.56)$$

es decir, $U, V \sim U(0, 1)$.

Demostración. La propiedad (i) es consecuencia de como esta definida cualquier cópula, mientras que la propiedad (ii) ya quedo probada en la demostración del teorema 1.39 inciso b).

□

Por lo cual, cualquier cópula C puede ser interpretada como una función de distribución conjunta de probabilidades para un v.a. (U, V) de uniformes cero-uno. Después de introducir el *Teorema de Sklar (1959)*, se notará que las cópulas permiten medir la dependencia de cualquier v.a. (X, Y) primeramente transformando su función de distribución conjunta de probabilidades F_{XY} en términos de una cópula.

Los siguientes resultados se establecen en general para cualquier subcópula, y por lo tanto también son válidos para cualquier cópula.

Teorema 1.41. *Sea C' una subcópula. Entonces para todo $(u, v) \in \text{Dom } C'$,*

$$\max\{u + v - 1, 0\} \leq C'(u, v) \leq \min\{u, v\} \quad (1.57)$$

Demostración. Para la parte superior de la desigualdad a demostrar, al ser C' creciente para cualquier entrada

$$C'(u, v) \leq C'(u, 1) = u$$

$$C'(u, v) \leq C'(1, v) = v$$

por lo cual, $C'(u, v) \leq \min\{u, v\}$. Para la desigualdad inferior, se considera el C' -volumen del rectángulo $B = [u, 1] \times [v, 1]$,

$$V_{C'}(B) = C'(1, 1) - C'(1, v) - C'(u, 1) + C'(u, v) \geq 0$$

entonces,

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

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

por lo cual se cumple que $C'(u, v) \geq \max\{u + v - 1, 0\}$.

□

En la desigualdad 1.57 se establecen dos funciones, $M := \min\{u, v\}$ y $W(u, v) := \max\{u + v - 1, 0\}$, que cumplen la siguiente condición para cualquier cópula C ,

$$W(u, v) \leq C(u, v) \leq M(u, v) \quad (1.58)$$

la desigualdad 1.58 es la versión en cópulas de las *cotas de Fréchet-Hoeffding*. A M se le denomina *cota superior de Fréchet-Hoeffding* y a W como la *cota inferior de Fréchet-Hoeffding*. Se puede comprobar fácilmente que M y W cumplen las propiedades de una cópula. Además, otra cópula de especial interés es la *cópula producto* $\Pi(u, v) := uv$. En la figura 1.1 se tienen mapas de calor para las cópulas W, Π y M .

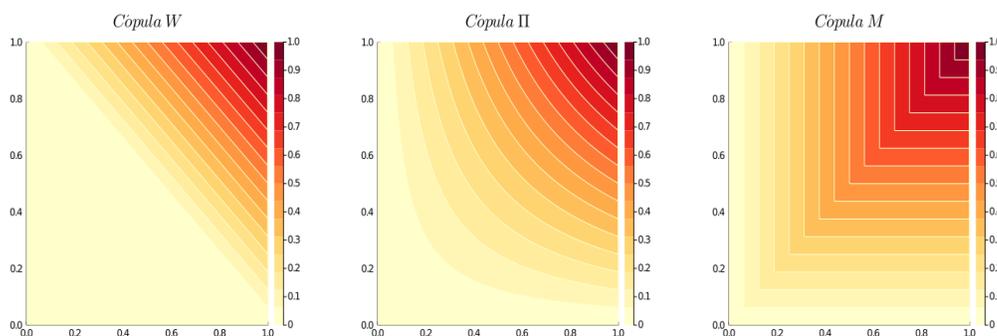


Figura 1.1: Mapas de calor cópulas W, Π y M .

Ejemplo 1.42. Sean $(U_M, V_M), (U_\Pi, V_\Pi), (U_W, V_W)$ vectores aleatorios con función de distribución conjunta de probabilidades $M_{\mathbb{R}^2}, \Pi_{\mathbb{R}^2}, W_{\mathbb{R}^2}$, respectivamente. Se obtienen las siguientes funciones de distribución condicionales para cualquier $(u, v) \in]0, 1[^2$

$$\begin{aligned} F_1(v | u) &:= F_{V_M | U_M}(v | u) \\ F_2(v | u) &:= F_{V_\Pi | U_\Pi}(v | u) \\ F_3(v | u) &:= F_{V_W | U_W}(v | u) \end{aligned}$$

(i) Para el caso de (U_M, V_M) , utilizando la ecuación 1.26,

$$\begin{aligned}
 F_1(v | u) &= \lim_{\delta \rightarrow 0^+} \mathbb{P}(V_M \leq v | u - \delta \leq U_M \leq u) \\
 &= \lim_{\delta \rightarrow 0^+} \frac{M_{\mathbb{R}^2}(u, v) - M_{\mathbb{R}^2}((u - \delta), v)}{\delta} \\
 &= \begin{cases} 0; & \text{Si } v < u \\ 1; & \text{Si } v \geq u \end{cases} = \mathbf{1}_{\{v \geq u\}}
 \end{aligned}$$

por lo que $V_M | \{U_M = u\}$ es una variable aleatoria discreta tal que,

$$\mathbb{P}(V_M = u | U_M = u) = 1$$

(ii) Para el caso de (U_{Π}, V_{Π}) , utilizando la ecuación 1.26 y 1.28,

$$\begin{aligned}
 F_2(v | u) &= \lim_{\delta \rightarrow +0} \mathbb{P}(V_{\Pi} \leq v | u \leq U_{\Pi} \leq u + \delta) = \frac{\frac{\partial}{\partial u} \Pi_{\mathbb{R}^2}(u, v)}{\frac{d}{du} u} \\
 &= v \mathbf{1}_{\{0 < v < 1\}} + \mathbf{1}_{\{v \geq 1\}}
 \end{aligned}$$

por lo que $V_{\Pi} | \{U_{\Pi} = u\} \sim U(0, 1)$.

(iii) Por último, para el caso de que (U_W, V_W) , utilizando la ecuación 1.26,

$$\begin{aligned}
 F_3(v | u) &= \lim_{\delta \rightarrow 0^+} \mathbb{P}(V_{\Pi} \leq v | u \leq U_{\Pi} \leq u + \delta) \\
 &= \lim_{\delta \rightarrow 0^+} \frac{W_{\mathbb{R}^2}(u + \delta, v) - W_{\mathbb{R}^2}(u-, v)}{u + \delta - u} \\
 &= \begin{cases} 0; & \text{Si } v < 1 - u \\ 1; & \text{Si } v \geq 1 - u \end{cases} = \mathbf{1}_{\{v \geq 1 - u\}}
 \end{aligned}$$

por lo que $V_W | \{U_W = u\}$ es una variable aleatoria discreta tal que,

$$\mathbb{P}(V_W = 1 - u | U_W = u) = 1$$

Este ejemplo hace ver como los v.a. $(U_M, V_M), (U_W, V_W)$ cumplen la siguiente condición,

$$\mathbb{P}(U_M = V_M) = 1 = \mathbb{P}(U_W = 1 - V_W)$$

mientras que en el caso de (U_{Π}, V_{Π}) se tiene que U, V son v.a. independientes.

El siguiente teorema es consecuencia inmediata del lema 1.32 y de la proposición 1.35, por lo cual la demostración se omite.

Teorema 1.43. *Sea C' una subcópula. Entonces para todo $(u_1, u_2), (v_1, v_2) \in \text{Dom } C'$,*

$$\left| C'(u_2, v_2) - C'(u_1, v_1) \right| \leq |u_2 - u_1| + |v_2 - v_1| \quad (1.59)$$

Por lo que C' es uniformemente continua en $\text{Dom } C'$.

En la siguiente definición, se establecen ciertas funciones de interés que representan secciones de una cópula C cualquiera.

Definición 1.44. *Sea C una cópula, y a cualquier número en \mathbf{I} . Se definen las siguientes funciones:*

(i) *La sección horizontal de C en el punto a es la función $h_a^C : \mathbf{I} \mapsto \mathbf{I}$ definida como sigue:*

$$h_a^C(t) := C(t, a) \quad (1.60)$$

(ii) *La sección vertical de C en el punto a es la función $v_a^C : \mathbf{I} \mapsto \mathbf{I}$ definida como sigue:*

$$v_a^C(t) := C(a, t) \quad (1.61)$$

(iii) *La sección diagonal de C es la función $\delta_C : \mathbf{I} \mapsto \mathbf{I}$ definida como:*

$$\delta_C(t) := C(t, t) \quad (1.62)$$

Corolario 1.44.1. *Las secciones horizontal, vertical y diagonal de una cópula C son no decrecientes y uniformemente continuas en \mathbf{I} .*

Demostración. Consecuencia inmediata del lema 1.30 y del teorema 1.43. □

Se concluye el capítulo con dos teoremas relacionados con las derivadas parciales de una cópula y un ejemplo. La palabra “casi” es usada en el sentido de la medida de Lebesgue¹³(una condición que se cumple “casi” en cualquier lugar significa que se cumple en un conjunto de medida uno, aunque dicho conjunto no necesariamente sea todo el espacio muestral).

¹³Para mas información al respecto, véase, Klenke *op. cit.*

Teorema 1.45. *Sea C una cópula. Para cualquier $v \in \mathbf{I}$, la derivada parcial $\partial C(u, v)/\partial u$ existe para casi todo $u \in \mathbf{I}$. Y para los u y v que exista se cumple que,*

$$0 \leq \frac{\partial}{\partial u} C(u, v) \leq 1 \quad (1.63)$$

De forma similar, para cualquier $u \in \mathbf{I}$, la derivada parcial $\partial C(u, v)/\partial v$ existe para casi todo $v \in \mathbf{I}$, y para esos u y v se cumple que,

$$0 \leq \frac{\partial}{\partial v} C(u, v) \leq 1 \quad (1.64)$$

Además las funciones,

$$\begin{aligned} u &\mapsto \frac{\partial}{\partial v} C(u, v) \\ v &\mapsto \frac{\partial}{\partial u} C(u, v) \end{aligned}$$

están definidas y son no-decrecientes casi en todo \mathbf{I} .

Demostración. Véase Nelsen *op. cit.* □

Ejemplo 1.46. *Las derivadas parciales de las cópulas M, Π, W están dadas por las siguientes ecuaciones:*

(i) *Para el caso de M , se cumple lo siguiente:*

$$\begin{aligned} \frac{\partial}{\partial v} M(u, v) &= \frac{\partial}{\partial v} \min \{u, v\} = \mathbf{1}_{\{u > v\}}; & u \neq v \\ \frac{\partial}{\partial u} M(u, v) &= \frac{\partial}{\partial u} \min \{u, v\} = \mathbf{1}_{\{v > u\}}; & u \neq v \end{aligned}$$

(ii) *Para el caso de Π , se cumple lo siguiente:*

$$\begin{aligned} \frac{\partial}{\partial v} \Pi(u, v) &= \frac{\partial}{\partial v} uv = u \\ \frac{\partial}{\partial u} \Pi(u, v) &= \frac{\partial}{\partial u} uv = v \end{aligned}$$

(iii) Y para el caso de W , se cumple lo siguiente:

$$\begin{aligned}\frac{\partial}{\partial v}W(u, v) &= \frac{\partial}{\partial v} \max\{u + v - 1, 0\} = \mathbf{1}_{\{u > 1-v\}}; & u \neq 1 - v \\ \frac{\partial}{\partial u}W(u, v) &= \frac{\partial}{\partial u} \max\{u + v - 1, 0\} = \mathbf{1}_{\{v > 1-u\}}; & u \neq 1 - v\end{aligned}$$

por lo cual, en los puntos que las derivadas siguientes estén definidas, se cumple lo siguiente (en relación al ejemplo 1.42),

$$\begin{aligned}F_1(v | u) &= \frac{\partial}{\partial u}M(u, v) \\ F_2(v | u) &= \frac{\partial}{\partial u}\Pi(u, v) \\ F_3(v | u) &= \frac{\partial}{\partial u}W(u, v)\end{aligned}$$

El anterior ejemplo, es un caso particular del siguiente resultado,

Teorema 1.47. Sea (U, V) un vector aleatorio cualquiera y sea $C_{\mathbb{R}^2}$ su respectiva función de distribución conjunta de probabilidades para una cópula C cualquiera. Entonces se cumplen lo siguiente,

$$F_{V|U}(v | u) = \frac{\partial}{\partial u}C_{\mathbb{R}^2}(u, v) \quad (1.65)$$

$$F_{U|V}(u | v) = \frac{\partial}{\partial v}C_{\mathbb{R}^2}(u, v) \quad (1.66)$$

para todo $(u, v) \in \mathbf{I}^2$ donde las derivadas 1.65 y 1.66 estén definidas.

Demostración. Sea (U, V) un v.a. que cumpla las condiciones del teorema 1.47. Entonces por 1.40 se cumple que para todo $z \in (0, 1)$

$$F_U(z) = z = F_V(z)$$

entonces se cumple que,

$$\begin{aligned}F_{V|U}(v | u) &= \lim_{\delta \rightarrow +0} \frac{C(u + \delta, v) - C(u, v)}{F_U(u + \delta) - F_U(u)} = \lim_{\delta \rightarrow +0} \frac{C(u + \delta, v) - C(u, v)}{u + \delta + u} \\ &= \lim_{\delta \rightarrow +0} \frac{C(u + \delta, v) - C(u, v)}{\delta} = \frac{\partial}{\partial u}C(u, v)\end{aligned}$$

y se puede desarrollar un resultado similar para $F_{U|V}(u|v)$.

□

Teorema 1.48. *Sea C una cópula cualquiera. Si $\partial C(u, v)/\partial v$ y $\partial C(u, v)/\partial u$ son continuas en \mathbf{I}^2 y $\partial C(u, v)/\partial u$ existe para todo $u \in]0, 1[$ cuando $v = 0$ entonces $\partial C(u, v)/\partial u$ y $\partial^2 C(u, v)/\partial v \partial u$ existen en $]0, 1[^2$ y*

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

Demostración. La prueba se puede ver en Seeley (1961).

□

1.3. El Teorema de Sklar y sus consecuencias

Se inicia esta sección estableciendo dos definiciones importantes e inmediatamente se establece el teorema de Sklar, el cual permite relacionar las funciones de distribución marginales F_X y F_Y , con su respectiva función de distribución conjunta de probabilidades F_{XY} . Luego, se extienden las implicaciones de dicho resultado, así como un ejemplo que nos permita comprender mas a fondo como las cópulas permiten construir modelos multivariados de forma sencilla.

Definición 1.49. *Una función de distribución es una función F con dominio $\overline{\mathbb{R}}$ tal que,*

(i) F es no-decreciente.

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

Definición 1.50. *Una distribución conjunta es una función H con dominio $\overline{\mathbb{R}}^2$ tal que*

(i) H es 2-creciente.

(ii) $H(x, -\infty) = 0 = H(-\infty, y)$ y $H(\infty, \infty) = 1$.

Nótese que si H es una función de distribución, entonces H es anulable (para $-\infty$ y $-\infty$) y como $Dom H = \overline{\mathbb{R}}^2$, H tiene las siguientes marginales

$$\begin{aligned} F(x) &= H(x, \infty) \\ G(y) &= H(\infty, y) \end{aligned}$$

además, como consecuencia del corolario 1.44.1 se tiene que las funciones F y G son no-crecientes y por la definición 1.50,

$$\begin{aligned} F(-\infty) &= H(-\infty, \infty) = 0; & F(\infty) &= H(\infty, \infty) = 1 \\ G(-\infty) &= H(\infty, -\infty) = 0; & G(\infty) &= H(\infty, \infty) = 1 \end{aligned}$$

por lo que F y G cumplen la definición 1.49 y pueden ser consideradas como funciones de distribución.

Una vez establecidas las dos definiciones anteriores, es necesario establecer el siguiente resultado.

Teorema 1.51. *Sea F_{XY} una función de distribución conjunta de probabilidades asociada a un vector aleatorio (X, Y) cualquiera, entonces se cumple lo siguiente*

(i) F_{XY} está definida en todo \mathbb{R}^2 y podemos extender F_{XY} a $\overline{\mathbb{R}}^2$ como sigue

$$\begin{aligned} F_{XY}(x, -\infty) &:= \lim_{y \rightarrow -\infty} F_{XY}(x, y) \\ &= \lim_{y \rightarrow -\infty} \mathbb{P}\left(\left\{\omega \in \Omega : X(\omega) \leq x\right\} \cap \left\{\omega \in \Omega : Y(\omega) \leq y\right\}\right) \\ &= \mathbb{P}(\emptyset) = 0 \\ F_{XY}(-\infty, y) &:= \lim_{x \rightarrow -\infty} F_{XY}(x, y) \\ &= \lim_{x \rightarrow -\infty} \mathbb{P}\left(\left\{\omega \in \Omega : X(\omega) \leq x\right\} \cap \left\{\omega \in \Omega : Y(\omega) \leq y\right\}\right) \\ &= \mathbb{P}(\emptyset) = 0 \\ F_{XY}(-\infty, -\infty) &:= 0 \\ F_{XY}(\infty, \infty) &:= \lim_{(x,y) \rightarrow (\infty, \infty)} F_{XY}(x, y) \\ &= \lim_{(x,y) \rightarrow (\infty, \infty)} \mathbb{P}\left(\left\{\omega \in \Omega : X(\omega) \leq x\right\} \cap \left\{\omega \in \Omega : Y(\omega) \leq y\right\}\right) \\ &= \mathbb{P}(\Omega) = 1 \end{aligned}$$

donde todos los anteriores límites existen gracias a que las medidas de probabilidad son continuas.¹⁴

¹⁴Para mas información con respecto a la continuidad de las medidas de probabilidad, véase Klenke *op. cit.*

(ii) F_{XY} es una función 2-creciente.

(iii) Las marginales de F_{XY} descritas a continuación

$$F_X(x) = \lim_{y \rightarrow \infty} F_{XY}(x, y)$$

$$F_Y(y) = \lim_{x \rightarrow \infty} F_{XY}(x, y)$$

cumplen las condiciones de la definición 1.49.

(iv) Se tiene que también las marginales F_X y F_Y son continuas por la derecha.¹⁵

Por lo cual, toda función de distribución conjunta de probabilidades F_{XY} es también una función de distribución, y sus respectivas marginales F_X y F_Y son también funciones de distribución. Esto es importante ya que cualquier resultado posterior relacionado con funciones de distribución y funciones de distribución conjunta también será válido para cualquier modelos probabilístico (X, Y) en relación a su función de distribución conjunta de probabilidades asociada F_{XY} .

En preparación para la demostración central de esta sección (Sklar (1959)), se establece el concepto de *cuasi-inversa* para una función de distribución y se establecen los lemas 1.53 y 1.54, así como el corolario 1.53.1. Las demostraciones de los lemas 1.53 y 1.54 se pueden encontrar en Schweizer y Sklar (1974).

Definición 1.52. Sea F una función de distribución. La *cuasi-inversa* de F es cualquier función $F^{(-1)}$ tal que,

(i) Si $t \in \text{Ran } F$,

$$F^{(-1)}(t) = F^{-1}(t) \tag{1.67}$$

(ii) Si $t \notin \text{Ran } F$,

$$F^{(-1)}(t) = \inf\{x \mid F(x) \geq t\} \tag{1.68}$$

Si F es estrictamente creciente, entonces F solo tiene una función cuasi-inversa y se cumple que,

$$F^{(-1)} = F^{-1} \tag{1.69}$$

¹⁵Consecuencia inmediata de la continuidad de las medidas de probabilidad y los límites de eventos, se cita nuevamente Klenke *op. cit.*

Lema 1.53. *Sea H una función de distribución conjunta con marginales F y G . Entonces existe una única subcópula C' tal que*

$$(i) \text{ Dom } C' = \text{Ran } F \times \text{Ran } G.$$

$$(ii) \text{ Para todo } x, y \in \overline{\mathbb{R}}, H(x, y) = C'(F(x), G(y)).$$

Corolario 1.53.1. *Sean H, F, G , y C' tal como en el lema 1.53, y sean $F^{(-1)}$ y $G^{(-1)}$ cuasi-inversas de F y G , respectivamente. Entonces para cualquier $(u, v) \in \text{Dom } C'$,*

$$C'(u, v) = H(F^{(-1)}(u), G^{(-1)}(v)) \quad (1.70)$$

y en el caso de que F y G sean continuas, C' es una cópula denotada como C y

$$C(u, v) = H(F^{-1}(u), G^{-1}(v)) \quad (1.71)$$

Demostración. Por el lema 1.53 se cumple que,

$$H(x, y) = C'(F(x), G(y)) \quad (1.72)$$

y realizando el cambio de variable siguiente,

$$\begin{aligned} u = F(x) &\Rightarrow x = F^{(-1)}(u) \\ v = G(y) &\Rightarrow y = G^{(-1)}(v) \end{aligned}$$

sustituyendo en la ecuación 1.72 se obtiene la igualdad siguiente,

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

En el caso de que F y G sean continuas, entonces C' es una cópula y además F y G solo tienen una cuasi-inversa, la cuales son sus funciones inversas F^{-1} , G^{-1} , respectivamente.

□

Lema 1.54. *Sea C' una subcópula. Entonces existe una cópula C tal que $C(u, v) = C'(u, v)$; $\forall (u, v) \in \text{Dom } C'$; es decir, cualquier subcópula puede ser extendida a una cópula. La extensión en general no es única.*

Teorema 1.55 (Sklar (1959)). *Sea H una función de distribución conjunta con marginales F y G . Entonces existe una cópula C tal que para todo $x, y \in \overline{\mathbb{R}}$,*

$$H(x, y) = C(F(x), G(y)) \quad (1.73)$$

Si F y G son continuas, entonces C es única; en otro caso, C está determinada de forma única en $\text{Ran } F \times \text{Ran } G$. De la misma forma, si C es una cópula y F y G son funciones de distribución, entonces la función H definida en 1.73 es una función de distribución conjunta con marginales F y G .

Demostración. Para la primera implicación, utilizando el lema 1.53 existe una única subcópula C' con $\text{Dom } C' = \text{Ran } F \times \text{Ran } G$ que puede extenderse a una cópula mediante el lema 1.54; en el caso particular que F y G sean continuas, como consecuencia del corolario 1.53.1 se tiene que $\text{Ran } F = [0, 1] = \text{Ran } G$ y la cópula en cuestión es única. Ahora bien, sea C una cópula y F y G funciones de distribución cualesquiera, definamos la función H como sigue,

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

la función H cumple con las propiedades de la definición 1.50, esto como consecuencia de la definición 1.49 y por las propiedades establecidas para una cópula en la anterior sección. Por último se cumple lo siguiente,

$$\begin{aligned} F(x) &= C(F(x), 1) = \lim_{y \rightarrow \infty} H(x, y) \\ G(y) &= C(1, G(y)) = \lim_{x \rightarrow \infty} H(x, y) \end{aligned}$$

Cumpléndose lo deseado.

□

Tal y como dice (Nelsen, 2006),

The name “copula” was chosen to emphasize the manner in which a copula “couples” a joint distribution function to its univariate margins.

por lo cual, se puede considerar a una cópula como un *enlace* o *conector* entre distribuciones univariadas marginales y distribuciones conjuntas.

En relación con una función de distribución conjunta de probabilidades F_{XY} asociada a un v.a. (X, Y) , el teorema de Sklar dice dos cosas:

(i) Siempre se puede obtener una cópula C que cumple con la condición

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

(ii) Para cualquier par de v.a. X y Y , se pueden crear muchos modelos bivariados utilizando cualquier cópula C que se desee. Esto se hace utilizando las funciones de distribución de probabilidades F_X y F_Y ; y luego escogiendo la cópula C que represente la relación de dependencia que se desea exista entre X y Y . Así se obtiene un posible modelo H para el v. a. definido como sigue,

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

El corolario 1.53.1 permite obtener la cópula que representa la dependencia en un v.a. (X, Y) . Gracias a esto, se da la definición.

Definición 1.56. *Sea (X, Y) un vector aleatorio y F_{XY} su respectiva función de distribución conjunta de probabilidades. Si F_X y F_Y son funciones continuas, entonces se define como C_{XY} a la siguiente cópula,*

$$C_{XY}(u, v) := F_{XY}(F_X^{-1}(u), F_Y^{-1}(v)) \quad (1.74)$$

A la cópula C_{XY} la denominamos como la cópula del modelo (X, Y) . La unicidad de la función C_{XY} requiere de la continuidad de las funciones F_X y F_Y .

Una cópula especial ya antes mencionada es la cópula producto $\Pi(u, v) = uv$; la importancia de esta cópula se da en el siguiente resultado

Teorema 1.57. *Sea X y Y dos v.a. continuas cualesquiera y sea C_{XY} la cópula del modelo (X, Y) . Entonces X y Y son variables aleatorias independientes si y sólo si,*

$$C_{XY}(u, v) = uv = \Pi(u, v)$$

Demostración. Esto es consecuencia inmediata del teorema 1.16.

□

Otro resultado interesante e útil en relación a cópulas es el comportamiento que estas tienen con respecto a funciones continuas estrictamente monótonas.

Lema 1.58. *Sea X una variable aleatoria continua y sea $\alpha : \text{Ran } X \mapsto \mathbb{R}$, una función continua estrictamente monótona, entonces la variable aleatorio $\alpha(X)$ es también continua y se cumple que*

(i) *Si α es estrictamente creciente,*

$$F_{\alpha(X)}(t) = F_X(\alpha^{-1}(t))$$

(ii) *Si α es estrictamente decreciente,*

$$F_{\alpha(X)}(t) = 1 - F_X(\alpha^{-1}(t))$$

Demostración. Sean α_1 y α_2 funciones continuas estrictamente monótonas, la primera creciente y la segunda decreciente. Entonces se cumple lo siguiente,

$$\begin{aligned} F_{\alpha_1(X)}(t) &= \mathbb{P}(\alpha_1(X) \leq t) = \mathbb{P}(X \leq \alpha_1^{-1}(t)) = F_X(\alpha_1^{-1}(t)) \\ F_{\alpha_2(X)}(t) &= \mathbb{P}(\alpha_2(X) \leq t) = \mathbb{P}(X \geq \alpha_2^{-1}(t)) = 1 - F_X(\alpha_2^{-1}(t)) \end{aligned}$$

y lo demás es consecuencia inmediata de que α_1 y α_2 son continuas.

□

Con el anterior resultado, se establece el siguiente teorema.

Teorema 1.59. *Sean X y Y dos variables aleatorias continuas con cópula C_{XY} . Si α y β son funciones continuas estrictamente crecientes en $\text{Ran } X$ y $\text{Ran } Y$, respectivamente, entonces $C_{\alpha(X)\beta(Y)} = C_{XY}$. Es decir, C_{XY} es invariante con respecto a transformaciones continuas y estrictamente crecientes en X y Y .*

Demostración. Sean F_X , F_α , F_Y y G_β las funciones de distribución de X , $\alpha(X)$, Y y $\beta(Y)$, respectivamente. Como α y β son continuas y estrictamente crecientes, entonces se tiene que $\alpha(X)$ y $\beta(Y)$ también son variables aleatorias continuas y además,

$$\begin{aligned} F_\alpha(x) &= F_X(\alpha^{-1}(x)) \\ G_\beta(y) &= F_Y(\beta^{-1}(y)) \end{aligned}$$

por lo que $C_{\alpha(X)\beta(Y)}$ cumple lo siguiente,

$$\begin{aligned}
C_{\alpha(X)\beta(Y)}(F_\alpha(x), G_\beta(y)) &= \mathbb{P}(\alpha(X) \leq x, \beta(Y) \leq y) \\
&= \mathbb{P}(X \leq \alpha^{-1}(x), Y \leq \beta^{-1}(y)) \\
&= C_{XY}(F_X(\alpha^{-1}(x)), F_Y(\beta^{-1}(y))) \\
&= C_{XY}(F_\alpha(x), F_\beta(y))
\end{aligned}$$

Y como X y Y son v.a. continuas, $\text{Ran } F_\alpha = \text{Ran } F_\beta = \mathbf{I}$, entonces se sigue que $C_{\alpha(X)\beta(Y)} = C_{XY}$ en \mathbf{I}^2 .

□

El resultado en 1.59 se puede extender con el siguiente teorema.

Teorema 1.60. *Sea X y Y variables continuas con cópula C_{XY} . Sea α y β son funciones continuas estrictamente monótonas en $\text{Ran } X$ y $\text{Ran } Y$, respectivamente.*

(i) *Si α es estrictamente creciente y β es estrictamente decreciente, entonces:*

$$C_{\alpha(X)\beta(Y)}(u, v) = u - C_{XY}(u, 1 - v) \quad (1.75)$$

(ii) *Si α es estrictamente decreciente y β es estrictamente creciente, entonces:*

$$C_{\alpha(X)\beta(Y)}(u, v) = v - C_{XY}(1 - u, 1) \quad (1.76)$$

(iii) *Si α y β son estrictamente decrecientes, entonces:*

$$C_{\alpha(X)\beta(Y)}(u, v) = u + v - 1 + C_{XY}(1 - u, 1 - v) \quad (1.77)$$

a las cópulas 1.75, 1.76 y 1.77 se les denota como C^H , C^V y C^{HV} , respectivamente, y se les denomina rotaciones horizontales, verticales y completas de C , respectivamente. Por lo tanto, para cualquier cópula C se tiene lo siguiente,

$$\begin{aligned}
C^H(u, v) &:= u - C(u, 1 - v) \\
C^V(u, v) &:= v - C(1 - u, v) \\
C^{HV}(u, v) &:= u + v - 1 + C(1 - u, 1 - v)
\end{aligned}$$

Demostración. Sean F_X , F_α , F_Y y G_β las funciones de distribución de X , $\alpha(X)$, Y y $\beta(Y)$, respectivamente. Se demuestran los casos (i) y (iii),

(i) Si α es estrictamente creciente y β es estrictamente decreciente, entonces:

$$\begin{aligned}
C_{\alpha(X)\beta(Y)}(F_\alpha(x), G_\beta(y)) &= \mathbb{P}(\alpha(X) \leq x, \beta(Y) \leq y) \\
&= \mathbb{P}(X \leq \alpha^{-1}(x), Y \geq \beta^{-1}(y)) \\
&= \mathbb{P}(X \leq \alpha^{-1}(x)) - \mathbb{P}(X \leq \alpha^{-1}(x), Y \leq \beta^{-1}(y)) \\
&= F_X(\alpha^{-1}(x)) - C_{XY}(F_X(\alpha^{-1}(x)), F_Y(\beta^{-1}(x))) \\
&= F_\alpha(x) - C_{XY}(F_\alpha(x), 1 - F_\beta(y))
\end{aligned}$$

es decir, para todo $(u, v) \in \mathbf{I}^2$,

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

(iii) Si α y β son estrictamente decrecientes, entonces:

$$\begin{aligned}
C_{\alpha(X)\beta(Y)}(F_\alpha(x), G_\beta(y)) &= \mathbb{P}(\alpha(X) \leq x, \beta(Y) \leq y) \\
&= \mathbb{P}(X \geq \alpha^{-1}(x), Y \geq \beta^{-1}(y)) \\
&= 1 - \mathbb{P}(X \leq \alpha^{-1}(x)) - \mathbb{P}(Y \leq \beta^{-1}(y)) \\
&\quad + \mathbb{P}(X \leq \alpha^{-1}(x), Y \leq \beta^{-1}(y)) \\
&= 1 - F_X(\alpha^{-1}(x)) - F_Y(\beta^{-1}(y)) \\
&\quad + C_{XY}(1 - F_\alpha(x), 1 - F_\beta(y)) \\
&= F_\alpha(x) + F_\beta(y) - 1 \\
&\quad + C_{XY}(1 - F_\alpha(x), 1 - F_\beta(y))
\end{aligned}$$

es decir, para todo $(u, v) \in \mathbf{I}^2$,

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

mientras que el caso (ii) es equivalente a (i). □

Por lo cual, toda medida de dependencia que se defina sobre C_{XY} es invariante con respecto a transformaciones continuas monótonas crecientes. A continuación, se establece una clasificación para cópulas que tiene una interpretación probabilística bastante interesante y esta información se analiza mediante un ejemplo.

Definición 1.61. Sea C una cópula cualquiera, se denota,

$$C(u, v) = A_C(u, v) + S_C(u, v) \tag{1.78}$$

donde,

$$A_C(u, v) := \int_0^u \int_0^v \frac{\partial^2}{\partial s \partial t} C(s, t) dt ds \quad (1.79)$$

$$S_C(u, v) := C(u, v) - A_C(u, v) \quad (1.80)$$

- (i) Si $C \equiv A_C$ en \mathbf{I}^2 , entonces se dice que C es absolutamente continua.
- (ii) Si $C \equiv S_C$ en \mathbf{I}^2 , entonces se dice que C es singular. Y para este caso se cumple que $\partial^2 / \partial u \partial v C(u, v) = 0$ casi en todo \mathbf{I}^2 .
- (iii) En otro caso, se dice que C tiene un componente absolutamente continuo A_C y un componente singular S_C .

Ejemplo 1.62. Para este ejemplo se toman las siguiente cópulas y se analizan sus derivadas parciales mixtas,

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

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

$$MO(u, v) = \min\{u^{1/2}v, uv^{1/3}\}$$

- (i) Para la cópula M se cumple que,

$$\frac{\partial^2}{\partial u \partial v} M(u, v) = \begin{cases} 0 & \text{Si } u \neq v \\ \text{no existe} & \text{Si } u = v \end{cases}$$

- (ii) Para la cópula Π se cumple que,

$$\frac{\partial^2}{\partial u \partial v} \Pi(u, v) = 1$$

- (iii) Para la cópula MO se cumple que,

$$\frac{\partial^2}{\partial u \partial v} MO(u, v) = \begin{cases} u^{-1/2} & \text{Si } u^{1/2} > v^{2/3} \\ v^{-1/2} & \text{Si } u^{1/2} < v^{2/3} \\ \text{no existe} & \text{Si } u^{1/2} = v^{2/3} \end{cases}$$

Si los v.a. (U_M, V_M) , (U_Π, V_Π) y (U_{MO}, V_{MO}) tienen como funciones de distribución $M_{\mathbb{R}^2}$, $\Pi_{\mathbb{R}^2}$ y $MO_{\mathbb{R}^2}$, respectivamente, entonces se cumple lo siguiente

(i) Para el v.a. (U_M, V_M) , se tiene que la cópula M es singular y además:

$$\mathbb{P}(U_M = V_M) = 1$$

(ii) Para el v.a. (U_Π, V_Π) , se tiene que la cópula Π es absolutamente continua.

(iii) Para el v.a. (U_{MO}, V_{MO}) , se tiene que la cópula MO tiene un componente absolutamente continuo A_{MO} , un componente singular S_{MO} y además:

$$0 < \mathbb{P}(U_{MO}^{1/2} = V_{MO}^{2/3}) < 1$$

Se termina esta sección dando dos resultados que son de especial importancia para la simulación o regresión de vectores aleatorios.

Lema 1.63. Sea X una variable aleatoria cualquiera, F_X su respectiva función de distribución de probabilidades y $U \sim U(0, 1)$; entonces se cumple que si $Y := F_X^{(-1)}(U)$,

$$F_Y(y) = F_X(y) \tag{1.81}$$

En el caso particular de que F_X sea estrictamente continua se cumple que

$$Y = F_X^{-1}(U) \tag{1.82}$$

Demostración. Sea $Y := F_X^{(-1)}(U)$, entonces F_Y queda como sigue,

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leq y) = \mathbb{P}(F_X^{(-1)}(U) \leq y) = \mathbb{P}\left(F_X(F_X^{(-1)}(U)) \leq F_X(y)\right) \\ &= \mathbb{P}(U \leq F_X(y)) = F_U(F_X(y)) = F_X(y) \end{aligned}$$

Si F_X es estrictamente creciente, se tiene que la única cuasi-inversa de F_X es su inversa F_X^{-1} , y

$$Y = F_X^{-1}(U)$$

□

Teorema 1.64. Sea (X, Y) un vector aleatorio absolutamente continuo y C la cópula asociada al modelo, entonces se cumple para el vector aleatorio (U, V) con $U := F_X(X)$; $V := F_Y(Y)$,

(i) $U, V \sim U(0, 1)$. Es decir, ambas variables aleatorias son transformadas en uniformes cero-uno, sin importar que tipo de variables aleatorias sean X y Y .

(ii) $F_{UV} = C_{\mathbb{R}^2}$. Es decir, el vector (U, V) mantiene la dependencia existente en la cópula C .

Demostración. Realizando la demostración por partes,

(i) Para la primera parte, como (X, Y) es un v.a. abs. continuo, entonces X y Y son v.a. abs. continuas y se cumple que,

$$\begin{aligned} F_U(u) &= \mathbb{P}(U \leq u) = \mathbb{P}(F_X(X) \leq u) = \mathbb{P}(X \leq F_X^{-1}(u)) = F_X(F_X^{-1}(u)) \\ &= u\mathbf{1}_{\{0 < u < 1\}} + \mathbf{1}_{\{u \geq 1\}} \\ F_V(v) &= \mathbb{P}(V \leq v) = \mathbb{P}(F_Y(Y) \leq v) = \mathbb{P}(Y \leq F_Y^{-1}(v)) = F_Y(F_Y^{-1}(v)) \\ &= v\mathbf{1}_{\{0 < v < 1\}} + \mathbf{1}_{\{v \geq 1\}} \end{aligned}$$

por lo cual, $U, V \sim U(0, 1)$.

(ii) Ya que F_X y F_Y son funciones continuas y estrictamente crecientes se cumple, gracias al teorema 1.59, que,

$$C_{UV} = C_{F_X(X)F_Y(Y)} = C_{XY}$$

□

1.4. Simulación y regresión mediante cópulas

En esta sección se establece una generalización para el concepto de regresión establecido en la definición 1.18, esto se hace utilizando el teorema 1.64. Se hace notar la ventaja que implican las cópulas a la hora de generalizar procesos de regresión y simulación de variables aleatorias.

1.4.1. Simulación mediante cópulas

Para la parte de simulación, se inicia con el concepto de muestra aleatoria.

Definición 1.65. Sean $\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n\}$ un conjunto de variables aleatorias (o vectores aleatorios). Decimos que \mathbf{X} es una muestra aleatoria (denotado como *m.a.*), si se cumplen las siguientes dos condiciones.

(i) Si para $i \neq j$:

$$F_{\mathbf{X}_i} = F_{\mathbf{X}_j} \quad (1.83)$$

(ii) La distribución conjunta de probabilidades de \mathbf{X} es de la siguiente forma:

$$F_{\mathbf{X}}(\mathbf{x}) = F_{\mathbf{X}_1 \mathbf{X}_2 \dots \mathbf{X}_n}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) = \prod_{i=1}^n F_{\mathbf{X}_i}(\mathbf{x}_i) \quad (1.84)$$

es decir, el conjunto \mathbf{X} es un conjunto de variables aleatorias (o vectores aleatorios) independientes e idénticamente distribuidos.

Una vez dada la definición 1.65, utilizando el Teorema de Sklar 1.55 se establece el siguiente resultado.

Teorema 1.66. Sea (X_1, X_2) un vector aleatorio cualquiera, $F_{X_1 X_2}$ su respectiva función de distribución conjunta de probabilidades y C su cópula asociada.

Si $U_1, U_2 \sim U(0, 1)$ tales que $F_{U_1 U_2} = C_{\mathbb{R}^2}$, entonces para las variables aleatorias $Y_i := F_{X_i}^{(-1)}(U_i)$; $i = 1, 2$,

$$F_{Y_1 Y_2}(y_1, y_2) = F_{X_1 X_2}(y_1, y_2) \quad (1.85)$$

En el caso particular de que F_{X_i} sea estrictamente continua se cumple que

$$Y_i = F_{X_i}^{-1}(U_i) \quad (1.86)$$

Demostración. Tomemos $Y_i := F_{X_i}^{(-1)}(U_i)$; $i = 1, 2$ y encontremos $F_{Y_1 Y_2}$,

$$\begin{aligned} F_{Y_1 Y_2}(y_1, y_2) &= \mathbb{P}(Y_1 \leq y_1, Y_2 \leq y_2) = \mathbb{P}(F_{X_1}^{(-1)}(U_1) \leq y_1, F_{X_2}^{(-1)}(U_2) \leq y_2) \\ &= \mathbb{P}(U_1 \leq F_{X_1}(y_1), U_2 \leq F_{X_2}(y_2)) = C_{\mathbb{R}^2}(F_{X_1}(y_1), F_{X_2}(y_2)) \end{aligned}$$

como $(F_{X_1}(y_1), F_{X_2}(y_2)) \in \mathbf{I}^2$ y por como está definida $C_{\mathbb{R}^2}$ en la definición 1.38, se cumple que,

$$F_{Y_1 Y_2}(y_1, y_2) = C_{\mathbb{R}^2}(F_{X_1}(y_1), F_{X_2}(y_2)) = C(F_{X_1}(y_1), F_{X_2}(y_2))$$

donde C es la cópula asociada al vector (X, Y) . Por el teorema de Sklar se cumple que,

$$F_{Y_1 Y_2}(y_1, y_2) = C(F_{X_1}(y_1), F_{X_2}(y_2)) = F_{X_1 X_2}(y_1, y_2)$$

En el caso de que F_{X_i} sea estrictamente creciente, entonces también se cumple que la única cuasi-inversa de F_{X_i} es $F_{X_i}^{-1}$, por lo cual,

$$Y_i = F_{X_i}^{-1}(U_i)$$

□

El teorema 1.66 permite crear el siguiente algoritmo para la creación de una muestra aleatoria observada de cualquier vector aleatorio (X, Y) .

Algoritmo 1.66.1. *Sea (X, Y) un vector aleatorio cualquiera con función de distribución conjunta de probabilidades F_{XY} . F_X y F_Y son las funciones de distribución de probabilidades marginales y sea C_{XY} la cópula asociada al vector aleatorio (X, Y) . Se desea obtener una muestra aleatoria de tamaño n del vector aleatorio (X, Y) . Esto se puede hacer con los siguientes pasos:*

1. *Se simula una m.a. de tamaño n para una v.a. $U \sim U(0, 1)$. Dicha muestra se denota como $\{u_1, u_2, \dots, u_n\}$.*
2. *Se simula un valor aleatorio para cada una de las siguiente n v.a. $\{V_1, V_2, \dots, V_n\}$, cuyas respectivas funciones de distribución de probabilidades son las siguientes.*

$$F_{V_i|U_i}(v_i | u_i) = \frac{\partial}{\partial u} C_{XY}(u, v) \Big|_{u=u_i} \quad i \in \{1, 2, \dots, n\}$$

Dichos valores se denotan como $\{v_1, v_2, \dots, v_n\}$.

3. *Se realizan las siguientes transformaciones para la m.a. $\{(u_1, v_1), (u_2, v_2), \dots, (u_n, v_n)\}$,*

$$x_i = F_X^{(-1)}(u_i)$$

$$y_i = F_Y^{(-1)}(v_i)$$

obteniendo así el conjunto $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ que representa una m.a. del vector (X, Y) .

La implementación del algoritmo 1.66.1 se da en el siguiente ejemplo.

Ejemplo 1.67. *Sean (X_1, Y_1) y (Y_1, Y_2) dos vectores aleatorios con las siguientes funciones de distribución conjuntas de probabilidades.*

$$F_{X_1 Y_1}(x, y) = (1 - e^{-x}) \varphi_1(y) \exp \left\{ \left((-\log(1 - e^{-x}))^8 + (-\log \varphi_1(y))^8 \right)^{-1/8} \right\} \mathbf{1}_{\{x, y > 0\}}$$

$$F_{X_2 Y_2}(x, y) = \varphi_2(x) (1 - 1/y^3) \exp \left\{ \left((-\log \varphi_2(x))^8 + (-\log(1 - 1/y^3))^8 \right)^{-1/8} \right\} \mathbf{1}_{\{x, y > 0\}}$$

donde φ_1 y φ_2 están definidas como sigue,

$$\varphi_1(y) = \int_{-\infty}^y \frac{1}{t\sqrt{2\pi}} \exp\left\{-\frac{t^2}{2}\right\} dt$$

$$\varphi_2(x) = \int_0^x \frac{1}{t\sqrt{2\pi}} \exp\left\{-\frac{(\log t)^2}{2}\right\} dt$$

Para estos dos modelos, se nota lo siguiente,

$$C_{X_1Y_1}(u, v) = uv \exp\left\{\left((-\log u)^8 + (-\log v)^8\right)^{-1/8}\right\} = C_{X_2Y_2}(u, v)$$

por lo cual, ambos vectores aleatorios cuentan con la misma cópula asociada pero distintas marginales, por lo que para la simulación de la muestra aleatoria, los primeros 2 pasos del algoritmo 1.66.1 son iguales para ambos modelos y el único que difiere es el paso 3. En la figura 1.2 se puede ver los resultados de una simulación de tamaño 1000 para cada vector aleatorio.

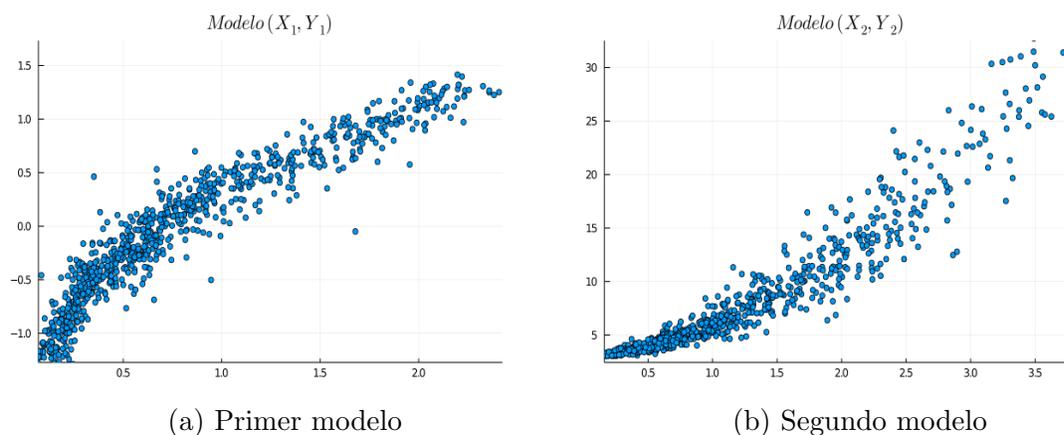


Figura 1.2: Simulación para los modelos (X_1, Y_1) y (X_2, Y_2) .

Se puede notar como las distribuciones univariadas $F_{X_1}, F_{X_2}, F_{Y_1}, F_{Y_2}$ afectan el mapa de dispersión de los datos. A simple vista, podría llegarse a la conclusión errónea de que las figuras 1.2a y 1.2b tienen dependencia distinta, aunque se sabe por construcción del ejemplo que esto no es el caso, este ejemplo hace notar como las distribuciones univariadas agregan “ruido” a los mapas de dispersión.

1.4.2. Regresión mediante cópulas

Para esta sección, se extiende el concepto de regresión a cualquier cuantil $\alpha \in]0, 1[$.

Definición 1.68. Sea (X, Y) un vector aleatorio cualquiera donde F_{XY} es la función de distribución conjunta de probabilidades del vector aleatorio anterior y sea $\alpha \in]0, 1[$ un número cualquiera. Se define la curva de regresión de nivel α para Y dado el evento $\{X = x\}$ para cualquier $x \in \mathbb{R}$ (siempre y cuando la función esté bien definida en dicho punto),

$$\mathcal{RC}_{Q_\alpha}(Y | X = x) := Q_\alpha(Y | X = x) \quad (1.87)$$

donde el valor $Q_\alpha(Y | X = x)$ queda definida como el punto que cumple la siguiente condición,

$$\mathbb{P}(Y \leq Q_\alpha(Y | X = x) | X = x) = \alpha \quad (1.88)$$

Observación 1.68.1. Para el caso en el que $\alpha = \frac{1}{2}$ (siempre y cuando la función esté bien definida en dicho punto),

$$\mathcal{RC}_{Q_{1/2}}(Y | X = x) \equiv \mathcal{RC}_M(Y | X = x)$$

y en el caso de $0 < \alpha < \beta < 1$ se puede crear un intervalo de confianza con probabilidad $(\beta - \alpha)$ de la siguiente forma,

$$\left] Q_\alpha(Y | X = x), Q_\beta(Y | X = x) \right]$$

ya que se cumple que,

$$\begin{aligned} & \mathbb{P}\left(Y \in \left] Q_\alpha(Y | X = x), Q_\beta(Y | X = x) \right] \middle| X = x\right) \\ &= F_{Y|X}\left(Q_\beta(Y | X = x) \middle| x\right) - F_{Y|X}\left(Q_\alpha(Y | X = x) \middle| x\right) \\ &= \beta - \alpha \end{aligned}$$

El siguiente resultado crea una relación entre las curvas de regresión y la cópula del modelo de interés.

Teorema 1.69. *Sea (X, Y) un vector aleatorio absolutamente continuo, $C = C_{XY}$ la cópula perteneciente al modelo (X, Y) y sea $\alpha \in (0, 1)$ un número cualquiera. Entonces se cumple (para los puntos donde la curva de regresión esté definida),*

$$Q_\alpha(Y | X = x) = F_Y^{-1}\left(Q_\alpha(V | U = F_X(x))\right) \quad (1.89)$$

donde (U, V) es un vector aleatorio de uniformes cero-uno tales que $F_{UV} = C_{\mathbb{R}^2}$.

Demostración. Sea (X, Y) el v. a. abs. continuo en cuestión, entonces por consecuencia del teorema 1.64,

$$\begin{aligned} \alpha &= F_{Y|X}(Q_\alpha(Y | X = x) | x) = \mathbb{P}(Y \leq Q_\alpha(Y | X = x) | X = x) \\ &= \mathbb{P}\left(F_Y(Y) \leq F_Y(Q_\alpha(Y | X = x)) \mid F_X(X) = F_X(x)\right) \\ &= \mathbb{P}\left(V \leq F_Y(Q_\alpha(Y | X = x)) \mid U = F_X(x)\right) \end{aligned}$$

por lo cual, se cumple que,

$$F_Y(Q_\alpha(Y | X = x)) = Q_\alpha(V | U = F_X(x))$$

y sabiendo que F_Y^{-1} existe,

$$Q_\alpha(Y | X = x) = F_Y^{-1}\left(Q_\alpha(V | U = F_X(x))\right)$$

□

Ejemplo 1.70. *Sean (X_1, Y_1) y (X_2, Y_2) dos v.a. abs. continuos y sean H_1 y H_2 , sus funciones de distribución conjunta de probabilidades, respectivamente.*

$$\begin{aligned} H_1(x, y) &= \left[(1 - e^{-x})^{-5} + \left(1 - \frac{1}{y^3}\right)^{-5} - 1 \right]^{-1/5} \mathbf{1}_{\{(x, y) \in (0, \infty) \times (1, \infty)\}} \\ H_2(x, y) &= \frac{1}{8} \log \left[1 + \frac{(e^{8(1-e^{-x})} - 1)(e^{8(1-1/y^3)} - 1)}{e^8 - 1} \right] \mathbf{1}_{\{(x, y) \in (0, \infty) \times (1, \infty)\}} \end{aligned}$$

para estos dos modelos, las distribuciones univariadas son las siguientes,

$$\begin{aligned} F_{X_1}(z) &= H_1(z, +\infty) = (1 - e^{-z}) \mathbf{1}_{\{z > 0\}} = H_2(z, +\infty) = F_{X_2}(z) \\ F_{Y_1}(z) &= H_1(+\infty, z) = \left(1 - \frac{1}{z^3}\right) \mathbf{1}_{\{z > 1\}} = H_2(+\infty, z) = F_{Y_2}(z) \end{aligned}$$

es decir, marginalmente $X_1 = X_2$ y $Y_1 = Y_2$. Las cópulas pertenecientes a los v.a. (X_1, Y_1) y (X_2, Y_2) son las siguientes,

$$C^1(u, v) = \left(u^{-5} + v^{-5} - 1\right)^{-1/5}$$

$$C^2(u, v) = \frac{1}{8} \log \left[1 + \frac{(e^{8u} - 1)(e^{8v} - 1)}{e^8 - 1} \right]$$

Ahora bien, se desea construir para ambos modelos curvas de regresión que representen los cuantiles 25%, 50% y 75%, respectivamente. Esto se puede hacer utilizando el teorema 1.69, primero se obtienen las derivadas parciales de C^1 y C^2 .

$$\frac{\partial}{\partial u} C^1(u, v) = u^{-6} \left(u^{-5} + v^{-5} - 1\right)^{-6/5}$$

$$\frac{\partial}{\partial u} C^2(u, v) = \frac{e^{8u}(e^{8v} - 1)}{e^8 - 1 + (e^{8u} - 1)(e^{8v} - 1)}$$

y utilizando el teorema 1.47, si (U_1, V_1) y (U_2, V_2) son v.a. abs. continuos con funciones de distribución conjuntas de probabilidades $C_{\mathbb{R}^2}^1$ y $C_{\mathbb{R}^2}^2$, respectivamente. Entonces se cumple que,

$$F_{V_1|U_1}(v|u) = u^{-6} \left(u^{-5} + v^{-5} - 1\right)^{-6/5}$$

$$F_{V_2|U_2}(v|u) = \frac{e^{8u}(e^{8v} - 1)}{e^8 - 1 + (e^{8u} - 1)(e^{8v} - 1)}$$

por lo que se puede obtener $Q_\alpha(V_1|U_1 = u)$ y $Q_\alpha(V_2|U_2 = u)$ encontrando las funciones inversas de $F_{V_1|U_1}(v|u)$ y $F_{V_2|U_2}(v|u)$, respectivamente, quedando como sigue,

$$Q_\alpha(V_1|U_1 = u) = \left[1 + \left(\alpha^{-5/6} - 1\right)u^{-5} \right]^{-1/5}$$

$$Q_\alpha(V_2|U_2 = u) = \frac{1}{8} \log \left[1 + \frac{(e^8 - 1)\alpha}{1 + (e^{8u} - 1)(1 - \alpha)} \right]$$

ahora bien, se necesita obtener $F_Y^{-1} = F_{Y_1}^{-1} = F_{Y_2}^{-1}$,

$$F_Y^{-1}(z) = \left(\frac{1}{1 - z}\right)^{1/3}$$

por lo cual se cumple que,

$$Q_\alpha(Y_1 | X_1 = x) = \left(1 - \left[1 + (\alpha^{-5/6} - 1)(1 - e^{-x})^{-5} \right]^{-1/5} \right)^{-1/3}$$

$$Q_\alpha(Y_2 | X_2 = x) = \left(1 - \frac{1}{8} \log \left[1 + \frac{(e^8 - 1)\alpha}{1 + (e^8(1 - e^{-x}) - 1)(1 - \alpha)} \right] \right)^{-1/3}$$

En la figura 1.3, se puede ver como se comportan las curvas de regresión con respecto a una simulación de tamaño 1000 para cada v.a. (X_1, Y_1) y (X_2, Y_2) .

En la figura 1.3a, se puede notar como conforme el valor x incrementa, entonces $Q_\alpha(Y_1 | X_1 = x)$ incrementa, mientras que en el caso de 1.3b el valor de $Q_\alpha(Y_2 | X_2 = x)$ decrece.

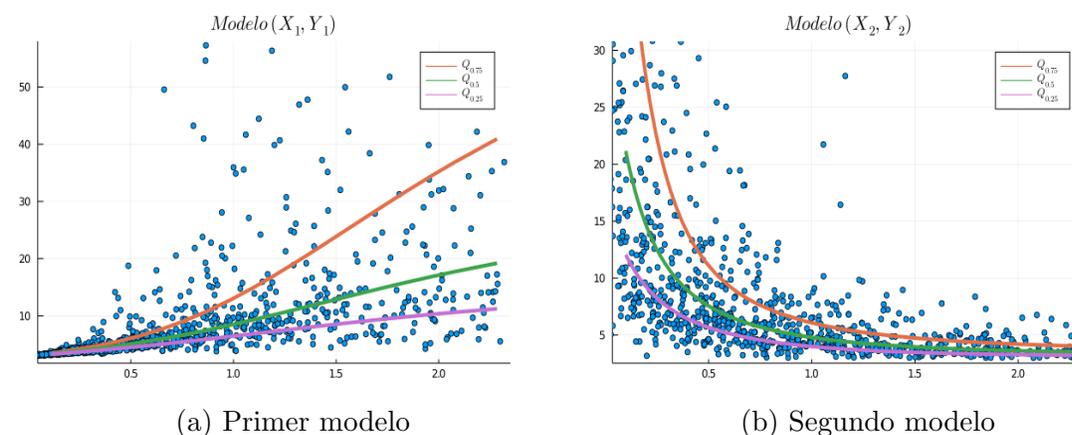


Figura 1.3: Curvas de regresión $Q_\alpha(Y | X = x)$ para los v.a. abs. continuos (X_1, Y_1) y (X_2, Y_2) .

1.5. Cópulas Arquimedianas

En esta sección se construye una familia paramétrica de cópulas denominada *arquimediana*, la cual permite modelar distintos grados de dependencia y de distintas formas. Esta sección esta fundamentada en información extraída de Nelsen (2006) y Genest *et. al.* (1986).

Se inicia esta sección planteando un tipo de funciones que se convertirán en “generadoras” de cópulas bivariadas. Para esto se establece el concepto de funciones *pseudo-inversas*.

Definición 1.71. Sea $\varphi : \mathbf{I} \mapsto [0, \infty]$ una función continua, estrictamente decreciente tal que $\varphi(1) = 0$. La pseudo-inversa de φ es la función $\varphi^{[-1]} : [0, \infty] \mapsto \mathbf{I}$ dada por,

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t); & 0 \leq t \leq \varphi(0) \\ 0; & \varphi(0) \leq t \leq +\infty \end{cases} \quad (1.90)$$

Observación 1.71.1. Nótese que $\varphi^{[-1]}$ es continua y no creciente en $[0, \infty]$, y estrictamente decreciente en $[0, \varphi(0)]$. Además, $\varphi^{[-1]}(\varphi(u)) = u$ en \mathbf{I} , y

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

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

Ejemplo 1.72. Sean φ_1 y φ_2 las siguientes dos funciones definidas sobre \mathbf{I} ,

$$\begin{aligned} \varphi_1(t) &= (1 - t)^2 \\ \varphi_2(t) &= (-\log t)^2 \end{aligned}$$

para las funciones φ_1 y φ_2 se cumple que,

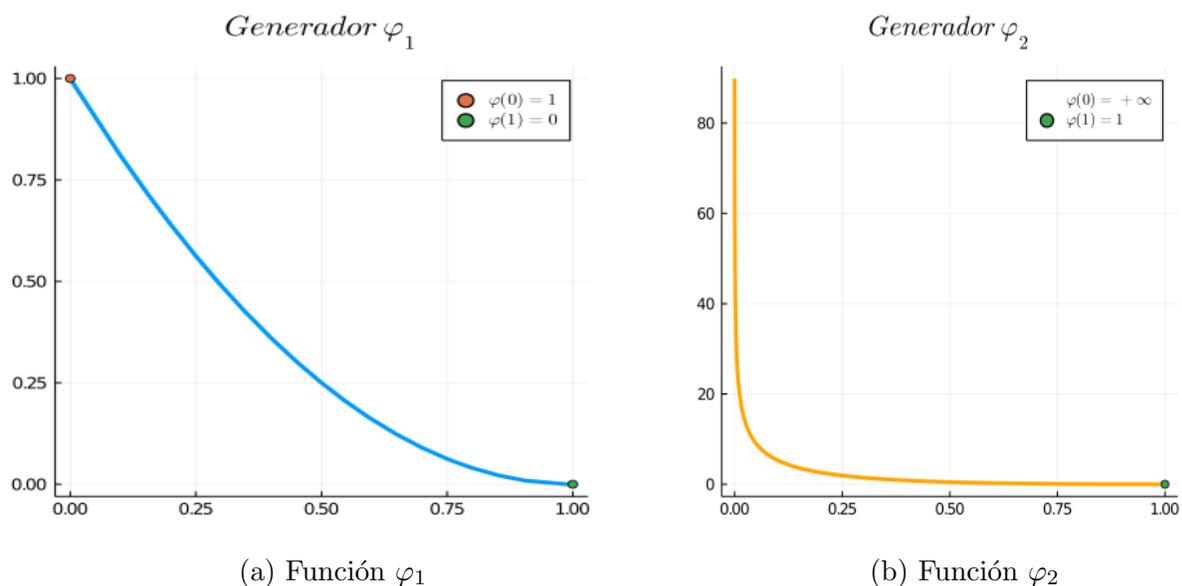
$$\varphi_1(1) = 0; \quad \varphi_2(1) = 0; \quad \varphi_1(0) = 1; \quad \varphi_2(0) = +\infty$$

por lo cual las pseudo-inversas de φ_1 y φ_2 , respectivamente, son las siguientes,

$$\begin{aligned} \varphi_1^{[-1]}(t) &= \begin{cases} 1 - \sqrt{t}; & 0 \leq t \leq 1 \\ 0; & 1 \leq t \leq \infty \end{cases} \\ \varphi_2^{[-1]}(t) &= \exp\left\{-\sqrt{t}\right\} \end{aligned}$$

las imágenes de las funciones φ_1 y φ_2 se muestran en la figura 1.4; mientras que la figura 1.5 muestra las gráficas de $\varphi_1^{[-1]}$ y $\varphi_2^{[-1]}$.

Se puede notar como $\varphi_1(0) = 1$ y $\varphi_2(0) = \infty$, por esa causa la gráfica 1.5a se comporta tal y como se ve.

Figura 1.4: Funciones φ_1 y φ_2

Se establecen en los lemas 1.73 y 1.74 las propiedades que deben cumplir las funciones φ descritas en la definición 1.71 para poder generar cópulas bivariadas.

Lema 1.73. Sea $\varphi : \mathbf{I} \mapsto [0, \infty]$ una función continua, estrictamente decreciente tal que $\varphi(1) = 0$, y sea $\varphi^{[-1]}$ la pseudo-inversa de φ definida en 1.72. Sea $C : \mathbf{I}^2 \mapsto \mathbf{I}$ una función dada por,

$$C(u, v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)) \quad (1.91)$$

Entonces C satisface las siguientes dos condiciones

$$C(u, 0) = 0; \quad C(0, v) = 0 \quad (1.92)$$

$$C(u, 1) = u; \quad C(1, v) = v \quad (1.93)$$

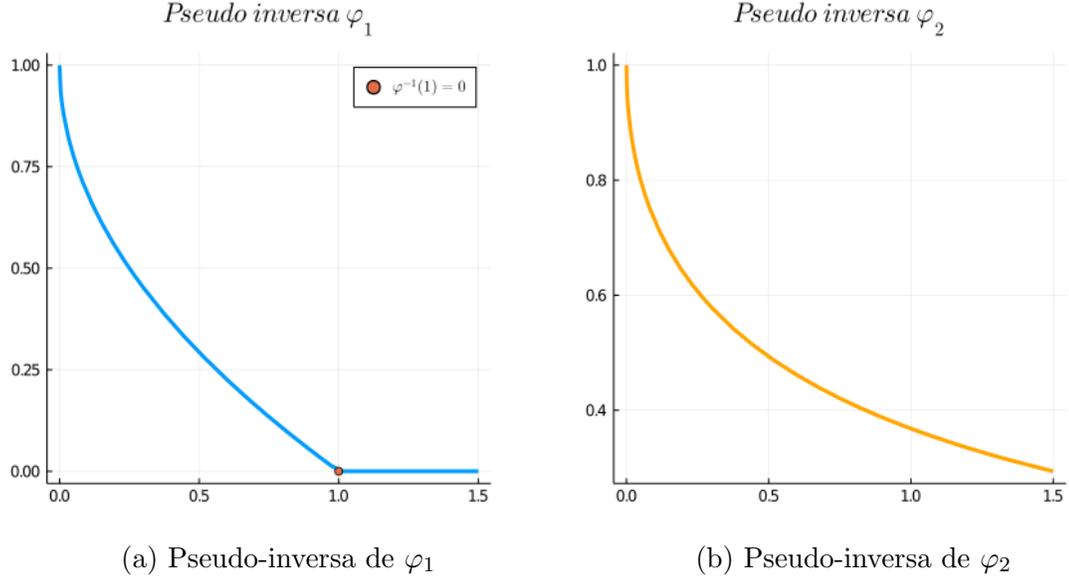
es decir, C satisface la propiedad (i) establecida en la proposición 1.35.

Demostración. Para el primer caso, como $\varphi(u), \varphi(v) \geq 0; \forall u, v \in \mathbf{I}$,

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

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

y para el segundo caso,

Figura 1.5: Pseudo-inversas φ_1 y φ_2

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

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

quedando demostrado el lema 1.73. □

Lema 1.74. Sean φ , $\varphi^{[-1]}$ y C definidas como en el lema 1.73. Entonces C es 2-creciente si y sólo si para todo $v \in \mathbf{I}$,

$$C(u_2, v) - C(u_1, v) \leq u_2 - u_1 \tag{1.94}$$

siempre que $u_1 \leq u_2$.

Demostración. Nótese que 1.94 es equivalente a la siguiente condición,

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

por lo cual la primera implicación del lema 1.74 es clara.

Ahora bien, si C satisface la ecuación 1.94. Se escojen $v_1, v_2 \in \mathbf{I}$ tales que $v_1 \leq v_2$, y como consecuencia del lema 1.73,

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

Pero C es continua (ya que φ y $\varphi^{[-1]}$ lo son), y entonces existe un $t \in \mathbf{I}$ tal que $C(t, v_2) = v_1$ y como $\varphi^{[-1]}$ es estrictamente decreciente en $[0, \varphi(0)]$,

$$\begin{aligned} C(t, v_2) &= v_1 \\ \varphi^{[-1]}(\varphi(t) + \varphi(v_2)) &= v_1 = \varphi^{[-1]}(\varphi(v_1)) \end{aligned}$$

si y sólo si,

$$\varphi(t) + \varphi(v_2) = \varphi(v_1)$$

Entonces,

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

cumpléndose que,

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

por lo cual C es 2-creciente. □

Los lemas 1.73 y 1.74 permiten probar el teorema mas importante de esta sección.

Teorema 1.75. *Sea $\varphi : \mathbf{I} \mapsto [0, \infty]$ una función continua, estrictamente decreciente tal que $\varphi(1) = 0$, y sea $\varphi^{[-1]}$ la pseudo-inversa de φ definida en 1.73. Entonces la función $C : \mathbf{I}^2 \mapsto \mathbf{I}$ dada por 1.74 es una cópula si y sólo si φ es convexa.*

Demostración. La demostración se puede ver en Nelsen *op. cit.* □

El teorema 1.75 permite crear una gran familia de cópulas, y a las funciones φ se les denomina “generadoras aditivos”. Para terminar esta sección se mencionan algunas de las familias de cópulas arquimedianas más conocidas.

Ejemplo 1.76 (Familia Frank). *Esta familia fue mencionada por primera vez en Frank (1979) en un entorno no estadístico. La cópula queda definida como sigue,*

$$C_\theta(u, v) := -\frac{1}{\theta} \log \left[1 + \frac{(e^{-u\theta} - 1)(e^{-v\theta} - 1)}{e^{-\theta} - 1} \right] \quad (1.95)$$

y esta familia está definida para todo $\theta \in \mathbb{R}$; $\theta \neq 0$. En este caso se tiene que la función generadora está dada por,

$$\varphi_\theta(t) = -\log \left(\frac{e^{-t\theta} - 1}{e^{-\theta} - 1} \right) \quad (1.96)$$

la cual es un generador aditivo estricto. Para esta familia se cumple que $C_{-\infty} = W$, $C_0 = \Pi$ y $C_\infty = M$. Además la familia Frank cumple la siguiente propiedad,

$$v - C_\theta(1 - u, v) = C_{-\theta}(u, v) = u - C_\theta(u, 1 - v) \quad (1.97)$$

Es decir, si (X, Y) es un vector aleatorio cualquiera con cópula asociada perteneciente a la familia Frank denotada como C_θ , y además α y β son funciones continuas estrictamente decrecientes en $\text{Ran } X$ y $\text{Ran } Y$, respectivamente. Gracias a la ecuación 1.97,

$$\begin{aligned} C_{\alpha(X)Y} &= C_{-\theta} \\ C_{X\beta(Y)} &= C_{-\theta} \\ C_{\alpha(X)\beta(Y)} &= C_\theta \end{aligned}$$

es decir, la familia Frank es cerrada bajo transformaciones continuas monótonas.

En la figura 1.6 se ven ejemplos de simulaciones de uniformes cero-uno que tienen como distribución conjunta la extensión a \mathbb{R}^2 de una cópula de Frank.

Otra parametrización de la cópula Frank se puede encontrar en Salvadori et. al. (2007), aunque dicha parametrización no cumple la ecuación establecida en 1.97, por lo cual se prefiere la parametrización establecida en la ecuación 1.96.

Ejemplo 1.77 (Familia Gumbel-Hougaard). *Esta familia fue discutida por primera vez en Gumbel (1960) y es la cópula subyacente de los modelos establecidos en*

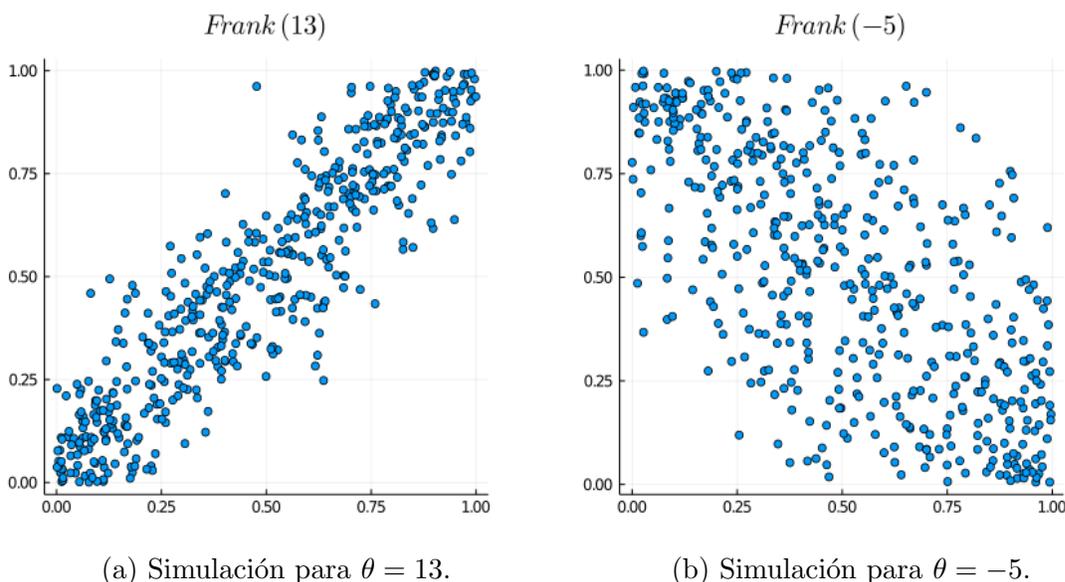


Figura 1.6: Simulaciones para cópulas de la familia Frank.

Hougaard (1986), de ahí viene su nombre. Además de que en Genest et. al. (1989) se establece la relación de esta familia de cópulas con la teoría de distribuciones de valores extremos. La familia de Gumbel-Hougaard queda definida como sigue,

$$C_{\theta}(u, v) := \exp \left(- \left[(-\log u)^{\theta} + (-\log v)^{\theta} \right]^{1/\theta} \right) \quad (1.98)$$

y esta familia está definida para todo $\theta \in [1, \infty[$. En este caso tenemos que la función generadora esta dada por,

$$\varphi_{\theta}(t) = (-\log t)^{\theta} \quad (1.99)$$

el cual es un generador aditivo estricto. Para esta familia se cumple que $C_1 = \Pi$ y $C_{\infty} = M$.

Aunque la familia Gumbel-Hougaard no cubre todo el espectro deseado (al no pasar por W), se puede realizar una extensión de esta familia de la siguiente forma, esto se hace en base a la ecuación 1.98,

$$C_{\theta^*}(u, v) := u - C_{-\theta^*}(u, 1 - v); \quad \theta^* \in [-\infty, -1[\quad (1.100)$$

la ecuación 1.100 permite extender el modelo y se cumple que $C_{-1} = \Pi$ y $C_{-\infty} = W$. En la figura 1.7 se muestran simulaciones de uniformes cero-uno que tienen como distribución conjunta la extensión a \mathbb{R}^2 de una cópula de Gumbel-Hougaard.

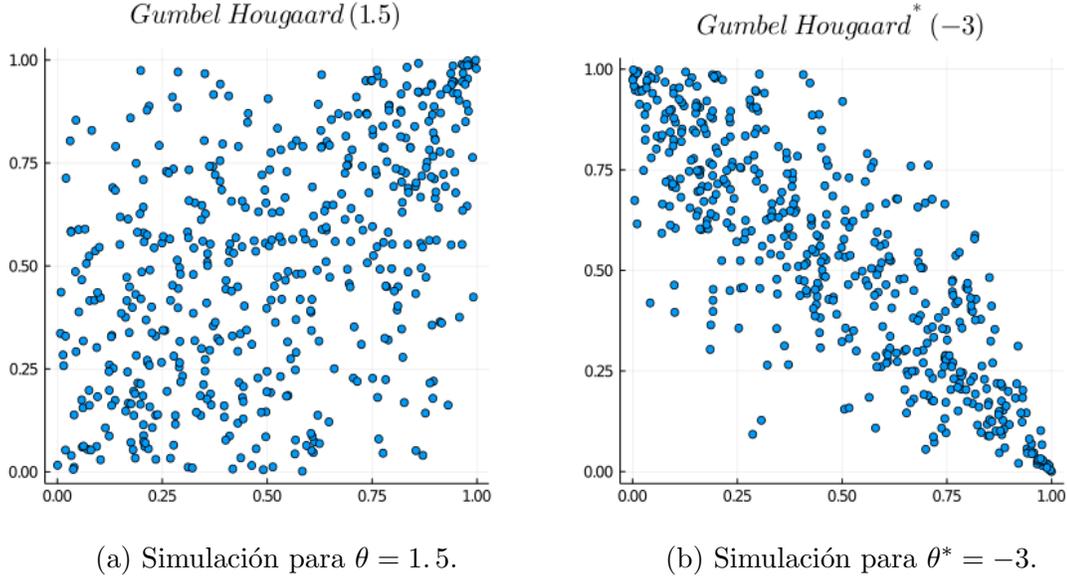


Figura 1.7: Simulaciones para cópulas de la familia Gumbel Hougaard (extendida).

Ejemplo 1.78 (Familia Clayton). *Cópulas de esta forma fueron sugeridas por Clayton (1978) y Oakes (1982) para modelar la asociación de tablas de vida bivariadas, además de ser mencionada en Genest et. al. (1986). La familia de Gumbel-Hougaard queda definida como sigue,*

$$C_{\theta}(u, v) = \left[\max \left(u^{-\theta} + v^{-\theta} - 1, 0 \right) \right]^{-1/\theta} \quad (1.101)$$

ya esta familia está definida para todo $\theta \in [-1, \infty)$ tal que $\theta \neq 0$. En este caso tenemos que la función generadora esta dada por,

$$\varphi_{\theta}(t) = \frac{1}{\theta} (t^{-\theta} - 1) \quad (1.102)$$

el cual es un generador aditivo estricto para $\theta \geq 0$. Para esta familia se cumple que $C_{-1} = W$, $C_0 = \Pi$ y $C_{\infty} = M$.

En la figura 1.8 se muestran simulaciones de uniformes cero-uno que tienen como distribución conjunta la extensión a \mathbb{R}^2 de una cópula de Clayton.

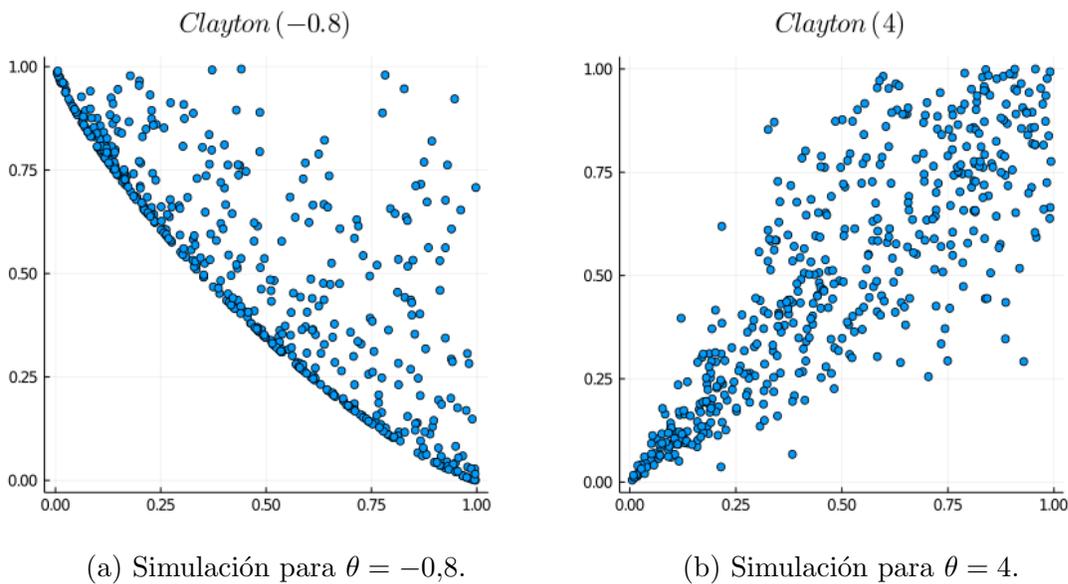


Figura 1.8: Simulaciones para cópulas de la familia Clayton.

Para mayores ejemplos en relación a cópulas arquimedianas se puede checar Nelsen *op. cit.* y el apéndice de Salvadori *et. al. op. cit.* Por último, existen extensiones de cópulas bivariadas arquimedianas al caso multivariado, y estos desarrollos requieren de extender las propiedades de las funciones φ (a un concepto denominado funciones *completamente monotónicas*¹⁶), aunque cabe mencionar el siguiente comentario de Nelsen *op. cit.*:

Although it is fairly simple to generate Archimedean n -copulas, they do have their limitations. First of all, in general all the k -margins of an Archimedean n -copula are identical. Secondly, the fact that there are usually only one or two parameters limits the nature of the dependence structure in these families.

por lo que esta extensión aunque matemáticamente es posible, no permite tener mucha libertad a la hora de cuantificar dependencia entre múltiples variables aleatorias.

¹⁶Para más información al respecto puede verse Nelsen *op. cit.*

1.6. Medidas de concordancia y dependencia

Para cualquier v.a. (X, Y) compuesto por variables aleatorias continuas X y Y (para asegurar la unicidad de C_{XY}), se puede expresar la función de distribución conjunta de probabilidades F_{XY} mediante las siguientes tres funciones:

- (i) La función F_X que representa la incertidumbre univariada de la v.a. X .
- (ii) La función F_Y que representa la incertidumbre univariada de la v.a. Y .
- (iii) La función C_{XY} que representa la relación existente entre X y Y .

creando así la siguiente relación,

$$F_{XY} \longleftrightarrow F_X, F_Y, C_{XY} \tag{1.103}$$

Gracias al resultado en el teorema 1.64 así como al ejemplo 1.67, se puede notar como F_X y F_Y “distorsionan” la dependencia contenida en el modelo (X, Y) (el ejemplo 1.67 contenía distintos modelos bivariados (X, Y) pero con una la misma cópula C_{XY}), por lo que es de interés crear medidas de dependencia que no dependan de F_X y F_Y en ningún sentido.

Esta sección se enfoca en establecer dos tipos de medidas de asociación numéricas entre pares de variables aleatorias continuas; una de ellas es un caso particular de medida de *concordancia* y la otra es un ejemplo de una medida de *dependencia*, dichas medidas se llaman *rho de Spearman*¹⁷ y *sigma de Schweizer*¹⁸, respectivamente.

1.6.1. Ejemplo de medida de concordancia: Rho de Spearman

Definición 1.79. *Una medida numérica de asociación κ entre dos variables aleatorias continuas X y Y cuya cópula es C es una medida de concordancia si se satisfacen las siguientes propiedades (donde se escribe $\kappa_{X,Y}$ o κ_C cuando sea conveniente):*

- (i) κ está definida para cada par de variables aleatorias continuas (X, Y) .

¹⁷Véase Kruskal (1958).

¹⁸Establecida por Schweizer *et. al.* (1981).

- (ii) $-1 \leq \kappa_{X,Y} \leq 1$, $\kappa_{X,X} = 1$, y $\kappa_{X,-X} = -1$.
- (iii) $\kappa_{X,Y} = \kappa_{Y,X}$.
- (iv) Si X y Y son independientes, entonces $\kappa_{X,Y} = \kappa_{\Pi} = 0$.
- (v) $\kappa_{-X,Y} = \kappa_{X,-Y} = -\kappa_{X,Y}$.
- (vi) Si $\{(X_n, Y_n)\}_{n \in \mathbb{N}}$ es una secuencia de variables aleatorias continuas con cópula C_n , respectivamente, y si $\{C_n\}_{n \in \mathbb{N}}$ converge puntualmente a C , entonces,

$$\lim_{n \rightarrow \infty} \kappa_{C_n} = \kappa_C \quad (1.104)$$

Una vez hecho esto, la rho de Spearman se define como sigue.

Definición 1.80. Sean X, Y dos variables aleatorias continuas cuya cópula es C . Entonces la rho de Spearman queda denotada como ρ_{XY} (cuando sea conveniente se usa la notación ρ_C) y está definida como sigue,

$$\begin{aligned} \rho_{XY} := \rho_C &= 12 \int_0^1 \int_0^1 (C(u, v) - uv) \, dudv \\ &= 12 \int_0^1 \int_0^1 C(u, v) \, dudv - 3 \end{aligned} \quad (1.105)$$

Si se toma un vector aleatorio compuesto por v.a. continuas (X, Y) donde F_X y F_Y son sus respectivas funciones de distribución de probabilidades univariadas y C es la cópula asociada del modelo (X, Y) . Entonces se denota como $Q_{X,Y}$ al siguiente valor

$$Q_{X,Y} := \mathbb{P}((X_1 - X_2)(Y_1 - Y_2) > 0) - \mathbb{P}((X_1 - X_2)(Y_1 - Y_2) < 0) \quad (1.106)$$

donde (X_1, Y_1) y (X_2, Y_2) son vectores aleatorios independientes con funciones de distribución conjuntas de probabilidades H_1 y H_2 , respectivamente,

$$\begin{aligned} H_1(x, y) &= C(F_X(x), F_Y(y)) \\ H_2(x, y) &= F_X(x)F_Y(y) \end{aligned}$$

es decir, el término $Q_{X,Y}$ representa la *diferencia de probabilidades de concordancia y discordancia* entre los vectores (X_1, Y_1) (que tiene dependencia de por medio) y (X_2, Y_2) (el cual no tiene dependencia de por medio).

Teorema 1.81. *Sea (X, Y) un vector aleatorio compuesto por v.a. continuas, donde F_X y F_Y son sus respectivas funciones de distribución de probabilidades univariadas y C es la cópula asociada del modelo (X, Y) . Entonces $Q_{X,Y}$ tal como se definió en 1.106 es equivalente a lo siguiente,*

$$Q_{X,Y} \equiv 4 \int \int_{\mathbf{I}^2} C(u, v) \, dudv - 1 = 4 \int \int_{\mathbf{I}^2} uv \, dC(u, v) - 1 \quad (1.107)$$

Demostración. Para la prueba, véase Nelsen *op. cit.* □

Usando el teorema 1.81 se cumple lo siguiente para la rho de Spearman para un vector aleatorio continuo (X, Y) ,

$$\begin{aligned} \rho_{XY} &= 3Q_{X,Y} = 12 \int_0^1 \int_0^1 C(u, v) \, dudv - 3 \\ &= 12 \int_0^1 \int_0^1 uv \, dC(u, v) - 3 \end{aligned}$$

Sea (X, Y) un vector aleatorio continuo donde F_X y F_Y son las respectivas funciones de distribución de probabilidades univariadas y C es la cópula asociada del modelo (X, Y) , entonces de acuerdo a Nelsen *op. cit.* se cumple lo siguiente en relación al término ρ_{XY} ,

$$\begin{aligned} \rho_{XY} &= 12 \int_0^1 \int_0^1 uv \, dC(u, v) - 3 = 12 \mathbb{E}(UV) - 3 \\ &= \frac{\mathbb{E}(UV) - 1/4}{1/12} = \frac{\mathbb{E}(UV) - \mathbb{E}(U)\mathbb{E}(V)}{\sqrt{\mathbb{V}(U)}\sqrt{\mathbb{V}(V)}} \\ &= \text{Corr}(U, V) \end{aligned}$$

para el vector aleatorio (U, V) tal que $F_{UV} = C_{\mathbb{R}^2}$. Por el teorema 1.64 se cumple que $U = F_X(X)$ y $V = F_Y(Y)$, por lo que se cumple que,

$$\rho_{XY} = \text{Corr}(F_X(X), F_Y(Y)) \quad (1.108)$$

es decir, la rho de Spearman representa el coeficiente de correlación de Pearson de las v.a. X y Y una vez no se cuenta con la “distorsión” provocada por sus respectivas funciones de distribución de probabilidades univariadas F_X y F_Y .

Gracias al resultado establecido en la ecuación 1.108, es fácil concluir como el coeficiente de Pearson no es invariante al realizar transformaciones continuas estrictamente crecientes a las v.a. en cuestión (recuerdesé que la cópula si es invariante para dichas transformaciones pero las distribuciones univariadas no lo són en general). El siguiente ejemplo fue extraído de Embrechts *et. al.* (2001) y refleja lo antes explicado,

Ejemplo 1.82. Sean X y Y dos variables aleatorias continuas con distribución exponencial estandar, y sea C_θ la familia de cópulas Farlie-Gumbel-Morgenstern (FGM) definida para $\theta \in [-1, 1]$ como sigue,

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

con la anterior familia de cópulas se define los siguientes modelos bivariados,

$$H_\theta(x, y) = C_\theta(1 - e^{-x}, 1 - e^{-y})\mathbf{1}_{\{x, y > 0\}}; \quad \theta \in [-1, 1]$$

En este caso,

$$\mathbb{E}(XY) = \int_0^\infty \int_0^\infty xy \, dH_\theta(x, y) = \dots = 1 + \frac{\theta}{4}$$

y como $\mathbb{E}(X) = \mathbb{V}(X) = 1 = \mathbb{E}(Y) = \mathbb{V}(Y)$, se tiene que el coeficiente de correlación de Pearson queda como sigue,

$$\text{Corr}(X, Y) = \frac{\theta}{4}$$

Y ahora obtengamos $\text{Corr}(F_X(X), F_Y(Y)) = \rho_{XY}$,

$$\begin{aligned} \text{Corr}(F_X(X), F_Y(Y)) &= \rho_{XY} = 12 \int \int_{\mathbb{I}^2} C_\theta(u, v) \, dudv - 3 \\ &= 12 \int \int_{\mathbb{I}^2} [uv + \theta uv(1 - u)(1 - v)] \, dudv - 3 \\ &= 12 \left(\frac{1}{4} + \frac{\theta}{36} \right) = \frac{\theta}{3} \end{aligned}$$

Por lo tanto, en este ejemplo $\text{Corr}(X, Y) = \rho_{XY}$ solamente en el caso de que $\theta = 0$ (es decir cuando X y Y son independientes).

El teorema 1.83 establece la relación entre la rho de Spearman con la definición 1.79, mientras que la observación 1.84 establece propiedades de la rho de Spearman,

Teorema 1.83. *La rho de Spearman es una medida de concordancia.*

Demostración. La demostración puede verse en Nelsen *op. cit.* □

Proposición 1.84. *Sea (X, Y) un vector aleatorio continuo con cópula C , entonces se cumple lo siguiente,*

(i) *Si $C \geq \Pi$ entonces,*

$$\rho_{XY} \geq 0$$

(ii) *Si $C \leq \Pi$ entonces,*

$$\rho_{XY} \leq 0$$

(iii) *Para las cópulas W , Π y M se cumple lo siguiente,*

$$\rho_M = 1$$

$$\rho_\Pi = 0$$

$$\rho_W = -1$$

Demostración. Recordando la definición de ρ_{XY} ,

$$\rho_{XY} = 12 \int_0^1 \int_0^1 (C(u, v) - uv) \, dudv$$

Para el inciso (i), como $C \geq \Pi$ entonces $C(u, v) \geq \Pi(u, v)$ y por lo tanto $\rho_{XY} \geq 0$. Y el caso (ii) es equivalente. Para la cópula M ,

$$\begin{aligned} \rho_M &= 12 \int_0^1 \int_0^1 M(u, v) \, dudv - 3 = 12 \int_0^1 \int_0^1 \min\{u, v\} \, dudv - 3 \\ &= 12 \left(\int_0^1 \int_0^u v \, dvdu + \int_0^1 \int_0^v u \, dudv \right) - 3 \\ &= 12 \left(\frac{1}{6} + \frac{1}{6} \right) - 3 = \frac{12}{3} - 3 = 4 - 3 = 1 \end{aligned}$$

para W ,

$$\begin{aligned}\rho_W &= 12 \int_0^1 \int_0^1 W(u, v) \, dudv - 3 = 12 \int_0^1 \int_0^1 \max\{0, u + v - 1\} \, dudv - 3 \\ &= 12 \int_0^1 \int_{1-v}^1 (u + v - 1) \, dudv - 3 = \frac{12}{6} - 3 = 2 - 3 = -1\end{aligned}$$

y el caso de Π es claro. □

1.6.2. Ejemplo de medida de dependencia: Sigma de Schweizer

Sea C una cópula cualquiera, se define la cópula C^* como sigue,

$$C^*(u, v) := \frac{1}{2}(u - C(u, 1 - v) + C(u, v)) \quad (1.109)$$

la sigma de Spearman para la cópula C^* queda como sigue,

$$\begin{aligned}\rho_{C^*} &= 12 \int \int_{\mathbb{I}^2} (C^*(u, v) - uv) \, dudv \\ &= \frac{12}{2} \int \int_{\mathbb{I}^2} (C(u, v) - uv) \, dudv + \frac{12}{2} \int \int_{\mathbb{I}^2} (u - C(u, 1 - v) - uv) \, dudv \\ &= \frac{12}{2} \int \int_{\mathbb{I}^2} (C(u, v) - uv) \, dudv - \frac{12}{2} \int \int_{\mathbb{I}^2} (C(u, v) - uv) \, dudv \\ &= \frac{1}{2}\rho_C - \frac{1}{2}\rho_C = 0\end{aligned}$$

y esto se cumple para cualquier cópula C , es decir, la rho de Spearman se anula para cópulas que pueden llegar a tener dependencia. Si se desea crear una medida numérica ψ_C tal que $\psi_C = 0$ si y sólo si $C = \Pi$, se debe usar un concepto distinto al de medida de concordancia. Este concepto se define a continuación,

Definición 1.85. *Una medida numérica de asociación δ entre dos variables aleatorias continuas X y Y cuya cópula es C es una medida de dependencia si se satisfacen las siguientes propiedades (donde se escribe $\delta_{X,Y}$ o δ_C cuando sea conveniente):*

(i) δ está definida para cada par de variables aleatorias continuas (X, Y) .

- (ii) $\delta_{X,Y} = \delta_{Y,X}$.
- (iii) $0 \leq \delta_{X,Y} \leq 1$.
- (iv) $\delta_{X,Y} = 0$ si y sólo si X y Y son independientes.
- (v) $\delta_{X,Y} = 1$ si y sólo si X y Y son casi seguramente funciones monótonas una de otra.
- (vi) Si α y β son funciones casi seguramente monótonas crecientes sobre $\text{Ran } X$ y $\text{Ran } Y$, respectivamente, entonces $\delta_{\alpha(X),\beta(Y)} = \delta_{X,Y}$.
- (vii) Si $\{(X_n, Y_n)\}_{n \in \mathbb{N}}$ es una secuencia de variables aleatorias continuas con cópula C_n , respectivamente, y si $\{C_n\}_{n \in \mathbb{N}}$ converge puntualmente a C , entonces,

$$\lim_{n \rightarrow \infty} \delta_{C_n} = \delta_C \quad (1.110)$$

Se puede crear una medida de dependencia basándose en la rho de Spearman definida en la anterior subsección, esto se hace reemplazando la diferencia entre las cópulas C y Π por una diferencia en valor absoluto (esta diferencia se anula si y sólo si $C = \Pi$). Esta medida de dependencia fue establecida en Schweizer *et. al.* (1981) y es conocida como la sigma de Schweizer y Wolff (por simpleza se le denomina como sigma de Schweizer),

Definición 1.86. Sean X, Y dos variables aleatorias continuas cuya cópula es C . Entonces la sigma de Schweizer queda denotada como σ_{XY} (cuando sea conveniente se denota como σ_C) y está definida como sigue,

$$\sigma_{XY} = \sigma_C = 12 \int \int_{\mathbb{I}^2} |C(u, v) - uv| \, du dv \quad (1.111)$$

Proposición 1.87. La sigma de Schweizer cumple las siguientes propiedades para una cópula cualquiera C ,

$$|\rho_C| \leq \sigma_C$$

además,

- (i) Si $C \geq \Pi$, entonces

$$\sigma_C = \rho_C$$

(ii) Si $C \leq \Pi$, entonces

$$\sigma_C = -\rho_C$$

Demostración. Recordando las definiciones de ρ_C y σ_C ,

$$\begin{aligned}\rho_C &= 12 \int \int_{\mathbb{I}^2} (C(u, v) - uv) \, dudv \\ \sigma_C &= 12 \int \int_{\mathbb{I}^2} |C(u, v) - uv| \, dudv\end{aligned}$$

y por propiedades de la integral,

$$|\rho_C| = 12 \left| \int \int_{\mathbb{I}^2} (C(u, v) - uv) \, dudv \right| \leq 12 \int \int_{\mathbb{I}^2} |C(u, v) - uv| \, dudv = \sigma_C$$

para las propiedades (i) y (ii) es solamente necesario ver que,

(i) Si $C \geq \Pi$ entonces $|C(u, v) - uv| = C(u, v) - uv$

(ii) Si $C \leq \Pi$ entonces $|C(u, v) - uv| = uv - C(u, v)$

□

Gracias a la proposición 1.87 se puede ver que mientras el valor $|\rho_C|$ se aleje más del término σ_C , entonces la dependencia que modela la cópula C tiene un alto grado de “combinación” de información por encima de Π y por debajo de Π . Si se da el caso de que $\rho_C = 0$ mientras que $\sigma_C \neq 0$, entonces se puede decir que la concordancia por encima de Π y por debajo de Π se “cancelan mutuamente”.

Para terminar esta subsección, el teorema 1.88 establece la relación entre la sigma de Schweizer y la definición 1.85.

Teorema 1.88. Sean X y Y dos variables aleatorias continuas con cópula C . Entonces la cantidad σ_C cumple las propiedades de una medida de dependencia.

Demostración. Para la demostración véase Nelsen *op. cit.*

□

1.7. Tail Dependence

Un concepto de dependencia basados en cópulas que resultan de mucho interés es el concepto de *tail dependence* (que en español significa, dependencia de colas), dicho concepto permite relacionar probabilidades “límites” en los extremos superiores e inferiores de un v.a. (X, Y) cualquiera. El concepto de *tail dependence* muestra la relación que tiene un v.a. (X, Y) en casos “extremos”, formalmente esto queda definido como sigue,

Definición 1.89. Sean X y Y dos variables aleatorias continuas con funciones de distribución de probabilidades univariadas F y G , respectivamente. El parámetro de dependencia en la cola superior λ_U es el límite (si existe) de la probabilidad condicional siguiente,

$$\lambda_U = \lim_{t \rightarrow 1^-} \mathbb{P}[Y > G^{(-1)}(t) \mid X > F^{(-1)}(t)] \quad (1.112)$$

De forma similar, el parámetro de dependencia en la cola inferior λ_L es el límite (si existe) de la probabilidad condicional siguiente,

$$\lambda_L = \lim_{t \rightarrow 0^+} \mathbb{P}[Y \leq G^{(-1)}(t) \mid X \leq F^{(-1)}(t)] \quad (1.113)$$

El siguiente resultado establece la relación entre los valores λ_U y λ_L con respecto a la cópula del modelo bivariado (X, Y) .

Teorema 1.90. Sean X, Y, F, G, λ_U y λ_L establecidas como en la definición 1.89 y sea C la cópula del modelo bivariado (X, Y) con sección diagonal δ_C . Si los límites definidos en las ecuaciones 1.112 y 1.113 existen, entonces,

$$\begin{aligned} \lambda_U &= 2 - \delta'_C(1^-) \\ \lambda_L &= \delta'_C(0^+) \end{aligned}$$

Demostración. Se define $U = F(X)$ y $V = G(Y)$ y por el teorema 1.64 se cumple que

$C_{UV} = C_{\mathbb{R}^2}$. Para el caso de λ_U ,

$$\begin{aligned}
\lambda_U &= \lim_{t \rightarrow 1^-} \mathbb{P}[Y > G^{(-1)}(t) \mid X > F^{(-1)}(t)] = \lim_{t \rightarrow 1^-} \mathbb{P}[V > t \mid U > t] \\
&= \lim_{t \rightarrow 1^-} \frac{\mathbb{P}(U > t, V > t)}{\mathbb{P}(U > t)} = \lim_{t \rightarrow 1^-} \frac{1 - \mathbb{P}(\{U \leq t\} \cup \{V \leq t\})}{1 - t} \\
&= \lim_{t \rightarrow 1^-} \frac{1 - \mathbb{P}(U \leq t) - \mathbb{P}(V \leq t) + \mathbb{P}(U \leq t, V \leq t)}{1 - t} \\
&= \lim_{t \rightarrow 1^-} \frac{1 - 2t + C(t, t)}{1 - t} = 2 - \lim_{t \rightarrow 1^-} \frac{C(t, t) - 1}{t - 1} \\
&= 2 - \lim_{h \rightarrow 0^-} \frac{C(1 + h, 1 + h) - C(1, 1)}{h} = 2 - \lim_{h \rightarrow 0^-} \frac{\delta_C(1 + h) - \delta_C(1)}{h} \\
&= 2 - \delta'_C(1^-)
\end{aligned}$$

Y para el caso de λ_U se tiene que,

$$\begin{aligned}
\lambda_L &= \lim_{t \rightarrow 0^+} \mathbb{P}[Y \leq G^{(-1)}(t) \mid X \leq F^{(-1)}(t)] = \lim_{t \rightarrow 0^+} \mathbb{P}[V \leq t \mid U \leq t] \\
&= \lim_{t \rightarrow 0^+} \frac{\mathbb{P}(U \leq t, V \leq t)}{\mathbb{P}(U \leq t)} = \lim_{t \rightarrow 0^+} \frac{C(t, t)}{t} = \lim_{h \rightarrow 0^+} \frac{C(0 + h, 0 + h) - C(0, 0)}{h} \\
&= \lim_{h \rightarrow 0^+} \frac{\delta_C(0 + h) - \delta_C(0)}{h} = \delta'_C(0^+)
\end{aligned}$$

□

Si $\lambda_U \in (0, 1]$, entonces se dice que C tiene dependencia de cola superior; si $\lambda_U = 0$, entonces se dice que C no tiene dependencia de cola superior; y de forma equivalente para λ_L .

En el siguiente ejemplo se muestra como se comporta la dependencia de cola superior e inferior de manera gráfica, notamos que dichas dependencias se ven como “picos” pronunciados en los gráficos de dispersion siguientes en las esquinas superior derecha e inferior izquierda, respectivamente.

Ejemplo 1.91. Sean C_θ^1 y C_α^2 las siguientes cópulas,

$$\begin{aligned}
C_\theta^1(u, v) &= uv + \theta uv(1 - u)(1 - v); & \theta &\in [1-, 1] \\
C_\alpha^2(u, v) &= (u^{-\alpha} + v^{-\alpha} - 1)^{-1/\alpha}; & \alpha &\in]0, \infty[
\end{aligned}$$

Primeramente se consiguen los parámetros de dependencia en las colas superior e inferior para las cópulas C_θ^1 y C_α^2 . Para esto si δ_θ y ψ_α son las secciones diagonales de las cópulas C_θ^1 y C_α^2 , respectivamente, entonces,

$$\begin{aligned}\delta_\theta(t) &= t^2 + \theta t^2(1-t)^2 \\ \psi_\alpha(t) &= (2t^{-\alpha} - 1)^{-1/\alpha}\end{aligned}$$

y las derivadas de las funciones δ_θ y ψ_α son las siguientes,

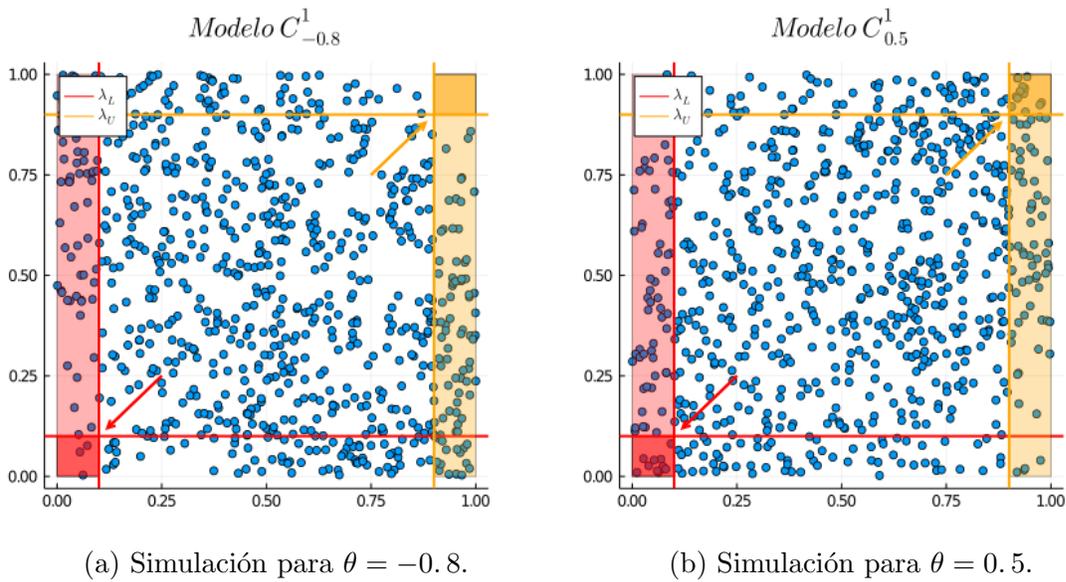
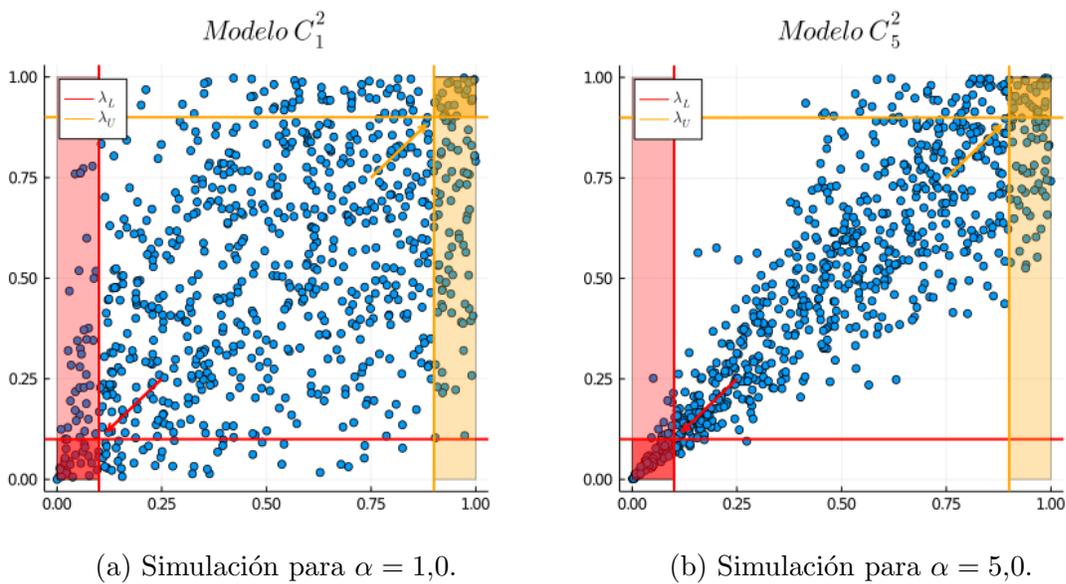
$$\begin{aligned}\delta'_\theta(t) &= 2t + 2t\theta(1-t)^2 - 2\theta t^2(1-t) \\ \psi'_\alpha(t) &= 2t^{-1-\alpha}(2t^{-\alpha} - 1)^{-1-\alpha}\end{aligned}$$

por lo que los parámetros de dependencia en las colas superior e inferior para las cópulas son los siguientes,

	C_θ^1	C_α^2
λ_U	0	0
λ_L	0	$2^{-1/\alpha}$

Cuadro 1.1: Tail dependence superior e inferior.

Ahora bien, en las figuras 1.9 y 1.10 se muestran simulaciones para las cópulas C_θ^1 y C_α^2 y se compara de forma gráfica los resultados en la tabla 1.1. Nótese como los valores λ_U y λ_L se pueden interpretar como la proporción de puntos acumulados en las esquinas superior derecha e inferior izquierda, respectivamente, en relación a las áreas pintadas en naranja y rojo, respectivamente.

Figura 1.9: Simulaciones para la cópula C_θ^1 .Figura 1.10: Simulaciones para la cópula C_α^2 .

En las figuras 1.9a y 1.9b se puede ver como se cumple que $\lambda_U = 0 = \lambda_L$ sin importar el valor de θ , se ve como la acumulación en las esquinas superior derecha e inferior

izquierda no es lo suficiente alta como para considerarse “importante”, por lo cual decimos que C_θ^1 no tiene dependencia de colas de ningún tipo.

En las figuras 1.10a y 1.10b se puede ver como se cumple que $\lambda_U = 0$ y que λ_L es una función creciente con respecto al valor de α , y se nota como la dependencia en colas se ve como una “alta” acumulación de puntos en la esquina inferior izquierda. Por lo cual se dice que C_θ^2 tiene dependencia de colas inferior y no tiene dependencia de colas superior.

1.8. Gluing Copulas

En la mayoría de las ocasiones, el analizar dependencia en un modelo bivariado mediante solo un modelo paramétrico de cópulas es una limitante muy fuerte. Una de las formas en las que se puede eliminar esta limitante es mediante la combinación de cópulas para formar un modelo mas robusto y una estrategia para realizar este cometido es mediante el proceso en inglés denominado como *gluing copula*, el cual se establece a continuación.

Definición 1.92. Sean C_1 y C_2 dos cópulas cualesquiera y sea $0 < \theta < 1$ un número cualquiera que se denomina “gluing point”. Se definen las funciones $C_1 \overset{H}{\otimes}_\theta C_2$ y $C_1 \overset{V}{\otimes}_\theta C_2$ en el cuadrado unitario como siguen,

$$\begin{aligned} \left(C_1 \overset{H}{\otimes}_\theta C_2 \right) (u, v) &= \begin{cases} \theta C_1\left(\frac{u}{\theta}, v\right); & 0 \leq u \leq \theta \\ (1 - \theta)C_2\left(\frac{u-\theta}{1-\theta}, v\right) + \theta v; & \theta \leq u \leq 1 \end{cases} \\ \left(C_1 \overset{V}{\otimes}_\theta C_2 \right) (u, v) &= \begin{cases} \theta C_1\left(u, \frac{v}{\theta}\right); & 0 \leq v \leq \theta \\ (1 - \theta)C_2\left(u, \frac{v-\theta}{1-\theta}\right) + \theta u; & \theta \leq v \leq 1 \end{cases} \end{aligned}$$

a las funciones $C_1 \overset{H}{\otimes}_\theta C_2$ y $C_1 \overset{V}{\otimes}_\theta C_2$ se les denomina *gluing-cópulas horizontales y verticales en el punto θ con componentes C_1 y C_2 , respectivamente.*

Proposición 1.93. Las funciones definidas en el teorema 1.92 son cópulas.

Demostración. Para la demostración, véase el teorema 2.1 en Siburg y Stoimenov (2008). \square

El objetivo de una gluing-cópula es el de combinar la dependencia de dos cópulas (denotadas como C_1 y C_2 en el teorema 1.92) sobre un punto $0 < \theta < 1$ de dos formas posibles:

- 1) De forma horizontal como en el caso de $C_1 \overset{H}{\otimes}_{\theta} C_2$.
- 2) De forma vertical como en el caso de $C_1 \overset{V}{\otimes}_{\theta} C_2$.

en las figuras 1.11a y 1.11b se puede ver un mapa de calor y de dispersión de una gluing cópula vertical para el caso en el que $C_1 = Frank(3)$, $C_2 = Clayton(-0.5)$ y $\theta = 0.4$, respectivamente,

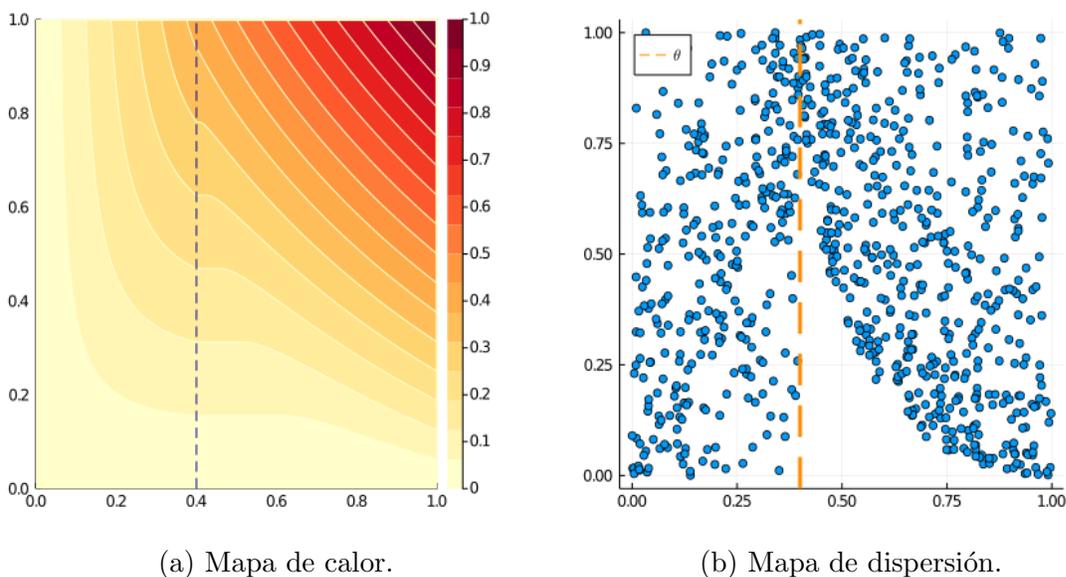


Figura 1.11: Gráficas gluing-cópula.

Se puede notar como el nodo horizontal provoca un corte vertical en las dos gráficas, así como un claro cambio de dependencia que tiene como frontera al nodo θ . En el caso de la sección diagonal de la cópula, la figura 1.12 muestra su comportamiento,

así como una propiedad esencial de la misma, TODA gluing-cópula con nodo θ cumple con la condición siguiente,

$$\delta_C(\theta) = \theta^2 = \delta_{\Pi}(\theta)$$

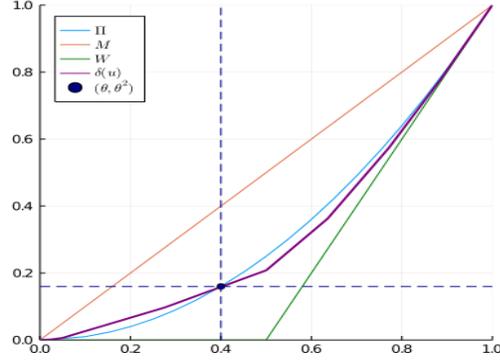


Figura 1.12: Sección Diagonal.

y esto es consecuencia inmediata de la definición 1.92. Además, la sección diagonal no puede diferenciar que tipo de corte tiene la gluing-cópula (horizontal o vertical) como consecuencia de que esta función es una reducción a una dimensión de la información contenida en la cópula completa.

Una propiedad interesante para una gluing-cópula es su funcionamiento en relación a sus rotaciones horizontales y verticales respectivas y como éstas coinciden con otras gluing-cópulas. Dichos razonamientos se establecen en el siguiente teorema,

Teorema 1.94. Sean C_1, C_2 y θ tal y como se establecieron en la definición 1.92 y sea $C = C_1 \overset{H}{\otimes}_{\theta} C_2$, entonces se cumple lo siguiente para las rotaciones horizontales y verticales de C ,

$$C^H = C_1^H \overset{H}{\otimes}_{\theta} C_2^H$$

$$C^V = C_2^V \overset{H}{\otimes}_{1-\theta} C_1^V$$

en el caso de que $C = C_1 \overset{V}{\circledast}_{\theta} C_2$,

$$C^H = C_2^H \overset{H}{\circledast}_{1-\theta} C_1^H$$

$$C^V = C_1^V \overset{H}{\circledast}_{\theta} C_2^V$$

Demostración. Para el caso de C^H , primeramente tomemos $0 \leq u \leq \theta$,

$$C^H(u, v) = u - C(u, 1 - v) = u - \theta C_1\left(\frac{u}{\theta}, 1 - v\right) = \theta \left[\frac{u}{\theta} - C_1\left(\frac{u}{\theta}, 1 - v\right) \right]$$

$$= \theta C_1^H\left(\frac{u}{\theta}, v\right)$$

y si $\theta \leq u \leq 1$,

$$C^H(u, v) = u - C(u, 1 - v) = u - (1 - \theta)C_2\left(\frac{u - \theta}{1 - \theta}, 1 - v\right) - \theta(1 - v)$$

$$= (1 - \theta) \left[\frac{u - \theta}{1 - \theta} - C_2\left(\frac{u - \theta}{1 - \theta}, 1 - v\right) \right] + \theta v = (1 - \theta)C_2\left(\frac{u - \theta}{1 - \theta}, v\right) + \theta v$$

por lo tanto $C^H = C_1^H \overset{H}{\circledast}_{\theta} C_2^H$. Para el caso de C^H , si se toma $0 \leq u \leq 1 - \theta$ entonces $\theta \leq 1 - u \leq 1$ y así,

$$C^V(u, v) = v - C(1 - u, v) = v - (1 - \theta)C_2\left(\frac{1 - u - \theta}{1 - \theta}, v\right) - \theta v$$

$$= (1 - \theta) \left[v - C_2\left(1 - \frac{u}{1 - \theta}, v\right) \right] = (1 - \theta)C_2^V\left(\frac{u}{1 - \theta}, v\right)$$

ahora bien, si se toma $1 - \theta \leq u \leq 1$ se tiene que $0 \leq 1 - u \leq \theta$ y así,

$$C^V(u, v) = v - C(1 - u, v) = v - \theta C_1\left(\frac{1 - u}{\theta}, v\right) + (1 - \theta)v - (1 - \theta)v$$

$$= (1 - \theta)v + \theta \left[v - C_1\left(1 - \frac{u - (1 - \theta)}{\theta}, v\right) \right]$$

$$= (1 - \theta)v + \theta C_1^V\left(\frac{u - (1 - \theta)}{\theta}, v\right)$$

por lo tanto $C^V = C_2^V \underset{1-\theta}{\otimes}^H C_1^V$. El caso en el que $C = C_1 \underset{\theta}{\otimes}^V C_2$ es equivalente. □

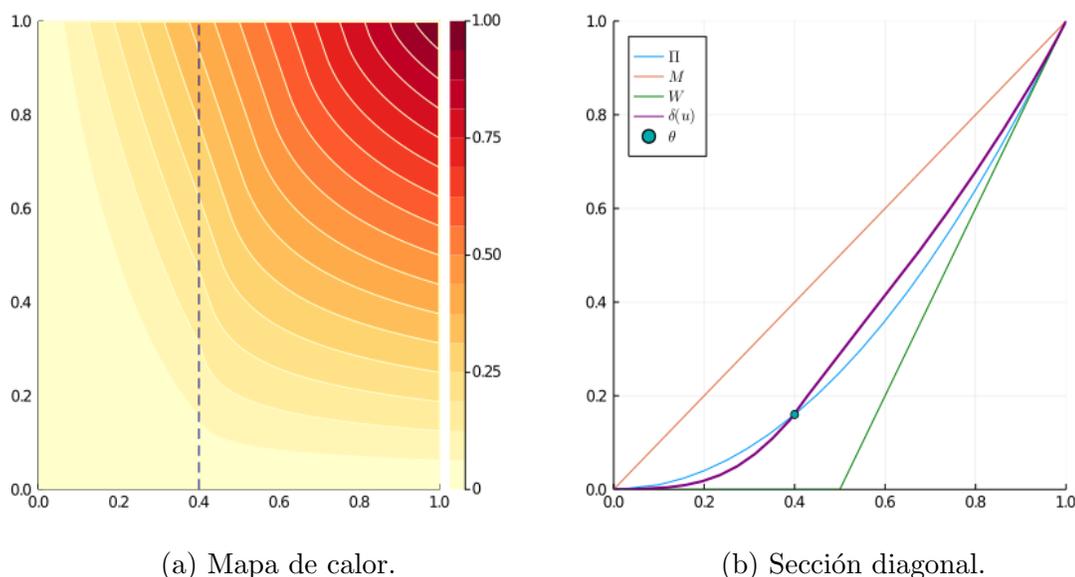


Figura 1.13: Gráficas gluing-cópula rotación horizontal.

En la figuras 1.13 y 1.14 se puede ver como el mapa de calor y la sección diagonal de una gluing-cópula se ven afectadas al realizar rotaciones. La gluing cópula $C = C_1 \underset{\theta}{\otimes}^H C_2$ es la misma que se utilizó en la figura 1.11 (con sus respectivas rotaciones horizontales y verticales),

Ahora bien, es de interés conocer un algoritmo que permita generar valores aleatorios para un vector aleatorio (X, Y) donde la cópula subyacente C_{XY} sea una gluing-cópula cualquiera, dicho algoritmo se menciona a continuación,

Algoritmo 1.94.1. *Sea (X, Y) un vector aleatorio cualquiera con función de distribución conjunta de probabilidades F_{XY} . F_X y F_Y son las funciones de distribución de probabilidades marginales y sea C_{XY} la cópula asociada al vector aleatorio (X, Y) y se tiene que C_{XY} es una gluing-cópula en el punto θ con componentes C_1 y C_2 , es decir, $C_{XY} = C_1 \underset{\theta}{\otimes}^H C_2$. Se desea obtener una muestra aleatoria de tamaño n del vector aleatorio (X, Y) . Esto se puede hacer con los siguientes pasos:*

1. Se simula una m.a. de tamaño n para una v.a. $U \sim U(0, 1)$. Dicha muestra se denota como $\{u_1, u_2, \dots, u_n\}$.
2. Se simula un valor aleatorio para cada una de las siguiente n v.a. $\{V_1, V_2, \dots, V_n\}$, cuyas respectivas funciones de distribución de probabilidades son las siguientes.

2.1 Si $0 \leq u_i \leq \theta$,

$$F_{V_i|U_i}(v_i | u_i) = \frac{\partial}{\partial u} C_{XY}(u, v) \Big|_{u=u_i/\theta} \quad i \in \{1, 2, \dots, n\}$$

2.2 Si $\theta \leq u_i \leq 1$,

$$F_{V_i|U_i}(v_i | u_i) = \frac{\partial}{\partial u} C_{XY}(u, v) \Big|_{u=(u_i-\theta)/(1-\theta)} \quad i \in \{1, 2, \dots, n\}$$

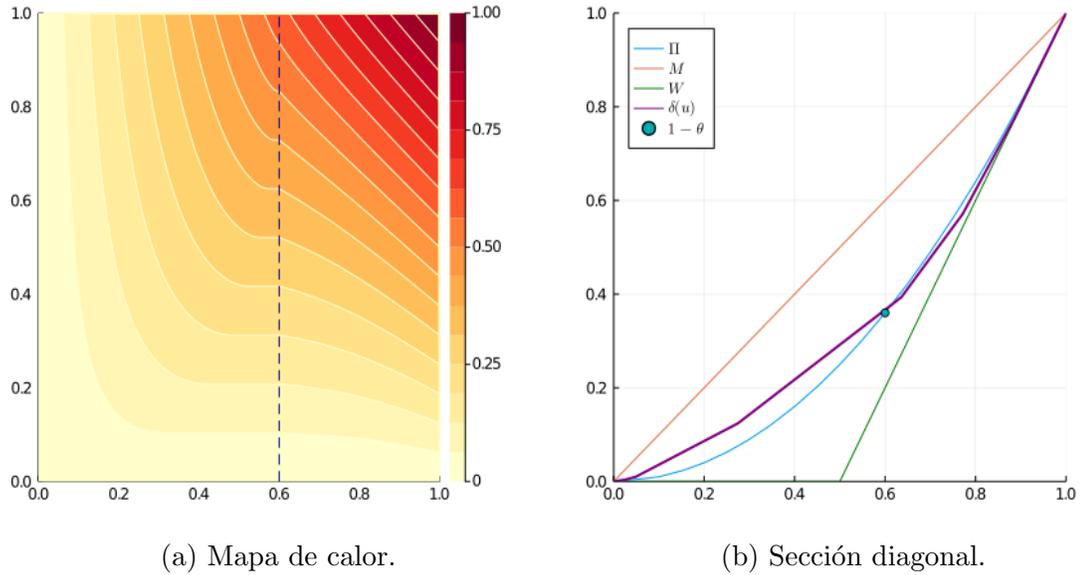
Dichos valores se denotan como $\{v_1, v_2, \dots, v_n\}$.

3. Se realizan las siguientes transformaciones para la m.a. $\{(u_1, v_1), (u_2, v_2), \dots, (u_n, v_n)\}$,

$$\begin{aligned} x_i &= F_X^{(-1)}(u_i) \\ y_i &= F_Y^{(-1)}(v_i) \end{aligned}$$

obteniendo así el conjunto $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ que representa una m.a. del vector (X, Y) .

el caso en el que $C_{XY} = C_1 \overset{V}{\otimes}_{\theta} C_2$ es equivalente, modificando el papel de U y V entre sí.



(a) Mapa de calor.

(b) Sección diagonal.

Figura 1.14: Gráficas giuing-cópula rotación horizontal.

Se termina esta sección con resultados relacionados con la ρ de Spearman y la σ Schweizer para el caso de una giuing-cópula cualquiera,

Teorema 1.95. Sean C_1 y C_2 dos cópulas cualquiera tales que ρ_1, σ_1 y ρ_2, σ_2 son sus respectivas ρ de Spearman y la σ Schweizer, además sea $0 < \theta < 1$ un número cualquiera. Para una giuing-cópula $C_{XY} = C_1 \underset{\theta}{\overset{*}{\otimes}} C_2$ (donde $*$ es H o V) se cumple lo siguiente,

$$\begin{aligned}\rho_C &= \theta^2 \rho_1 + (1 - \theta)^2 \rho_2 \\ \sigma_C &= \theta^2 \sigma_1 + (1 - \theta)^2 \sigma_2\end{aligned}$$

Demostración. Para la demostración se utiliza $C_{XY} = C_1 \underset{\theta}{\overset{H}{\otimes}} C_2$, el caso $C_{XY} = C_1 \underset{\theta}{\overset{V}{\otimes}} C_2$ es equivalente, se divide la integral de interés con respecto al punto en el que la giuing-cópula cambia de dependencia (es decir, el punto θ),

$$\begin{aligned}\int_0^1 \int_0^1 (C(u, v) - uv) \, dudv &= \int_0^1 \int_0^\theta (C(u, v) - uv) \, dudv \\ &\quad + \int_0^1 \int_\theta^1 (C(u, v) - uv) \, dudv\end{aligned}$$

para la primera integral,

$$\begin{aligned} \int_0^1 \int_0^\theta (C(u, v) - uv) \, dudv &= \int_0^1 \int_0^\theta \left(\theta C_1\left(\frac{u}{\theta}, v\right) - uv \right) \, dudv \\ &= \theta^2 \int_0^1 \int_0^1 (C_1(u, v) - uv) \, dudv = \theta^2 \rho_1 \end{aligned}$$

y esto se realiza con el cambio de variable $z = \frac{u}{\theta}$, $w = v$ y el determinante $\left| \frac{\partial(u, v)}{\partial(z, w)} \right| = \theta$.

Para el caso de la segunda integral,

$$\begin{aligned} \int_0^1 \int_\theta^1 (C(u, v) - uv) \, dudv &= \int_0^1 \int_\theta^1 \left(\theta v + (1 - \theta) C_2\left(\frac{u - \theta}{1 - \theta}, v\right) - uv \right) \, dudv \\ &= (1 - \theta)^2 \int_0^1 \int_0^1 (C_2(u, v) - uv) \, dudv = (1 - \theta)^2 \rho_2 \end{aligned}$$

el caso para σ_C es equivalente.

□

Capítulo 2

Instrumentos Financieros Derivados

En este capítulo se describen los resultados básicos de la teoría de instrumentos financieros derivados, iniciando con el modelo a tiempo discreto y concluyendo con el modelo continuo que es un caso “límite” del modelo discreto. Además, se establece el principio de arbitraje, así como las consecuencias que tiene este en la valuación de activos financieros. Para este capítulo, se utiliza principalmente Cherubini *et. al.* (2004) y Shreve (2004a, 2004b).

Este capítulo comienza con acontecimientos históricos de gran interés para el área de los Instrumentos Financieros Derivados, esto se debe a que los avances teóricos en la valuación de Instrumentos Derivados (denominados como derivados o derivados financieros) son consecuencia de su amplia utilización en los mercados financieros (si no se utilizaron en el mundo real, saber como valuarlos carecería de un sentido práctico).

2.1. Antecedentes históricos de la teoría de Derivados

Los derivados financieros básicos para que se van a utilizar en este trabajo son denominados **opciones**. Una opción financiera garantiza al dueño de la misma el derecho, mas no la obligación, de comprar o vender un activo financiero denominado

subyacente, dicho activo puede ser acciones de empresas, commodities¹, etc. Las opciones cuentan con una fecha de madurez preestablecida (tiempo máximo de vida del derivado); si la opción es de compra se denomina opción tipo *Call* y si es de venta se denomina de tipo *Put*; por último, si la opción puede ser ejercida solamente cuando cumpla la madurez, se denomina una opción *Europea*, mientras que si puede ser ejercida en cualquier momento entre el presente y su fecha de madurez, se dice que la opción es de tipo *Americana*.

Otro tipo de derivados financieros de gran utilización e importancia son los conocidos como contratos *forward*, estos permiten la compra o venta de un producto a un tiempo futuro preestablecido y su precio queda pactado en el tiempo presente (cuando se pacta el contrato), para este caso existen contratos forward sobre acciones de empresas, commodities, etc. Para evitar el riesgo de contrapartida o contraparte², se crean los derivados financieros denominados como *futuros*, los cuales cuentan con una contraparte central que se encarga de la liquidación periódica de pérdidas y ganancias del instrumento en cuestión.

Tal y como menciona Poitras (2009), existe evidencia de la implementación de contratos de compras futuras en la civilización griega, esta implementación viene descrita en el libro *Política* de Aristóteles, donde la especulación de dichos contratos fue llevada a cabo por el filósofo Tales de Mileto al alquilar de forma anticipada prensas de aceitunas en Mileto y Quíos; y al comprarlas a un valor anticipado y con una muy buena cosecha de aceitunas ese año, Tales acumuló riqueza por una inversión que se ganó en el futuro, por eso se dice que este contrato es equivalente a un instrumento derivado actual de tipo forward implementado de forma rústica.

Como se puede notar, los primeros derivados que existieron en el mundo estaban conectados de una a otra forma a commodities, esto es lógico ya que la incertidumbre en la agricultura, ganadería y el valor de materias primas era uno de los factores de mayor importancia económica en civilizaciones preindustriales (y actualmente, todavía esto sucede hasta cierto grado) por lo cual, se puede notar como los principales centros financieros de la antigüedad estaban fundamentados en commodities varios, un claro ejemplo de esto son los centros financieros de Bélgica entre los siglos XII y XV. En Bélgica, los centros financieros más importantes, en esa época, eran los siguientes:

¹Bienes básicos destinados para uso comercial como materias primas.

²Riesgo financiero en el cual no se liquidan las pérdidas o ganancias de un activo financiero por alguna de las partes aunque ya hayan sido pactadas.

- 1) La ciudad de Brujas³ (con gran apogeo en el siglo XIV).
- 2) El mercado de concentración de Amberes⁴ (con gran apogeo en el siglo XV).

como medidas interesantes establecidas en estos centros financieros se tienen las siguientes:

- 1) Transmisión de contratos para entrega a plazo (compra-venta de contratos forward ya estipulados).
- 2) Se eliminó la necesidad de hacer entrega de bienes subyacentes a los acuerdos pactados, sino que bastaba con la compensación de pérdidas y ganancias entre las partes (se elimina la necesidad del subyacente físico en el pacto del derivado).

esto en conjunto permite ver como los derivados financieros empezaron a adquirir una especie de “independencia”, en términos de negociación, con respecto a los subyacentes de los contratos; esto dio el inicio a mercados financieros enfocados a dichos instrumentos.

En años posteriores, se puede notar un incremento de interés en la negociación de estos instrumentos, ejemplo de esto se puede ver en los contratos de opciones establecidos por la *Dutch East India Company* y la *Dutch West India Company* entre los años 1602 y 1621 (Poitras (2009)); y el crecimiento de Inglaterra como el principal mercado de commodities en los 1650, esto gracias al dominio inglés sobre el transporte marítimo, lo cual permitió a Inglaterra adquirir y actualizar las prácticas comerciales y financieras adquiridas por comerciantes belgas y holandeses entre los siglos XV y XVII, respectivamente.

Aunque también se tiene que tomar en consideración aspectos que han afectado la expansión de este mercado en Europa, un caso muy conocido de esto fue el mercado de futuros creado por la burbuja de precios en tulipanes, conocido como “Tulipmanía” que tuvo lugar en Holanda durante la década de 1630, este acontecimiento provocó un gran efecto negativo a la economía neerlandesa; otro caso de interés es el llamado “crack de 1720”, provocado por la especulación existente en las acciones de la *South Sea Company*, empresa marítima inglesa.

Los anteriores eventos fueron provocados por que los instrumentos derivados de esa época no contaban con los característicos de los derivados actuales; propiedades como las siguientes:

³Actualmente capital de la provincia de Flandes Occidental, Bélgica.

⁴Actualmente capital de la homónima provincia de Amberes, Bélgica.

- 1) Negociación con contraparte central.
- 2) Liquidación periódica de pérdidas y ganancias.
- 3) Contratos estandarizados.

En los siguientes años, se puede ver indicios de mercados de derivados que empezaron a considerar los anteriores puntos de forma más rigurosa, ejemplo de esto se puede ver en los contratos sobre el arroz establecidos en el *Dojima Rice Exchange* hacia el año 1730, información al respecto se puede ver en Kummer *et. al.* (2012) y en la página oficial del *Dojima Rice Exchange*⁵.

El avance de mayor interés a nivel global en el área de derivados financieros se puede considerar en la creación y expansión del centro financiero de Chicago, que actualmente es la meca de la industria de los derivados financieros, esto inició con la creación del *Chicago Board of Trade* (CBOT) donde a partir de 1851 se comenzaron a pactar contratos forward con maíz como producto subyacente, hasta el punto que empezaron a existir actores económicos que no estaban involucrados en la producción o consumo del activo subyacente sino que solamente fijaban el precio de estos derivados financieros, creando nuevamente una especie de “independencia” entre los mercados de derivados y de commodities, tal y como se vio en Bélgica anteriormente.

Para el año 1865, el CBOT formalizó la negociación de contratos con entrega a plazo que se operaban en este mercado. De este modo nacen los futuros en los Estados Unidos. Mientras que para el siglo XX empiezan a surgir mercados de derivados financieros alrededor de Estados Unidos, ejemplo de esto es el *Commercial Exchange* de New York o el *Chicago Board of Options Exchange* en Chicago.

El impulso en los derivados financieros que tiene mayor interés para este escrito fue en los 1970, ya que para estas fechas se empezaron a publicar algunos de los trabajos más relevantes en el área de las finanzas cuantitativas. En este área se hace referencia a los trabajos de Black *et. al.* (1973) y al trabajo de Merton (1973), ambos enfocados en la valuación de opciones financieras europeas suponiendo log-rendimientos normales con varianza y media no estocásticas.

Con toda esta información, se puede ver como el mercado de derivados financieros es algo que se ha desarrollado a través del tiempo. Este proceso inició en el mercado de commodities, donde creció en penetración de mercado, y terminó

⁵Véase <https://www.jpx.co.jp/dojima/en/index.html>

convirtiéndose en mercados financieros especializados en la compra/venta de estos instrumentos, tal es el caso del CBOT, entre otros.

Las siguientes secciones de este capítulo se encargan de plantear resultados enfocados en la valuación de derivados a tiempo discreto y continuo basándose en el principio de arbitraje.

2.2. Procesos estocásticos y esperanza condicional

Para modelar la incertidumbre que se va a tener respecto a un activo financiero cualquiera, se necesita un concepto que permita cuantificar incertidumbre a través del tiempo, y esto se logra gracias al siguiente concepto.

Definición 2.1. *Sea T un conjunto no vacío cualquiera y sea $\mathbf{X} = (X, t \in T) \equiv (X_t)_{t \in T}$ una familia de variables aleatorias definidas en un mismo espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$. Entonces se dice que \mathbf{X} es un proceso estocástico indexado al tiempo T .*

Se crean las siguientes dos notaciones particulares,

(i) *Si $T = \mathbb{R}^+$; entonces de forma equivalente,*

$$\mathbf{X} \equiv (X_t)_{t \geq 0}$$

(ii) *Si $T = \{1, \dots, n\}$ para $n \in \mathbb{N}$ cualquiera; entonces de forma equivalente,*

$$\mathbf{X} \equiv (X_k)_{k \in I_n}$$

El siguiente ejemplo es un ejemplo clásico de un proceso estocástico a tiempo discreto.

Ejemplo 2.2. *Sea $(X_k)_{k \in \mathbb{N}}$ un proceso estocástico compuesto por variables aleatorias independientes e idénticamente distribuidas tales que,*

$$\begin{aligned} \mathbb{P}(X_j = 1) &= p \\ \mathbb{P}(X_j = -1) &= 1 - p \end{aligned}$$

se define un nuevo proceso estocástico $\mathbf{M} = (M_k)_{k \in \mathbb{N}_0}$ de la siguiente forma,

$$M_0 = 0; \quad M_k := \sum_{j=1}^k X_j \quad k \in \mathbb{N}$$

al proceso \mathbf{M} se le denota como una caminata aleatoria con probabilidad de subida p . Si $p = 0.5$ entonces se dice que \mathbf{M} es una caminata aleatoria simétrica. En la figura 2.1 se ve posibles caminos que tomaría el proceso estocástico \mathbf{M} en el caso de que fuera una caminata aleatoria simétrica para $j \in \{1, \dots, 5\}$.

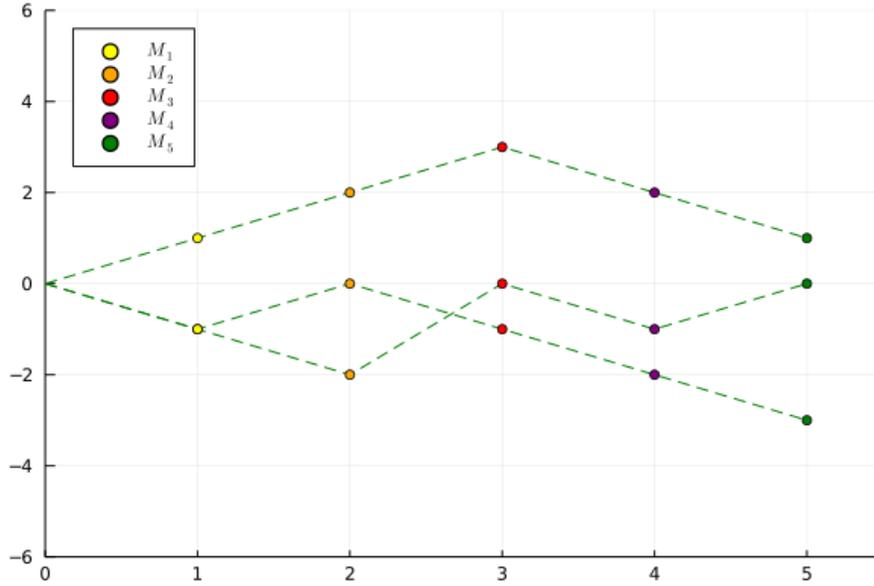


Figura 2.1: Caminata Aleatoria Simétrica

Una pregunta de interés para este ejemplo es obtener la distribución de M_k para cualquier $k \in \mathbb{N}$, esto se hace gracias a la siguiente relación,

$$X_j = 2Y_j - 1$$

donde $Y_j \sim \text{Ber}(p)$. Por lo que sustituyendo esto en la definición de M_k ,

$$M_k = \sum_{j=1}^k X_j = \sum_{j=1}^k (2Y_j - 1) = 2 \sum_{j=1}^k Y_j - k \quad (2.1)$$

y como consecuencia de que $\{X_1, \dots, X_k\}$ es una muestra aleatoria, entonces $\{Y_1, \dots, Y_k\}$ también lo es, y se cumple que,

$$Y^k := \sum_{j=1}^k Y_j \sim \text{Binom}(k, p)$$

donde la función de masa de probabilidades de la v.a. Y^k es la siguiente,

$$p_{Y^k}(y) = \mathbb{P}(Y^k = y) = \binom{k}{y} p^y (1-p)^{k-y}$$

por lo tanto la función de masa de probabilidades de la v.a. M_k queda como sigue,

$$\begin{aligned} p_{M_k}(m) &= \mathbb{P}(M_k = m) = \mathbb{P}(2Y^k - k = m) = \mathbb{P}\left(Y^k = \frac{m+k}{2}\right) \\ &= \binom{k}{(m+k)/2} p^{(m+k)/2} (1-p)^{(k-m)/2} \end{aligned}$$

y esto es válido para todo $m \in \text{Ran } M_k = \{-k, 2-k, 4-k, \dots, k-4, k-2, k\}$.

Una vez introducido el concepto de un proceso estocástico. Si $(X_t)_{t \in T}$ es el proceso estocástico en cuestión, existen dos preguntas de gran interés en relación a este proceso estocástico:

- 1) Para $t \in T$ cualquiera, ¿cómo se distribuye X_t ?
- 2) Si $t_1, \dots, t_m \in T$ son elementos cualesquiera, ¿cómo se distribuye el vector aleatorio $(X_{t_1}, \dots, X_{t_m})$?

Por lo desarrollado en el ejemplo 2.2, las distribuciones univariadas del proceso estocástico $\mathbf{M} = (M_k)_{k \in \mathbb{N}_0}$ están completamente identificadas, en el ejemplo siguiente se analizan distribuciones bivariadas para la caminata aleatoria.

Ejemplo 2.3. (Continuación ejemplo 2.2) Sea $\mathbf{M} = (M_k)_{k \in \mathbb{N}_0}$ una caminata aleatoria con probabilidad de subida p cualquiera. Se analiza la distribución conjunta de probabilidades para el vector aleatorio (M_i, M_j) donde $i < j$. Nótese como M_j se puede poner en términos de M_i de la siguiente forma,

$$M_j = \sum_{k=1}^j X_k = \sum_{k=1}^i X_k + \sum_{k=i+1}^j X_k = M_i + \sum_{k=i+1}^j X_k$$

Sea $M^* := \sum_{k=i+1}^j X_k$; entonces por como se definió el proceso estocástico $(X_k)_{k \in \mathbb{N}}$ en el ejemplo 2.2, se tiene que M_i y M^* son variables aleatorias independientes y además,

$$M^* = M_{j-i}$$

por lo que,

$$M_j | \{M_i = m\} = M^* + m = M_{j-i} + m$$

Con esta información, se concluye lo siguiente en relación a la función de masa de probabilidades conjunta para el v.a. (M_i, M_j) para cualquier $m_1 \in \text{Ran } M_i$, $m_2 \in \text{Ran } M_j$ (denotada en este caso particular como p_{ij}),

$$\begin{aligned} p_{ij}(m_1, m_2) &= \mathbb{P}(M_i = m_1, M_j = m_2) = \frac{\mathbb{P}(M_i = m_1, M_j = m_2)\mathbb{P}(M_i = m_1)}{\mathbb{P}(M_i = m_1)} \\ &= \mathbb{P}(M_i = m_1)\mathbb{P}(M_j = m_2 | M_i = m_1) \\ &= \mathbb{P}(M_i = m_1)\mathbb{P}(M_{j-i} + m_1 = m_2) \\ &= \mathbb{P}(M_i = m_1)\mathbb{P}(M_{j-i} = m_2 - m_1) \end{aligned}$$

y esta probabilidad es distinta a cero en el caso de que $m_2 - m_1 \in \text{Ran } M_{j-i}$.

Un proceso estocástico $(X_t)_{t \in T}$ cualquiera permite cuantificar la incertidumbre de un fenómeno o experimento aleatorio a través del tiempo, el tiempo es descrito mediante el conjunto T y la incertidumbre a tiempo $t \in T$ es medida mediante la v.a. X_t . Para $s, t \in T$ dos elementos cualesquiera tales que $s < t$, la v.a. $X_t | \{X_s = x_s\}$ permite conocer la incertidumbre que se tiene respecto al fenómeno o experimento aleatorio a tiempo t , sabiendo o suponiendo que acontecimientos han ocurrido hasta el tiempo s , esto cobra mucho interés en aplicaciones financieras, ya que se desea cuantificar la incertidumbre de rendimientos futuros conociendo el proceso estocástico hasta un tiempo pasado s .

La siguiente definición extiende el concepto de variable aleatoria condicional y lo relaciona con el concepto de esperanza.

Definición 2.4. Sea X una v.a. definida en un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ tal que X tiene esperanza finita y sea $A \in \mathcal{F}$. Entonces se define $\mathbb{E}(X | A)$ como sigue,

$$\mathbf{E}(X | A) := \begin{cases} \frac{\mathbb{E}(\mathbf{1}_A X)}{\mathbb{P}(A)} & \text{si } \mathbb{P}(A) > 0 \\ 0 & \text{en otro caso} \end{cases} \quad (2.2)$$

se dice que $\mathbb{E}(X | A)$ es la esperanza condicional de X dado el evento A .

Sea I un conjunto numerable y sea $(B_i)_{i \in I}$ conjuntos disjuntos a pares tales que $\biguplus_{i \in I} B_i = \Omega$. Si X una v.a. con esperanza finita y $\mathcal{F} := \sigma(B_i, i \in I)$, entonces se define el siguiente mapeo $\mathbb{E}(X | \mathcal{F}) : \Omega \mapsto \mathbb{R}$ como:

$$\mathbb{E}(X | \mathcal{F})(\omega) = \mathbb{E}(X | B_i) \iff B_i \ni \omega \quad (2.3)$$

Para la generalización de la segunda parte de la definición 2.4, nótese que la única incertidumbre que existe con el término $\mathbb{E}(X | B)$ está dada por la esperanza $\mathbb{E}(X \mathbf{1}_B)$, por lo que se puede ampliar este concepto con la siguiente definición y con el resultado establecido en el teorema 2.6.

Definición 2.5. Sea X una v.a. definida sobre un espacio de probabilidad $(\Omega, \mathcal{A}, \mathbb{P})$ donde X tiene una esperanza finita y sea $\mathcal{F} \subset \mathcal{A}$ una sigma-álgebra cualquiera. A la variable aleatoria Y se le llama “esperanza condicional de X dada \mathcal{F} ”, simbólicamente $\mathbb{E}(X | \mathcal{F}) := Y$, si:

(i) Y es \mathcal{F} -medible.

(ii) Para cualquier $A \in \mathcal{F}$, se tiene que $\mathbb{E}(X \mathbf{1}_A) = \mathbb{E}(Y \mathbf{1}_A)$.

Para $B \in \mathcal{A}$, se tiene que $\mathbb{P}(B | \mathcal{F}) := \mathbb{E}(\mathbf{1}_B | \mathcal{F})$ es llamada la “probabilidad condicional de B dada la σ -álgebra \mathcal{F} ”.

Teorema 2.6. Tomando X y \mathcal{F} descritos en la definición 2.5, se cumple que $\mathbb{E}(X | \mathcal{F})$ existe y es única (hasta una igualdad casi segura).

Demostración. La demostración se puede consultar en Klenke *op. cit.* en el teorema 8.12. □

Definición 2.7. Sea $\mathbb{F} = (\mathcal{F}_t, t \in T)$ una familia de σ -álgebras con $\mathcal{F}_t \subset \mathcal{F}$ para todo $t \in T$ donde \mathcal{F} representa un espacio de eventos cualquiera. Se dice que \mathbb{F} es una filtración si $\mathcal{F}_s \subseteq \mathcal{F}_t$ para todo $s, t \in T$ con $s \leq t$.

Definición 2.8. Un proceso estocástico $\mathbf{X} = (X_t, t \in T)$ es llamado adaptable a la filtración \mathbb{F} si X_t es \mathcal{F}_t -medible para toda $t \in T$. Si $\mathcal{F}_t = \sigma(X_s, s \leq t)$ para toda $t \in T$, entonces denotamos a $\mathbb{F} = \sigma(\mathbf{X})$ y decimos que \mathbb{F} es la filtración generada por el proceso estocástico \mathbf{X} .

El concepto de filtración establecido en la definición 2.7 tiene el objetivo de plasmar como la información disponible que se tiene a un tiempo fijo $s \in T$ va en aumento una se avanza al futuro ($\mathcal{F}_s \subset \mathcal{F}_t, t \geq s$), mientras que el concepto de adaptabilidad del proceso \mathbf{X} con respecto a la filtración \mathbb{F} , establece que para cualquier tiempo $t \in T$, se cuenta con información **suficiente** para cuantificar toda la incertidumbre que se tiene con respecto a la variable aleatoria X_t (X_t es \mathcal{F}_t -medible).

Ejemplo 2.9. (Continuación ejemplo 2.2) En este ejemplo, se utiliza un subconjunto finito de tamaño k de la caminata aleatoria \mathbf{M} simétrica definida en el ejercicio 2.2. Para cualquier $k \in \mathbb{N}$ se define el proceso aleatorio \mathbf{M}^k como sigue:

$$\mathbf{M}^k := \{M_0, M_1, \dots, M_k\}$$

para el proceso \mathbf{M}^k se desea obtener lo siguiente:

- 1) Establecer la filtración $\mathbb{F} := \sigma(\mathbf{M}^k)$ para la cual el proceso estocástico \mathbf{M}^k es adaptable.
- 2) Tomando $i \leq j \leq k$, encontrar una v.a. que cumpla ser la esperanza condicional siguiente,

$$\mathbb{E}(M_j | \mathcal{F}_i)$$

Para resolver los incisos 1) y 2), se necesita primero definir claramente el espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ para el cual todas las funciones $M \in \mathbf{M}^k$ son variables aleatorias. Si $M \in \mathbf{M}^k$ entonces M representa una transformación de interés para algún tipo de v.a. binomial Z , por lo que se considera a Ω (el espacio muestral) como en el caso de una v.a. binomial, es decir, Ω queda definido como sigue:

$$\Omega = \{(\omega_1, \dots, \omega_k) : \omega_j \in \{e, f\}\}$$

donde e representa un estado de “éxito” (acompañado de la probabilidad $\frac{1}{2}$ de éxito) mientras que f representa un estado de “fallo” (acompañado de la probabilidad $\frac{1}{2}$ de fallo).

Ahora bien, el conjunto Ω es finito, por lo que se puede considerar $\mathcal{F} = 2^\Omega$ (el conjunto potencia de Ω) como el espacio de eventos. Para la medida de probabilidad \mathbb{P} , se crea la siguiente regla de correspondencia para cada una de las singuletas que tiene Ω (elementos de Ω):

$$\mathbb{P}(\{(\omega_1, \dots, \omega_k)\}) = \prod_{j=1}^k \left(\frac{1}{2}\right)^{\mathbf{1}_{\{\omega_j=e\}}} \left(\frac{1}{2}\right)^{\mathbf{1}_{\{\omega_j=f\}}}$$

A continuación se pone en práctica la definición de \mathbb{P} . Tómesese $M_2 \in \mathbf{M}^k$ y calcúlese la probabilidad siguiente,

$$\mathbb{P}(M_2 = 0)$$

recuérdese que gracias a la ecuación 2.1,

$$M_2 = 2Y^2 - 2$$

por lo que,

$$\mathbb{P}(M_2 = 0) = \mathbb{P}(Y^2 = 1)$$

el evento $\{Y^2 = 1\}$ es equivalente al siguiente conjunto,

$$\{Y^2 = 1\} = \{(e, f, \omega_2, \dots, \omega_k) : \omega_j \in \{e, f\}\} \cup \{(f, e, \omega_2, \dots, \omega_k) : \omega_j \in \{e, f\}\}$$

por lo cual,

$$\begin{aligned} \mathbb{P}(Y^2 = 1) &= \mathbb{P}\left(\{(e, f, \omega_2, \dots, \omega_k) : \omega_j \in \{e, f\}\}\right) \\ &\quad + \mathbb{P}\left(\{(f, e, \omega_2, \dots, \omega_k) : \omega_j \in \{e, f\}\}\right) \\ &= \frac{1}{4} + \frac{1}{4} \\ &= \frac{1}{2} = p_{Y^2}(1) \end{aligned}$$

cumplíndose la función de masa antes descrita para el modelo binomial.

Una vez establecido el espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, se puede trabajar con los conjuntos $\sigma(M)$ para cualquier $M \in \mathbf{M}^k$. Para $M_j \in \mathbf{M}^k$ para $1 \leq j \leq k$ cualquiera, los únicos conjuntos $A \in \mathcal{B}(\mathbb{R})$ para los cuales se cumple que $M_j^{-1}(A) \neq \emptyset$ son los conjuntos que incluyen cualquiera de las siguientes singuletas,

$$s_i^j := \{i\} \quad i \in \text{Ran } M_j$$

entonces,

$$M_j^{-1}(s_i^j) = \left\{ \omega \in \Omega : \sum_{k=1}^j \mathbf{1}_{\{w_k=e\}} = \frac{i+j}{2} \right\}$$

por lo que $\sigma(M_j)$ queda definido simplemente como sigue,

$$\sigma(M_j) = \sigma(\{M_j^{-1}(s_i^j) : i \in \text{Ran } M_j\}) \subseteq \mathcal{F}$$

con esto y sabiendo que todo $\mathcal{F}_j \in \mathbb{F}$ esta definido simplemente como sigue,

$$\mathcal{F}_j = \sigma(\sigma(M_k), k \leq j)$$

contamos con una clara definición de \mathbb{F} y con esto queda cumplido el inciso 1).

Para el inciso 2); si $Z = M_i$, se comprueba que $Z = \mathbb{E}(M_j | \mathcal{F}_i)$. Primeramente, Z es \mathcal{F}_i -medible; ahora bien, para $A \in \mathcal{F}_i$ cualquiera se cumple la siguiente propiedad,

$$M_j \mathbf{1}_A = M_i \mathbf{1}_A + M^* \mathbf{1}_A$$

tal y como fue definido M^* en el ejemplo 2.3 y además $\mathbb{E}(M^*) = 0$ (ya que \mathbf{M} es una caminata aleatoria simétrica). Como consecuencia de que $\mathbf{1}_A$ y M^* son variables independientes, se cumple lo siguiente,

$$\mathbb{E}(M_j \mathbf{1}_A) = \mathbb{E}(M_i \mathbf{1}_A)$$

y eso es para cualquier $A \in \mathcal{F}_i$. Cumpliéndose así la definición 2.5, por lo tanto,

$$M_i = \mathbb{E}(M_j | \mathcal{F}_i) \quad (2.4)$$

Teorema 2.10. (Propiedades de la esperanza condicional) Sea $(\Omega, \mathcal{A}, \mathbb{P})$ un espacio de probabilidad y sea X una v.a. perteneciente a dicho espacio. Sean $\mathcal{G} \subset \mathcal{F} \subset \mathcal{A}$ σ -álgebras y sea Y una v.a. con esperanza finita también perteneciente a dicho espacio. Entonces:

- (i) (Linealidad) $\mathbb{E}(\lambda X + Y | \mathcal{F}) = \lambda \mathbb{E}(X | \mathcal{F}) + \mathbb{E}(Y | \mathcal{F})$.
- (ii) (Monotocidad) Si $X \geq Y$ casi seguramente, entonces, $\mathbb{E}(X | \mathcal{F}) \geq \mathbb{E}(Y | \mathcal{F})$.
- (iii) Si $\mathbb{E}(|XY|) < \infty$ y Y es \mathcal{F} -medible, entonces,

$$\mathbb{E}(XY | \mathcal{F}) = Y \mathbb{E}(X | \mathcal{F}) \quad \mathbb{E}(Y | \mathcal{F}) = \mathbb{E}(Y | Y) = Y$$

- (iv) (Esperanza iterada) $\mathbb{E}(\mathbb{E}(X | \mathcal{F}) | \mathcal{G}) = \mathbb{E}(\mathbb{E}(X | \mathcal{G}) | \mathcal{F}) = \mathbb{E}(X | \mathcal{G})$.
- (v) (Desigualdad del triángulo) $\mathbb{E}(|X| | \mathcal{F}) \geq |\mathbb{E}(X | \mathcal{F})|$.
- (vi) (Independencia⁶) Si $\sigma(X)$ y \mathcal{F} son independientes, entonces $\mathbb{E}(X | \mathcal{F}) = \mathbb{E}(X)$.
- (vii) Si $\mathbb{P}(A) \in \{0, 1\}$ para cualquier $A \in \mathcal{F}$, entonces $\mathbb{E}(X | \mathcal{F}) = \mathbb{E}(X)$.

⁶Para conocer el concepto de independencia entre σ -álgebras puede verse Klenke *op. cit.* definición 2.14.

(viii) (Convergencia dominada) Si $Y \geq 0$ y $(X_n)_{n \in \mathbb{N}}$ es una sucesión de variables aleatorias con $|X_n| \leq Y$ para cualquier $n \in \mathbb{N}$ tales que,

$$X_n \xrightarrow{n \rightarrow \infty} X$$

entonces,

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_n | \mathcal{F}) = \mathbb{E}(X | \mathcal{F})$$

casi seguramente.

Una vez se cuenta con un proceso estocástico $(X_t)_{t \in T}$ cualquiera, es de gran interés conocer el concepto de *martingala* y las consecuencias que tiene esta propiedad en un proceso estocástico. En términos coloquiales, el concepto de martingala tiene su origen en estrategias de apuestas, información de esto se puede encontrar en el diccionario francés Académie Française⁷; aunque el término martingala también tiene significados no relacionados con las apuestas, ejemplo de esto es el arnés utilizado en los caballos.

Como consecuencia de esto, tiene sentido que el término martingala haya sido utilizado en el área de la probabilidad en relación a procesos estocásticos que pueden modelar apuestas *justas* o *equilibradas* de cierta forma, tal y como menciona Mansuy (2005),

“A financial market is viable (i.e., does not offer arbitrage opportunities) if and only if there exists a probability measure under which the realized prices are martingales.” ... For probabilists, martingales are first of all integrable processes, satisfying a particular conditional expectation property. Aside from their role in finance ... they have applications to various stochastic and analytic problems and represent, with Markov processes, one of the most important types of processes depending on the past. The notion seems to arise quite directly from the idea of strategy in a game of chance.

La definición probabilística de martingala es la siguiente,

Definición 2.11. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, $I \in \mathbb{R}$, y sea \mathbb{F} una filtración. Sea $\mathbf{X} = (X_t)_{t \in I}$ un procesos estocástico adaptable a \mathbb{F} tal que $\mathbb{E}(|X_t|) < \infty$ para toda $t \in I$. X es llamado (con respecto a \mathbb{F}) una:

⁷Para más información, véase <https://www.academie-francaise.fr/> y búsquese el término “martingale”.

- (i) (Martingala) Si $\mathbb{E}(X_t | \mathcal{F}_s) = X_s$ para toda $s, t \in I$ con $t \geq s$.
- (ii) (Submartingala) Si $\mathbb{E}(X_t | \mathcal{F}_s) \geq X_s$ para toda $s, t \in I$ con $t \geq s$.
- (iii) (Supermartingala) Si $\mathbb{E}(X_t | \mathcal{F}_s) \leq X_s$ para toda $s, t \in I$ con $t \geq s$.

Con la anterior definición y gracias al resultado obtenido en el ejemplo 2.9, se comprueba que la caminata aleatoria simétrica es una martingala, entonces de forma práctica se considera un “juego justo”. Además, tiene sentido que esto solo se cumpla en el caso que la caminata aleatoria sea simétrica (cuando $p = \frac{1}{2}$), ya que dicho caso es cuando el modelo está completamente equilibrado.

Se termina esta sección con un resultado en relación al concepto de martingala.

Teorema 2.12. Sea $\mathbf{X} = (X_t)_{t \in T}$ una martingala con respecto a la filtración \mathbb{F} , entonces se cumple lo siguiente,

- (i) Si existe $t^* := \min\{T\}$ tal que X_{t^*} es constante, entonces para todo $t \in T$ el mapeo $t \rightarrow \mathbb{E}(X_t)$ es constante y,

$$\mathbb{E}(X_t) = X_{t^*}$$

- (ii) Si $\mathcal{G} \subset \mathbb{F}$ es una σ -álgebra de subconjuntos de Ω y Y es una v.a. que es \mathcal{G} -medible, entonces se cumple para toda $t \in T$ tal que $\mathcal{G} \subset \mathcal{F}_t$,

$$\mathbb{E}(Y X_t | \mathcal{G}) = Y \mathbb{E}(X_t | \mathcal{G})$$

Demostración. Para la propiedad (i) se tiene que si $Y := \mathbb{E}(X_t | \mathcal{F}_s)$ para $s \leq t$ entonces por definición de esperanza condicional,

$$\mathbb{E}(X_t \mathbf{1}_A) = \mathbb{E}(Y \mathbf{1}_A)$$

para cualquier $A \in \mathcal{F}_s$ si se toma $A = \Omega$ entonces $X_t \mathbf{1}_\Omega = X_t$ y $Y \mathbf{1}_\Omega = Y$, entonces

$$\mathbb{E}\left(\mathbb{E}(X_t | \mathcal{F}_s)\right) = \mathbb{E}(Y) = \mathbb{E}(Y \mathbf{1}_\Omega) = \mathbb{E}(X_t \mathbf{1}_\Omega) = \mathbb{E}(X_t)$$

y como \mathbf{X} es una martingala, entonces $\mathbb{E}(X_t | \mathcal{F}_s) = X_s$ para cualquier $s < t$ y si $s = t^*$ entonces para cualquier $t \in T$,

$$\mathbb{E}(X_t) = \mathbb{E}\left(\mathbb{E}(X_t | \mathcal{F}_{t^*})\right) = \mathbb{E}(X_{t^*}) = X_{t^*}$$

Para la propiedad (ii) se puede ver la demostración en Klenke *op. cit.*

□

2.3. El Modelo Binomial

El *modelo binomial de valuación de activos* es un modelo de valuación a tiempo discreto que permite crear una herramienta poderosa a la hora de comprender el *Principio de Arbitraje* y sus consecuencias en el ámbito de la probabilidad. Primeramente, se analiza el modelo binomial a un periodo, y una vez comprendidos todos los conceptos que se pueden rescatar del modelo, se generaliza a un modelo binomial multiperiodo.

2.3.1. El modelo binomial a un periodo

Para iniciar esta subsección, se establece el principio de arbitraje,

Principio 2.1 (Principio de Arbitraje). *Un arbitraje es cualquier tipo de inversión en los mercados financieros que tiene las siguientes características:*

- (i) *El capital inicial de la inversión es cero.*
- (ii) *La probabilidad de perder dinero con la inversión es cero.*
- (iii) *Existe una probabilidad estrictamente positiva de generar dinero con la inversión en un plazo futuro.*

el principio de arbitraje establece que en la economía, **no** pueden existir arbitrajes.

Una vez establecido el principio de arbitraje, es de interés ver como este principio se ve reflejado en un modelo binomial a un periodo.

Definición 2.13. *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad donde el espacio muestral Ω solo cuenta con dos estados posibles, es decir, $\Omega = \{H, T\}$; dicho espacio de probabilidad representa el estado de la economía un periodo en el futuro. Sea \mathbf{S} un activo cualquiera existente en el mercado financiero y sea S_0 el valor actual de dicho activo. Se define a S_1 como la siguiente variable aleatoria,*

$$S_1 \equiv \text{Valor del activo } \mathbf{S} \text{ un periodo en el futuro.} \quad (2.5)$$

Como consecuencia de la definición 2.13, S_1 es una variable aleatoria discreta con dos puntos de masa. Entonces $\text{Ran } S_1 = \{S_1(H), S_1(T)\}$, tales que,

$$\begin{aligned} \mathbb{P}(S_1 = S_1(H)) &= p \\ \mathbb{P}(S_1 = S_1(T)) &= 1 - p \end{aligned}$$

para que el modelo sea interesante (ambos eventos sean posibles), se considera $0 < p < 1$ y además se denota u y d como las siguientes cantidades,

$$u := \frac{S_1(H)}{S_0}; \quad d := \frac{S_1(T)}{S_0}$$

y para evitar confusión, establecemos que $S_1(H)$ es el escenario “próspero” para el activo \mathbf{S} , por lo que establecemos sin pérdida de generalidad que $S_1(T) \neq S_1(H)$ (si son iguales, el modelo no es interesante).

También se introduce una tasa de interés r . Dicha tasa de interés refleja el interés que generaría una inversión en el mercado de deuda iniciada a tiempo cero y que termina un periodo en el futuro cuyo capital inicial es de $C = 1$; y de forma inversa, refleja el interés a pagar si se pide prestado a tiempo cero una cantidad de $C = 1$ (en la moneda que estemos trabajando). A la tasa r se le llama *tasa libre de riesgo*, ya que refleja una inversión “segura”⁸ en el mercado de deuda.

Como consecuencia del principio de arbitraje se tiene el siguiente resultado,

Teorema 2.14. *Sea \mathbf{S} un activo financiero que cumple la definición 2.13 y sea r la tasa de interés establecida en el párrafo anterior, entonces se debe cumplir la siguiente ecuación como consecuencia del principio de arbitraje,*

$$d < 1 + r < u \tag{2.6}$$

Demostración. Para la demostración, véase Shreve (2004b). □

El resultado anterior es evidente, si por alguna causa no se cumple la ecuación 2.6 entonces uno nunca invertiría a la tasa r (en el caso de que $d \geq 1 + r$) o nunca invertiría en el activo \mathbf{S} (en el caso de que $u \leq 1 + r$), por lo que la incertidumbre del activo \mathbf{S} debe *girar alrededor* de r para ser considerado de interés para analizarlo.

Ahora bien, se considera un instrumento financiero derivado cualquiera \mathbf{V} . En las palabras de Hull (2011):

A derivative can be defined as a financial instrument whose value depends on (or derives from) the values of other, more basic underlying variables. Very often the variables underlying derivatives are the prices of traded assets. (...) However, derivatives can be dependent on almost any variable, from the price of hogs to the amount of snow falling at a certain ski resort.

⁸Cabe mencionar que el término “segura” debería ser sustituido por *altamente probable de cumplirse*, ya que en realidad ninguna inversión en los mercados financieros es segura completamente.

Por lo que un instrumento derivado puede depender del activo \mathbf{S} . Si este es el caso, en el siguiente desarrollo se concluirá que el valor de \mathbf{V} , como consecuencia del principio de arbitraje, depende completamente del activo \mathbf{S} y la tasa de interés r .

En efecto, ya que \mathbf{V} es un derivado financiero que cuenta con \mathbf{S} como subyacente, se busca construir un portafolio de inversión \mathbf{X} que cumpla la siguiente condición,

$$X_1 = V_1 \quad (2.7)$$

si este es el caso, entonces $X_0 = V_0$ como consecuencia del principio de arbitraje. Para esto, se representa a X_0 como sigue,

$$X_0 = \Delta_0 S_0 + (X_0 - \Delta_0 S_0) \quad (2.8)$$

donde Δ_0 representa la cantidad de acciones que se compran del activo \mathbf{S} a tiempo cero, mientras que $(X_0 - \Delta_0 S_0)$ representa una inversión que se va a realizar con respecto a la tasa de interés r .

Una vez ha pasado un periodo, X_1 esta constituido de la siguiente forma,

$$X_1 = \Delta_0 S_1 + (X_0 - \Delta_0 S_0)(1 + r) \quad (2.9)$$

y como consecuencia de la ecuación 2.7, se deben cumplir las siguientes dos ecuaciones,

$$V_1(H) = X_1(H) = \Delta_0 S_1(H) + (X_0 - \Delta_0 S_0)(1 + r) \quad (2.10)$$

$$V_1(T) = X_1(T) = \Delta_0 S_1(T) + (X_0 - \Delta_0 S_0)(1 + r) \quad (2.11)$$

restando 2.11 a 2.10 y despejando a Δ_0 , se obtiene lo siguiente,

$$\Delta_0 = \frac{V_1(H) - V_1(T)}{S_1(H) - S_1(T)} \quad (2.12)$$

donde como ya se había mencionado antes, Δ_0 representa la cantidad de acciones a comprar del activo \mathbf{S} a tiempo cero. Volviendo a las ecuaciones 2.10 y 2.11; si se multiplica 2.10 por un valor \tilde{p} , a 2.11 por $\tilde{q} = 1 - \tilde{p}$ y se suman, se obtiene lo siguiente,

$$X_0(1 + r) + \Delta_0 \left(S_1(H)\tilde{p} + S_1(T)\tilde{q} - S_0(1 + r) \right) = V_1(H)\tilde{p} + V_1(T)\tilde{q} \quad (2.13)$$

si se escoge \tilde{p} tal que,

$$S_0(1 + r) = S_1(H)\tilde{p} + S_1(T)\tilde{q} \quad (2.14)$$

entonces $X_0 = V_0$ equivale a lo siguiente,

$$V_0(1+r) = X_0(1+r) = V_1(H)\tilde{p} + V_1(T)\tilde{q} \quad (2.15)$$

como consecuencia de la ecuación 2.14 se tiene que \tilde{p} y \tilde{q} son equivalentes a las siguientes fracciones,

$$\tilde{p} = \frac{1+r-d}{u-d}; \quad \tilde{q} = \frac{u-1-r}{u-d} \quad (2.16)$$

y como consecuencia del teorema 2.14 se tiene que,

$$0 < \tilde{p}, \tilde{q} < 1 \quad (2.17)$$

La ecuación 2.17 da la pauta para crear el siguiente resultado,

Teorema 2.15. Sean \mathbf{S} y \mathbf{V} un activo en los mercados financieros y un derivado que tiene como subyacente al activo \mathbf{S} , respectivamente. Además se define sobre el espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ establecido en la definición 2.13 una nueva medida de probabilidad $\tilde{\mathbb{P}}$ que cumpla con las siguientes dos condiciones:

$$\begin{aligned} \tilde{\mathbb{P}}(S_1 = S_1(H)) &= \tilde{p} \\ \tilde{\mathbb{P}}(S_1 = S_1(T)) &= \tilde{q} \end{aligned}$$

Donde \tilde{p} y \tilde{q} son los valores escritos en la ecuación 2.16. Entonces se cumplen las siguientes dos ecuaciones:

$$S_0 = \frac{1}{1+r} \tilde{\mathbb{E}}(S_1) \quad (2.18)$$

$$V_0 = \frac{1}{1+r} \tilde{\mathbb{E}}(V_1) \quad (2.19)$$

es decir, el valor esperado descontado un periodo de S_1 y V_1 equivalen a sus valores presentes S_0 y V_0 , respectivamente.

Demostración. Consecuencia inmediata de las ecuaciones 2.14 y 2.15. □

Se puede notar gracias al teorema 2.15 como se realizó un *cambio de medida* o una *transformación* a la incertidumbre, esto con el objetivo de poder cumplir con el

principio de arbitraje. Y en esta nueva medida de probabilidad $\tilde{\mathbb{P}}$ se cumple que la tasa de interés r es constante, es decir:

$$\begin{aligned}\tilde{\mathbb{E}}\left(\frac{1}{1+r}S_1\right) &= S_0 = \frac{\tilde{\mathbb{E}}(S_1)}{1+r} \\ \tilde{\mathbb{E}}\left(\frac{1}{1+r}V_1\right) &= V_0 = \frac{\tilde{\mathbb{E}}(V_1)}{1+r}\end{aligned}$$

Por lo cual para cualquier activo \mathbf{A} que se ha definido sobre el espacio $(\Omega, \mathcal{F}, \tilde{\mathbb{P}})$ (en este caso \mathbf{S} y \mathbf{V}), si se denota como ${}_pA_1$ a la incertidumbre que tenemos respecto al activo \mathbf{A} un periodo en el futuro de forma descontada, entonces,

$$A_0 = \tilde{\mathbb{E}}({}_pA_1) \quad (2.20)$$

donde A_0 representa el valor del activo \mathbf{A} a tiempo cero.

La relación entre las medidas de probabilidad \mathbb{P} y $\tilde{\mathbb{P}}$ se hace mediante la siguiente variable aleatoria,

Definición 2.16. Sea Ω un espacio muestral finito y sea $\mathcal{F} = 2^\Omega$. Además sean \mathbb{P} y $\tilde{\mathbb{P}}$ dos medidas de probabilidad definidas sobre \mathcal{F} y si para todo $\omega \in \Omega$,

$$\mathbb{P}(\{\omega\}), \tilde{\mathbb{P}}(\{\omega\}) > 0 \quad (2.21)$$

es decir, \mathbb{P} y $\tilde{\mathbb{P}}$ dan probabilidad estrictamente positiva para cualquier elemento de Ω . Se denota como Z a la derivada de Radon-Nikodým de $\tilde{\mathbb{P}}$ con respecto a \mathbb{P} y queda definida como sigue para todo $\omega \in \Omega$,

$$Z(\omega) := \frac{\tilde{\mathbb{P}}(\{\omega\})}{\mathbb{P}(\{\omega\})} \quad (2.22)$$

la cual es una v.a. definida sobre los espacios de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ y $(\Omega, \mathcal{F}, \tilde{\mathbb{P}})$.

Teorema 2.17. Sean $\Omega, \mathcal{F}, \mathbb{P}, \tilde{\mathbb{P}}$ y Z establecidas como en la definición 2.16, entonces se cumplen las siguientes propiedades,

(i) $\mathbb{P}(Z > 0) = 1$.

(ii) $\mathbb{E}(Z) = 1$.

(iii) Para cualquier v.a. Y definida sobre los espacios de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ y $(\Omega, \mathcal{F}, \tilde{\mathbb{P}})$.

$$\tilde{\mathbb{E}}(Y) = \mathbb{E}(ZY) \quad (2.23)$$

Demostración. Sin pérdida de generalidad, se supone que $\Omega = \{\omega_1, \dots, \omega_n\}$ para un $n \in \mathbb{N}$ cualquiera. La propiedad (i) se sigue de forma inmediata, ya que se asume que $\mathbb{P}(\{\omega\}), \tilde{\mathbb{P}}(\{\omega\}) > 0$ para cada $\omega \in \Omega$.

Para la propiedad (ii), se asume que $\text{Ran } Z = \{z_1, \dots, z_k\}$ donde claramente $k \leq n$. Y además se definen los siguientes conjuntos disjuntos a pares,

$$A_j := \{\omega \in \Omega : Z(\omega) = z_j\}; \quad 1 \leq j \leq k$$

el valor $\mathbb{E}(Z)$ quedacomo sigue,

$$\mathbb{E}(Z) = \sum_{j=1}^k z_j \mathbb{P}(Z = z_j) = \sum_{j=1}^k \sum_{\omega \in A_j} z_j \mathbb{P}(\{\omega\})$$

nótese que para todo $z_j \in \text{Ran } Z$ existe un $\omega_{z_j} \in \Omega$ tal que,

$$z_j = \frac{\tilde{\mathbb{P}}(\{\omega_{z_j}\})}{\mathbb{P}(\{\omega_{z_j}\})}$$

entonces si $\omega \in A_j$ se cumple que $Z(\omega) = z_j$ y,

$$\begin{aligned} \frac{\tilde{\mathbb{P}}(\{\omega_{z_j}\})}{\mathbb{P}(\{\omega_{z_j}\})} &= \frac{\tilde{\mathbb{P}}(\{\omega\})}{\mathbb{P}(\{\omega\})} \\ \tilde{\mathbb{P}}(\{\omega\}) &= \frac{\tilde{\mathbb{P}}(\{\omega_{z_j}\})}{\mathbb{P}(\{\omega_{z_j}\})} \mathbb{P}(\{\omega\}) = z_j \mathbb{P}(\{\omega\}) \end{aligned}$$

por lo que $z_j \mathbb{P}(\{\omega\}) = \tilde{\mathbb{P}}(\{\omega\})$ y sustituyendo en el desarrollo de $\mathbb{E}(Z)$ se tiene lo siguiente,

$$\mathbb{E}(Z) = \sum_{j=1}^k \sum_{\omega \in A_j} \tilde{\mathbb{P}}(\{\omega\}) = \sum_{j=1}^k \tilde{\mathbb{P}}(A_j) = \tilde{\mathbb{P}}\left(\biguplus_{j=1}^k A_j\right) = \tilde{\mathbb{P}}(\Omega) = 1$$

Para (iii), se asume que $\text{Ran } Y = \{y_1, \dots, y_m\}$ donde claramente $m \leq n$. Y además se definen los siguientes conjuntos disjuntos a pares,

$$B_i := \{\omega \in \Omega : Y(\omega) = y_i\}; \quad 1 \leq i \leq m$$

y los siguientes $k \times m$ conjuntos disjuntos a pares para $1 \leq j \leq k$ y $1 \leq i \leq m$,

$$C_j^i := A_j \cap B_i$$

$\tilde{\mathbb{E}}(Y)$ queda como sigue,

$$\tilde{\mathbb{E}}(Y) = \sum_{i=1}^m y_i \tilde{\mathbb{P}}(Y = y_i) = \sum_{i=1}^m \sum_{\omega \in B_i} y_i \tilde{\mathbb{P}}(\{\omega\}) = \sum_{i=1}^m \sum_{\omega \in B_i} y_i Z(\omega) \mathbb{P}(\{\omega\})$$

para el caso de B_i se cumple lo siguiente,

$$\sum_{\omega \in B_i} y_i Z(\omega) \mathbb{P}(\{\omega\}) = \sum_{j=1}^k \sum_{\omega \in C_j^i} y_i z_j \mathbb{P}(\{\omega\})$$

y en la anterior suma, los valores y_i y z_j son constantes en C_j^i , por lo que,

$$\sum_{j=1}^k \sum_{\omega \in C_j^i} y_i z_j \mathbb{P}(\{\omega\}) = \sum_{j=1}^k y_i z_j \mathbb{P}(C_j^i)$$

y reemplazando la suma anterior en el desarrollo de $\tilde{\mathbb{E}}(Y)$ se tiene lo siguiente,

$$\tilde{\mathbb{E}}(Y) = \sum_{i=1}^m \sum_{\omega \in B_i} y_i Z(\omega) \mathbb{P}(\{\omega\}) = \sum_{i=1}^m \sum_{j=1}^k y_i z_j \mathbb{P}(C_j^i)$$

y por como se definio C_j^i se cumple que,

$$C_j^i = \{\omega \in \Omega : ZY(\omega) = Z(\omega)Y(\omega) = z_j y_i\}$$

por lo tanto,

$$\tilde{\mathbb{E}}(Y) = \sum_{i=1}^m \sum_{j=1}^k z_j y_i \mathbb{P}(ZY = z_j y_i) = \mathbb{E}(ZY)$$

□

Este resultado permite concluir que la valuación de activos mediante la medida neutral al riesgo $\tilde{\mathbb{P}}$ se puede sustituir mediante la medida \mathbb{P} y la derivada de Radon-Nikodým de $\tilde{\mathbb{P}}$ con respecto a \mathbb{P} .

2.3.2. El modelo binomial multiperiodo

En esta subsección, se establece el modelo binomial multiperiodo y como este se ve afectado por las medidas de \mathbb{P} y $\tilde{\mathbb{P}}$ (donde $\tilde{\mathbb{P}}$ de forma indirecta incluye el principio de arbitraje).

Definición 2.18. Sea $\mathbf{M} = (M_n)_{n \in \mathbb{N}_0}$ una caminata aleatoria cualquiera con probabilidad de subida p tal y como fué establecida en el ejemplo 2.2. Además, se toman dos valores u y d constantes tales que cumplan el teorema 2.14 para una tasa de interés r fija, se define el modelo binomial multiperiodo (también llamado árbol binomial multiperiodo) $\mathbf{S} = (S_n)_{n \in \mathbb{N}_0}$ de la siguiente forma,

(i) S_0 es un valor constante establecido inicialmente tal que $S_0 > 0$.

(ii) Para cualquier $n \geq 1$ se tiene lo siguiente,

$$S_n = S_0 u^{\frac{n+M_n}{2}} d^{\frac{n-M_n}{2}} \quad (2.24)$$

Una representación visual del proceso estocástico \mathbf{S} puede verse en la figura 2.2, se ve como la representación visual de dicho proceso estocástico tiene la forma de un árbol (con raíces horizontales), donde cada rama se biparte (de ahí el nombre de binomial).

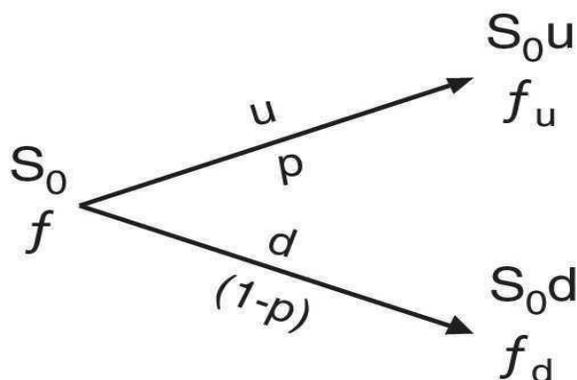


Figura 2.2: Modelo binomial multiperiodo

Se inicia el análisis de \mathbf{S} estableciendo las distribuciones de $\frac{n+M_n}{2}$ y $\frac{n-M_n}{2}$, respectivamente,

Teorema 2.19. *Sea \mathbf{S} el modelo binomial multiperiodo establecido en la definición 2.18, se definen los siguientes dos procesos estocásticos $\mathbf{H} = (H_n)_{n \in \mathbb{N}}$ y $\mathbf{T} = (T_n)_{n \in \mathbb{N}}$ como siguen,*

$$H_n := \frac{n + M_n}{2} \quad (2.25)$$

$$T_n := \frac{n - M_n}{2} \quad (2.26)$$

entonces se cumple las siguientes tres propiedades,

(i) $H_n \sim \text{Bin}(n, p)$ y $T_n \sim \text{Bin}(n, 1 - p)$.

(ii) Para todo $n \in \mathbb{N}$ se cumple que,

$$\mathbb{P}(T_n = n - h_n \mid H_n = h_n) = 1 \quad (2.27)$$

$$\mathbb{P}(H_n = n - t_n \mid T_n = t_n) = 1 \quad (2.28)$$

y por lo tanto,

$$\mathbb{P}(H_n + T_n = n) = 1 \quad (2.29)$$

(iii) Para todo $n \in \mathbb{N}$, $\text{Ran } S_n = \{S_0 u^i d^{n-i} : i \in \{0, 1, \dots, n\}\}$.

Demostración. Para la propiedad (i), la v.a. H_n equivale a la v.a. Y^n definida en el ejemplo 2.2, por lo que se cumple que $H_n \sim \text{Binom}(n, p)$. Para el caso de T_n ,

$$T_n = n - H_n$$

por lo que su función de masa de probabilidades queda como sigue,

$$\begin{aligned} p_{T_n}(t) &= \mathbb{P}(T_n = t) = \mathbb{P}(n - H_n = t) = \mathbb{P}(H_n = n - t) \\ &= \binom{n}{n-t} p^{n-t} (1-p)^t = \binom{n}{t} (1-p)^t p^{n-t} \end{aligned}$$

por lo tanto $T_n \sim \text{Binom}(n, 1 - p)$. Las propiedades (ii) y (iii) son evidentes ya que $H_n + T_n = n$.

□

Definición 2.20. Sea $\mathbf{S} = (S_n)_{n \in \mathbb{N}_0}$ el modelo binomial multiperiodo establecido en la definición 2.18, se define el proceso estocástico ${}_p\mathbf{S} = ({}_pS_n)_{n \in \mathbb{N}_0}$ como sigue,

$$(i) \quad {}_pS_0 = S_0$$

(ii) Para $n \in \mathbb{N}$,

$${}_pS_n := \frac{S_n}{(1+r)^n} \quad (2.30)$$

se dice que ${}_p\mathbf{S}$ es el modelo binomial multiperiodo descontado a tiempo cero (tiempo presente).

Si \mathbf{S} es un activo financiero cuya incertidumbre se modela mediante el árbol binomial multiperiodo establecido en la definición 2.18, entonces el proceso estocástico ${}_p\mathbf{S}$ representa la incertidumbre descontada a tiempo presente de dicho modelo. El siguiente lema sirve para demostrar el teorema 2.22.

Lema 2.21. Sean \mathbf{S} y ${}_p\mathbf{S}$ los modelos establecidos en las definiciones 2.18 y 2.20, respectivamente. Entonces bajo la medida neutral al riesgo $\tilde{\mathbb{P}}$ se cumple lo siguiente,

$$\tilde{\mathbb{E}}(S_n) = S_0(1+r)^n \quad (2.31)$$

$$\tilde{\mathbb{E}}({}_pS_n) = S_0 \quad (2.32)$$

Demostración. La demostración se hace por inducción. Sea $k = 1$, y se busca obtener $\tilde{\mathbb{E}}(S_1)$. Recordando que $\text{Ran } S_1 = \{S_0u, S_0d\}$ y que $\tilde{p} = \frac{1+r-d}{u-d}$,

$$\begin{aligned} \tilde{\mathbb{E}}(S_1) &= S_0u \tilde{\mathbb{P}}(S_1 = S_0u) + S_0d \tilde{\mathbb{P}}(S_1 = S_0d) \\ &= S_0u \tilde{p} + S_0d(1 - \tilde{p}) \\ &= S_0u \frac{1+r-d}{u-d} + S_0d \frac{u-(1+r)}{u-d} \\ &= \frac{(1+r)(S_0u - S_0d) + S_0ud - S_0ud}{u-d} \\ &= \frac{S_0(1+r)(u-d)}{u-d} = S_0(1+r) \end{aligned}$$

Se supone válido para $k = n$, y con esto se demuestra válido para $n + 1$. Los rangos de las v.a. S_n y S_{n+1} son los siguientes,

$$1) \quad \text{Ran } S_n = \{S_0u^i d^{n-i} : i \in \{0, 1, \dots, n\}\}.$$

$$2) \text{Ran } S_{n+1} = \{S_0 u^i d^{n-i} : i \in \{0, 1, \dots, n, n+1\}\}.$$

y además se cumple lo siguiente para la variable aleatoria $S_{n+1} | \{S_n = s_n\}$,

$$1) \text{Ran } S_{n+1} | \{S_n = s_n\} = \{s_n u, s_n d\}.$$

$$2) \tilde{\mathbb{P}}(S_{n+1} = s_n u | S_n = s_n) = \tilde{p}.$$

esto como consecuencia de la definición 2.18. Ahora bien, para obtener la esperanza $\tilde{\mathbb{E}}(S_{n+1})$ se utiliza el concepto de “esperanza iterada”⁹ de la siguiente forma,

$$\tilde{\mathbb{E}}(S_{n+1}) = \tilde{\mathbb{E}}\left(\tilde{\mathbb{E}}(S_{n+1} | S_n)\right)$$

Para la v.a. $S_{n+1} | \{S_n = s_n\}$ se cumple lo siguiente,

$$\tilde{\mathbb{E}}(S_{n+1} | S_n = s_n) = s_n(1+r)$$

por lo tanto $\tilde{\mathbb{E}}(S_{n+1} | S_n) = S_n(1+r)$ y,

$$\tilde{\mathbb{E}}(S_{n+1}) = \tilde{\mathbb{E}}\left(\tilde{\mathbb{E}}(S_{n+1} | S_n)\right) = \tilde{\mathbb{E}}(S_n(1+r)) = S_0(1+r)^n(1+r) = S_0(1+r)^{n+1}$$

quedando así demostrado 2.31. Para demostrar 2.32 es suficiente notar que ${}_p S_n = \frac{S_n}{(1+r)^n}$ por lo que,

$$\tilde{\mathbb{E}}({}_p S_n) = \tilde{\mathbb{E}}\left(\frac{S_n}{(1+r)^n}\right) = \frac{S_0(1+r)^n}{(1+r)^n} = S_0$$

quedando así demostrado el lema 2.21. □

Teorema 2.22. Sean \mathbf{S} y ${}_p \mathbf{S}$ los modelos establecidos en las definiciones 2.18 y 2.20, respectivamente. Entonces bajo la medida neutral al riesgo $\tilde{\mathbb{P}}$ se cumple que ${}_p \mathbf{S}$ es una martingala, donde \tilde{p} y \tilde{q} están definidas como sigue,

$$\tilde{p} := \frac{1+r-d}{u-d} \tag{2.33}$$

$$\tilde{q} := \frac{u-1-r}{u-d} \tag{2.34}$$

y esto es sobre la filtración $\mathbb{F} = \sigma({}_p \mathbf{S}) = \sigma(\mathbf{S})$.

⁹El concepto de “esperanza iterada” se puede consultar en Casella *et. al.* (2002) en el Teorema 4.3.3.

Demostración. Para la demostración del teorema 2.22 es suficiente ver que para cualquier $i < j$ se cumple lo siguiente,

$$\tilde{\mathbb{E}}({}_pS_j | \mathcal{F}_i) = {}_pS_i \quad \mathcal{F}_i \in \mathbb{F}$$

Para esto, se escribe ${}_pS_j$ de alguna forma equivalente que sea mas apropiada para analizar. Recordando lo establecido en el ejercicio 2.3,

$$M_j = M_i + M^*$$

donde M_i y M^* son v.a. independientes y $M^* = M_{j-i}$. Por lo que H_j y T_j quedan como sigue,

$$\begin{aligned} H_j &= \frac{j + M_j}{2} = \frac{i + M_i + (j - i) + M^*}{2} = \frac{i + M_i}{2} + \frac{(j - i) + M^*}{2} \\ T_j &= \frac{j - M_j}{2} = \frac{i - M_i + (j - i) - M^*}{2} = \frac{i - M_i}{2} + \frac{(j - i) - M^*}{2} \end{aligned}$$

por lo tanto,

$$\begin{aligned} {}_pS_j &= \frac{1}{(1+r)^j} S_0 u^{\frac{j+M_j}{2}} d^{\frac{j-M_j}{2}} = \frac{1}{(1+r)^j} S_0 u^{\frac{i+M_i}{2} + \frac{(j-i)+M^*}{2}} d^{\frac{i-M_i}{2} + \frac{(j-i)-M^*}{2}} \\ &= \frac{1}{(1+r)^i} S_0 u^{\frac{i+M_i}{2}} d^{\frac{i-M_i}{2}} \frac{1}{(1+r)^{j-i}} S_0 u^{\frac{(j-i)+M^*}{2}} d^{\frac{(j-i)-M^*}{2}} \frac{1}{S_0} \\ &= \frac{1}{S_0} {}_pS_i {}_pS^* \end{aligned}$$

donde la v.a. ${}_pS_i$ y ${}_pS^*$ son independientes (ya que M_i y M^* lo son) y ${}_pS^*$ esta definida como sigue,

$${}_pS^* := \frac{1}{(1+r)^{j-i}} S_0 u^{\frac{(j-i)+M^*}{2}} d^{\frac{(j-i)-M^*}{2}}$$

Ahora bien, se toma un $A \in \mathcal{F}_i$ cualquiera, como ${}_pS_i$ y ${}_pS^*$ son independientes entre si, entonces también $\mathbf{1}_A$ y ${}_pS^*$ lo son, por lo que,

$$\tilde{\mathbb{E}}({}_pS_j \mathbf{1}_A) = \frac{1}{S_0} \tilde{\mathbb{E}}({}_pS_i {}_pS^* \mathbf{1}_A) = \frac{1}{S_0} \tilde{\mathbb{E}}({}_pS_i \mathbf{1}_A) \tilde{\mathbb{E}}({}_pS^*)$$

y ya que $M^* = M_{j-i}$ entonces ${}_pS^* = {}_pS_{j-i}$ y por consecuencia del lema 2.21 se cumple que $\tilde{\mathbb{E}}({}_pS^*) = S_0$ concluyendo que,

$$\tilde{\mathbb{E}}({}_pS_j \mathbf{1}_A) = \frac{S_0}{S_0} \tilde{\mathbb{E}}({}_pS_i \mathbf{1}_A) = \tilde{\mathbb{E}}({}_pS_i \mathbf{1}_A)$$

y por como se estableció la definición 2.5 para todo $i < j$,

$$\tilde{\mathbb{E}}({}_pS_j | \mathcal{F}_i) = {}_pS_i \quad \mathcal{F}_i \in \mathbb{F}$$

□

Con estos resultados se puede hacer la siguiente observación,

Observación 2.22.1. *Para la caminata aleatoria \mathbf{M} con probabilidad de subida p se demostró que la propiedad de martingala solo se cumplía si y sólo si $p = 1/2$ (es decir, la caminata aleatoria era justa).*

Mientras que para el proceso ${}_p\mathbf{S}$, la propiedad de martingala se consigue haciendo un “cambio de medida” el cual transforma el valor de p a un valor \tilde{p} . Además, toda esta información esta disponible en la derivada de Radon-Nikodým de $\tilde{\mathbb{P}}$ con respecto a \mathbb{P} establecida en la definición 2.16. Si dicha derivada la denotamos como Z (para simplificar la notación), entonces se tiene lo siguiente:

- 1) *El proceso ${}_p\mathbf{S} = ({}_pS_n)_{n \in \mathbb{N}_0}$ es una martingala con respecto a $\tilde{\mathbb{P}}$.*
- 2) *El proceso $\mathbf{Z}_p\mathbf{S} := (Z {}_pS_n)_{n \in \mathbb{N}_0}$ es una martingala con respecto a \mathbb{P} .*

y esto sobre la filtración $\mathbb{F} = \sigma(\mathbf{S})$.

Una vez completado el análisis del árbol binomial multiperiodo \mathbf{S} y su respectivo árbol binomial descontado ${}_p\mathbf{S}$, se considera un derivado \mathbf{V} que cuente como subyacente al activo \mathbf{S} y que tenga como fecha de madurez un tiempo $n \in \mathbb{N}$ en el futuro. Se establecen las siguiente funciones:

- (i) *Función *Payoff**: Dicha función se denota como G_V y representa el valor del derivado a tiempo n (fecha de madurez) dado el valor de S_n , es decir,

$$V_n = G_V(S_n) \tag{2.35}$$

- (ii) Función de valuación del derivado \mathbf{V} : Dicha función se denota con la letra g_V y representa el valor del derivado a tiempo $m \leq n$ dado el valor de S_m , es decir,

$$V_m = g_V(S_m, m) \quad (2.36)$$

Desde que queda establecido el derivado \mathbf{V} , la función G_V (el Payoff) queda completamente definida, lo único que falta es encontrar la función g_V para todo $m < n$, y para el caso de $m = n$,

$$g_V(S_n, n) = G_V(S_n) \quad (2.37)$$

Cabe aclarar, que las funciones G_V y g_V no son variables aleatorias, ya que el valor del derivado debe quedar completamente definido gracias al principio de arbitraje así como por los valores admisibles \mathbf{S} .

Para obtener los valores admisibles para las variables aleatorias V_0, V_1, \dots, V_{n-1} (donde n sigue representando la fecha de madurez del derivado), se utiliza el principio de arbitraje y el árbol binomial analizado en la subsección anterior. La siguiente información es conocida desde el principio:

- (i) Se supone que el activo \mathbf{S} es modelado mediante un árbol binomial multiperiodo, y en este caso dicho proceso estocástico nos interesa hasta tiempo n , ya que el derivado termina de hacer efecto después de ese periodo.
- (ii) Se conoce la función G_V y por lo tanto los valores que adquiere el derivado a tiempo n . Es decir, bajo la medida neutral al riesgo, la función de masa de V_n queda definida como sigue para cualquier $s_n \in \text{Ran } S_n$,

$$\tilde{\mathbb{P}}(V_n = G_V(s_n)) = \tilde{\mathbb{P}}(S_n = s_n)$$

Ahora bien, para el caso V_{n-1} , se sabe que para cualquier $s^* \in S_{n-1}$ se debe cumplir lo siguiente,

$$\tilde{\mathbb{E}}\left(\frac{S_n}{1+r} \mid S_{n-1} = s^*\right) = s^* \quad (2.38)$$

esto como consecuencia del lema 2.21.

Nótese que las v.a. $\{S_{n-1} \mid \{S_{n-1} = s^*\}, S_n \mid \{S_{n-1} = s^*\}\}$ representan un árbol binomial a un periodo, por lo cual se debe cumplir las siguientes condiciones como consecuencia del principio de arbitraje,

- (i) La variable aleatoria $S_{n-1} \mid \{S_{n-1} = s^*\}$ es constante y se cumple que,

$$S_{n-1} \mid \{S_{n-1} = s^*\} = s^* \quad (2.39)$$

- (ii) La variable aleatoria $S_n | \{S_{n-1} = s^*\}$ tiene dos valores admisibles $\{s^*u, s^*d\}$ tales que,

$$\begin{aligned}\tilde{\mathbb{P}}(S_n = s^*u | \{S_{n-1} = s^*\}) &= \tilde{p} \\ \tilde{\mathbb{P}}(S_n = s^*d | \{S_{n-1} = s^*\}) &= \tilde{q}\end{aligned}$$

y como resultado de lo establecido en el árbol binomial a un periodo, se debe cumplir que,

$$\tilde{\mathbb{E}}\left(\frac{V_n}{1+r} \middle| S_{n-1} = s^*\right) = V_{n-1} | \{S_{n-1} = s^*\} \quad (2.40)$$

donde la anterior esperanza neutral al riesgo esta dada por la siguiente ecuación,

$$\tilde{\mathbb{E}}\left(\frac{V_n}{1+r} \middle| S_{n-1} = s^*\right) = \frac{g_V(s^*u, n)\tilde{p} + g_V(s^*d, n)\tilde{q}}{1+r} \quad (2.41)$$

esto se hace para cualquier $s^* \in \text{Ran } S_{n-1}$ y la v.a. $V_{n-1} | \{S_{n-1} = s^*\}$ debe ser constante (por como está definido el árbol binomial a un periodo). Por lo cual,

$$g_V(s^*, n-1) = V_{n-1} | \{S_{n-1} = s^*\} \quad (2.42)$$

por lo que todos los valores de $g_V(s^*, n-1)$ para $s^* \in \text{Ran } S_{n-1}$ quedan definidos por la ecuación 2.42 y se cumple que,

$$\tilde{\mathbb{P}}(V_{n-1} = g_V(s^*, n-1)) = \tilde{\mathbb{P}}(S_{n-1} = s^*) \quad (2.43)$$

Una vez obtenidos los valores de V_{n-1} se pueden obtener los valores de V_{n-2} siguiendo el mismo argumento ya establecido pero ahora usando a V_{n-1} . Por lo que para conocer todo el proceso estocástico \mathbf{V} se realiza un análisis partiendo de “adelante hacia atrás”, es decir, conociendo primeramente V_n , luego V_{n-1} y así sucesivamente hasta llegar a V_0 el cual representa el valor al que debe venderse/comprarse el derivado a tiempo cero (valor presente del derivado).

El siguiente ejemplo de un árbol binomial a dos periodos muestra de forma práctica la valuación de un derivado \mathbf{V} .

Ejemplo 2.23 (Árbol binomial a dos periodos). Sean $u = 1.3$, $d = 0.8$, $r = 0.1$, $S_0 = 10$ los valores necesarios para construir un árbol binomial a dos periodos ($n = 2$) que represente la incertidumbre que se tiene respecto a un activo \mathbf{S} . Para dicho activo

se establece un derivado \mathbf{V} que tiene como fecha de madurez $n = 2$ y que cuenta con el siguiente Payoff,

$$G_V(S_2) = \max\{S_2 - 9, 0\}$$

Con esta información es suficiente para analizar el comportamiento de \mathbf{V} . Primeramente se resume el comportamiento de \mathbf{S} , sabiendo que $\tilde{p} = 0.6$,

(i) Para el activo S_0 (valor actual del activo).

$$\tilde{\mathbb{P}}(S_0 = 10) = 1$$

(ii) Para el activo S_1 (valor del activo un periodo en el futuro).

$$\tilde{\mathbb{P}}(S_1 = S_0u) = \tilde{\mathbb{P}}(S_1 = 13) = \tilde{p} = 0.6$$

$$\tilde{\mathbb{P}}(S_1 = S_0d) = \tilde{\mathbb{P}}(S_1 = 8) = \tilde{q} = 0.4$$

(iii) Para el activo S_2 (valor del activo dos periodos en el futuro).

$$\tilde{\mathbb{P}}(S_2 = S_0u^2) = \tilde{\mathbb{P}}(S_2 = 16.9) = \tilde{p}^2 = 0.36$$

$$\tilde{\mathbb{P}}(S_2 = S_0ud) = \tilde{\mathbb{P}}(S_2 = 10.4) = 2\tilde{p}\tilde{q} = 0.48$$

$$\tilde{\mathbb{P}}(S_2 = S_0d^2) = \tilde{\mathbb{P}}(S_2 = 6.4) = \tilde{q}^2 = 0.16$$

Ahora bien, el análisis de \mathbf{V} se inicia obteniendo los valores admisibles para V_2 lo cuales están completamente especificados por la función G_V (el Payoff del derivado),

$$\tilde{\mathbb{P}}(V_2 = G_V(16.9)) = \tilde{\mathbb{P}}(V_2 = 7.9) = \tilde{\mathbb{P}}(S_2 = 16.9) = 0.36$$

$$\tilde{\mathbb{P}}(V_2 = G_V(10.4)) = \tilde{\mathbb{P}}(V_2 = 1.4) = \tilde{\mathbb{P}}(S_2 = 10.4) = 0.48$$

$$\tilde{\mathbb{P}}(V_2 = G_V(6.4)) = \tilde{\mathbb{P}}(V_2 = 0.0) = \tilde{\mathbb{P}}(S_2 = 6.4) = 0.16$$

para los valores admisibles de V_1 se cumple lo siguiente,

$$g_V(13, 1) = \tilde{\mathbb{E}}\left(\frac{V_2}{1+r} \middle| S_1 = 13\right) = \frac{7.9\tilde{p} + 1.4\tilde{q}}{1+r} = 4.8181$$

$$g_V(8, 1) = \tilde{\mathbb{E}}\left(\frac{V_2}{1+r} \middle| S_1 = 8\right) = \frac{1.4\tilde{p} + 0.0\tilde{q}}{1+r} = 0.7636$$

por lo que,

$$\begin{aligned}\tilde{\mathbb{P}}(V_1 = g_V(13, 1)) &= \tilde{\mathbb{P}}(V_1 = 4.8181) = \tilde{\mathbb{P}}(S_1 = 13) = 0.6 \\ \tilde{\mathbb{P}}(V_1 = g_V(8, 1)) &= \tilde{\mathbb{P}}(V_1 = 0.7636) = \tilde{\mathbb{P}}(S_1 = 8) = 0.4\end{aligned}$$

por último, se obtiene el valor de V_0 ,

$$V_0 = \tilde{\mathbb{E}}\left(\frac{V_1}{1+r} \mid S_0 = 10\right) = \frac{4.8181\tilde{p} + 0.7636\tilde{q}}{1+r} = 2.9057$$

y como V_0 representa el valor actual del derivado, entonces claramente,

$$\tilde{\mathbb{P}}(V_0 = 2.9057) = 1$$

Para terminar esta subsección, es importante generalizar el portafolio que replica al derivado financiero \mathbf{V} , esto se hace obteniendo el proceso estocástico $\{\Delta_i\}_{I_{n-1}}$ que representa la cantidad de dinero invertido a cada tiempo $i \in \{0, 1, \dots, n-1\}$ en el activo subyacente S_i , esto permite crear un portafolio que replique perfectamente a \mathbf{V} , dicho portafolio se denota como \mathbf{X} y debe cumplir con la siguiente condición para cualquier $i \in \{0, 1, \dots, n-1\}$,

$$X_i = \Delta_i S_i + (X_i - \Delta_i S_i) \quad (2.44)$$

con un argumento equivalente al que se utiliza para obtener $\{V_i\}_{I_{n-1}}$, se debe cumplir lo siguiente para Δ_i ,

$$\Delta_i | \{S_i = s\} = \frac{g_V(su, i+1) - g_V(sd, i+1)}{su - sd} \quad (2.45)$$

y esto es válido para cualquier $s \in \text{Ran } S_i$; esto permite tener que $X_k = V_k$ para todo $k \in \{0, \dots, n\}$. En el siguiente ejemplo se ve de forma práctica como se comporta el portafolio \mathbf{X} con relación a los datos dados en el ejemplo 2.23.

Ejemplo 2.24. *Utilizando los valores establecidos de u , d , r , S_0 y G_V establecidos en el ejemplo 2.23, encontrar como se distribuyen los procesos estocásticos $\{\Delta_i\}_{I_{n-1}}$ y $\{X_i\}_{I_{n-1}}$.*

(i) Para el caso de Δ_0 se tiene que $S_0 = 10$, por lo que,

$$\Delta_0 = \frac{g_V(13, 1) - g_V(8, 1)}{13 - 8} = \frac{4.8181 - 0.7636}{13 - 8} = 0.8109$$

(ii) Para el caso de Δ_1 existen dos posibilidades,

$$\Delta_1 | \{S_1 = 13\} = \frac{g_V(16.9, 2) - g_V(10.4, 2)}{16.9 - 10.4} = \frac{7.9 - 1.4}{16.9 - 10.4} = 1.0$$

$$\Delta_1 | \{S_1 = 8\} = \frac{g_V(10.4, 2) - g_V(6.4, 2)}{10.4 - 6.4} = \frac{1.4 - 0.0}{10.4 - 6.4} = 0.35$$

y cada uno con las siguientes probabilidades de ser escogidas (de forma equivalente a lo establecido para \mathbf{V}),

$$\tilde{\mathbb{P}}(\Delta_1 = 1.0) = \tilde{\mathbb{P}}(S_1 = 13) = \tilde{p}$$

$$\tilde{\mathbb{P}}(\Delta_1 = 0.35) = \tilde{\mathbb{P}}(S_1 = 8) = \tilde{q}$$

Nótese como $\Delta_1 | \{S_1 = 13\} = 1$, esto se debe a que $V_1 | \{S_1 = 13\} > 0$, es decir, no se tiene ninguna pérdida en el derivado \mathbf{V} a tiempo 2 una vez haya sucedido el evento $\{S_1 = 13\}$, por lo cual se invierte todo el portafolio en el subyacente \mathbf{X} .

Para el comportamiento del portafolio \mathbf{X} se analizan los posibles eventos a continuación. Se inicia con $X_0 = V_0 = 2.9057$, dicho dinero se divide de la siguiente forma,

$$X_0 = \Delta_0 S_0 + (X_0 - \Delta_0 S_0)$$

$$X_0 = 0.8109 S_0 + (-5.2033)$$

(i) Para el caso de que $\{S_1 = 13\}$, entonces,

$$X_1 | \{S_1 = 13\} = 0.8109 * 13 + (1.1)(-5.2033) = 4.8181 = V_1 | \{S_1 = 13\}$$

una vez se tiene esto, entonces $\Delta_1 | \{S_1 = 13\} = 1$ y se “rebalancea” el portafolio \mathbf{X} de la siguiente forma,

$$X_1 | \{S_1 = 13\} = 1S_1 | \{S_1 = 13\} + (-8.1819)$$

a continuación, existen dos posibilidades,

(ia) Para el caso de que $\{S_2 = 16.9\}$,

$$X_2 | \{S_2 = 16.9\} = 1 * 16.9 + (1.1)(-8.1819) = 7.9 = V_2 | \{S_2 = 16.9\}$$

(ib) Para el caso de que $\{S_2 = 10.4\}$,

$$X_2 | \{S_2 = 10.4\} = 1 * 10.4 + (1.1)(-8.1819) = 1.4 = V_2 | \{S_2 = 10.4\}$$

(ii) Para el caso de que $\{S_1 = 8\}$, entonces,

$$X_1 | \{S_1 = 8\} = 0.8109 * 8 + (1.1)(-5.2033) = 0.7636 = V_1 | \{S_1 = 8\}$$

una vez se tiene esto, entonces $\Delta_1 | \{S_1 = 8\} = 0.35$ y se “rebalancea” el portafolio \mathbf{X} de la siguiente forma,

$$X_1 | \{S_1 = 8\} = 0.35S_1 | \{S_1 = 8\} + (-2.0364)$$

a continuación, existen dos posibilidades,

(iia) Para el caso de que $\{S_2 = 10.4\}$,

$$X_2 | \{S_2 = 10.4\} = 0.35 * 10.4 + (1.1)(-2.0364) = 1.4 = V_2 | \{S_2 = 10.4\}$$

(iib) Para el caso de que $\{S_2 = 6.4\}$,

$$X_2 | \{S_2 = 6.4\} = 0.35 * 6.4 + (1.1)(-2.0364) = 0.0 = V_2 | \{S_2 = 6.4\}$$

Claramente se nota como el portafolio \mathbf{X} replica perfectamente al derivado \mathbf{V} sin importar cualquier escenario del activo \mathbf{S} .

2.4. El modelo a tiempo continuo

Esta sección entrelaza el modelo binomial multiperiodo establecido en la sección 2.3 con el modelo continuo. Se ve como el modelo continuo es un modelo “límite” del modelo discreto, esto en términos de convergencia en distribución de variables aleatorias. Al terminar esta sección se cuenta con el modelo continuo utilizado en la fórmula de Balck, Sholes y Merton para la valuación de opciones europeas.

Para iniciar este capítulo, se necesita “aumentar la velocidad” con la que los pasos de la caminata aleatoria se dan, cuando la velocidad de dichos pasos tiende a infinito, se llega como consecuencia al modelo continuo con pasos “infinitesimalmente” continuos. Para esto se inicia con la siguiente definición,

Definición 2.25. Sea \mathbf{M} una caminata aleatoria con probabilidad de subida p . Se define la caminata aleatoria escalada n veces con probabilidad de subida p como el proceso estocástico $\mathbf{W}^{(n)} = (W_t^{(n)})_{t \geq 0}$ siguiente,

1) Si $nt \in \mathbb{N}$ entonces,

$$W_t^{(n)} := \frac{1}{\sqrt{n}} M_{nt}$$

2) Si $nt \notin \mathbb{N}$ entonces,

$$\begin{aligned} W_t^{(n)} &:= W_s^{(n)} + \frac{W_u^{(n)} - W_s^{(n)}}{u - s} (t - s) \\ &= \frac{1}{\sqrt{n}} M_{ns} + \frac{1}{\sqrt{n}} \frac{t - s}{u - s} (M_{nu} - M_{ns}) \end{aligned}$$

donde s y u son los valores mas cercanos a t , por izquierda y derecha respectivamente, tales que $ns, nu \in \mathbb{N}$.

Si $p = \frac{1}{2}$ (es decir, \mathbf{M} es una caminata aleatoria simétrica), se dice que $\mathbf{W}^{(n)}$ es una caminata aleatoria simétrica escalada n veces.

Proposición 2.26. Sea $n \in \mathbb{N}$ cualquiera y sea $\mathbf{W}^{(n)}$ la caminata aleatoria escalada n veces con probabilidad de subida p , entonces se cumple lo siguiente:

1) $\mathbb{E}(W_t^{(n)}) = t \sqrt{n} (2p - 1)$

2) El proceso $\mathbf{W}^{(n)}$ **no** es una martingala.

Demostración. Para el inciso 1),

1) Si $tn \in \mathbb{N}$,

$$\begin{aligned} \mathbb{E}(W_t^{(n)}) &= \mathbb{E}\left(\frac{1}{\sqrt{n}} M_{nt}\right) = \frac{1}{\sqrt{n}} \mathbb{E}(M_{nt}) = \frac{2ntp - nt}{\sqrt{n}} \\ &= \frac{nt(2p - 1)}{\sqrt{n}} = t \sqrt{n} (2p - 1) \end{aligned}$$

2) Si $tn \notin \mathbb{N}$, entonces se escoge s y u como los valores mas cercanos a t , por izquierda y derecha respectivamente, tales que $ns, nu \in \mathbb{N}$. Entonces,

$$\begin{aligned} \mathbb{E}(W_t^{(n)}) &= \mathbb{E}\left(\frac{1}{\sqrt{n}} M_{ns} + \frac{t - s}{u - s} \frac{1}{\sqrt{n}} (M_{nu} - M_{ns})\right) \\ &= s \sqrt{n} (2p - 1) + \frac{t - s}{u - s} \sqrt{n} (u - s) (2p - 1) \\ &= s \sqrt{n} (2p - 1) + (t - s) \sqrt{n} (2p - 1) \\ &= t \sqrt{n} (2p - 1) \end{aligned}$$

Para el inciso 2) se da un fundamento lógico por el cual no se cumple la propiedad de martingala. Sean $s, u \in \mathbb{R}$ tales que $s = u - 1/n$ y $un \in \mathbb{N}$, se toma un $t \in \mathbb{R}$ tal que $s < t < u$, entonces se cumple lo siguiente,

$$W_t^{(n)} = W_s^{(n)} + \frac{W_u^{(n)} - W_s^{(n)}}{u - s}(t - s)$$

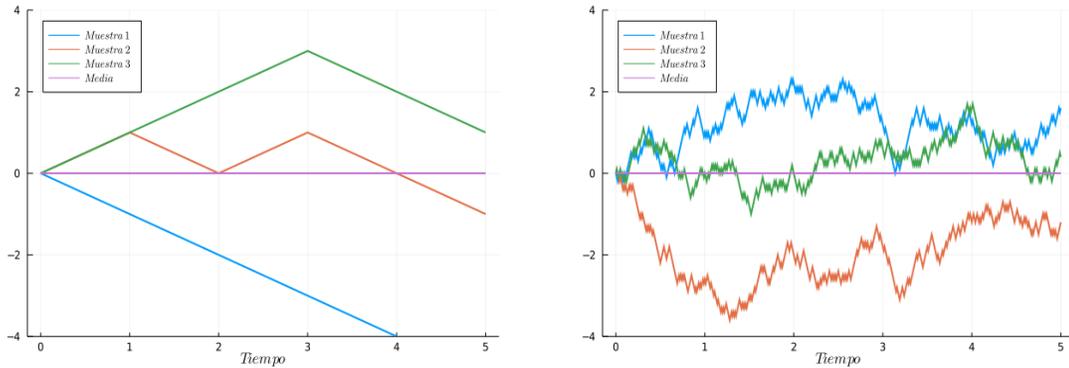
con esta información, es evidente que $W_t^{(n)}$ no es independiente de $W_u^{(n)}$ por lo cual la σ -álgebra \mathcal{F}_s (información disponible hasta el tiempo $s < t$) no es suficiente para conocer toda la incertidumbre que se tiene con respecto a $W_t^{(n)}$, se necesita información “futura” ($u > t$) para analizar la incertidumbre de $W_t^{(n)}$. Bajo este razonamiento se ve que $\mathbb{E}(W_t^{(n)} | \mathcal{F}_s) \neq W_s^{(n)}$.

□

En la figura 2.3 se puede ver como se comporta la caminata aleatoria simétrica y como esta se compara con la caminata aleatoria simétrica escalada. Para el proceso $\mathbf{W}^{(n)}$, el valor n representa “cuantos saltos” se dan en el intervalo $[0, 1]$, dichos saltos se dan en los siguientes valores,

$$\left\{ \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1 \right\}$$

y gracias al término $\frac{1}{\sqrt{n}}$ utilizado en la definición de la caminata aleatoria simétrica escalada, los saltos son cada vez mas “pequeños”.



(a) No escalada

(b) Escalada 100 veces

Figura 2.3: Caminatas Aleatorias Simétricas

Con lo descrito visualmente en la figura 2.3, una pregunta natural al respecto sería la siguiente: ¿El proceso estocástico $\mathbf{W}^{(n)}$ converge a algún proceso en específico? y la respuesta a esta pregunta la da el teorema 2.31. Para llegar a este resultado, se establecen a continuación definiciones y resultados con respecto a convergencia de v.a. en distribución y como estos resultados se relacionan con la función “generadora de momentos”.

Definición 2.27. (Convergencia en distribución) Sea $\mathbf{X} = (X_n)_{n \in \mathbb{N}}$ una sucesión de variables aleatorias cualquiera. Se dice que \mathbf{X} converge en distribución a una v.a. X , si para todo punto x en donde la función F_X es continua, se cumple que,

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x)$$

Esto se denota como $X_n \xrightarrow{d} X$, o bien, $F_{X_n} \xrightarrow{d} F_X$.

Definición 2.28. Sea X una v.a. cualquiera, se define la función generadora de momentos de X como sigue,

$$M_X(t) = \mathbb{E}(e^{tX})$$

definida para $t \in \mathbb{R}$ donde la anterior esperanza existe y es finita.

Proposición 2.29. Sea X y Y v.a. independientes y cuyas funciones generadoras de momentos existen en una vecindad no trivial alrededor del cero. Entonces para cualquier $t \in (-s, s)$ para algún $s > 0$,

$$M_{X+Y}(t) = M_X(t) M_Y(t)$$

Demostración. Ya que X y Y son independientes se cumple lo siguiente,

$$M_{X+Y}(t) = \mathbb{E}(e^{t(X+Y)}) = \mathbb{E}(e^{tX} e^{tY}) \stackrel{ind}{=} \mathbb{E}(e^{tX}) \mathbb{E}(e^{tY}) = M_X(t) M_Y(t)$$

demostrándose lo deseado. □

Teorema 2.30. Sea $\mathbf{X} = (X_n)_{n \in \mathbb{N}}$ una sucesión de variables aleatorias cuyas funciones generadoras de momentos existen todas ellas en algún intervalo no trivial alrededor del cero, y sea X una v.a. con función generadora de momentos M_X . Entonces $X_n \xrightarrow{d} X$ si, y sólo si, $M_{X_n}(t) \xrightarrow{d} M_X(t)$.

Demostración. Este teorema es una versión mas simple del *Teorema de continuidad de Lévy*. Para la demostración de dicho teorema puede verser *Klenke op. cit.* capítulo 15.3. \square

Con esta información, se puede establecer la convergencia en distribución del proceso estocástico $\mathbf{W}^{(n)}$.

Teorema 2.31. *Para t fijo. Mientras $n \rightarrow \infty$, la distribución de la caminata aleatoria simétrica escalada $\mathbf{W}^{(n)}$ evaluada a tiempo t converge en distribución a una v.a. normal con media cero y varianza t . Es decir, para t fijo,*

$$W_t^{(n)} \xrightarrow{d} W_t$$

donde $W_t \sim N(0, \sqrt{t})$.

Demostración. Para esta demostración, se analiza las funciones generadoras de momentos del modelo $W_t^{(n)}$ para $n \in \mathbb{N}$ y de W_t . Se sabe que M_{W_t} esta dada por la siguiente ecuación,

$$M(u) := M_{W_t}(u) = e^{\frac{1}{2}u^2t}$$

Ahora bien, si t es tal que $nt \in \mathbb{N}$ entonces la función generadora de momentos para $W_t^{(n)}$ es,

$$\begin{aligned} \varphi_n(u) &:= M_{W_t^{(n)}}(u) = \mathbb{E}\left(e^{uW_t^{(n)}}\right) = \mathbb{E}\left(\exp\left(\frac{u}{\sqrt{n}}M_{nt}\right)\right) \\ &= \mathbb{E}\left(\exp\left(\frac{u}{\sqrt{n}}\sum_{j=1}^{nt}X_j\right)\right) = \mathbb{E}\left(\prod_{j=1}^{nt}\exp\left(\frac{u}{\sqrt{n}}X_j\right)\right) \end{aligned}$$

donde X_j son las variables que conforman la caminata aleatoria simétrica establecida en el ejemplo 2.2 (recordando que las X_j son independientes entre sí). Por lo que,

$$\begin{aligned} \varphi_n(u) &\stackrel{ind}{=} \prod_{j=1}^{nt} \mathbb{E}\left(\exp\left(\frac{u}{\sqrt{n}}X_j\right)\right) = \prod_{j=1}^{nt} \left(\frac{1}{2}e^{\frac{u}{\sqrt{n}}} + \frac{1}{2}e^{-\frac{u}{\sqrt{n}}}\right) \\ &= \left(\frac{1}{2}e^{\frac{u}{\sqrt{n}}} + \frac{1}{2}e^{-\frac{u}{\sqrt{n}}}\right)^{nt} \end{aligned}$$

Se necesita demostrar que,

$$\lim_{n \rightarrow \infty} \varphi_n(u) = M(u) = e^{\frac{1}{2}u^2t}$$

Para obtener dicho límite, es suficiente considerar el logaritmo de $\varphi_n(u)$ y demostrar que,

$$\log \varphi_n(u) = nt \log \left(\frac{1}{2} e^{\frac{u}{\sqrt{n}}} + \frac{1}{2} e^{-\frac{u}{\sqrt{n}}} \right) \xrightarrow{n \rightarrow \infty} \frac{1}{2} u^2 t$$

para esto, se hace el cambio de variable $x = \frac{1}{\sqrt{n}}$ y como $n \rightarrow \infty$ entonces $x \rightarrow 0^+$ y el límite queda como sigue,

$$\lim_{n \rightarrow \infty} \varphi_n(u) = t \lim_{x \rightarrow 0^+} \frac{\log \left(\frac{1}{2} e^{\frac{u}{\sqrt{n}}} + \frac{1}{2} e^{-\frac{u}{\sqrt{n}}} \right)}{x^2} = \frac{0}{0}$$

y usando la regla de L'Hôpital,

$$\lim_{n \rightarrow \infty} \varphi_n(u) = t \lim_{x \rightarrow 0^+} \frac{\frac{u}{2} e^{ux} - \frac{u}{2} e^{-ux}}{2x \left(\frac{1}{2} e^{ux} + \frac{1}{2} e^{-ux} \right)} = \frac{t}{2} \lim_{x \rightarrow 0^+} \frac{\frac{u}{2} e^{ux} - \frac{u}{2} e^{-ux}}{x} = \frac{0}{0}$$

y usando la regla de L'Hôpital nuevamente,

$$\lim_{n \rightarrow \infty} \varphi_n(u) = \frac{t}{2} \lim_{x \rightarrow 0^+} \left(\frac{u^2}{2} e^{ux} + \frac{u^2}{2} e^{-ux} \right) = \frac{1}{2} u^2 t$$

Ahora bien, para el caso $nt \notin \mathbb{N}$, entonces se escoge s_1 y s_2 como los valores mas cercanos a t , por izquierda y derecha respectivamente, tales que $ns_1, ns_2 \in \mathbb{N}$ y se cumple que,

$$W_t^{(n)} = W_{s_1}^{(n)} + \left(W_{s_2}^{(n)} - W_{s_1}^{(n)} \right) \frac{t - s_1}{s_2 - s_1}$$

recordando que $W_{s_1}^{(n)}$ y $W_{s_2}^{(n)} - W_{s_1}^{(n)}$ son v.a. independientes (como consecuencia de la construcción de la caminata aleatoria) y además sabiendo que $W_{s_2}^{(n)} - W_{s_1}^{(n)} = W_{s_2 - s_1}^{(n)}$, $\varphi_n(u) := M_{W_t^{(n)}}(u)$ queda como sigue,

$$\begin{aligned} \varphi_n(u) &= M_{W_{s_1}^{(n)}}(u) M_{W_{s_2 - s_1}^{(n)}} \left(\frac{t - s_1}{s_2 - s_1} u \right) \\ &= \left(\frac{1}{2} e^{\frac{u}{\sqrt{n}}} + \frac{1}{2} e^{-\frac{u}{\sqrt{n}}} \right)^{ns_1} \left(\frac{1}{2} e^{\frac{u(t-s_1)}{(s_2-s_1)\sqrt{n}}} + \frac{1}{2} e^{-\frac{(t-s_1)u}{(s_2-s_1)\sqrt{n}}} \right)^{n(s_2-s_1)} \end{aligned}$$

por lo que aplicando logaritmo a φ_n se cumple lo siguiente,

$$\begin{aligned}\log \varphi_n(u) &= ns_1 \log \left(\frac{1}{2} e^{\frac{u}{\sqrt{n}}} + \frac{1}{2} e^{-\frac{u}{\sqrt{n}}} \right) + n(s_2 - s_1) \left(\frac{1}{2} e^{\frac{u(t-s_1)}{(s_2-s_1)\sqrt{n}}} + \frac{1}{2} e^{-\frac{(t-s_1)u}{(s_2-s_1)\sqrt{n}}} \right) \\ \log \varphi_n(u) &\xrightarrow{n \rightarrow \infty} \frac{1}{2} u^2 s_1 + \frac{1}{2} u^2 \left(\frac{t-s_1}{s_2-s_1} \right)^2 (s_2 - s_1) \\ &= \frac{1}{2} \left(1 - \frac{t-s_1}{s_2-s_1} \right) u^2 s_1 + \frac{1}{2} \left(\frac{t-s_1}{s_2-s_1} \right) u^2 s_2\end{aligned}$$

una consecuencia de que $n \rightarrow \infty$ es que $t \rightarrow s_1$ o $t \rightarrow s_2$ (la distancia entre s_1 y s_2 se reduce cada vez mas) por lo que existen los dos casos siguientes

$$\begin{aligned}M(u) &= \lim_{t \rightarrow s_1} \left(\frac{1}{2} \left(1 - \frac{t-s_1}{s_2-s_1} \right) u^2 s_1 + \frac{1}{2} \left(\frac{t-s_1}{s_2-s_1} \right) u^2 s_2 \right) = \frac{1}{2} u^2 s_1 \\ M(u) &= \lim_{t \rightarrow s_2} \left(\frac{1}{2} \left(1 - \frac{t-s_1}{s_2-s_1} \right) u^2 s_1 + \frac{1}{2} \left(\frac{t-s_1}{s_2-s_1} \right) u^2 s_2 \right) = \frac{1}{2} u^2 s_2\end{aligned}$$

como se deseaba. Con esto se sabe que $\varphi_n(u) \xrightarrow{n \rightarrow \infty} M(u)$ y gracias al teorema 2.30 se comprueba que,

$$W_t^{(n)} \xrightarrow{d} W_t$$

para cualquier $t \in \mathbb{R}$. □

En la figura 2.4 se pueden ver histogramas de las variables aleatorias $W_1^{(70)}$ y $W_1^{(100)}$, respectivamente. Además, dichos histogramas se comparan con una función de densidad perteneciente a un modelo $Normal(0, 1)$. Se puede notar como el ajuste mejora conforme n crece.

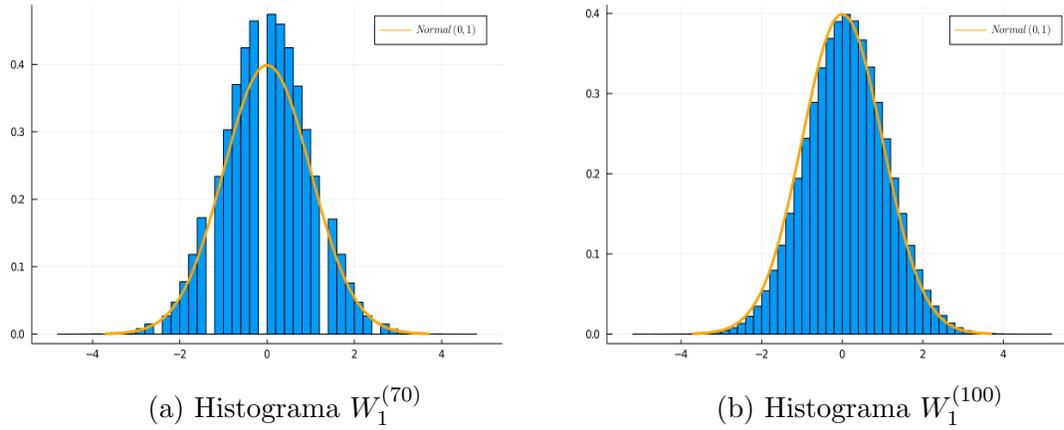


Figura 2.4: Histogramas de la caminata aleatoria escalada simétrica

Definición 2.32. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Se define el movimiento browniano $\mathbf{W} := (W_t)_{t \geq 0}$ como un proceso estocástico que cumple para cualesquiera $0 = t_0 < t_1 < \dots < t_m$ las siguientes propiedades:

1) Los incrementos,

$$W_{t_1} - W_{t_0}, W_{t_2} - W_{t_1}, \dots, W_{t_m} - W_{t_{m-1}}$$

son independientes.

2) La distribución del incremento $W_{t_{i+1}} - W_{t_i}$ es normal con media y varianza siguientes,

$$\begin{aligned} \mathbb{E}(W_{t_{i+1}} - W_{t_i}) &= 0 \\ \mathbb{V}(W_{t_{i+1}} - W_{t_i}) &= t_{i+1} - t_i \end{aligned}$$

Con esta definición se ve como el proceso estocástico $\mathbf{W}^{(n)}$ converge en distribución al movimiento browniano (la demostración se omite pero es consecuencia inmediata de la demostración del teorema 2.31). Además si se toman $0 < s < t$ cualesquiera y se define $\mathbb{F} = \sigma(\mathbf{W})$,

$$\begin{aligned} \mathbb{E}(W_t | \mathcal{F}_s) &= \mathbb{E}(W_t - W_s + W_s | \mathcal{F}_s) \\ &= \mathbb{E}(W_t - W_s | \mathcal{F}_s) + \mathbb{E}(W_s | \mathcal{F}_s) \\ &\stackrel{ind}{=} \mathbb{E}(W_t - W_s) + W_s \\ &= W_s \end{aligned}$$

por lo cual se cumple que \mathbf{W} es una martingala para la filtración \mathcal{F} . Lo cual no era en general para el proceso $\mathbf{W}^{(n)}$ (como consecuencia de la interpolación cuando $tn \notin \mathbb{N}$).

Una vez identificada la convergencia en distribución para el modelo $\mathbf{W}^{(n)}$, es de especial interés conocer si existe un tipo de convergencia similar para el árbol binomial multiperiodo \mathbf{S} cuando aumenta la velocidad de los saltos que da la caminata aleatoria subyacente. Este análisis nuevamente se inicia con una definición,

Definición 2.33. Sea S_0 tal y como se establecio en la definición 2.18 y además sean d_n, u_n valores de salto (indexados a $n \in \mathbb{N}$). Para cualquier valor $t \geq 0$ se define el arbol binomial escalado n veces, como el proceso estocástico $\mathbf{S}^{(n)} = (S_t^{(n)})_{t \geq 0}$ siguiente,

1) Si $tn \in \mathbb{N}$ entonces,

$$S_t^{(n)} := S_0 u_n^{\frac{nt+M_{nt}}{2}} d_n^{\frac{nt-M_{nt}}{2}}$$

2) Si $tn \notin \mathbb{N}$ entonces,

$$S_t^{(n)} := S_s^{(n)} + \frac{S_u^{(n)} - S_s^{(n)}}{u - s}(t - s)$$

donde s y u son los valores mas cercanos a t , por izquierda y derecha respectivamente, tales que $ns, nu \in \mathbb{N}$.

donde \mathbf{M} es una caminata aleatoria con probabilidad de subida p .

En la figura 2.5 se puede ver el comportamiento del proceso estocástico $\mathbf{S}^{(n)}$ para el caso simétrico ($p = \frac{1}{2}$) y de igual forma que con $\mathbf{W}^{(n)}$, se puede encontrar una distribución límite para el proceso $\mathbf{S}^{(n)}$, este resultado se consigue en el siguiente teorema,

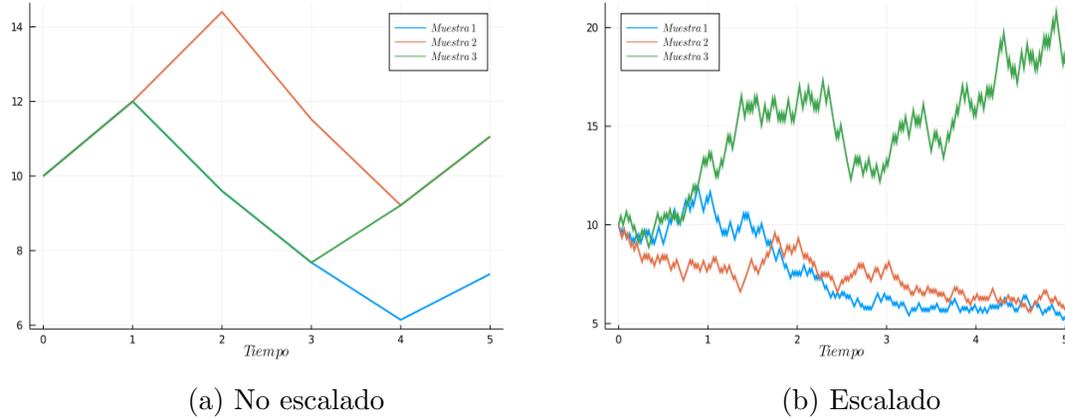


Figura 2.5: Arbol binomial simétrico a 5 periodos

Teorema 2.34. Para cualquier $n \in \mathbb{N}$ sean d_n y u_n los siguientes valores,

$$u_n := \frac{\sigma}{\sqrt{n}}$$

$$d_n := -\frac{\sigma}{\sqrt{n}}$$

se definen los valores \tilde{p}_n y \tilde{q}_n como sigue,

$$p_n := \frac{1 - d_n}{u_n - d_n}$$

$$q_n := \frac{u_n - 1}{u_n - d_n}$$

Además para t fijo, se tiene que si $\mathbf{S}^{(n)}$ representa la incertidumbre del modelo con probabilidad de subida \tilde{p}_n , entonces se cumple que,

$$S_t^{(n)} \xrightarrow{d} S_0 \exp \left\{ \sigma W_t - \frac{1}{2} \sigma^2 t \right\}$$

donde W_t pertenece al movimiento browniano establecido en la definición 2.32.

Demostración. Se cumple lo siguiente para \tilde{p}_n y \tilde{q}_n ,

$$\tilde{p} = \frac{1 - d_n}{u_n - d_n} = \frac{\sigma/\sqrt{n}}{2\sigma/\sqrt{n}} = \frac{1}{2} \tag{2.46}$$

$$\tilde{q} = \frac{u_n - 1}{u_n - d_n} = \frac{\sigma/\sqrt{n}}{2\sigma/\sqrt{n}} = \frac{1}{2} \tag{2.47}$$

ahora bien, para cumplir la demostración, es suficiente mostrar que la distribución de $\log S_t^{(n)}$, dada por la siguiente ecuación,

$$\log S_t^{(n)} = \log S_0 + \frac{1}{2}(nt + M_{nt}) \log \left(1 + \frac{\sigma}{\sqrt{n}} \right) + \frac{1}{2}(nt - M_{nt}) \log \left(1 - \frac{\sigma}{\sqrt{n}} \right) \quad (2.48)$$

converge a la distribución siguiente,

$$\log S_t := \log S_0 + \sigma W_t - \frac{1}{2}\sigma^2 t$$

donde $W_t \sim N(0, t)$ y se supone que $nt \in \mathbb{N}$. Para hacer esto, se establece la expansión de Talyor para $f(x) = \log(x + 1)$. Recuérdese que la expansión de Taylor centrada en cero para una función f es dada por la siguiente ecuación (siempre que $f \in C^\infty$),

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$$

donde $f^{(n)}$ es la n -ésima derivada de f . Por lo que si $f(x) = \log(x + 1)$ es la función en cuestión, entonces la expansión de Taylor centrada en cero queda como sigue,

$$\begin{aligned} f(x) &= f(0) + f'(0)x + \frac{1}{2}f''(0)x^2 + O(x^3) \\ f(x) &= x - \frac{1}{2}x^2 + O(x^3) \end{aligned} \quad (2.49)$$

donde $O(x^3)$ indica términos de orden x^3 o superior. Ahora bien, se utiliza 2.49 para los casos de $x = \frac{\sigma}{\sqrt{n}}$ y $x = -\frac{\sigma}{\sqrt{n}}$ en la ecuación 2.48. Quedando $\log S_t^{(n)}$ expresado de la siguiente forma,

$$\begin{aligned} \log S_t^{(n)} &= \log S_0 + \frac{1}{2}(nt + M_{nt}) \left(\frac{\sigma}{\sqrt{n}} - \frac{\sigma^2}{2n} + O(n^{-\frac{3}{2}}) \right) \\ &\quad + \frac{1}{2}(nt - M_{nt}) \left(-\frac{\sigma}{\sqrt{n}} - \frac{\sigma^2}{2n} + O(n^{-\frac{3}{2}}) \right) \\ &= \log S_0 + nt \left(-\frac{\sigma^2}{2n} + O(n^{-\frac{3}{2}}) \right) + M_{nt} \left(\frac{\sigma}{\sqrt{n}} + O(n^{-\frac{3}{2}}) \right) \\ &= \log S_0 - \frac{1}{2}\sigma^2 t + O(n^{-\frac{1}{2}}) + \sigma W_t^{(n)} + O(n^{-1})W_t^{(n)} \end{aligned}$$

Como consecuencia del teorema 2.31, se cumple que $W_t^{(n)} \xrightarrow{d} W_t$, pero el término $O(n^{-1})W_t^{(n)}$ converge en distribución a la v.a. constante 0 gracias al término $O(n^{-1})$

y esto de igual forma para $O(n^{-\frac{1}{2}})$. Concluyendo que,

$$\log S_t^{(n)} \xrightarrow{d} \log S_0 + \sigma W_t - \frac{1}{2}\sigma^2 t$$

demostrando así lo deseado. En el caso de que $nt \notin \mathbb{N}$, se puede realizar un desarrollo similar al utilizado en la demostración del teorema 2.31. \square

Recuérdese que los valores \tilde{p}_n y \tilde{q}_n mencionados en las ecuaciones 2.46 y 2.47, respectivamente, son los valores establecidos como neutrales al riesgo en el análisis del árbol binomial establecido en la sección 2.1 de este capítulo. Por lo cual se concluye que si $S_t^{(n)}$ cumple las propiedades del teorema 2.34, entonces

$$S_t^{(n)} \xrightarrow{d} S_0 \exp \left\{ \sigma W_t - \frac{1}{2}\sigma^2 t \right\}$$

y esto para el espacio de probabilidad $(\Omega, \mathcal{F}, \tilde{\mathbb{P}})$.

Si se desea extender al caso en el que $\tilde{p}_n \neq \frac{1}{2} \neq \tilde{q}_n$, esto se puede hacer con el siguiente teorema,

Teorema 2.35. *Para cualquier $n \in \mathbb{N}$ sean d_n y u_n los siguientes valores,*

$$\begin{aligned} u_n &:= e^{\sigma/\sqrt{n}} \\ d_n &:= e^{-\sigma/\sqrt{n}} \end{aligned}$$

se definen los valores \tilde{p}_n y \tilde{q}_n como sigue,

$$\begin{aligned} \tilde{p}_n &:= \frac{1 + \frac{r}{n} - d_n}{u_n - d_n} \\ \tilde{q}_n &:= \frac{u_n - 1 - \frac{r}{n}}{u_n - d_n} \end{aligned}$$

Además para t fijo, se tiene que si $\mathbf{S}^{(n)}$ representa la incertidumbre del modelo con probabilidad de subida \tilde{p}_n , entonces se cumple que,

$$S_t^{(n)} \xrightarrow{d} S_0 \exp \left\{ \sigma W_t + \left(r - \frac{1}{2}\sigma^2 \right) t \right\} \quad (2.50)$$

donde W_t pertenece al movimiento browniano establecido en la definición 2.32.

Demostración. Véase el ejercicio 3.8 en Shreve (2004a).

□

La figura 2.6 permite analizar de forma visual el resultado del teorema 2.34 para los casos de $n = 70$ (figura 2.6a) y $n = 90$ (figura 2.6b) con los valores $S_0 = 10$, $\sigma = 0.01$ y $r = 0,0$.

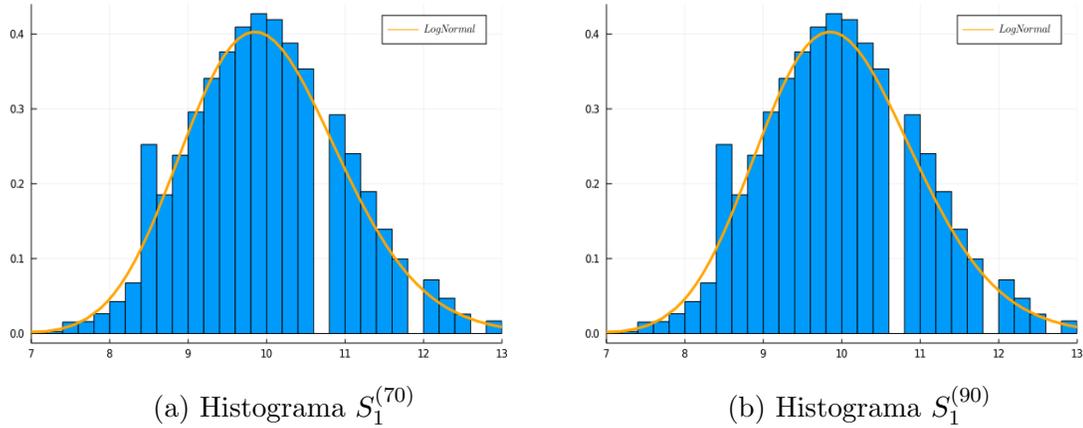


Figura 2.6: Histogramas para $S_1^{(n)}$

Para el caso de $r \neq 0$, en la figura 2.7 se ve como se comporta $S_2^{(n)}$ para los casos $n = 100$ (figura 2.7a) y $n = 500$ (figura 2.7b). En este caso, $S_0 = 10$, $\sigma = 0.1$ y $r = 0.1$.

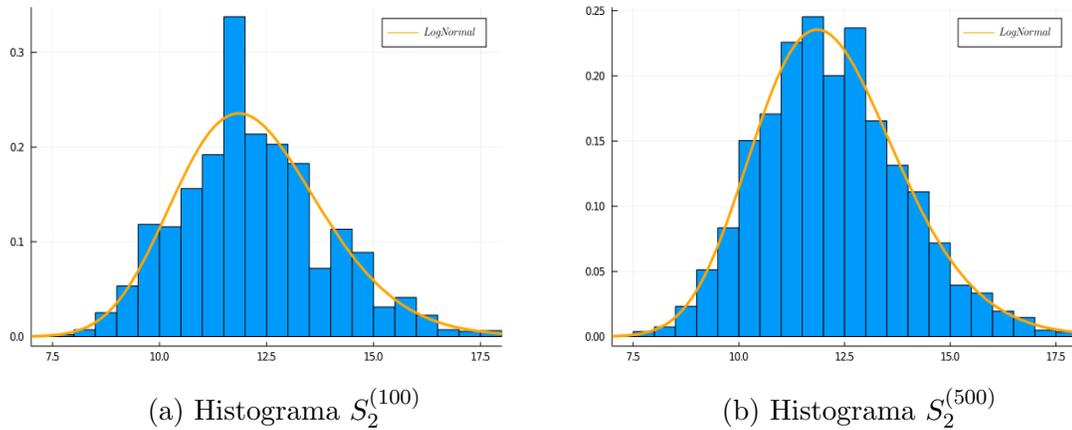


Figura 2.7: Histogramas para $S_2^{(n)}$

2.5. Cálculo estocástico

Utilizando los resultados de la sección anterior, se establece la fórmula para valoración de opciones europeas establecida por Black, Sholes y Merton en 1973. Primero se establece una introducción al Cálculo de Itô¹⁰ basando el desarrollo en Shreve (2004a) y la fórmula de Black-Sholes-Merton se puede encontrar en Black *et. al.* (1973) y Merton (1973).

Se inicia con un conjunto de definiciones básicas del Cálculo de Itô,

Definición 2.36. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y sea $\Delta = (\Delta_t)_{t \geq 0}$ un proceso estocástico. Se dice que Δ es un proceso simple sobre $[0, T] \subset \mathbb{R}$ si existe una partición $\Pi = \{t_0, t_1, \dots, t_n\}$ de $[0, T]$ sobre la cual se cumple que Δ es constante en t para cada subintervalo $[t_j, t_{j+1}) \subset [0, T]$.

Definición 2.37. (Integral de Itô) Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad sobre el cual se definen los siguientes dos procesos estocásticos:

- 1) Un proceso estocástico cualquiera $\mathbf{X} = (X_t)_{t \geq 0}$.
- 2) Un proceso simple $\Delta = (\Delta_t)_{t \geq 0}$ sobre $[0, T]$.

se define el proceso $\mathbf{I} = (I_t)_{0 \leq t \leq T}$ de la siguiente forma,

$$I_t := \sum_{j=0}^{k-1} \Delta_{t_j} (X_{t_{j+1}} - X_{t_j}) + \Delta_{t_k} (X_t - X_{t_k}); \quad t_{k-1} \leq t \leq t_k$$

ambos adaptables a \mathbb{F} y donde $\Pi = \{t_0, t_1, \dots, t_n\}$ es la partición establecida en la definición 2.36. El proceso \mathbf{I} se denota como sigue,

$$I_t \equiv \int_0^t \Delta_s dX_s$$

Además, si \mathbf{X} es un movimiento Browniano, se dice que \mathbf{I} es la Integral de Itô con respecto al proceso simple Δ .

Proposición 2.38. Sea \mathbf{I} una integral de Itô con respecto al proceso simple Δ , entonces se cumplen las siguientes propiedades:

¹⁰Construido por Itô Kiyoshi, matemático japonés; siendo el lema de Itô el resultado más importante de dicha área.

- 1) La integral de Itô es una martingala.
- 2) $\mathbb{E}(I_t) = 0$; $t \geq 0$.
- 3) La varianza de I_t es la siguiente,

$$\mathbb{V}(I_t) = \mathbb{E}\left(\int_0^t \Delta_u^2 du\right)$$

a esta propiedad también se lo conoce como “Isometría de Itô”.

- 4) La integral I_t , como función de t , tiene caminos continuos.
- 5) \mathbf{I} es un proceso adaptable a \mathbb{F} .
- 6) Si $I_t = \int_0^t \Delta_s dW_s$ y $J_t = \int_0^t \Gamma_s dW_s$, entonces $I_t \pm J_t = \int_0^t (\Delta_s \pm \Gamma_s) dW_s$; además para cualquier constante c , $cI_t = \int_0^t c \Delta_s dW_s$

Demostración. Véase Shreve (2004a). □

Observación 2.38.1. Para los objetivos de este texto, la notación,

$$I_t = I_0 + \int_0^t \Delta_s dX_s \quad (2.51)$$

donde I_0 es un valor constante; es equivalente a la siguiente notación,

$$dI_t = \Delta_t dX_t \quad (2.52)$$

se dice que 2.51 es la “forma integral” de 2.52, mientras que 2.52 es la “forma diferencial” de 2.51. La ecuación 2.52 es conocida en el cálculo estocástico como una **ecuación diferencial estocástica**, esto por su asemejo a una ecuación diferencial, y por que se agrega un proceso estocástico como factor de “incertidumbre” o “variabilidad”.

Un sentido práctico a los procesos Δ y \mathbf{X} establecidos en la definición 2.37 se da en el siguiente ejemplo,

Ejemplo 2.39. Sean u, d, S_0 tal y como fueron establecidos en el ejemplo 2.23 y sea $\{\Delta_i\}_{I_{n-1}}$ el proceso estocástico obtenido en el ejemplo 2.24. Además para cualquier $t \notin \{0, 1, 2\}$ se define a S_t como sigue

$$S_t := S_{\lfloor t \rfloor}$$

Con esta información, se calcula la siguiente integral,

$$I_t = \int_0^1 \Delta_s dS_s \quad 0 \leq t \leq 2$$

Para $0 \leq t < 1$,

$$I_t = \Delta_0 S_t = \Delta_0 S_0 = 0.8109 * 10 = 8.109$$

para $1 \leq t < 2$,

$$\begin{aligned} I_t &= \Delta_0 S_1 + \Delta_1 (S_t - S_1) = \Delta_0 S_1 + \Delta_1 (S_1 - S_1) \\ &= \Delta_0 S_1 = \begin{cases} 10.54 & \text{si } S_1 = 13 \\ 6.48 & \text{si } S_1 = 8 \end{cases} \end{aligned}$$

y para $t = 2$,

$$I_2 = \Delta_0 S_1 + \Delta_1 (S_2 - S_1) = \begin{cases} 14.44 & \text{si } S_1 = 13, S_2 = 16.9 \\ 7.94 & \text{si } S_1 = 13, S_2 = 10.4 \\ 7.32 & \text{si } S_1 = 8, S_2 = 10.4 \\ 5.92 & \text{si } S_1 = 8, S_2 = 6.4 \end{cases}$$

además, los puntos de masa de I_t quedaron completamente definidos por la dependencia entre Δ y \mathbf{S} .

Se concluye que la integral I_t puede representar el valor de un portafolio de inversión a través del tiempo, donde:

- 1) $\Pi = \{t_0, t_1, \dots, t_n\} \equiv$ Fechas de trading.
- 2) $\Delta \equiv$ La posición que se toma sobre el activo \mathbf{X} a tiempo t .
- 3) $\mathbf{X} \equiv$ Precio por acción a tiempo t .

La integral I_t se va a utilizar para crear un portafolio de inversión que replique un derivado financiero que dependa de un activo \mathbf{S} y la tasa libre de riesgo r . Para esto, primero se necesita extender el concepto de \mathbf{I} para el caso en el que Δ no es un proceso simple.

Se asume una condición de integrabilidad para el proceso Δ (donde Δ puede ser cualquier tipo de proceso estocástico):

$$\mathbb{E} \left(\int_0^T \Delta_s^2 ds \right) < \infty$$

Ahora bien, para poder definir $I_t = \int_0^t \Delta_s^2 dX_s$, se aproxima Δ mediante procesos simples que lo asemejen, esto se hace tomando una partición $\Pi \subset [0, t]$ donde $t_0 = 0$ y $t_n = t$, con la cual se crea un proceso simple Δ^n sobre $[0, t]$ que aproxima al proceso Δ definiendo Δ_{t_j} para cada t_j y luego se deja el proceso constante para el intervalo $[t_j, t_{j+1})$.

Conforme $n \rightarrow \infty$ (es decir, el número de puntos en la partición incremental) el proceso estocástico simple Δ^n converge al proceso Δ y la convergencia es en el sentido siguiente,

$$\lim_{n \rightarrow \infty} \mathbb{E} \left(\int_0^t |\Delta_s^n - \Delta_s|^2 dt \right) = 0$$

entonces, para este caso se define la integral I_t como sigue,

$$I_t := \int_0^t \Delta_s^2 dX_s = \lim_{n \rightarrow \infty} \left(\int_0^t \Delta_u^n dX_u \right)$$

que también se denomina Integral de Itô en el caso de que \mathbf{X} sea un proceso Browniano.

Ejemplo 2.40. *Obtégase la integral de Itô siguiente:*

$$I_T = \int_0^T W_t dW_t$$

Para esto, se crea el siguiente proceso simple $\Delta^n = (\Delta_t^n)_{0 \leq t < T}$ para $n \in \mathbb{N}$:

$$\Delta_t^n = \begin{cases} W_0 & \text{si } 0 \leq t < \frac{T}{n} \\ W_{\frac{T}{n}} & \text{si } \frac{T}{n} \leq t < \frac{2T}{n} \\ \vdots & \\ W_{\frac{(n-1)T}{n}} & \text{si } \frac{(n-1)T}{n} \leq t < T \end{cases}$$

entonces $\mathbb{E} \left(\int_0^t |\Delta_s^n - W_s|^2 dt \right) \xrightarrow[n \rightarrow \infty]{} 0$. Por definición,

$$I_T = \int_0^T W_t dW_t = \lim_{n \rightarrow \infty} \int_0^T \Delta_t^n dW_t = \lim_{n \rightarrow \infty} \sum_{j=0}^{n-1} W_{jT/n} (W_{(j+1)T/n} - W_{jT/n}) \quad (2.53)$$

para encontrar una equivalencia para 2.53, se desarrolla la ecuación 2.54 simplificando la notación de la siguiente forma $W_j^* := W_{jT/n}$,

$$\begin{aligned}
\frac{1}{2} \sum_{j=0}^{n-1} (W_{j+1}^* - W_j^*)^2 &= \frac{1}{2} \sum_{j=0}^{n-1} W_{j+1}^{*2} - \sum_{j=0}^{n-1} W_j^* W_{j+1}^* + \frac{1}{2} \sum_{j=0}^{n-1} W_j^{*2} \\
&= \frac{1}{2} \sum_{k=1}^n W_k^{*2} - \sum_{j=0}^{n-1} W_j^* W_{j+1}^* + \frac{1}{2} \sum_{j=0}^{n-1} W_j^{*2} \\
&= \frac{1}{2} W_n^{*2} + \frac{1}{2} \sum_{k=0}^{n-1} W_k^{*2} - \sum_{j=0}^{n-1} W_j^* W_{j+1}^* + \frac{1}{2} \sum_{j=0}^{n-1} W_j^{*2} \\
&= \frac{1}{2} W_n^{*2} + \sum_{k=0}^{n-1} W_k^{*2} - \sum_{j=0}^{n-1} W_j^* W_{j+1}^* \\
&= \frac{1}{2} W_n^{*2} + \sum_{j=0}^{n-1} W_j^* (W_j^* - W_{j+1}^*) \tag{2.54}
\end{aligned}$$

de 2.54 se concluye que,

$$\sum_{j=0}^{n-1} W_j^* (W_j^* - W_{j+1}^*) = \frac{1}{2} W_n^{*2} - \frac{1}{2} \sum_{j=0}^{n-1} (W_{j+1}^* - W_j^*)^2$$

y en la notación original,

$$\sum_{j=0}^{n-1} W_{jT/n} (W_{(j+1)T/n} - W_{jT/n}) = \frac{1}{2} W_T^2 - \frac{1}{2} \sum_{j=0}^{n-1} (W_{(j+1)T/n} - W_{jT/n})^2$$

haciendo que $n \rightarrow \infty$ y usando la anterior ecuación, se concluye que,

$$I_T = \int_0^T W_t dW_t = \frac{1}{2} W_T^2 - \frac{1}{2} [W, W]_T = \frac{1}{2} W_T^2 - \frac{1}{2} T$$

donde $[W, W]_T$ es conocida como la variación cuadrática¹¹ del movimiento Browniano W .

Una vez hecho esto, es de interés encontrar una regla equivalente a la “regla de cadena” del Cálculo Diferencial tradicional. Recordando dicha regla, sean f y g

¹¹Para mas información véase Shreve (2004a) sección 3.4.2

dos funciones reales de variable real y sea $c \in \mathbb{R}$ un punto cualquiera tal que g es diferenciable en c y f es diferenciable en $g(c)$, entonces se cumple lo siguiente,

$$(f \circ g)'(c) = f'(g(c))g'(c)$$

ahora bien, es de interés obtener un resultado equivalente en el caso de que g sea un movimiento browniano, y este resultado se plasma a continuación.

Teorema 2.41. (*Fórmula de Itô-Doeblin para un movimiento Browniano*) Sea $f(t, x)$ una función para la cual las derivadas parciales,

$$\begin{aligned} f_t(t, x) &:= \frac{\partial f(t, x)}{\partial t} \\ f_x(t, x) &:= \frac{\partial f(t, x)}{\partial x} \\ f_{xx}(t, x) &:= \frac{\partial^2 f(t, x)}{\partial x^2} \end{aligned}$$

existen, son todas continuas y sea $\mathbf{W} = (W_t)_{t \geq 0}$ un movimiento browniano. Entonces, para cualquier $T \geq 0$,

$$f(T, W_T) = f(0, W_0) + \int_0^T f_t(s, W_s) ds + \int_0^T f_x(s, W_s) dW_s + \frac{1}{2} \int_0^T f_{xx}(s, W_s) ds$$

o escrito en forma diferencial,

$$df(T, W_T) = f_t(T, W_T) + f_x(T, W_T) dW_T + \frac{1}{2} f_{xx}(T, W_T) dT$$

Demostración. Para la demostración, véase Shreve (2004a). □

La fórmula de Itô-Doeblin puede simplificar el cálculo de integrales de Itô. Por ejemplo, sea $f(t, x) = \frac{1}{2}x^2$, entonces $f_t(t, x) = 0$, $f_x(t, x) = x$ y $f_{xx}(t, x) = 1$,

$$\begin{aligned} \frac{1}{2}W_T^2 &= f(T, W_T) - f(0, W_0) \\ &= \int_0^T f_t(s, W_s) ds + \int_0^T f_x(s, W_s) dW_s + \frac{1}{2} \int_0^T f_{xx}(s, W_s) ds \\ &= \int_0^T W_s dW_s + \frac{1}{2}T \end{aligned}$$

por lo que,

$$\int_0^T W_s dW_s = \frac{1}{2}W_T^2 - \frac{1}{2}T$$

tal y como se obtuvo en el ejemplo 2.39. Ahora se extiende el concepto de “regla de la cadena” para procesos mas generales que el movimiento Browniano.

Definición 2.42. Sea $\mathbf{W} = (W_t)_{t \geq 0}$ un movimiento browniano adaptable con respecto a la filtración $\mathbb{F} = (\mathcal{F})_{t \geq 0}$. Un proceso de Itô es un proceso estocástico \mathbf{X} descrito de la siguiente forma,

$$X_t = X_0 + \int_0^t \Delta_s dW_s + \int_0^t \Theta_s ds$$

donde X_0 es una v.a. constante y Δ y Θ son procesos estocásticos adaptables a \mathbb{F} . Escrito de forma diferencial queda como sigue,

$$dX_t = \Delta_t dW_t + \Theta_t dt$$

Definición 2.43. Sea $\mathbf{X} = (X_t)_{t \geq 0}$ un proceso de Itô y sea $\Gamma = (\Gamma)_{t \geq 0}$ un proceso adaptable. Se define la integral de Γ con respecto al proceso \mathbf{X} como el proceso estocástico siguiente,

$$\int_0^T \Gamma_s dX_s := \int_0^T \Gamma_s \Delta_s dW_s + \int_0^T \Gamma_s \Theta_s dX_s$$

y escrito de otro modo,

$$\Gamma_T dX_T = \Gamma_T \Delta_T dW_T + \Gamma_T \Theta_T dX_T$$

Teorema 2.44. (Fórmula de Itô-Doebelin para un proceso de Itô) Sea $\mathbf{X} = (X_t)_{t \geq 0}$ un proceso de Itô y sea $f(t, x)$ una función para la cual las derivadas parciales,

$$\begin{aligned} f_t(t, x) &:= \frac{\partial f(t, x)}{\partial t} \\ f_x(t, x) &:= \frac{\partial f(t, x)}{\partial x} \\ f_{xx}(t, x) &:= \frac{\partial^2 f(t, x)}{\partial x^2} \end{aligned}$$

existen, están definidas y son continuas. Entonces, para cualquier $T \geq 0$,

$$\begin{aligned} f(T, X_T) &= f(0, X_0) + \int_0^T f_t(s, X_s) ds + \int_0^T f_x(s, X_s) \Delta_s dW_s \\ &\quad + \int_0^T f_x(s, X_s) \Theta_s ds + \frac{1}{2} \int_0^T f_{xx}(s, X_s) \Delta_s^2 ds \end{aligned}$$

y escrito en su forma diferencial,

$$\begin{aligned} df(T, X_T) &= f_t(T, X_T) dT + f_x(T, X_T) \Delta_T dW_T \\ &\quad + f_x(T, X_T) \Theta_T dT + \frac{1}{2} f_{xx}(T, X_T) \Delta_T^2 dT \\ &\equiv f_t(T, X_T) dT + f_x(T, X_T) dX_T + \frac{1}{2} f_{xx}(T, X_T) \Delta_T^2 dT \end{aligned}$$

Demostración. Para la demostración, véase Shreve (2004a). □

Ejemplo 2.45. Sea $\mathbf{W} = (W_t)_{t \geq 0}$ un movimiento browniano adaptable a \mathbb{F} , además sean $\boldsymbol{\alpha} = (\alpha_t)_{t \geq 0}$ y $\boldsymbol{\sigma} = (\sigma_t)_{t \geq 0}$ procesos estocásticos adaptables a \mathbb{F} . Se define el proceso de Itô siguiente,

$$X_t = \int_0^t \sigma_s dW_s + \int_0^t \left(\alpha_s - \frac{1}{2} \sigma_s^2 \right) ds$$

entonces,

$$dX_t = \sigma_t dW_t + \left(\alpha_t - \frac{1}{2} \sigma_t^2 \right) dt$$

y bajo la notación del teorema 2.42,

$$\begin{aligned} \Delta_t &:= \sigma_t \\ \Theta_t &:= \alpha_t - \frac{1}{2} \sigma_t^2 \end{aligned}$$

si se define $S_t := f(t, X_t) = S_0 e^{X_t}$ para un valor S_0 constante, entonces se cumple que,

$$S_t = S_0 \exp \left\{ \int_0^t \sigma_s dW_s + \int_0^t \left(\alpha_s - \frac{1}{2} \sigma_s^2 \right) ds \right\} \quad (2.55)$$

lo cual es un proceso estocástico que generaliza el obtenido en la ecuación 2.50 (la ecuación 2.50 es un caso particular en el que α y σ son procesos estocásticos constantes con valores r y σ , respectivamente).

Con esta información, se obtiene la “forma diferencial” de 2.55 utilizando el resultado del teorema 2.44. Se tiene que $f(t, x) = S_0 e^x$ entonces $f_t(t, x) = 0$, $f_x(t, x) = S_0 e^x = f_{xx}(t, x)$ y de acuerdo al teorema 2.44,

$$\begin{aligned} dS_t &= df(t, X_t) = f_t(t, X_t) dt + f_x(t, X_t) \Delta_t dW_t \\ &\quad + f_x(t, X_t) \Theta_t dt + \frac{1}{2} f_{xx}(t, X_t) \Delta_t^2 dt \\ &= S_t \left(\alpha_t - \frac{1}{2} \sigma_t^2 \right) dt + \frac{1}{2} S_t \sigma_t^2 dt + S_t \sigma_t dW_t \\ &= S_t \alpha_t dt + S_t \sigma_t dW_t \end{aligned}$$

o lo que es equivalente en “forma integral”,

$$S_t = S_0 + \int_0^t S_s \alpha_s ds + \int_0^t S_s \sigma_s dW_s$$

Con estos resultados se está listo para establecer el resultado más importante de esta sección, la denominada *fórmula de Black-Sholes-Merton* para la valuación de opciones europeas.

Teorema 2.46. (*Valuación de Black-Sholes-Merton*) Sea $\mathbf{S} = (S_t)_{t \geq 0}$ un activo financiero que cumple la siguiente ecuación diferencial estocástica,

$$dS_t = \alpha S_t dt + \sigma S_t dW_t \quad (2.56)$$

y sea r la tasa libre de riesgo. Además, se supone una opción europea tipo Call que tiene como subyacente al activo \mathbf{S} , si $\mathbf{C} = (C_t)_{0 \leq t \leq T}$ representa a dicha opción donde T es la madurez del derivado, entonces se cumple lo siguiente para \mathbf{C} ,

$$C_t = S_t \Phi(d_+(T-t, S_t)) - K e^{-r(T-t)} \Phi(d_-(T-t, S_t))$$

donde Φ es la función de distribución de una v.a. Normal(0,1) y,

$$d_{\pm}(\tau, x) := \frac{1}{\sigma \sqrt{\tau}} \left[\log \frac{x}{K} + \left(r \pm \frac{\sigma^2}{2} \right) \tau \right]$$

Demostración. Para esta prueba, se sigue un procedimiento equivalente al realizado en el caso discreto, se crea un portafolio de inversión que replique el derivado en cuestión. Para $0 \leq t \leq T$, el portafolio replicante $\mathbf{X} = (X_t)_{0 \leq t \leq T}$, va a estar compuesto por dos inversiones distintas:

- 1) Una parte se invertirá en el activo S_t y la posición que se tome sobre dicha inversión será denotada como Δ_t .
- 2) El valor restante de X_t va a estar invertido en una tasa libre de riesgo r , esta posición es la cantidad $X_t - \Delta_t S_t$.

Por lo cual la ecuación diferencial estocástica dX_t que representa la “ganancia instantánea” del portafolio es la siguiente,

$$dX_t = \Delta_t dS_t + r(X_t - \Delta_t S_t)dt$$

y como consecuencia de la ecuación 2.56,

$$\begin{aligned} dX_t &= \Delta_t (\alpha S_t dt + \sigma S_t dW_t) + r(X_t - \Delta_t S_t)dt \\ &= rX_t dt + (\alpha - r)\Delta_t S_t dt + \sigma \Delta_t S_t dW_t \end{aligned}$$

Ahora bien, se considera el valor descontado del activo \mathbf{S} y el portafolio \mathbf{X} a tiempo cero, es decir, se calcula la ecuación diferencial estocástica para $e^{-rt} S_t$ y $e^{-rt} X_t$. Utilizando la fórmula de Itô-Doeblin y tomando $f(t, x) = e^{-rt}x$,

$$\begin{aligned} d(e^{-rt} S_t) &= d(f(t, S_t)) \\ &= f_t(t, S_t) dt + f_x(t, S_t) dS_t + \frac{1}{2} f_{xx}(t, S_t) dS_t dS_t \\ &= -re^{-rt} S_t dt + e^{-rt} dS_t \\ &= (\alpha - r)e^{-rt} S_t dt + \sigma e^{-rt} S_t dW_t \end{aligned}$$

y para el caso de $e^{-rt} X_t$,

$$\begin{aligned} d(e^{-rt} X_t) &= d(f(t, X_t)) \\ &= f_t(t, X_t) dt + f_x(t, X_t) dX_t + \frac{1}{2} f_{xx}(t, X_t) dX_t dX_t \\ &= -re^{-rt} X_t dt + e^{-rt} dX_t \\ &= -re^{-rt} X_t dt + e^{-rt} \Delta_t dS_t + e^{-rt} \Delta_t (\alpha S_t dt + \sigma S_t dW_t) \\ &= \Delta_t (\alpha - r)e^{-rt} S_t dt + \Delta_t \sigma e^{-rt} S_t dW_t \\ &= \Delta_t d(f(t, S_t)) \end{aligned}$$

Ahora bien, una opción Call (denotada como $\mathbf{C} = (C_t)_{0 \leq t \leq T}$) tiene un *payoff* G_C definido como $G_C(S_T) = \max\{S_T - K\}$ a tiempo T , donde K es denominado el *strike* y es un valor constante que se estipula a tiempo $t = 0$. Se desea conseguir una función $c(t, x)$ que represente el valor de la opción call a tiempo t si la acción tiene precio $S_t = x$, ya que X_t es el portafolio replicante de la opción, entonces se debe cumplir la condición siguiente:

$$c(t, S_t) = X_t \quad 0 \leq t \leq T \quad (2.57)$$

la ecuación 2.57 es equivalente a las siguientes dos condiciones,

- 1) $d(f(t, X_t)) = d(f(t, c(t, S_t))) \quad 0 < t \leq T$
- 2) $X_0 = c(0, S_0)$.

ya que integrando (1),

$$f(t, X_t) - X_0 = f(t, c(t, S_t)) - c(0, S_0)$$

y si $X_0 = c(0, S_0)$ entonces se cumple la ecuación 2.57. Para cumplir (1), se obtiene $d(c(t, S_t))$,

$$\begin{aligned} d(c(t, S_t)) &= c_t(t, S_t) dt + c_x(t, S_t) dS_t + \frac{1}{2} c_{xx}(t, S_t) dS_t dS_t \\ &= c_t(t, S_t) dt + c_x(t, S_t) (\alpha S_t dt + \sigma S_t dW_t) + \frac{1}{2} c_{xx}(t, S_t) \sigma^2 S_t^2 dt \\ &= \left(c_t(t, S_t) + \alpha S_t c_x(t, S_t) + \frac{1}{2} \sigma^2 S_t^2 c_{xx}(t, S_t) \right) dt \\ &\quad + \sigma S_t c_x(t, S_t) dW_t \end{aligned}$$

y $d(e^{-rt} c(t, S_t))$,

$$\begin{aligned} d(e^{-rt} c(t, S_t)) &= d(f(t, c(t, S_t))) \\ &= f_t(t, c(t, S_t)) dt + f_x(t, c(t, S_t)) dc(t, S_t) \\ &\quad + \frac{1}{2} f_{xx}(t, c(t, S_t)) dc(t, S_t) dc(t, S_t) \\ &= -r e^{-rt} c(t, S_t) dt + e^{-rt} dc(t, S_t) \\ &= e^{-rt} \left(-rc(t, S_t) + c_t(t, S_t) + \alpha S_t c_x(t, S_t) \right. \\ &\quad \left. + \frac{1}{2} \sigma^2 S_t^2 c_{xx}(t, S_t) \right) dt + e^{-rt} \sigma S_t c_x(t, S_t) dW_t \end{aligned}$$

se debe cumplir que,

$$\begin{aligned} & \Delta_t (\alpha - r) e^{-rt} S_t dt + \Delta_t \sigma e^{-rt} S_t dW_t \\ &= e^{-rt} \left(-rc(t, S_t) + c_t(t, S_t) + \alpha S_t c_x(t, S_t) + \frac{1}{2} \sigma^2 S_t^2 c_{xx}(t, S_t) \right) dt \\ &+ e^{-rt} \sigma S_t c_x(t, S_t) dW_t \end{aligned} \quad (2.58)$$

se igualan los términos en 2.58 que tienen dW_t y se concluye que,

$$\Delta_t = c_x(t, S_t) \quad 0 < t \leq T \quad (2.59)$$

a la ecuación 2.59 se le denomina *delta-hedging*. Esta ecuación permite saber a tiempo t la cantidad de acciones que se necesitan para hacer una cobertura perfecta para el derivado \mathbf{C} , usando 2.59 en 2.58, se obtiene lo siguiente para los términos con dt ,

$$(a - r) S_t c_x(t, S_t) = -rc(t, S_t) + c_t(t, S_t) + \alpha S_t c_x(t, S_t) + \frac{1}{2} \sigma^2 S_t^2 c_{xx}(t, S_t)$$

entonces,

$$rc(t, S_t) = c_t(t, S_t) + r S_t c_x(t, S_t) + \frac{1}{2} \sigma^2 S_t^2 c_{xx}(t, S_t)$$

es decir, la función $c(t, x)$ debe cumplir la siguiente condición,

$$c_t(t, x) + r x c_x(t, x) + \frac{1}{2} \sigma^2 x^2 c_{xx}(t, x) = rc(t, x) \quad t \in [0, T] \quad (2.60)$$

a la ecuación 2.60, se le conoce como la ecuación diferencial parcial de Black-Scholes-Merton. Dicha ecuación debe satisfacer la condición terminal,

$$c(T, x) = (x - K)^+$$

la función que soluciona dicha ecuación diferencial parcial esta descrita en Shreve (2004a), *et. al.* (1973) y Merton (1973). Dicha función esta dada por la siguiente fórmula,

$$c(t, x) = x\Phi(d_+(T - t, x)) - Ke^{-r(T-t)}\Phi(d_-(T - t, x)) \quad 0 \leq t < T$$

donde,

$$d_{\pm}(\tau, x) = \frac{1}{\sigma\sqrt{\tau}} \left[\log \frac{x}{K} + \left(r \pm \frac{\sigma^2}{2} \right) \tau \right]$$

y Φ es la función de distribución asociada a una $N(0, 1)$. Para terminar, se encuentra la función $c_x(t, x)$ que permite construir el portafolio \mathbf{X} ,

$$c_x(t, x) = \Phi(d_+(T-t, x)) + x\phi(d_+(T-t, x))d'_+(T-t, x) \\ - Ke^{-r(T-t)}\phi(d_-(T-t, x))d'_-(T-t, x)$$

donde

$$d'_\pm(\tau, x) = \frac{1}{x\sigma\sqrt{\tau}}$$

y ϕ es la función de densidad asociada a una $N(0, 1)$. Por último, se puede comprobar¹² que,

$$\lim_{t \rightarrow T^-} c(t, x) = (x - K)^+ \\ \lim_{x \rightarrow 0^+} c(t, x) = 0$$

permitiendo tener una definición completa para $c(t, x)$ y por lo tanto se cumple que el derivado \mathbf{C} queda definido de la siguiente forma para todo $0 \leq t \leq T$,

$$C_t := c(t, S_t)$$

□

Corolario 2.46.1. *Como consecuencia del teorema 2.46, si $\mathbf{P} = (P_t)_{0 \leq t \leq T}$ representa el precio de un Put a tiempo t que tiene como subyacente al activo \mathbf{S} , entonces se cumple lo siguiente para \mathbf{P} ,*

$$P_t = Ke^{-r(T-t)}\Phi(-d_-(T-t, S_t)) - S_t\Phi(-d_+(T-t, S_t))$$

donde Φ es la función de distribución de una v.a. Normal(0,1) y,

$$d_\pm(\tau, x) := \frac{1}{\sigma\sqrt{\tau}} \left[\log \frac{x}{K} + \left(r \pm \frac{\sigma^2}{2} \right) \tau \right]$$

Demostración. Como consecuencia del principio de arbitraje se tiene la siguiente relación entre el subyacente \mathbf{S} , el Call europeo \mathbf{C} y el Put europeo \mathbf{P} ,

$$C_t - P_t = S_t - Ke^{-r(T-t)}$$

¹²Véase Shreve (2004a) para esa afirmación

por lo tanto el precio de \mathbf{P} se puede saber como consecuencia del precio de \mathbf{C} ,

$$P_t = C_t - S_t + K e^{-r(T-t)} = K e^{-r(T-t)} \Phi(-d_-(T-t, S_t)) - S_t \Phi(-d_+(T-t, S_t))$$

□

Corolario 2.46.2. *De acuerdo a la del teorema 2.46 y el corolario 2.46.1, las deltas de una opción europea tipo call y put, denotadas como ΔC_t y ΔP_t , respectivamente, están dadas por las siguientes ecuaciones,*

$$\begin{aligned} \Delta C_t(T-t, S_t) &= \left. \frac{\partial c(t, x)}{\partial x} \right|_{x=S_t} = \Phi(d_+(T-t, S_t)) \\ \Delta P_t(T-t, S_t) &= \left. \frac{\partial p(t, x)}{\partial x} \right|_{x=S_t} = -\Phi(-d_+(T-t, S_t)) \end{aligned}$$

donde $c(t, x)$ y $p(t, x)$ representan las funciones de precio para las opciones call y put, respectivamente.

Demostración. Se parte de la delta encontrada para la opción tipo call en el teorema 2.46,

$$\begin{aligned} \Delta C_t = c_x(t, S_t) &= \Phi(d_+(T-t, S_t)) + S_t \phi(d_+(T-t, S_t)) d'_+(T-t, S_t) \\ &\quad - K e^{-r(T-t)} \phi(d_-(T-t, S_t)) d'_-(T-t, S_t) \end{aligned}$$

Se hacen notar dos cosas,

- 1) Las derivadas parciales de $\frac{\partial d_+}{\partial x}$ y $\frac{\partial d_-}{\partial x}$ son iguales.
- 2) Se cumple que $K e^{-r(T-t)} \phi(d_-(T-t, S_t)) = x \phi(d_+(T-t, S_t))$.

El inciso 1 es evidente como consecuencia de la definición de los términos d_+ y d_- . Para el caso de 2), se necesita desarrollar las ecuaciones, esto se hace a continuación,

$$\begin{aligned} K e^{-r(T-t)} \phi(d_-) &= K e^{-r(T-t)} \phi(d_+) \exp \left\{ -\frac{1}{2} \left[\frac{1}{\sigma^2(T-t)} \left(-2(T-t)\sigma^2 \log \frac{S_t}{K} \right. \right. \right. \\ &\quad \left. \left. \left. - 2(T-t)^2 r \sigma^2 \right) \right] \right\} \\ &= K e^{-r(T-t)} \phi(d_+) \exp \left\{ \log \frac{S_t}{K} \right\} e^{r(T-t)} = x \phi(d_+(T-t, x)) \end{aligned}$$

por lo que para el caso de ΔC_t ,

$$\begin{aligned}\Delta C_t &= \Phi(d_+(T-t, S_t)) + S_t \phi(d_+(T-t, S_t)) d'_+(T-t, S_t) \\ &\quad - K e^{-r(T-t)} \phi(d_-(T-t, S_t)) d'_-(T-t, S_t) \\ &= \Phi(d_+(T-t, S_t)) + S_t \phi(d_+(T-t, S_t)) d'_+(T-t, S_t) \\ &\quad - K e^{-r(T-t)} \phi(d_-(T-t, S_t)) d'_-(T-t, S_t) = \Phi(d_+(T-t, S_t))\end{aligned}$$

esto como consecuencia del inciso 1). Para el caso de ΔC_t se utiliza la paridad put-call establecida en el corolario 2.46.1,

$$\begin{aligned}\Delta P_t &= \frac{\partial}{\partial S_t} (C_t - S_t + K) = \Delta C_t - S_t = \Phi(d_+(T-t, S_t)) - 1 \\ &= -\Phi(-d_+(T-t, S_t))\end{aligned}$$

quedando demostrado lo deseado. □

Ejemplo 2.47. Para los parámetros $S_0 = 10$, $\mu = 0.001$, $\sigma = 0.01$, $r = 0.0005$, $T = 20$ y $K = S_0 e^{0.014}$. En las figuras 2.9 se tienen la simulación de 4 caminos aleatorios para los procesos estocásticos \mathbf{S} , \mathbf{C} y \mathbf{P} .

Se puede notar como $C_t, P_t \geq 0$ para cualquier $t \in \{0 \leq t \leq T\}$, esto como consecuencia del principio de arbitraje y que $C_T, P_T \geq 0$.

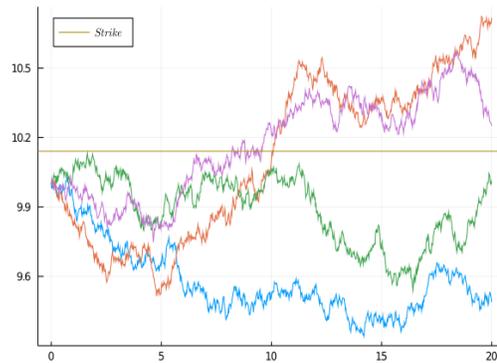
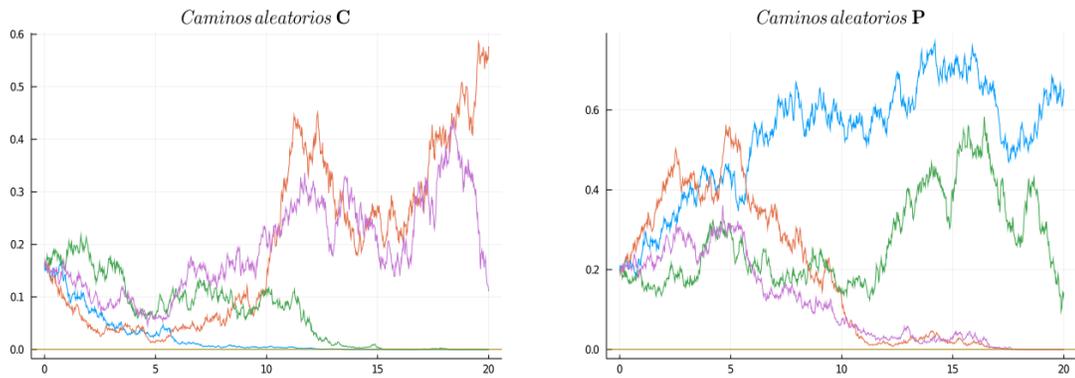
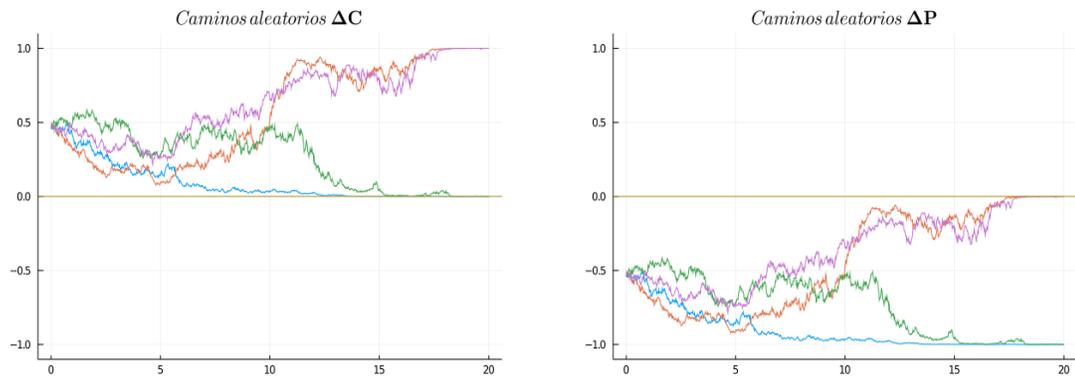


Figura 2.8: Procesos Estocástico de \mathbf{S} .

Figura 2.9: Procesos estocásticos de \mathbf{C} y \mathbf{P} .

Con los instrumentos de este mismo ejemplo, se grafican las deltas de las opciones en la figura 2.10, se comprueba como la delta de un call se encuentra siempre en el intervalo $(0,1)$ y la delta de un put en el intervalo $(-1,0)$. Además, la delta de un call crece conforme el valor del subyacente incrementa, y en el caso del put la delta aumenta conforme el valor del subyacente decrece.

Figura 2.10: Procesos estocásticos de $\Delta\mathbf{C}$ y $\Delta\mathbf{P}$.

2.6. Sobre una extensión al Cálculo de Itô.

En la sección anterior se estableció que un activo financiero está descrito mediante el proceso estocástico $\mathbf{S} = (S_t)_{t \geq 0}$ de la siguiente forma bajo la medida neutral

al riesgo,

$$S_t = S_0 \exp \left\{ \left(r - \frac{1}{2} \sigma^2 \right) t + \sigma W_t \right\} \quad (2.61)$$

donde S_0 es un valor constante que representa el valor presente del activo financiero, σ^2 es la varianza de los log-rendimientos a $t = 1$ de \mathbf{S} y $\mathbf{W} = (W_t)_{t \geq 0}$ es un movimiento browniano¹³.

Se tiene que el proceso estocástico \mathbf{S} sigue la siguiente ecuación diferencial como consecuencia del teorema 2.44,

$$dS_t = r S_t dt + \sigma S_t dW_t$$

el término $-\frac{1}{2}\sigma^2 t$ incluido en la ecuación 2.61 es consecuencia de la variación cuadrática¹⁴ del movimiento browniano \mathbf{W} .

Una generalización de interés para el cálculo de Itô sería modificar la suposición de normalidad en el proceso \mathbf{W} , esto se hace en la siguiente definición y se analizan las consecuencias de esto en los siguientes resultados.

Definición 2.48. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y Φ una distribución de probabilidades continua definida en todos los reales, con esperanza igual a cero y segundo momento finito¹⁵, se define el movimiento pseudo-browniano $\mathbf{W}_\Phi := (W_t)_{t \geq 0}$ como un proceso estocástico que cumple para cualesquiera $0 = t_0 < t_1 < \dots < t_m$ las siguientes propiedades:

1) Los incrementos,

$$W_{t_1} - W_{t_0}, W_{t_2} - W_{t_1}, \dots, W_{t_3} - W_{t_2}, W_{t_m} - W_{t_{m-1}}$$

son independientes.

2) La distribución del incremento $W_{t_{i+1}} - W_{t_i}$ es igual a $\sqrt{(t_{i+1} - t_i)} X$ donde $X \sim \Phi$.

A la varianza de X se le denota como σ_Φ^2 y además $W_0 = 0$.

¹³Tradicionalmente, el tiempo $t = 1$ denota movimientos diarios, pero esto no es una condición necesaria, depende de que temporalidad sea medida por t (días, semanas, meses, etc.).

¹⁴Dicho término se define formalmente en 2.49

¹⁵Para ser técnicos, la esperanza y el segundo momento hacen referencia a la v.a. X tal que $F_X = \Phi$.

La definición anterior permite generalizar ampliamente la distribuciones del proceso estocástico \mathbf{W}_Φ , haciendo posible la introducción de distribuciones con colas pesadas, no simétricas, etc. Además, es evidente que un movimiento browniano es un movimiento pseudo-browniano en el cual Φ es la distribución normal estándar y en este caso particular $\sigma_\Phi^2 = 1$.

Se define a continuación la variación de orden n para un movimiento pseudo-browniano cualquiera además de resultados de interés para dicha definición,

Definición 2.49. Sea $\mathbf{W}_\Phi := (W_t)_{t \geq 0}$ un movimiento pseudo-browniano y sea $n \in \mathbb{N}$. La variación de orden n de \mathbf{W}_Φ hasta el tiempo T queda definida como sigue,

$$\lim_{\|\Pi\| \rightarrow 0} \sum_{j=0}^{n-1} |W_{t_{j+1}} - W_{t_j}|^n$$

donde $\Pi = \{t_0, t_1, \dots, t_n\}$ y $0 = t_0 < t_1 < \dots < t_n = T$. A la variación de orden 2 se le denomina variación cuadrática y se denota especialmente como $[W, W]_T$. En su forma diferencial, la variación cuadrática tiene dos notaciones equivalentes,

$$d[W, W]_T \equiv dW_T dW_T$$

Teorema 2.50. Sea $\mathbf{W}_\Phi := (W_t)_{t \geq 0}$ un movimiento pseudo-browniano, entonces se cumple las siguientes propiedades:

1) La variación cuadrática $[W, W]_T$ es finita y es igual a lo siguiente,

$$[W, W]_T = \lim_{\|\Pi\| \rightarrow 0} \sum_{j=0}^{n-1} |W_{t_{j+1}} - W_{t_j}|^2 = T \sigma_\Phi^2$$

2) La varianción de orden 1 es infinita, es decir,

$$\sum_{j=0}^{n-1} |W_{t_{j+1}} - W_{t_j}| \xrightarrow{\|\Pi\| \rightarrow 0} \infty$$

3) La varianción de orden $n \geq 3$ es cero, es decir,

$$\sum_{j=0}^{n-1} |W_{t_{j+1}} - W_{t_j}|^n \xrightarrow{\|\Pi\| \rightarrow 0} 0$$

Demostración. Sea $\Pi = \{t_0, t_1, \dots, t_n\}$ una partición del intervalo $[0, T]$ tal que $t_{i+1} - t_i = \frac{T}{n}$ para cualquier i , entonces se tiene que $W_{t_{i+1}} - W_{t_i} = \frac{T}{\sqrt{n}}X$, por lo que

$$[W, W]_t = \lim_{n \rightarrow \infty} \sum_{j=0}^{n-1} (W_{t_{j+1}} - W_{t_j})^2 = T \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} X^2 = T \mathbb{E}(X^2) = T \sigma_{\Phi}^2$$

quedando demostrado 1). Ahora bien, si $\Pi = \{t_0, t_1, \dots, t_n\}$ es una partición cualquiera, se cumple lo siguiente,

$$\sum_{j=0}^{n-1} (W_{t_{j+1}} - W_{t_j})^2 \leq \max_{0 \leq k \leq n-1} |W_{t_{k+1}} - W_{t_k}| \sum_{j=0}^{n-1} |W_{t_{j+1}} - W_{t_j}| \quad (2.62)$$

como consecuencia de que \mathbf{W}_{Φ} es un proceso continuo definido para todo $t \in [0, T]$, se tiene que $\max_{0 \leq k \leq n-1} |W_{t_{k+1}} - W_{t_k}| \xrightarrow{\|\Pi\| \rightarrow 0} 0$, pero por el punto 1) se tiene que $[W, W]_T$ es un valor finito mayor a cero, por lo que se tiene que cumplir lo siguiente,

$$\sum_{j=0}^{n-1} |W_{t_{j+1}} - W_{t_j}|^n \xrightarrow{\|\Pi\| \rightarrow 0} \infty$$

cumpléndose así 2). Para el caso de 3) nótese lo siguiente para $n = 3$,

$$\sum_{j=0}^{n-1} |W_{t_{j+1}} - W_{t_j}|^3 \leq \max_{0 \leq k \leq n-1} |W_{t_{k+1}} - W_{t_k}| \sum_{j=0}^{n-1} |W_{t_{j+1}} - W_{t_j}|^2 \quad (2.63)$$

como consecuencia de que $[W, W]_T$ es un valor finito y que $\max_{0 \leq k \leq n-1} |W_{t_{k+1}} - W_{t_k}| \xrightarrow{\|\Pi\| \rightarrow 0} 0$ se cumple que,

$$\sum_{j=0}^{n-1} |W_{t_{j+1}} - W_{t_j}|^3 \xrightarrow{\|\Pi\| \rightarrow 0} 0$$

para los casos $n > 3$ se realiza un proceso inductivo para llegar a la conclusión deseada. \square

Corolario 2.50.1. Sea $\mathbf{W}_{\Phi} := (W_t)_{t \geq 0}$ un movimiento pseudo-browniano, entonces se cumple las siguientes propiedades:

1) La variación de W_t con t es cero, es decir,

$$\sum_{j=0}^{n-1} (W_{t_{j+1}} - W_{t_j})(t_{j+1} - t_j) \xrightarrow{\|\Pi\| \rightarrow 0} 0$$

2) La variación de t consigo mismo es cero, es decir,

$$\sum_{j=0}^{n-1} (t_{j+1} - t_j)^2 \xrightarrow{\|\Pi\| \rightarrow 0} 0$$

3) En general, para $n \geq 1$ y $m \geq 1$ se cumple lo siguiente,

$$\sum_{j=0}^{n-1} (W_{t_{j+1}} - W_{t_j})^n (t_{j+1} - t_j)^m \xrightarrow{\|\Pi\| \rightarrow 0} 0$$

Las propiedades 1) y 2) se escriben de forma diferencial como sigue,

$$dW_t dt = 0 = dt dW_t \quad y \quad dt dt = 0$$

Demostración. Para 1), se tiene lo siguiente,

$$\begin{aligned} \left| \sum_{j=0}^{n-1} (W_{t_{j+1}} - W_{t_j})(t_{j+1} - t_j) \right| &\leq \max_{0 \leq k \leq n-1} |W_{t_{k+1}} - W_{t_k}| \sum_{j=0}^{n-1} (t_{j+1} - t_j) \\ &= \max_{0 \leq k \leq n-1} |W_{t_{k+1}} - W_{t_k}| T \end{aligned}$$

Como consecuencia de que \mathbf{W}_Φ es un proceso estocástico continuo se tiene que $\max_{0 \leq k \leq n-1} |W_{t_{k+1}} - W_{t_k}| \xrightarrow{\|\Pi\| \rightarrow 0} 0$, cumpliéndose así 1). Para 2), se tiene lo siguiente,

$$\sum_{j=0}^{n-1} (t_{j+1} - t_j)^2 \leq \max_{0 \leq k \leq n-1} (t_{k+1} - t_k) \sum_{j=0}^{n-1} (t_{j+1} - t_j) = \|\Pi\| T$$

el cual converge a cero conforme $\|\Pi\| \rightarrow 0$. El caso 3) es una consecuencia del teorema 2.50 en conjunto a los puntos 1) y 2) de este corolario.

□

Los resultados anteriores son suficientes para establecer la fórmula de Itô-Doeblin para un movimiento pseudo-Browniano cualquiera,

Teorema 2.51. (Fórmula de Itô-Doeblin para un movimiento pseudo-Browniano) Sea $f(t, x)$ una función de orden C^∞ en ambas entradas y sea $\mathbf{W}_\Phi = (W_t)_{t \geq 0}$ un movimiento pseudo-browniano. Entonces, para cualquier $T \geq 0$,

$$f(T, W_T) = f(0, W_0) + \int_0^T f_t(s, W_s) ds + \int_0^T f_x(s, W_s) dW_s + \frac{\sigma_\Phi^2}{2} \int_0^T f_{xx}(s, W_s) ds$$

o escrito en forma diferencial,

$$df(T, W_T) = f_t(T, W_T) + f_x(T, W_T) dW_T + \frac{\sigma_\Phi^2}{2} f_{xx}(T, W_T) dT$$

donde f_t, f_x, f_{xx}, f_{tx} y f_{tt} están definidas como sigue,

$$\begin{aligned} f_t(t, x) &:= \frac{\partial f(t, x)}{\partial t} \\ f_x(t, x) &:= \frac{\partial f(t, x)}{\partial x} \\ f_{xx}(t, x) &:= \frac{\partial^2 f(t, x)}{\partial x^2} \\ f_{tx}(t, x) &:= \frac{\partial^2 f(t, x)}{\partial t \partial x} \\ f_{tt}(t, x) &:= \frac{\partial^2 f(t, x)}{\partial t^2} \end{aligned}$$

Demostración. Sea $\Pi = \{t_0, t_1, \dots, t_n\}$ una partición cualquiera, como consecuencia de la expansión de Taylor para funciones de dos variables se cumple lo siguiente,

$$\begin{aligned} f(t_{k+1}, x_{k+1}) - f(t_k, x_k) &= f_t(t_k, x_k)(t_{k+1} - t_k) + f_x(t_k, x_k)(x_{k+1} - x_k) \\ &+ \sum_{i+j \geq 2} \frac{1}{i!j!} \frac{\partial^{(i+j)} f(t_k, x_k)}{\partial t^{(i)} \partial x^{(j)}} (t_{k+1} - t_k)^i (x_{k+1} - x_k)^j \\ &= f_t(t_k, x_k)(t_{k+1} - t_k) + f_x(t_k, x_k)(x_{k+1} - x_k) + \frac{1}{2} f_{tt}(t_k, x_k)(t_{k+1} - t_k)^2 \\ &+ \frac{1}{2} f_{tx}(t_k, x_k)(t_{k+1} - t_k)(x_{k+1} - x_k) + \frac{1}{2} f_{xx}(t_k, x_k)(x_{k+1} - x_k)^2 \\ &+ R(t_k, t_{k+1}, x_k, x_{k+1}) \end{aligned}$$

donde $R(t_k, t_{k+1}, x_k, x_{k+1})$ representa la suma de todas las derivadas parciales mixtas de orden mayor a 2. Sumando todos los intervalos se tiene lo siguiente,

$$\begin{aligned}
f(T, W_T) - f(0, W_0) &= \sum_{k=0}^{n-1} f(t_{k+1}, W_{k+1}) - f(t_k, W_k) \\
&= \sum_{k=0}^{n-1} \left(f_t(t_k, W_k)(t_{k+1} - t_k) + f_x(t_k, W_k)(W_{k+1} - W_k) + \frac{1}{2} f_{xx}(t_k, W_k)(W_{k+1} - W_k)^2 \right) \\
&+ \sum_{k=0}^{n-1} \left(\frac{1}{2} f_{tt}(t_k, W_k)(t_{k+1} - t_k)^2 + \frac{1}{2} f_{tx}(t_k, W_k)(t_{k+1} - t_k)(W_{k+1} - W_k) \right) \\
&+ \sum_{k=0}^{n-1} R(t_k, t_{k+1}, W_k, W_{k+1}) \tag{2.64}
\end{aligned}$$

Para la ecuación anterior se toma el límite conforme $\|\Pi\| \rightarrow 0$, la primera parte de la ecuación 2.64 queda inalterada, además,

$$\begin{aligned}
\sum_{k=0}^{n-1} f_{tx}(t_k, W_k)(t_{k+1} - t_k)(W_{k+1} - W_k) &\xrightarrow{\|\Pi\| \rightarrow 0} \int_0^T f_{tx}(s, W_s) dW_s ds = 0 \\
\sum_{k=0}^{n-1} f_{tt}(t_k, W_k)(t_{k+1} - t_k)^2 &\xrightarrow{\|\Pi\| \rightarrow 0} \int_0^T f_{tt}(s, W_s) ds ds = 0 \\
\sum_{k=0}^{n-1} R(t_k, t_{k+1}, W_k, W_{k+1}) &\xrightarrow{\|\Pi\| \rightarrow 0} 0
\end{aligned}$$

como consecuencia del teorema 2.50 y el corolario 2.50.1, por lo tanto,

$$\begin{aligned}
f(T, W_T) - f(0, W_0) &= \lim_{\|\Pi\| \rightarrow 0} \sum_{k=0}^{n-1} \left(f_t(t_k, W_k)(t_{k+1} - t_k) + f_x(t_k, W_k)(W_{k+1} - W_k) \right. \\
&\quad \left. + \frac{1}{2} f_{xx}(t_k, W_k)(W_{k+1} - W_k)^2 \right) \\
&= \int_0^T f_t(s, W_s) ds + \int_0^T f_x(s, W_s) dW_s + \frac{\sigma_{\Phi}^2}{2} \int_0^T f_{xx}(s, W_s) ds
\end{aligned}$$

esto como consecuencia de que $d[W, W]_s \equiv dW_s dW_s = \sigma_{\Phi}^2 ds$ (véase teorema 2.50). Por tanto la forma diferencial queda como sigue,

$$df(T, W_T) = f_t(T, W_T) + f_x(T, W_T) dW_T + \frac{\sigma_{\Phi}^2}{2} f_{xx}(T, W_T) dT$$

□

Ahora bien, se extiende el concepto de movimiento pseudo-browniano al concepto de pseudo-proceso de Itô y como se generaliza la fórmula de Itô-Doeblin para dichos procesos,

Definición 2.52. Sea $\mathbf{W}_\Phi := (W_t)_{t \geq 0}$ un movimiento pseudo-browniano adaptable con respecto a la filtración $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$. Un pseudo-proceso de Itô es un proceso estocástico $\mathbf{X}_\Phi = (X_t)_{t \geq 0}$ descrito de la siguiente forma,

$$X_t = X_0 + \int_0^t \Delta_s dW_s + \int_0^t \Theta_s ds$$

donde X_0 es una v.a. constante y Δ y Θ son procesos estocásticos adaptables a \mathbb{F} . Escrito de forma diferencial queda como sigue,

$$dX_t = \Delta_t dW_t + \Theta_t dt$$

Definición 2.53. Sea $\mathbf{X}_\Phi = (X_t)_{t \geq 0}$ un pseudo-proceso de Itô. Se define la variación de orden n de \mathbf{X}_Φ hasta el tiempo T como el siguiente límite,

$$\lim_{\|\Pi\| \rightarrow 0} \sum_{j=0}^{n-1} |X_{t_{j+1}} - X_{t_j}|^n$$

donde $\Pi = \{t_0, t_1, \dots, t_n\}$ y $0 = t_0 < t_1 < \dots < t_n = T$. A la variación de orden 2 se le denomina variación cuadrática y se denota especialmente como $[X, X]_T$. En su forma diferencial, la variación cuadrática tiene dos notaciones equivalentes,

$$d[X, X]_T \equiv dX_T dX_T$$

Teorema 2.54. Sea $\mathbf{X}_\Phi := (X_t)_{t \geq 0}$ un pseudo-proceso de Itô, entonces se cumple las siguientes propiedades:

1) La variación cuadrática $[X, X]_T$ es finita y es igual a lo siguiente,

$$[X, X]_T = \lim_{\|\Pi\| \rightarrow 0} \sum_{j=0}^{n-1} |X_{t_{j+1}} - X_{t_j}|^2 = \sigma_\Phi^2 \int_0^T \Delta_s^2 ds$$

2) La varianción de orden 1 es infinita, es decir,

$$\sum_{j=0}^{n-1} |X_{t_{j+1}} - X_{t_j}| \xrightarrow{\|\Pi\| \rightarrow 0} \infty$$

3) La varianción de orden $n \geq 3$ es cero, es decir,

$$\sum_{j=0}^{n-1} |X_{t_{j+1}} - X_{t_j}|^n \xrightarrow{\|\Pi\| \rightarrow 0} 0$$

Demostración. En representación diferencial, calculemos $d[X, X]_T$,

$$\begin{aligned} d[X, X]_s &\equiv dX_s dX_s = (\Delta_s dW_s + \Theta_s ds)(\Delta_s dW_s + \Theta_s ds) \\ &= \Delta_s^2 dW_s dW_s + 2\Delta_s \Theta_s ds dW_s + \Theta_s^2 ds ds \end{aligned}$$

como consecuencia del teorema 2.50 se cumple que $ds dW_s = 0 = ds ds$, por lo que,

$$d[X, X]_s = \Delta_s^2 dW_s dW_s = \sigma_{\Phi}^2 \Delta_s^2 ds$$

por lo tanto,

$$[X, X]_T = \int_0^T d[X, X]_s = \sigma_{\Phi}^2 \int_0^T \Delta_s^2 ds$$

cumpléndose así el punto 1). Los puntos 2) y 3) son consecuencia de que \mathbf{X}_{Φ} es un proceso estocástico continuo¹⁶, además las demostraciones son equivalentes a las dadas en el teorema 2.50 por lo cual se omiten. □

Corolario 2.54.1. Sea $\mathbf{X}_{\Phi} := (X_t)_{t \geq 0}$ un pseudo-proceso de Itô, entonces se cumple las siguientes propiedades:

1) La variación de X_t con t es cero, es decir,

$$\sum_{j=0}^{n-1} (X_{t_{j+1}} - X_{t_j})(t_{j+1} - t_j) \xrightarrow{\|\Pi\| \rightarrow 0} 0$$

2) En general, para $n \geq 1$ y $m \geq 1$ se cumple lo siguiente,

$$\sum_{j=0}^{n-1} (X_{t_{j+1}} - X_{t_j})^n (t_{j+1} - t_j)^m \xrightarrow{\|\Pi\| \rightarrow 0} 0$$

¹⁶Para mas información, véase *Shreve, op. cit.*

Las propiedad 1) se escribe de forma diferencial como sigue,

$$dX_t dt = 0 = dt dX_t$$

Demostración. La demostración es equivalente a la del corolario 2.50.1.

□

Definición 2.55. Sea \mathbf{X}_Φ un pseudo-proceso de Itô tal y como fue descrito en la definición 2.52 y sea $\mathbf{\Gamma} = (\Gamma_t)_{t \geq 0}$ un proceso adaptable a la filtración \mathbb{F} . Se define la integral con respecto al pseudo-proceso de Itô \mathbf{X}_Φ como sigue,

$$\int_0^t \Gamma_s dX_s := \int_0^t \Gamma_s \Delta_s dW_s + \int_0^t \Gamma_s \Theta_s ds$$

Teorema 2.56. (Fórmula de Itô-Doeblin para un pseudo-proceso de Itô) Sea $f(t, x)$ una función de orden \mathcal{C}^∞ en ambas entradas y sea $\mathbf{W}_\Phi = (W_t)_{t \geq 0}$ un movimiento pseudo-browniano. Entonces, para cualquier $T \geq 0$,

$$\begin{aligned} f(T, X_T) &= f(0, X_0) + \int_0^T f_t(s, X_s) ds + \int_0^T f_x(s, X_s) dX_s \\ &\quad + \frac{1}{2} \int_0^T f_{xx}(s, X_s) d[X, X]_s \\ &= f(0, X_0) + \int_0^T f_t(s, X_s) ds + \int_0^T f_x(s, X_s) \Delta_s dW_s \\ &\quad + \int_0^T f_x(s, X_s) \Theta_s ds + \frac{\sigma_\Phi^2}{2} \int_0^T f_{xx}(s, X_s) \Delta_s^2 ds \end{aligned}$$

o escrito en forma diferencial,

$$\begin{aligned} df(T, X_T) &= f_t(T, X_T) + f_x(T, X_T) dX_T + \frac{1}{2} f_{xx}(T, X_T) d[X, X]_s \\ &= f_t(T, X_T) + f_x(T, X_T) \Delta_T dW_T + f_x(T, X_T) \Theta_T dT \\ &\quad + \frac{\sigma_\Phi^2}{2} f_{xx}(T, X_T) \Delta_T^2 dT \end{aligned}$$

donde f_t, f_x, f_{xx}, f_{tx} y f_{tt} están definidas como sigue,

$$\begin{aligned} f_t(t, x) &:= \frac{\partial f(t, x)}{\partial t} \\ f_x(t, x) &:= \frac{\partial f(t, x)}{\partial x} \\ f_{xx}(t, x) &:= \frac{\partial^2 f(t, x)}{\partial x^2} \\ f_{tx}(t, x) &:= \frac{\partial^2 f(t, x)}{\partial t \partial x} \\ f_{tt}(t, x) &:= \frac{\partial^2 f(t, x)}{\partial t^2} \end{aligned}$$

Demostración. La demostración es similar a la del teorema 2.51, como consecuencia del teorema 2.56 y el corolario 2.54.1 se cumple lo siguiente para $f(T, X_T)$,

$$\begin{aligned} f(T, X_T) &= f(0, X_0) + \int_0^T f_t(s, X_s) ds + \int_0^T f_x(s, X_s) dX_s \\ &\quad + \frac{1}{2} \int_0^T f_{xx}(s, X_s) d[X, X]_s \end{aligned}$$

el teorema 2.51 establece que $d[X, X]_s = \sigma_{\Phi}^2 \Delta_s^2 ds$, y gracias a la definición 2.55 se tiene una equivalencia para la anterior formula, la cual queda como sigue,

$$\begin{aligned} f(T, X_T) &= f(0, X_0) + \int_0^T f_t(s, X_s) ds + \int_0^T f_x(s, X_s) \Delta_s dW_s \\ &\quad + \int_0^T f_x(s, X_s) \Theta_s ds + \frac{\sigma_{\Phi}^2}{2} \int_0^T f_{xx}(s, X_s) \Delta_s^2 ds \end{aligned}$$

y la forma diferencial queda como sigue,

$$\begin{aligned} df(T, X_T) &= f_t(T, X_T) + f_x(T, X_T) dX_T + \frac{1}{2} f_{xx}(T, X_T) d[X, X]_T \\ &= f_t(T, X_T) + f_x(T, X_T) \Delta_T dW_T + f_x(T, X_T) \Theta_T dT \\ &\quad + \frac{\sigma_{\Phi}^2}{2} f_{xx}(T, X_T) \Delta_T^2 dT \end{aligned}$$

□

Ejemplo 2.57. (*Movimiento Geométrico pseudo-Browniano*)

Recordando la ecuación 2.61, es importante ver su equivalencia en el caso de que $\mathbf{W}_\Phi = (W_t)_{t \geq 0}$ sea un movimiento pseudo-browniano y Φ no sea una distribución normal estandar. Para esto, se define el siguiente pseudo-proceso de Itô,

$$X_t = \mu(t) + \sigma W_t$$

donde $\mu(t)$ es una función continua no aleatoria, con esto se define el proceso estocástico $\mathbf{S}_\Phi = (S_t)_{t \geq 0}$ de la siguiente forma,

$$S_t := S_0 e^{X_t} = S_0 \exp \left\{ \mu(t) + \sigma W_t \right\}$$

al proceso \mathbf{S}_Φ se le denomina un movimiento geométrico pseudo-browniano. Implementando la fórmula del teorema 2.56 con $f(t, x) = S_0 e^x$, $f_t(t, x) = 0$,

$$f_x(t, x) = S_0 e^x = f_{xx}(t, x),$$

$$\begin{aligned} dS_t &= f_t(t, X_t) + f_x(t, X_t) \sigma dW_t + f_x(t, X_t) \mu(t) dt + \frac{\sigma_\Phi^2}{2} f_{xx}(t, X_t) \sigma^2 dt \\ &= \left(\mu(t) + \frac{\sigma^2 \sigma_\Phi^2}{2} \right) S_t dt + \sigma S_t dW_t \end{aligned}$$

Ahora bien, para la medida neutral al riesgo se desea que S_t tenga una esperanza igual al rendimiento continuo libre de riesgo sobre la tasa r , es decir,

$$\tilde{\mathbb{E}}(S_t) = S_0 e^{rt} = S_0 \tilde{\mathbb{E}} \left(\exp \left\{ \mu(t) + \sigma W_t \right\} \right) = S_0 e^{\mu(t)} \tilde{\mathbb{E}} \left(\exp \left\{ \sigma W_t \right\} \right)$$

por lo tanto,

$$\mu(t) = rt - \log \left\{ \tilde{\mathbb{E}} \left(\exp \left\{ \sigma W_t \right\} \right) \right\} \quad (2.65)$$

En el caso de que \mathbf{W}_Φ sea un movimiento browniano, se tiene que $W_t \sim N(0, \sqrt{t})$ y además,

$$\begin{aligned} \mu(t) &= rt - \log \left\{ \tilde{\mathbb{E}} \left(\exp \left\{ \sigma W_t \right\} \right) \right\} = rt - \log \left\{ \exp \left(\frac{1}{2} \sigma^2 t \right) \right\} \\ &= \left(r - \frac{1}{2} \sigma^2 \right) t \end{aligned}$$

y se cumple que $dS_t = r S_t dt + \sigma S_t dW_t$ tal y como sucede en la ecuación 2.61.

Una vez se cuenta con la distribución del activo S_t bajo la medida neutral al riesgo $\tilde{\mathbb{P}}$, se puede conocer la distribución de cualquier derivado $\mathbf{V} = (V_t)_{0 \leq t \leq T}$ con fecha de madurez $T > 0$ mediante la siguiente fórmula,

$$V_t = e^{-r(T-t)} \tilde{\mathbb{E}} \left(V_T \mid \mathcal{F}_t \right)$$

donde $V_T = G(S_T)$ y G es denominado el *payoff* del derivado. El teorema de Black-Scholes-Merton encuentra una función $H(t, x)$ la cual valua a V_t para cualquier valor del subyacente S_t , dicha función es la siguiente:

- 1) En el caso de que $G(S_T) = (S_T - K)^+$, es decir, \mathbf{V} representa una opción europea tipo Call, entonces $H = C$ y,

$$C(t, x) = x \text{Normal}(d_+(T-t, x)) - Ke^{-r(T-t)} \text{Normal}(d_-(T-t, x))$$

- 2) En el caso de que $G(S_T) = (K - S_T)^+$, es decir, \mathbf{V} representa una opción europea tipo Put, entonces $H = P$ y,

$$P(t, x) = xe^{-r(T-t)} \text{Normal}(-d_-(T-t, x)) - x \text{Normal}(-d_+(T-t, x))$$

esto lo hace suponiendo que \mathbf{S} es un movimiento geométrico browniano y resolviendo una ecuación diferencial parcial.

Algo muy importante a mencionar es que la función $H(t, x)$ **no** es una v.a., el factor de aleatoriedad se consigue en el modelo una vez se hay compuesto la función H con la v.a. S_t , este argumento es esencial para comprender el resultado siguiente.

Teorema 2.58. (*Extensión de la valuación Balck-Sholes-Merton para un movimiento geométrico pseudo-browniano*) Sea $\mathbf{S}_\Phi = (S_t)_{t \geq 0}$ un movimiento geométrico pseudo-browniano el cual cumple la siguiente ecuación diferencial estocástica,

$$dS_t = \left(\mu(t) + \frac{\sigma^2 \sigma_\Phi^2}{2} \right) S_t dt + \sigma S_t dW_t$$

y sea r la tasa libre de riesgo. Además, sea $\mathbf{V} = (V_t)_{0 \leq t \leq T}$ un derivado financiero que tiene a \mathbf{S}_Φ como subyacente y a $T > 0$ como fecha de madurez, entonces se cumple los siguientes dos casos:

- 1) En el caso de que $V_T = (S_T - K)^+$, es decir, \mathbf{V} representa una opción europea tipo Call con strike K , entonces $V_t = C(t, S_t)$ con C definido como sigue,

$$C(t, x) = x \text{Normal}(d_+(T-t, x)) - Ke^{-r(T-t)} \text{Normal}(d_-(T-t, x))$$

2) En el caso de que $V_T = (K - S_T)^+$, es decir, \mathbf{V} representa una opción europea tipo Put con strike K , entonces $V_t = P(t, S_t)$ con P definido como sigue,

$$P(t, x) = x e^{-r(T-t)} \text{Normal}(-d_-(T-t, x)) - x \text{Normal}(-d_+(T-t, x))$$

donde Normal es la función de distribución de una v.a. $\text{Normal}(0, 1)$ y,

$$d_{\pm}(\tau, x) := \frac{1}{\sigma \sigma_{\Phi} \sqrt{\tau}} \left[\log \frac{x}{K} + \left(r \pm \frac{\sigma^2 \sigma_{\Phi}^2}{2} \right) \tau \right]$$

Demostración. Para iniciar la demostración, se define $\alpha(t) = \left(\mu(t) + \frac{\sigma^2 \sigma_{\Phi}^2}{2} \right)$ y se tiene que $dS_t = \alpha(t) S_t dt + \sigma S_t dW_t$. Para cualquier derivado \mathbf{V} se busca lo siguiente,

$$dV_t = \Delta_t dS_t + r b_t e^{rt} dt \quad (2.66)$$

es decir, el rendimiento instantaneo del derivado \mathbf{V} es una combinación lineal de rendimientos instantaneos en el subyacente \mathbf{S}_{Φ} y la tasa libre de riesgo r , permitiendo así su replicación en el mercado. La ecuación 2.66 es equivalente a lo siguiente,

$$dV_t = \Delta_t S_t dt + \Delta_t \sigma S_t dW_t + r b_t e^{rt} dt \quad (2.67)$$

Ahora bien, si $H(t, x)$ denota la función de precio para el derivado \mathbf{V} a tiempo t , se tiene lo siguiente gracias al teorema 2.56,

$$dH(t, S_t) = H_t(t, S_t) dt + H_x(t, S_t) dS_t + \frac{1}{2} H_{xx}(t, S_t) d[S, S]_t \quad (2.68)$$

donde $d[S, S]_t$ es la variación cuadrática para el proceso \mathbf{S}_{Φ} la cual se define de forma equivalente a las definición 2.49. Como consiguiente, $d[S, S]_t$ está dado por la siguiente función,

$$\begin{aligned} d[S, S]_t &\equiv dS_t dS_t = (\alpha(t) S_t dt + \sigma S_t dW_t) (\alpha(t) S_t dt + \sigma S_t dW_t) \\ &= \sigma^2 S_t^2 dW_t dW_t = \sigma^2 \sigma_{\Phi}^2 S_t^2 dt \end{aligned}$$

con esto la ecuación 2.68 es equivalente a lo siguiente,

$$\begin{aligned} dH(t, S_t) &= \left(H_t(t, S_t) + \alpha(t) S_t H_x(t, S_t) + \frac{1}{2} \sigma^2 \sigma_{\Phi}^2 S_t^2 H_{xx}(t, S_t) \right) dt \\ &\quad + \sigma S_t H_x(t, S_t) dW_t \end{aligned} \quad (2.69)$$

igualando los términos con dW_t de las ecuaciones 2.67 y 2.69 se debe cumplir lo siguiente,

$$\left(\Delta_t \sigma S_t - \sigma S_t H_x(t, S_t) \right) dW_t = 0$$

o lo que es lo mismo,

$$\Delta_t = H_x(t, S_t)$$

y como consiguiente,

$$b_t = e^{-rt} \left\{ H(t, S_t) - H_x(t, S_t) S_t \right\}$$

Con esta información, se igualan los términos con dt en las ecuaciones 2.67 y 2.69, quedando la siguiente ecuación,

$$r H(t, S_t) = H_t(t, S_t) + r S_t H_x(t, S_t) + \frac{1}{2} \sigma^2 \sigma_{\Phi}^2 S_t^2 H_{xx}(t, S_t) \quad (2.70)$$

la ecuación 2.70 es una ecuación diferencial parcial, y se nota que la solución de dicha ecuación solo depende de la función H y no de la v.a. S_t , por lo tanto la solución es igual a la dada en el teorema 2.46 solo con un agregado de varianza en la forma de σ_{Φ}^2 .

Los puntos 1) y 2) dependen de las condiciones de frontera a tiempo T para la función H , las cuales son $G_C(S_T) = (S_T - K)^+$ y $G_P(S_T) = (K - S_T)^+$, respectivamente.

Como consecuencia del corolario 2.46.2, las deltas del portafolio son equivalentes a las del modelo Black-scholes Merton tradicional, solo con un agregado de varianza en la forma de σ_{Φ}^2 .

□

Observación 2.58.1. *Se puede notar que el precio del derivado V_{Φ} depende principalmente de los siguientes términos:*

- 1) *La tasa de rendimiento libre de riesgo r .*
- 2) *La variabilidad del subyacente σ^2 .*
- 3) *La variabilidad proveniente del movimiento pseudo browniano σ_{Φ}^2 .*
- 4) *El valor presente del subyacente S_0 .*
- 5) *El Strike o precio de ejercicio K .*

6) La madurez del derivado T .

Esto nos hace notar que si \mathbf{W}_{Φ_1} y \mathbf{W}_{Φ_2} son dos movimientos pseudo-brownianos tales que $\sigma_{\Phi_1}^2 = \sigma_{\Phi_2}^2$, entonces los precios de los derivados descritos en el teorema 2.58 son iguales sin importar las distribuciones de los movimientos pseudo-brownianos antes descrito.

Lo que si se va a ver afectado son los valores de Δ_t y V_t a través del tiempo, esto como consecuencia de que dichos procesos estocásticos dependen de \mathbf{S}_{Φ_1} (o en su caso \mathbf{S}_{Φ_2}), y dichas v.a. si contienen toda la información disponible en \mathbf{W}_{Φ_1} y \mathbf{W}_{Φ_2} (asimetría, colas pesadas, sesgo, kurtosis, etc.).

Ejemplo 2.59. Sean Φ_1 y Φ_2 las funciones de distribución de probabilidades siguientes,

$$\Phi_1(t) = \int_{-\infty}^t \frac{1}{\sqrt{2t\pi}} e^{-\frac{w^2}{2t}} dw$$

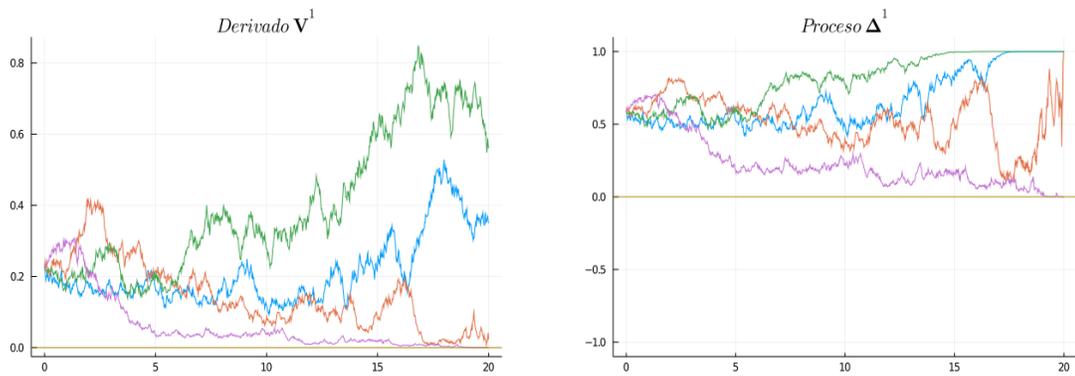
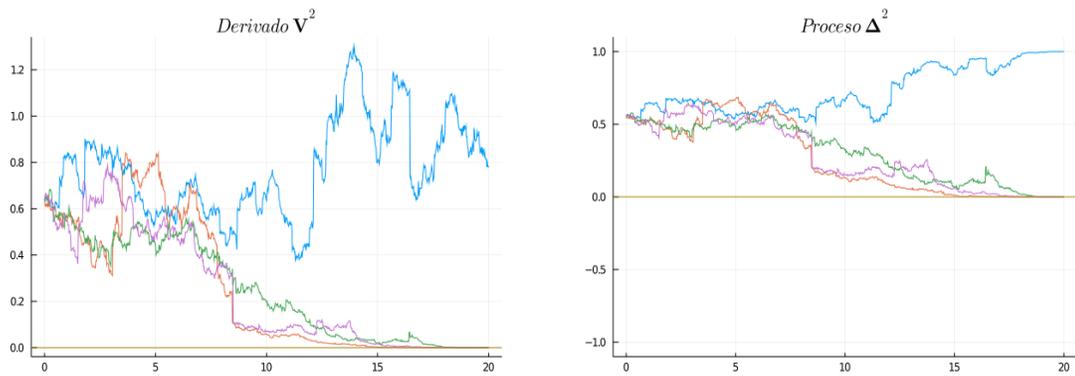
$$\Phi_2(t) = \int_{-\infty}^t \frac{\Gamma\left(\frac{8}{5}\right)}{\sqrt{2 \cdot 2t\pi} \Gamma\left(\frac{11}{10}\right)} \left(1 + \frac{w^2}{2 \cdot 2t}\right)^{-\frac{8}{5}} dw$$

es decir, Φ_1 es una normal estandar y Φ_2 es una distribución t con 2.2 grados de libertad. Para estas dos funciones se crean los movimientos geométricos pseudo-brownianos siguientes,

$${}_1S_t := S_0 \exp \left\{ \alpha_1(t) + \sigma_1 W_t \right\}$$

$${}_2S_t := S_0 \exp \left\{ \alpha_2(t) + \sigma_2 W_t \right\}$$

donde ${}_1W_t \sim \sqrt{t}\Phi_1$, ${}_2W_t \sim \sqrt{t}\Phi_2$ y $\alpha_1(t)$, $\alpha_2(t)$ son escogidos de tal forma que $\tilde{\mathbb{E}}({}_1S_t) = S_0 e^{rt} = \tilde{\mathbb{E}}({}_2S_t)$. Para ambos procesos, se definen opciones europeas tipo call con el mismo strike y madurez con $\sigma = 0.01$, $r = 0.0005$, $T = 20$ y $K = 10$, las figuras 2.11 y 2.12 representan el comportamiento del derivado y su delta respectiva.

Figura 2.11: Modelo bajo la distribución Φ_1 .Figura 2.12: Modelo bajo la distribución Φ_2 .

Para la figura 2.11, se nota como la distribución del subyacente cumple las propiedades del modelo Black-Scholes Merton tradicional, por lo que estas figuras son similares a las obtenidas en el ejemplo 2.47. Para el caso de la figura 2.12 se escogió una distribución con colas pesadas (distribución t) para notar claramente como los procesos estocásticos \mathbf{V} y $\mathbf{\Delta}$ se ven afectados por la distribución del subyacente \mathbf{S} .

Capítulo 3

Modelos discretos de valuación y resultados

3.1. Enfoque no paramétrico

El objetivo de la estadística no paramétrica es modelar la incertidumbre de un conjunto de v.a. sin suposiciones iniciales (sin suponer un modelo paramétrico). Esto se hace con el objetivo de modelar, con el mayor grado de libertad posible, la incertidumbre con la que se cuenta en el fenómeno o experimento aleatorio en cuestión.

Una desventaja de la estadística no paramétrica es que esta se apoya en gran medida de resultados asintóticos (es decir, cuando el tamaño de la muestra es grande). Por lo que si no se cuenta con una muestra aleatoria lo suficiente grande, los resultados hechos bajo este enfoque pueden no ser los mejores. Tal y como dice Shao (2003),

The theoretical justification for estimators in nonparametric models, however, relies more on asymptotics than that in parametric models . . . Thus, to choose between a parametric method and a nonparametric method, we need to balance the advantage of requiring weaker model assumptions (robustness) against the drawback of losing efficiency, which results in requiring a larger sample size.

En esta sección se plantean resultados básicos de estadística no paramétrica. Tales como la distribución empírica univariada y bivariada utilizando Shao *op. cit.*, así como la cópula empírica bivariada desarrollada en Nelsen *op. cit.*

3.1.1. Distribución empírica univariada

Para el resto de esta subsección, $\mathbf{X} := (X_1, \dots, X_n)$ es una muestra aleatoria de tamaño n .

Definición 3.1. Para la m.a. \mathbf{X} se define la función de distribución empírica como la siguiente colección de v.a.

$$\hat{F}_n := \left\{ \hat{F}_n(x) := \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_i \leq x\}} : x \in \mathbb{R} \right\}$$

Proposición 3.2. La v.a. $\hat{F}_n(x)$ establecida en la definición 3.1 es un estimador insesgado y consistente para $F_X(x)$.

Demostración. Para esta demostración se definen las siguientes n colecciones de v.a.,

$$\left\{ Z_i(x) := \mathbf{1}_{\{X_i \leq x\}} : x \in \mathbb{R} \right\}_{i \in \{1, \dots, n\}} \quad (3.1)$$

entonces para $x \in \mathbb{R}$ fija, $Z_i(x) \sim \text{Ber}(F_X(x))$ y además $Z_1(x), \dots, Z_n(x)$ es una m.a. de tamaño n , por lo que se cumple lo siguiente,

$$n\hat{F}_n(x) = \sum_{i=1}^n Z_i(x) \sim \text{Binom}(n, F_X(x))$$

Para demostrar que $\hat{F}_n(x)$ es insesgado,

$$\begin{aligned} n\mathbb{E}\left(\hat{F}_n(x)\right) &= \mathbb{E}\left(n\hat{F}_n(x)\right) = nF_X(x) \\ &\implies \mathbb{E}\left(\hat{F}_n(x)\right) = F_X(x) \end{aligned}$$

por lo cual $\hat{F}_n(x)$ es un estimador insesgado. Y para la consistencia,

$$\begin{aligned} n^2\mathbb{V}\left(\hat{F}_n(x)\right) &= \mathbb{V}\left(n\hat{F}_n(x)\right) = nF_X(x)(1 - F_X(x)) \\ \mathbb{V}\left(\hat{F}_n(x)\right) &= \frac{F_X(x)(1 - F_X(x))}{n} \end{aligned}$$

y ya que $\hat{F}_n(x)$ es un estimador insesgado la consistencia solo depende de la varianza del estadístico y se cumple que,

$$\mathbb{V}\left(\hat{F}_n(x)\right) \xrightarrow{n \rightarrow \infty} 0$$

por lo cual $\hat{F}_n(x)$ es un estimador puntual consistente para $F_X(x)$.

□

Observación 3.2.1. Para n fijo, la variabilidad de $\hat{F}_n(x)$ es mayor en las partes “centrales” de F_X (alrededor de la mediana) y además,

$$\lim_{x \rightarrow +\infty} \mathbb{V}\left(\hat{F}_n(x)\right) = 0 = \lim_{x \rightarrow -\infty} \mathbb{V}\left(\hat{F}_n(x)\right)$$

es decir, la variabilidad del estimador $\hat{F}_n(x)$ depende de x y disminuye conforme $|x| \rightarrow \infty$.

Una vez se cuenta con la muestra aleatoria observada¹ $\mathbf{x} := (x_1, \dots, x_n)$ (se denota como m.a.obs.) se da la siguiente definición,

Definición 3.3. Para una m.a.obs. \mathbf{x} se define la función de distribución empírica observada $F_n : \mathbb{R} \rightarrow [0, 1]$ como sigue,

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

Proposición 3.4. La función F_n cumple las siguientes propiedades:

- 1) $\{0, 1\} \subseteq \text{Ran } F_n \subset \left\{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\right\}$.
- 2) $F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{x_{(i)} \leq x\}}$ donde $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ son las estimaciones ordenadas de \mathbf{x} .
- 3) $\lim_{x \rightarrow -\infty} F_n(x) = 0$.
- 4) $\lim_{x \rightarrow +\infty} F_n(x) = 1$.
- 5) F_n es una función monótona creciente y continua por la derecha.

¹Es decir, una vez se cuenta con los datos reales del fenómeno o experimento aleatorio en cuestión.

Gracias a las propiedades (1)-(5) se cumple que F_n es una función de distribución de probabilidades perteneciente a una v.a. discreta denotada como X_n^{emp} . Además,

$$\text{Ran } X_n^{emp} = \mathbf{z}_x$$

donde $\mathbf{z}_x = \{z_1, \dots, z_m\}$ es el conjunto de elementos únicos de \mathbf{x} ordenadas de menor a mayor (claramente $m \leq n$).

Demostración. La propiedad (1) es clara por como está definida la función F_n . Si $\text{Ran } F_n = \left\{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\right\}$ entonces la m.a.obs. está compuesta por n elementos distintos (en este caso $m = n$), en el caso de que $\text{Ran } F_n \neq \left\{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\right\}$, la m.a.obs. está compuesta por elementos repetidos (en este caso $m < n$).

La propiedad (2) es solamente un reordenamiento de la m.a.obs. \mathbf{x} , en el caso de (3) y (4) si $x < x_{(1)}$ entonces $F_n(x) = 0$ y $F_n(x) = 1$ para $x \geq x_{(n)}$ por lo tanto,

$$\begin{aligned} \lim_{x \rightarrow -\infty} F_n(x) &= 0 \\ \lim_{x \rightarrow +\infty} F_n(x) &= 1 \end{aligned}$$

Para (5), si $t_1 \leq t_2$ entonces,

$$F_n(t_1) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{x_i \leq t_1\}} \leq \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{x_i \leq t_2\}} = F_n(t_2)$$

Ahora bien, los puntos de discontinuidad para la función F_n son los valores $t \in \mathbf{z}_x$ (como sucede en el caso de cualquier v.a. discreta), si se toma $t \in \mathbf{z}_x$ cualquiera entonces existe un $i^* \in \{1, \dots, n\}$ tal que,

$$F_n(t) = \frac{i^*}{n}$$

y esto se cumple también $\forall t \in [x_{(i^*)}, x_{(i^*)+1}[$, por lo tanto,

$$\lim_{x \rightarrow t^+} F_n(x) = \frac{i^*}{n}$$

asi F_n es continua por la derecha y monótona creciente.

□

Observación 3.4.1. Sean X_n^{emp} y \mathbf{z}_x los términos definidos en la proposición 3.4. Entonces la esperanza de X_n^{emp} es igual a la media muestral observada, es decir,

$$\mathbb{E}(X_n^{emp}) = \frac{1}{n} \sum_{i=1}^n x_i \equiv \bar{x}_n \quad (3.2)$$

Demostración. Para esta demostración solo es necesario obtener $\mathbb{P}(X_n^{emp} = z_j)$,

$$\mathbb{P}(X_n^{emp} = z_j) = F_n(z_j) - \lim_{\delta \rightarrow 0^-} F_n(z_j + \delta) = \sum_{i=1}^n \mathbf{1}_{\{x_i = z_j\}}$$

y como $\mathbf{z}_x = \text{Ran } X_n^{emp}$, entonces la esperanza de X_n^{emp} queda como sigue,

$$\mathbb{E}(X_n^{emp}) = \sum_{j=1}^m z_j \mathbb{P}(X_n^{emp} = z_j) = \frac{1}{n} \sum_{j=1}^m \sum_{i=1}^n z_j \mathbf{1}_{\{x_i = z_j\}} = \frac{1}{n} \sum_{i=1}^n x_i$$

cumpléndose así la ecuación 3.2. □

El resultado mas importante con respecto a la distribución empírica es el teorema de Gilvenko-Cantelli el cual se establece a continuación,

Teorema 3.5. (*Gilvenko-Cantelli*) Sea \hat{F}_n la distribución empírica construida mediante una m.a. \mathbf{X} perteneciente a un modelo F_X cualquiera. Se define la v.a. D_n siguiente,

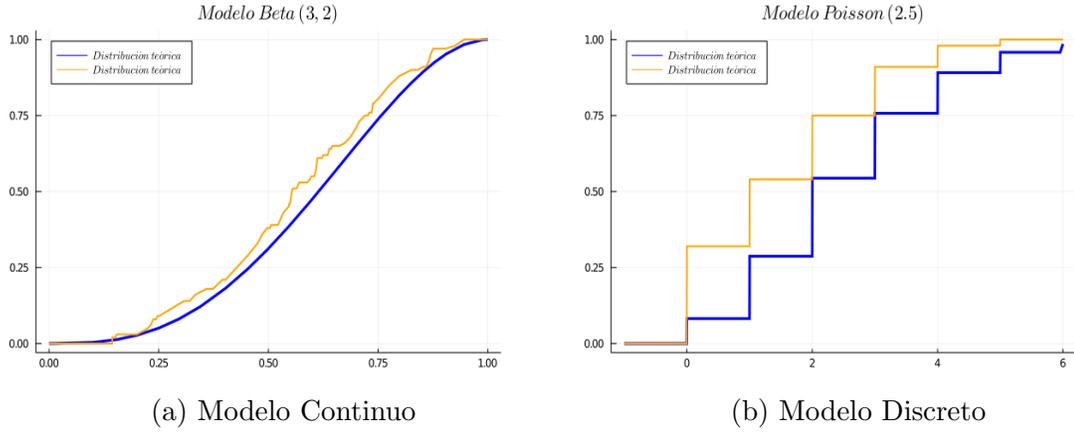
$$D_n := \sup_{x \in \mathbb{R}} |\hat{F}_n(x) - F_X(x)|$$

entonces se cumple que $D_n \xrightarrow{c.s.} 0$, es decir,

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} D_n = 0\right) = 1$$

El teorema 3.5 establece que conforme el tamaño de una m.a. incrementa ($n \rightarrow \infty$) la distribución empírica hace un mejor ajuste a la distribución F_X . Esto es gracias a que “la peor” (mayor) desviación que tiene \hat{F}_n con respecto a F_X (representada por D_n) es cada vez menor conforme el tamaño de la m.a. incrementa ($D_n \xrightarrow{c.s.} 0$).

En la figura 3.1 se puede ver como se comporta la función F_n para distintos modelos paramétricos. Se nota como el ajuste es mejor cuando $|x| \rightarrow \infty$, además se puede ver como la función F_n es escalonada (tal como es el caso en una v.a. discreta).


 Figura 3.1: Distribución empírica observada para $n = 100$.

3.1.2. Distribución y Cópula empírica bivariada

Sea $\mathcal{M}_n \equiv (\mathbf{X}, \mathbf{Y}) := ((X_1, Y_1), \dots, (X_n, Y_n))$ una m.a. bivariada perteneciente a un modelo bivariado H_{XY} , compuesta por m.a. univariadas $\mathbf{X} = (X_1, \dots, X_n)$ y $\mathbf{Y} = (Y_1, \dots, Y_n)$ con funciones de distribución univariadas F_X y G_Y , respectivamente. Para el resto de esta subsección X y Y son modelos marginalmente continuos, esto con el objetivo de asegurar la unicidad de C_{XY} .

Definición 3.6. Para la m.a. \mathcal{M}_n se define la función de distribución empírica bivariada como la siguiente colección de v.a.,

$$\hat{H}_n := \left\{ \hat{H}_n(x, y) := \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_i \leq x, Y_i \leq y\}} : (x, y) \in \mathbb{R}^2 \right\}$$

Proposición 3.7. La v.a. $\hat{H}_n(x, y)$ establecida en la definición 3.6 es un estimador insesgado y consistente para $H_{XY}(x, y)$.

Demostración. La demostración es equivalente al caso univariado, por lo cual se omite. Solo se escribe la esperanza y varianza del estimador $\hat{H}_n(x, y)$ a continuación,

$$\begin{aligned} \mathbb{E}(\hat{H}_n(x, y)) &= H_n(x, y) \\ \mathbb{V}(\hat{H}_n(x, y)) &= \frac{H_n(x, y)(1 - H_n(x, y))}{n} \end{aligned}$$

□

Observación 3.7.1. Para n fijo, la variabilidad de $\hat{H}_n(x, y)$ para los puntos $(x, y) \in \mathbb{R}^2$ tales que $H_{XY}(x, y) \approx \frac{1}{2}$ es mayor, y además,

$$\lim_{(x,y) \rightarrow (+\infty, +\infty)} \mathbb{V}(\hat{H}_n(x, y)) = 0 = \lim_{(x,y) \rightarrow (-\infty, -\infty)} \mathbb{V}(\hat{H}_n(x, y))$$

es decir, la variabilidad del estimador $\hat{H}_n(x, y)$ decrece conforme $\max\{|x|, |y|\} \rightarrow \infty$.

Una vez se cuenta con la m.a.obs. bivariada $(\mathbf{x}, \mathbf{y}) := ((x_1, y_1), \dots, (x_n, y_n))$ perteneciente a los modelos univariados continuos X y Y , respectivamente; se da la siguiente definición,

Definición 3.8. Para una m.a.obs. (\mathbf{x}, \mathbf{y}) se define la función de distribución empírica observada $H_n : \mathbb{R}^2 \rightarrow [0, 1]$ como sigue,

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

Proposición 3.9. La función H_n cumple las siguientes propiedades:

- 1) $\{0, 1\} \subseteq \text{Ran } H_n \subseteq \left\{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\right\}$.
- 2) Las distribuciones empíricas observadas univariadas de las v.a. X y Y , denotadas como F_n y G_n , respectivamente, se pueden obtener mediante los siguientes límites:

$$F_n(x) = \lim_{y \rightarrow \infty} H_n(x, y)$$

$$G_n(y) = \lim_{x \rightarrow \infty} H_n(x, y)$$

- 3) La función H_n cumple las propiedades de una función de distribución bivariada.
- 4) Las m.a.obs. \mathbf{x} y \mathbf{y} están compuestas marginalmente por valores distintos, es decir, para cualquier $i, j \in I_n$ distintos se cumple que $x_i \neq x_j$ y $y_i \neq y_j$, por lo tanto,

$$\text{Ran } F_n = \text{Ran } G_n = \left\{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\right\}$$

Demostración. Las propiedades (1) es equivalente al caso univariado y por esa razón se omite, en el caso de (2) se cumple lo siguiente,

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

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

cumpléndose así (2). Para el caso de (3) se demuestran los puntos establecidos en la proposición 1.37. Para (i) se toma un $\delta < 0$ cualquiera,

$$H_n(x, -\infty) = H_n(x, y_{(1)} + \delta) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{x_i \leq x, y_i \leq y_{(1)} + \delta\}} = 0$$

$$H_n(-\infty, y) = H_n(x_{(1)} + \delta, y) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{x_i \leq x_{(1)} + \delta, y_i \leq y\}} = 0$$

la propiedad (ii) ya fue demostrada donde el v.a. en cuestión es $(X_n^{\text{emp}}, Y_n^{\text{emp}})$, la propiedad (iii) es consecuencia inmediata de (ii) mientras que las propiedades (v) y (vi) son propiedades equivalentes al caso discreto. Para la propiedad (iv) se toman $t_1 < t_2$ y $s_1 < s_2$,

$$\begin{aligned} & H_n(t_2, s_2) - H_n(t_1, s_2) - H_n(t_2, s_1) + H_n(t_1, s_1) \\ &= \frac{1}{n} \sum_{i=1}^n \left(\mathbf{1}_{\{x_i \leq t_2, y_i \leq s_2\}} - \mathbf{1}_{\{x_i \leq t_1, y_i \leq s_2\}} \right. \\ & \quad \left. - \mathbf{1}_{\{x_i \leq t_2, y_i \leq s_1\}} + \mathbf{1}_{\{x_i \leq t_1, y_i \leq s_1\}} \right) \\ &= \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{t_1 < x_i \leq t_2, s_1 < y_i \leq s_2\}} \geq 0 \end{aligned}$$

quedando así demostrado (iv) y la propiedad (3) en conjunto. La propiedad (4) es consecuencia de que ninguna muestra observada en \mathbf{x} o \mathbf{y} se repite, ya que las v.a. X y Y son continuas. □

Se sabe que la función de distribución empírica \hat{H}_n es un estadístico que permite “aproximar” la distribución conjunta de probabilidades H_{XY} ; entonces se busca bajo un argumento similar encontrar una familia de variables aleatorias \hat{C}_n que de igual forma permitan “aproximar” a la cópula C_{XY} . Esto lo hace la siguiente definición,

Definición 3.10. Para la m.a. \mathcal{M}_n se define la cópula empírica como la siguiente familia de v.a.,

$$\mathcal{C}_n = \left\{ \hat{C}_n \left(\frac{j_1}{n}, \frac{j_2}{n} \right) := \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_i \leq X_{(j_1)}, Y_i \leq Y_{(j_2)}\}} : j_1, j_2 \in \{1, \dots, n\} \right\}$$

Nótese como \mathcal{C}_n no está definida para todo $(u, v) \in I^2$, esto se debe a que para cualquier $(x, y) \in \mathbb{R}^2$ se cumple que,

$$\text{Ran } F_n = \text{Ran } G_n = \left\{ 0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1 \right\}$$

gracias a la proposición 3.9. Por lo que nunca se cubre todo el intervalo unitario I con las distribuciones empíricas univariadas \hat{F}_n y \hat{G}_n , respectivamente².

Como consecuencia de la proposición 3.9 se tiene que H_n es válida como distribución conjunta de probabilidades para el vector aleatorio $(X_n^{\text{emp}}, Y_n^{\text{emp}})$. Y gracias al Teorema de Sklar (teorema 1.55), se cumple la siguiente definición,

Definición 3.11. Para las funciones H_n, F_n y G_n establecidas en la proposición 3.9 se cumple que existe una única función $C_n : \text{Ran } F_n \times \text{Ran } G_n \rightarrow [0, 1]$ tal que,

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

a la función C_n se le denomina cópula empírica observada. Además, se define la diagonal empírica observada $\delta_n : \left\{ 0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1 \right\} \rightarrow [0, 1]$ como sigue,

$$\delta_n(t) := C_n(t, t)$$

Observación 3.11.1. Sean H_n, C_n, F_n, G_n las funciones usadas en la definición 3.11. Gracias a la propiedad (4) en la proposición 3.9 se cumple que,

$$\text{Ran } F_n = \left\{ 0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1 \right\} = \text{Ran } G_n$$

Como \mathbf{x} y \mathbf{y} tienen valores distintos entonces se cumple lo siguiente,

$$\begin{aligned} F_n(x_{(j_1)}) &= \frac{j_1}{n} \\ G_n(y_{(j_2)}) &= \frac{j_2}{n} \end{aligned}$$

²Para más información con respecto a la cópula empírica, véase Deheuvels (1979).

por lo que C_n queda completamente definida por la siguiente ecuación,

$$C_n\left(\frac{j_1}{n}, \frac{j_2}{n}\right) = H_n\left(F_n^{(-1)}\left(\frac{j_1}{n}\right), G_n^{(-1)}\left(\frac{j_2}{n}\right)\right) = H_n(x_{(j_1)}, y_{(j_2)})$$

es decir,

$$C_n\left(\frac{j_1}{n}, \frac{j_2}{n}\right) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{x_i \leq x_{(j_1)}, y_i \leq y_{(j_2)}\}} \quad (3.3)$$

en el caso de que $j_1, j_2 \in \{1, \dots, n\}$. Si alguna de las entradas se anula ($j_1 = 0$ ó $j_2 = 0$) la cópula empírica observada es cero.

Una vez se cuenta con la cópula empírica observada C_n , se desea cuantificar concordancia y dependencia mediante la m.a.obs. (\mathbf{x}, \mathbf{y}) , esto se puede hacer con estimaciones muestrales de la ρ de Spearman y la σ de Schweizer, respectivamente. El origen de los dos siguientes estadísticos se puede encontrar en Nelsen *op. cit.* (en el caso de la ρ de Spearman) y en Poczos *et. al.* (2018) (en el caso de la σ de Schweizer).

Definición 3.12. Sea C_n la cópula empírica observada perteneciente a una m.a. (\mathbf{X}, \mathbf{Y}) ; donde X y Y son v.a. continuas y (\mathbf{X}, \mathbf{Y}) cuenta con cópula subyacente C_{XY} . Se definen ρ_n y σ_n como los siguientes dos estadísticos:

$$\rho_n := \frac{12}{n^2 - 1} \sum_{i=1}^n \sum_{j=1}^n \left(\hat{C}_n\left(\frac{i}{n}, \frac{j}{n}\right) - \frac{i}{n} \times \frac{j}{n} \right)$$

$$\sigma_n := \frac{12}{n^2 - 1} \sum_{i=1}^n \sum_{j=1}^n \left| \hat{C}_n\left(\frac{i}{n}, \frac{j}{n}\right) - \frac{i}{n} \times \frac{j}{n} \right|$$

A ρ_n y σ_n se les considera los estadísticos canónicos para la ρ de Spearman y la σ de Schweizer, respectivamente, para la cópula C_{XY} .

Además, se definen r_n y s_n como las versiones muestrales de ρ y σ , respectivamente; quedando del siguiente modo,

$$r_n := \frac{12}{n^2 - 1} \sum_{i=1}^n \sum_{j=1}^n \left(C_n\left(\frac{i}{n}, \frac{j}{n}\right) - \frac{i}{n} \times \frac{j}{n} \right)$$

$$s_n := \frac{12}{n^2 - 1} \sum_{i=1}^n \sum_{j=1}^n \left| C_n\left(\frac{i}{n}, \frac{j}{n}\right) - \frac{i}{n} \times \frac{j}{n} \right|$$

donde C_n es la cópula empírica observada creada mediante la m.a.obs. (\mathbf{x}, \mathbf{y}) .

Ejemplo 3.13. Sean \mathcal{M}_n^1 y \mathcal{M}_n^2 m.a. de tamaño 300 pertenecientes a los modelos bivariados H_{XY}^1 , H_{XY}^2 , respectivamente,

$$\begin{aligned} H_{XY}^1(x, y) &= C^1(F_X(x), F_Y(y)) \\ H_{XY}^2(x, y) &= C^2(F_X(x), F_Y(y)) \end{aligned}$$

donde X y Y son variables aleatorias absolutamente continuas con funciones de densidad,

$$\begin{aligned} f_X(x) &= \frac{x^{6.5} e^{-x}}{\Gamma(7.5)} \mathbf{1}_{\{x > 0\}} \\ f_Y(y) &= \frac{y^{2.5} e^{-y/2}}{\Gamma(3.5) 2^{3.5}} \mathbf{1}_{\{x > 0\}} \end{aligned}$$

es decir, X y Y son v.a. $\text{Gamma}(7.5, 1)$ y $\text{Gamma}(3.5, 2)$, respectivamente. Además, las cópulas C^1 y C^2 son las siguientes,

$$\begin{aligned} C^1(u, v) &= \left(\max \{ (u^{0.5} + v^{0.5} - 1), 0 \} \right)^2 \\ C^2(u, v) &= uv \exp \left\{ \left[(-\log u)^{-2} + (-\log v)^{-2} \right]^{-1/2} \right\} \end{aligned}$$

es decir, C^1 es una cópula de Clayton con parámetro 4 y C^2 es una cópula de Gamlambos con parámetro 2.

En la figuras 3.2b y 3.2a se puede apreciar mapas de dispersión de los modelos H_{XY}^1 y H_{XY}^2 . Nótese que aunque las distribuciones univariadas son las mismas, el modelo se ve completamente distinto como consecuencia del cambio de dependencia,

En las figuras 3.3 y 3.4 se compara los conjuntos de nivel para las cópulas empíricas observadas (izquierda) y las cópulas teóricas (derecha). Nótese como el mayor “ruido” presente en las cópulas empíricas observadas C_n^1, C_n^2 se alcanza principalmente en las partes centrales del cuadrado unitario. Esto como consecuencia de la varianza de la distribución empírica obtenida en la proposición 3.7.

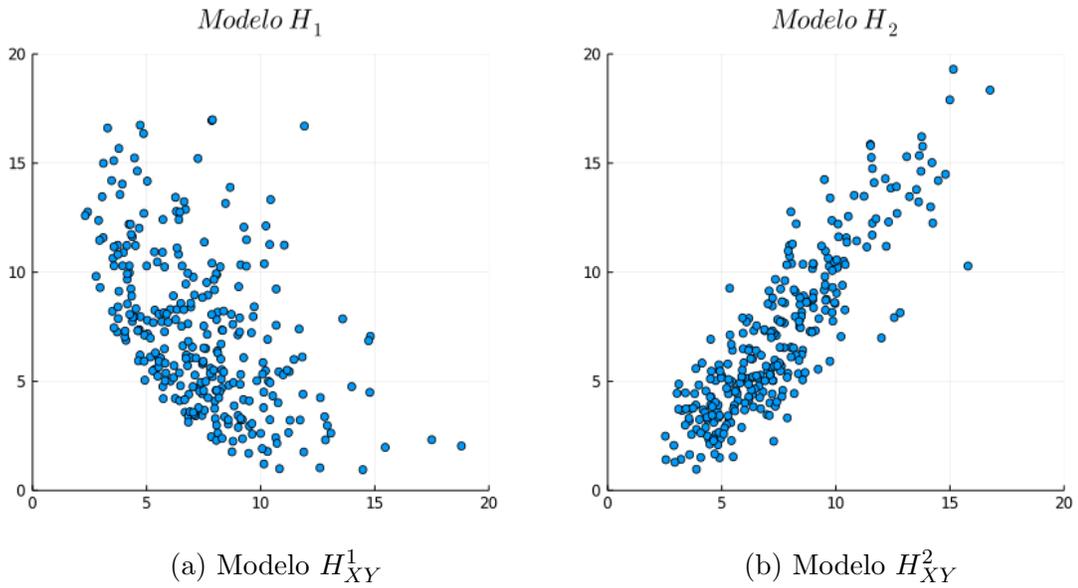


Figura 3.2: Simulación de tamaño 300.

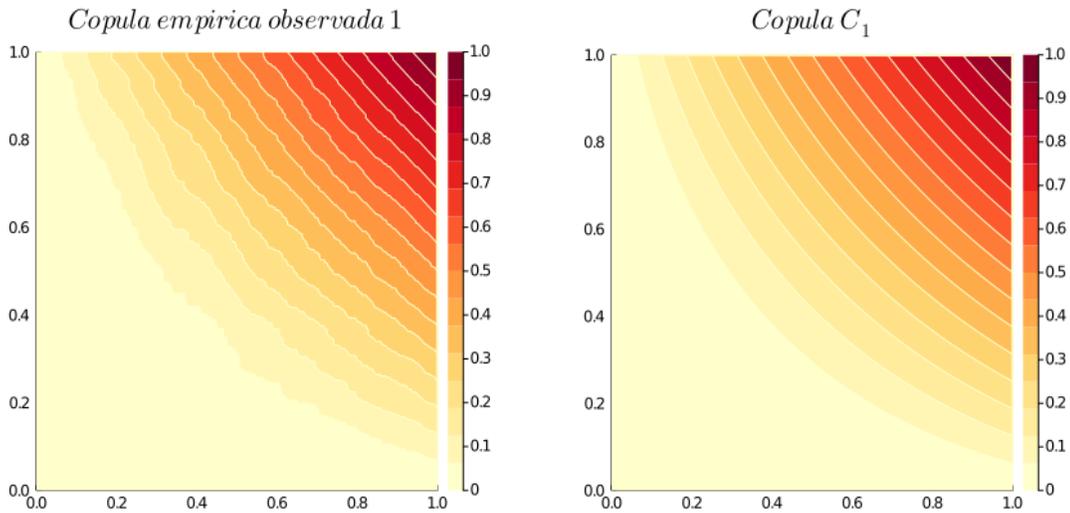


Figura 3.3: Conjuntos de nivel para el Modelo C_1 .

En la figura 3.5 se pueden ver las diagonales empíricas observadas δ_n^1 y δ_n^2 y su parecido con las diagonales teóricas δ^1 y δ^2 , respectivamente. Nótese como el mayor “ruido” de estas funciones se consigue alrededor del 0.5.

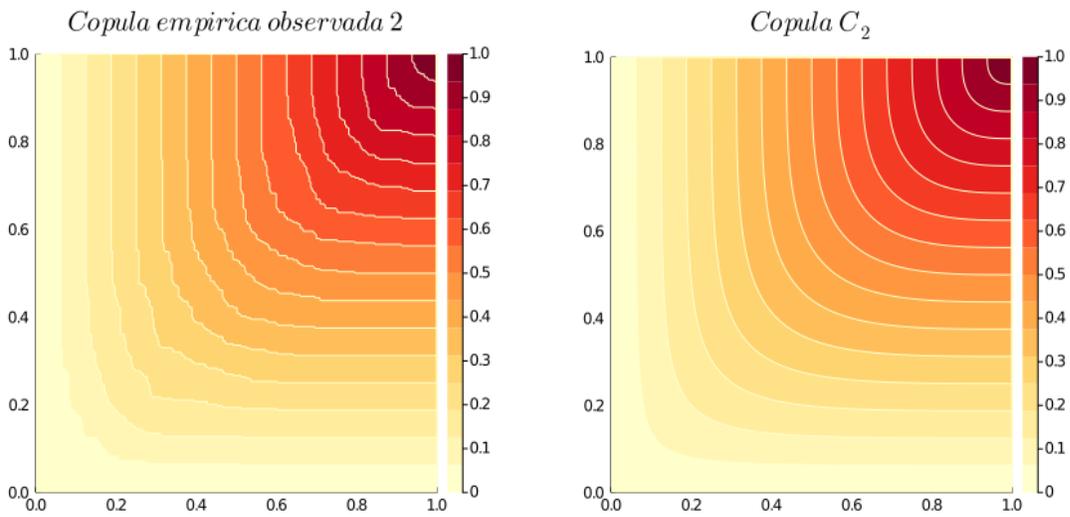


Figura 3.4: Conjuntos de nivel para el Modelo C_2 .

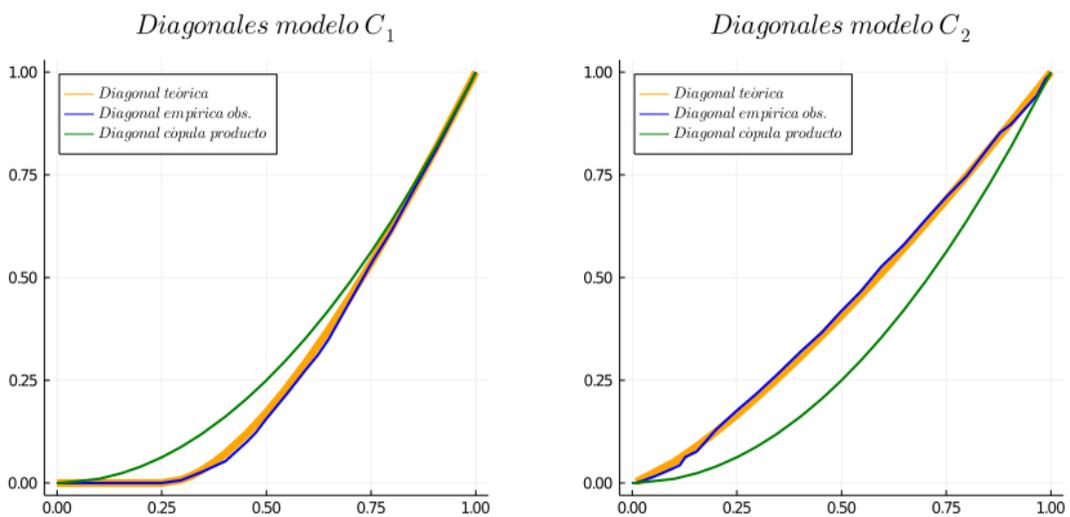


Figura 3.5: Diagonales empíricas observadas y teóricas.

Para terminar, se calculan r_n y s_n y se comparan dichos valores con los valores teóricos de la ρ de Spearman y la σ de Schweizer, respectivamente.

Modelo	r_n	ρ	s_n	σ
H_1	-0.384	-0.466	0.384	0.466
H_2	0.837	0.818	0.837	0.818

Cuadro 3.1: Estimación muestral de la ρ de Spearman y la σ de Schweizer.

Dada una m.a. (\mathbf{X}, \mathbf{Y}) donde X y Y son v.a. continuas, es de interés conocer si la cópula subyacente de dicho modelo (denotada como C_{XY}) es simétrica o no. Si la cópula C_{XY} es una cópula simétrica, restringir el modelo paramétrico a la familia arquimediana no suena como una idea tan mala (toda cópula arquimediana es simétrica por construcción). En caso contrario, en automático se debe descartar a la familia arquimediana como una opción viable en el ámbito paramétrico.

Para probar simetría, se propone una prueba de hipótesis basada en los siguientes dos comentarios:

- 1) La cópula producto $\Pi(u, v) = uv$, cuenta con la mayor variabilidad con respecto a la familia de cópulas simétricas. Esto gracias a que dicha cópula modela la falta completa de dependencia en un modelo bivariado.
- 2) Cualquier cópula C (sin contar Π) cuenta con mayor información disponible que la cópula Π para hacer notar si es o no simétrica.

Gracias a los anteriores dos puntos, se considera a la distribución empírica de la cópula producto como base para la creación de una prueba de hipótesis en relación a simetría de cópulas. El estadístico de la prueba es el siguiente,

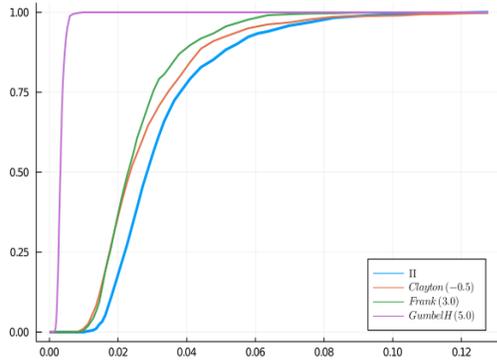
$$\Xi_n := \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \left(\hat{\Pi}_n \left(\frac{i}{n}, \frac{j}{n} \right) - \hat{\Pi}_n \left(\frac{j}{n}, \frac{i}{n} \right) \right)^2 \quad (3.4)$$

Conforme n crece, se debería cumplir que $\Xi_n \xrightarrow{c.s.} 0$. Además si para cualquier cópula empírica perteneciente a un modelo simétrico C , se define Ξ_n^C como sigue,

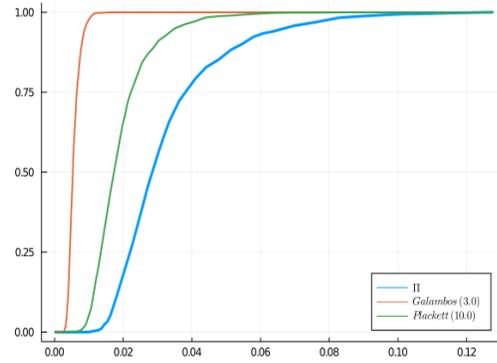
$$\Xi_n^C := \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \left(\hat{C}_n \left(\frac{i}{n}, \frac{j}{n} \right) - \hat{C}_n \left(\frac{j}{n}, \frac{i}{n} \right) \right)^2$$

gracias a la propiedades (1) y (2) se debería cumplir la siguiente desigualdad,

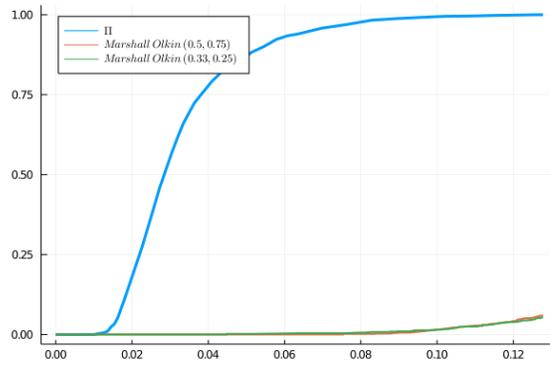
$$\mathbb{P}(\Xi_n^C \leq t) \geq \mathbb{P}(\Xi_n \leq t) \quad (3.5)$$



(a) Cópulas arquimedianas



(b) Cópulas simétricas no arquimedianas



(c) Cópulas asimétricas

Figura 3.6: Distribución del estadístico Ξ_n^C .

en el caso de que C sea una cópula simétrica, es decir, el estadístico Ξ_n^C converge de forma “mas rápida” a 0 que el estadístico Ξ_n . Y en el caso de que C sea una cópula asimétrica,

$$\mathbb{P}(\Xi_n^C \leq t) \ll \mathbb{P}(\Xi_n \leq t) \quad (3.6)$$

Gracias a esto, se hace la siguiente propuesta de prueba de hipótesis no paramétrica,

Propuesta 3.1. (Prueba de simetría) Sea (\mathbf{X}, \mathbf{Y}) una m.a. donde X y Y son v.a. continuas, se propone una prueba de hipótesis no paramétrica para la simetría de C_{XY} . La hipótesis nula queda como sigue:

$$\mathcal{H}_0 : C_{XY} \text{ es una cópula simétrica.}$$

la probabilidad de Error Tipo I (que sirve para rechazar \mathcal{H}_0) esta dada de la siguiente forma:

$$\mathbb{P}(\text{Error Tipo I}) = \mathbb{P}(\text{Rechazar } \mathcal{H}_0 \mid \mathcal{H}_0 \text{ es verdadera})$$

si \mathcal{H}_0 es verdadera entonces se cumple la ecuación 3.5. Por lo que ya una vez se tiene la m.a.obs. (\mathbf{x}, \mathbf{y}) , la probabilidad de Error Tipo I cumple lo siguiente,

$$\mathbb{P}(\text{Error Tipo I}) = \mathbb{P}(\Xi_n^C > \xi_n) \leq \mathbb{P}(\Xi_n > \xi_n)$$

así $\mathbb{P}(\Xi_n > \xi_n)$ se utiliza como cota superior de la probabilidad de Error Tipo I, donde ξ_n queda definido como sigue,

$$\xi_n := \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \left(C_n \left(\frac{i}{n}, \frac{j}{n} \right) - C_n \left(\frac{j}{n}, \frac{i}{n} \right) \right)^2$$

la probabilidad de Error Tipo I es útil para rechazar \mathcal{H}_0 así como para aceptar la hipótesis alternativa \mathcal{H}_1 , la cual queda como sigue:

$$\mathcal{H}_1 : C_{XY} \text{ es una cópula asimétrica.}$$

En la práctica la prueba de simetría funciona como sigue:

- 1) Fijar el valor α que se desea como valor máximo para la probabilidad de Error Tipo I.
- 2) Se obtiene el valor ξ_n partiendo de la m.a.obs. (\mathbf{x}, \mathbf{y}) .
- 3) Se obtiene la probabilidad $p_{\xi_n} := \mathbb{P}(\Xi_n > \xi_n)$.
- 4) Se toma una decisión: Si $p_{\xi_n} \leq \alpha$, entonces se rechaza la hipótesis nula \mathcal{H}_0 y por lo tanto se acepta que C_{XY} es asimétrica con una probabilidad $(1 - p_{\xi_n}) \times 100\%$ de confianza. En caso contrario, no existe información suficiente para rechazar la hipótesis nula \mathcal{H}_0 .

En las figura 3.6a y 3.6b se puede notar como la ecuación 3.5 se cumple para la familia arquimediana de cópulas y también para cópulas simétricas fuera de dicha familia. Mientras que en la figura 3.6c se puede notar como la ecuación 3.6 se cumple para la familia de cópulas asimétricas.

3.2. Activos financieras a tiempo discreto

Para el resto de esta sección, \mathbf{D} es un derivado financiero que tiene como subyacente a los activos \mathbf{S}^1 y \mathbf{S}^2 . Además, N es la madurez del derivado y r es la tasa de rendimiento libre de riesgo con capitalización continua utilizada para el descuento de cualquier instrumento financiero en el mercado. Antes de poder valorar el derivado \mathbf{D} , se necesita tener la distribución de los subyacentes \mathbf{S}^1 y \mathbf{S}^2 la cual se establece a continuación.

3.2.1. Distribución discreta del subyacente

Se supone que los activos subyacentes \mathbf{S}^1 y \mathbf{S}^2 son árboles multinomiales a n periodos con $n \in \{0, \dots, N\}$ (cada periodo representa un rebalanceo en el portafolio de replicación del derivado \mathbf{D} , además los periodos son de la misma longitud temporal para poder ser representados por la misma v.a.). Es decir, S_n^1 y S_n^2 son iguales a lo siguiente,

$$S_n^1 = S_0^1 \prod_{j=1}^n U_j \quad n \in \{1, \dots, N\}$$

$$S_n^2 = S_0^2 \prod_{j=1}^n V_j \quad n \in \{1, \dots, N\}$$

con S_0^1 y S_0^2 los valores presentes de los activos \mathbf{S}^1 y \mathbf{S}^2 , respectivamente. Las v.a. $\{(U_1, V_1), \dots, (U_N, V_N)\}$ son una m.a. con distribución en común (U^m, V^m) , donde m representa los m -posibles escenarios en cada periodo para cada subyacente³.

Para asemejar al modelo Black-Scholes Merton, se tiene que el v.a. (U^m, V^m) representa los rendimientos intra-periodo para ambos subyacentes. Ahora bien, la distribución del v.a. (S_n^1, S_n^2) antes descrito es sobre la medida real \mathbb{P} , por lo que se necesita obtener la distribución de dicho modelo bajo la medida neutral al riesgo $\tilde{\mathbb{P}}$, esto se hace realizando el siguiente ajuste a las distribuciones de U^m y V^m ,

$$U^m \stackrel{\tilde{\mathbb{P}}}{=} U^m \frac{e^r}{\bar{u}_m}$$

$$V^m \stackrel{\tilde{\mathbb{P}}}{=} V^m \frac{e^r}{\bar{v}_m}$$

³Como se tienen 2 subyacentes, en cada periodo existen m^{2n} escenarios posibles de forma conjunta.

donde \bar{u}_m y \bar{v}_m son las esperanzas de las v.a. U^m y V^m bajo la medida \mathbb{P} , respectivamente. Dichas esperanzas se obtienen como sigue,

$$\bar{u}_m \equiv \mathbb{E}(U^m) = \frac{1}{n} \sum_{i=1}^n \mathbb{P}(U^m = u_m) u_m$$

$$\bar{v}_m \equiv \mathbb{E}(V^m) = \frac{1}{n} \sum_{i=1}^n \mathbb{P}(V^m = v_m) v_m$$

donde $\text{Ran } U^m = \{u_1, \dots, u_m\}$ y $\text{Ran } V^m = \{v_1, \dots, v_m\}$. Cabe mencionar que la subcópula C_{UV} del modelo (U^m, V^m) se mantiene bajo la medida neutral al riesgo $\tilde{\mathbb{P}}$, ya que la transformación es de tipo creciente⁴.

Ejemplo 3.14. Para los parámetros $S_0 = 10$, $\mu = 0.001$, $\sigma = 0.01$, $r = 0.0005$ y $N = 20$, en la figura 3.7 se comparan gráficamente los procesos estocásticos del subyacente bajo el modelo tradicional Black-Scholes Merton y el modelo discreto que se acaba de proponer, esto sobre la medida neutral al riesgo $\tilde{\mathbb{P}}$.

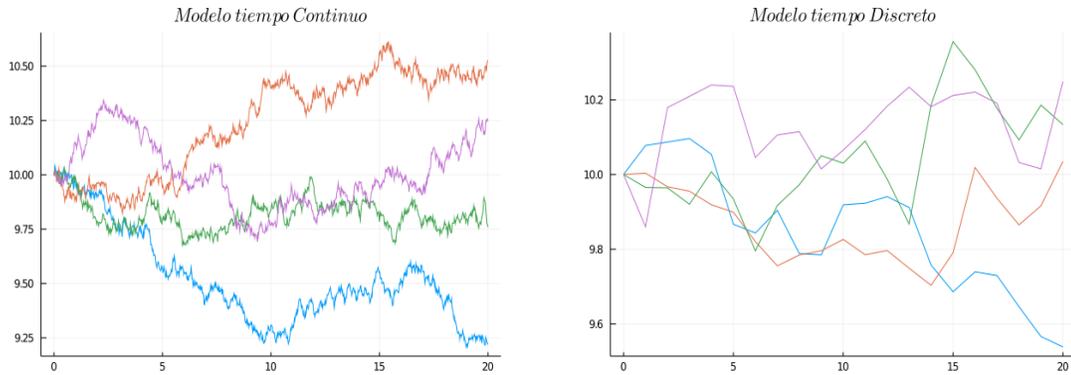


Figura 3.7: Comparación modelos subyacente bajo $\tilde{\mathbb{P}}$.

Como consecuencia de la distribución de (U^m, V^m) bajo la medida $\tilde{\mathbb{P}}$, se puede comprobar fácilmente que la esperanza neutral al riesgo para los activos \mathbf{S}^1 y \mathbf{S}^2 esta centrada en el rendimiento libre de riesgo de sus valores presentes respectivos con respecto a la tasa r . Para comprobar esto se obtiene la esperanza de S_n^1 bajo la medida $\tilde{\mathbb{P}}$,

$$\tilde{\mathbb{E}}(S_n^1) = \tilde{\mathbb{E}}\left(S_0^1 \prod_{j=1}^n U_j\right) = S_0^1 \prod_{j=1}^n \tilde{\mathbb{E}}(U_j) = S_0^1 \prod_{j=1}^n e^r = S_0^1 e^{rn}$$

⁴En este caso es una subcópula ya que el vector está conformado por v.a. discretas.

y el caso de S_n^2 es equivalente.

3.2.2. Distribución discreta del derivado

Una vez se sabe como se comportan los procesos estocásticos \mathbf{S}^1 y \mathbf{S}^2 de forma conjunta sobre la medida $\tilde{\mathbb{P}}$, se puede obtener el precio de cualquier derivado \mathbf{D} que depende de forma conjunta de dichos subyacentes. Sea $G(S_N^1, S_N^2)$ el payoff del derivado a tiempo N y sea $g_D(S_n^1, S_n^2, n)$ el precio del derivado a tiempo n dado los valores S_n^1 y S_n^2 , entonces,

$$g_D(S_n^1, S_n^2, n) = \tilde{\mathbb{E}}\left(D_N \mid \mathcal{F}_n\right) e^{-r(N-n)}$$

donde \mathcal{F}_n representa la filtración inducida a tiempo n para los procesos \mathbf{S}^1 y \mathbf{S}^2 de forma conjunta y claramente $D_N = G(S_N^1, S_N^2)$.

Ejemplo 3.15. (Derivado con un solo subyacente)

Para los parámetros $S_0 = 10$, $\mu = 0.001$, $\sigma = 0.01$, $r = 0.0005$, $N = 20$ y $K = S_0 e^{rN}$, se establece un Call europeo a tiempo N con strike igual a K y tasa libre de riesgo r . En la figuras 3.8 y 3.9 se comparan gráficamente los procesos estocásticos del subyacente y el derivado bajo el modelo tradicional Black-Scholes Merton y el modelo discreto que se acaba de proponer, ambos sobre la medida neutral al riesgo $\tilde{\mathbb{P}}$.

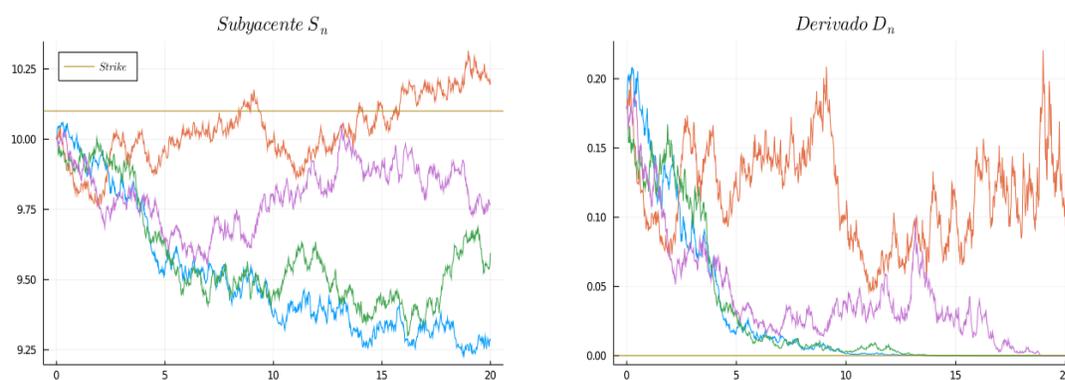


Figura 3.8: Modelo Black-Scholes Merton.

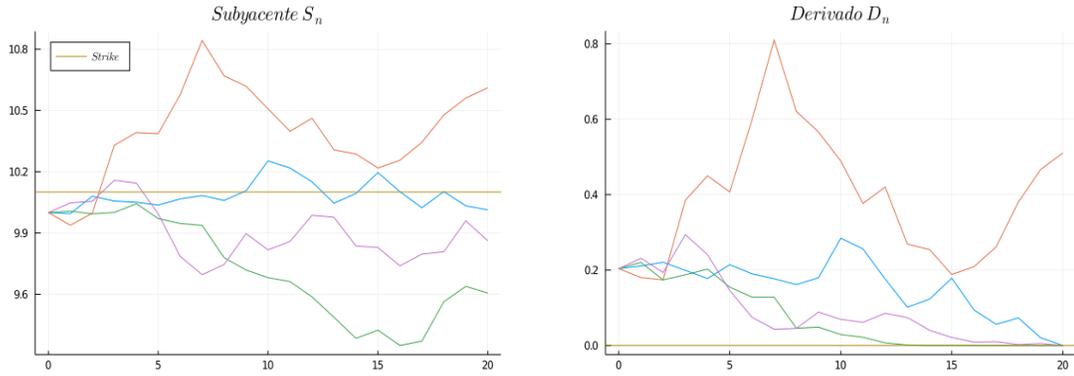


Figura 3.9: Modelo Discreto.

3.2.3. Distribución discreta del portafolio de replicación

Una vez se cuenta con el precio del derivado \mathbf{D} , se desea crear una estrategia de replicación para el derivado en cuestión, esto con el objetivo de poder replicar su valor mediante activos financieros disponibles en el mercado. Teóricamente, el portafolio de replicación \mathbf{X} está compuesto de la siguiente forma,

$$X_n = \Delta_n^1 S_n^1 + \Delta_n^2 S_n^2 + \Gamma_n$$

donde Δ_n^1 y Δ_n^2 representan las inversiones a tiempo n en los activos \mathbf{S}^1 y \mathbf{S}^2 , respectivamente, mientras que Γ_n representa una inversión libre de riesgo bajo la tasa r . En los casos en los que \mathbf{S} es un árbol binomial (Sección 2.3) o un modelo continuo que supone normalidad (Secciones 2.4-2.5), el portafolio de replicación \mathbf{X} quedo completamente definido, y por lo tanto el derivado \mathbf{D} es replicable.

Como consecuencia de que no existen replications instantaneas en el portafolio de replicación puede no llegar a existir una replicación exacta del derivado \mathbf{D} . Como consecuencia de esto se establece el concepto de superreplicación en la definición 3.16. Para poder llegar a esta definición, se parte intuitivamente de lo desarrollado en la secciones 2.4-2.6 y se intenta replicar una opción europea tradicional \mathbf{V} cualquiera.

Si T es la madurez del derivado, se supone que $\Pi = \{t_0, t_1, \dots, t_n\}$ con $0 = t_0 < t_1 < \dots < t_n = T$ representa los periodos de rebalanceo disponibles en el mercado.

Recordando la ecuación diferencial 2.66, se tiene que si \mathbf{V} representa un derivado financiero con un subyacente \mathbf{S} cualquiera, se debe cumplir lo siguiente,

$$dV_t = \Delta_t dS_t + r b_t e^{rt} dt = \Delta_t dS_t + r \left(V_t - \Delta_t S_t \right) dt$$

es decir, el rendimiento instantaneo del derivado \mathbf{V} es una combinación lineal de rendimientos instantaneos en el subyacente \mathbf{S} y la tasa libre de riesgo r , permitiendo así su replicación en el mercado. Con estos puntos en consideración, se crea la siguiente estrategia de replicación para el derivado \mathbf{V} ,

Definición 3.16. *Sea \mathbf{V} un derivado financiero disponible en el mercado con subyacente \mathbf{S} . Se define el proceso estocástico \mathbf{X} como sigue,*

$$X_{t_k} = X_0 + \sum_{j=0}^{k-1} \left\{ \Delta_{t_j} (S_{t_{j+1}} - S_{t_j}) + r (V_{t_j} - \Delta_{t_j} S_{t_j}) (t_{j+1} - t_j) \right\}$$

donde $\Pi = \{t_0, t_1, \dots, t_n\}$ con $0 = t_0 < t_1 < \dots < t_n = T$ son los periodos de rebalanceo disponibles en el mercado. Al proceso \mathbf{X} se le conoce de las siguientes formas:

1) Como una estrategia de superreplicación a condición final si se cumple que,

$$X_T \geq V_T$$

2) Como una estrategia de superreplicación completa si se cumple que,

$$X_{t_k} \geq V_{t_k}$$

para todo t_k .

3) Como una estrategia de superreplicación a condición final a nivel α de confianza si se cumple que,

$$\mathbb{P}(X_T - V_T > 0) \geq \alpha$$

4) Como una estrategia de superreplicación completa a nivel α de confianza si se cumple que,

$$\mathbb{P}(X_{t_k} - V_{t_k} > 0) \geq \alpha$$

para todo t_k .

El término Δ_{t_j} se obtiene mediante la delta de Black-Scholes Merton de la sección 2.6. Es decir,

$$\Delta_{t_j} = \text{Normal}(d_+(T - t_j, S_{t_j}))$$

Observación 3.16.1. En el caso de que $\|\Pi\| \rightarrow 0$ se cumple lo siguiente para \mathbf{X} ,

$$X_t = X_0 + \int_0^t \Delta_u dS_u + \int_0^t r (V_u - \Delta_u S_u) du$$

o escrito de forma diferencial,

$$dX_t = \Delta_t dS_t + r (V_t - \Delta_t S_t) dt$$

dicha ecuación es equivalente a la ecuación diferencial 2.66, por lo que si $X_0 = V_0$, entonces \mathbf{X} es un portafolio de replicación completo para \mathbf{V} .

La observación 3.16.1 aclara que conforme $\|\Pi\| \rightarrow 0$ (es decir, se tiene acceso a mayores periodos de rebalanceo para el portafolio de replicación \mathbf{X}), un mejor ajuste se hace al derivado \mathbf{V} , lo cual es un resultado lógico. La siguiente definición con su respectivo teorema dan un algoritmo práctico para desarrollar el portafolio descrito en la definición 3.16.

Definición 3.17. Sea \mathbf{I} el siguiente proceso estocástico a tiempo discreto sobre Π ,

$$I_{t_k} = V_0 + \sum_{j=0}^{k-1} \left\{ \Delta_{t_j} (S_{t_{j+1}} - S_{t_j}) + r (V_{t_j} - \Delta_{t_j} S_{t_j}) (t_{j+1} - t_j) \right\}$$

al portafolio \mathbf{I} se le conoce como portafolio canónico del derivado \mathbf{V} .

Teorema 3.18. Sea \mathbf{I} el portafolio canónico del derivado \mathbf{V} y sea \mathbf{X} el proceso estocástico definido en la definición 3.16, entonces se cumple lo siguiente:

- 1) \mathbf{X} es una estrategia de superreplicación a condición final a nivel α de confianza si se cumple que,

$$X_0 = I_0 - F_{Y_T}^{-1}(1 - \alpha) = V_0 - F_{Y_T}^{-1}(1 - \alpha)$$

- 2) \mathbf{X} es una estrategia de superreplicación completa a nivel α de confianza si se cumple que,

$$X_0 = I_0 - \min_{t_k} F_{Y_{t_k}}^{-1}(1 - \alpha) = V_0 - \min_{t_k} F_{Y_{t_k}}^{-1}(1 - \alpha)$$

donde $F_{Y_{t_k}}^{-1}$ es la distribución inversa para la v.a. $Y_{t_k} = V_{t_k} - I_{t_k}$.

Demostración. Para el caso de 1) se tiene que $X_T = I_T - F_{Y_T}^{-1}(1 - \alpha)$ por lo que,

$$\begin{aligned}\mathbb{P}(X_T - V_T > 0) &= \mathbb{P}(I_T - V_T > F_{Y_T}^{-1}(1 - \alpha)) = 1 - \mathbb{P}(I_T - V_T \leq F_{Y_T}^{-1}(1 - \alpha)) \\ &= 1 - \mathbb{P}(Y_T \leq F_{Y_T}^{-1}(1 - \alpha)) = 1 - (1 - \alpha) = \alpha\end{aligned}$$

cumpléndose así que \mathbf{X} una estrategia de superreplicación completa a nivel α de confianza. Para el caso de 2) se tiene lo siguiente,

$$\begin{aligned}\mathbb{P}(X_{t_k} - V_{t_k} > 0) &= \mathbb{P}(I_{t_k} - V_{t_k} > \min_{t_j} F_{Y_{t_j}}^{-1}(1 - \alpha)) \geq \mathbb{P}(I_{t_k} - V_{t_k} > F_{Y_{t_k}}^{-1}(1 - \alpha)) \\ &= 1 - \mathbb{P}(I_{t_k} - V_{t_k} \leq F_{Y_{t_k}}^{-1}(1 - \alpha)) = 1 - \mathbb{P}(Y_{t_k} \leq F_{Y_{t_k}}^{-1}(1 - \alpha)) \\ &= 1 - (1 - \alpha) = \alpha\end{aligned}$$

□

Algoritmo 3.18.1. (*Simulación del portafolio canónico de un derivado.*) Supóngase que \mathbf{V} es una opción europea cualquiera con su respectivo activo subyacente \mathbf{S} . Se desea generar una muestra de tamaño n para el portafolio canónico del derivado \mathbf{V} donde $\Pi = \{t_0, t_1, \dots, t_n\}$ y T es la fecha de madurez del derivado, para esto se siguen los siguientes pasos:

- 1) Se genera una m.a. $\mathbf{s} = \{{}_1\mathbf{s}, \dots, {}_n\mathbf{s}\}$ de tamaño n para el subyacente \mathbf{S} sobre \mathbb{P} , es decir,

$${}_j\mathbf{s} = ({}_j s_{t_0}, {}_j s_{t_1}, \dots, {}_j s_{t_n})$$

- 2) Si $H(t, x)$ representa la función de precio del derivado a tiempo t , se genera una m.a. $\mathbf{v} = \{{}_1\mathbf{v}, \dots, {}_n\mathbf{v}\}$ de tamaño n para el derivado \mathbf{V} sobre \mathbb{P} , es decir,

$${}_j\mathbf{v} = (H(t_0, {}_j s_{t_0}), H(t_1, {}_j s_{t_1}), \dots, H(t_n, {}_j s_{t_n}))$$

donde ${}_j v_{t_k} = H(t_k, {}_j s_{t_k})$.

- 4) Se genera una m.a. $\mathbf{\Delta} = \{{}_1\mathbf{\Delta}, \dots, {}_n\mathbf{\Delta}\}$ de la siguiente forma,

$${}_j\mathbf{\Delta} = (H_x(t_0, {}_j s_{t_0}), H_x(t_1, {}_j s_{t_1}), \dots, H_x(t_n, {}_j s_{t_n}))$$

donde ${}_j \Delta_{t_k} = H_x(t_k, {}_j s_{t_k})$.

- 3) Se genera una m.a. $\mathbf{w} = \{ {}_1\mathbf{w}, \dots, {}_n\mathbf{w} \}$ de tamaño n para el portafolio canónico \mathbf{I} como sigue,

$${}_jw_{t_k} := {}_jv_{t_0} + \sum_{i=0}^{k-1} \left\{ {}_j\Delta_{t_i} \left({}_js_{t_{i+1}} - {}_js_{t_i} \right) + r \left({}_jv_{t_i} - {}_j\Delta_{t_i} {}_js_{t_i} \right) \right\}$$

Algoritmo 3.18.2. (Obtención y simulación portafolio de superreplicación a condición final a nivel α) Supóngase que \mathbf{V} es una opción europea cualquiera con su respectivo activo subyacente \mathbf{S} . Se desea obtener y generar una muestra de tamaño n para un portafolio de replicación \mathbf{X} que sea una estrategia de superreplicación a condición final a nivel α de confianza para el derivado donde $\Pi = \{t_0, t_1, \dots, t_n\}$ y T es la fecha de madurez del derivado. Para obtener una estimación de X_0 se realiza lo siguiente:

- 1) Se genera una m.a. $(\mathbf{v}, \mathbf{w}) = \{({}_1\mathbf{v}, {}_1\mathbf{w}), \dots, ({}_n\mathbf{v}, {}_n\mathbf{w})\}$ conjunta de tamaño n (mientras mas grande, mejor) para el derivado \mathbf{V} y el portafolio canónico \mathbf{I} tal y como se definió en el algoritmo 3.18.1.

- 2) Se obtiene una m.a. $\mathbf{y} = (y_1, \dots, y_n)$ de tamaño n para la v.a. Y_T como sigue,

$$y_j = {}_jv_{t_n} - {}_jw_{t_n}$$

- 3) Se estima el valor de $F_{Y_T}^{-1}(1 - \alpha)$ mediante la muestra \mathbf{y} .

- 4) Se obtiene el valor de X_0 mediante la siguiente ecuación,

$$X_0 = V_0 - F_{Y_T}^{-1}(1 - \alpha)$$

Una vez se cuenta con la estimación de X_0 , para generar una m.a. $\mathbf{x} = \{ {}_1\mathbf{x}, \dots, {}_n\mathbf{x} \}$ de tamaño n para el portafolio \mathbf{X} se realiza lo siguiente:

- 1) Se genera una m.a. $\mathbf{w} = \{ {}_1\mathbf{w}, \dots, {}_n\mathbf{w} \}$ de tamaño n para el portafolio canónico \mathbf{I} .

- 2) Se resta el valor de $F_{Y_T}^{-1}(1 - \alpha)$ a toda la muestra, es decir,

$${}_jx_{t_k} = {}_jw_{t_k} - F_{Y_T}^{-1}(1 - \alpha)$$

Ejemplo 3.19. Sean Φ_1 , Φ_2 y Φ_3 las funciones de distribución de probabilidades siguientes,

$$\begin{aligned}\Phi_1(t) &= \int_{-\infty}^t \frac{1}{\sqrt{2t\pi}} e^{-\frac{w^2}{2t}} dw \\ \Phi_2(t) &= \int_{-\infty}^t \frac{\Gamma(\frac{11}{2})}{\sqrt{10t\pi} \Gamma(5)} \left(1 + \frac{w^2}{10t}\right)^{-\frac{11}{2}} dw \\ \Phi_3(t) &= \int_{-\infty}^t \frac{\Gamma(2)}{\sqrt{3t\pi} \Gamma(\frac{3}{2})} \left(1 + \frac{w^2}{3t}\right)^{-2} dw\end{aligned}$$

es decir, Φ_1 es una normal estanda, Φ_2 es una distribución t con 10 grados de libertad y Φ_3 es una distribución t con 3 grados de libertad. Para estas dos funciones se crean los movimientos geométricos pseudo-brownianos siguientes,

$$\begin{aligned}{}_1S_t &:= S_0 \exp \left\{ \alpha_1(t) + \sigma_1 W_t \right\} \\ {}_2S_t &:= S_0 \exp \left\{ \alpha_2(t) + \sigma_2 W_t \right\} \\ {}_3S_t &:= S_0 \exp \left\{ \alpha_3(t) + \sigma_3 W_t \right\}\end{aligned}$$

donde ${}_1W_t \sim \sqrt{t}\Phi_1$, ${}_2W_t \sim \sqrt{t}\Phi_2$, ${}_3W_t \sim \sqrt{t}\Phi_3$ y $\alpha_1(t)$, $\alpha_2(t)$ y $\alpha_3(t)$ son escogidos de tal forma que la media es el rendimiento libre de riesgo $S_0 e^{rt}$. Para los tres procesos anteriores, se definen opciones europeas tipo call con el mismo strike y madurez con $\sigma = 0.01$, $r = 0.0005$, $T = 20$ y $K = 10$, las figuras 3.10, 3.11 y 3.12 se ven gráficas para los derivados creados y sus respectivos portafolios de replicación \mathbf{X} que cumplen con ser estrategias de superreplicación a condición final a nivel 0.75.

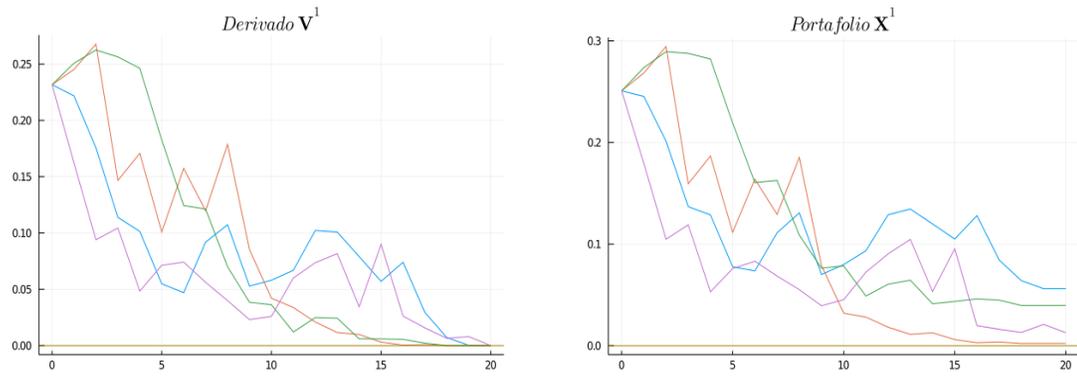


Figura 3.10: Distribuciones de \mathbf{V} y \mathbf{X} bajo Φ_1

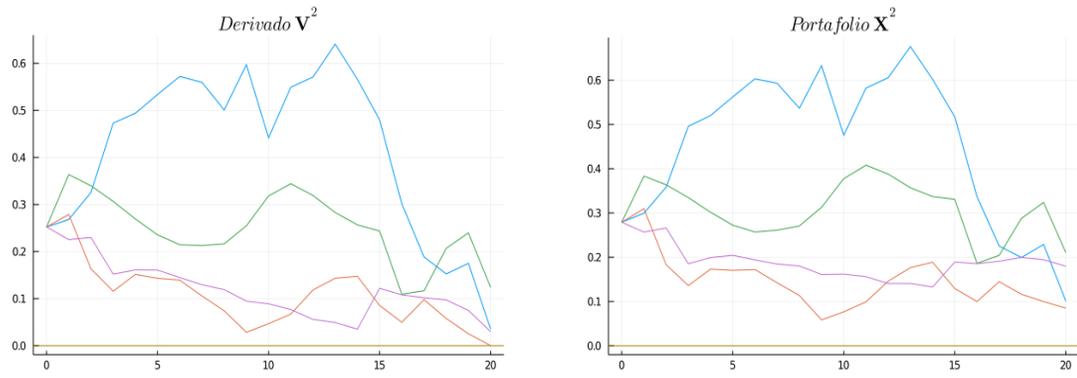


Figura 3.11: Distribuciones de \mathbf{V} y \mathbf{X} bajo Φ_2

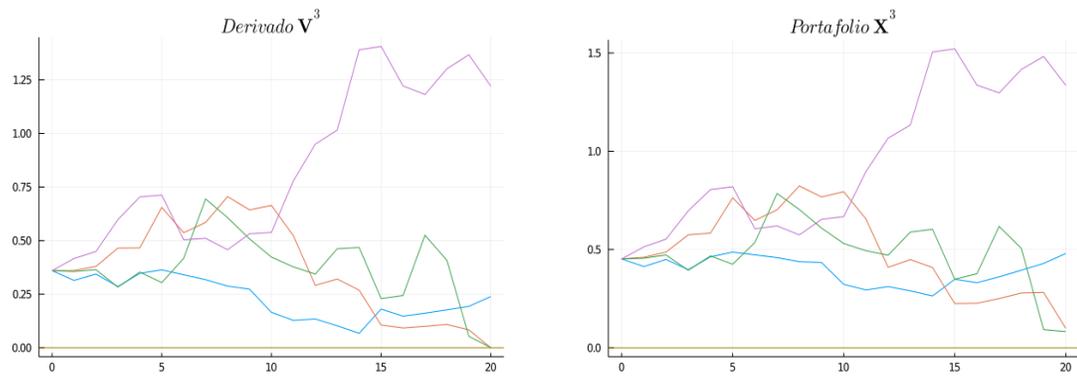


Figura 3.12: Distribuciones de \mathbf{V} y \mathbf{X} bajo Φ_3

En la tabla 3.2 se muestran los precios del portafolio \mathbf{X} para distintos periodos de replicación y las distribuciones Φ_1 , Φ_2 y Φ_3 para el mismo nivel de confianza α . Se pueden notar como la observación 3.16.1 sucede en la práctica al aumentar los periodos de rebalanceo del portafolio de replicación.

Precio V_0	Distribución Φ	Replicación diaria	Replicación 15 veces al día
0.23168	Φ_1	0.25065	0.23610
0.25216	Φ_2	0.27861	0.25817
0.35971	Φ_3	0.41539	0.39343

Cuadro 3.2: Precios portafolio \mathbf{X} .

Ejemplo 3.20. Sean Φ_1 , Φ_2 y Φ_3 las funciones de distribución de probabilidades utilizadas en el ejemplo anterior. Para $m.a.$ de tamaño $m = 300$ para dichas distribuciones denotadas como $\Theta_1 = \{1\theta_1, \dots, 1\theta_m\}$, $\Theta_2 = \{2\theta_1, \dots, 2\theta_m\}$ y $\Theta_3 = \{3\theta_1, \dots, 3\theta_m\}$, se definene los activos subyacentes a tiempo discreto como siguen,

$$S_n^1 = S_0 \prod_{j=1}^n U_j \quad n \in \{1, \dots, N\}$$

$$S_n^2 = S_0 \prod_{j=1}^n V_j \quad n \in \{1, \dots, N\}$$

$$S_n^3 = S_0 \prod_{j=1}^n W_j \quad n \in \{1, \dots, N\}$$

Para los parámetros $S_0 = 10$, $\mu = 0.001$, $\sigma = 0.01$, $r = 0.0005$, $N = 20$ y $K = S_0 e^{rN}$, se establecen opciones europeas tipo Call a tiempo N con strike igual a K y tasa libre de riesgo r , esto para cada uno de los subyacentes descritos anteriormente. En la figuras 3.13, 3.14 y 3.15 se ven gráficas para los derivados creados y sus respectivos portafolios de replicación \mathbf{X} que cumplen con ser estrategias de superreplicación a condición final a nivel 0.75.

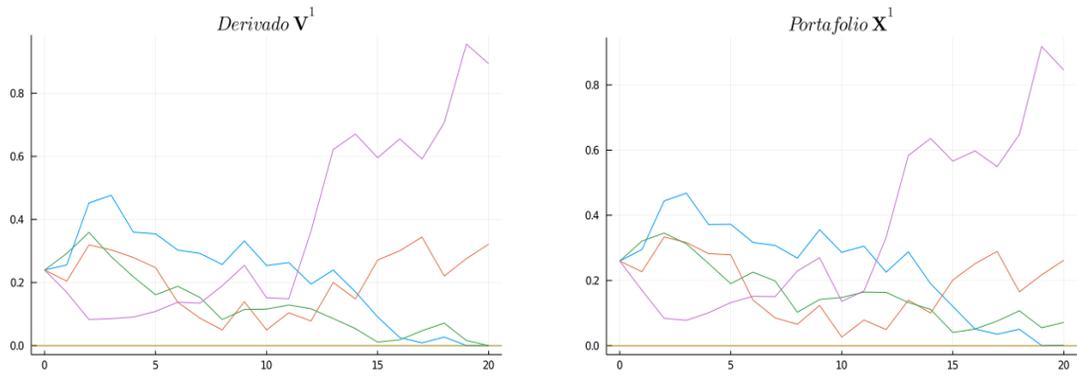


Figura 3.13: Distribuciones de \mathbf{V} y \mathbf{X} bajo Φ_1

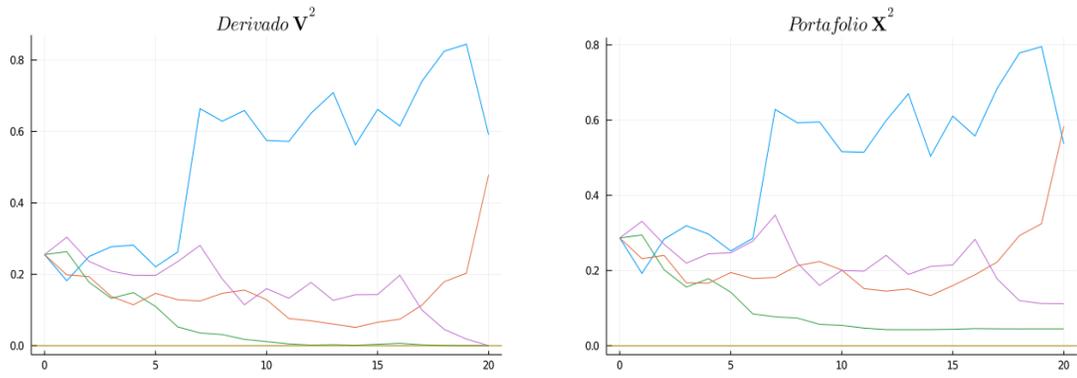


Figura 3.14: Distribuciones de \mathbf{V} y \mathbf{X} bajo Φ_2

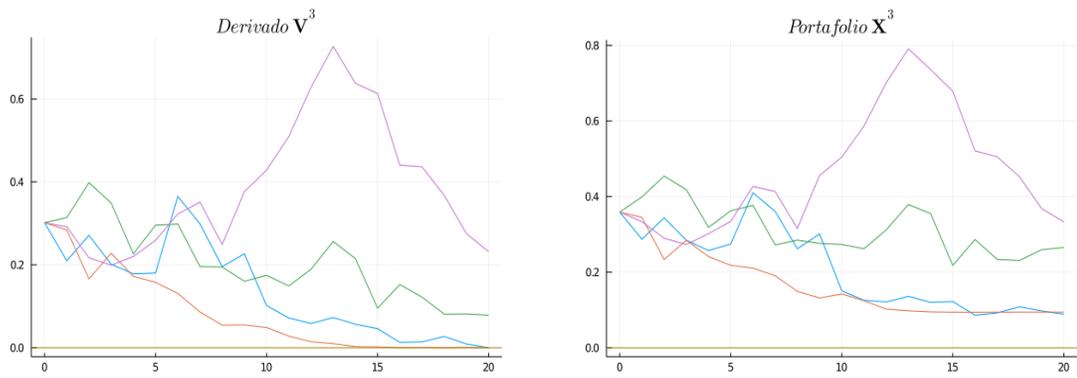


Figura 3.15: Distribuciones de \mathbf{V} y \mathbf{X} bajo Φ_3

En la tabla 3.3 se muestran los precios del portafolio \mathbf{X} para distintos periodos de replicación y los distintos subyacentes $\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3$ para el mismo nivel de confianza $\alpha = 0.75$ y tamaños de muestra $m = 300$. Se pueden notar como la observación 3.16.1 sucede en la práctica al aumentar los periodos de rebalanceo del portafolio de replicación.

Distribucion Φ	Replicación diaria		Replicación 5 veces al día	
	Precio V_0	Precio X_0	Precio V_0	Precio X_0
Φ_1	0.22086	0.23672	0.22467	0.23486
Φ_2	0.23771	0.25998	0.24576	0.25819
Φ_3	0.32541	0.39149	0.30964	0.32857

Cuadro 3.3: Precios portafolio \mathbf{X} .

3.3. Análisis de datos y resultados

En este capítulo se realiza el ajuste semi-paramétrico del modelo subyacente a utilizar, además se plantea y valua el derivado propuesto mediante la metodología desarrollada en el capítulo 3 así como el método de valuación Black-Scholes-Merton.

3.3.1. Ajuste de la información

La información ha utilizar es los log-rendimientos del Índice de Precios y Cotizaciones (de ahora en adelante denotado como IPC) y los log-rendimientos del tipo de cambio dolar estadounidense peso mexicano (de ahora en adelante denotado como tipo de cambio USD/MXN). Los log-rendimientos son obtenidos mediante los precios de cierre del IPC y el promedio del precio de cierre de venta y compra para el tipo de cambio USD/MXN; además, dicha información tiene temporalidad semanal para los periodos entre 10/01/2000 hasta 31/12/2021⁵

Con esta información, se obtienen los procesos estocásticos $\mathbf{R}^1 = (R_t^1)_{1 \leq t \leq 5485}$, $\mathbf{R}^2 = (R_t^2)_{1 \leq t \leq 5485}$ siguientes,

$$R_t^1 := \log \left(\frac{\text{IPC}_t}{\text{IPC}_{t-7}} \right) \quad R_t^2 := \log \left(\frac{\text{USD/MXN}_t}{\text{USD/MXN}_{t-7}} \right)$$

⁵La información fue obtenida de forma gratuita mediante el portal del Banco de México. Fecha de consulta: 19/01/2021.

donde $\mathbf{IPC} := (IPC_t)_{1 \leq t \leq T}$ y $\mathbf{USD/MXN} := (USD/MXN_t)_{1 \leq t \leq T}$ son los procesos estocásticos que representan los precios de cierre antes mencionados y t representa el t -ésimo día iniciando en 10/01/2000 hasta 31/12/2021.

Dependencia Serial

Es importante considerar la dependencia serial para los procesos estocásticos \mathbf{R}^1 y \mathbf{R}^2 , ya que considerar dichos procesos estocásticos como muestras aleatorias sin realizar pruebas de ningún tipo en la información pueden provocar problemas en ajustes futuros. En las figuras 3.16 y 3.17 se analizan la cópula empírica de la dependencia serial a un periodo para el IPC y el tipo de cambio USD/MXN; esto se hace mediante un mapa de calor de la cópula empírica y una gráfica de la diagonal empírica comparada con las diagonales de la cópulas W , M y Π .

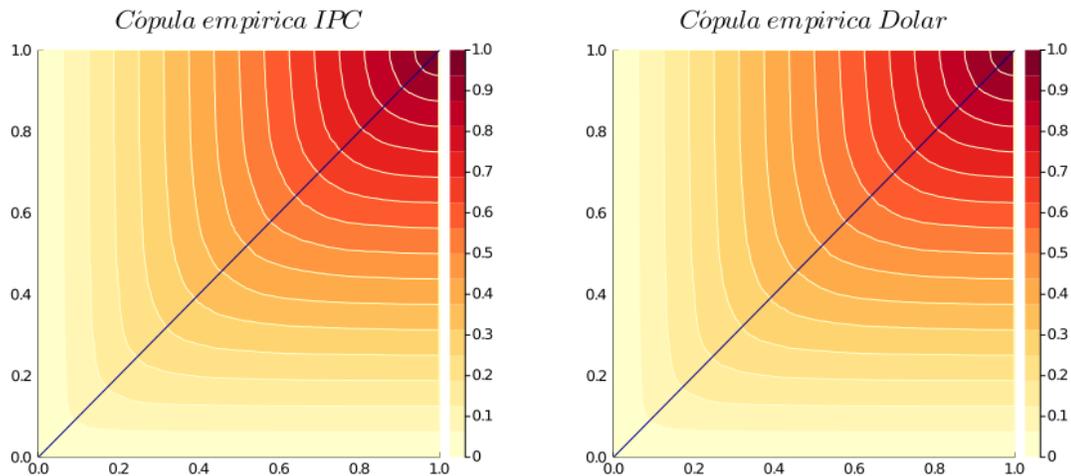


Figura 3.16: Cópula empírica serial del IPC y el tipo de cambio USD/MXN.

Se puede notar claramente como el ajuste empírico expresa un cierto grado de dependencia serial y este resultado visual se complementa con la tabla 3.4 la cual establece las estimaciones respectivas de la Rho de Spearman y la Sigma de Schewiezer, en ambos casos se tiene que $\rho \approx \sigma$ por lo tanto se considera dependencia monótona creciente.

	Rho de Spearman	Sigma de Schweizer
IPC	0.85872	0.85872
Dolar	0.81803	0.81803

Cuadro 3.4: Concordancia y dependencia

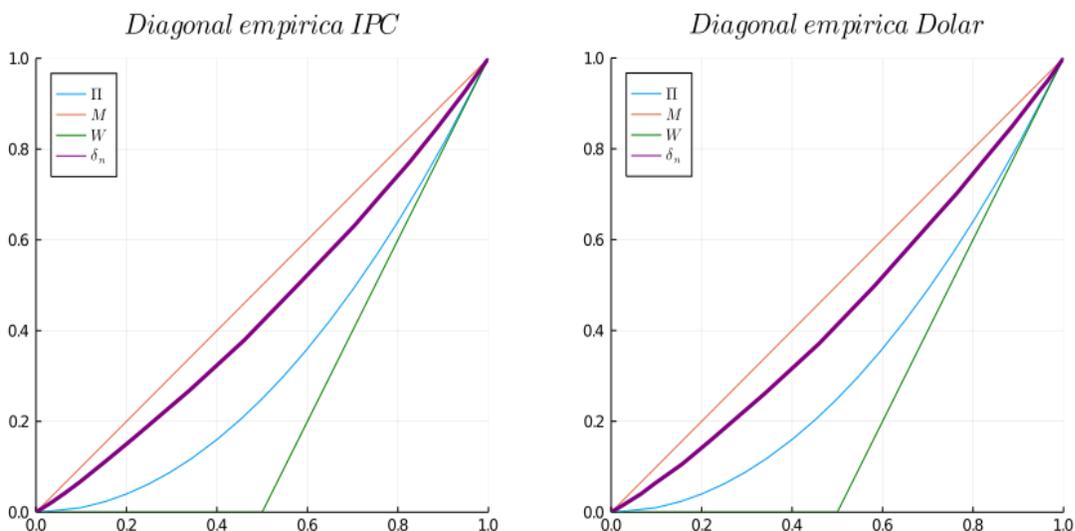


Figura 3.17: Diagonal empírica

	C	C^H	C^V
IPC	1.0	1.0	1.0
Dolar	1.0	1.0	1.0

Cuadro 3.5: P-value prueba de simetría

En la tabla 3.5 se muestran los p-values de la prueba de simetría para las cópulas C y las rotaciones horizontales y verticales C^H y C^V . Por lo que las pruebas de bondad y ajuste que se van a implementar a continuación solo van a tomar en consideración familias de cópulas simétricas, en particular la familia arquimediana de cópulas es una buena opción para realizar este tipo de ajustes (por construcción, toda cópula arquimediana es simétrica). Por último, en las tablas 3.6 y 3.7 se tienen los p-values de bondad y ajuste para la cópula en el caso de dependencia serial y se tienen como

mejores ajustes los siguientes⁶,

$$C^1(u, v) = Frank(u, v | 10.4041)$$

$$C^2(u, v) = Frank(u, v | 9.98985)$$

Cópula	Familia	Parámetro	P-value
C	Copula 17	14.418	0.16491
C^H	Frank	-10.4041	0.29068
C^V	Frank	-9.04749	0.004839
C	Normal	0.8389	0.004

Cuadro 3.6: Prueba de Bondad y Ajuste para el IPC

Cópula	Familia	Parámetro	P-value
C	Gumbel Hougard	3.166	0.51144
C^H	Frank	-9.86818	0.658
C^V	Frank	-9.98985	0.7818
C	Normal	0.8525	0.002

Cuadro 3.7: Prueba de Bondad y Ajuste para el Dolar

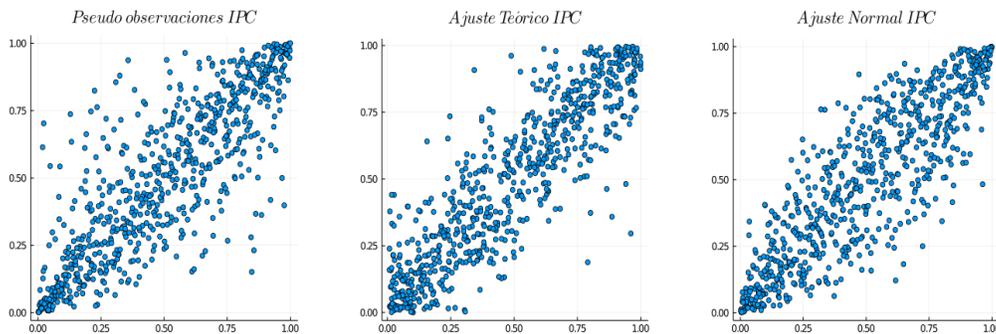


Figura 3.18: Mapas de dispersión IPC

⁶Las cópulas 13 y 17 mencionadas en las siguientes tablas fueron sacadas de la tabla en Nelsen *op. cit.* en el capítulo 4.3, página 116 con los números 13 y 17, respectivamente.

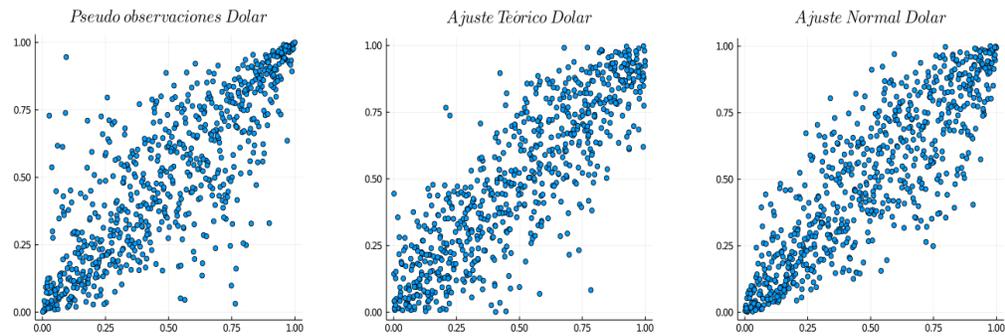


Figura 3.19: Mapas de dispersión Dolar

además, en las figuras 3.18 y 3.19 se comparan los ajustes teóricos C^1 y C^2 con los ajustes supuestos en el modelo Black-Scholes Merton en conjunto con pseudo observaciones de las cópulas que modelan la dependencia serial⁷.

Dependencia IPC y el tipo de cambio USD/MXN

Para el caso de la dependencia entre el IPC y el tipo de cambio USD/MXN, en las figuras 3.20 se ven la copula y diagonales empíricas respectivas y se nota como la diagonal empírica pasa por debajo de la diagonal producto $\Pi(u, u) = u^2$.

⁷Para más información en el concepto de pseudo observaciones, véase Genest *op. cit.*

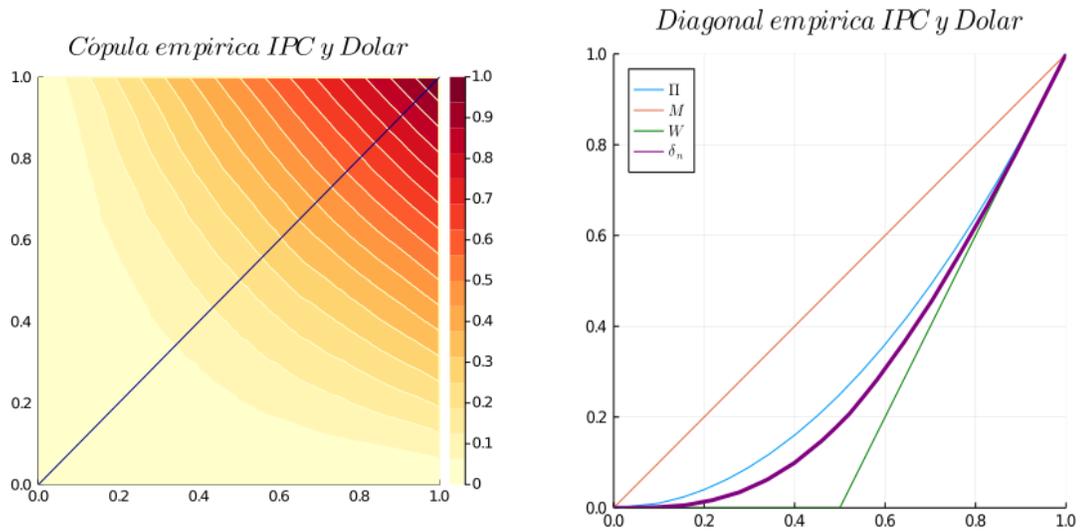


Figura 3.20: Cópula y diagonal empírica IPC y Dolar

En las tablas 3.8 y 3.9 se tienen los valores muestrales de la Rho de Spearman y Sigma de Schweizer además de los p-values para la prueba de simetría. Gracias a estos valores se concluye que la dependencia de C es de tipo monótona decreciente (ya que $-\rho \approx \sigma$) y existe cierta indecisión a la hora de suponer simetría o asimetría, por lo tanto se realizan dos ajustes (uno simétrico y otro asimétrico) a los datos.

Rho de Spearman	Sigma de Schweizer
-0.32518	0.32518

Cuadro 3.8: Concordancia y dependencia entre el IPC y Dolar

C	C^H	C^V
0.3219	0.1410	0.1410

Cuadro 3.9: P-value prueba de simetría entre el IPC y Dolar

Ajuste simétrico

Para el ajuste simétrico la tabla 3.10 muestra los p-values de la prueba de bondad y ajuste mencionada en Genest *op. cit.* y se tiene que el mejor ajuste para C es el siguiente,

<i>Copula</i>	<i>Familia</i>	<i>Parametro</i>	<i>P – value</i>
C	<i>Frank</i>	-3.0038	0.63145
C^H	<i>Copula 17</i>	2.96971	0.82045
C^V	<i>Gumbel Houggard</i>	1.43576	0.19035
C	<i>Normal</i>	-0.3895	0.003

Cuadro 3.10: Prueba de Bondad y Ajuste entre el IPC y Dolar (Usando cópula simétrica)

$$C(u, v) = u - Copula17(u, 1 - v | 2.96971) \tag{3.7}$$

En la figura 3.21 se muestra una submuestra de tamaño 1000 de las pseudo-observaciones para dicho ajuste, además del ajuste bajo el modelo Black-Scholes Merton y el ajuste teórico C .

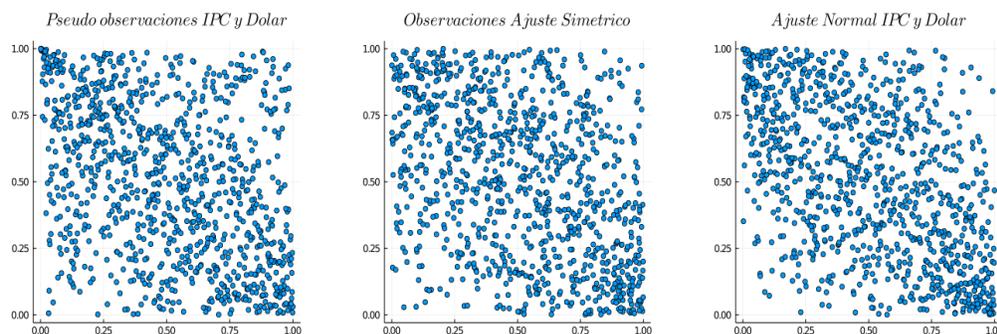


Figura 3.21: Mapas de dispersión IPC y Dolar

Ajuste asimétrico

Para el ajuste asimétrico, tomando $0 < \alpha < 1$, la cópula C se ajuste con tres cópulas simétricas C^1, C^2 y C^3 de las siguientes dos formas posibles,

$$C(u, v) = \alpha \left(C^1 \underset{\theta}{\otimes}^H C^2 \right) (u, v) + (1 - \alpha) C^3(u, v)$$

$$C(u, v) = \alpha \left(C^1 \underset{\theta}{\otimes}^V C^2 \right) (u, v) + (1 - \alpha) C^3(u, v)$$

es decir, se logra un ajuste asimétrico con componentes simétricos los cuales pueden ser ajustados facilmente mediante los resultados de Genest *op. cit.*, esto se realiza de la siguiente forma:

- 1) Dividir la m.a.obs. original $(\mathbf{r}^1, \mathbf{r}^2)$ en submuestras $(\mathbf{r}_1^1, \mathbf{r}_1^2)$, $(\mathbf{r}_2^1, \mathbf{r}_2^2)$, $(\mathbf{r}_3^1, \mathbf{r}_3^2)$ que permitan p-value's altos en las pruebas de simetría individuales (se desea mayor a 0.5).
- 2) Realizar un ajuste asimétrico a la m.a.obs. original $(\mathbf{r}^1, \mathbf{r}^2)$ como alternativa al ajuste teórico simétrico de la subsección anterior. Esto mediante ajustes simétricos C^1 , C^2 y C^3 aplicados a las submuestras $(\mathbf{r}_1^1, \mathbf{r}_1^2)$, $(\mathbf{r}_2^1, \mathbf{r}_2^2)$, $(\mathbf{r}_3^1, \mathbf{r}_3^2)$, respectivamente.

Mediante este argumento, en las tablas 3.11, 3.12 y 3.13 se tienen los ajustes de las cópulas C^1 , C^2 y C^3 , además de θ y α para las cópulas C , C^H y C^V , respectivamente,

<i>Copula</i>	<i>Familia</i>	<i>Parametro</i>	<i>P – value</i>	θ tipo <i>H</i>	α
C^1	<i>AMH</i>	-0.617944	0.64935		
C^2	<i>Clayton</i>	-0.0819845	0.64682	0.503	0.54372
C^3	<i>Frank</i>	-3.58523	0.66454		

Cuadro 3.11: Prueba de Bondad y Ajuste para C (Usando cópula asimétrica)

<i>Copula</i>	<i>Familia</i>	<i>Parametro</i>	<i>P – value</i>	θ tipo <i>H</i>	α
C^1	<i>Copula 13</i>	1.29074	0.26208		
C^2	<i>AMH</i>	0.44965	0.34724	0.6	0.2334
C^3	<i>Frank</i>	2.95629	0.93176		

Cuadro 3.12: Prueba de Bondad y Ajuste para C^H (Usando cópula asimétrica)

<i>Copula</i>	<i>Familia</i>	<i>Parametro</i>	<i>P – value</i>	θ tipo <i>H</i>	α
C^1	<i>FGM</i>	0.374985	0.083		
C^2	<i>AMH</i>	0.3819	0.6688	0.3	0.3009
C^3	<i>Copula 17</i>	3.90791	0.39054		

Cuadro 3.13: Prueba de Bondad y Ajuste para C^V (Usando cópula asimétrica)

Con esta información, se pueden crear los posibles ajustes asimétricos para la muestra $(\mathbf{r}^1, \mathbf{r}^2)$ y la figura 3.22 muestra observaciones de tamaño 1000 de dichos modelos. En la tabla 3.14 se tienen los p-values promedio (promedio ponderado por tamaño de submuestra) y el p-value completo de los ajustes asimétricos correspondientes.

<i>Ajuste</i>	P-value promedio	P-value completo
C	0.65562	0.00008
C^H	0.78129	0.13732
C^V	0.40382	0.17894

Cuadro 3.14: P-value prueba de bondad y ajuste entre el IPC y Dolar

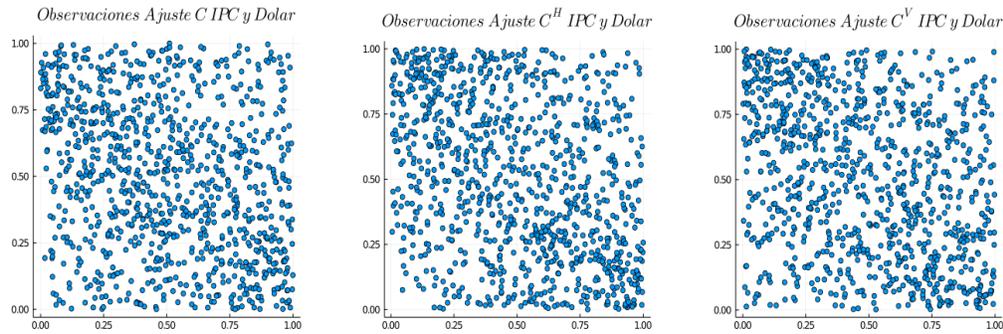


Figura 3.22: Mapas de dispersión IPC y Dolar

Ya que los p-values completos no superan al 0.2, se realiza una nueva combinación lineal convexa con estos ajustes junto con el ajuste simétrico de la subsección anterior, el mejor ajuste de este tipo tiene un p-value del 0.385 y está dado de la siguiente forma,

$$C(u, v) = 0.3 \bar{C}^H(u, v) + 0.7(u - Copula17(u, 1 - v | 2.96971)) \quad (3.8)$$

donde \bar{C}^H es el ajuste asimétrico de C^H dada en la tabla 3.13. En la figura 3.23 se tiene el ajuste asimétrico final dado por la ecuación 3.8 y se compara con las pseudo-observaciones de los datos y el ajuste bajo el modelo Black-Scholes Merton.

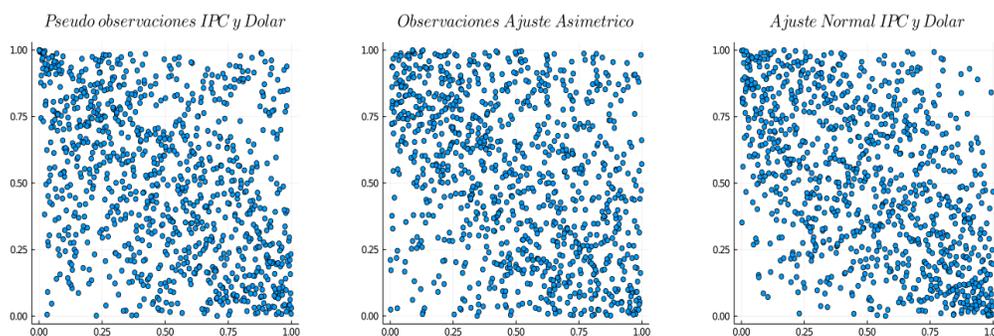


Figura 3.23: Mapas de dispersión Dolar

3.3.2. Derivados propuestos y resultados

Una vez se cuenta con el ajuste paramétrico de la cópula para los log-rendimientos deseados, se establecen los derivados que se desean valorar mediante el método propuesto en el capítulo 3 y el modelo tradicional, su comparación y resultados de importancia. Para poder proponer derivados de interés, se inicia esta sección proponiendo un portafolio de inversión tradicional de la siguiente forma,

$$P_\alpha = \alpha X + (1 - \alpha)Y \quad (3.9)$$

donde X y Y representan los log-rendimientos de los subyacentes de interés, es decir, el IPC y el tipo de cambio USD/MXN, respectivamente. Además, se supone $\alpha \in]0, 1[$ para facilitar la optimización del portafolio⁸.

En la metodología tradicional⁹, el portafolio 3.9 se puede optimizar de distintas formas, una de ellas es el llamado “portafolio de mínimo riesgo”, el cual minimiza la varianza con respecto a α , es decir,

$$\min_{\alpha \in]0, 1[} \mathbb{V}(P_\alpha) \quad (3.10)$$

Para poder realizar dicha optimización, se necesita conocer la distribución bivariada (X, Y) . Como alternativa a esta optimización, se propone una minimización de la siguiente cantidad,

$$\min_{\alpha \in]0, 1[} \mathbb{P}(P_\alpha < 0) \quad (3.11)$$

⁸Esto se establece con el objetivo de no permitir inversiones cortas en el portafolio, para las cuales no siempre se tiene acceso en el mercado.

⁹Para mas información, véase Markowitz *op. cit.*

dicha propuesta se da como consecuencia de que la cantidad $\mathbb{P}(P_\alpha < 0)$ representa de mejor forma la situación no favorable para el portafolio en cuestión, mientras que la varianza considera dañinas de igual forma las desviaciones a la izquierda y derecha de la esperanza, lo cual en la práctica no es apropiado.

Para obtener estos portafolios, primeramente se realizan ajustes univariados a la información. En las tablas 3.15 y 3.16 se tienen los mejores ajustes univariados para los log-rendimientos semanales del IPC y el tipo de cambio USD/MXN.

Familia	Parámetros	P-value
Student	$\mu_1 = 0.00041$ $\sigma_1 = 0.02185$ $\nu_1 = 4.54122$	0.59888
Normal	$\mu_1 = 0.00041$ $\sigma_1 = 0.02922$	0.32426

Cuadro 3.15: Prueba de Bondad y Ajuste IPC univariada.

Familia	Parámetros	P-value
Normal	$\mu_1 = 0.00040$ $\sigma_1 = 0.02379$	0.97724
Student	$\mu_1 = 0.00040$ $\sigma_1 = 0.01588$ $\nu_1 = 3.84373$	0.20862

Cuadro 3.16: Prueba de Bondad y Ajuste USD/MXN univariada.

Para este caso particular, es claro que el ajuste bivariado entre el IPC y el tipo de

cambio USD/MXN no cumple la suposición del modelo tradicional Black-Scholes Merton, esto por dos razones:

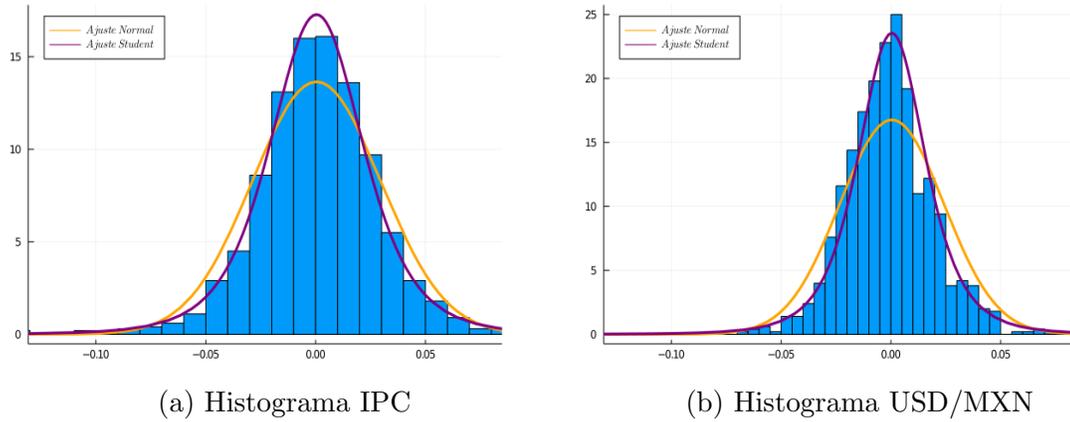


Figura 3.24: Histogramas y ajustes teóricos univariados

- 1) Se pudo encontrar un ajuste mediante la distribución t-Student el cual es mejor¹⁰ para el caso del log-rendimiento del IPC.
- 2) Los ajustes simétrico y asimétrico para la cópula dado por las ecuaciones 3.7 y 3.8, respectivamente, son mejores ajustes que el ajuste supuesto por el modelo tradicional (cópula normal).

Histogramas de las m.a.obs. univariadas con los respectivos ajustes teóricos se pueden ver en las figuras 3.24a y 3.24b, respectivamente.

Una vez se tienen los ajustes univariados de X y Y en conjunto con los resultados para la cópula C_{XY} en la sección anterior, mediante simulación, se obtienen los valores α que optimizan 3.10 y 3.11. Estos resultados se establecen en la tabla 3.17 y se realizan para las cópulas 3.7 y 3.8.

¹⁰Es mejor en términos del p-value dado por la prueba de Bondad y Ajuste de Anderson-Darling. Para mayor información de dicha prueba, véase Anderson y Darling (1952) y Anderson y Darling (1954).

Cópula	Optimización tradicional	Optimización propuesta
Simétrica	0.42159	0.6351
Asimétrica	0.41982	0.4474

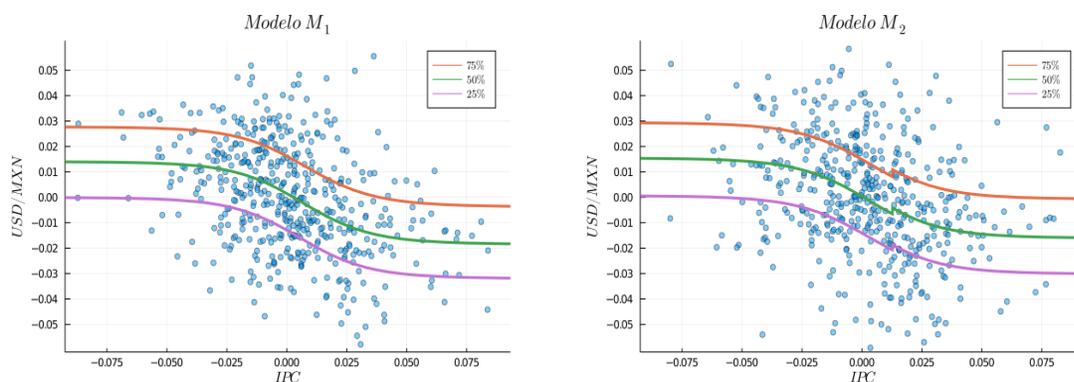
Cuadro 3.17: Pesos de optimización para el portafolio 3.9

Observación 3.20.1. *Los pesos obtenidos en la tabla 3.17 están dadas para los siguientes dos modelos bivariados,*

$$M_1 \equiv F_{XY}^1(x, y) = C_1\left(F_X(x), F_Y(y)\right)$$

$$M_2 \equiv F_{XY}^2(x, y) = C_2\left(F_X(x), F_Y(y)\right)$$

donde $X \sim t_{4.54}(0.00041, 0.02185)$ y $Y \sim Normal(0.00040, 0.02379)$ son las distribuciones univariadas obtenidas en la tablas 3.15 y 3.16 y C_1 y C_2 son las cópulas en las ecuaciones 3.7 y 3.8 (simétrica y asimétrica, respectivamente).

Figura 3.25: Modelos de dispersión y curvas de regresión para M_1 y M_2 .

En la figura 3.25 se tienen gráficos de dispersión para ambos modelos que incluyen curvas de regresión a nivel 25%, 50% (mediana condicional) y 75%; se puede notar como el modelo M_2 espera mejores resultados positivos de la relación entre el IPC y el tipo de cambio USD/MXN, por lo que al crear un portafolio que minimiza el riesgo mediante $\mathbb{P}(P_\alpha < 0)$ tiene sentido que α_2 sea mas cercano a 0.5 que en el caso de α_1 . Las diferencias mas importantes entre C_1 y C_2 son las siguientes:

- 1) *Asimetría de C_2 : La cópula C_1 es una cópula simétrica mientras que C_2 no lo es, por lo que existe intercambialidad para el caso de F^1 (es decir, $F_{XY}^1 = F_{YX}^1$) este no es el caso para F^2 .*

- 2) *El gluing point existente en la cópula C_2* : Recordando la sección 1.8, una *gluing* cópula cualquiera C con *gluing point* θ sufre un cambio abrupto de dependencia en dicho punto, creando modificaciones en el comportamiento del modelo al que pertenece. En nuestro ejemplo, esto se puede ver claramente en la figura 3.25 en los cambios abruptos en las curvas de regresión a cualquier nivel. En la figura 3.26 se puede ver las diferencias entre C_1 y C_2 .

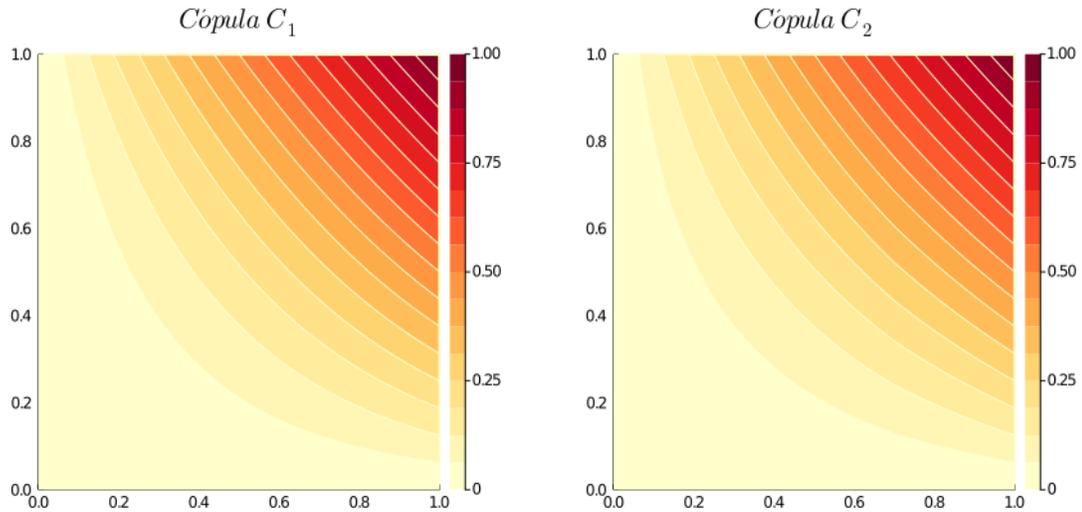


Figura 3.26: Conjuntos de nivel para las cópulas C_1 y C_2 .

Ahora bien, si se tiene (o se desea tener) una posición **larga** en el portafolio P_α (es decir, existe ganancia cuando el portafolio sube de valor), se podría desear estar cubierto ante rendimientos por debajo de un valor r_0 fijo, esto para asegurar una ganancia mínima, se proponen dos formas de realizar esto:

- 1) Si se desea tener una ganancia mínima y beneficiarse de la incertidumbre de P_α se plantea una posición larga en el derivado con el siguiente payoff,

$$\max \left\{ e^{P_\alpha}, e^{r_0} \right\} \quad (3.12)$$

- 2) Si se desea cubrirse ante pérdidas de P_α cuando ya se cuenta con una posición larga en el portafolio, esto se puede hacer mediante una posición larga en el

derivado con el siguiente payoff,

$$\text{máx} \left\{ e^{r_0} - e^{P_\alpha}, 0 \right\} \quad (3.13)$$

Para poder valuar los derivados 3.12 y 3.13 mediante el modelo Black-Scholes Merton se establecen una definición y los siguientes dos resultados,

Definición 3.21. Sea $\mathbf{X} = (X_1, \dots, X_n)$ un v.a. cualquiera, la función generadora de momentos del v.a. \mathbf{X} queda definida como sigue, en el caso de que exista, de la siguiente forma,

$$M_{\mathbf{X}}(t) \equiv M_{\mathbf{X}}(t_1, \dots, t_n) := \mathbb{E} \left(e^{\sum_{i=1}^n t_i X_i} \right)$$

donde $t = (t_1, \dots, t_n)$ y $t_i \in \mathbb{R}$.

Lema 3.22. Sea (X, Y) un v.a. normal multivariado con matriz de medias $\mu = [\mu_x, \mu_y]^T$ y matriz de covarianzas,

$$\Sigma = \begin{bmatrix} \sigma_x^2 & \rho_{xy} \sigma_x \sigma_y \\ \rho_{xy} \sigma_x \sigma_y & \sigma_y^2 \end{bmatrix}$$

Si $Z = a_1 X + a_2 Y$ para valores a_1 y a_2 cualesquiera, entonces se cumple que Z sigue una distribución normal univariada con los siguientes parametros,

$$\begin{aligned} \mathbb{E}(Z) &= A\mu \\ \mathbb{V}(Z) &= A\Sigma A^T \end{aligned}$$

donde $A = [a_1, a_2]$.

Demostración. Denotemos $\mathbf{W} = (X, Y)$, la función generadora de momentos para el v.a. W está dada como sigue,

$$M_{\mathbf{W}}(t_1, t_2) = \mathbb{E} \left(e^{t_1 X + t_2 Y} \right) = \exp \left(t \mu + \frac{1}{2} t \Sigma t^T \right)$$

donde $t = (t_1, t_2)$. Con dicho resultado, la función generadora de momentos para la v.a. $Z = a_1 X + a_2 Y$ es igual a lo siguiente (considerando $A = [a_1, a_2]$),

$$\begin{aligned} M_Z(t) &= \mathbb{E} \left(e^{tZ} \right) = \mathbb{E} \left(e^{t(a_1 X + a_2 Y)} \right) = \mathbb{E} \left(e^{ta_1 X + ta_2 Y} \right) = M_{\mathbf{W}}(ta_1, ta_2) \\ &= \exp \left([ta_1, ta_2] \mu + \frac{1}{2} [ta_1, ta_2] \Sigma [ta_1, ta_2]^T \right) \\ &= \exp \left(t A \mu + \frac{1}{2} t A \Sigma (t A)^T \right) = \exp \left(t (A \mu) + \frac{1}{2} t (A \Sigma A^T) t^T \right) \end{aligned}$$

y como la función generadora de momentos define de forma unívoca a una v.a., la distribución de Z es una normal univariada con media y varianza siguientes,

$$\begin{aligned}\mathbb{E}(Z) &= A\mu \\ \mathbb{V}(Z) &= A\Sigma A^T\end{aligned}$$

□

Teorema 3.23. *Sea (X, Y) un v.a. que represente los log-rendimientos de dos subyacentes a un periodo, además sean \mathbf{V}^1 y \mathbf{V}^2 los derivados financieros dados por las funciones pay-off 3.12 y 3.13, respectivamente, a tiempo T . Es decir,*

$$\begin{aligned}V_T^1 &= \text{máx} \left\{ e^{P_\alpha T}, e^{r_0 T} \right\} \\ V_T^2 &= \text{máx} \left\{ e^{r_0 T} - e^{P_\alpha T}, 0 \right\}\end{aligned}$$

Bajo el modelo clásico Black-Scholes Merton para la valuación de opciones europeas, se tiene que los procesos estocásticos $\mathbf{V}^1 = \{V_t^1\}_{0 \leq t \leq T}$ y $\mathbf{V}^2 = \{V_t^2\}_{0 \leq t \leq T}$ son equivalentes a lo siguiente,

$$\begin{aligned}V_t^1 &= e^{r_0 T} e^{-r(T-t)} \Phi\left(-d_-(T-t, e^{P_\alpha t})\right) + e^{P_\alpha t} \Phi\left(d_+(T-t, e^{P_\alpha t})\right) \\ V_t^2 &= e^{r_0 T} e^{-r(T-t)} \Phi\left(-d_-(T-t, e^{P_\alpha t})\right) - e^{P_\alpha t} \Phi\left(-d_+(T-t, e^{P_\alpha t})\right)\end{aligned}$$

donde Φ es la función de distribución de probabilidades para una Normal estandar,

$$d_\pm(\tau, x) := \frac{1}{\sigma\sqrt{\tau}} \left[\left(r \pm \frac{\sigma^2}{2} \right) \tau + \log x - r_0 T \right]$$

r es la tasa libre de riesgo a un periodo y σ^2 está dada por la siguiente ecuación,

$$\sigma^2 = \alpha^2 \sigma_x^2 + (1 - \alpha)^2 \sigma_y^2 + 2\alpha(1 - \alpha)\sigma_x \sigma_y \rho_{xy}$$

con $\sigma_x^2 = \text{Var}(X)$, $\sigma_y^2 = \text{Var}(Y)$ y $\rho_{xy} = \text{Corr}(X, Y)$.

Demostración. Bajo el modelo tradicional y sobre la medida neutral al riesgo, se tiene que (X, Y) se distribuye como una normal bivariada con los parámetros establecidos en el lema 3.22 y como consecuencia del mismo lema se tiene que $\mathbb{V}(P_\alpha)$ esta dada por la siguiente ecuación,

$$\mathbb{V}(P_\alpha) \equiv \sigma^2 = \alpha^2 \sigma_x^2 + (1 - \alpha)^2 \sigma_y^2 + 2\alpha(1 - \alpha)\sigma_x \sigma_y \rho_{xy}$$

Ahora bien, se tiene la siguiente equivalencia entre V_T^1 y V_T^2 ,

$$V_T^1 = \text{máx} \left\{ e^{P_\alpha T}, e^{r_0 T} \right\} = \text{máx} \left\{ e^{r_0 T} - e^{P_\alpha T}, 0 \right\} + e^{P_\alpha T} = V_T^2 + e^{P_\alpha T} \quad (3.14)$$

por lo que la valuación de V_T^2 es suficiente para obtener la valuación de V_T^1 . El derivado V_T^2 se puede interpretar como una opción europea tipo put con strike $K = e^{r_0 T}$ y con $S_0 = 1$, por lo que V_t^2 está dado por la siguiente ecuación,

$$V_t^2 = e^{r_0 T} e^{-r(T-t)} \Phi \left(-d_-(T-t, e^{P_\alpha t}) \right) - e^{P_\alpha t} \Phi \left(-d_+(T-t, e^{P_\alpha t}) \right)$$

gracias al corolario 2.46.1, donde $d_\pm(\tau, x)$ queda definido como sigue,

$$d_\pm(\tau, x) = \frac{\left(r \pm \frac{\sigma^2}{2}\right)\tau + \log x - \log e^{r_0 T}}{\sigma\sqrt{\tau}} = \frac{1}{\sigma\sqrt{\tau}} \left[\left(r \pm \frac{\sigma^2}{2}\right)\tau + \log x - r_0 T \right]$$

una vez se sabe el valor de V_t^2 , es fácil obtener el valor V_t^1 utilizando la ecuación 3.14,

$$V_t^1 = V_t^2 + e^{P_\alpha t} = e^{r_0 T} e^{-r(T-t)} \Phi \left(-d_-(T-t, e^{P_\alpha t}) \right) + e^{P_\alpha t} \Phi \left(d_+(T-t, e^{P_\alpha t}) \right)$$

quedando así demostrado lo deseado. □

Observación 3.23.1. Si H_1 y H_2 representan las funciones pay-off de los derivados V_1 y V_2 , respectivamente, para los derivados valuados en el teorema 3.23, es decir,

$$\begin{aligned} H_1(t, z, w) &= e^{r_0 T} e^{-r(T-t)} \Phi \left(-d_-(T-t, z^\alpha w^{1-\alpha}) \right) \\ &\quad + z^\alpha w^{1-\alpha} \Phi \left(d_+(T-t, z^\alpha w^{1-\alpha}) \right) \\ H_2(t, z, w) &= e^{r_0 T} e^{-r(T-t)} \Phi \left(-d_-(T-t, z^\alpha w^{1-\alpha}) \right) \\ &\quad - z^\alpha w^{1-\alpha} \Phi \left(-d_+(T-t, z^\alpha w^{1-\alpha}) \right) \end{aligned}$$

donde $V_t^1 = H_1(t, S_t^1, S_t^2)$, $V_t^2 = H_2(t, S_t^1, S_t^2)$ y S_t^1, S_t^2 están definidos como sigue,

$$\begin{aligned} S_t^1 &= \exp \left\{ \left(r - \frac{1}{2} \sigma_1^2 \right) t + \sigma_1 W_t^1 \right\} \\ S_t^2 &= \exp \left\{ \left(r - \frac{1}{2} \sigma_2^2 \right) t + \sigma_2 W_t^2 \right\} \end{aligned}$$

se tiene que ${}_1\Delta_t^1$, ${}_1\Delta_t^2$, ${}_2\Delta_t^1$ y ${}_2\Delta_t^2$ representan las siguientes derivadas parciales,

$$\begin{aligned} {}_1\Delta_t^1 &\equiv \left. \frac{\partial H_1(t, z, w)}{\partial z} \right|_{(z,w)=(S_t^1, S_t^2)} \\ {}_1\Delta_t^2 &\equiv \left. \frac{\partial H_1(t, z, w)}{\partial w} \right|_{(z,w)=(S_t^1, S_t^2)} \\ {}_2\Delta_t^1 &\equiv \left. \frac{\partial H_2(t, z, w)}{\partial z} \right|_{(z,w)=(S_t^1, S_t^2)} \\ {}_2\Delta_t^2 &\equiv \left. \frac{\partial H_2(t, z, w)}{\partial w} \right|_{(z,w)=(S_t^1, S_t^2)} \end{aligned}$$

Para obtener ${}_i\Delta_t^j$ $i, j \in \{1, 2\}$, se utiliza la siguiente parametrización para d_{\pm} ,

$$\begin{aligned} d_{\pm}(\tau, z, w) &= \frac{1}{\sigma\sqrt{\tau}} \left[\left(r \pm \frac{\sigma^2}{2} \right) \tau + \log(z^\alpha w^{1-\alpha}) - r_0 T \right] \\ &= \frac{1}{\sigma\sqrt{\tau}} \left[\left(r \pm \frac{\sigma^2}{2} \right) \tau + \alpha \log z + (1 - \alpha) \log w - r_0 T \right] \end{aligned}$$

por lo que las derivadas parciales de d_{\pm} quedan como sigue,

$$\begin{aligned} \frac{\partial d_{\pm}(\tau, z, w)}{\partial z} &= \frac{\alpha}{z\sigma\sqrt{\tau}} \\ \frac{\partial d_{\pm}(\tau, z, w)}{\partial w} &= \frac{1 - \alpha}{w\sigma\sqrt{\tau}} \end{aligned}$$

si se considera $\phi = \frac{d\Phi}{dx}$, las derivadas parciales de H_1 y H_2 quedan como sigue,

$$\begin{aligned}
\frac{\partial H_1(t, z, w)}{\partial z} &= -\frac{\alpha e^{r_0 T} e^{-r(T-t)}}{z\sigma\sqrt{\tau}} \phi\left(-d_-(T-t, z, w)\right) \\
&\quad + \frac{\alpha z^{\alpha-1} w^{1-\alpha}}{\sigma\sqrt{\tau}} \phi\left(d_+(T-t, z, w)\right) \\
&\quad + \alpha z^{\alpha-1} w^{1-\alpha} \Phi\left(d_+(T-t, z, w)\right) \\
\frac{\partial H_1(t, z, w)}{\partial w} &= -\frac{(1-\alpha) e^{r_0 T} e^{-r(T-t)}}{w\sigma\sqrt{\tau}} \phi\left(-d_-(T-t, z, w)\right) \\
&\quad + \frac{(1-\alpha) z^\alpha w^{-\alpha}}{\sigma\sqrt{\tau}} \phi\left(d_+(T-t, z, w)\right) \\
&\quad + (1-\alpha) z^\alpha w^{-\alpha} \Phi\left(d_+(T-t, z, w)\right) \\
\frac{\partial H_2(t, z, w)}{\partial z} &= -\frac{\alpha e^{r_0 T} e^{-r(T-t)}}{z\sigma\sqrt{\tau}} \phi\left(-d_-(T-t, z, w)\right) \\
&\quad + \frac{\alpha z^{\alpha-1} w^{1-\alpha}}{\sigma\sqrt{\tau}} \phi\left(-d_+(T-t, z, w)\right) \\
&\quad - \alpha z^{\alpha-1} w^{1-\alpha} \Phi\left(-d_+(T-t, z, w)\right) \\
\frac{\partial H_2(t, z, w)}{\partial w} &= -\frac{(1-\alpha) e^{r_0 T} e^{-r(T-t)}}{w\sigma\sqrt{\tau}} \phi\left(-d_-(T-t, z, w)\right) \\
&\quad + \frac{(1-\alpha) z^\alpha w^{-\alpha}}{\sigma\sqrt{\tau}} \phi\left(-d_+(T-t, z, w)\right) \\
&\quad - (1-\alpha) z^\alpha w^{-\alpha} \Phi\left(-d_+(T-t, z, w)\right)
\end{aligned}$$

De forma equivalente al corolario 2.46.2 se concluye lo siguiente,

$$\begin{aligned}
{}_1\Delta_t^1 &= \alpha (S_t^1)^{\alpha-1} (S_t^2)^{1-\alpha} \Phi\left(d_+(T-t, S_t^1, S_t^2)\right) \\
{}_1\Delta_t^2 &= (1-\alpha) (S_t^1)^\alpha (S_t^2)^{-\alpha} \Phi\left(d_+(T-t, z, w)\right) \\
{}_1\Delta_t^1 &= -\alpha (S_t^1)^{\alpha-1} (S_t^2)^{1-\alpha} \Phi\left(-d_+(T-t, S_t^1, S_t^2)\right) \\
{}_1\Delta_t^2 &= -(1-\alpha) (S_t^1)^\alpha (S_t^2)^{-\alpha} \Phi\left(-d_+(T-t, z, w)\right)
\end{aligned}$$

y por lo tanto,

$$\begin{aligned} {}_1\Delta_t^1 S_t^1 + {}_1\Delta_t^2 S_t^2 &= (S_t^1)^\alpha (S_t^2)^{1-\alpha} \Phi\left(d_+(T-t, S_t^1, S_t^2)\right) = {}_1\Delta S_t^3 \\ {}_2\Delta_t^1 S_t^1 + {}_2\Delta_t^2 S_t^2 &= -(S_t^1)^\alpha (S_t^2)^{1-\alpha} \Phi\left(-d_+(T-t, S_t^1, S_t^2)\right) = {}_2\Delta S_t^3 \end{aligned}$$

es decir, la inversión con incertidumbre en el portafolio de replicación es igual a la inversión bajo un nuevo subyacente S_t^3 con incertidumbre definido como sigue,

$$S_t^3 := (S_t^1)^\alpha (S_t^2)^{1-\alpha}$$

y los respectivos deltas ${}_1\Delta$ y ${}_2\Delta$ definidos como sigue,

$$\begin{aligned} {}_1\Delta &:= \left. \frac{\partial H_1(t, s)}{\partial s} \right|_{s=S_t^3} = \Phi\left(d_+(T-t, S_t^3)\right) \\ {}_2\Delta &:= \left. \frac{\partial H_2(t, s)}{\partial s} \right|_{s=S_t^3} = -\Phi\left(-d_+(T-t, S_t^3)\right) \end{aligned}$$

Una vez se cuenta con las deltas del portafolio de replicación para los derivados \mathbf{V}^1 y \mathbf{V}^2 , se necesita construir el portafolio de replicación a tiempo discreto equivalente al desarrollado en el capítulo 3 para dichos derivados. Estos portafolios se construyen de la siguiente forma,

$$\begin{aligned} X_{t_k}^1 &= X_0^1 + \sum_{j=0}^{k-1} \left\{ {}_1\Delta_{t_j}^1 \left(S_{t_{j+1}}^1 - S_{t_j}^1 \right) + {}_1\Delta_{t_j}^2 \left(S_{t_{j+1}}^2 - S_{t_j}^2 \right) \right. \\ &\quad \left. + r \left(V_{t_j} - {}_1\Delta_{t_j}^1 S_{t_j}^1 - {}_1\Delta_{t_j}^2 S_{t_j}^2 \right) \left(t_{j+1} - t_j \right) \right\} \\ X_{t_k}^2 &= X_0^2 + \sum_{j=0}^{k-1} \left\{ {}_2\Delta_{t_j}^1 \left(S_{t_{j+1}}^1 - S_{t_j}^1 \right) + {}_2\Delta_{t_j}^2 \left(S_{t_{j+1}}^2 - S_{t_j}^2 \right) \right. \\ &\quad \left. + r \left(V_{t_j} - {}_2\Delta_{t_j}^1 S_{t_j}^1 - {}_2\Delta_{t_j}^2 S_{t_j}^2 \right) \left(t_{j+1} - t_j \right) \right\} \end{aligned}$$

donde $\Pi = \{t_0, t_1, \dots, t_n\}$ son los periodos disponibles de rebalanceo. Los valores X_0^1 y X_0^2 son escogidos para que se cumpla el teorema 3.18 (escogiendo que condiciones de dicho teorema se desean cumplir).

Para comparar los modelos disponibles (modelo continuo Black-Scholes-Merton vs. modelo discreto mediante cópulas) se comparan los precios Black-Scholes-Merton

de los derivados propuestos en el teorema 3.23 y las esperanzas neutrales al riesgo descontadas para dichos derivados. Se utilizan muestras móviles de log-rendimientos diarios de tamaño 100 y se valúan derivados a 20 días donde $r = 0.0002$, $r_0 = 0.00025$ son la tasa libre de riesgo y la tasa de rendimiento mínima deseada diaria, respectivamente, y $\alpha = 0.65$ es el peso que se le da al rendimiento del IPC.

En las figuras 3.27-3.35 se comparan los precios para distintos periodos, donde la leyenda BSM representa el precio Black-Scholes-Merton y las leyendas C_1 y C_2 representan las esperanzas neutrales al riesgo descontadas implementando las cópulas 3.7 y 3.8, respectivamente.

Es fácil notar que los precios obtenidos mediante las cópulas 3.7, 3.8 y el modelo de Black-Scholes-Merton siguen la misma tendencia en todos los escenarios, la esperanza neutral al riesgo que más se parece al precio Black-Scholes-Merton es con la cópula 3.7. Recuérdese que dicha cópula es **simétrica**, por lo que su similitud con el precio Black-Scholes-Merton (que implica una cópula normal que también es simétrica) es de esperarse. Para el caso de la cópula 3.8, el precio puede variar por encima (o por debajo) de los otros dos precios, aunque también existen escenarios en los que los tres precios son lo suficiente “similares” para considerarlos iguales en la práctica¹¹.

Otro factor de importancia es el p-value, dichas gráficas se crearon basándose en la prueba de bondad y ajuste para cópulas planteada por Genest *et. al.* (2009). Recuérdese que un p-value chico (normalmente menor o igual a 0.05) representa un rechazo de la hipótesis nula (que en este caso es $\mathcal{H}_0 : C = C_0$ para una cópula C_0 cualquiera), en caso contrario se decide por no poder rechazar la hipótesis nula. De forma práctica, en el caso de que el p-value sea menor o igual a 0.05, ninguna de las cópulas es apta para captar la dependencia entre los subyacentes, por lo que los precios de los derivados hay que considerarlos con cuidado; en caso contrario, si el p-value es mayor a 0.05, se escoje el precio para la cópula con el mayor p-value.

Existen tres escenarios de interés en relación a los precios mostrados en la gráficas 3.27-3.35:

- i) **Derivado subvaluado:** Esto sucede cuando el precio y el p-value para las cópulas C_1 y C_2 es mayor en comparación al modelo Black-Scholes-Merton. Esto se puede ver reflejado en la figuras 3.33 y 3.34, por ejemplo. Para este

¹¹Recuérdese que la esperanza neutral al riesgo no representa el precio al cuál la entidad emisora va a vender el derivado, dicho precio debe obtenerse mediante la creación de un portafolio de superreplicación a tiempo discreto tal y como se desarrolló en la sección 3.2.3. No se incluyen dichas gráficas para evitar redundancia en los precios (dichos precios incrementan un poco en relación a las esperanzas descontadas).

caso, el modelo tradicional está subestimando el riesgo del derivado y existe una oportunidad de arbitraje por parte del **comprador** del derivado.

- ii) **Derivado sobrevaluado:** Esto sucede cuando el precio y el p-value para las cópulas C_1 y C_2 es menor y mayor, respectivamente, en comparación al modelo Black-Scholes-Merton. Esto se puede ver reflejado en la figuras 3.33 y 3.34, por ejemplo. Para este caso, el modelo tradicional está sobrestimando el riesgo del derivado y existe una oportunidad de arbitraje por parte del **emisor** (vendedor) del derivado.
- iii) **Derivado valuado correctamente:** Esto sucede cuando el precio y el p-value para las cópulas C_1 y C_2 son “similares” en comparación al modelo Black-Scholes-Merton. Esto se puede ver reflejado en la figura 3.35, por ejemplo. Para este caso, el modelo tradicional cuantifica de buena forma el riesgo del derivado y no existen oportunidades de arbitraje para ninguna de las partes.

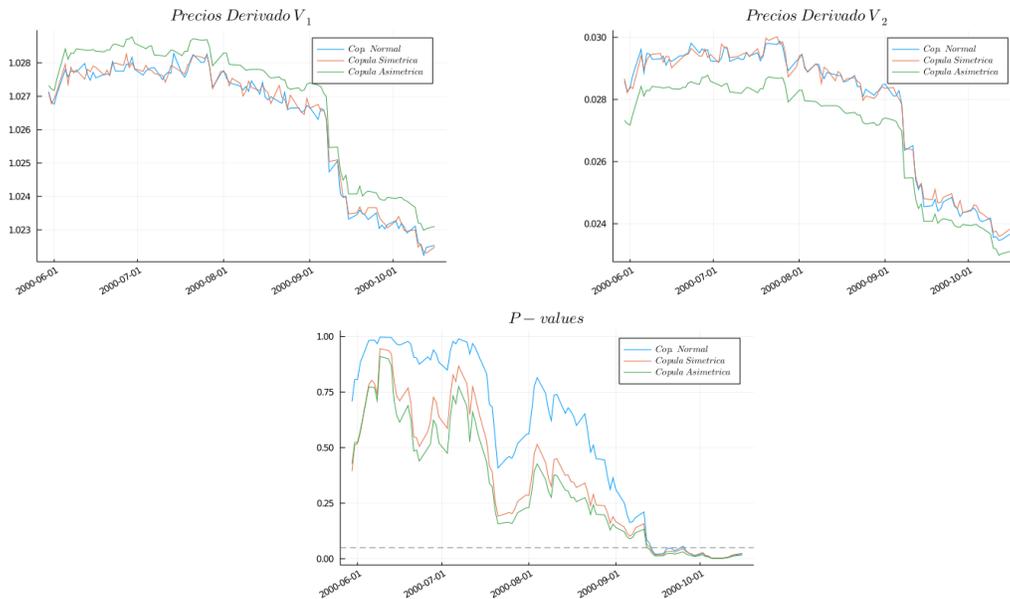


Figura 3.27: Precios y P-values entre 30/05/2000-16/10/2000.

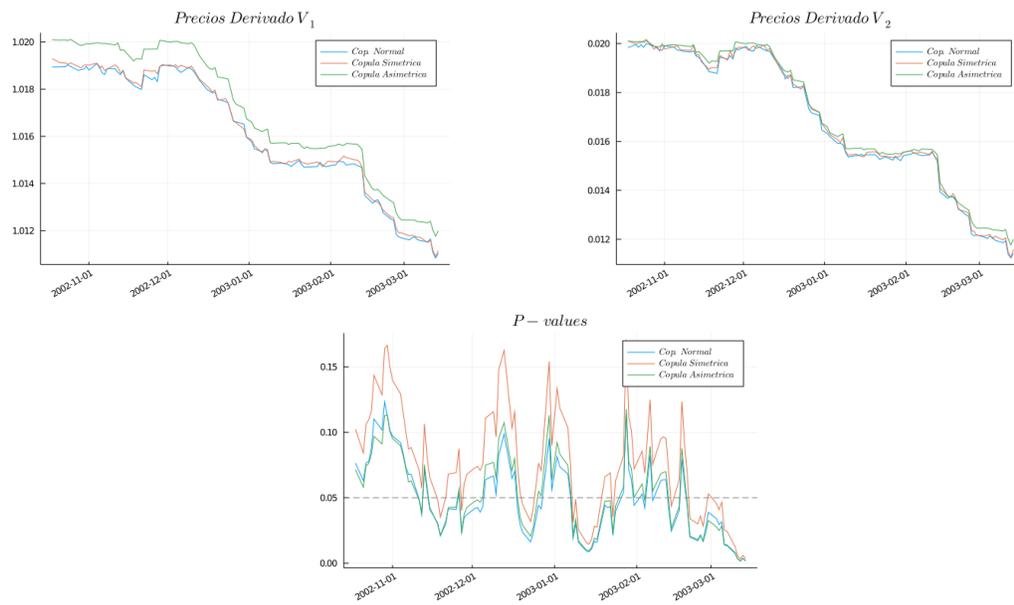


Figura 3.28: Precios y P-values entre 18/10/2002-14/03/2003.

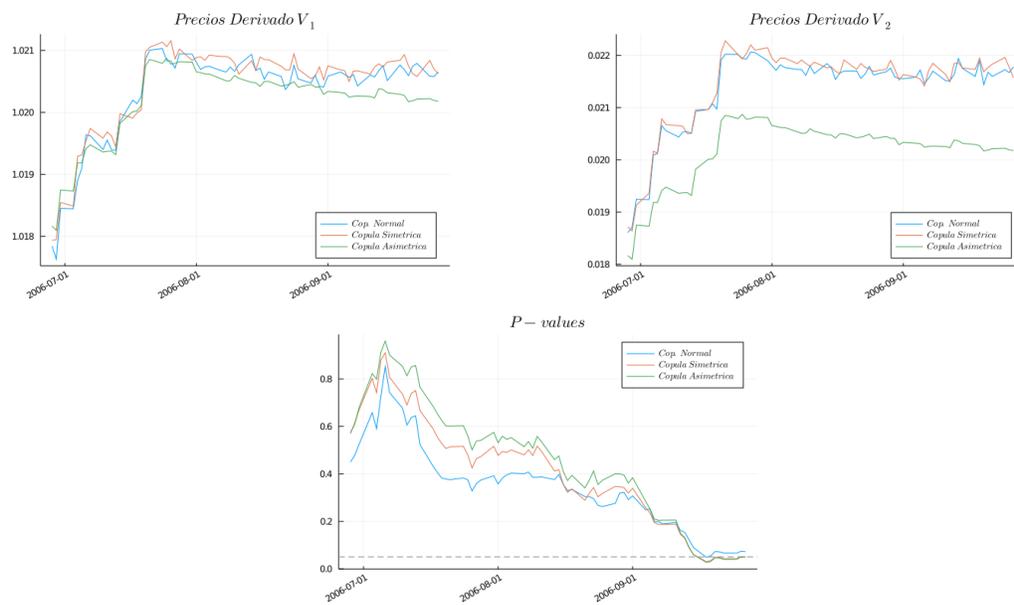


Figura 3.29: Precios y P-values entre 28/06/2006-27/09/2006.

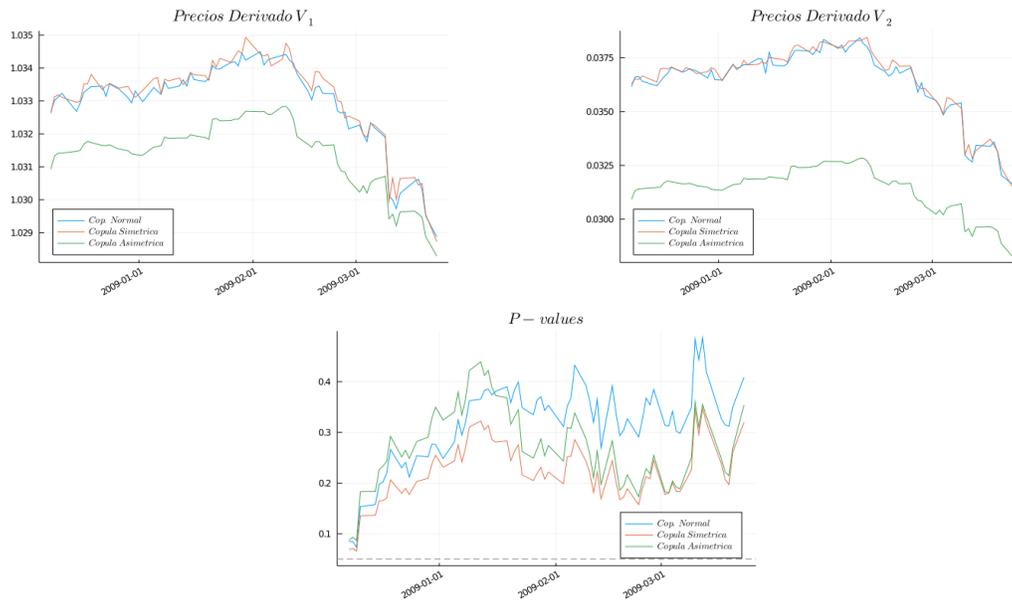


Figura 3.30: Precios y P-values entre 08/12/2008-23/03/2009.

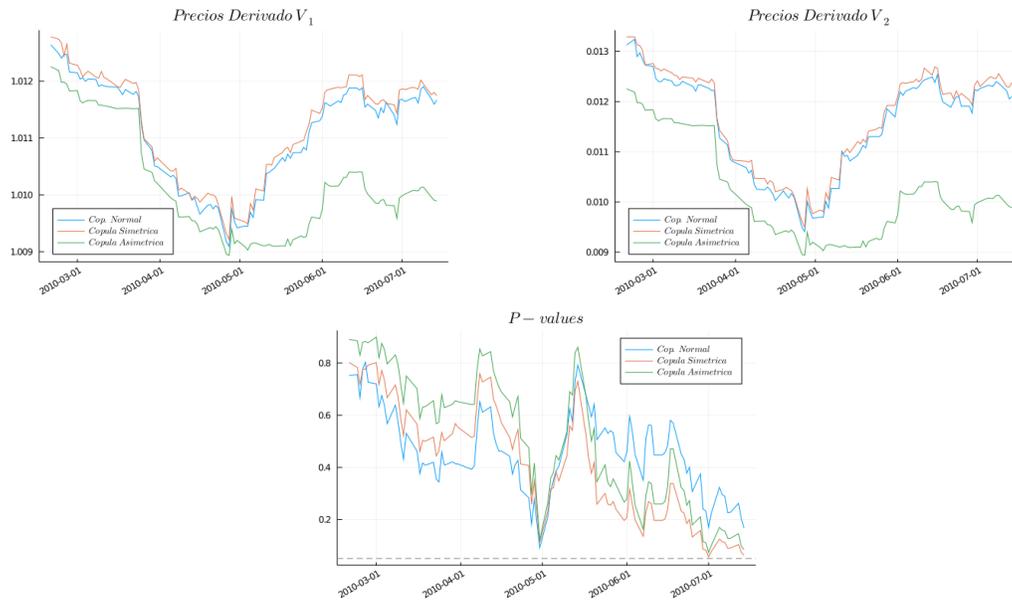


Figura 3.31: Precios y P-values entre 19/02/2010-14/07/2010.

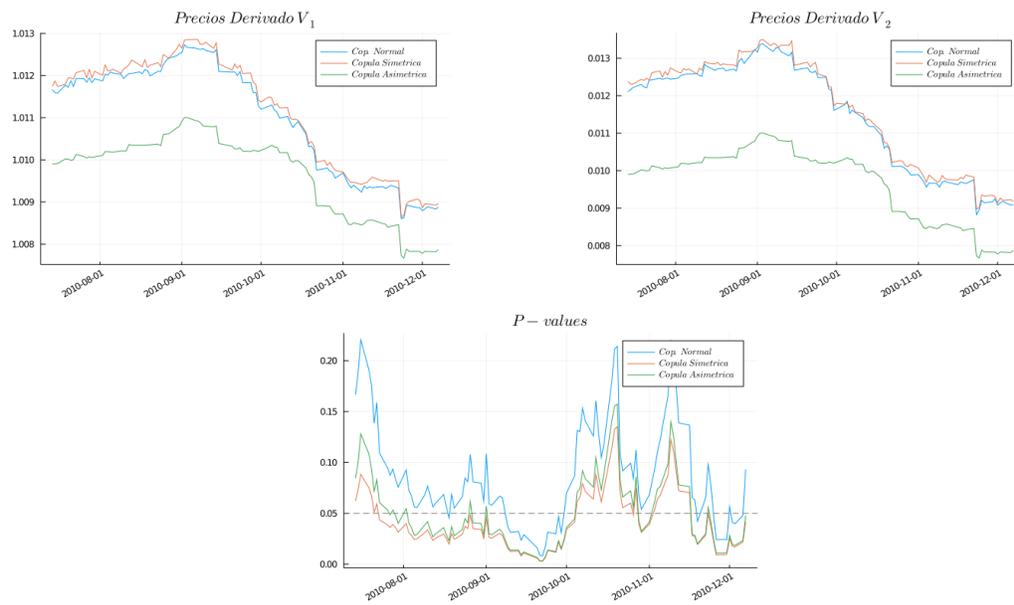


Figura 3.32: Precios y P-values entre 14/07/2010-07/12/2010.

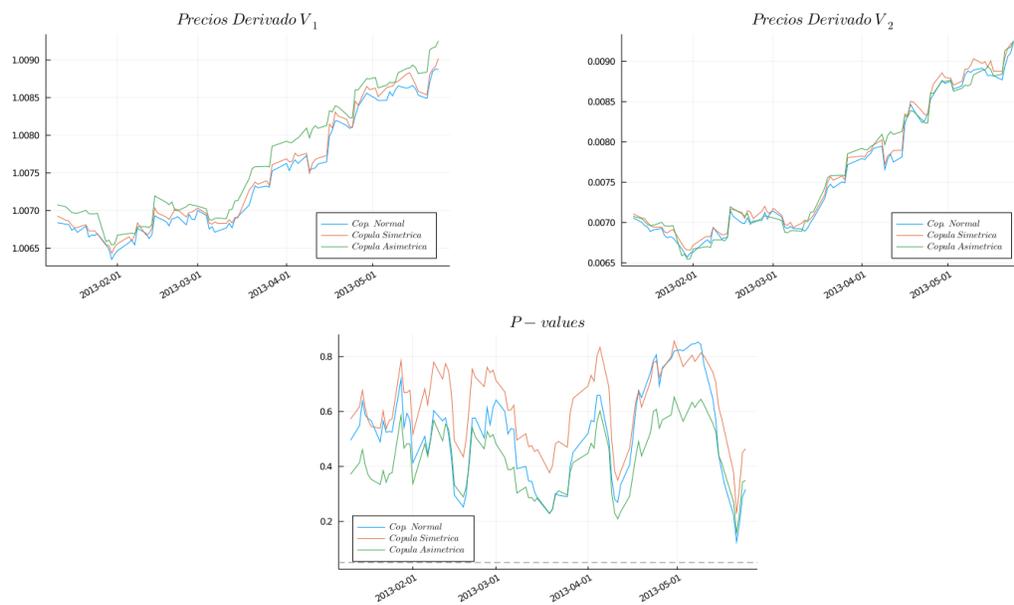


Figura 3.33: Precios y P-values entre 11/01/2013-24/05/2013.

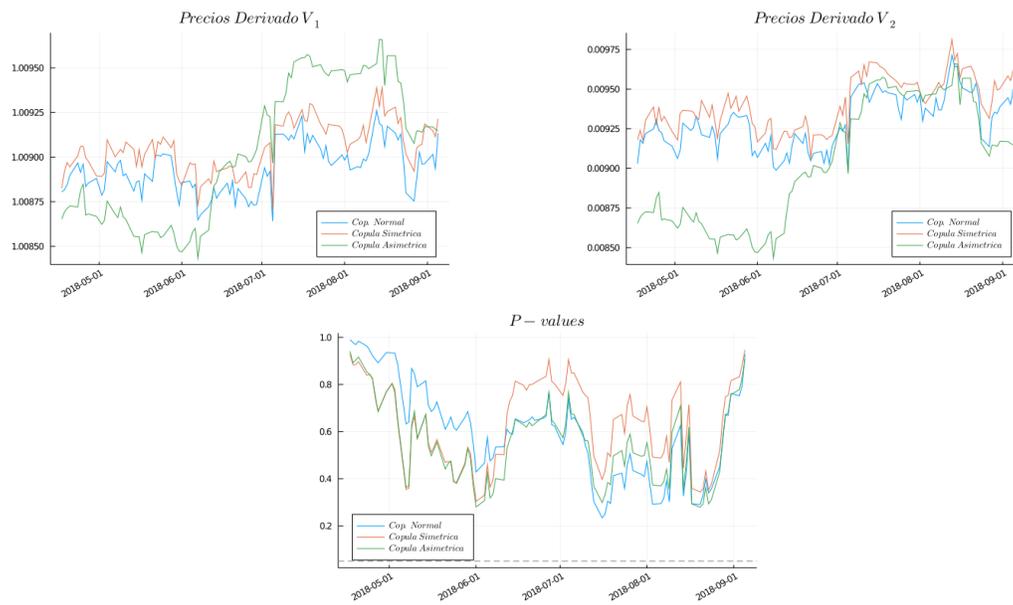


Figura 3.34: Precios y P-values entre 17/04/2018-05/09/2018.

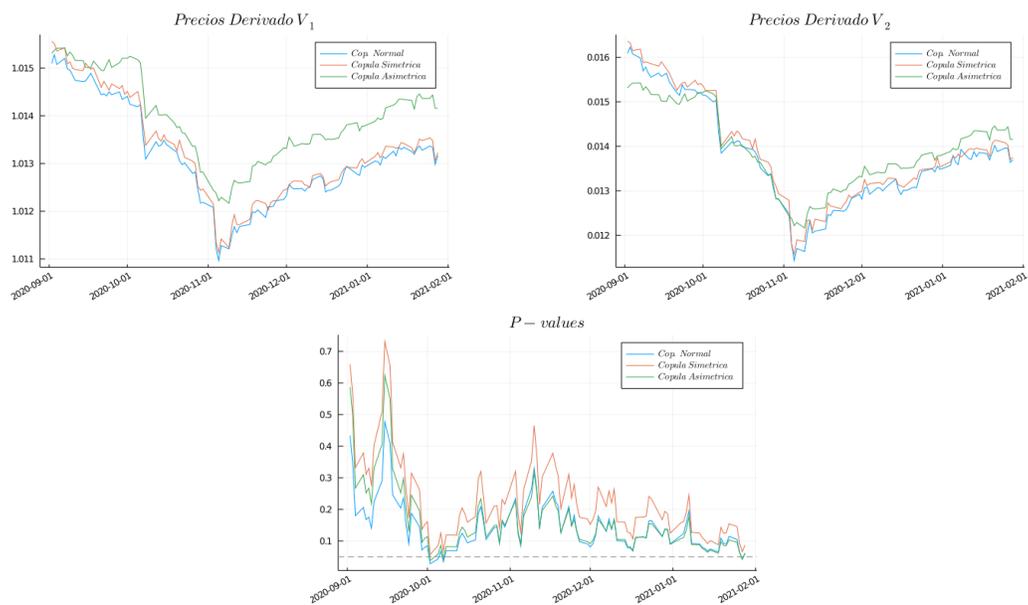


Figura 3.35: Precios y P-values entre 02/09/2020-28/01/2021.

Conclusiones

En el caso de un modelo financiero univariado para acciones, la extensión de un modelo log-normal a un modelo generalizado se puede implementar de forma directa mediante el ajuste paramétrico (o no paramétrico, dependiendo de lo que se desee) de la distribución de los log-rendimientos, y como consecuencia en el subyacente mismo. Esto se describió de forma paramétrica y no paramétrica en las secciones 2.6 y 3.1, respectivamente.

En el caso de un modelo financiero univariado para derivados sobre acciones, la extensión se da como consecuencia natural de lo planteado en el inciso 1), esto se da de forma paramétrica y no paramétrica en las secciones 2.6 y 3.2, respectivamente.

En el caso de un modelo financiero bivariado (para el caso de subyacente o de derivado conjunto), la respuesta no paramétrica se da de igual forma en la sección 3.2, esto implementando un modelo mediante cópulas, lo que permite flexibilizar la modelación de dependencia y ajustarla de acuerdo a la información disponible.

Las pruebas de bondad y ajuste aplicadas de forma univariada sobre los datos (Índice de Precios y Cotizaciones y tipo de cambio peso mexicano-dolar estadounidense) muestran que el ajuste normal para los log-rendimientos **NO** es el modelo adecuado para el caso del IPC, por lo tanto, no se cumple el supuesto univariado del modelo Black-Scholes-Merton.

Las pruebas de bondad y ajuste aplicadas a las tres cópulas seleccionadas (simétrica, no simétrica y normal) muestran que la cópula normal **NO** es siempre el modelo adecuado para cuantificar la dependencia entre los subyacentes. Esto en conjunto al inciso 1), demuestran que el modelo Black-Scholes-Merton no es adecuado para la valuación de derivados con estos subyacentes.

La cópula del subyacente bivariado es cambiante a través del tiempo, eso es claro como consecuencia de los p-values presentes en las figuras 3.27-3.35, dejando en claro que la dependencia es cambiante y se debe analizar continuamente.

El ajuste simétrico y el ajuste normal (valuación Black-Scholes-Merton) conllevarían precios “similares” y difieren de gran forma al precio mediante el ajuste asimétrico, por lo que las zonas de mayor interés son las que muestran dependencia asimétrica y por lo tanto precios distintos.

Por último, el trabajo realizado muestra la eficacia de las cópulas paramétricas y no paramétricas para cuantificar la dependencia entre variables aleatorias, y como éstas se pueden combinar de múltiples formas (combinación lineal convexa, gluing copulas, etc.) para crear modelos de dependencia complejos. El ajuste a datos financieros es eficaz y permite comprender de mejor forma la relación entre indicadores del mercado. Además, la valuación de instrumentos derivados puede extenderse mediante cópulas y mediante modelos univariados paramétricos y no paramétricos que no cumplan las hipótesis del modelo Black-Scholes-Merton tradicional.

Referencias

- Anderson, T., y Darling, D. (1952). Asymptotic theory of certain “Goodness of Fit” criteria based on stochastic processes. *Annals of Mathematical Statistics*, 23(2), 193–212.
- Anderson, T., y Darling, D. (1954). A Test of Goodness of Fit. *Journal of the American Statistical Association*, 49(268), 765–769.
- Bezanson, J., Edelman, A., Karpinski, S., y Shah, V. (2017). Julia: A Fresh Approach to Numerical Computing. *Society of Industrial and Applied Mathematics*, 59(1), 65–98.
- Black, F., y Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3), 637–654.
- Casella, G., y Berger, R. (2002). *Statistical inference*. Duxbury: Thomson Learning.
- Cherubini, U., Gobbi, F., Mulinacci, S., y Romagnoli, S. (2012). *Dynamic Copula Methods in Finance*. John Wiley and Sons, Ltd.
- Cherubini, U., Luciano, E., y Vecchiato, W. (2004). *Copula Methods in Finance*. John Wiley and Sons, Ltd.
- Clayton, D. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*(65).
- Dall’Aglío, G. (1956). Sugli estremi dei momenti delle funzioni di ripartizione doppia. *Annali della Scuola Normale Superiore di Pisa 3rd série*, 10.
- Dall’Aglío, G. (1961). Sulle distribuzioni con margini assegnati soggette a delle limitazioni. *Giorn. Ist. Ital. Attuari*, 29, 94–108.
- Dall’Aglío, G. (1991). Frechet classes: the beginnings. *Advances in Probability Distributions with Given Marginals; Dall’Aglío, Kotz, and Salinetti editors (Kluwer Academic Publishers, Dordrecht)*, 1–12.
- Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés. un test non paramétrique d’indépendance. *Bulletins de l’Académie Royale de Belgique*, 65, 274–292.

- Denuit, M., Dhaene, J., Goovaerts, M., y Kaas, R. (2005). *Actuarial Theory for Dependent Risks Measures, Orders and Models*. John Wiley and Sons, Ltd.
- Domínguez, J. (2001). *Diseño y análisis de modelos de probabilidad*. Grupo Editorial Iberoamérica (México).
- Embrechts, P., Lindskog, F., y McNeil, A. (2001). Modelling Dependence with Copulas and Applications to Risk Management. *Handbook of Heavy Tailed Distributions in Finance*, Rachev, S. (editor) Elsevier, capítulo 8, 329–384.
- Erdelyi, A. (2009). Cópulas y dependencia de variables aleatorias: Una introducción. *Miscelánea Matemática*, 48, 7–28.
- Erdelyi, A. (2017). Vectores aleatorios Notas(incompletas).
- Frank, M. (1979). On the simultaneous associativity of $F(x, y)$ and $x+y-F(x, y)$. *Aequationes Mathematicae*, 197–226.
- Franke, J., Härdle, W., y Hafner, C. (2008). *Statistics of Financial Markets*. Springer.
- Fréchet, M. (1951). Sur les tableaux de corrélation dont les marges sont données. *Annales de l' Université de Lyon*, 4, 53–84.
- Fréchet, M. (1957). Les tableaux de corrélation dont les marges et des bornes sont données. *Ann. Univ. Lyon Sc.*, 20, 13–31.
- Genest, C., y MacKay, J. (1986). The Joy of Copulas: Bivariate Distributions with Uniform Marginals. *The American Statistician*, 40(4), 280–283.
- Genest, C., y Rivest, L. (1989). A Characterization of Gumbel's Family of Extreme Value Distributions. *Statistics and Probability Letters*, 207–211.
- 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, 199–213.
- Gumbel, E. (1960). Distributions des valeurs extremes en plusieurs dimensions. *Publ Inst Statist Univ Paris*, 9, 171–173.
- Hougaard, P. (1986). A Class of Multivariate Failure Time Distributions. *Biometrika*, 73, 671–678.
- Hull, J. (2011). *Options, Futures, and Other Derivatives*. Pearson.
- Kellerer, H. G. (1961). Funktionen auf Producträumen mit vorgegeben Marginal Funktionen. *Math. Annalen*, 153, 323–344.
- Klenke, A. (2008). *Probability Theory A Comprehensive Course*. Springer.
- Kruskal, W. (1958). Ordinal Measures of Association. *Journal of the American Statistical Association*, 53(284), 814–861.
- Kummer, S., y Pauletto, C. (2012). The History of Derivatives: A Few Milestones.
- Mansuy, R. (2005). Histoire de martingales. *Mathématiques et Sciences Humaines/Mathematical Social Sciences*, 43(169), 105–113.
- Merton, R. (1973). Theory of Rational Option Pricing. *The Bell Journal of Economics and Management Science*, 4(1), 141–183.

- Moynihan, R., y Schweizer, B. (1979). Betweenness relations in Probabilistic Metric Spaces. *Pacific Journal of Mathematics*, 81(1), 175–196.
- Nelsen, R. B. (2006). *An introduction to copulas*. Ed. Springer-Verlag (New York) 2nd edition (first edition dates from 1999).
- Oakes, D. (1982). A model for Association in Bivariate Survival Data. *Journal of the Royal Statistical Society*, 44(3), 414–422.
- Poitras, G. (2009). The Early History of Option Contracts. En W. Hafner y H. Zimmermann (Eds.), (pp. 487–518). Springer.
- Póczos, B., Krishner, S., Pal, D., Szepervari, C., y Schneider, J. (2018). Robust Non-parametric Copula Based Dependence Estimators. *Carnegie Mellon University. Journal contribution*.
- Rincón, L. (2007). *Curso intermedio de probabilidad*. Facultad de Ciencias de la UNAM (México).
- Salvadori, G., De Michele, C., Kottegoda, N., y Rosso, R. (2007). *Water Science and Technology Library: Extremes in Nature An Approach Using Copulas* (Vol. 56). Springer Complexity.
- Schweizer, B. (1991). Thirty years of Copulas. *Advances in Probability Distributions with Given Marginals; Dall’Aglia, Kotz, and Salinetti editors (Kluwer Academic Publishers, Dordrecht)*, 13–50.
- Schweizer, B., y Sklar, A. (1960). Statistical Metric Spaces. *Pacific Journal of Mathematics*, 10(1), 313–334.
- Schweizer, B., y Sklar, A. (1974). Operations on distribution functions not derivable form operations on random variables. *Studia Mathematica*, 52, 43–52.
- Schweizer, B., y Sklar, A. (1983). *Probabilistic Metric spaces*. Elsevier North-Holland, New York.
- Schweizer, B., y Wolff, E. F. (1981). On Nonparametric Measures of dependence for Random Variables. *The Annals of Statistics*, 9, 879–885.
- Seeley, R. (1961). Fubini Implies Leibniz Implies $F_{xy} = F_{yx}$. *The American Mathematical Monthly*, 68, 56–57.
- Shao, J. (2003). *Mathematical statistics*. Springer.
- Shreve, S. (2004a). *Stochastic Calculus for Finance II Continuous-Time Models*. Springer.
- Shreve, S. (2004b). *Stochastic Calculus for Finance I The Binomial Asset Pricing Model*. Springer.
- Siburg, K., y Stoimenov, P. (2008). Gluing Copulas. *Communications in Statistics-Theory and Methods*, 37(19), 3124–3134.
- Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Inst. Statist. Univ. Paris Publ.*, 8, 229-231.

- Sklar, A. (1996). Random variables, distribution functions, and copulas - A personal look backward and Forward. *Distributions with Fixed Marginals and Related Topics IMS Lecture Notes - Monograph Series*, 28.
- Spivak, M. (1988). *Cálculo en variedades*. Reverté Ediciones, S.A. de C.V.
- Spivak, M. (1996). *Cálculo Infinitesimal*. Reverté Ediciones, S.A. de C.V.
- Trivedi, P., y Zimmer, D. (2005). Copula Modeling: An Introduction for Practitioners. *Foundations and Trends in Econometrics*, 1, 1–111.
- Vitale, R. (1990).
Henry W. Block, Allan R. Sampson, and Thomas H. Savits, eds. Topics in statistical dependence (Hayward, CA: Institute of Mathematical Statistics, 1990), 16, 459–469.