

162-21

INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE MONTERREY

CAMPUS CIUDAD DE MÉXICO



**Valuación de un bono hipotecario mediante cópulas
arquimedianas con densidades marginales de valores
extremos**

DOCTORADO EN CIENCIAS FINANCIERAS

TESIS PRESENTADA POR

Salvador Cruz Aké

Marzo de 2009

TESIS

HG4655

C78

2009

CVV

513005492





**TECNOLÓGICO
DE MONTERREY**

Hacemos constar que en la Ciudad de México, el día 25 de marzo de 2009, el alumno:

SALVADOR CRUZ AKÉ

sustentó el Examen de Grado en defensa de la Tesis titulada:

Valuación de un bono hipotecario mediante cópulas arquimedianas con densidades marginales de valores extremos.

Presentada como requisito final para la obtención del Grado de:

DOCTOR EN CIENCIAS FINANCIERAS

Ante la evidencia presentada en el trabajo de tesis y en este examen, el *Comité Examinador*, presidido por el **DR. FRANCISCO ORTIZ ARANGO**, ha tomado la siguiente resolución:

APROBADO

Dr. Francisco Venegas Martínez
Director de Tesis

Dr. José Antonio Núñez Mora
Codirector

Dr. Igor Patricio Rivera González
Lector

Dr. Francisco Ortiz Arango
Lector

Dr. José Antonio Núñez Mora
Director del Programa Doctoral

Índice

Parte I. Marco Teórico

Introducción	1
Resumen	1
1.1 El valor del dinero en el tiempo	1
1.2 Concepto de <i>Mortgage Backed Securities</i>	9
1.3 Conceptos básicos de cópulas	25
1.4 Medidas de dependencia no paramétricas	34
1.5 Distribuciones marginales y formación de cópulas	41
1.6 Máxima verosimilitud en dos etapas	43

Parte II Los datos y las estimaciones

2.1 Fuente de los datos	51
2.2 Relaciones empíricas, la primera aproximación	58
2.3 Anlisis de las distribuciones marginales con funciones de valores extremos, ajustes preeliminares	66
2.4 Estimaciones de las funciones marginales de densidad por mes.	74
2.5 Breve interpretación de las marginales	81
2.6 Estimación preeliminar de los parámetros de dependencia.	85
2.7 Estimación formal de los parámetros de dependencia.	88

Parte III Aproximación del valor del bono y conclusiones.

3.1 Aproximación final al valor de los BRH.	94
--	----

Parte IV Apéndices.

A.1 Las cópulas y el teorema de Sklar.	102
A.2 Métodos para la generación de cópulas.	119
A.3 Teoría de valores extremos.	125
A.4 Máxima Verosimilitud en dos Etapas (2SML).	131
A.5 Algoritmos usados para las estimaciones.	146

Parte IV Bibliografía

Bibliografía 163

AGRADECIMIENTOS

Este trabajo se lo quiero dedicar a:

Mi madre, María del Carmen Aké, debió de haber sido difícil criar a dos varones como Enrique y yo. (y yo me quejo de la tesis)

Enrique, Brenda y Victor, son mis contactos con el mundo real, siempre el motivo para seguir adelante.

El “Doc” Francisco Venegas, por ser mi amigo, mentor, director de tesis y profesor. No imagina como cambió mi vida y mi concepto del mundo.

Pepina, Sus, Cuca, Lulú y Salvador Cruz (Ya se que no les gustan esos diminutivos, pero ahora es cuando) por confiar en mi y aceptarme como parte de su familia.

César Contreras Piedragil, gracias por todo lo que te he aprendido, las horas de consejo, clase y en ocasiones consuelo.

Araceli Ortega Díaz, fuiste un parteaguas en mi vida adulta, eres una mujer excepcional y lo mejor de todo, mi amiga.

Araceli Elguea, eres una estupenda amiga y un ejemplo de constancia y determinación.

Francisco Aguirre y Lourdes Cortés; sin su ayuda no hubiera acabado la licenciatura, me devolvieron mucha de la confianza perdida.

Juan Jesús Vargas, Concepción Covián y personal de caja de la UP; por los años de amistad, convivencia y trabajo.

Edwin Rivera, gracias por ayudarme a controlar mi carácter, los consejos y ayuda en la UP, incluso por correrme de tu clase.

Mis amigos de la UP, del Tec y del Poli.

Resumen

El objetivo principal de este trabajo consiste en proponer un método que permita valorar bonos hipotecarios tomando en cuenta la dependencia no lineal que presentan los parámetros de prepago e incumplimiento con la sobretasa de refinanciamiento. Es común encontrar en la literatura de riesgos (*v.gr.* Fabozzi) dos métodos sencillos aunque inexactos para valorar esta clase de instrumentos. El primero de ellos considera al bono colateralizado (*Mortgage Backed*) como un bono con riesgo crédito que tiene inmersa una posición corta en una opción de compra americana donde la parte sin amortizar del bono (*outstanding*) es el precio de ejercicio en cada momento, mientras que el subyacente es el bono mismo. Asimismo, es común encontrar una simplificación de este modelo cuando se supone que la opción a valorar no es americana sino europea, con lo que se abandona la naturaleza misma del instrumento.

Dentro de la literatura especializada también se encuentran métodos para valorar el bono en cuestión tomando en cuenta su prepago y siguiendo las tasas PSA (*Post Specific Allowance Rates*), las cuales son una función lineal escalonada sobre el prepago que experimentan estos bonos según un promedio histórico dado por la “Bond Dealers Association”. Sobre esta misma línea se encuentra el método CPR (*Conditional Prepayment Rate*), el cual se basa en un parámetro de supervivencia que permanece constante a lo largo de todos los periodos y que modela el prepago del instrumento.

Dentro de los avances más recientes por valorar estos instrumentos, se encuentra el uso de funciones de distribución multivariadas para modelar la dependencia de los parámetros, destacando el uso de las distribuciones normales multivariadas, aunque por su naturaleza, no reflejan las realizaciones extremas que pueden apreciarse en estos instrumentos. A lo largo de esta tesis, se estudiará el mecanismo a través del cual los distintos parámetros

que afectan al precio del bono interactúan entre sí, a saber: la tasa de refinanciamiento, el parámetro de incumplimiento y el de prepago. Se inicia el trabajo con una explicación de lo que es un bono colateralizado o bono respaldado por hipotecas (*Mortgage backed securities*) y se delimita el trabajo al campo de los bonos (*pass through*), los cuales son vendidos a los inversionistas como un instrumento que paga conjuntamente los intereses y la amortización del capital de los acreditados.

Una vez establecido el problema en que se concentrará la presente investigación, se procederá a explicar las ventajas que tienen estos instrumentos para sus emisores, así como la forma en que usualmente son constituidos y valuados, para después revisar el concepto de cópula y la forma en que se construyen las familias más comunes de cópulas, con especial énfasis en la familia arquimediana, así como mostrar la conveniencia del uso de éstas para la valuación de dichos bonos.

Para la valuación de estos bonos “de traspaso total” (*pass trough*), se describirán las cópulas que serán usadas en esta investigación para modelar las relaciones de dependencia; haciendo referencia a varios trabajos recientes dentro de la literatura de cópulas y funciones de probabilidad multivariadas. La forma funcional explícita que seguirán las cópulas son: una cópula Weibull para la relación entre la tasa de refinanciamiento y el parámetro de prepago, una cópula Ali-Mikhail-Haq para la relación entre la tasa de refinanciamiento y el parámetro de incumplimiento y, por último, una cópula FGM para la relación entre los parámetros de incumplimiento y prepago.

Posteriormente, se discute sobre la medida de dependencia que se empleará en el trabajo empírico. Para ello y en vista de la limitación que una medida paramétrica como la desviación estándar representa, se hace una breve exploración de las medidas de dependencia no paramétrica, pasando por la “rho” de Spearman, la “tau” de Kendall y la “beta” de Blomqvist. Asimismo, se reviza la forma en que son calculados cada uno de estos parámetros y se desechan los dos primeros por su complejidad computacional. Después se lleva a cabo un análisis más profundo sobre la “beta” de Blomqvist y de su relación con el concepto de cópula con base en una investigación reciente que demuestra la convergencia

asintótica de este parámetro a la distribución normal, lo que resulta de vital importancia para su instrumentación dentro del proceso de máxima verosimilitud en dos etapas. el cual es necesario para hacer un ajuste empírico del modelo.

Después de la revisión los fundamentos de la estadística no paramétrica, se hace un breve resumen de la teoría de valores extremos, específicamente de las distribuciones acumuladas del máximo. Se parte de la parametrización “alfa” de las distribuciones Weibull y Fréchet, para después hacer uso de la transformación “gamma” resaltando que ambas distribuciones presentan la misma forma funcional, las cuales difieren únicamente en el dominio del parámetro “gamma”. Una vez hecha esta transformación, se obtienen las funciones de cuantiles respectivas y se incluyen dentro de las cópulas previamente escogidas, dando como resultado las formas funcionales explícitas de éstas.

Una vez finalizado el aspecto teórico, se hace un breve repaso del método de Máxima Verosimilitud en dos etapas referido previamente, en cuyo caso se muestra la convergencia de los parámetros a la distribución normal y, posteriormente, se procederá a la estimación del modelo para datos de “Fannie Mae” (Agencia del gobierno norteamericano dedicada a facilitar créditos para la vivienda). Para ello se ajustan funciones de valores extremos para las funciones marginales y después para la cópula, tomando como parámetros iniciales los valores obtenidos en la primer etapa y el valor muestral de la “Beta” de Blomqvist para los factores de riesgo.

Introducción

1.1 El valor del dinero en el tiempo

Uno de los conceptos fundamentales de las finanzas, tal vez su *leitmotiv*, es el concepto del valor del dinero a través del tiempo. Este surge de la existencia de las tasas de interés y del costo de oportunidad que ellas reflejan, pues el no usar hoy los recursos que se poseen, significa la pérdida potencial de los mismos (si no se invierten) o de la “utilidad” obtenida de ellos (si no se consumen). La tasa de interés es definida¹ como el “precio” del dinero. Lo anterior implica la existencia de un mercado del mismo, lo que a su vez supone la existencia de una oferta y una demanda de este “activo”.

Más allá de la discusión microeconómica que esta afirmación podría suscitar, pues algunos economistas dirán que es el Banco Central (la autoridad monetaria) la única capaz de ofrecer dinero, el trabajo se enfocará al mercado de los fondos prestables. Este, puede ser tratado como un mercado, con algunas intervenciones de participantes importantes (como el gobierno).

Si se intentan calcular las funciones de oferta y demanda por estos fondos prestables haciendo uso de herramientas econométricas, el lector se daría cuenta que sólo conoce los puntos de equilibrio del mercado, *i.e.*, las realizaciones de estos equilibrios. Y aunque econométricamente el problema es abordado “identificando” cada una de las ecuaciones con una única variable independiente, aun cuando sólo se observan equilibrios de mercado. Afortunadamente éstos son los que importan para este trabajo.

¹ Véase Nicholson W, *Microeconomic Theory, basic principles and extentions*, South-Western College Publications, 9 Ed, 2004.

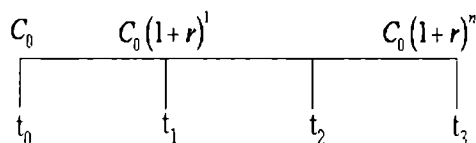
Estos equilibrios observados en el mercado de fondos prestables (nodos), que aún no incluyen otras fuentes de incertidumbre que afectan al bono, son realizaciones de una variable aleatoria conducida por un proceso estocástico cuya fuente de incertidumbre se supondrá como un movimiento browniano. Este proceso estocástico se puede utilizar para modelar la tasa de interés de distintas maneras, teniendo siempre en mente que el modelo será usado, entre otras cosas, para generar una estructura de plazos, la cual proporciona las tasas de interés $r(t, T)$, aplicable en $[t, T]$, mismas que se utilizan para los fondos que permanecerán comprometidos a distintos periodos de tiempo y que reflejan los costos de oportunidad que en conjunto perciben los integrantes del mercado de fondos prestables². Antes de seguir adelante con el planteamiento del modelo, es necesario revisar la noción de valor del dinero en el tiempo con base en un préstamo con interés simple.

Suponga que se parte de una estructura de plazos plana y se establece la condición de no arbitraje en el mercado de fondos prestables. Asimismo, suponga que se poseen N unidades monetarias y que éstas se depositan durante un año en un sistema bancario en el que los agentes pueden prestar y pedir prestado a una misma tasa libre de riesgo (de incumplimiento), r . En este caso, después de un año, se tiene que el valor futuro del depósito inicial N es de

$$V_f = N(1 + r), \quad (1.1.1)$$

lo cual define el valor futuro de N unidades monetarias. Este sencillo ejemplo donde la tasa de interés es constante y determinista y el valor futuro tiene como referencia exactamente un año puede generalizarse suponiendo que el valor futuro buscado se refiere a un tiempo arbitrario de $T - t$ años.

² Aún no se hace explícita la segregación por riesgo crédito de este mercado de fondos prestables. Se parte del supuesto de que se modela la tasa libre de riesgo crédito y después se añade un diferencial para reflejar la probabilidad de no pago.



Gráfica 1.1 El valor del dinero en el tiempo, la capitalización

Esta consideración lleva a un valor futuro $V_f = N(1+r)^n$, donde $n \in \mathbb{N}$ es medido en años. Este ejemplo sencillo introduce un concepto importante, a saber, la capitalización, es decir, el capital inicial recibe un monto de intereses que después de un año son agregados a éste para conformar un nuevo capital, el cual a su vez es reinvertido a la misma tasa.

Obviamente, el ejemplo anterior se puede complicar un poco si se supone que el periodo donde se aplica la tasa es $T - t \in \mathbb{R} \cup \{0\}$, es decir, que aunque la tasa de interés es expresada en términos anuales, el periodo de capitalización buscado no es un múltiplo exacto de un año, *v.gr.* un mes o un semestre. Esto conduce a

$$V_f = N \left(1 + \frac{r(T-t)}{n} \right)^n, \quad (1.1.2)$$

lo cual sigue siendo un valor futuro.

El siguiente paso en el proceso de generalización es permitir que el nominal, N , capitalice intereses de forma continua, lo que implica la búsqueda del siguiente límite

$$\lim_{n \rightarrow \infty} \left(1 + \frac{r(T-t)}{n} \right)^n = e^{r(T-t)}. \quad (1.1.3)$$

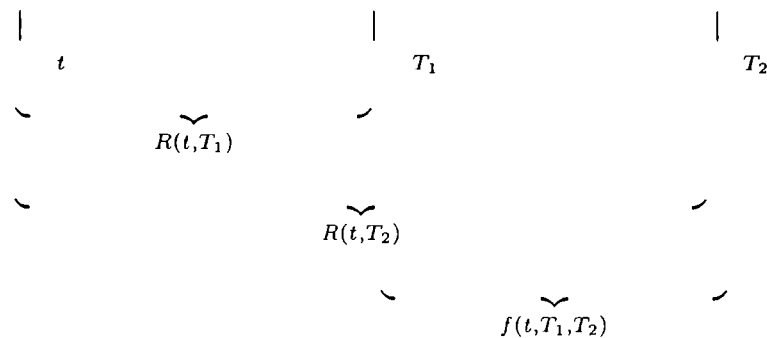
Este último resultado indica que si la capitalización de los intereses es continua estos seguirán un crecimiento exponencial con la tasa de interés como tasa de crecimiento. Con la expresión anterior un nominal N se puede llevar a valor futuro o traerlo a valor presente. Lo anterior permite valuar bonos cupón cero cuando la capitalización es continua, la tasa es constante y no hay riesgo crédito.

Una vez que es conocida la forma de valuar un bono cupón cero, se puede buscar la estructura de plazos de la tasa de interés. Para hacerlo se parte del hecho de que

$B = e^{-R(t,T)(T-t)}$, lo cual implica que se puede obtener la estructura de plazos al despejarla del precio del bono, esto es: $R(t,T) = -\frac{1}{T-t} [\ln(B(t,T))]$. Una vez que se tiene esta estructura de plazos, es posible estimar las tasas forward³. Su construcción toma como premisa la inexistencia de oportunidades de arbitraje, esto es:

$$1 + f(t, T, T + \Delta) = \frac{1 + r(t, T + \Delta)}{1 + r(t, T)} \quad (1.1.4)$$

lo que puede verse gráficamente como:



Gráfica 1.1.1 Estructura de plazos, la tasa forward y el concepto de no arbitraje.

Uno de los puntos relevantes del supuesto de no arbitraje es que éste (el arbitraje) sólo dura por espacios muy breves de tiempo, pues en cuanto una de estas oportunidades es detectada, es aprovechada por algún participante del mercado; no siendo siempre el mismo participante el que aprovecha esta oportunidad.

La sutileza anterior puede llevar a la vieja discusión sobre la eficiencia de mercado, no siendo ésta el motivo de la tesis. Para evitar este escollo, se supondrá como válida la eficiencia débil del mercado, lo que significa que los participantes no pueden realizar ganancias extraordinarias usando información pública (estados financieros, precios de activos, datos macroeconómicos u otro tipo de información) o alguna clase de filtro (lineal o no lineal) sobre la información pública disponible. Por lo que la única fuente de ganancias extraordinarias es la información privada de la firma u organización, cuyo uso está prohibido por la mayor parte de los mercados organizados.

³ Una tasa forward, $f(t, T, T + \Delta)$, es aquella tasa que en el tiempo t se compromete a pagar la tasa $f(t, T, T + \Delta)$ a partir del tiempo T y concluyendo en $T + \Delta$; siendo t el tiempo actual.

Si se regresa al concepto y método para obtener las tasas “forward”, es importante mencionar que también es posible obtenerlas usando tasas que capitalizan de forma continua, lo que significa que $e^{f(t,T,T+\Delta)} = e^{r(t,T+\Delta)} - e^{r(t,T)}$, por lo que es posible inferir que

$$f(t, T, T + \Delta) = r(t, T + \Delta) - r(t, T). \quad (1.1.5)$$

Por el momento, se dejará de lado la estructura de plazos recién obtenida para focalizar la atención en uno de los conceptos fundamentales en el desarrollo de esta tesis. Este concepto es el de “anualidad vencida”.

Los libros básicos de matemáticas financieras⁴ definen las anualidades vencidas como “una sucesión de pagos iguales a intervalos regulares sobre un periodo de tiempo específico”. La ecuación que usualmente se emplea para valorar una anualidad supone que la estructura de plazos de la tasa de interés libre de riesgo y anualizada es plana.

En la valuación de bonos se plantea el problema de calcular una serie geométrica donde cada término es el valor futuro del flujo común A (iniciando los flujos en el periodo $t + 1$), donde la tasa libre de riesgo se utiliza para obtener el valor futuro de cada pago, A , en el tiempo $t + i, i = 1, \dots, n$. Lo anterior lleva a:

$$F = A \left[(1 + r)^{n-1} + (1 + r)^{n-2} + \dots + (1 + r)^1 + 1 \right].$$

En este caso, se sabe que es posible obtener el valor de una serie geométrica usando

$$S_n = \frac{a_1 - ba_n}{1 - b}, \quad (1.1.6)$$

donde $a_n = (1 + r)^{n-1}$ es el último término, $b = 1/(1 + r)$ es la razón común y $a_1 = 1$ es el primer término. Después de realizar las sustituciones correspondientes se encuentra que el valor futuro de una serie de n pagos iguales por un monto de A , suponiendo una tasa

⁴ Simon, C & Blume, L, *Mathematics for Economists*, Norton & Company, New York, 1994, pp 98.

de interés constante de r , es

$$\begin{aligned}
 F &= A \begin{bmatrix} (1+r)^{n-1} - \frac{1}{(1+r)} \\ 1 - \frac{1}{(1+r)} \end{bmatrix} \\
 &= A \begin{bmatrix} (1+r)^{n-1} - (1+r)^{-1} \\ \frac{(1+r)^{-1}}{(1+r)} \end{bmatrix} \\
 &= A \begin{bmatrix} (1+r)^{n-1} - (1+r)^{-1} \\ r(1+r)^{-1} \end{bmatrix} \\
 &= A \begin{bmatrix} (1+r)^n - 1 \\ r \end{bmatrix}.
 \end{aligned} \tag{1.1.7}$$

Para propósitos de esta tesis se necesita que el valor presente de los flujos sea el del préstamo recibido, por lo que se debe adecuar la ecuación anterior para llegar a un valor presente. Esto se logra descontando el valor futuro y trayéndolo al presente, t , esto es $F = P(1+i)^n$. Si se sustituye la F de la ecuación anterior en el valor presente y se despeja P , se obtiene

$$\begin{aligned}
 P &= A \begin{bmatrix} (1+r)^n - 1 \\ r \end{bmatrix} (1+r)^{-n} \\
 &= A \begin{bmatrix} (1+r)^n - 1 \\ r(1+r)^n \end{bmatrix}.
 \end{aligned} \tag{1.1.8}$$

Aunque a lo largo de la tesis se trabajará con estas anualidades compuestas de forma discreta, es posible encontrar una expresión para cuando la composición es continua. Para ello también se parte de una serie de flujos. Por simplicidad, dichos flujos se traerán a valor presente, $P = Ae^{-r} + Ae^{-2r} + \dots + Ae^{-nr} = A(e^{-r} + e^{-2r} + \dots + e^{-nr})$.

De nuevo, se usa un resultado conocido para valuar una serie que muestra un término común, se inicia buscando el factor común de la serie, éste es e^{-r} . Una vez identificado, se multiplica toda la serie por ese factor común, para después restar ambas series (la original y la multiplicada por el factor constante). Este procedimiento conduce a $(a + \dots + a^n)(1-a) = a - a^{n+1}$, reagrupando se llega a

$$a + \dots + a^n = \frac{a(1-a^n)}{1-a}. \tag{1.1.9}$$

Después de hacer las sustituciones necesarias, es posible concluir que

$$\begin{aligned} P &= A \left[\frac{e^{-r} (1 - e^{-nr})}{1 - e^{-r}} \right] \\ &= \frac{A(1 - e^{-nr})}{e^r - 1} \end{aligned} \quad (1.1.10)$$

Dado que en esta tesis se busca encontrar un método más eficiente para valorar MBS (*Mortgage Backed Securities*) que se ven como un tipo de notas estructuradas, es necesario partir de la definición de anualidad proporcionada por la ecuación 1.1.8 para de allí apegar el modelo un poco más a la realidad de los bonos respaldados por hipotecas.

En general, una hipoteca es cubierta haciendo pagos de manera periódica, los cuales pueden ser fijos o variables, a lo largo de un periodo de tiempo durante el cual se pagan intereses sobre el monto de capital (el valor del inmueble) no amortizado. Esta situación recuerda, haciendo algunas simplificaciones, a la ecuación del valor presente de una anualidad.

Como se mencionó antes, esta forma de calcular el valor presente de una anualidad supone que la estructura de plazos de la tasa de interés es plana, *i.e.*, que la tasa de interés es la misma a lo largo de toda la vida del instrumento. El siguiente paso dentro del desarrollo de ésta tesis será relajar este supuesto, por lo que la valuación de la anualidad se convertirá en

$$\begin{aligned} B &= A(1 + r(t, t + 1))^{-1} + \dots + A(1 + r(t, t + n))^{-n} \\ &= A \left[(1 + r(t, t + 1))^{-1} + \dots + (1 + r(t, t + n))^{-n} \right] \\ &= A \sum_{i=1}^n d_i, \end{aligned} \quad (1.1.11)$$

donde d_i es el factor de descuento obtenido de aplicar la tasa de interés correspondiente al plazo que comienza en el periodo de valuación t y termina en el tiempo $t + i$, tiempo en el cual, el flujo de efectivo es recibido. Por lo que es el factor de descuento apropiado para cada periodo de pago de la anualidad.

La forma en que se obtendrá esa estructura de plazos, y la explicación detallada de la valuación del instrumento, son temas que serán discutidos posteriormente a lo largo del

trabajo. Por lo pronto, en la siguiente sección, se tratará el concepto de *Mortgaged Backed Securities*, así como la forma en que actualmente la literatura sugiere su valuación.

1.2 Concepto de Mortgage Backed Securities

Los Bonos Respaldados por Hipotecas (BRH) son una subclase de los bonos respaldados por activos, cuyos flujos de efectivo son garantizados por los pagos de capital e intereses de préstamos hipotecarios. Generalmente estos pagos son mensuales y llegan a presentar incumplimientos que tradicionalmente son absorbidos por el agente emisor, quien conserva un segmento de las hipotecas tal que, a su juicio, sea suficiente para garantizar el pago puntual del bono y sus intereses. Estos instrumentos también suelen presentar prepagos que generalmente son absorbidos por el tenedor del mismo.

Aunque su reciente popularidad en los mercados mundiales fue adquirida a raíz de la reciente crisis financiera que afectó el capital contable de bancos del tamaño de Citigroup, Merrill Lynch, Barclay's Capital, *et. al.*, su existencia en el mercado no es nueva. Aun cuando su valuación puede resultar sumamente complicada, el mercado ha aceptado el método de valuación que los considera como un bono que enfrenta riesgo crédito con una opción de prepago. En otras palabras, un caso especial de nota estructurada.

Como parte de la reciente expansión del mercado inmobiliario en México y de su inclusión en los mercados financieros globales, el país ha tropicalizado estos instrumentos llamándolos BORHIS (Bonos Respaldados por Hipotecas), los cuales son similares a los *MBS* del mercado norteamericano. En todos los casos, los BRH responden a la necesidad de liquidez por parte de los agentes que otorgaron financiamientos hipotecarios y recibieron como colateral un bien inmueble. Este instrumento otorga la liquidez de los mercados secundarios de deuda a los emisores que originalmente mantenían sus recursos en el relativamente ilíquido mercado de bienes raíces, lo que les permite acelerar el ciclo de recuperación del efectivo y continuar con su negocio.

Otra ventaja para los emisores es la posibilidad de sacar de su balance algunos activos, lo que permite mejorar las medidas de desempeño financiero que entregarán a sus

inversionistas, obteniendo además formas de financiamiento alternativas a los préstamos bancarios o las rondas de captación de capital en cualquiera de sus formas. En algunos casos, la emisión de estos instrumentos resulta incluso más barata que las formas de financiamiento antes mencionadas, pues independientemente de la calidad crediticia del emisor, los BRH se consideran “AA” o “AAA” si el mismo se queda con un segmento de las hipotecas de manera tal que absorba los posibles incumplimientos que, a juicio de una agencia calificadora externa, se presentarán durante la vida del instrumento con un margen de confianza determinado.

Estos activos también permiten a los bancos o entidades dedicadas a este tipo de financiamientos obtener parámetros de referencia (*benchmarks*) sobre el precio a cobrar por concepto de riesgo crédito y de prepago a los individuos que acuden a ellos para obtener un crédito hipotecario.

A pesar de todas sus bondades, estos instrumentos híbridos han demostrado no estar exentos de peligro, e incluso, ser potencialmente dañinos en caso de que las condiciones del mercado sean adversas. Las principales fuentes de riesgo nacen de la incertidumbre relativa al tiempo y monto de los pagos e incumplimientos. En ambos casos, los eventos no son siempre independientes entre aquellos que tienen una hipoteca, lo que afecta la duración y la convexidad del instrumento.

El mecanismo a través del cual opera el riesgo de prepago en un crédito hipotecario será esbozado mediante un ejemplo y después de una manera más formal. Suponga que un individuo representativo decide comprar una casa, y para ello contrata un crédito hipotecario. Por simplicidad suponga que el crédito contratado por éste individuo está pactado a una tasa de 15% a un plazo de 15 años.

Tiempo después, 4 años, el mismo individuo representativo recibe una oferta de empleo en otro estado del país, por lo que decide mudarse. Con esta decisión se ve forzado a refinanciar su préstamo al adquirir una nueva casa con una hipoteca a una tasa de 16% anual a un nuevo plazo de 10 años. Para financiar esta operación, decide vender su casa actual y con el dinero obtenido paga la primera hipoteca, dejando con ello al

tenedor del bono que incluía su primera hipoteca con un bono que teóricamente tenía una duración cercana a los 12 años, la cifra es una mera ilustración, con un bono cuya duración real es de aproximadamente 3.5 años. En este caso el prepago puede verse como un evento independiente y no representa mayor problema para el tenedor pues esta hipoteca representa una fracción pequeña del total de su activo.

¿Qué pasa si 4 años después, el mismo individuo representativo se da cuenta de que las tasas de interés han bajado y que puede refinanciar su crédito? Suponga que después de valorar todos los gastos colaterales, el nuevo crédito puede ser obtenido a una tasa del 10% con un nuevo plazo de 10 años, por lo que el agente económico analizado decide usar el nuevo préstamo para pagar el primero y de esta forma mejorar su perfil de pagos.

Si fuese un único individuo el que realiza esta operación, el tenedor del bono no sería afectado de manera significativa, el problema radica en que no sólo el agente económico abstracto usado en este análisis aprovechará la oportunidad. Es posible que otros individuos que estén en posibilidad de refinanciar y cuya hipoteca esté por encima del 10% actual decidan hacerlo, por lo que estos eventos, relacionados entre sí, implican un prepago masivo del bono, lo que deja al tenedor del primero bono con un instrumento que en teoría tenía una duración cercana a los 12 años, la cifra sigue siendo meramente ilustrativa, con un activo que en realidad tiene una duración de 3.5 años.

Para mostrar el problema del riesgo crédito se seguirá utilizando el caso de un individuo representativo. Se supone ahora que, por problemas personales, el sujeto teórico analizado pierde su empleo y no consigue otro sino hasta cuatro meses después, tiempo durante el cual no realiza los pagos de su hipoteca. Siendo este un evento aislado, *i.e.*, independiente, su ocurrencia estaba prevista en la emisión del bono y fue absorbida por el emisor dentro del segmento de hipotecas que conservó, por lo que el tenedor del bono no fue afectado.

Ahora suponga que es el país el que entra en una recesión grave y que ésta provoca un aumento drástico en la tasa de desempleo. En este caso, no sólo el individuo analizado perderá el trabajo, al igual que él, miles de personas quedarán cesantes. A esto es necesario añadir aquellas personas que verán reducidos sus ingresos y serán incapaces de cumplir con

sus obligaciones financieras, por lo que iniciarán los incumplimientos masivos.

En ciertos casos de concentración geográfica o concentración en el tipo de acreditado, los incumplimientos pueden sobrepasar la previsión del emisor del bono (las hipotecas que se quedaron con el emisor para minimizar el riesgo crédito), afectando al tenedor más allá de los prepagos efectuados por contrato ante incumplimiento. Este fenómeno implica la existencia del riesgo crédito en el bono.

Este parece un momento adecuado para explicar el mecanismo a través del cual el prepagado afecta a los tenedores del bono tomando como ejemplo el prepagado de una hipoteca. El ejemplo puede luego extenderse tanto como se desee ponderando el valor de la hipoteca dentro del bono. Suponga que el inversionista adquirió la hipoteca dentro del bono al 95% de su valor nominal, la compró bajo par. Si los individuos representativos prepagan, tendrán que hacerlo a valor “par”, lo que significa cubrir el total del crédito, beneficiando al tenedor. El detalle problemático con este supuesto es que si el bono fue vendido a descuento es porque en ese momento las tasas de interés se encontraban “altas”, *i.e.*, mayores a la tasa cupón, por lo que los individuos tienen pocos incentivos para prepagar.

Ahora suponga el caso contrario, en el que el tenedor compró la hipoteca dentro del bono “sobre par”, *V.gr.* 1.05 veces su valor nominal. En este caso, se puede pensar que las tasas de interés que prevalecen en el mercado son “bajas”, *i.e.*, menores a la tasa cupón, por lo que los individuos que tienen una hipoteca tienen incentivos para prepagar.

El quebranto al tenedor del bono surge porque el prepagado de los deudores de las hipotecas será hecho a valor “par”, por lo que el inversionista perderá la diferencia entre el precio de compra (\$105.00) y el valor “par”(\$100.00) para cada crédito hipotecario prepagado.

No sólo los factores mencionados en el relato pueden afectar el riesgo de prepagado o el riesgo crédito del bono, entendido como un conjunto de hipotecas. Existen otros factores que pueden afectarlos, alterando con ello el precio del instrumento, a saber:

- 1) La concentración de la cartera en una zona geográfica.

- 2) El tipo de crédito contratado (tasa fija o variable, en moneda nacional o extranjera, residencial o de interés social. *etc.*).
- 3) La concentración de la cartera en un tipo específico de deudor.
- 4) El aforo de la hipoteca.
- 5) La cantidad y calidad de los ingresos del individuo que solicitó el crédito.
- 6) El porcentaje cubierto del total de la hipoteca al tiempo t .
- 7) El cambio en los precios de los inmuebles.

Se iniciará el análisis de los factores de riesgo con el parámetro de incumplimiento, λ , que modela el riesgo crédito. Este parámetro será representado por el valor de los incumplimientos en el periodo estudiado, *i.e.*, $\lambda \in \mathbb{R}$. Se supondrá que este parámetro dependerá de la sobretasa de interés por refinanciamiento, r , dado que a altas sobretasas, el valor de los inmuebles cae y el agente económico se da cuenta de que es mayor el valor del crédito que el del inmueble, por lo que decide dejar que el acreedor se quede con el colateral.

Otro factor a considerar en el parámetro de incumplimiento es el cambio en el PIB⁵, ΔY , dado que una recesión implica una tasa de desempleo alta o por lo menos un ingreso personal menor, lo cual puede dificultar o impedir el pago de las mensualidades del crédito.

Con una lógica similar al mecanismo de la sobretasa por refinanciamiento, aunque independiente, es necesario tomar en cuenta el aforo presente en el crédito, E , dado que los individuos son más proclives a incumplir con los pagos y permitir que el acreedor se quede con el colateral si no tienen dinero propio comprometido en la propiedad. El tener dinero propio comprometido en la operación hace más costosa la decisión de incumplir, es posible notar que este parámetro permanece fijo una vez que es conocido al inicio del crédito.

Por último, aunque no menos importante, se debe tomar en cuenta la relación entre el ingreso mensual del deudor con el pago mensual de la hipoteca, C , pues éste indica

⁵ Producto Interno Bruto.

la flexibilidad del deudor para cubrir cabalmente sus obligaciones cuando se presentan gastos inesperados o el ingreso es reducido por cualquier razón. La relación entre todos los factores antes descritos puede ser resumida en:

$$\lambda = f(r, \Delta Y, \bar{E}, C). \quad (1.2.1)$$

Por otro lado, el parámetro de prepago, ψ , es afectado por la sobretasa en la que se incurre al refinanciar, r . El mecanismo de transmisión fue explicado en párrafos anteriores mediante la ilustración del individuo representativo.

Otro factor que incide en el parámetro de prepago es el porcentaje cubierto del total de la hipoteca al tiempo de valuación, χ , puesto que el agente económico no refinanciará su hipoteca si el monto a refinanciar es pequeño y los costos fijos de hacerlo son grandes, pues esto le generará un valor presente negativo.

Es importante aclarar que este parámetro no es dependiente de la trayectoria pues los prepagos por este concepto se harán sólo tomando en cuenta el último monto adeudado, por lo que se trata con un proceso Markoviano aunque el proceso que sigue el bono en su conjunto no lo sea.

Un factor que incide en el deseo de prepagar o de incumplir en los BRH es el cambio en el precio de los inmuebles, aunque lo hace por movimientos en la tasa por refinanciamiento y cambios en el PIB, por lo que resulta una variable dependiente de las anteriores y por lo tanto estadísticamente inútil⁶, razón por la cual será excluida del análisis. La relación entre todos estos factores de riesgo, y sus mecanismos de acción pueden ser resumidos como:

$$\psi = \psi(r, \chi). \quad (1.2.2)$$

Hasta el momento se han visto los factores de riesgo y los mecanismos a través de los cuales un individuo decide incumplir o prepagar su hipoteca. Ahora es el momento de

⁶ Teóricamente es posible construir el efecto de una variable dependiente de otras a través de una combinación lineal de las variables de las cuales depende, ésta es la idea básica de todo el análisis econométrico. Más aún, su presencia provocará problemas de multicolinealidad en los ejercicios econométricos

considerar la forma en que los agentes colocadores “arman” los lotes (en inglés *pools*) de hipotecas para poder emitir el bono. La forma en que son estructurados estos paquetes determina las características de riesgo de los bonos.

Es aquí donde factores como la concentración geográfica o en un tipo de deudor, pueden afectar el rendimiento del bono. Antes de seguir, se comentarán los mecanismos en que los factores de riesgo anteriores afectan al instrumento visto como un lote.

El primer factor a analizar en el prepago es la concentración en una zona geográfica, κ . La influencia de este factor se hace sentir en el caso de eventos extremos como un desastre natural o un hallazgo importante de petróleo o minas, *etc.*, eventos poco probables pero de gran impacto que al presentarse vulneran el valor del instrumento por las grandes oleadas de prepagos que se producen como respuesta a ellos.

También la concentración del instrumento en un único tipo de deudor, ζ , afecta al parámetro de prepago pues concentra el riesgo de prepago en individuos cuyo comportamiento no es totalmente independiente. Este factor de riesgo exhibe un mecanismo de transmisión y un tratamiento similar al de concentración en la zona geográfica.

Después de considerar los factores anteriores, es posible reescribir al parámetro de prepago como:

$$\psi = \psi(r, \chi, \varepsilon, \bar{\kappa}, \bar{\zeta}), \quad (1.2.3)$$

donde $\bar{\kappa}$ representa la concentración geográfica y $\bar{\zeta}$ la concentración en un solo tipo de acreditado, se recuerda que ambos parámetros permanecen constantes una vez integrados los “lotes”. En este punto, antes de revisar la integración de los mismos, es prudente revisar el modelo de prepago de la *Public Securities Association*⁹ (PSA por sus siglas en inglés), el cual establece el “100% PSA”, creado de forma empírica, como una forma de modelar los prepagos en un BRH.

⁹ Actualmente llamada *Bond Market Association*. Esta organización, con oficinas en Londres, Nueva York y Washington D.C., funge como portavoz global de emisores y negociadores de bonos, también sirve para coordinar a sus miembros con gobiernos, corporaciones e inversionistas de todo el orbe. En Noviembre de 2006 se fusionó con la *Securities Industry Association* para formar la *Securities Industry and Financial Markets Association*.

El modelo establece que, empezando con un prepago de cero al inicio de la vida del instrumento, los pagos anticipados aumentarán un 0.2% mensualmente hasta alcanzar un 6% en el mes treinta, para después mantenerse constantes a lo largo del resto de la vida del bono, esto es:

$$\psi = \begin{cases} t = 0 & \Rightarrow \psi = 0, \\ t \leq 30 & \Rightarrow \psi = .002t, \\ t > 30 & \Rightarrow \psi = .06. \end{cases} \quad (1.2.4)$$

Existen también una variación del modelo, llamada 150% PSA, en la que se permite que la tasa de prepago llegue hasta un 9% a través de un incremento mensual de 0.3%, lo que puede ser leído como:

$$\psi = \begin{cases} t = 0 & \Rightarrow \psi = 0, \\ t \leq 30 & \Rightarrow \psi = .003t, \\ t > 30 & \Rightarrow \psi = .09. \end{cases} \quad (1.2.5)$$

Además de este modelo, existe el llamado *Conditional Prepayment Rate*, CPR por sus siglas en inglés, el cual es una aplicación de los modelos de supervivencia usados en el análisis de riesgo crédito. El CPR indica, para cualquier año, la fracción de los adeudos por capital dentro del bono que no han sido prepagados y que lo serán a lo largo del mismo.

El modelo parte del supuesto de que se conoce la tasa de mortalidad mensual (SMM) o *Single Monthly Mortality*, la cual indica la tasa de adeudos por capital que no sobrevivirán al prepago a lo largo del mes, por lo que el modelo anualizado puede resumirse como:

$$\text{CPR} = 1 - [(1 - \text{SMM}_1)(1 - \text{SMM}_2) \cdots (1 - \text{SMM}_{12})], \quad (1.2.6)$$

lo que puede ser aproximado, si se supone que la SMM es la misma para todos los meses, por:

$$\text{CPR} = 1 - (1 - \text{SMM})^{12}. \quad (1.2.7)$$

Una vez que se conoce la forma en que es evaluado el parámetro de prepago, se revisará la forma en que se integran los “lotes” de hipotecas, de esta forma se hará patente la necesidad de replantear, como se hace en la tesis, la modelación del riesgo crédito y de prepago.

El proceso comienza cuando el originador (acredor) de las hipotecas elige a los acreditados en función a la capacidad que estima que éstos tienen para repagar el crédito, enfrentando obviamente un problema de selección adversa¹⁰. Para solucionarlo, típicamente se vale de señales como la posibilidad del acreditado para comprobar ingresos, tiempo en el mismo trabajo, número de dependientes, estado civil, ingresos libres, *etc.*

Una vez que el crédito es otorgado, el originador solicita la asistencia de algún intermediario financiero encargado de recolectar y organizar los pagos de las hipotecas creando un fideicomiso¹¹ para tal fin. Este fideicomiso es el nuevo dueño de esos activos y es el encargado de realizar los pagos de cupón y principal a los tenedores del bono.

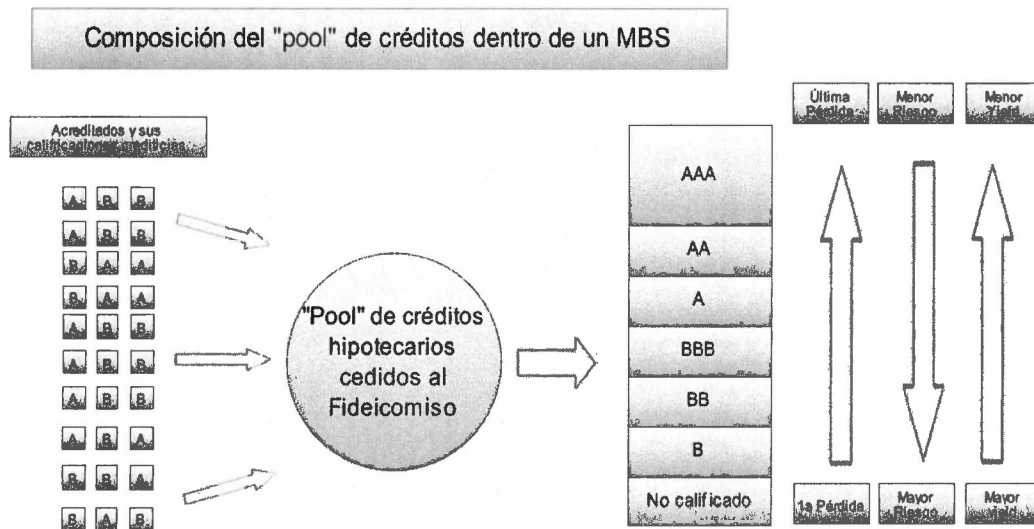
Para minimizar el riesgo de incumplimiento del instrumento, independientemente de su calidad crediticia, el originador se queda con una parte de los créditos, por lo que los primeros n incumplimientos, dependiendo de la calidad deseada para el bono, son tomados por el segmento (o *tranch* en la literatura de habla inglesa) de los créditos que quedan bajo custodia del originador.

El fideicomiso, como encargado de los pagos, hace entrega a los inversionistas del dinero producto de intereses, capital o ambos, dependiendo del tipo de MBS del que se trate, pero no es responsable por los prepagos existentes puesto que por contrato, el tenedor asume ese riesgo.

¹⁰ El problema de selección adversa es un problema de agencia estudiado por la Microeconomía, en este problema el “principal” (el originador) tiene menos información que el “agente” (el acreditado) para evaluar su posible desempeño. La solución microeconómica es la separación de mercados a través de señales de mercado, lo que se conoce como “discriminación en la asignación de crédito”.

¹¹ Llamado Vehículo de un Sólo Propósito o *Single Purpose Vehicle* en la literatura de Bursatilización.

Aunque existen “seguros” (instrumentos de riesgo emitido por algún otra institución) contra el prepago, éstos son vendidos por separado, con lo que se hace patente su naturaleza de mero “vehículo” a través del cual es hecha la bursatilización del instrumento. Este proceso puede ser descrito en el siguiente gráfico.



Gráfica 1.2.1 Integración de un MBS a partir de varios créditos hipotecarios

En este punto es prudente aclarar que las altas calificaciones crediticias ostentadas por estos instrumentos son producto tanto del respaldo gubernamental a algunas de las agencias que los emiten, como de la desagregación de la cartera, siempre y cuando el originador se quede con un número suficiente de créditos para absorber los n primeros incumplimientos y con ello proteger los niveles superiores de la desagregación.

En caso de incumplimiento (lo que es distinto de mora), el originador no pierde todo su dinero, pues el inmueble queda siempre como colateral y aunque su valor de mercado puede variar y su valor de recuperación puede caer por los gastos del proceso de realización de garantías, se mantiene como un porcentaje razonable de recuperación después del incumplimiento del bono i al tiempo j , δ_{ij} .

Una vez conocido el proceso de integración de los “lotes”, se procederá a detallar algunas de las clases de Bonos Respaldados por Hipotecas (BRH) existentes en los mercados.

Esta no pretende ser una lista exhaustiva, sólo pretende ejemplificar la gran variedad de instrumentos que se derivan de esta estructura básica.

- 1) *Pass through mortgage backed security*: Éste es el instrumento anteriormente descrito, el cual sólo junta los flujos derivados de las hipotecas y los entrega a los tenedores de los bonos, se puede subdividir, por el tipo de colateral en: *Residential mortgage backed security* o RMBS por sus siglas en inglés ¹² y *Commercial mortgage backed security* o CMBS por sus siglas en inglés ¹³. Estos instrumentos generalmente son divididos en segmentos (*tranches*) y empíricamente han mostrado tener una menor cantidad de prepagos que sus contrapartes residenciales. En español se les conoce como Certificados Bursátiles respaldados por Hipotecas.
- 2) *Collateralized mortgage obligation* o CMO: En este caso el MBS es dividido en segmentos (*tranches*) los cuales son hechos a partir de algún criterio en especial, *v.gr.* tiempo estimado de prepago. Cada sección es vendida por separado como un instrumento diferente.
- 3) *Stripped mortgage backed security* o SMBS: En este caso, cada pago es dividido en flujos para capital o intereses, para luego vender cada grupo de flujos como instrumentos diferentes. El primero (Certificado Bursátil Respaldado por Intereses Hipotecarios), que está formado por flujos de interés es llamado *Interest only stripped mortgage backed security* o IO. Mientras que el integrado por flujos para capital es llamado *Principal only stripped mortgage backed security* o PO, su nombre puede ser castellanizado como Certificado Bursátil respaldado por Capital Hipotecario.
- 4) *Adjustable rate mortgages* o ARM: Este BRH, es dirigido a segmentos de mercado con calidad crediticia menor a la óptima o *subprime* como se le denomina en el mercado estadounidense. Este tipo de instrumentos da a los acreditados una tasa baja de interés durante los primeros *n* años de vida del crédito para luego ajustarlo a una tasa variable más un diferencial por riesgo crédito. Es posible que al llegar al periodo de

¹² Éste es un *pass through* respaldado sólo por hipotecas residenciales.

¹³ Éste es un *pass through* respaldado sólo por hipotecas de inmuebles comerciales.

ajuste, el acreditado sea incapaz de cubrir sus pagos dada la alta tasa variable, según los expertos, éste fue el mecanismo de construcción de la crisis hipotecaria. Las clases más usuales de este instrumento son las “2-28”, “3-27” y “5-25”, sus nombres indican en la primer posición el periodo de intereses fijos y bajos, mientras que la segunda cifra indica periodos de pago a tasa variable.

Otra forma en que el mercado clasifica estos instrumentos es por la clase de sujetos de crédito a los que se les otorgó financiamiento. Esta clasificación es la clásica respuesta del mercado al problema de selección adversa antes explicado, que en última instancia, es la separación del mercado crediticio hipotecario en tres: el *prime*, el *Alt-A* y el *subprime*. A cada segmento se le carga un mayor diferencial por riesgo crédito que al segmento inmediato anterior para compensar las pérdidas esperadas. Por esta razón es que el mercado *subprime* atrae a individuos que no siempre tienen la intención o la capacidad de pagar los créditos, tal y como lo predice la teoría microeconómica en los problemas de selección adversa y separación de mercados.

El mercado *prime* y el *Alt-A* son los segmentos superiores del mercado hipotecario, siendo el *Alt-A* el segmento ubicado entre el *prime* y el *subprime*. Se considera una hipoteca *prime* como aquella que fue contratada por un individuo con un fuerte historial crediticio, *i.e.*, que no ha sufrido quiebras en los últimos siete años, cuenta con un cociente “pagos-ingreso” inferior al 75%, puede comprobar sus ingresos a través de documentación y no ha incumplido con los pagos de sus créditos vigentes en los últimos noventa días.¹⁴

Por otro lado, se cataloga a un acreditado como *Alt-A* cuando cuenta con un fuerte historial crediticio, pero no es capaz de comprobar sus ingresos o no cumple con alguno de los requisitos necesarios para alcanzar la categoría superior.

Por otro lado, las hipotecas *subprime* o de “segunda vuelta”, son aquellas que no cumplen con dos o más de los requisitos anteriormente mencionados. Por ésta razón, se supone que tienen una mayor probabilidad de incumplimiento y, como consecuencia, tienen

¹⁴ Estos son los estándares en los mercados norteamericanos, en nuestros mercados crediticios pueden cambiar algunos de los criterios.

una sobretasa por riesgo que incrementa la tasa de interés final. Aunque estos préstamos han sido fuente de controversia a raíz de la crisis hipotecaria reciente, suponen una única vía de acceso a bienes “mayores” para individuos que de otra forma no podrían acceder a un crédito hipotecario. Se podría pensar que éste es un mercado de crédito predatorio que genera un círculo vicioso, pues las altas tasas de interés hacen más difíciles los pagos del préstamo, con lo que se “obliga” a los acreditados a incumplir si las condiciones económicas son adversas, siendo muy comunes en el segmento instrumentos como los *ARM*'s.

En general, dentro del segmento *subprime* se encuentran créditos por montos menores a los otros segmentos, tasas de interés más altas y penalizaciones más severas en caso de retraso o incumplimiento con el pago del crédito. Estas condiciones en un entorno recesivo llevan a los acreditados a penalizaciones que a la larga los lleva a perder el inmueble adquirido a través del crédito.

Una vez revizados los tipos de BRH más comunes en el mercado, las fuentes de riesgo y sus mecanismos de transmisión, además de algunas particularidades del mercado, se explicará como es que el mercado ha valuado estos instrumentos hasta el momento.

Como se mencionó antes, es usual considerar al BRH como una nota estructurada, *i.e.*, como un bono que enfrenta riesgo crédito más una posición corta en una opción de compra americana; aunque en ocasiones, por simplicidad, se supone que la opción es europea, esto es:

$$\begin{aligned}
 B_{BRH} &= B + c_{americana} \\
 &= \sum_{i=1}^n d_i A + c_{americana} \\
 &\approx \sum_{i=1}^n d_i A + c_{europea} \\
 &\approx \sum_{i=1}^n d_i A + E \left[e^{-rt} \max(S_T - N, 0) \right],
 \end{aligned} \tag{1.2.8}$$

donde el precio de ejercicio de la opción es el nominal del valor del inmueble, N , y d_i es el factor de descuento apropiado para el i -ésimo periodo. La prima de la opción de compra se suma al valor del bono debido a que es el tenedor del bono el que se encuentra corto en

esta opción¹⁵ por lo que recibe la prima de la misma.

Aun cuando se suponga que el factor de descuento incluye un diferencial por riesgo crédito y que se realice la valuación de la opción americana, se deja de lado la estructura de dependencia que tienen los factores de riesgo, lo que ocasiona una valuación errónea del instrumento. Esta posición ha sido defendida en algunos artículos de la literatura existente¹⁶ por su simplicidad, aunque se reconocen las limitaciones del modelo. Lo que se propone en esta tesis como modelo simple es ver al bono como una función de sus factores de riesgo, esto es:

$$B = (r, \lambda(r, \Delta Y, \bar{E}, C), \psi(r, \chi(t), \varepsilon, \bar{\kappa}, \bar{\zeta})), \quad (1.2.9)$$

donde el aforo, E , es un parámetro que permanece fijo una vez que se ha otorgado el crédito, los cambios en el PIB (ΔY) y la razón de cobertura, C , son variables exógenas al modelo que por simplicidad se tomarán como incluidas dentro de los parámetros de prepago e incumplimiento. Estos dos factores de riesgo pueden ser modelados usando distribuciones “Beta” o distribuciones de valores extremos, por otro lado, factores como la concentración en una zona geográfica ($\bar{\kappa}$) o un sólo tipo de acreditado ($\bar{\zeta}$), permanecen constantes una vez otorgados los créditos y también serán modelados junto con los parámetros de prepago e incumplimiento.

Al seguir con el análisis de las fuentes de incertidumbre que afectan a los factores de riesgo de interés para esta tesis, se discutirá la importancia del monto restante sobre el pago de la hipoteca. Si se supone que el “lote” (*pool*) se elaboró con créditos con tiempos a vencimiento similares, entonces el capital insoluto es una función conocida del tiempo.

Después de haber analizado todos los factores anteriores, es posible pensar que la sobretasa por refinanciamiento es el factor de riesgo más importante para el modelo una vez que se han visto las realizaciones (exógenas) de las variables aleatorias que influyen

¹⁵ El tenedor del bono está obligado a vender su bono al acreditado, en el momento que éste lo desee, al valor nominal del inmueble que se quedó como colateral, liberando con esto al acreditado de cualquier obligación futura con el tenedor del bono.

¹⁶ Kalotay, A., D. Yang, and F. Fabozzi (2005). An Option-Theoretic Prepayment Model for Mortgages and Mortgage-Backed Securities. *International Journal of Theoretical and Applied Finance*.

en λ y ψ , aunque desafortunadamente las características del instrumento lo vuelven un proceso dependiente del punto en donde inicia, *i.e.*, no es Markoviano.

Una vez planteado el modelo, se procede a descomponer el BRH en bonos cupón cero, \bar{B} , que a su vez son descompuestos en sus factores de riesgo, usando una aproximación por serie de Taylor alrededor del cero dado que las tasas aplicables son pequeñas. Se toma como ejemplo un solo bono cupón cero, cuyos términos de primer orden representan la descomposición de los flujos del bono en los pagos a sus factores de riesgo. Los términos cruzados de segundo orden y superiores son sustituidos por cópulas formadas por distribuciones marginales de valores extremos. La familia de la cópulas será discutida más adelante y modelarán las dependencias entre los factores de riesgo. Lo anterior puede ser resumido en:

$$\begin{aligned} \bar{B}(\mathbf{r}, \lambda(\mathbf{r}, \Delta Y, \bar{\mathbf{E}}, C), \psi(r, \chi, \varepsilon, \bar{\kappa}, \bar{\zeta})) &= \bar{B}(\mathbf{r}, \lambda(\cdot), \psi(\cdot)) + \bar{B}_t(t - t^*) + \bar{B}_r(r - r^*) + \\ &\bar{B}_\lambda(\lambda - \lambda^*) + \bar{B}_\psi(\psi - \psi^*) + \\ &\frac{1}{2} \left[\bar{B}_{tt}(t - t^*)^2 + \bar{B}_{rr}(r - r^*)^2 + \bar{B}_{\lambda\lambda}(\lambda - \lambda^*)^2 + \bar{B}_{\psi\psi}(\psi - \psi^*)^2 \right] + \\ &\bar{B}_{tr}(t - t^*)(r - r^*) + \bar{B}_{t\lambda}(t - t^*)(\lambda - \lambda^*) + \bar{B}_{t\psi}(t - t^*)(\psi - \psi^*) + \\ &\bar{B}_{r\lambda}(r - r^*)(\lambda - \lambda^*) \\ &+ \bar{B}_{r\psi}(r - r^*)(\psi - \psi^*) + \bar{B}_{\psi\lambda}(\psi - \psi^*)(\lambda - \lambda^*) + R^3. \end{aligned}$$

Evidentemente, los términos cuadráticos asociados con t desaparecen por no tratarse de una variable aleatoria, por lo que el bono cupón cero en cuestión puede ser aproximado por:

$$\begin{aligned} d\bar{B}(\mathbf{r}, \lambda(\cdot), \psi(\cdot)) &= \bar{B}_t(t - t^*) + \bar{B}_r(r - r^*) + \bar{B}_\lambda(\lambda - \lambda^*) + \bar{B}_\psi(\psi - \psi^*) + \\ &\mathbf{E}[C(r, \lambda)] + \mathbf{E}[C(r, \psi)] + \mathbf{E}[C(\lambda, \psi)]. \end{aligned} \quad (1.2.10)$$

Si se reúnen todos los bonos cupón cero que representan cada uno de los pagos en el instrumento original, se tiene que:

$$\begin{aligned} dB(\cdot) &= \sum_{i=1}^n B_t dt + \sum_{i=1}^n B_r dr + \sum_{i=1}^n B_\lambda d\lambda + \sum_{i=1}^n B_\psi d\psi + \sum_{i=1}^n \mathbf{E}[C(r, \lambda)] + \\ &\sum_{i=1}^n \mathbf{E}[C(r, \psi)] + \sum_{i=1}^n \mathbf{E}[(\lambda, \psi)]. \end{aligned} \quad (1.2.11)$$

Una vez que se ha hecho explícita la ecuación de la aproximación, es necesario discutir las cópulas, su naturaleza y el tipo de cópula a usar. Esto será hecho en la próxima sección.

1.3 Conceptos básicos de Cópulas

Tal y como se puede observar al final de la sección anterior, es necesario encontrar una distribución conjunta que modele la interdependencia entre los parámetros de incumplimiento, prepago y la tasa por refinanciamiento. Estos factores de riesgo no parecen ser variables aleatorias que sigan una distribución normal, se hace esta presunción ya que se ha visto de forma empírica que las combinaciones de estas tres variables presentan más valores extremos que los pronosticados por la normal bivariada.

Una vez que se acepta la no normalidad de estas variables, se acepta también que la estructura de dependencia no es caracterizada por la matriz de varianza covarianza y que por lo tanto va más allá de lo que una medida lineal como la correlación puede mostrar. Para encontrar una función de probabilidad multivariada que cumpla con estas características es que se propone el uso de cópulas.

La noción de cópula fue introducida por A. Sklar en 1959 como respuesta a una pregunta formulada por el matemático francés M. Fréchet sobre la relación entre una función de probabilidad multivariada y sus marginales. Pronto pasó de ser un desarrollo en teoría de espacios de medida a un método que permite obtener distribuciones conjuntas de probabilidad, $F(X_1, X_2)$, a partir de funciones de distribución marginales conocidas, $F(X_1)$, $F(X_2)$, aun cuando las marginales no son funciones de densidad normales.¹⁷

Dicho de una forma sucinta, las cópulas son funciones que conectan las distribuciones multivariadas con sus marginales, esto es: si existe una función acumulada de probabilidad de dimensión m con funciones marginales unidimensionales, F_1, F_2, \dots, F_m , entonces existe una cópula m -dimensional, C , tal que $F(X_1, \dots, X_m) = C(F_1(X_1), \dots, F_m(X_m))$, siendo

¹⁷ Es prudente recordar que si se parte de una función de distribución de probabilidad normal multivariada, las distribuciones marginales que surjan de ella, serán funciones de distribución de probabilidad normales, aunque no a la inversa. Para más detalles se puede consultar: Ross, S.M.(1985), Introduction to Probability Models, Academic Press Inc., 3rd Ed., Florida, USA.

el caso bivariado el más estudiado.¹⁸ El término cópula fue introducido en 1959 por Sklar¹⁹, aunque la idea había aparecido en algunos artículos previos, *v.gr.* Hoeffding (1940, 1941), quien estableció cotas para estas funciones.

A continuación se presentará una rápida recapitulación del teorema de Sklar y de uno de los métodos de formación de cópulas más simples, dejando para el apéndice AI el tratamiento formal del teorema.

1.3.1 Teorema de Sklar: Dentro de la literatura de cópulas, este teorema es la piedra angular sobre la cual se levanta toda la teoría. En este momento sólo se hará un esbozo del trabajo original de 1959, el cual establece que una cópula de dimensión m es una función, C , que mapea un cubo unitario de dimensión m en el intervalo cerrado $[0, 1]$. La función, C , debe satisfacer las siguientes condiciones:

$$1) C(1, \dots, 1, a_n, 1, \dots, 1) = a_n \quad \forall n \leq m \text{ que además cumple con que } a_n \in [0, 1].$$

Esto es, dado que la cópula es una función acumulada multivariada de probabilidad, cumple con las características de éstas. Entre ellas, la de ser igual a su distribución acumulada marginal cuando ha acumulado toda la probabilidad en el resto de sus dimensiones.

$$2) C(a_1, \dots, a_m) = 0 \text{ si alguna de sus entradas } a_n = 0 \quad \forall n \leq m. \text{ Esto es que la cópula, al igual que cualquier función acumulada multivariada de probabilidad, no acumula densidad si alguna de sus entradas es cero pues no acumula volúmen en todo el } m\text{-cubo, i.e., se queda como un objeto de } m - 1 \text{ dimensiones. Coloquialmente se puede entender como "incapaz de levantarse".}$$

$$3) C \text{ es } m\text{-creciente. Esto indica que el volúmen del cubo } C \text{ en cualquier intervalo } m\text{-dimensional es no negativo.}$$

¹⁸ La mayor parte del apartado referente a cópulas se obtuvo de: Pravin K. T., D. Zimmer, Copula modeling: An introduction for Practitioners, Foundations and Trends in Econometrics (2005), Vol 1, No. 1, 1 - 111. y de Frees, E.W. Valdez, E., Understanding relationships using copulas (2001), North American Actuarial Journal, Vol 2, 1-25.

¹⁹ Sklar, A. (1959), Fonctions de repartition 'a n dimensions et leurs marges, Publ.Inst. Statist. Univ. Paris 8, 229-231.

Estas propiedades implican que una cópula es una función acumulada, $F(x_1, \dots, x_m)$, cuyo dominio está dado en el cubo m -dimensional $[0, 1]$ y que cada una de sus funciones acumuladas de distribución marginales, $F_1(x_1), \dots, F_m(x_m)$, tiene su dominio sobre el intervalo real $[0, 1]$, esto es:

$$F(x_1, \dots, x_m) = \Pr[U_1 \leq u_1, \dots, U_m \leq u_m] = C(u_1, \dots, u_m). \quad (1.3.1)$$

lo que indica que la cópula es única dadas la distribución acumulada multivariada y las marginales que se escogieron si estas son continuas, esto es $(F_1(x_1), \dots, F_m(x_m)) \sim C$ además de que si $U \sim C$, entonces $(F_1^{-1}(u_1), \dots, F_m^{-1}(u_m)) \sim F$.

1.3.2 Cópulas para modelado de factores de riesgo: Después de hacer un breve repaso de lo que es una cópula, (se recomienda revisar el apéndice AI para mayores detalles) se está en condiciones de preguntar por una cópula en específico que modele las relaciones entre los factores de riesgo. Para ello fue necesario repasar las propiedades y alcances de las cópulas (hecho en AI) y escoger las que permitan modelar mejor el problema en cuestión.

Se iniciará tomando el problema de la dependencia entre el incumplimiento y la sobretasa por refinanciamiento. En este caso, la teoría establece una relación de dependencia positiva, pues se ha visto de forma empírica que en general los instrumentos que presentan riesgo crediticio, incumplen poco, pero que cuando lo hacen, es como respuesta a entornos económicos donde la tasa de interés es alta, además de presentar, en ocasiones, incumplimientos masivos.

En el caso de los BRH, se tiene un “lote” de hipotecas que difícilmente se pueden suponer independientes pues en última instancia todas dependen de la misma economía y en algunos casos, por el tipo de cliente buscado por los agentes colocadores de los créditos, dependientes del mismo sector de la propia economía.

Sobre esta línea se puede suponer que en ambientes macroeconómicos de tasas de interés bajas, sean pocos los deudores que dejan de cumplir con sus hipotecas, casos aislados y fortuitos, que generan pocas minusvalías en el valor del bono.

Ahora suponga que ese mismo conglomerado de hipotecas es colocado en un ambiente macroeconómico de tasas de interés altas, donde es más probable que algún deudor incumpla con sus obligaciones. Si a ese ambiente se añade una recesión, se genera una situación adecuada para que buena parte de los deudores integrados en el “lote” sean incapaces de honrar sus compromisos, generando con ello una oleada de incumplimientos que provocan una fuerte pérdida por este concepto a los tenedores del bono.

La distribución empírica generada por este fenómeno es consistente con lo que en Estadística se conoce como “colas pesadas” y puede ser modelado haciendo uso de las distribuciones de valores extremos. Por lo tanto, toda la historia antes relatada, puede ser representada por una función de distribución de probabilidad de valores extremos. En este caso, debido a la comonotonicidad de las variables, *i.e.*, si baja la tasa de refinanciamiento disminuye la de incumplimiento y viceversa, se usará una distribución Weibull bivariada.

Estas distribuciones son conocidas por su capacidad de generar valores extremos, dependiendo del parámetro de forma, aunque sus contrapartes bivariadas son menos conocidas. La distribución Weibull bivariada es más usada en análisis epidemiológicos o biomédicos²⁰ y fue desarrollada en primer lugar por Freund²¹ en 1961 quien propuso una función Weibull de densidad de probabilidad bivariada de la forma

$$f(x_1, x_2) = \begin{cases} \theta_1 \theta_{22} \sigma^2 (x_1 x_2)^{\sigma-1} e^{-\theta_{22} x_2^\sigma - (\theta - \theta_{22}) x_1^\sigma} & 0 < x_1 < x_2 < \infty \\ \theta_2 \theta_{11} \sigma^2 (x_1 x_2)^{\sigma-1} e^{-\theta_{11} x_1^\sigma - (\theta - \theta_{11}) x_2^\sigma} & 0 < x_2 < x_1 < \infty \\ \theta_3 \sigma x^{\sigma-1} e^{-\theta x^\sigma} & 0 < x_1 = x_2 = x < \infty, \end{cases}$$

donde además $\theta = \theta_1 + \theta_2 + \theta_3$. Es posible comprobar que esta expresión constituye una función de densidad, esto se hace integrando hasta infinito cada uno de los segmentos sobre

²⁰ En particular, durante el desarrollo de este trabajo, se encontró la distribución en el artículo de Hanagal, D. A. (2005), Bivariate Weibull Regression Model, *Economic Quality Control*, Vol 20, No. 1, 1-8, Helder mann Verlag. En el que se hace una regresión por máxima verosimilitud para encontrar los parámetros de la Weibull bivariada que se usa para modelar tiempo de fallo de dos órganosopartes de sistemas que trabajan juntos.

²¹ Freund, J.E. (1961): A bivariate extension of the exponential distribution. *Journal of Amer. Statist. Assoc.* 56, 971-77.

la variable. Esto es:

$$\begin{cases} \int_0^{\infty} \int_0^{x_2} \theta_1 \theta_{22} \sigma^2 (x_1 x_2)^{\sigma-1} e^{-\theta_{22} x_2^{\sigma} - (\theta - \theta_{22}) x_1^{\sigma}} dx_1 dx_2 & 0 < x_1 < x_2 < \infty \\ \int_0^{\infty} \int_0^{x_1} \theta_2 \theta_{11} \sigma^2 (x_1 x_2)^{\sigma-1} e^{-\theta_{11} x_1^{\sigma} - (\theta - \theta_{11}) x_2^{\sigma}} dx_2 dx_1 & 0 < x_2 < x_1 < \infty \\ \int_0^{\infty} \theta_3 \sigma x^{\sigma-1} e^{-\theta x^{\sigma}} dx & 0 < x_1 = x_2 = x < \infty. \end{cases}$$

Este conjunto de integrales se resuelve con una sencilla sustitución de tal forma que $u = x_i^{\sigma}$, lo que implica que $du = \sigma x_i^{\sigma-1} dx_i$, con el consecuente cambio en los límites de la integral. Para hacer más claro el proceso mediante el cual se comprobó que esta función suma uno, se resolvera la primera integral:

$$\begin{aligned} F(x_1, x_2)_{0 < x_1 < x_2 < \infty} &= \int_0^{\infty} \int_0^{x_2} \theta_1 \theta_{22} \sigma^2 (x_1 x_2)^{\sigma-1} e^{-\theta_{22} x_2^{\sigma} - (\theta - \theta_{22}) x_1^{\sigma}} dx_1 dx_2 \\ &= \theta_1 \theta_{22} \sigma^2 \int_0^{\infty} x_2^{\sigma-1} e^{-\theta_{22} x_2^{\sigma}} \int_0^{x_2} x_1^{\sigma-1} e^{-(\theta - \theta_{22}) x_1^{\sigma}} dx_1 dx_2 \\ &= \theta_1 \theta_{22} \sigma^2 \int_0^{\infty} x_2^{\sigma-1} e^{-\theta_{22} x_2^{\sigma}} \int_0^{x_2^{\sigma}} x_1^{\sigma-1} e^{-(\theta - \theta_{22}) u} \frac{du}{\sigma x_1^{\sigma-1}} dx_2 \\ &= \theta_1 \theta_{22} \sigma \int_0^{\infty} x_2^{\sigma-1} e^{-\theta_{22} x_2^{\sigma}} \left[\left(\frac{-1}{\theta - \theta_{22}} \right) e^{-(\theta - \theta_{22}) u} \Big|_0^{x_2^{\sigma}} \right] dx_2 \\ &= \left(\frac{-\theta_1 \theta_{22} \sigma}{\theta - \theta_{22}} \right) \int_0^{\infty} x_2^{\sigma-1} e^{-\theta_{22} x_2^{\sigma}} \left[e^{-(\theta - \theta_{22}) x_2^{\sigma}} - 1 \right] dx_2 \\ &= \left(\frac{-\theta_1 \theta_{22} \sigma}{\theta - \theta_{22}} \right) \left[\int_0^{\infty} x_2^{\sigma-1} e^{-\theta x_2^{\sigma}} dx_2 - \int_0^{\infty} x_2^{\sigma-1} e^{-\theta_{22} x_2^{\sigma}} dx_2 \right] \\ &= \left(\frac{-\theta_1 \theta_{22} \sigma}{\theta - \theta_{22}} \right) \left(\frac{1}{\theta} - \frac{1}{\theta_{22}} \right) = \frac{\theta_1}{\theta}. \end{aligned}$$

Si se sigue este mismo proceso para cada una de las integrales, se llega a:

$$F(\cdot) = \begin{cases} \int_0^{\infty} \int_0^{x_2} \theta_1 \theta_{22} \sigma^2 (x_1 x_2)^{\sigma-1} e^{-\theta_{22} x_2^{\sigma} - (\theta - \theta_{22}) x_1^{\sigma}} dx_1 dx_2 = \frac{\theta_1}{\theta} & 0 < x_1 < x_2 < \infty \\ \int_0^{\infty} \int_0^{x_1} \theta_2 \theta_{11} \sigma^2 (x_1 x_2)^{\sigma-1} e^{-\theta_{11} x_1^{\sigma} - (\theta - \theta_{11}) x_2^{\sigma}} dx_2 dx_1 = \frac{\theta_2}{\theta} & 0 < x_2 < x_1 < \infty \\ \int_0^{\infty} \theta_3 \sigma x^{\sigma-1} e^{-\theta x^{\sigma}} dx = \frac{\theta_3}{\theta} & 0 < x_1 = x_2 = x < \infty, \end{cases}$$

lo que suma uno si se recuerda que $\theta = \theta_1 + \theta_2 + \theta_3$. Desafortunadamente esta función de densidad está segmentada, lo que llevaría a la no unicidad de la cópula. véase AI.

Otro intento por generar una distribución Weibull multivariada está dado por Hougaard (1986)²² quien propone una nueva clase de distribuciones continuas de supervivencia, es importante hacer notar que su investigación es consistente con la teoría de cópulas. El autor comienza por definir variables aleatorias T_{ij} que denotan el tiempo de supervivencia del individuo j que pertenece al grupo i . Se supone que, condicionando al grupo de pertenencia, las variables aleatorias son independientes, siendo su riesgo $Z_i \lambda_{ij}(t)$, el cual sigue una distribución estable²³ positiva con parámetro α de estabilidad. El autor llega a la transformada de Laplace para poder, usando la propiedad de la torre de las esperanzas, llegar al valor esperado incondicional de esta distribución. Esto es $E[e^{-sZ}] = e^{-s^\alpha}$, con un parámetro de estabilidad dentro del intervalo $(0, 1]$, siendo el 1 el caso de independencia y el caso de $\alpha = .5$ una normal inversa. Estas funciones resultan interesantes por presentar, en función al parámetro α , asimetría y exceso de curtosis.

A continuación, Hougaard hace explícito el parámetro acumulado de mortalidad como $\Lambda_{ij}(t) = \int_0^t \lambda_{ij}(u) du$, por lo que la probabilidad conjunta de supervivencia está representada por: $P(T_1 > t_1, \dots, T_n > t_n | Z = z) = e^{-z\Lambda}$. Para después suponer que los riesgos de mortalidad son aditivos, *i.e.* $\Lambda = \Lambda_1(t_n) + \dots + \Lambda_n(t_n)$ y obtener el valor esperado de la probabilidad condicional de los tiempos de muerte, con lo que queda:

$$P(T_1 > t_1, \dots, T_n > t_n) = e^{-\Lambda^\alpha}. \quad (1.3.2)$$

Aunque el artículo sigue con la función de verosimilitud para estas funciones, además de algunos lemmas para obtener sus cumulantes y momentos, se pasará al caso especial cuando el parámetro de mortalidad instantánea es $\lambda_j(t) = \epsilon_j \gamma t^{\gamma-1}$.

²² Philip, H. (1986), Survival Models for Heterogeneous Populations Derived from Stable Distributions, *Biometrika*, Vol. 73, No. 2., pp. 387-396.

²³ Se define como función estable si para copias independientes X_1, X_2 de la variable aleatoria X y cualquier par de constantes positivas a, b se tiene que $aX_1 + bX_2 \stackrel{d}{=} cX + d$ para algunas $c, d \in \mathbb{R}$. Por otro lado se dice que es estrictamente estable si $d=0$ y simétricamente estable si además de ser estable está simétricamente distribuida alrededor del cero. Las más conocidas por tener una función de densidad cerrada son las distribuciones de Cauchy, Lévy y Normal.

Si Z sigue una función de distribución estable, $P(\alpha)$, las distribuciones marginales son también Weibull pero con un parámetro de forma $\alpha\gamma$ y una función de mortalidad $\epsilon_j^\alpha \alpha \gamma t^{\alpha\gamma-1}$. Con lo que la ecuación 1.3.2 se reduce a $P(T_1 > t_1, \dots, T_n > t_n) = e^{-(\epsilon t_1^\gamma + \epsilon t_2^\gamma)^\alpha}$ que puede ser reescrita como:

$$\bar{F}(x_1, x_2) = e^{-\left[\left(\frac{x_1}{\lambda_1}\right)^{\frac{\gamma_1}{\alpha}} + \left(\frac{x_2}{\lambda_2}\right)^{\frac{\gamma_2}{\alpha}}\right]^\alpha}, \quad (1.3.3)$$

donde $\alpha \in (0, 1]$ es el parámetro de dependencia de la función, $0 < \lambda_1, \lambda_2 < \infty$ son las tasas individuales de mortalidad instantánea y $0 < \gamma_1, \gamma_2 < \infty$ son los parámetros de forma de las marginales. Lo que hace especial a esta función de densidad más allá de su continuidad es el hecho de que es coherente con la metodología de generación de cópulas explicada en el apéndice AII, pues se puede llegar a ella si se usa como generador la función $(-\ln(f(x)))^\alpha$.

En este punto, es posible saber que si sustituye la palabra muerte por incumplimiento en el desarrollo anterior, se tiene un modelo que ajusta a la medida al problema que atañe a esta tesis. También es posible notar que la ecuación 1.3.3 considera de una función acumulada de supervivencia \bar{F} (también llamada cópula de supervivencia), que será reemplazada por su cópula complemento, la cual está dada por:

$$F(x_1, x_2, \alpha) = 1 - e^{-\left[\left(\frac{x_1}{\lambda_1}\right)^{\frac{\gamma_1}{\alpha}} + \left(\frac{x_2}{\lambda_2}\right)^{\frac{\gamma_2}{\alpha}}\right]^\alpha}.$$

Se puede comprobar por mera inspección que la función anterior es una función de probabilidad acumulada, basta con tomar límites a la función cuando sus argumentos tienden a infinito y verificar que ésta tiende a uno, esto es: $\lim_{x_1 \rightarrow \infty, x_2 \rightarrow \infty} [F(x_1, x_2)] = 1$.

Dada toda la explicación anterior, se usará esta función Weibull bivariada como base para desarrollar la cópula, donde las distribuciones marginales de los parámetros de riesgo quedan por definir, lo que en realidad no afecta la elección de la cópula.

Aunque los detalles técnicos de esta irrelevancia son explicados con detenimiento en AI, basta con decir que la cópula es una función que tiene como argumentos las funciones

de probabilidad de las marginales, cuyo rango siempre está en $[0, 1]$. Por este motivo es que la marginal de la cual provenga esa función de cuantiles resulta irrelevante, lo que le da una gran flexibilidad al modelado por cópulas y permite dejar para después la elección de las marginales y la conformación final de la misma.

Ahora se hará referencia a la segunda relación funcional de importancia en el problema de modelado, esta es la dependencia del prepago con la sobretasa por refinanciamiento. Tal y como se explicó en el ejemplo del individuo representativo, existen incentivos para los individuos a prepagar si la tasa de interés cobrada por su hipoteca cae por debajo de la que fue contratada inicialmente, y esto se da, *ceteris paribus*, si cae la tasa libre de riesgo crédito. Lo que conlleva un fuerte quebranto para el tenedor del bono si el prepago es masivo.

Un primer intento de modelación implicó el uso de otra cópula Weibull, lo que no es posible pues el parámetro de dependencia está confinado al intervalo $(0, 1]$ lo que conlleva a una dependencia positiva. El segundo intento consistió en pensar en una cópula de Frank, pero su incapacidad de generar valores extremos imposibilitó sus uso.

Revizando la bibliografía existente, se abrió la posibilidad de usar la cópula propuesta por Ali-Mikhail-Haq, la cual fue usada en el artículo de Shiau, Wang y Tsai (2006) en una aplicación de un artículo de hidrología²⁴, ciencia que es la cuna de la teoría de los valores extremos, en la que se ajustaban registros de lluvias e inundaciones a esta distribución bivariada, obteniendo excelentes ajustes.

Para el desarrollo de la investigación se ha optado por esta distribución bivariada dada su capacidad de generar valores extremos aún cuando la estructura de dependencias sea negativa²⁵, pues el parámetro de dependencia está confinado al intervalo $[-1, 1]$. A continuación sólo se mostrará esta distribución bivariada, relegando el desarrollo de la

²⁴ Shiau, J., Wang, H., and Tsai, C. (2006), Bivariate frequency analysis of floods using copulas, *Journal of the American Water Resources Association, American Water Resources Association*.

²⁵ Lo que no es el caso de las lluvias ni de las inundaciones.

misma al apéndice AII.

$$C(x_1, x_2; \alpha) = \frac{x_1 x_2}{1 - \alpha((1 - x_1)(1 - x_2))} \quad (1.3.4)$$

Por inspección es posible saber que el caso de independencia se presenta cuando el parámetro de dependencia es igual a cero, $\theta = 0$, pues la cópula se reduce al producto de las marginales cuando esto ocurre, *i.e.* $C(x_1, x_2, 0) = x_1 x_2$. También es prudente hacer patente el que esta cópula no alcanza las cotas de Fréchet-Hoeffding. De nuevo se deja para después la elección de las distribuciones marginales.

Ahora, sólo queda por definir la cópula que modelará la relación entre los parámetros de prepago y de incumplimiento. Sus mecanismos de actuación ya han sido explicados a lo largo del trabajo por lo que se puede suponer que no existe una relación especialmente fuerte entre ellos y por lo tanto la cópula necesaria deberá reflejar este hecho.

La cópula elegida es la cópula FGM (véase AII para mayor referencia) que no es otra cosa sino una perturbación a la cópula independiente, esto es:

$$C(x_1, x_2, \alpha) = x_1 x_2 (1 + \alpha(1 - x_1)(1 - x_2)). \quad (1.3.5)$$

Un punto importante dentro del modelado de estas relaciones, y que no ha sido tocado hasta el momento, es el grado de dependencia entre las variables. Para conocerlo, se hará uso de medidas de dependencia no paramétricas que buscan establecer, en términos de cópulas, el grado de comonotonidad o contramonotonidad entre las variables usando las realizaciones muestrales con las que se cuentan.

1.4 Medidas de dependencia no paramétricas.

Una vez definidas las funciones acumuladas bivariadas que darán origen a las cópulas, es necesario conocer el grado de dependencia que existe entre las variables. Una primera mirada al problema llevaría a pensar que buscar la correlación muestral entre los datos y luego confiar en su convergencia asintótica, sería un buen inicio para conocer el grado de

dependencia, pues esta es la forma en que tradicionalmente se resuelve el problema dentro del paradigma de distribución normal multivariada.

Desafortunadamente, la naturaleza extrema de los datos ha alejado la investigación de este paradigma, llevandola a la teoría de cópulas y de funciones de valores extremos, por lo que nuevas medidas de dependencia son necesarias.

Antes de iniciar la búsqueda de medidas de dependencia, es necesario hacer una breve pausa para mencionar algunas de las cualidades que dentro de la literatura (Cherubin *et. al.*, 2004) se piden para una buena medida, a saber:

- a) **Simetría:** $\mu(X, Y) = \mu(Y, X)$. Esto es que la medida no cambie si lo hace el orden de las variables aleatorias.
- b) **Normalización:** $-1 \leq \mu(X, Y) \leq 1$. La medida es igual a -1 para el caso de contramonotonidad perfecta y de 1 para comonotonidad perfecta, lo que la hace comparable en escala.
- c) **Invariabilidad:** Para una transformación estrictamente monótonica, $T : \mathbb{R} \rightarrow \mathbb{R}$, de X , se tiene que: $\mu(T(X), Y) = \begin{cases} \mu(Y, X) T & \text{es creciente} \\ -\mu(Y, X) T & \text{es decreciente.} \end{cases}$

Después de revizar la literatura, tres medidas de dependencia surgen repetidamente, a saber: La “Tau” de Kendall, la “Rho” de Spearman y la “Beta” de Blomqvist. Estas medidas resultan interesantes para los propósitos de la tesis pues arrojan luz sobre el grado de dependencia de dos variables aleatorias sin importar la naturaleza de sus funciones de distribución marginales, pues se trata de medidas de dependencia no paramétricas.

Básicamente estas medidas capturan el grado de comonotonidad, concordancia si se desea ver en términos más simples, de dos variables aleatorias al obtener una “frecuencia relativa” del numero de veces que van en el mismo sentido, aunque sus metodologías difieren y el número de cálculos requerido para obtenerlas puede ser prohibitivo en algunos casos.

1.4.1 Tau de Kendall: Se iniciará el recorrido por las medidas de dependencia con la “Tau” de Kendall. Esta medida fue inicialmente intuida por G.T. Fechner alrededor de 1900, aunque desarrollada por M.G. Kendall en 1938. Básicamente es una medida de

concordancia entre los elementos (x_j, y_j) y (x_k, y_k) de la muestra $\{(x_i, y_i)\}_{i=1}^n$ que fue obtenida a su vez de una población bivariada.

Se dice que los elementos son concordantes si se cumple al mismo tiempo con $x_j < x_k$ y $y_j < y_k$ o con $x_j > x_k$ y $y_j > y_k$, lo que implica que $(x_j - x_k)(y_j - y_k) > 0$. Siguiendo la misma idea se dice que los elementos son discordantes si se cumple que $x_j < x_k$ y $y_j > y_k$ o con $x_j > x_k$ y $y_j < y_k$, lo que implica que $(x_j - x_k)(y_j - y_k) < 0$. La medida requiere que se verifique la concordancia o discordancia para cada par de elementos en la muestra, lo que lleva a calcular $C \binom{n}{2}$ pares de elementos para una muestra de tamaño n .

Para calcular la “Tau”, se le llama s a la diferencia entre observaciones concordantes, c , menos las discordantes, d , siendo el estimador

$$\tau_n = \frac{c - d}{c + d} = \frac{s}{C \binom{n}{2}} = \frac{2s}{n(n-1)}. \quad (1.4.1)$$

Aunque en el caso de encontrar observaciones iguales o “empates”, u , el estimador se ajusta a:

$$\tau_n = \frac{s}{\sqrt{n(n-1)/2} - T \sqrt{n(n-1)/2} - u}$$

Para dimensionar el problema, se puede verificar fácilmente que de una muestra de cien observaciones bivariadas, se desprende la necesidad de calcular la concordancia de $\binom{99}{2} \binom{100}{2}$ pares de observaciones. Lo que hace que su cálculo sea prohibitivo para muestras relativamente grandes, y aunque se ha probado la convergencia de este estimador a una distribución normal, esta sólo se alcanza de forma asintótica.²⁶

Por otro lado, existe una contraparte poblacional para el estimador, la cual es calculada a través de²⁷

$$\tau_n = 4 \int_0^1 \int_0^1 C_{X,Y}(u, v) dC_{X,Y}(u, v) - 1, \quad (1.4.2)$$

²⁶ En algunos documentos como: Hervé Abdi (2007), The Kendall Rank Correlation Coefficient, *Encyclopedia of Measurement and Statistics*, Neil Salkind (Ed.). Thousand Oaks (CA), se plantea que la convergencia es aceptable a partir de diez observaciones y no las tradicionales treinta usualmente requeridas.

²⁷ Vea el trabajo de R.B. Nelsen (1999).

dicho estimador es generalmente reportado en los textos académicos que tratan sobre cópulas. La doble integración, referida al área de dominio de la cópula, puede representar una complicación técnica, aunque heurísticamente se puede entender como:

$$\begin{aligned}\tau &= P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0] \\ &= \text{corr}[\text{sgn}(X_1 - X_2), \text{sgn}(Y_1 - Y_2)].\end{aligned}$$

Que no es otra cosa que la probabilidad de concordancia menos la de discordancia o la correlación de los signos de los pares de observaciones. Tal como se puede observar, esta medida es invariante bajo transformaciones monótonas no decrecientes, propiedad que comparte con las funciones de cópula y que la hace deseable como medida de dependencia poblacional más no muestral.

Esta medida también es simétrica, pues su computo no se altera si se invierte el orden de las variables aleatorias ya que usa todas las combinaciones posibles en ella. También se dice que esta es normalizada, pues en su definición se ve que es una comparación entre el exceso de pares coincidentes sobre el total de combinaciones posibles.

Buscando facilitar su uso, algunos autores²⁸ han buscado relaciones que simplifiquen su cálculo, ejemplo de esto es el trabajo de Lindskog, McNeil, Schmock y Kendall (2001) en el que obtienen un estimador de la “Tau” de Kendall fuertemente relacionado con la correlación para cualquier distribución elíptica²⁹, y aunque esta familia de distribuciones puede ser usada en el modelado de colas pesadas, la extensión propuesta está lejos de hacer simple el cálculo de la medida propuesta por Kendall. A continuación se muestra el resultado de estos investigadores.

$$\tau(X_i, X_j) = \left(1 - \sum_{x \in \mathbb{R}} (P\{X_i = x\})^2\right) \frac{2}{\pi} \arcsin \rho_{ij}$$

²⁸ F. Lindskog, McNeil, A., Schmock, U., Kendall, R.B. (2001), Kendalls Tau for elliptical distributions, Research supported by Credit Suisse, Swiss Re and UBS through RiskLab, Switzerland.

²⁹ Se dice que un vector aleatorio $(p \times 1)$ tiene una distribución elíptica con un vector de medias $\mu(p \times 1)$ y una matriz de dispersión, definida y no negativa, $\Sigma(p \times p)$ además de una función escalar, que sólo existe si hay una función característica, $\phi: [0, \infty) \rightarrow \mathbb{R}$ cuando su función característica puede ser expresada como $\psi(t) = e^{it^T \mu} \phi(t^T \Sigma t)$, manteniendo esta propiedad cualquier combinación lineal de variables elípticas.

1.4.2 Rho de Spearman: Esta es una medida muy similar a la Tau de Kendall, y puede ser vista como un caso especial del coeficiente de correlación de Pearson, en donde no es necesario suponer que la relación entre las variables aleatorias es lineal. Aunque si supone que las observaciones son equidistantes entre sí, lo que la hace inviable para los propósitos de esta investigación.

Con la finalidad de proporcionar una visión completa de la medida, se mostrará la fórmula con la cual se obtiene y se aclara que para llevar a cabo los cálculos de la “rho” es necesario jerarquizar las observaciones, siendo la más pequeña la número uno, $d = 1$, con lo que la medida muestral está dada por: $\rho = 1 - \frac{6 \sum d_i^2}{n(n^2-1)}$, mientras que su contraparte poblacional es:

$$\rho = 12 \int_0^1 \int_0^1 \{C(u_1, u_2) - u_1 u_2\} du_1 du_2.$$

Para enfatizar su similitud con la “Tau” de Kendall, se mencionará que Durbin y Stuart (1951) demostraron que la relación existente entre estas dos medidas es:

$$\frac{3}{2}\tau - \frac{1}{2} \leq \rho \leq \frac{1}{2} + \tau - \frac{1}{2}\tau^2 \quad \forall \tau \geq 0$$

$$\frac{1}{2}\tau^2 + \tau - \frac{1}{2} \leq \rho \leq \frac{3}{2}\tau + \frac{1}{2} \quad \forall \tau \leq 0.$$

1.4.3 Beta de Blomqvist: Dado lo constrictivo de los supuestos sobre los que se basa la “Rho” de Spearman y la dificultad computacional que representa la “Tau” de Kendall, que sirvió como referencia para la dependencia en la literatura de cópulas, fue necesario profundizar en la literatura para encontrar otra medida de dependencia que conjugara un cálculo relativamente simple con la plasticidad necesaria para capturar relaciones no lineales. Esta medida es la “Beta” de Blomqvist.

En su forma muestral, la β es una proporción de comonotonicidad sobre el total de las observaciones. Para su cálculo, es necesario obtener la media de ambas variables aleatorias, dividiendo con estas medias el plano cartesiano en cuatro cuadrantes. Siendo los cuadrantes I y III los cuadrantes comonotónicos, esto significa que ambas variables aleatorias estuvieron por encima de sus medias (en el primer cuadrante) o que ambas variables estuvieron debajo de sus medias (tercer cuadrante).

Una vez hecha esta división, se suma el número de observaciones dentro del primer y tercer cuadrantes, n_1 , y por otro lado se suman las que cayeron dentro del segundo y cuarto cuadrantes, contramonotónicos, para después compararlo con el número total de observaciones, esto es:

$$\beta = \frac{n_1 - n_2}{n_1 + n_2} = \frac{2n_1}{n_1 + n_2} - 1, \quad (1.4.3)$$

mientras que su contraparte poblacional es:

$$\beta = P\{(X - \bar{x})(Y - \bar{y}) > 0\} - P\{(X - \bar{x})(Y - \bar{y}) < 0\}. \quad (1.4.4)$$

Como se puede observar, la β es computacionalmente más simple que el resto de las medidas, siendo capaz de captar dependencias no lineales. Por la forma en que es obtenida, también se le conoce como coeficiente de correlación medial, aunque el usar la media como punto de comparación para las observaciones es arbitrario.

Siguiendo esta idea, y aunque la mayor parte de la literatura está hecha tomando a la media como punto de referencia, la β ha sido usada como medida de dependencia para las colas cuando se establece arbitrariamente un umbral después del cual se dice que aparecen valores extremos, aunque esa no será la línea de investigación de la tesis pues implica el uso de distribuciones de la familia de Pareto para el modelado de los picos después del umbral.

Otra de las grandes ventajas que ofrece la “beta” es la posibilidad de ser expresada en función de las propias cópulas, para luego obtener intervalos de confianza que servirán para validarlas.³⁰ En su artículo, Schmid y Schmidt (2007) parten de la definición poblacional de la Beta de Blomqvist, ecuación (1.4.4), de la cual sustituyen el segundo sumando por su complemento, obteniendo:

$$\begin{aligned} \beta &= 2P[(X - \bar{x})(Y - \bar{y}) > 0] - 1 \\ &= 2\{P[X < \bar{x}, Y < \bar{y}] + P[X > \bar{x}, Y > \bar{y}]\} - 1 \\ &= 2\{C(\bar{x}, \bar{y}) + [1 - F(\bar{x}) - G(\bar{y}) + C(\bar{x}, \bar{y})]\} - 1 \\ &= 4C(\bar{x}, \bar{y}) - 1, \end{aligned}$$

³⁰ Friedrich Schmid & Rafael Schmidt (2007), Nonparametric inference on multivariate versions of Blomqvists beta and related measures of tail dependence, *Metrika*, 66, 323 - 354.

lo cual no es otra cosa que observar a la probabilidad expresada en el primer sumando como dos veces la masa de probabilidad acumulada en la cópula bivariada en el quinto cuantil. El salto cualitativo de su investigación se da cuando relacionan este concepto con la cópula independiente, véase AI para mayor referencia, lo que lleva a:

$$\beta = \frac{C(1/2) - \Pi(1/2) + \bar{C}(1/2) - \bar{\Pi}(1/2)}{M(1/2) - \Pi(1/2) + \bar{M}(1/2) - \bar{\Pi}(1/2)}. \quad (1.4.5)$$

Esto hace a la β de Blomqvist una medida de distancia entre la cópula y la función bivariada independiente. En su artículo, los autores generalizan el estimador para que este sea útil en funciones multivariadas, además de hacer constar el hecho de que la cota superior de Fréchet, $M(\cdot) = \min(F_1, \dots, F_n)$, valuada en $1/2$ es precisamente $.5$, y que la cópula independiente, función de bivariada que supone independencia, es igual a 2^{-d} , donde d es el número de dimensiones de la cópula y h_d es una constante de normalización. Esto es el número de marginales que acaban de acumular la mitad de su densidad. Por lo que el estimador queda como:

$$\begin{aligned} \beta &= \frac{2^{d-1}}{2^{d-1} - 1} \{C(1/2) + \bar{C}(1/2) - 2^{1-d}\} \\ &= h_d(1/2) \{C(1/2) + \bar{C}(1/2) - 2^{1-d}\}. \end{aligned} \quad (1.4.6)$$

Como se dijo anteriormente, el tomar las medias de las marginales, *i.e* valorar la cópula en el vector $\mathbf{1}/2$, es un procedimiento arbitrario, pues es posible usar la β para valorar dependencia en las colas, *v.gr.* en el vector \mathbf{u}, \mathbf{v} . Esto lleva las constantes de normalización a: $h_d(u, v) = \{\min(u_1, \dots, u_d) + \min(1 - v_1, \dots, 1 - v_d) - \prod_{i=1}^d u_i - \prod_{i=1}^d (1 - v_i)\}^{-1}$ y $g_d(u, v) = \prod_{i=1}^d u_i - \prod_{i=1}^d (1 - v_i)$. Lo que en realidad no altera el parámetro, él cual sólo se modifica para hacer explícito el nuevo umbral, esto es:

$$\beta = h_d(u, d) \{C(1/2) + \bar{C}(1/2) - g_d(u, v)\}. \quad (1.4.7)$$

Con esto en mente, los autores dan un estimador del parámetro de dependencia y demuestran convergencia asintótica a una normal, $\sqrt{n}(\hat{\beta}_n - \beta) \xrightarrow{d} Z \sim N(0, \sigma^2)$, siendo

el parámetro muestral:

$$\hat{\beta}_n = h_d \left\{ \frac{1}{n} \sum_{j=1}^n \left(\prod_{i=1}^d 1_{\{U_{i,j} \leq \frac{1}{2}\}} + \prod_{i=1}^d 1_{\{U_{i,j} > \frac{1}{2}\}} \right) - 2^{1-d} \right\}. \quad (1.4.8)$$

Afortunadamente, en este trabajo sólo se usarán cópulas bivariadas de las cuales se conocen, por supuesto, las distribuciones marginales y la distribución bivariada. Esto permite simplificar el estimador a: $\hat{\beta}_n = 4 \sum_{j=1}^n 1_{\{U_{1,j} \leq \frac{1}{2}, U_{2,j} \leq \frac{1}{2}\}} - 1$. Lo que implica una varianza asintótica de: $4 \left[C(1/2) - \{C(1/2)\}^2 \right]$.

1.5 Distribuciones marginales y formación de cópulas

Una vez establecido el método a través del cual se determinará la dependencia muestral y las funciones bivariadas (cópulas), es momento de establecer que funciones de densidad marginales serán ajustadas a los datos. para hacerlo, es necesario considerar tanto el sesgo de los datos como la existencia de observaciones extremas.

Se propone para modelar las distribuciones marginales funciones Fréchet y Weibull (para datos con sesgo a la derecha e izquierda respectivamente), las cuales son comunes dentro de la literatura de valores extremos, para mayor referencia véase AIII, las cuales son respectivamente:

$$\begin{aligned} \text{Fréchet : } & G_{1,\alpha}(x) = e^{-x^{-\alpha}} & x \geq 0 \\ \text{Weibull : } & G_{2,\alpha}(x) = e^{-(-x)^{-\alpha}} & x \leq 0. \end{aligned} \quad (1.5.1)$$

Se puede observar que las funciones marginales son similares y que de hecho sólo difieren sustancialmente en el dominio. Este mismo hecho fue notado por Von Mises (1936) y por Jenkinson (1955) de forma independiente, lo que desembocó en una parametrización que es capaz de captar este hecho. La parametrización gamma es básicamente una sustitución en los parámetros de ubicación, $\mu = 1$, y escala, $\alpha = 1/\sigma$, de las funciones acumuladas, donde $\gamma = 1/\alpha$, lo que conduce a:

$$G_\alpha(x) = e^{-\left(1 + \frac{1}{\alpha}x\right)^{-\alpha}} = e^{-\left(\frac{\alpha+x}{\alpha}\right)^{-\alpha}} = G_\gamma(x) = \exp \left[-\left(1 + \gamma x\right)^{-\frac{1}{\gamma}} \right].$$

siendo el mismo procedimiento para la distribución acumulada Weibull. En este caso, se salva el inconveniente del segundo signo negativo al suponer que gamma es negativa lo que nos deja con una misma forma para las dos distribuciones, a saber:

$$\begin{aligned} \text{Frechet} \quad G_\gamma(x) &= \exp \left[- (1 + \gamma x)^{-\frac{1}{\gamma}} \right] \quad \gamma > 0 \\ \text{Weibull} \quad G_\gamma(x) &= \exp \left[- (1 + \gamma x)^{-\frac{1}{\gamma}} \right] \quad \gamma < 0. \end{aligned}$$

Siguiendo la metodología de cópulas, es necesario sustituir las funciones de densidad marginales acumuladas en las funciones bivariadas, lo que conduce a:

$$C(u_1, u_2; \alpha) = \exp \left\{ - \left[\left(\frac{\Omega_1}{\lambda_1} \right)^{\frac{\gamma_1}{\alpha}} + \left(\frac{\Omega_2}{\lambda_2} \right)^{\frac{\gamma_2}{\alpha}} \right] \right\}, \quad (1.5.3)$$

donde $\Omega_i = \exp \left\{ - \left(1 + \gamma_i \left(\frac{x_i - \mu_i}{\sigma_i} \right) \right)^{-\frac{1}{\gamma_i}} \right\}$.

La ecuación anterior representa la cópula que se usará para modelar la dependencia entre la sobretasa por refinanciamiento, r , y el parámetro de prepago, ψ , modeladas con distribuciones marginales de valores extremos.

Para la dependencia existente entre la sobretasa por refinanciamiento y el parámetro de incumplimiento, se usarán distribuciones marginales Fréchet pues en ambos casos las variables pueden experimentar grandes saltos positivos y nunca valores negativos. Esto es especialmente cierto en el caso de la sobretasa por refinanciamiento, la cual mantiene sus valores positivos aún cuando la economía experimenta periodos recesivos, mostrando en algunos casos tasas bajas como parte de una política económica contracíclica de la autoridad monetaria, o tasa altas en periodos de hiper inflación, cuando las tasas nominales “compiten” con la inflación para mantener atractivos los instrumentos de renta fija.

Por otro lado, el mecanismo de mercado que se supone para el incumplimiento, λ , ya ha sido explicado con anterioridad por lo que la exposición se limitará a usar la función de probabilidad Fréchet (ecuación 1.5.2) dentro de la cópula Ali-Mikhail-Haq (ecuación 1.3.4), lo que lleva a:

$$C(u_1, u_2; \alpha) = \frac{\Omega_1 \Omega_2}{1 - (\alpha (1 - \Omega_1) (1 - \Omega_2))}, \quad (1.5.4)$$

de nuevo $\Omega_i = \exp \left\{ - \left(1 + \gamma_i \left(\frac{x_i - \mu_i}{\sigma_i} \right) \right)^{-\frac{1}{\gamma_i}} \right\}$.

Siguiendo con el orden de aparición de las cópulas, es el turno de la última de ellas, la FGM (ecuación 1.3.5). Esta será usada para modelar la relación entre los parámetros de prepago e incumplimiento y su forma definitiva será construida a partir del mismo método que las dos anteriores, es decir sustituyendo las funciones de probabilidad marginales, lo que conduce a:

$$C(.) = \Omega_1 \Omega_2 (1 + \alpha (1 - \Omega_1) (1 - \Omega_2)). \quad (1.5.5)$$

al igual que en los casos anteriores, $\Omega_i = \exp \left\{ - \left(1 + \gamma_i \left(\frac{x_i - \mu_i}{\sigma_i} \right) \right)^{-\frac{1}{\gamma_i}} \right\}$.

Con estos avances, sólo resta explicar el método de estimación (máxima verosimilitud en dos etapas) para estas funciones, lo que se hará en la siguiente sección.

1.6 Máxima verosimilitud en dos etapas.

A lo largo de este trabajo, se ha podido notar que uno de los principales problemas en la aproximación planteada es la parametrización de las interdependencias entre los distintos factores de riesgo. Aunque ya se ha determinado que esta será realizada usando cópulas arquimedeanas, quedaba pendiente la búsqueda econométrica de las mismas.

La dificultad econométrica radica en la existencia de colas pesadas en las distribuciones marginales, así como la presencia de valores extremos dentro de las propias funciones bivariadas. Estas formas funcionales excluyen por definición las estimaciones por Mínimos Cuadrados Ordinarios y a través del Método Generalizado de Momentos, ya que el primero supone normalidad en regresores y perturbaciones para hacer cualquier clase de pruebas de hipótesis que ayudarían a validar las inferencias. En el caso del Método de Momentos, la existencia de colas demasiado pesadas haría imposible obtener estimaciones más allá del $1 - j/\alpha$ momento muestral³¹.

³¹ En el apéndice AIII se menciona que para las distribuciones de valores extremos sólo existen momentos muestrales mientras $\Gamma(1 - \frac{j}{\alpha})$ exista, lo que implica que para colas demasiado pesadas, sería imposible calcular incluso medias.

Partiendo del supuesto de que las distribuciones de valores extremos usadas a lo largo del trabajo se ajustan de manera adecuada a la distribuciones marginales empíricas, y que lo mismo sucede con las distribuciones conjuntas, queda el Método de Máxima Verosimilitud para llevar a cabo las estimaciones. Aunque sería posible estimar directamente las cópulas, esto implicaría una gran cantidad de poder computacional y podría llevar a estimaciones degeneradas dada la existencia de valores extremos y condiciones iniciales inadecuadas.

La solución a este problema fue propuesta por Joe y Xu (1996)³² y consiste en realizar las estimaciones en dos etapas. Primero se estiman de forma independiente las distribuciones marginales para después estimar la cópula tomando los valores de las marginales como datos (sustituyéndolos). Se ha comprobado³³ que los estimadores se distribuyen asintóticamente como normales si se cuenta con suficientes observaciones. En el trabajo de László. Gyorgy & Zoltán (2002) se demuestra la convergencia de los parámetros obtenidos por Máxima Verosimilitud en Dos Etapas (2SML) partiendo de procesos ARMA cuya estimación es sensible a valores aberrantes, *outliers*, sin hacer ningún supuesto sobre la distribución de las innovaciones.

El artículo se centra en la tasa de convergencia de las predicciones sobre los errores cuando el proceso es estimado usando 2SML, se demuestra que la predicción de los errores decae en el orden de $O_M(N^{-1})$, lo cual implica convergencia asintótica.

Se deja para el apéndice IV la explicación puntual de la metodología y la demostración de consistencia y convergencia de los parámetros a la distribución normal. Para continuar con la exposición, se supondrá que el lector ha revizado el apéndice o está familiarizado con la técnica.

El primer paso en la estimación de los parámetros a través de 2SML es la construcción de las funciones marginales que servirán de insumo para la segunda estimación (la cópula). Al igual que en cualquier proceso de verosimilitud, se parte de las funciones de densidad

³² Joe, H., and Xu, J. J. (1996). The estimation method of inference functions for margins for multivariate models. Technical Report 166. Vancouver: University of British Columbia, Department of Statistics.

³³ Patton, A. 2006a. Estimation of Multivariate Models for Time Series of Possibly Different Lengths. Journal of Applied Econometrics 21 (2): 147-173.

de probabilidad que mejor ajustan a los datos³⁴.

Es posible que el lector recuerde que en apartados anteriores se determinó el uso de funciones marginales de valores extremos y que estas fueron usadas como insumos de las cópulas. Para obtener las estimaciones de sus parámetros de forma, ubicación y escala es necesario conocer las funciones marginales correspondientes, problema que se simplifica al recordar que estas marginales fueron expresadas usando la parametrización de Von Mises (gamma), por lo que todas las funciones de la familia son iguales a excepción del parámetro de forma cuyo signo identifica a la función. a saber:

$$g_{x,\gamma} = \frac{dG_\gamma}{dx} = \left[e^{-(1+\gamma\left(\frac{x-\mu}{\sigma}\right))^{-1/\gamma}} \right] \left[\left(\left(1 + \gamma \left(\frac{x-\mu}{\sigma} \right) \right)^{-\frac{1+\gamma}{\gamma}} \right) \frac{1}{\sigma} \right]. \quad (1.6.1)$$

Suponiendo que los datos de la muestra son independientes e idénticamente distribuidos, *i.i.d.*, es viable generar una función de verosimilitud de la forma:

$$L = \prod_{i=1}^n \left[e^{-(1+\gamma\left(\frac{x-\mu}{\sigma}\right))^{-1/\gamma}} \right] \left[\frac{1}{\sigma} \right] \left[\left(1 + \gamma \left(\frac{x-\mu}{\sigma} \right) \right)^{-\frac{1+\gamma}{\gamma}} \right],$$

que a su vez implica la siguiente función de Log - verosimilitud:

$$\begin{aligned} \ln(L) = & \sum_{i=1}^n - \left(1 + \gamma \left(\frac{x-\mu}{\sigma} \right) \right)^{\frac{-1}{\gamma}} + n \ln \left(\frac{1}{\sigma} \right) + \\ & \sum_{i=1}^n \left(\frac{1+\gamma}{-\gamma} \right) \ln \left(1 + \gamma \left(\frac{x-\mu}{\sigma} \right) \right). \end{aligned} \quad (1.6.2)$$

Después de haber realizado la estimación de las marginales, se está en posibilidad de sustituir los parámetros de la función de densidad obtenidos en la primera etapa en la función acumulada de probabilidad que será usada como argumento de la cópula bivariada.

Es en este momento cuando, recurriendo a las cópulas halladas en la sección 1.5, se pueden obtener las funciones de verosimilitud. Para ello es necesario obtener las funciones de densidad de probabilidad y posteriormente formar las ecuaciones de verosimilitud y log - verosimilitud.

³⁴ Las preeliminarios de ajuste sobre varias distribuciones serán comentadas en el siguiente capítulo del trabajo.

Es conveniente recordar al lector que en el caso de las funciones de densidad multivariadas continuas, la función acumulada de probabilidad está dada por:

$$F(x_1, \dots, x_n) = \int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_1} f(x_1, \dots, x_n) dx_1 \cdots dx_n, \quad (1.6.3)$$

por lo que de estar construidas sobre dominios continuos y suaves en cada una de sus marginales, la función de densidad está dada por: $f(x_1, \dots, x_n) = \frac{\partial^n F(x_1, \dots, x_n)}{\partial x_1 \cdots \partial x_n}$, que en el caso de la cópula bivariada es:

$$c(u_1, u_2) = \frac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2}. \quad (1.6.4)$$

Recordando que la cópula Weibull mostrada en (1.5.3) está dada por:

$$C(\cdot) = e^{-\left\{ \left[(\lambda_1^{-1} \Omega_1)^{\frac{\gamma_1}{\alpha}} + (\lambda_2^{-1} \Omega_2)^{\frac{\gamma_2}{\alpha}} \right]^\alpha \right\}},$$

donde $\Omega_1 = e^{-(1+\gamma_1(\frac{x_1-\mu_1}{\sigma_1}))^{-\frac{1}{\gamma_1}}}$ y $\Omega_2 = e^{-(1+\gamma_2(\frac{x_2-\mu_2}{\sigma_2}))^{-\frac{1}{\gamma_2}}}$, se tiene una primer derivada parcial, por comodidad se eligió el primer argumento, de:

$$\frac{\partial C(\cdot)}{\partial u_1} = -C(\cdot) \left(\left(\frac{\Omega_1}{\lambda_1} \right)^{\frac{\gamma_1}{\alpha}} + \left(\frac{\Omega_2}{\lambda_2} \right)^{\frac{\gamma_2}{\alpha}} \right)^{\alpha-1} \left(\frac{\gamma_1 \Upsilon_1}{\lambda_1} \right) \left(\frac{\Omega_1}{\lambda_1} \right)^{\frac{\gamma_1}{\alpha}-1}$$

De cumplirse con el teorema de Young, se obtendrá una segunda derivada cruzada, $c(\cdot)$, de la forma:

$$c(\cdot) = \left(\frac{C(\cdot) \gamma_1 \Upsilon_1 \gamma_2 \Upsilon_2}{\lambda_1 \lambda_2} \right) \left(\frac{\Omega_1}{\lambda_1} \right)^{\frac{\gamma_1}{\alpha}-1} \left(\frac{\Omega_2}{\lambda_2} \right)^{\frac{\gamma_2}{\alpha}-1} \left[\binom{\alpha-1}{\alpha} (Z)^{\alpha-2} - C(\cdot) (Z)^{2(\alpha-1)} \right],$$

donde $Z = \left[\left(\frac{\Omega_1}{\lambda_1} \right)^{\frac{\gamma_1}{\alpha}} + \left(\frac{\Omega_2}{\lambda_2} \right)^{\frac{\gamma_2}{\alpha}} \right]$, $\Upsilon_1 = \frac{\Omega_1}{\sigma_1} \left(1 + \gamma_1 \left(\frac{x_1-\mu_1}{\sigma_1} \right) \right)^{-\frac{1}{\gamma_1}-1}$ y

$\Upsilon_2 = \frac{\Omega_2}{\sigma_2} \left(1 + \gamma_2 \left(\frac{x_2-\mu_2}{\sigma_2} \right) \right)^{-\frac{1}{\gamma_2}-1}$. Observe que $\frac{d\Omega_i}{du_i} = \Upsilon_i$. También dada la continuidad de las funciones marginales y de la cópula, se puede garantizar que las derivadas cruzadas son iguales independientemente del orden en que se realicen.

Una vez obtenida la función de densidad de probabilidad, véase Reiss & Thomas (2001), es posible transformarla en una función de verosimilitud, $\Lambda(\cdot) = \prod_{i=1}^n c_i(u_1, u_2; \alpha)$,

de la cual se obtendrán logaritmos para llegar a la función de log - verosimilitud. esto es: $\ln \Lambda (\cdot) = \ln \left(\prod_{i=1}^n c_i (u_1, u_2; \alpha) \right) = \sum_{i=1}^n \ln c_i (u_1, u_2; \alpha)$, lo que en el caso particular de la cópula Weibull implica:

$$\begin{aligned} \ln \Lambda (\cdot) = & \sum_{i=1}^n \ln C (\cdot) + n \ln (\gamma_1) + n \ln (\gamma_2) - n \ln (\lambda_1) - n \ln (\lambda_2) + \sum_{i=1}^n \ln \Upsilon_1 + \\ & \sum_{i=1}^n \ln \Upsilon_2 + \left(\frac{\gamma_2}{\alpha} - 1 \right) \sum_{i=1}^n \ln \left(\frac{\Omega_1}{\lambda_1} \right) + \left(\frac{\gamma_1}{\alpha} - 1 \right) \sum_{i=1}^n \ln \left(\frac{\Omega_2}{\lambda_2} \right) + \\ & + \sum_{i=1}^n \ln \left(\left(\frac{\alpha - 1}{\alpha} \right) Z^{\alpha-1} - C (\cdot) Z^{2(\alpha-1)} \right). \end{aligned} \quad (1.6.6)$$

Para obtener la función de densidad de la la cópula Ali-Mikhail-Haq se considera oportuno mencionar que la ecuación (1.5.4) puede ser reexpresada como:

$$C (u_1, u_2; \alpha) = \frac{\Omega_1 \Omega_2}{1 - \alpha [(1 - \Omega_1) (1 - \Omega_2)]},$$

lo que implica que $\frac{\partial C(\cdot)}{\partial u_1} = \frac{\Omega_2 \Upsilon_1 (1 - \alpha [(1 - \Omega_1) (1 - \Omega_2)]) - \alpha \Omega_1 \Omega_2 (1 - \Omega_2) \Upsilon_1}{(1 - \alpha [(1 - \Omega_1) (1 - \Omega_2)])^2}$ es su primer diferencial parcial. Esta función puede ser simplificada como:

$$\frac{\partial C (\cdot)}{\partial u_1} = \frac{\Omega_2 \Upsilon_1 [1 - \alpha (1 - \Omega_2) (1 - 2\Omega_1)]}{(1 - \alpha [(1 - \Omega_1) (1 - \Omega_2)])^2}.$$

para obtener el valor de la función de densidad de la cópula es necesario obtener la segunda derivada cruzada. Dado lo extenso de esta derivada, sólo se presentará la última simplificación, a saber:

$$\frac{\partial^2 C (\cdot)}{\partial u_1 \partial u_2} = c(\cdot) = \frac{\Upsilon_1 \Upsilon_2 [1 - \alpha [(1 - 2\Omega_1)H^2 - 2\Omega_2(1 - \Omega_1)H]]}{H^4}, \quad (1.6.7)$$

en este caso, por cuestiones de espacio, se sustituyó $H = 1 - \alpha (1 - \Omega_2) (1 - 2\Omega_1)$.

Finalmente, se obtendrá la función de log - verosimilitud siguiendo el procedimiento habitual, lo que conduce a:

$$\begin{aligned} \ln \Lambda (\cdot) = & \sum_{i=1}^n \ln \Upsilon_1 \Upsilon_2 - \sum_{i=1}^n 4 \ln (1 - \alpha (1 - \Omega_1) (1 - \Omega_2)) + \\ & \sum_{i=1}^n \ln \left(1 - \left[\alpha (1 - 2\Omega_1) (H)^2 - 2\Omega_2 (1 - \Omega_1) H \right] \right). \end{aligned} \quad (1.6.8)$$

Para la tercer y última cópula, la FGM, se procederá de la misma manera. En primera instancia se recordará al lector que la cópula acumulada FGM puede ser reescrita, usando la notación usada en las dos primeras cópulas, como

$$C(\cdot) = \Omega_1 \Omega_2 (1 + \alpha (1 - \Omega_1) (1 - \Omega_2)).$$

y que su primer derivada parcial con respecto de la primer marginal está dada por

$$\frac{\partial C(\cdot)}{\partial u_1} = \Upsilon_1 \Omega_2 [1 + \alpha (1 - \Omega_2) (1 - 2\Omega_1)].$$

mientras que la función de densidad, dada por la segunda parcial cruzada, está dada por:

$$\frac{\partial^2 C(\cdot)}{\partial u_1 \partial u_2} = \Upsilon_1 \Upsilon_2 [1 + \alpha (1 - 2\Omega_2) (1 - 2\Omega_1)]. \quad (1.6.9)$$

lo que implica que la función de log - verosimilitud tiene la forma:

$$\ln \Lambda(\cdot) = \sum_{i=1}^n \ln \Upsilon_1 \Upsilon_2 + \sum_{i=1}^n \ln (1 + \alpha (1 - 2\Omega_2) (1 - 2\Omega_1)). \quad (1.6.10)$$

Ambas etapas del 2SML son conducidas, al igual que todo proceso de máxima verosimilitud, buscando maximizar el valor de $\ln \Lambda(\cdot)$, que no es otra cosa más que una transformación monótona de la función de verosimilitud. El método tradicionalmente usado, aunque no el único, para llevar a cabo la maximización es el algoritmo de Newton - Raphson (N-R). Dado que es el algoritmo usado en las estimaciones del modelo, se hará un breve repaso del mismo.

El algoritmo de N-R hace uso del vector gradiente (primeras derivadas) para aproximar el valor de la “pendiente” de la función haciendo uso del conjunto de parámetros θ_t dados, e irá cambiándolos hasta que el cambio en la función de verosimilitud dado un cambio en los parámetros y ponderado por el inverso del Hessiano, \mathbf{H}_t^{-1} , sea menor que un valor de paro, ϵ , predeterminado, esto es:

$$\theta_{t+1} = \theta_t - \mathbf{H}_t^{-1} \mathbf{g}_t. \quad (1.6.11)$$

Como puede observar el lector, es necesario contar con el vector gradiente y la matriz Hessiana. Algunos paquetes como el STATA[®] calculan numéricamente estos vectores y matrices a través del método de la secante, por lo que sólo es necesario introducir la función de log - verosimilitud. De ahí la necesidad de contar con funciones suaves dos veces derivables en la vecindad del punto donde se evalúan los parámetros.

Al igual que otros métodos numéricos, es necesario proveer al algoritmo de valores iniciales confiables que se encuentren relativamente cerca del máximo global con la finalidad de evitar que el mismo se detenga en un máximo local o en un punto silla. En el caso de las funciones marginales, esta posibilidad es remota pues las funciones de densidad son generalmente convexas y el máximo es único. El problema se presenta en funciones multivariadas que pueden tener varios máximos locales e incluso puntos silla que serán alcanzados si no se provee al algoritmo de buenos valores iniciales.

Los puntos iniciales provistos a los algoritmos fueron obtenidos a partir de muestras aleatorias que fueron trasladadas al X-tremes[®] y modeladas con las funciones establecidas previamente en el programa. Incluso se hizo uso de este programa para verificar el buen funcionamiento de los algoritmos presentados en este trabajo al analizar las distribuciones marginales.

Cuando las estimaciones se hacen difíciles debido a la carencia de valores iniciales adecuados, siempre es posible recurrir a otros algoritmos de maximización, entre los más socorridos están los de Davidson - Fletcher - Powell (DFP) y el de Broyden, Fletcher, Goldfarb y Shanno (BFGS). Estos algoritmos aproximan el valor de la matriz Hessiana usando el método de la secante, reduciendo con ello el poder de cómputo consumido y convergiendo con cierta regularidad al valor óptimo aún cuando los valores iniciales se encuentran lejos de los valores óptimos. El costo de esta aproximación es una matriz de varianzas-covarianza aproximada cuyas inexactitudes pueden alterar las pruebas de hipótesis.

También es posible recurrir, cuando el cálculo de la matriz Hessiana es muy complicado, al método de Berndt, Hall, Hall y Hausman (BHHH), el cual omite el cálculo de esta a cambio de un mayor consumo de poder de computadora y la necesidad de buenos valores

iniciales.

Una vez mostradas las funciones de log - verosimilitud de las tres cópulas a usar en la aproximación, la convergencia de los estimadores hallados con estas cópulas y la posibilidad de elaborar pruebas de hipótesis basadas en la certeza de que estos estimadores son variables aleatorias que convergen a una distribución normal (estos dos últimos puntos aclarados en el apéndice AIV). Es viable pasar al análisis empírico de los datos y a la aplicación del modelo.

Antes de pasar a la siguiente sección. resulta útil hacer patente al lector de que siempre es posible generar intervalos de confianza y pruebas de hipótesis para una serie de parámetros cuya distribución no se conoce haciendo uso del método de Bootstrap. Por lo que la verificación de la validez de las pruebas de hipótesis a través de este método es una posibilidad abierta, su validez asintótica está demostrada en el apéndice AIV.

Tecnológico de Monterrey, Campus Ciudad de México
Biblioteca

Parte II

Los Datos y las estimaciones

2.1 Fuente de los datos

Al inicio de la investigación se tenía contemplado realizar todo el análisis propuesto en la sección anterior al mercado de bonos hipotecarios mexicano que aunque incipiente y no tan grande como el Norteamericano, tiene un atractivo especial por su proximidad. Lamentablemente, los datos necesarios no se encuentran al alcance de todo el público, y aunque era posible obtenerlos de forma privada, hacerlo implicaría anular la posibilidad de replicación del análisis por parte de terceros.

Ante esta imposibilidad, el análisis empírico se restringió a los emisores gubernamentales (amparados o no por el gobierno federal) del mercado hipotecario norteamericano cuya misión es aportar liquidez al mismo (no necesariamente otorgar créditos) a saber: *Federal National Mortgage Association (FNMA)*, *Government National Mortgage Association (GNMA)* y *Federal Home Loan Mortgage Corporation (FHLMC)*³⁵.

A pesar de ser tres organismos independientes, dos de ellos manejados como compañías públicas (FNMA y FHLMC), cada uno maneja su propio formato y disposición de los datos referentes a cada crédito otorgado, lo que dificulta su comparación y análisis. Es importante aclarar que estos dos entes compran, empaquetan y bursatilizan los créditos hipotecarios que distintas compañías ofrecen al público, es decir, son meros facilitadores de mercado, a diferencia del INFONAVIT que, bajo algunas circunstancias, otorga créditos directamente.

De cualquier modo, todas estas entidades están obligadas a revisar la documentación presentada por el acreditado, así como el procedimiento de otorgamiento del crédito antes

³⁵ Comúnmente conocidas como: Fannie Mae, Ginnie Mae y Freddie Mac.

de aceptar las hipotecas de manos de las empresas colocadoras (pueden ser bancos o sociedades especializadas). Por lo que, en teoría, las hipotecas bursatilizadas cumplen con los estándares de solvencia amparados por la calificación crediticia bajo la cual son otorgados y los riesgos asociados al instrumento son claramente especificados.

Después de realizar una profunda búsqueda en las bases de datos públicas puestas a disposición por parte de los tres organismos a los posibles inversionistas, se decidió hacer uso de la base de datos mensual publicada por la FNMA sobre bonos emitidos a tasa fija que no habían caído en incumplimiento hasta el momento del análisis.

Ante la inmensa gama de instrumentos respaldados por hipotecas existentes en el mercado³⁶, los bonos a tasa fija que aún no habían caído en cumplimiento y son pasados íntegramente a los inversionistas fueron considerados sujetos idóneos de investigación dada la simplicidad relativa en la construcción del instrumento frente a las alternativas a tasa variable o segregados donde los intereses y el capital son considerados instrumentos independientes.

También fueron seleccionados por la amplitud de la muestra vista como una serie de cortes transversales, aunque vistos como serie de tiempo fuesen un conjunto corto de datos, lo que le confería validez estadística a una muestra que por su propia naturaleza, como se mostrará en AIII, es incapaz de converger a una distribución normal.³⁷

Esta no es una cuestión menor, puesto que la naturaleza del instrumento implica un proceso Markoviano de varios pasos (depende de la tasa cupón a la que fue pactado), lo que a su vez implica que aún con series de tiempo largas, no es sencillo cumplir con el teorema de ergodicidad necesario para la estabilidad de la serie de tiempo, véase Greene (2003), es decir:

$$\begin{aligned} & \lim_{k \rightarrow \infty} |\mathbf{E}[f(x_t, x_{t+1}, \dots, x_{t+a}), g(x_{t+k}, x_{t+k+1}, \dots, x_{t+k+b})]| \\ &= |\mathbf{E}[g(x_{t+k}, x_{t+k+1}, \dots, x_{t+k+b})]| |\mathbf{E}[f(x_t, x_{t+1}, \dots, x_{t+a})]|, \end{aligned}$$

³⁶ Se da una breve explicación de las principales variedades cuando se delimitan los alcances del trabajo.

³⁷ La explicación detallada se desarrolla como introducción al apéndice AIII, por el momento basta con decir que existen varianzas que dominan la muestra e impiden la convergencia a la distribución normal.

no puede cumplirse ya que todas las observaciones dependerán del valor de la tasa cupón pactada al inicio del contrato.

Cuando se decide tratar al problema como una colección de cortes transversales, se elude el problema de la ergodicidad y es posible suponer que cada contratante de crédito se hace de una hipoteca de manera independiente y que esencialmente es evaluado de la misma forma que cualquier otro solicitante. Todo esto puede entenderse como independencia e identidad en la distribución de cada variable aleatoria (crédito).

Una vez justificado el hecho de usar series de cortes transversales como forma de eludir el problema de la falta de ergodicidad y justificación para usar la idea de que cada crédito es una variable independiente e idénticamente distribuida, es necesario aclarar que la elección de instrumentos emitidos por la FNMA³⁸ es meramente por la disposición de los datos que agrupan en un único informe detalles sobre la tasa de cupón promedio pagado por cada crédito, el incumplimiento a treinta, sesenta y noventa días y el pago a capital regular y anticipado hecho en el mes por cada uno de los créditos.

Sobre este conjunto de información, se construyeron diferencias entre la tasa cupón pagada por cada crédito y un promedio móvil a seis meses de la tasa de refinanciamiento promedio de hipotecas similares en plazo al inicio del contrato (30 años o menos). Estos datos fueron obtenidos de la FHLMC³⁹ y son consideradas como las tasas promedio representativas en el sistema dada la “competencia” entre estas entidades, esto es:

$$dr_i = r_{WAC} - \bar{r}. \quad (2.1.1)$$

Una vez construida la base de datos sobre la cual se realizarán las estimaciones, es posible analizar la estructura de los mismos. El conjunto de datos a analizar está dividido en ocho subconjuntos de datos en corte transversal, cada uno conformado inicialmente por ocho mil ciento setenta y seis datos que representan el mismo número de créditos otorgados.

³⁸ La información fue obtenida originalmente de <http://www.fanniemae.com/mbs/data/remic/wholeloan/pseudopoolprintable.jhtml?p=Mortgage-Backed+Securities&s=Monthly+Reporting+Data&t=REMICs&q=Whole+Loan>

³⁹ Específicamente, los datos fueron obtenidos de <http://www.freddiemac.com/pmms/pmms30.htm>

Todos los créditos están inicialmente pactados a 30 años o menos. Al momento de ser otorgados, se les consideró créditos “prime” (los acreditados contaban con buenas referencias crediticias) y todos son instrumentos a tasa fija que fueron bursatilizados de manera íntegra, además fueron excluidos de la muestra todos aquellos créditos que habían caído en incumplimiento total y fueron renegociados, dejando sólo aquellos cuya mora fuera menor o igual a noventa días.

La siguiente tabla muestra los promedios por submuestra (mensuales) para cada una de las variables obtenidas inicialmente de la página de la agencia:

Fecha	Saldo insoluto	Pago a capital	ψ	Pago a interés	λ	r
Dic-07	\$1,632,374.28	\$2,932.37	\$17,890	\$9,168.82	\$78,249	1.390%
Ene-08	\$1,619,860.55	\$2,937.54	\$16,161	\$9,087.34	\$79,846	1.546%
Feb-08	\$1,606,491.64	\$2,912.75	\$15,769	\$9,009.10	\$79,043	1.653%
Mar-08	\$1,590,520.29	\$2,905.65	\$16,932	\$8,926.36	\$76,215	1.720%
Abr-08	\$1,576,763.84	\$2,898.76	\$17,228	\$8,851.37	\$76,053	1.795%
May-08	\$1,561,867.85	\$2,897.16	\$17,355	\$8,768.98	\$74,294	1.822%
Jun-08	\$1,549,221.13	\$2,885.39	\$16,971	\$8,694.72	\$74,452	1.782%
Jul-08	\$1,537,888.69	\$2,874.38	\$13,768	\$8,613.25	\$32,067	1.673%

Cuadro 2.2.1 Promedios mensuales nominales de las variables analizadas.

Como se puede apreciar en la tabla anterior, la muestra toma datos desde Diciembre de 2007 a Julio de 2008 (momento de la investigación). Este era el mejor conjunto de datos disponibles de forma pública de las tres agencias aún después de solicitar mayor información vía correo electrónico con el departamento de atención a los inversionistas.

El primer detalle que puede llamar la atención es que los rubros de pago a capital y prepago no representan exactamente, aunque lo aproximan razonablemente, el cambio en el saldo insoluto que se muestra al mes siguiente. Esto se debe a los movimientos contables que mes a mes realiza la agencia por cargos que se hacen a las cuentas de pérdida reconocida sobre algunos préstamos, pagos que estaban dados como pérdida y fueron recuperados mediante cobranza judicial y extra judicial, además de refinanciamientos parciales solicitados por el deudor que fueron aceptados por los acreedores a fin de no llevar a pérdidas

los montos renegociados. Todas estas aclaraciones puede encontrarse en la página de red de la agencia.

Más allá de estos ajustes contables, el cuadro 2.1.1 muestra una amortización del capital más o menos constante a lo largo del tiempo y pagos a capital e intereses prácticamente estables, lo que es propio de una anualidad. Tal vez los datos más interesantes estén dados por la sobretasa que tendrían que enfrentar los acreditados si decidieran refinanciar sus hipotecas.

El lector podrá observar, en la siguiente tabla, una sobretasa creciente hasta mayo de 2008 que se mantiene alta por el resto de la muestra, lo que señala el deterioro de la condiciones crediticias en el mercado a pesar de que la tasa de interés gubernamental de corto plazo estaba bajando en ese mismo periodo como parte de la respuesta del gobierno federal para paliar el problema hipotecario que a la postre contagiaría a todo el sistema financiero.

Es común observar en la literatura una vinculación teórica de las tasas y montos de prepago e incumplimiento con la tasa libre de riesgo pues en general, para muestras lo suficientemente grandes, siguen la misma dirección. Sin embargo lo que de inicio parece un contra sentido, ver pocos refinanciamientos con una tasa libre de riesgo cercana a cero, es una respuesta perfectamente comprensible de los acreditados pues aunque las tasas de corto plazo estén cercanas a cero, las tasas a 30 años permanecen a niveles similares a los de inicios del periodo estudiado.

Un refinanciamiento implicaría adquirir un riesgo de tasa de interés, en el refinanciamiento periódico, si se decidiera y fuese posible financiar un pasivo hipotecario de largo plazo con créditos de corto plazo, esto sin contar el resto de los riesgos por prepago e incumplimiento. Visto desde el punto de vista de los contadores, es necesario contabilizar los productos del negocio a precio de reposición, la tasa para nuevas hipotecas, pues de lo contrario se adquieren nuevas exposiciones, esto es expresado en la siguiente tabla.

Fecha	Tasa a 1 mes	Tasa a 30 años	Tasa de refinanciamiento
Ene-07	4.938095238	4.851904762	6.22
Feb-07	5.181052632	4.821578947	6.29
Mar-07	5.206363636	4.721818182	6.16
Abr-07	4.988571429	4.866190476	6.18
May-07	4.817727273	4.901363636	6.26
Jun-07	4.521904762	5.203333333	6.66
Jul-07	4.816190476	5.108095238	6.7
Ago-07	4.196956522	4.932173913	6.57
Sep-07	3.784210526	4.793157895	6.38
Oct-07	3.810909091	4.773636364	6.38
Nov-07	3.6825	4.52	6.21
Dic-07	2.858	4.527	6.1
Ene-08	2.738571429	4.33047619	5.76
Feb-08	2.273	4.517	5.92
Mar-08	1.3505	4.393	5.97
Abr-08	1.071818182	4.443181818	5.921
May-08	1.758571429	4.596190476	6.04
Jun-08	1.723809524	4.689047619	6.32
Jul-08	1.606190476	4.57	6.43
Ago-08	1.55	4.59	6.48

Cuadro 2.2.2 Desvinculación de las tasas gubernamentales de corto plazo con la de refinanciamiento hipotecario en los Estados Unidos de Norte América.

Para ilustrar este punto, imagine por un momento que el ficticio Banco Nacional de Comercio decide que tiene el suficiente tamaño como para pedir dinero de corto plazo a una tasa 200 puntos base sobre la tasa libre de riesgo del mismo periodo, con ese dinero, decide prestarle a sus clientes quienes le pagarán, en un crédito hipotecario a treinta años una tasa de 400 puntos base sobre la tasa libre de corto plazo.

Suponga también que el banco decide contratar futuros sobre la tasa de interés ajustando los vencimientos de estos a sus pagos y anulando con ello el riesgo de mercado del negocio, parecería que el banco tiene una ganancia segura de 200 puntos base. Si el lector acepta la existencia de los riesgos de prepago e incumplimiento además de la hipótesis de eficiencia débil de mercado, aprobará la idea de que el precio de ambos riesgos está inclu-

ido en los 200 puntos base de supuesta ganancia del banco como el valor promedio de los incumplimientos y prepagos. Por lo que el banco únicamente está comprando exposiciones a estos riesgos con su estrategia.

Es precisamente la percepción de estos riesgos, la incertidumbre actual, la que ocasiona que el costo de obtener una nueva hipoteca sea positivo y aún con tasas de interés de corto plazo cercanas a cero, los agentes económicos no prepaguen en masa sus créditos. Esta observación es la que valida empíricamente el uso del diferencial para ser vinculada con los riesgos de incumplimiento y prepago.

2.2 Relaciones empíricas, la primera aproximación.

Como se mencionó al inicio de la investigación, los primeros intentos por establecer relaciones causales entre los tres mayores parámetros de riesgo de un bono hipotecario incluyeron análisis de regresión, que aunque probaban una relación estadísticamente significativa entre las variables, no resultaban los mejores predictores para modelos más elaborados, véase la introducción de Kalotay, Yang y Fabozzi(2003), precisamente por la existencia de dependencias de orden superior no captadas por los análisis de regresión y que son el objeto de estudio de este trabajo.

En los siguientes párrafos se reproducirán algunas variantes de los primeros intentos de vinculación, véase Curley and Guttentag (1974), a fin de establecer la relación empírica entre los datos, aunque para su medición se hará uso de los modelos de cópulas. En esta variante del modelo se regresa la tasa de prepago como función del diferencial para refinanciamiento a través de Mínimos Cuadrados Ordinarios (MCO).

En este caso se hizo el análisis sobre los ocho subconjuntos de datos sin respetar ninguna estructura temporal, eludiendo de este modo el problema de la falta de ergodici-

dad, obteniendo los siguientes resultados⁴⁰ :

*regress prepago sobretasa					
Source	SS	df	MS	Number of obs = 64597	
Model	2.0201e+12	1	2.0201e+12	F(1, 64595) = 457.91	
Residual	2.8496e+14	64595	4.4115e+09	Prob > F = 0.0000	
				R-squared = 0.007	
				Adj R-squared = 0.007	
Total	2.8698e+14	64596	4.4427e+09	Root MSE = 66419	
prepago	Coef.	Std. Err.	t	P > t	[95% Conf. Interval]
sobretasa	-338616.2	15823.99	-21.40	0.000	-369631.2 -307601.1
cons	22336.82	377.256	59.21	0.000	21597.4 23076.24

Cuadro 2.2.3 Regresión del prepago como función del diferencial por refinanciamiento.

Estos resultados muestran que existe una relación negativa, como establece la teoría, entre el parámetro de prepago y el diferencial por refinanciamiento. La constante representa un monto de prepago no explicado por el diferencial que se puede enmarcar bajo el título de prepago friccional. Siguiendo con el análisis de la regresión, se puede observar que los parámetros, así como la regresión en su conjunto, son estadísticamente significativos.

Para analizar con mayor detalle este modelo, que aunque ya superado, es capaz de indicar la dirección de la búsqueda, es necesario analizar los residuales de la mismo. En la siguiente tabla se muestran pruebas de Kolmogorov - Smirnoff (K-S) y de curtosis practicadas a los residuales a fin de probar normalidad, que de darse (junto a la no correlación de las observaciones) sugeriría que se trata de un ruido blanco, lo que a su vez indicaría el fin de las pruebas de regresión, pues se habría extraído toda la información posible de los datos.

Después de inspeccionar los residuales y usar su media y varianza como insumos para la prueba K-S, cuya hipótesis nula es la normalidad de los residuales, es necesario rechazar la nula, lo que conlleva la no normalidad de los mismos.

A manera de comprobación, se realizó una prueba de sesgo y kurtosis para normalidad,

⁴⁰ Los datos, así como las regresiones están a disposición de los lectores interesados a vuelta de correo electrónico.

cuya hipótesis nula también es la normalidad de los residuales. De nuevo, esta hipótesis es rechazada.

Finalmente se realizó una prueba de Breusch - Pagan (B-P) para detectar un posible problema de heterocedasticidad, muy probable dada la no normalidad de los residuales, la cual tiene como hipótesis nula homocedasticidad en la regresión. Nuevamente la nula es rechazada. A continuación se muestran estas pruebas.

```
.predict double resid, residuals
summarize resid
  Variable | Obs | Mean |  $\sigma$  | Min | Max
  resid   | 64597 | 1.18e-10 | 66440.15 | -190100 | 1576561
ksmirnov resid = normal(( resid-(0))/(66428.16))
One-sample Kolmogorov-Smirnov test against theoretical distribution
normal(( resid-(0))/(66428.16))
  Smaller group | D | P-value | Corrected |
  resid:        | 0.3677 | 0.000 |
  Cumulative:   | -0.3227 | 0.000 |
  Combined K-S: | 0.3677 | 0.000 | 0.000 |
Note: ties exist in dataset;
there are 32630 unique values out of 64597 observations.
. sktest resid
.Skewness/Kurtosis tests for Normality
  Variable | Obs | Pr(Skewness) | Pr(Kurtosis) | adj  $\chi^2$  | Prob>  $\chi^2$ 
  resid   | 6.5e+04 | 0.000 | 0.000 |
. estat hettest
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
Ho: Constant variance
Variables: fitted values of prepago
chi2(1) = 8243.9
Prob > chi2 = 0.0000
```

Cuadro 2.2.4 Pruebas adicionales a la regresión anterior
para refinanciamiento

Después de haber realizado esta batería de pruebas, es creíble considerar una relación empírica negativa entre el parámetro de prepago y el diferencial en el que se incurriría si se refinanciara la hipoteca. Más aún, es posible decir que esta relación implica una

estructura de dependencia a todas luces más allá de la normal. A continuación se realizará la misma batería de análisis para probar empíricamente la relación entre el incumplimiento y el diferencial para refinanciamiento.

. regress incump sobretasa					
Source	SS	df	MS	Number of obs = 64597	
Model	6.9100e+13	1	6.9100e+13	F(1, 64595) = 227.69	
Residual	1.9604e+16	64595	3.0349e+11	Prob > F = 0.0000	
				R-squared = 0.0035	
				Adj R-squared = 0.0035	
Total	1.9673e+16	64596	3.0455e+11	Root MSE = 5.5e+05	
incump	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]
sobretasa	- 1980446	131248.7	15.09	0.000	-2237693 -1723198
cons	105414.1	3129.068	33.69	0.000	99281.13 111547.1

Cuadro 2.2.5 Regresión del incumplimiento como función del diferencial.

.summarize resid					
Variable	Obs	Mean	σ	Min	Max
resid	64597	3.19e-11	550891	-232755	1.64e+07
.ksmirnov resid = normal((resid-(0))/(550891.9))					
One-sample Kolmogorov-Smirnov test against theoretical distribution normal((resid-(0))/(550891.9))					
Smaller group	D	P-value	Corrected		
resid:	0.3899	0.000			
Cumulative:	-0.4009	0.000			
Combined K-S:	0.4009	0.000	0.000		
Note: ties exist in dataset; there are 9470 unique values out of 64597 observations.					
.sktest resid					
Skewness/Kurtosis tests for Normality					
Variable	Pr(Skewness)	Pr(Kurtosis)	adj χ^2	----- joint ----- Prob>chi2	
resid	0.000	0.000			
.estat hettes					
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity					
Ho: Constant variance					
Variables: fitted values of incump					
chi2(1) = 5009.76					
Prob > chi2 = 0.0000					

Cuadro 2.2.6 Pruebas adicionales a la regresión anterior.

Al igual que en la regresión anterior, se observan parámetros estadísticamente significativos aunque en esta ocasión no son acordes con la teoría (una relación directa). Es posible que esto sea provocados por la recompra y prepago por parte de la agencia de aquellos créditos cuyas moras son mayores a cuatro meses, lo que al final del día se traduce en saldos incumplidos que son sacados de este rubro al ser prepagados por la agencia. Hay que recordar que se sacaron de la muestra los créditos que al inicio del periodo de observación presentaban incumplimientos, no los que incumplieron durante el mismo.

De igual forma es posible observar en las pruebas "F" que la regresión es globalmente significativa, aunque el ajuste también es pobre. Mucha de esta falta de poder de explicación está dada por las dependencias de orden superior que no son captadas por la regresión lineal, por lo que estos resultados son un indicador de la necesidad de profundizar con el análisis.

Después de la regresión lineal, se presentó la batería de pruebas complementarias en las que es necesario rechazar la hipótesis nula de normalidad sugerida por las pruebas K-S y de sesgo - curtosis. También es necesario rechazar la hipótesis nula de homocedasticidad de la prueba de Breusch - Pagan.

De nuevo se ha mostrado de forma empírica que la estructura de dependencia entre estas dos variables está definida por más de dos momentos muestrales, es decir, está dada por una función de densidad bivariada que no es una normal conjunta.

Sólo por completitud en los análisis realizados, se buscará probar empíricamente la relación entre el prepago y el incumplimiento. Para ello se hará uso de una regresión lineal por Mínimos Cuadrados Ordinarios (al igual que el resto). En esta regresión (mostrada en la tabla 2.2.7), se puede encontrar la existencia de una relación entre los prepagos y el incumplimiento dado que tanto los coeficientes como la regresión en su conjunto son estadísticamente significativos.

En este caso, la teoría coloca estos dos eventos como competitivos entre sí, pues en un ambiente de bonanza con tasas de refinanciamiento bajas, es poco probable que se den incumplimientos, aunque en una recesión como la actual, es común encontrar pagos

anticipados asociados a incumplimientos.

. regress incump prepago					
Source	SS	df	MS	Number of obs = 64597	
Model	1.3864e+15	1	1.3864e+15	F(1, 64595) = 4897.17	
Residual	1.8286e+16	64595	2.8309e+11	Prob > F = 0.0000	
				R-squared = 0.0705	
				Adj R-squared = 0.0705	
Total	1.9673e+16	64596	3.0455e+11	Root MSE = 5.3e+05	
incump	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]
prepago	2.197929	.0314081	-69.98	0.000	2.136369 2.259489
cons	35063.73	2156.734	16.26	0.000	30836.53 39290.93

Cuadro 2.2.7 Regresión del incumplimiento como función del prepago.

.summarize resid					
Variable	Obs	Mean	σ	Min	Max
resid	64597	-2.24e-10	532061	-2886967	1.58e+07
.ksmirnov resid = normal((resid-(0))/(532061.6))					
One-sample Kolmogorov-Smirnov test against theoretical distribution					
normal((resid-(0))/(532061.6))					
Smaller group	D	P-value	Corrected		
resid:	0.4153	0.000			
Cumulative:	-0.3604	0.000			
Combined K-S:	0.4153	0.000	0.000		
Note: ties exist in dataset;					
there are 25582 unique values out of 64597 observations.					
.sktest resid					
Skewness/Kurtosis tests for Normality					
Variable	Pr(Skewness)	Pr(Kurtosis)	adj χ^2	P> χ^2 — joint —	
resid	0.000	0.000	.		
.estat hettes					
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity					
Ho: Constant variance					
Variables: fitted values of incump					
chi2(1) = 468951.76					
Prob > chi2 = 0.0000					

Cuadro 2.2.8 Pruebas adicionales a la regresión del incumplimiento como función del prepago.

Es importante aclarar que todos estos análisis solamente establecen la existencia de una relación estadísticamente significativa, no una causalidad entre ellas. Para establecerla es necesario realizar un análisis de causalidad de Granger, para su estudio se remite al lector al libro de Econometría de Hayashi (2000)..

Nuevamente se practicó la batería de pruebas a los residuales cuya hipótesis nula es la normalidad de los mismos y al igual que en oportunidades anteriores, los estadísticos de las pruebas de K - S y sesgo - kurtosis caen en la región de rechazo de la hipótesis nula.

De nuevo es posible observar la presencia de heterocedasticidad en los residuales, probablemente producto de la no normalidad de los mismos. En este punto, es lógico preguntarse si la causa de la no normalidad de todos los residuales es la no normalidad de cada uno de los factores analizados,dejando con ello abierta la puerta a distribuciones bivariadas distintas a la normal. Se proponen como solución a estas preguntas las pruebas de Kolmogorov - Smirnov y sesgo - kurtosis. Las siguientes tablas muestran los resultados de las mismas aplicadas a los factores de riesgo (regresores) analizados en el trabajo.

.summarize sobretasa					
Variable	Obs	Mean	Std. Dev.	Min	Max
sobretasa	64597	.0171944	.0165147	-.0643	.1174
.ksmirnov prepag = normal((sobretasa-(.0171944))/(.0165147))					
One-sample Kolmogorov-Smirnov test against theoretical distribution					
normal((sobretasa-(.0171944))/(.0165147))					
Smaller group	D	P-value	Corrected		
resid:	0.0954	0.000			
Cumulative:	-0.0681	0.000			
Combined K-S:	0.0954	0.000	0.000		
Note: ties exist in dataset; there are 3677 unique values out of 64597 observations.					
.sktest prepag					
Skewness/Kurtosis tests for Normality					
Variable	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2)	—— joint —— Prob>chi2	
prepag	0.000	0.000			

Cuadro 2.2.9 Pruebas de normalidad practicadas a la sobretasa.

```
.summarize prepago
```

Variable	Obs	Mean	σ	Min	Max
prepago	64597	16514.5	66653	-172576	1596650

```
.ksmirnov prepago = normal(( prepago-(16514.5))/(66653.26))
One-sample Kolmogorov-Smirnov test against theoretical distribution
normal(( prepago-(16514.5))/(66653.26))
```

Smaller group	D	P-value	Corrected
resid:	0.4616	0.000	
Cumulative:	-0.3605	0.000	
Combined K-S:	0.4616	0.000	0.000

Note: ties exist in dataset;
there are 22160 unique values out of 64597 observations.

```
.sktest prepago
Skewness/Kurtosis tests for Normality
```

Variable	Pr(Skewness)	Pr(Kurtosis)	adj χ^2	P > χ^2	joint
prepago	0.000	0.000			

```
.summarize incumplimiento
```

Variable	Obs	Mean	σ	Min	Max
incumplimiento	64597	71361.44	551861.9	0	1.65e+07

```
.ksmirnov incumplimiento = normal(( incumplimiento-(71361.44))/(551861.9))
One-sample Kolmogorov-Smirnov test against theoretical distribution
(( incumplimiento-(71361.44))/(551861.9))
```

Smaller group	D	P-value	Corrected
resid:	0.4606	0.000	
Cumulative:	-0.4486	0.000	
Combined K-S:	0.4606	0.000	0.000

Note: ties exist in dataset;
there are 5960 unique values out of 64597 observations.

```
.sktest prepago
Skewness/Kurtosis tests for Normality
```

Variable	Pr(Skewness)	Pr(Kurtosis)	adj χ^2	P > χ^2	joint
prepago	0.000	0.000	.		

Cuadro 2.2.10 Pruebas de normalidad practicadas a incumplimiento y prepago.

En esta serie de pruebas se rechaza la hipótesis nula de normalidad marginal de los datos y con ello la posibilidad de una distribución normal bivariada como función de distribución conjunta, lo que también excluye a la cópula normal.

2.3 Análisis de las distribuciones marginales con funciones de valores extremos, ajustes preliminares.

Una vez probada la no normalidad de las observaciones, queda la necesidad de encontrar una función de distribución de probabilidad que ajuste con suficiente fidelidad a los datos observados. La primera peculiaridad que puede observar el lector en el conjunto de datos analizados es la abrumadora cantidad de observaciones cuyo valor es cero, además de la existencia, no del todo aislada, de valores de gran magnitud que pueden ser calificados como “aberrantes”, cuya varianza domina la muestra.

Al escuchar esta descripción, es posible que se recuerde la definición de observaciones producidas por un proceso de valores extremos. Para constatar si los datos observados son aceptablemente ajustados por una de estas distribuciones, se recurrirá a una primera aproximación haciendo uso de la estadística descriptiva, poniendo énfasis en el exeso de curtosis de las submuestras y del conjunto de los datos.

	Prepago muestral							
	Dic	Ene	Feb	Mar	Abr	May	Jun	Jul
\bar{x}	18858.1	20084.1	17978.5	17223.8	14701	17157	18354	18424
$S_{\bar{x}}$	2256.99	2452.64	2197.49	2208.95	1980	2311	2248	2395
Me	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Md	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
σ	71372.2	77559.4	69490.8	69853.0	62617	73078	71085	75727
k	43.83	75.04	38.40	56.61	75.84	55.18	60.06	81
γ_1	5.84	7.36	5.71	6.65	7.58	6.63	6.70	7.78
n	1000	1000	1000	1000	1000	1000	1000	1000

Cuadro 2.3.1 Estadística descriptiva de una muestra del prepago.

En la tabla anterior, se observa que tanto la estadística descriptiva de la muestra obtenida de la base de datos del prepago, así como la del conjunto de la información presenta un exeso de kurtosis que invita a pensar más que en colas anchas, en colas pesadas (la diferencia y el tratamiento están dados en el apéndice AIII).

Prepago global								
	Dic	Ene	Feb	Mar	Abr	May	Jun	Jul
\bar{x}	17890.8	16161.8	15769.1	16932.2	17228	17356	16972	13769
$S_{\bar{x}}$	802.72	720.50	726.38	767.71	772.2	740.4	758.3	628.4
Me	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mo	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
σ	72582.8	65017.3	65438.8	69076.6	69339	66373	67821	56104
k	67.70	73.35	61.67	86.40	77.27	44.85	62.21	64.85
γ_1	7.03	7.11	6.89	7.56	7.25	5.89	6.76	6.65
n	8176	8143	8116	8096	8063	8035	7998	7970

Cuadro 2.3.2 Estadística descriptiva de la base de datos del prepago.

También es posible observar un aumento dramático de la kurtosis entre la muestra y el conjunto de los datos, lo que sugiere la existencia de observaciones “aberrantes” que la dominan. Esto constituye otra evidencia de la necesidad del modelado usado EVT.

Es posible que, después de observar la estadística descriptiva, se cuestione el motivo de una kurtosis tan alta y poco robusta. Un indicio sobre la causa puede encontrarse en la mediana y moda iguales a cero, lo cual indica de entrada que las observaciones más comunes (moda) son ceros, y que por lo menos hasta el quinto decil (50%) integran toda la muestra, por lo que algunas observaciones distintas de cero son tan importantes que la dominan, y por tanto, son una parte integral de ella que no puede ser desechada u omitida.

Para el incumplimiento se tiene que:

Incumplimiento global								
	Dic	Ene	Feb	Mar	Abr	May	Jun	Jul
\bar{x}	78250	79846	79043	76216	76053	74295	74452	32067
$S_{\bar{x}}$	6572	6677	6626	6393	6124	6007	5951	4371
Me	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mo	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
σ^*	59425	60255	59695	57518	54988	53849	53220	39017
k	316.11	315.78	319.77	330.33	312.84	309.32	299.25	760.79
γ_1	15.87	15.84	15.90	16.03	15.57	15.51	15.23	25.06
n	8176	8143	8116	8096	8063	8035	7998	7970

σ^* está en decenas.

Cuadro 2.3.3 Estadística descriptiva de la base de datos de incumplimiento.

	Incumplimiento muestral							
	Dic	Ene	Feb	Mar	Abr	May	Jun	Jul
\bar{x}	56488	44542	59058	32971	44038	32791	36678	46369
$S_{\bar{x}}$	20617	18032	23526	14205	14443	9750	11033	17809
Me	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mo	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
σ^*	65195	57021	74395	44921	45671	30831	34888	56315
k	460.87	695.43	420.86	859.82	293.30	360.97	609.97	515.54
γ_1	20.26	24.94	20.00	28.41	16.18	17.61	22.56	21.44
n	1000	1000	1000	1000	1000	1000	1000	1000

σ^* está en decenas.

Cuadro 2.3.4 Estadística descriptiva de una muestra de los incumplimientos.

En las dos tablas anteriores se puede apreciar el mismo fenómeno de falta de robustez en la kurtosis y concentración de la muestra, por lo menos hasta el quinto decil, en el cero. Por lo que, al igual que en el prepago, se justifica el modelado de los incumplimientos con marginales de EVT que en teoría reproducen de manera fiel los datos.

A diferencia del prepago, los incumplimientos presentan un cambio en las medias muestral y global en Julio de 2008 que responde al reconocimiento como pérdida sujeta a castigo en capital para los emisores de algunos de los créditos con mayores moras.

Es importante recordar que los prospectos de inversión de la FNMA, véase el ejemplar citado en la bibliografía, aclaran que las moras mayores a cuatro meses son transmitidas a los tenedores, según el lote (*tranch*) al que pertenezcan, en forma de prepagos cuando éstos son recomprados por la FNMA. Como se puede ver, la recompra por incumplimiento es la causa de que el tamaño de la población decaiga a lo largo del tiempo.

Sólo resta presentar los cuadros de resumen mensual para la sobretasa por refinanciamiento. En ellas, el lector apreciará que se trata del factor de riesgo menos leptocúrtico de los tres analizados en el trabajo sin que por ello se pueda suponer que se trata de variables aleatorias distribuidas como normales puesto que las pruebas de normalidad que les fueron practicadas, y mostradas en páginas anteriores, rechazan esa hipótesis nula.

También se puede observar en estas tablas una mayor robustez de las kurtosis muestral y global a pesar del desplome de la tasa libre de riesgo crédito en el periodo de la

muestra. Esto responde a la relativa estabilidad, mostrada en el cuadro 2.2.2, de la tasa libre (de riesgo crédito) a 30 años. Ante esta evidencia empírica, la decisión de hacer uso de marginales provenientes de la EVT responde también a la necesidad simplificadora de mantener una homogeneidad en las marginales usadas aún cuando la propia teoría de cópulas permite casi cualquier clase de marginales a condición de ser continuas. A continuación se muestran estas tablas.

Sobretasa para refinanciamiento global								
	Dic	Ene	Feb	Mar	Abr	May	Jun	Jul
\bar{x}	0.0139	0.0155	0.0165	0.0172	0.0179	0.0182	0.0178	0.0167
$S_{\bar{x}}$	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
Me	0.0111	0.0127	0.0138	0.0144	0.0152	0.0155	0.0151	0.0140
Mo	0.0061	0.0077	0.0088	0.0094	0.0102	0.0105	0.0101	0.0090
σ	0.0164	0.0164	0.0163	0.0164	0.0164	0.0164	0.0164	0.0164
k	5.0219	5.0155	4.9543	5.0014	5.0378	5.0858	5.1492	5.0991
γ_1	1.4482	1.4688	1.4580	1.4438	1.4299	1.4165	1.3849	1.4376
n	8176	8143	8116	8096	8063	8035	7998	7970

Cuadro 2.3.5 Estadística descriptiva global de la sobretasa para refinanciamiento.

Sobretasa muestral para refinanciamiento								
	Dic	Ene	Feb	Mar	Abr	May	Jun	Jul
\bar{x}	0.0149	0.0166	0.01747	0.01876	0.0183	0.0198	0.0189	0.0184
$S_{\bar{x}}$	0.0005	0.0005	0.00050	0.00051	0.0005	0.0006	0.0005	0.0005
Me	0.0124	0.0139	0.0146	0.01568	0.0152	0.016	0.0151	0.0153
Mo	0.0011	0.0077	0.0088	0.00943	0.0252	0.011	0.0101	0.0140
σ	0.0161	0.0166	0.01579	0.01622	0.0157	0.0179	0.0172	0.0161
k	5.1401	5.3985	5.50150	5.37526	6.2397	5.0379	4.263	4.2964
γ_1	1.1027	1.3988	1.54783	1.50258	1.3927	1.5871	1.1616	1.4042
n	1000	1000	1000	1000	1000	1000	1000	1000

Cuadro 2.3.6 Estadística descriptiva muestral de la sobretasa para refinanciamiento.

Existe la posibilidad de que, después de observar las tablas y sus implicaciones, el lector cuestione la validez del muestro practicado a los datos y extrapole su duda a los intervalos de confianza surgidos del método de Bootstrap propuesto como alternativa para la realización de intervalos de confianza.

Para allanar estas dudas, se pone a disposición de los lectores en un quinto apéndice, (AV), el algoritmo de muestro con reemplazo creado en VBA para Excel© usado para crear las muestras. Este algoritmo se basa en la generación de números pseudo aleatorios, por lo que su validez depende del buen desempeño del generador usado. El principio básico es la generación de un número pseudoaleatorio que de ser menor a la proporción deseada de datos muestrales respecto a la muestra, provoca que el algoritmo copie el dato en una matriz de muestreo la cual es luego presentada en una hoja de cálculo distinta. La condicional que funciona como el corazón del algoritmo puede se resumida como:

$$f(x) = \begin{cases} 1 & \text{si } x < (n/N) \text{ (1.2)} \\ 0 & \text{si } x > (n/N) \text{ (1.2)}, \end{cases} \quad (2.3.1)$$

donde n es el tamaño deseado para la muestra, N es el tamaño de la población, x es el número pseudo aleatorio generado por la computadora y la constante, 1.2, está dada de forma arbitraria para dar un margen (subir el umbral de aceptación) suficiente al algoritmo para llenar la matriz en caso de no generar la cantidad suficiente de eventos que cumplan la condición con la proporción justa, n/N .

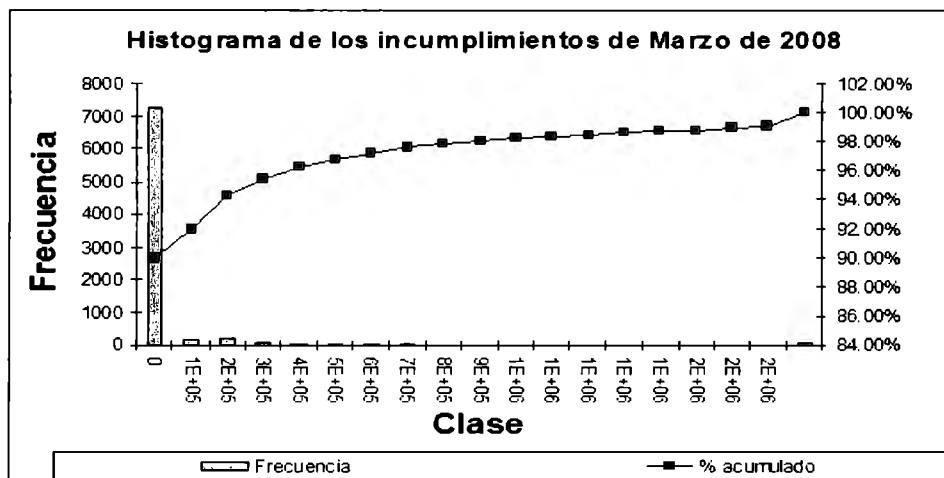
A fin de hacer más clara la necesidad de modelar las distribuciones marginales de los tres factores de riesgo analizados mediante funciones de EVT, se presentarán tres⁴¹ gráficos representativos (uno por factor de riesgo) escogidos arbitrariamente que muestran la distribución empírica acumulada (línea punteada) y la función empírica de cuantiles (histograma) asociados a los tres conjuntos de datos escogidos.

En los gráficos, el lector podrá observar que para los incumplimientos y prepagos, más del cincuenta por ciento de las observaciones son iguales a cero, tal y como la mediana y moda de la estadística descriptiva habían sugerido. Lo que convierte a estos tres factores claros candidatos a modelado usando EVT.

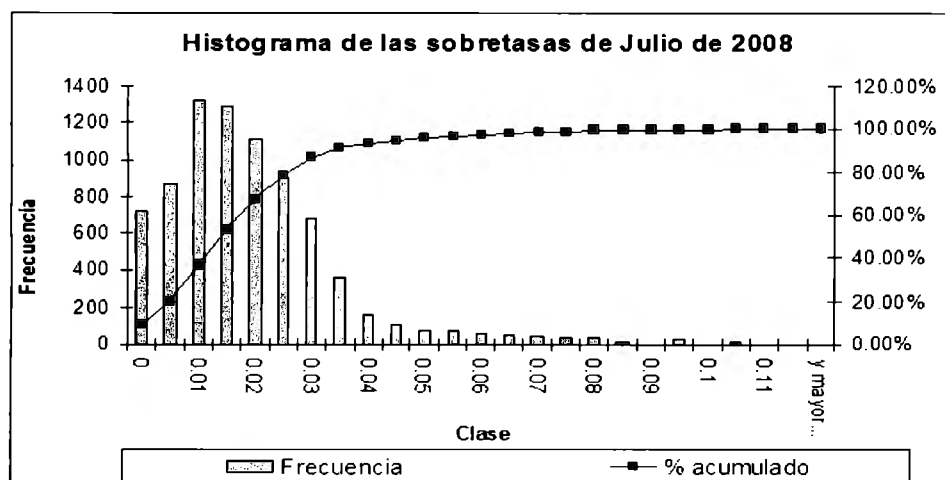
Del mismo modo se podrá observar que los histogramas, que son en última instancia el inverso de la función de probabilidad o la derivada de la función de cuantiles, muestran

⁴¹ Por simplicidad en la exposición y cuestiones de espacio, sólo se presentan tres gráficos. El lector puede constatar la similitud de las distribuciones acumulada y de cuantiles al analizar la estadística descriptiva antes mostrada. De ser necesario, el resto de los histogramas están a disposición del lector interesado a vuelta de correo electrónico.

colas pesadas y sesgos de tal magnitud que difícilmente podrán ser explicados por funciones que únicamente arrojen excesos de curtosis como la distribución *t de student*. Asociado a este hecho, se encontrarán funciones acumuladas de probabilidad que acumulan la mayor parte de su masa en las primeras particiones del dominio (no se usa la palabra decil o percentil puesto que el dominio no está dividido en diez o cien intervalos).



del factor de riesgo por sobretasa para refinanciamiento no presenta colas tan pesadas, incluso podrían se interpretadas como meras colas anchas, ni sesgos tan marcados como los factores de riesgo precedentes



Gráfica 2.3.3 Histograma de las sobretasas presentes en julio de 2008.

Después de analizar el histograma, el lector podría pensar en una distribución gamma o una Weibull como las idóneas para modelar estos datos. Para la realización de este trabajo se escogió la distribución Weibull por pertenecer a la familia de las distribuciones de la EVT cuya parametrización gamma la lleva a tener una distribución común con el resto de ellas.

Una vez mostrada la pertinencia del modelado escogido para los datos, se mostrarán las estimaciones formales realizadas sobre toda la información disponible (los ocho subconjuntos de corte transversal). Para realizar las estimaciones se hizo uso del programa STATA© versión SE 10.1 que es un programa econométrico de uso común con amplias capacidades de programación y un sólido algoritmo de optimización por máxima verosimilitud.

La elección del Stata sobre programas especialmente diseñados para el modelado de valores extremos como el Xtremes 3.01© responde a las limitaciones en la cantidad de datos aceptados por este último programa (mil en su versión académica) y la necesidad

de mostrar los algoritmos usados en la elaboración de las estimaciones (evitar una posible caja negra). Sin embargo el lector puede usar cualquier otro programa econométrico o algoritmo diseñado *ex professo* a condición de verificar su buen funcionamiento contra otro ya probado.

Sobre esta línea, los resultados del algoritmo usado en la construcción de las marginales fue contrastado contra los resultados proporcionados por el Xtremes 3.01 en un conjunto de datos distribuido como Frechét con media y varianza no estandarizada creado y estimado por el propio Xtremes©. Posteriormente, estos datos fueron trasladados al STATA© y sometidos al análisis del algoritmo de la primera etapa de la Máxima verosimilitud suponiendo una marginal de valores extremos, obteniendo los mismos resultados que el Xtremes©, lo que valida el algoritmo usado en la investigación.

A continuación se presentan las estimaciones de los parámetros de forma, γ , ubicación, μ , y escala, σ , para cada uno de los tres factores de riesgo agrupados en los ocho subconjuntos de corte transversal que comprenden las observaciones mensuales de Diciembre de 2007 a Julio de 2008. De nuevo, el algoritmo de verosimilitud, así como los archivos (.do) usados serán relegados al apéndice AV, mientras que los conjuntos de datos (archivos .dat) están a disposición del lector interesado a vuelta de correo electrónico.

2.4 Estimación mensual de las marginales.

Antes de iniciar, es necesario aclarar al lector que se han reemplazado los rótulos de los coeficientes de la regresión por su símbolo a fin de dar claridad a la exposición.

2.4.1 Estimaciones marginales para Diciembre de 2007.

					Number of obs= 8176	
					Wald chi2(0) =.	
					Prob > chi2 = .	
Log likelihood = -96631.365					[95% Conf. Interval]	
	Coef.	Std. Err.	z	P> z		
γ	.0251871	.0016615	15.16	0.000	.0219307	.0284436
μ	879.9421	337.4914	2.61	0.009	218.4711	1541.413
σ	29529.03	200.4065	147.35	0.000	29136.24	29921.82

Cuadro 2.4.1.1 Estimación de la marginal para el prepego de Diciembre de 2007.

					Number of obs= 8176
					Wald chi2(0) = .
					Prob > chi2 = .
Log likelihood = -64611.137					[95% Conf. Interval]
	Coef.	Std. Err.	z	P> z	
γ	.0817506	.0015484	52.80	0.000	.0787158 .0847854
μ	-45.08434	6.920621	-6.51	0.000	-58.64851 -31.52018
σ	605.3342	3.910721	154.79	0.000	597.6694 612.9991

Cuadro 2.4.1.2 Estimación de la marginal para el incumplimiento de Diciembre de 2007.

					Number of obs= 8176
					Wald chi2(0) = .
					Prob > chi2 = .
Log likelihood = 22378.868					[95% Conf. Interval]
	Coef.	Std. Err.	z	P> z	
γ	-.1110657	.002906	-38.22	0.000	-.1167614 -.10537
μ	.0075303	.0001735	43.40	0.000	.0071903 .0078704
σ	.0148452	.0001071	138.57	0.000	.0146352 .0150552

Cuadro 2.4.1.3 Estimación de la marginal para la sobretasa de Diciembre de 2007.

2.4.2 Estimaciones marginales para Enero de 2008.

					Number of obs= 8143
					Wald chi2(0) = .
					Prob > chi2 = .
Log likelihood = -92319.822					[95% Conf. Interval]
	Coef.	Std. Err.	z	P> z	
γ	.1173872	.0025271	46.45	0.000	.1124342 .1223403
μ	364.9465	182.6253	2.00	0.046	7.007571 722.8855
σ	15971.96	127.9142	124.86	0.000	15721.26 16222.67

Cuadro 2.4.2.1 Estimación de la marginal para el prepego de Enero de 2008.

					Number of obs= 8143
					Wald chi2(0) = .
					Prob > chi2 = .
Log likelihood = -64363.09					[95% Conf. Interval]
	Coef.	Std. Err.	z	P> z	
γ	.0820187	.0015537	52.79	0.000	.0789735 .0850638
μ	-45.17981	6.942984	-6.51	0.000	-58.7878 -31.57181
σ	606.0624	3.925156	154.40	0.000	598.3692 613.7555

Cuadro 2.4.2.2 Estimación de la marginal para el incumplimiento de Enero de 2008.

					Number of obs= 8143	
					Wald chi2(0) =.	
Log likelihood = 22329.793					Prob > chi2 = .	
	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
γ	-.1081657	.0029505	-36.66	0.000	-.1139486	-.1023828
μ	.0090888	.0001725	52.70	0.000	.0087508	.0094268
σ	.0147259	.0001068	137.85	0.000	.0145165	.0149353

Cuadro 2.4.2.3 Estimación de la marginal para la sobretasa de Enero de 2008.

2.4.3 Estimaciones marginales para Febrero de 2008.

					Number of obs= 8116	
					Wald chi2(0) =.	
Log likelihood = -91666.02					Prob > chi2 = .	
	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
γ	.1198164	.0025292	47.37	0.000	.1148592	.1247736
μ	233.1176	175.3105	1.33	0.184	-110.4847	576.7199
σ	15304.97	122.5796	124.86	0.000	15064.72	15545.22

Cuadro 2.4.3.1 Estimación de la marginal para el prepego de Febrero de 2008.

					Number of obs= 8116	
					Wald chi2(0) =.	
Log likelihood = -64247.914					Prob > chi2 = .	
	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
γ	.0841452	.0015748	53.43	0.000	.0810586	.0872317
μ	-44.96001	7.005468	-6.42	0.000	-58.69048	-31.22955
σ	610.5993	3.981387	153.36	0.000	602.796	618.4027

Cuadro 2.4.3.2 Estimación de la marginal para el incumplimiento de Febrero de 2008.

					Number of obs= 8116	
					Wald chi2(0) =.	
Log likelihood = 22268.363					Prob > chi2 = .	
	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
γ	-.1091369	.0029667	-36.79	0.000	-.1149515	-.1033224
μ	.0101729	.0001726	58.96	0.000	.0098347	.0105111
σ	.0147086	.0001069	137.62	0.000	.0144992	.0149181

Cuadro 2.4.3.3 Estimación de la marginal para la sobretasa de Febrero de 2008.

2.4.4 Estimaciones marginales para Marzo de 2008.

					Number of obs= 8096
					Wald chi2(0) = .
Log likelihood = -92465.838					Prob > chi2 = .
	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
γ	.0996033	.002305	43.21	0.000	.0950856 .1041211
μ	623.7535	202.1069	3.09	0.002	227.6312 1019.876
σ	17651.73	138.7811	127.19	0.000	17379.72 17923.73

Cuadro 2.4.4.1 Estimación de la marginal para el prepego de Marzo de 2008.

					Number of obs= 8096
					Wald chi2(0) = .
Log likelihood = -64295.257					Prob > chi2 = .
	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
γ	.0884827	.001617	54.72	0.000	.0853134 .091652
μ	-44.4496	7.116698	-6.25	0.000	-58.39807 -30.50113
σ	619.7717	4.091396	151.48	0.000	611.7527 627.7907

Cuadro 2.4.4.2 Estimación de la marginal para el incumplimiento de Marzo de 2008.

					Number of obs= 8096
					Wald chi2(0) = .
Log likelihood = 22169.299					Prob > chi2 = .
	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
γ	-.1119027	.0029268	-38.23	0.000	-.1176392 -.1061662
μ	.0108402	.0001742	62.21	0.000	.0104987 .0111817
σ	.0148358	.0001075	137.95	0.000	.014625 .0150466

Cuadro 2.4.4.3 Estimación de la marginal para la sobretasa de Marzo de 2008.

2.4.5 Estimaciones marginales para Abril de 2008.

					Number of obs= 8063
					Wald chi2(0) = .
Log likelihood = -92918.13					Prob > chi2 = .
	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
γ	.0842275	.0022162	38.01	0.000	.0798838 .0885712
μ	530.0229	229.2943	2.31	0.021	80.61446 979.4314
σ	19968.94	154.0548	129.62	0.000	19666.99 20270.88

Cuadro 2.4.5.1 Estimación de la marginal para el prepego de Abril de 2008.

					Number of obs= 8063
					Wald chi2(0) = .
					Prob > chi2 = .
Log likelihood = -64338.043					[95% Conf. Interval]
	Coef.	Std. Err.	z	P> z	
γ	.0939276	.0017019	55.19	0.000	.0905919 .0972634
μ	-49.54851	7.354506	-6.74	0.000	-63.96308 -35.13394
σ	638.5641	4.280836	149.17	0.000	630.1738 646.954

Cuadro 2.4.5.2 Estimación de la marginal para el incumplimiento de Abril de 2008.

					Number of obs= 8063
					Wald chi2(0) = .
					Prob > chi2 = .
Log likelihood = 22034.729					[95% Conf. Interval]
	Coef.	Std. Err.	z	P> z	
γ	-.1144707	.0028959	-39.53	0.000	-.1201465 -.1087949
μ	.01159	.0001761	65.83	0.000	.0112449 .0119351
σ	.01496	.0001083	138.11	0.000	.0147477 .0151723

Cuadro 2.4.5.3 Estimación de la marginal para la sobretasa de Abril de 2008.

2.4.6 Estimaciones marginales para Mayo de 2008.

					Number of obs= 8035
					Wald chi2(0) = .
					Prob > chi2 = .
Log likelihood = -94768.059					[95% Conf. Interval]
	Coef.	Std. Err.	z	P> z	
γ	.0269054	.0018035	14.92	0.000	.0233705 .0304403
μ	637.6985	327.8443	1.95	0.052	-4.86451 1280.262
σ	28426.83	199.209	142.70	0.000	28036.38 28817.27

Cuadro 2.4.6.1 Estimación de la marginal para el prepego de Mayo de 2008.

					Number of obs= 8035
					Wald chi2(0) = .
					Prob > chi2 = .
Log likelihood = -64167.813					[95% Conf. Interval]
	Coef.	Std. Err.	z	P> z	
γ	.0957523	.0016949	56.49	0.000	.0924303 .0990743
μ	-42.89972	7.323995	-5.86	0.000	-57.25449 -28.54496
σ	635.8592	4.2962	148.01	0.000	627.4388 644.2796

Cuadro 2.4.6.2 Estimación de la marginal para el incumplimiento de Mayo de 2008.

Log likelihood = 21921.514						Number of obs= 8035	
						Wald chi2(0) = .	
						Prob > chi2 = .	
	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]		
γ	-.1169182	.0028612	-40.86	0.000	-.1225261	-.1113103	
μ	.0118644	.0001777	66.78	0.000	.0115161	.0122126	
σ	.0150712	.000109	138.33	0.000	.0148577	.0152848	

Cuadro 2.4.6.3 Estimación de la marginal para la sobretasa de Mayo de 2008.

2.4.7 Estimaciones marginales para Junio de 2008.

Log likelihood = -90653.687						Number of obs= 7998	
						Wald chi2(0) = .	
						Prob > chi2 = .	
	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]		
γ	.1354991	.0027093	50.01	0.000	.130189	.1408092	
μ	239.0335	180.4671	1.32	0.185	-114.6756	592.7425	
σ	15634.63	129.2744	120.94	0.000	15381.26	15888	

Cuadro 2.4.7.1 Estimación de la marginal para el prepego de Junio de 2008.

Log likelihood = -63938.806						Number of obs= 7998	
						Wald chi2(0) = .	
						Prob > chi2 = .	
	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]		
γ	.0970564	.0017122	56.68	0.000	.0937005	.1004123	
μ	-42.42099	7.376699	-5.75	0.000	-56.87905	-27.96292	
σ	639.0283	4.342927	147.14	0.000	630.5163	647.5403	

Cuadro 2.4.7.2 Estimación de la marginal para el incumplimiento de Junio de 2008.

Log likelihood = 21747.003						Number of obs= 7998	
						Wald chi2(0) = .	
						Prob > chi2 = .	
	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]		
γ	-.1212998	.0028016	-43.30	0.000	-.1267909	-.1158088	
μ	.0114538	.0001806	63.42	0.000	.0110999	.0118078	
σ	.0152857	.0001102	138.74	0.000	.0150698	.0155017	

Cuadro 2.4.7.3 Estimación de la marginal para la sobretasa de Junio de 2008.

2.4.8 Estimaciones marginales para Julio de 2008.

					Number of obs= 7970
					Wald chi2(0) = .
					Prob > chi2 = .
Log likelihood = -89455.987					[95% Conf. Interval]
	Coef.	Std. Err.	z	P> z	
γ	.116462	.0025479	45.71	0.000	.1114682 .1214557
μ	-151.011	165.7792	-0.91	0.362	-475.9324 173.9103
σ	14333.69	115.4283	124.18	0.000	14107.45 14559.92

Cuadro 2.4.8.1 Estimación de la marginal para el prepego de Julio de 2008.

					Number of obs= 7970
					Wald chi2(0) = .
					Prob > chi2 = .
Log likelihood = -63847.65					[95% Conf. Interval]
	Coef.	Std. Err.	z	P> z	
γ	.0999914	.0017421	57.40	0.000	.096577 .1034057
μ	-42.25176	7.461329	-5.66	0.000	-56.8757 -27.62783
σ	645.3855	4.423774	145.89	0.000	636.7151 654.0559

Cuadro 2.4.8.2 Estimación de la marginal para el incumplimiento de Julio de 2008.

					Number of obs= 7970
					Wald chi2(0) = .
					Prob > chi2 = .
Log likelihood = 21783.773					[95% Conf. Interval]
	Coef.	Std. Err.	z	P> z	
γ	-.1145995	.0028969	-39.56	0.000	-.1202774 -.1089215
μ	.0103878	.0001771	58.65	0.000	.0100406 .0107349
σ	.0149633	.0001088	137.48	0.000	.01475 .0151766

Cuadro 2.4.8.3 Estimación de la marginal para la sobretasa de Julio de 2008.

Después de la maratónica presentación de los resultados de las marginales, es posible cerciorarse de la validez estadística de todos los estimadores de esta primera etapa, así como de los intervalos a un noventa y cinco por ciento de confianza. La prueba de validez conjunta de Wald no fue presentada por el programa pero siempre quedan las del Multiplicador de Lagrange o el cociente de verosimilitud para validar la regresión en su conjunto, aunque por requerir de nuevas estimaciones para los modelos restringidos, fueron obviadas.

2.5 Breve interpretación de las marginales

Es posible que haya sido notorio el que el parámetro de forma de la marginal del incumplimiento es consistentemente positivo a lo largo de los ocho cortes transversales, esto obedece al signo positivo con el que se registraron los incumplimientos dentro de la contabilidad de la agencia norteamericana (tal vez por simplicidad), lo que no afecta al análisis pues sólo se hace uso de la simetría entre las distribuciones Frechét y Weibull.

Para su mejor análisis, se presentan las esperanzas condicionales de los parámetros, hay que recordar que son variables aleatorias, en un cuadro que ayudará a comprender mejor la evolución de los tres factores de riesgo analizados a lo largo de los ocho meses de observación. Básicamente se busca observar el desarrollo de los eventos del mercado al mirar la trayectoria de los parámetros de forma de las marginales.

En el siguiente cuadro se presentan únicamente los parámetros de forma (en parametrización gamma) de las funciones marginales agrupados mensualmente. La decisión de sólo presentar los parámetros de forma obedece a que son estos los que determinan la existencia de colas pesadas y la cantidad de momentos muestrales susceptibles de ser encontrados analíticamente. Para mayor referencia, véase AIII.

Mes	Prepago	Parametrización γ		
		Incumplimiento	Sobretasa	Mortalidad
Dic-07	0.0251871	0.0817506	-0.1110657	N.D.
Ene-08	0.1173872	0.0820187	-0.1081657	-0.004036204
Feb-08	0.1198164	0.0841452	-0.1091369	-0.003315731
Mar-08	0.0996033	0.0884827	-0.1119027	-0.002464268
Abr-08	0.0842275	0.0939276	-0.1144707	-0.004076087
May-08	0.0269054	0.0957523	-0.1196182	-0.003472653
Jun-08	0.1354991	0.0970564	-0.1212998	-0.004604854
Jul-08	0.116462	0.0999914	-0.1145995	-0.003500875

Cuadro 2.5.1 Resumen de los parámetros de forma (gamma) de las marginales.

En general, los parámetros asociados a la sobretasa y a los incumplimientos son relativamente estables a lo largo de la muestra, presentando ligeras oscilaciones asociadas

a los eventos de mercado y a las intervenciones de la reserva federal en el caso de la sobretasa. Sin embargo se ve un paulatino pero constante deterioro de las condiciones de mercado que se puede observar en el aumento (disminución) del parámetro de forma gamma (alpha)⁴² que está fuertemente asociado con la “pesadez” de las colas, lo que a su vez está relacionado con un aumento de la probabilidad de aparición de valores extremos. Lo que invariablemente llevará a una caída en el valor de los instrumentos asociados a las hipotecas.

Las mayores variaciones en las estimaciones están presentes en los parámetros de forma de los prepagos, esta variabilidad está dada tanto por la turbulencia del mercado como por la construcción de la variable, pues incorpora las recompras producto de absorciones de incumplimientos de más de cuatro meses, y en general, reconocimientos de pérdidas que son absorbidas por la agencia y sacadas del “lote”.

Esta política de recompra puede ser apreciada por el salto ocurrido de Diciembre de 2007 a Enero de 2008 cuando muchas de las hipotecas que respaldaban a estos bonos entraron en un proceso judicial de adjudicación de colaterales y fueron sacadas del conjunto bursatilizado para ser prepagadas a los tenedores de los bonos (en muchos casos bancos) que se vieron en la necesidad de reconocer la pérdida asociada a estos prepagos masivos. De ahí la existencia de grandes cantidades de prepagos e incumplimientos aún cuando las tasas de largo plazo permanecieron estables.

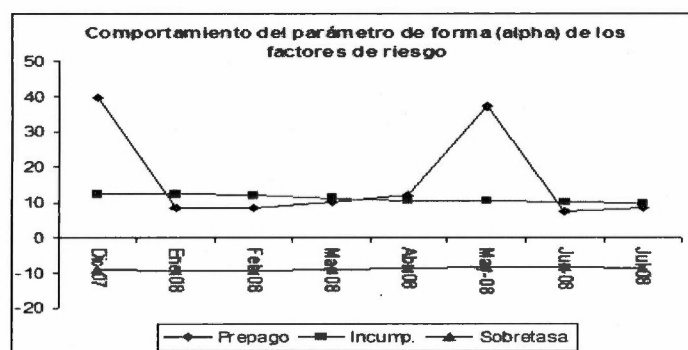
Más allá de esos grandes saltos producto del reconocimiento de pérdidas, es posible observar el mismo deterioro en las condiciones de mercado al analizar la disminución (aumento) del parámetro de forma gamma (alpha) que implica una menor cantidad de prepagos asociados al refinanciamiento lo que a su vez conlleva una menor probabilidad de ocurrencia de prepagos fuertes, lo que finalmente se traduce en una reducción de las colas que el lector observa en un alpha más grande.⁴³

⁴² Se le recuerda al lector que la parametrización gamma se construye bajo el supuesto de $\gamma=1/\alpha$.

⁴³ Es posible encontrar analíticamente valores para los momentos muestrales de las distribuciones de EVT hasta el $1-(j/\alpha)>0$ -ésimo momento.

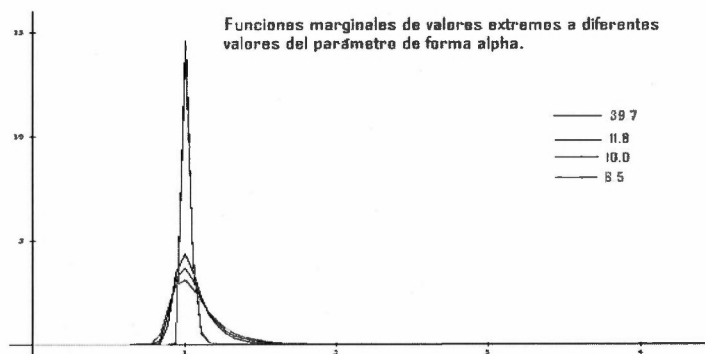
Esta “irrupción” en el mercado mediante la recompra de bonos con incumplimiento manifiesto, especificada por contrato, puede apreciarse mejor al graficar los parámetros de forma (alpha) como función del tiempo para los prepagos.

Es posible pensar en la posibilidad de modelar estos parámetros de forma como variables aleatorias sujetas al análisis de series de tiempo, lamentablemente la serie histórica con la que se elaboró esta investigación no es lo suficientemente larga como para soportar un análisis serio de esta clase con la subsecuente inferencia.



Gráfica 2.5.1 Seguimiento mensual de los parámetros α de forma por factor de riesgo.

Para hacer más clara la relación entre los parámetros de forma, la “pesadez” de las colas y el aumento de las probabilidades de encontrar valores extremos conforme las colas se hacen pesadas, lo que se refleja en una menor alpha, se presenta el siguiente gráfico.



Gráfica 2.5.2 Funciones de densidad EVT para distintos parámetros α .

Se puede observar una mayor concentración de masa de probabilidad en sitios alejados

de la media conforme el parámetro α se hace más pequeño. Parece claro que el lector preferirá encontrarse en poder de un bono cuyo prepago se modele mediante una α grande pues implica que los prepagos estarán mejor agrupados alrededor de la media y será menos probable que sufra un quebranto por un prepago masivo.

Por claridad y brevedad en la exposición, se presentará sólo una función de densidad de probabilidad marginal acumulada una vez que los parámetros de la primer verosimilitud han sido sustituidos en la forma funcional asociada a la función de cuantiles presentada en la ecuación (1.5.2). La función arbitrariamente escogida como función tipo es la usada para modelar el incumplimiento del mes de marzo de 2008.

Sustituyendo los parámetros de forma, $\gamma = .0884827$, ubicación, $\mu = -44.4496$, y escala, $\sigma = 619.7717$, descritos en el cuadro 2.4.4.2 se tiene que:

$$G_\gamma(x) = \exp \left\{ - \left(1 + .0884827 \left(\frac{x + 44.4496}{619.7717} \right) \right)^{-1/.0884827} \right\}, \quad (2.5.1)$$

Es en este momento, al sólo “sustituir” los valores de las regresiones en las formas funcionales de las marginales, cuando el lector puede apreciar los beneficios de haber construido todas las funciones marginales usando como base la familia de valores extremos en su parametrización gamma. Esto reduce dramáticamente la cantidad de esfuerzo necesaria para obtener las funciones de cuantiles que alimentarán a las cópulas.

Una vez estimadas todas las funciones marginales que constituyen las bases de las cópulas, es necesario hacer una estimación preliminar del parámetro de dependencia, *alpha*, (distinto de la parametrización de las funciones de la EVT). Para ello se hará uso de la Beta de Blomqvist analizada en la sección 1.4.3 de este trabajo.

2.6 Estimación preeliminar de los parámetros de dependencia.

El primer paso sugerido en la estimación de las dependencias es la construcción de la versión muestral de la Beta de Blomqvist para cada pareja posible (3 parejas distintas) en cada mes de observación (ocho meses), los cuales constituirán los valores iniciales para la segunda estimación de la verosimilitud.

Para estimar estos parámetros es necesario obtener las medias muestrales de cada uno de los factores, para posteriormente establecer una condicional que asignará un valor de uno a cada caso de comonmonotonicidad y un valor de menos uno para cada caso de contramonotonicidad entre las variables analizadas.

Se entiende por comonmonotonicidad al hecho de que ambos factores de riesgo analizados se encuentren por encima o por debajo de sus medias al mismo tiempo, *i.e.* se mueven en el mismo sentido alrededor de sus medias. Siguiendo la misma línea, se entiende por contramonotonicidad al hecho de que una de las variables analizadas este por debajo de su media mientras que la otra está por encima de la propia.

Básicamente, se está siguiendo el estimador sugerido en la ecuación (1.4.3), el cual es implementado mediante una condicional de la forma

$$f(x) = \begin{pmatrix} 1 & \text{si } ((x_i - \bar{x}) < 0 \cap (y_i - \bar{y}) < 0) \cup ((x_i - \bar{x}) > 0 \cap (y_i - \bar{y}) > 0) \\ -1 & \text{si } ((x_i - \bar{x}) < 0 \cap (y_i - \bar{y}) > 0) \cup ((x_i - \bar{x}) > 0 \cap (y_i - \bar{y}) < 0) \end{pmatrix}, \quad (2.6.1)$$

los resultados de esta función aplicada a cada una de las observaciones de la muestra arrojarán una serie de valores que al ser sumados constituirán el numerador de la ecuación (1.4.3), a saber: $\beta = \frac{n_1 - n_2}{n_1 + n_2}$. En cuanto al denominador, por definición, se puede resumir como el tamaño total de la muestra, pues $n_1 + n_2 = n$. En este caso, n_1 es el número de observaciones comonmonotónicas, mientras que n_2 es el número de observaciones contramonotónicas.

Una vez aplicada la metodología a todo el conjunto de los datos, se obtienen una serie de parámetros que constituyen una guía sobre la dependencia de las variables, estas

dependencias pueden ser “pegadas” de forma provisional en la cópula bajo la consigna de su validez muestral. Las dependencias obtenidas de esta forma son resumidas en la siguiente tabla:

Mes	Valores muestrales Beta Blomqvist		
	Prepago y sobretasa	Incumplimiento y sobretasa	Prepago e incumplimiento
Dic-07	0.102495108	0.082925636	0.651541096
Ene-08	0.1079455	0.082893283	0.656269188
Feb-08	0.098817151	0.08501725	0.6696649
Mar-08	0.091403162	0.089426877	0.6564970
Abr-08	0.087932531	0.082475505	0.6722064
May-08	0.095955196	0.085252022	0.6596142
Jun-08	0.096524131	0.08652163	0.6707927
Jul-08	0.093851945	0.13425345	0.7577164

Cuadro 2.6.1 Resumen de los parámetros de dependencia muestrales.

En el cuadro anterior, el lector puede observar una relativa estabilidad entre los pares de dependencias, lo que en sí mismo constituye un hallazgo pues implica que aún dentro de la turbulencia financiera vivida a lo largo de los meses analizados, las dependencias entre las variables permanecen más o menos constantes, lo que a su vez sugiere estabilidad estructural una vez iniciado el ajuste en los agrupamientos de las observaciones.

Un hecho todavía más llamativo es el que todas las dependencias sean positivas, lo que es contrario a la hipótesis originalmente planteada en la investigación, la cuál planteaba una relación negativa entre la sobretasa pagada por refinanciamiento y los prepagos observados.

La explicación a este fenómeno está dada por la construcción de la variable prepago, la cual incluye las recompras hechas por la agencia cuando el crédito es declarado incobrable y se decide ejecutar la garantía, lo que implica un aumento artificial en los prepagos asociado a un aumento en los incumplimientos lo que a su vez está asociado a una sobretasa alta, es decir, una dependencia positiva en todas las relaciones producto del estrés financiero al momento de tomar las observaciones.

Ahora sólo resta mostrar al lector las estimaciones formales realizadas para hallar los

parámetros de dependencia de las cópulas. Para ello se puede hacer uso de las funciones de verosimilitud mostradas en las ecuaciones (1.6.6), (1.6.8) y (1.6.10) o recurrir a la metodología de *Bootstrapping* en caso de que la función de verosimilitud sea tan plana que la vuelva incapaz de converger.

2.6.7 Estimación formal de los parámetros de dependencia.

Originalmente se había planteado hacer una estimación totalmente paramétrica del modelo, sin embargo la magnitud de las colas de ambas marginales y la fuerte dependencia exhibida en algunas de las relaciones, hacen imposible al algoritmo de máxima verosimilitud declarar convergencia a menos de forzarla⁴⁴ lo que invariablemente desnaturalizaría la estimación.

De cualquier modo, están disponibles en el apéndice AV las transcripciones de los archivos “do” elaborados en STATA© para el lector interesado. De ellos resulta importante hacer notar que fue necesario aproximar a cero punto nueve mil novecientos noventa y nueve y un cien milésimo respectivamente, los valores de uno y cero obtenidos como resultados de las funciones de densidad de probabilidad. Esto con el fin de ponerlas dentro del dominio de la función de cópula que contiene logaritmos naturales evaluando estos resultados (probabilidades).

Estos valores de cero y uno son resultado de las colas tan pesadas presentes en los últimos meses y aunque aproximarlas constituye algo cercano a la censura de los datos, es necesario modificarlas a fin de evaluar las funciones. Tal y como se puede imaginar, entre más dígitos se añada a la aproximación, menor será la “censura” practicada. Se usa la palabra censura entrecomillada dado que haciendo uso del concepto de límite al añadir decimales, es mínima la información perdida y es posible pensar que estos valores de cero y uno son producto del redondeo propio de la computadora.

⁴⁴ Esto puede lograrse en el STATA© usando el comando “ltolerance *valor*” al usar la instrucción de maximización “ml maximize”, sin embargo el hacerlo en una función casi plana implica sesgar el estimador al detenerla, posiblemente mucho antes de encontrar un óptimo.

Haciendo uso de las sugerencias planteadas por el artículo de Chen y Fan (2002), se realizó una segunda estimación (la del parámetro de dependencia) de forma semiparamétrica haciendo uso de la técnica del Bootstrapping. Dado lo conocido y común de la técnica, se omitirá su descripción para pasar de inmediato a los resultados obtenidos, no sin antes poner a disposición del lector interesado, en el apéndice AV, el algoritmo de BVA para Excel© usado en el proceso.

Resumen de las estimaciones obtenidas para el parámetro de dependencia α				
Mes	Concepto	int [.05,.95]	Media	Desv. est.
Diciembre	ψ, λ	0.626, 0.6961	0.6616032	0.021984114
	r, λ	0.088, 0.18	0.1329512	0.028275402
	r, ψ	0.07, 0.164	0.1169124	0.028485938
Enero	ψ, λ	0.638, 0.708	0.6733492	0.021399417
	r, λ	0.088, 0.178	0.132516	0.027831824
	r, ψ	0.074, 0.168	0.1206636	0.028612821
Febrero	ψ, λ	0.652, 0.722	0.6868034	0.02108027
	r, λ	0.092, 0.184	0.1362876	0.028188763
	r, ψ	0.072, 0.166	0.119033	0.028567865
Marzo	ψ, λ	0.65, 0.72	0.6844488	0.021574622
	r, λ	0.096, 0.188	0.1423636	0.02810649
	r, ψ	0.07, 0.164	0.1159652	0.028586299
Abril	ψ, λ	0.644, 0.716	0.680807	0.02176707
	r, λ	0.086, 0.176	0.130213	0.028168734
	r, ψ	0.068, 0.162	0.1153744	0.028455031
Mayo	ψ, λ	0.638, 0.706	0.6722604	0.021416827
	r, λ	0.086, 0.178	0.1305702	0.028327265
	r, ψ	0.07, 0.164	0.1159322	0.028668582
Junio	ψ, λ	0.642, 0.712	0.6769108	0.021141238
	r, λ	0.084, 0.176	0.1300626	0.028175221
	r, ψ	0.07, 0.162	0.116587	0.02825894
Julio	ψ, λ	0.666, 0.736	0.7009268	0.020926118
	r, λ	0.084, 0.178	0.130122	0.028749215
	r, ψ	0.076, 0.17	0.1225152	0.028140877

Cuadro 2.6.2 Resumen de los α 's con diez mil simulaciones de mil pares de datos.

En el cuadro anterior se presentaron los resultados promedio, desviaciones estándar e intervalos de confianza obtenidos por el algoritmo de Bootstrapping presentado en AV tomando diez mil muestras integradas cada una por mil pares de observaciones de cada

binomio de factores de riesgo (ψ , λ y r) analizado para un intervalo de confianza con el 5% de probabilidad de cometer un error de tipo II (no rechazar una hipótesis nula falsa).

De los resultados obtenidos, hay varios puntos interesantes en los cuales es necesario poner especial atención. El primero es el hecho de tener un intervalo de confianza al 95% extraído de los respectivos percentiles, (5% y 95%), muy similar al que se obtendría de sumar y restar dos veces la desviación estándar al valor promedio. Esta similitud implica la convergencia a la distribución normal de los parámetros de dependencia obtenidos siguiendo el algoritmo presentado en AV, tal y como se explicó en AIV. El siguiente cuadro resume los resultados.

Resumen de α 's para distintos tamaños de remuestreo.					
Mes	Concepto	10,000	1,000	100	10
Diciembre	ψ, λ	0.6616032	0.661022	0.66346	0.671
	r, λ	0.1329512	0.132996	0.13536	0.1348
	r, ψ	0.1169124	0.116478	0.12342	0.113
Enero	ψ, λ	0.6733492	0.67403	0.67908	0.6628
	r, λ	0.132516	0.131952	0.14184	0.1316
	r, ψ	0.1206636	0.120654	0.12716	0.1148
Febrero	ψ, λ	0.6868034	0.686968	0.68728	0.6888
	r, λ	0.1362876	0.137442	0.13334	0.1382
	r, ψ	0.119033	0.118726	0.12062	0.1194
Marzo	ψ, λ	0.6844488	0.68406	0.68282	0.6812
	r, λ	0.1423636	0.14049	0.14228	0.13
	r, ψ	0.1159652	0.115554	0.11594	0.1236
Abril	ψ, λ	0.680807	0.681032	0.68146	0.6788
	r, λ	0.130213	0.130322	0.12906	0.1282
	r, ψ	0.1153744	0.115386	0.11804	0.1066
Mayo	ψ, λ	0.6722604	0.6719	0.67264	0.6844
	r, λ	0.1305702	0.130354	0.1258	0.1196
	r, ψ	0.1159322	0.115426	0.117	0.1076
Junio	ψ, λ	0.6769108	0.677258	0.67828	0.6744
	r, λ	0.1300626	0.130354	0.13312	0.1408
	r, ψ	0.116587	0.116616	0.1202	0.1212
Julio	ψ, λ	0.7009268	0.701934	0.70506	0.6884
	r, λ	0.130122	0.131192	0.12798	0.124
	r, ψ	0.1225152	0.122666	0.1248	0.1072

Cuadro 2.6.3 Convergencia de los parámetros de dependencia.

El segundo punto a notar es la similaridad de los parámetros de dependencia obtenidos con diez, cien, mil y diezmil corridas (remuestreos) de mil parejas de datos cada uno, lo que implica convergencia de los parámetros muestrales a los poblacionales conforme crece la muestra. Esta consistencia empírica está respaldada por lo explicado en el apéndice AIV el cual se basa en el trabajo de László Gyrgy & Zoltán (2002).

Al igual que en cualquier estimación empírica, el ganar precisión conlleva un aumento en el poder de computadora requerido y el tiempo de estimación. Sólo como dato curioso, se comenta que con un procesador “Centrino Duo”[®] y una memoria DDR de 512 Mb, tomará alrededor de veinte minutos realizar mil remuestreos y cerca de doscientos minutos realizar diez mil. Se hace este comentario para que el lector ajuste sus necesidades de precisión a la demanda de tiempo que lleva implícita.

De la tabla anterior, es posible observar que la dependencia entre el prepago y el incumplimiento se va haciendo mayor conforme avanza el tiempo a lo largo de la muestra. De nuevo se recuerda que por contrato, los créditos con retrasos mayores a cuatro meses son prepagados al tenedor del bono por la institución emisora y posteriormente, si es posible, recobrados por la misma. Esta política de recompra ante incumplimiento ha forzado al parámetro de prepago a seguir al de incumplimiento.

Si se reviza la construcción de la Beta de Blomqvist (ecuación 1.4.3), se observará que es una medida normalizada cuyo valor es uno en perfecta comonmonotonicidad y menos uno en perfecta contramonotonicidad. Al encontrar dependencias cercanas a siete decimos se está encontrando que los prepagos están siguiendo de cerca a los incumplimientos, siendo estos la causa de la mayor parte de los primeros.

El signo contrario a la teoría planteada en la introducción es un indicador de alarma por su implicación de incumplimientos masivos, sin embargo el que el parámetro aumente de valor conforme avanza el tiempo en la muestra es una terrible señal de que en ese momento el problema estaba empeorando y aún no tocaba fondo, tal y como se puede recordar de los meses anteriores a Febrero de 2009.

Este hecho revela al parámetro de dependencia de la cópula FGM como un indicador

de la salud del mercado hipotecario y en la coyuntura de la recesión de 2009, como un indicador de la posibilidad de recuperación o cese del problema al menos en el mercado hipotecario que lo originó. Incluso puede ser usada como técnica base para la construcción de indicadores adelantados de la actividad económica al vincularla con factores precursores que el investigador considere pertinentes. Esto queda como una posible línea de investigación futura.

Siguiendo con el análisis de los parámetros de dependencia, se puede observar que la relación de los incumplimientos con la sobretasa se acerca a los ciento treinta milésimos y ha bajado lenta pero consistentemente, lo que indica una paulatina desvinculación de la sobretasa pagada por refinanciamiento, lo cual está correlacionado con la tasa de interés de largo plazo, con los incumplimientos.

Los niveles relativamente cercanos a cero de la dependencia establecen una desvinculación previa de estos parámetros, lo que a su vez sugiere que el problema de incumplimientos no está dado explicado actualmente por altas tasas de interés, relativamente constantes en la muestra, sino por cambios más profundos con raíces en la economía real, *v.gr.* caída del PIB.

Este contagio del mercado financiero, identificado como origen del problema, hacia la economía real cuyo “vector de contagio” fue la caída del valor de los hogares por la alza en las tasas de interés que a su vez provocó pérdidas no esperadas del segmento más riesgoso del mercado hipotecario (*subprime*), las cuales no habían sido correctamente evaluadas por los inversionistas ha dejado de ser el motor del problema aún cuando las tasas de largo plazo permanecen altas. Esto puede ser observado por el nivel bajo de dependencia entre estos dos factores de riesgo.

Esto no implica que el problema se esté solucionando, sólo muestra que se ha esparcido lo suficiente para que el motor del problema se haya trasladado a la economía real. La observación del cuadro anterior arroja que fue en Marzo de 2008 cuando se observó el pico de la dependencia entre estos dos parámetros de riesgo, es decir, fue en Marzo de 2008 cuando el contagio del sector financiero a la Economía real alcanzó su cúspide y una oleada

de pánico lo trasladó vía expectativas al sector real.

De tomarse como referencia el intervalo de confianza de Marzo de 2008, no puede establecerse un cambio estructural en el comportamiento de estos dos parámetros de riesgo. Aunque formalmente es necesaria una prueba de estabilidad estructural (de Chow⁴⁵), la inspección (comparación de los intervalos) sugiere que no ha habido cambios en la dependencia a lo largo de la muestra, lo que puede implicar que el problema viene de más atrás y que antes de iniciado el muestreo ya se habían dislocado las dependencias por la extensión del problema al resto de la Economía.

Por último, el parámetro de dependencia de la cópula Weibull, cuyo signo es contrario a lo esperado, únicamente muestra una dislocación en el mercado hipotecario pues los prepagos por incumplimiento han contaminado la variable de tal forma que su valor en niveles es similar a la dependencia entre la sobretasa y el incumplimiento, es posible que al pasar la turbulencia financiera, los parámetros regresen al comportamiento predicho por la teoría.

⁴⁵ Es posible realizar esta prueba que se distribuye como "F" dada la distribución normal de los parámetros.

Parte III

Aproximación del valor del bono y conclusiones.

3.1 Aproximación final al valor de los BRH

Después de una exhaustiva revisión de la literatura y de varios procesos de simulación e inferencia estadística, se está en posibilidades de calcular el valor promedio de las pérdidas esperadas por las interacciones de los factores de riesgo al multiplicar el valor promedio de las cópulas dados los parámetros obtenidos a lo largo de la tesis con el tamaño promedio de los créditos analizados. Para ello es necesario obtener el valor esperado de la cada una de las tres cópulas en cada uno de los ocho cortes transversales.

Esto es realizado mediante un algoritmo en BVA para Excel© que remuestreando diez mil veces los valores de la base de datos, genera submuestras de mil observaciones cada una, las cuales alimentan a las funciones marginales de densidad que a su vez son insumos de cada una de las cópulas. Todo este proceso arroja valores para cada una de las tres cópulas dadas las observaciones remuestreadas.

Es necesario realizar simultáneamente la simulación de las tres cópulas para garantizar que las observaciones que juegan el papel de insumos correspondan al mismo crédito y por tanto estén relacionadas, capturando con ello las relaciones entre los tres procesos. Es claro que toda la descripción anterior recuerda a un híbrido entre un método de integración numérica y un remuestreo sin reemplazo, véase Chung(1993).

Se realizó la estimación de promedios en este híbrido, transcrito en AV, dada la necesidad de replicar las interdependencias de la muestra, lo que no habría sido posible haciendo uso de un Montecarlo tradicional o cualquiera de sus versiones, pues este implica generar tres series independientes de números pseudoaleatorios distribuidos uniformemente que al

ser usados como insumos de la cópula habrían dado la misma probabilidad a cualquier valor. Más aún, hubieran destruido cualquier estructura de dependencia aún usando una matriz de Choleski, la cual sólo transmite dependencias de primer orden.

En el siguiente cuadro se presentan las probabilidades de obtener de manera conjunta las pérdidas promedio presentes en la muestra con un intervalo de confianza de 95%. De nuevo es posible observar que los intervalos obtenidos a través de los cuantiles es similar a los que se obtendrían de sumar y restar dos veces la desviación estándar, es decir se observa la convergencia a la distribución normal.

Probabilidades conjuntas de las pérdidas promedio.				
Mes	Concepto	int [.05,.95]	Media	Desv. Est
Diciembre	ψ, λ	0.211291738 , 0.222206984	0.21661739	0.003328087
	r, λ	0.212977685 , 0.224725401	0.218855932	0.003575617
	r, ψ	0.31659884 , 0.319114672	0.317877595	0.000762049
Enero	ψ, λ	0.211922907 , 0.223142329	0.217442022	0.003406042
	r, λ	0.21304869 , 0.224688005	0.218865168	0.003578363
	r, ψ	0.323208602 , 0.325814976	0.324504189	0.00079936
Febrero	ψ, λ	0.212310587 , 0.223341648	0.217705125	0.003342612
	r, λ	0.212822504 , 0.224612619	0.218724377	0.003548822
	r, ψ	0.323297703 , 0.325990484	0.324654026	0.000817757
Marzo	ψ, λ	0.208432848 , 0.218831458	0.21357853	0.003199015
	r, λ	0.212163743 , 0.2237423	0.217929948	0.003490732
	r, ψ	0.322079235 , 0.324785433	0.323451739	0.00082218
Abril	ψ, λ	0.212018186 , 0.223355601	0.217572026	0.003446355
	r, λ	0.213024433 , 0.22456891	0.218756546	0.003495601
	r, ψ	0.320622026 , 0.323317168	0.321987638	0.000827336
Mayo	ψ, λ	0.21016626 , 0.221206821	0.215667071	0.00337633
	r, λ	0.211070458 , 0.222605413	0.216830755	0.003500424
	r, ψ	0.314899426 , 0.317527746	0.316238337	0.000806059
Junio	ψ, λ	0.211779924 , 0.223380028	0.217508055	0.003522124
	r, λ	0.210768285 , 0.222297854	0.216499497	0.003506388
	r, ψ	0.321127944 , 0.324108289	0.322629511	0.000902397
Julio	ψ, λ	0.213059501 , 0.22389783	0.218352761	0.003311304
	r, λ	0.211029721 , 0.222661259	0.216854335	0.003519636
	r, ψ	0.321144167 , 0.323909448	0.322524854	0.000834686

Cuadro 3.1.1 Probabilidades promedio conjuntas de las pérdidas promedio.

Es posible percatarse del uso de la propiedad de la torre de las esperanzas en la construcción de estos intervalos, pues formalmente se ha tomado la “la probabilidad promedio de obtener de forma conjunta la pérdida promedio presente en el conjunto original de los datos”. Lo que puede traducirse como la probabilidad conjunta de observar las pérdidas promedio actuales.

Al igual que en cualquier intento de valuación de activos contingentes, únicamente se pueden obtener promedios e intervalos de confianza para las estimaciones dado el conjunto de datos disponibles, por lo que el valor de la pérdida conjunta a sustituir en la ecuación 1.2.11 será la esperanza condicional (a los datos) de cada una de las cópulas a formarse con los parámetros obtenidos a lo largo de la tesis. Esta esperanza puede obtenerse de forma analítica al resolver la doble integral del producto de las variables por la cópula, o de manera numérica al obtener los promedios acoplado para ello el algoritmo de solución de la cópula.

Como se puede imaginar, los resultados dependerán, al igual que en cualquier esperanza condicional, del conjunto de datos usado. Por lo que se recomienda mantener un conjunto de datos suficientemente amplio de manera tal que se pueda llevar a cabo un remuestreo eficiente. Por esta razón es que se han tomado todos los créditos disponibles en la base de datos.

Tal vez una de las características más importantes del método propuesto es que puede ser usado a nivel crédito al analizar cada uno de los créditos dentro del bono, o a nivel bonos en el mercado, analizando varios bonos de la misma clase y extrapolando su información al que se desea valorar. Todo esto, dependiendo de la cantidad de información disponible, coloca esta aproximación entre los modelos de segunda (basada en información de bonos) y tercer generación (basada en información de cada crédito) de valuación.

Lamentablemente la naturaleza extrema de las observaciones del problema provoca que las medias de las marginales no sean los mejores estimadores para el valor del bono, pues como se ha visto a lo largo del trabajo la dispersión es mucha y va más allá de lo que las simples media y varianza pueden explicar. Por ello se recomienda usar el algoritmo

de valuación de las cópulas y calcular la probabilidad conjunta de las pérdidas observadas para que, con ayuda de estos datos, sea posible calcular intervalos de confianza para los valores del bono.

Antes de seguir con la segunda aplicación de los resultados de la tesis, se presentan las probabilidades marginales de ocurrencia de los valores promedio de la muestra. De nuevo se hace uso de la propiedad de la torre de las esperanzas para su elaboración

Mes	Concepto	Probabilidades marginales de las pérdidas promedio.		
		int [.05, .95]	Media	Desv. Est
Diciembre	ψ	0.419545372 , 0.440429421	0.429195974	0.006214577
	λ	0.420582466 , 0.434487046	0.427660143	0.00448044
	τ	0.49684235 , 0.517698195	0.506546148	0.006772214
Enero	ψ	0.422171868 , 0.439725104	0.43194056	0.005561451
	λ	0.418799244 , 0.433331148	0.427096946	0.004525026
	τ	0.493831849 , 0.51705566	0.505841472	0.007168625
Febrero	ψ	0.421419137 , 0.439220984	0.430613785	0.005587398
	λ	0.417257061 , 0.430594433	0.424383054	0.003891426
	τ	0.492456311 , 0.518047629	0.505885959	0.007335406
Marzo	ψ	0.417964899 , 0.43553943	0.426981862	0.005429436
	λ	0.414092978 , 0.427480645	0.421358663	0.004143008
	τ	0.494952729 , 0.51625686	0.50644319	0.00632639
Abril	ψ	0.417663721 , 0.437596447	0.427937965	0.006478809
	λ	0.421018798 , 0.433777561	0.427606799	0.003978895
	τ	0.493655166 , 0.515410955	0.504713938	0.00713132
Mayo	ψ	0.419222162 , 0.43861128	0.428128455	0.006001622
	λ	0.414471091 , 0.430870281	0.423019975	0.004917929
	τ	0.495173354 , 0.518220435	0.506864107	0.007175375
Junio	ψ	0.424211674 , 0.44300295	0.433417274	0.005835851
	λ	0.415879191 , 0.429951391	0.422826172	0.004431206
	τ	0.495446679 , 0.515341781	0.504371414	0.006968364
Julio	ψ	0.424963117 , 0.44048818	0.432385164	0.005318043
	λ	0.41680405 , 0.431460024	0.424228021	0.004316809
	τ	0.493734825 , 0.516129958	0.505612337	0.0069720761

Cuadro 3.1.2 Probabilidades promedio marginales de las pérdidas promedio.

Con estas probabilidades en mano, es posible pasar a la segunda aplicación de este desarrollo, el control de riesgos. La primer forma sugerida es a través de los llamados

“periodos de retorno” que son una medida de riesgo usada en hidrología definida como el periodo de tiempo promedio transcurrido entre dos ventos de la misma magnitud, véase Shiau, Wang & Tsai (2006). Este concepto puede ser escrito como:

$$T_{\lambda,\psi,r} = \frac{1}{1 - F_{\lambda} - F_{\psi} - F_r + C_{\psi,r} + C_{\lambda,r} + C_{\lambda,\psi}}, \quad (3.1)$$

lo que implica la necesidad de usar las probabilidades de los valores promedio observados de las marginales dadas en el cuadro anterior. La forma de obtenerlas está incluida en el algoritmo transcrito en AV usado para conocer las probabilidades de ocurrencia de los valores promedio de la muestra.

La sustitución directa de los valores obtenidos a lo largo de la tesis lleva a los periodos de retorno, *i.e.* tiempos de espera entre eventos de igual magnitud, mostrados en la siguiente tabla:

Periodos de retorno para los ocho cortes transversales.	
Mes	Periodo de retorno
Diciembre	2.564440208
Enero	2.525683665
Febrero	2.498746067
Marzo	2.498897358
Abril	2.512199815
Mayo	2.559353809
Junio	2.525110953
Julio	2.528403911

Cuadro 3.1.3 Periodos de retorno mensual para los valores promedio de la muestra.

Los resultados anteriores parecen alarmantes, pues implican que de mantenerse los mercados como hasta hoy, se puede esperar una crisis similar en dos años y medio. Sin embargo, esta predicción supone que no hay un cambio en la estructura básica del mercado y el comportamiento de sus participantes, *i.e.* no se atenuan las colas, lo que es poco creíble dada la adaptabilidad de los mercados y su capacidad de aprendizaje.

Desafortunadamente, el resultado parece verificarse al realizar un sondeo rápido en internet con la palabra “*subprime*”, lo cual arroja resultados tan “antiguos” como informes

de la reserva federal de los Estados Unidos que datan de 2004,⁴⁶ posteriormente, se puede encontrar que el primer signo de alarma en los mercados hipotecarios se dió en agosto de 2007 cuando la reserva federal otorgó créditos puente sin precedentes a las agencias encargadas de velar por el mercado, lo cual tranquilizó a los inversionistas y sólo quedó como una alarma pasajera hasta 2008 cuando nuevos signos de alarma se dispararon en bancos de gran envergadura.

La otra forma de control de riesgo sugerida una vez conocidas las distribuciones conjuntas es el ajuste de éstas a la metodología del VaR_{1-q}^x , lo que consiste simplemente en obtener las pérdidas esperadas de las marginales como los cuantiles al 95% y ajustar estas pérdidas con las pérdidas conjuntas que se obtendrían con la misma confianza. Esto se puede lograr tomando el 5% peor percentil de la muestra y aplicando el mismo análisis sobre esa muestra sesgada, después aplicar promedios tomar esos promedios como el VaR_{1-q}^x .

La desventaja básica de esta aproximación está dada por la propia naturaleza del VaR_{1-q}^x , pues es “ciego” después del umbral, además de presentar todas las desventajas de la aproximación paramétrica basada en remuestreo pues al final de cuentas es una sofisticación de la aproximación histórica con ponderadores dados por las cópulas y las marginales, quedando su aplicación como una línea de investigación futura.

Además de los controles de riesgo propuestos, la principal aportación de este trabajo es la metodología para formar cópulas que capturan todas las interdependencias entre los factores de riesgo de instrumentos no normales, lo que abre la puerta para muchos análisis e investigaciones futuras. A través del trabajo se proporcionan algoritmos que permiten formar las cópulas a partir de los datos “crudos”. Aunque en este trabajo se han usado datos del mercado inmobiliario de los Estados Unidos de Norteamérica, la metodología puede ser aplicada casi sin ajustes a cualquier problema de valuación no normal que implique dependencia de dos o más factores.

Otro subproducto de la tesis son los indicadores de “salud” del mercado inmobiliario

⁴⁶ Véase: <http://www.federalreserve.gov/boarddocs/Speeches/2004/20040521/default.htm>

proporcionados por los parámetros de dependencia de la cópula, los cuales mostraron la historia y el “vector de contagio”⁴⁷ del sector financiero hacia la economía real cuando se desvincularon los factores de incumplimiento con la sobretasa, lo que implica que las personas están incumpliendo por falta de ingresos derivados de la pérdida de sus empleos o disminución en los ingresos familiares, no por el monto de los intereses.

Sobre la misma línea se puede usar al parámetro de dependencia entre el prepago y el incumplimiento como indicador de la salud del mercado inmobiliario pues mientras corran en paralelo se puede inferir que los prepagos están dados por los incumplimientos, por lo que se vive un momento crítico en el mercado. Cuando el parámetro regrese a niveles cercanos a cero con un signo negativo, entonces se sabrá que el mercado se habrá estabilizado.

Las magnitudes de las colas marginales, así como las relaciones entre los factores de riesgo dados por las cópulas quedan como base para la valuación de instrumentos de manejo de riesgo similares a los actuales *Credit Default Swaps* inmersos en los contratos estándar del mercado y culpables de la ola de prepagos asociados al incumplimiento.

Los nuevos instrumentos pueden ser adaptados a las necesidades de cada bono haciendo uso de sus funciones de cópula para cada pareja de factores de riesgo arrojando valuaciones de instrumentos similares al *Credit Default Swap* que hará uso de la marginal de incumplimiento y la cópula que asocia el incumplimiento con la sobretasa ó al *Credit Prepayment Swap* que se propone como un seguro para el manejo del riesgo de prepago ante cambios en la tasa por refinanciamiento de manera tal que el tenedor reciba pagos por intereses y capital similares a los originalmente pactados de una contraparte si el acreditado decide prepagar. Su valor puede estar dado por el valor esperado de la marginal del prepago más un ajuste dado por la cópula del prepago y la sobretasa. Esto también queda como futura línea de investigación.

Otro resultado interesante es la necesidad de abrir los registros de las bursatilizaciones

⁴⁷ En medicina se le denomina “vector de contagio” a la forma de transmisión de una enfermedad entre individuos que incluso pueden ser de distinta especie, si se acopla a las finanzas, se puede leer como la forma de transmisión de la turbulencia de un mercado a otro.

de BRH en el mercado mexicano a los inversionistas de la misma forma que está hecho en el mercado de los Estados Unidos de Norte América, aunque se debe asegurar una longitud adecuada de las series para garantizar robustez en las estimaciones. Esta no es una cuestión menor pues implica una solución al problema de selección adversa que tiene el inversionista al comprar cada uno de estos bonos pues así conocerá que es lo que en realidad compra. todo lo anterior redundará en una mayor demanda por estos instrumentos en un momento en que los inversionistas recelan de cualquier forma de riesgo y más del asociado a las hipotecas.

Como se mencionó tangencialmente al inicio del trabajo, el objetivo último es propiciar profundidad al mercado a través de la transparencia y buen flujo de la información, de tal manera que incluso los deciles de menores ingresos sean sujetos de crédito una vez que los riesgos asociados a prestarles han sido debidamente cuantificados. Esto los alejará de las practicas predatorias e integrará a la economía de mercado y la competencia. Para decirlo en palabras llanas: El objetivo de esta tesis es encontrar la forma de que ms personas, aún con historiales crediticios imperfectos, tengan casa.

Parte IV

Apéndice I

Las cópulas y el teorema de Sklar

AI.1 Consideraciones técnicas:

Se iniciará definiendo dos variables aleatorias X_1, X_2 con funciones marginales de distribución $F_1(X_1) = F_1(X_1 \leq x_1)$ y $F_2(x_2) = F_2(X_2 \leq x_2)$ además de una distribución conjunta $H(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2)$. Es útil notar que todas las funciones mapean al intervalo $[0, 1]$, lo que implica que cada par de coordenadas x_1, x_2 corresponde a un punto en el cuadrado unitario sobre el plano \mathbb{R}^2 lo cual a su vez corresponde a un número $H(x_1, x_2)$ en $[0, 1]$.

También se define $\bar{\mathbb{R}}$ como la recta real extendida, lo que implica $[-\infty, +\infty]$. Siguiendo sobre esta línea se define $\bar{\mathbb{R}}^2$ como el plano real extendido, esto es $\bar{\mathbb{R}} \times \bar{\mathbb{R}}$. Ahora se define un rectángulo B sobre este plano, lo que resulta en $B = [x_{1a}, x_{1b}] \times [x_{2a}, x_{2b}]$

Sobre el mismo $\bar{\mathbb{R}}^2$ defina el rectángulo unitario \bar{I}^2 , donde $I = [0, 1]$. También defina una función de dos variables H cuyo dominio está en un subconjunto del plano extendido $\bar{\mathbb{R}}^2$ y su rango es un subconjunto de la recta real extendida $\bar{\mathbb{R}}$.

Ahora se definen dos subconjuntos no vacíos sobre la recta real extendida, $S_1, S_2 \subset \bar{\mathbb{R}}$, y se hace que la función de dos variables H tenga como dominio el rectángulo formado por los dos subconjunto anteriores, $\text{Dom}H = S_1 \times S_2$. Por otro lado, suponga que el producto cartesiano $B = [x_{1a}, x_{1b}] \times [x_{2a}, x_{2b}]$ es un rectángulo cuyos vértices están en el dominio de H . Si se desea calcular el H-volumen, o masa, de B se tiene que

$$V_H(B) = H(x_{1b}, x_{2b}) - H(x_{1b}, x_{2a}) - H(x_{1a}, x_{2b}) + H(x_{1a}, x_{2a}). \quad (AI.1)$$

Esta ecuación puede ser vista como una aplicación de la desigualdad del triángulo aplicada a una función bivariada acumulada de probabilidad si se especifica que es mayor

o igual a cero y que los vértices del rectángulo B se encuentran dentro del dominio de la función H . Todo esto llevará a la conclusión de que la función es \mathcal{L} -creciente. Lo anterior no implica ó es implicado por la idea de que H es *no decreciente* en cada argumento. Para decir que la función H es creciente en cada uno de sus argumentos, se recurrirá al siguiente lemma:

Lemma 1: Sean S_1 y S_2 dos subconjuntos no vacíos de la recta real extendida $\bar{\mathbb{R}}$ y sea H una función \mathcal{L} -creciente con dominio en el producto cruz de los dos subconjuntos anteriores, $\text{Dom}H = S_1 \times S_2$. Ahora se establece que los puntos x_{2a} y x_{2b} están dentro del segundo subconjunto S_2 , $x_{2a}, x_{2b} \in S_2$, donde además $x_{2a} \leq x_{2b}$. Entonces la función $\xi(t) = H(t, x_{2b}) - H(t, x_{2a})$ es no decreciente. Por simetría, se pueden establecer dos puntos $x_{1a}, x_{1b} \in S_1$ que también cumplen con $x_{1a} \leq x_{1b}$ y concluir que existe una función $\eta(t) = H(x_{1b}, t) - H(x_{1a}, t)$ que es no decreciente en S_1 . Ahora se establecerán los “pisos” para las mismas funciones. Para hacerlo se supondrá que los subconjuntos S_1 y S_2 tienen un elemento ínfimo a_1 y a_2 respectivamente. Entonces la función H con dominio en el área formada por los dos subconjuntos está limitada por abajo, *i.e.* tiene un piso, si

$$H(x_1, a_2) = 0 = H(a_1, x_2) \forall (x_1, x_2) \in S_1 \times S_2.$$

Con estas nuevas definiciones se llega al segundo lemma.

Lemma 2: Sean S_1 y S_2 subconjuntos de la recta real extendida, $S_1, S_2 \subset \bar{\mathbb{R}}$, y sea H una función \mathcal{L} -creciente con “piso”, cuyo dominio está en el rectángulo dado por $S_1 \times S_2$. Entonces H es no decreciente en cada argumento, esto se observó hasta que se les puso un “piso” a las funciones sobre una sola variable aleatoria.

A continuación se establecerá un procedimiento similar para definir “márgenes” en las funciones. Suponga que S_1 y S_2 tienen un elemento supremo b_1 y b_2 respectivamente. Entonces la función H con dominio en el área formada por los dos subconjuntos tiene “márgenes” y estos “márgenes” son las funciones F y G dadas por:

$$\text{Dom}F = S_1 \text{ y } F(x_1) = H(x_1, b_2) \forall x_1 \in S_1,$$

$$\text{Dom}G = S_2 y G(x_2) = H(b_1, x_2) \forall x_2 \in S_2.$$

Para ejemplificar los conceptos de “piso” y “techo” se propone la función

$$H(x_1, x_2) = \frac{(x_1 + 1)(e^{x_2} - 1)}{x_1 + 2e^{x_2} - 1},$$

con dominio en el rectángulo $[-1, +1] \times [0, 1]$ y que presenta un volúmen positivo, esto es: $V_H(B) \geq 0$. Lo que se puede comprobar sustituyendo cualquier par de valores del dominio dentro de la función, la cual siempre arrojará un valor no negativo.

Para buscar el “piso” de la función, basta con sustituir los valores más pequeños permitidos por el dominio (-1 y 0 respectivamente), primero para x_1 y luego para x_2 . Esto lleva a: $H(x_1, 0) = H(-1, x_2) = 0$.

Siguiendo la misma lógica para los márgenes, basta con sustituir los valores más grandes permitidos por el dominio (1 y ∞ respectivamente), primero para x_1 y luego para x_2 . Lo que conduce a $G(x_2) = 1 - e^{-y}$ y al $\lim_{x_2 \rightarrow \infty} \left[H(x_1, x_2) = \frac{(x_1+1)(e^{x_2}-1)}{x_1+2e^{x_2}-1} \right] = F_{x_1}$. Dado que ese límite lleva a una indeterminación, $\frac{\infty}{\infty}$, es necesario aplicar la regla de L'Hopital, la que conduce a: $\lim_{x_2 \rightarrow \infty} \left[H(x_1, x_2) = \frac{(x_1+1)(e^{x_2}-1)}{x_1+2e^{x_2}-1} \right] = \lim_{x_2 \rightarrow \infty} \left[\frac{(x_1+1)e^{x_2}}{2e^{x_2}} \right] = \frac{(x_1+1)}{2} = F_{x_1}$.

Todo lo anterior conduce a probar el tercer lemma el cual es:

Lemma 3: Sean $S_1, S_2 \neq \emptyset, S_1, S_2 \subset \bar{\mathbb{R}}$, subconjuntos no vacíos de la recta real extendida, y H una función \mathcal{L} -creciente con dominio en el rectángulo $S_1 \times S_2$, con “pisos” y “márgenes” como se especificaron anteriormente. Además sean (x_{1a}, x_{2a}) y (x_{1b}, x_{2b}) cualquier par de puntos en el rectángulo $S_1 \times S_2$, entonces:

$$|H(x_{1b}, x_{2b}) - H(x_{1a}, x_{2a})| \leq |F(x_{1b}) - F(x_{1a})| + |G(x_{2b}) - G(x_{2a})|. \quad (AI.2)$$

La prueba de este lemma está dada por la desigualdad del triángulo. Suponga que $x_{1a} \leq x_{1b}$ y que $x_{2a} \leq x_{2b}$. Ya que la función H es \mathcal{L} creciente, tiene un “piso” y un “techo”, los lemmas 1 y 2 sirven de cotas y conducen a $0 \leq |H(x_{1b}, x_{2b}) - H(x_{1a}, x_{2a})| \leq |F(x_{1b}) - F(x_{1a})|$.

Comprobándolo para cualquier par de valores en la marginal cuyo dominio es S_1 , se tiene que $\forall x_{1a}, x_{1b} \in S_1$ entonces $|H(x_{1b}, x_{2b}) - H(x_{1a}, x_{2a})| \leq |F(x_{1b}) - F(x_{1a})|$. Para

comprobar la otra marginal, basta con establecer la simetría entre los subconjuntos, lo que lleva a $\forall x_{2a}, x_{2b} \in S_2$ entonces $|H(x_{1b}, x_{2b}) - H(x_{1a}, x_{2a})| \leq |G(x_{2b}) - G(x_{2a})|$.

Una vez realizadas estas demostraciones, las cuales aplican para funciones acumuladas de probabilidad bivariadas, continuas por la derecha, se pueden establecer que:

$$(a) \lim_{x_i \rightarrow -\infty} H(x_1, x_2) = 0,$$

$$(b) \lim_{x_i \rightarrow +\infty \forall i} H(x_1, x_2) = 1,$$

(c) por la desigualdad del rectángulo $\forall x_{1a}, x_{1b}$ y x_{2a}, x_{2b} donde $x_{1a} \leq x_{1b}$ y $x_{2a} \leq x_{2b}$ se tiene que: $0 \leq |H(x_{1b}, x_{2b}) - H(x_{1a}, x_{2a})| \leq |F(x_{1b}) - F(x_{1a})|$. Si la función es dos veces derivable, $H \in C^2$, la propiedad *2-creciente* es equivalente a $\frac{\partial^2 H}{\partial x_1 \partial x_2} \geq 0$.

Definición de cópula: Con estas definiciones, lemmas y propiedades en mano, se está en condiciones de establecer lo que es una subcópula.¹ Se define como subcópula de 2 dimensiones a una función, C' , que cumple con las siguientes características:

- 1 . El dominio de la subcópula es el subplano sobre $\bar{I}R^2$ formado por el producto cruz $S_1 \times S_2$ donde $S_i = [0, 1]$. Esto es, $\text{Dom}C' = S_1 \times S_2$.
- 2 . C' tiene un piso y es una función *2 creciente*.
- 3 . $\forall u \in S_1$ y $\forall v \in S_2$: $C'(u, 1) = u$ y $C'(1, v) = v$ lo que implica toda u, v dentro del dominio de la subcópula estará en el intervalo $[0, 1]$ y por lo tanto su rango será un subconjunto del intervalo $[0, 1]$, esto es: $\forall (u, v) \in \text{Dom}C', 0 \leq C'(u, v) \leq 1$, por lo que $\text{Ran}C' \subset I$.

Una vez definida la subcópula, se definirá a la cópula como un caso especial de esta, siendo una cópula de dos dimensiones una subcópula cuyo dominio está en I^2 y su rango está en I , $C : I^2 \rightarrow I$, que además cumple con:

- 1 . Para todo u, v dentro de I , $\forall u, v \in I$, la cópula tiene un valor de cero si alguno de sus argumentos es cero (esto ocurre por la presencia de los pisos), $C(u, 0) = 0 = C(0, v)$,

¹ Toda esta demostración acerca de la naturaleza de las cópulas está basada en la presentación del seminario de Cópulas de Natalia Baeva, quien a su vez lo tomó de: Nelsen R.B, An Introduction to Copulas, Lecture Notes in Statistics, 1999, pp. 5-23

y el valor de la cópula es el de su marginal si el otro valor es uno (por los márgenes),
 $C(u, 1) = u, C(1, v) = v$.

2 . Para todo u_1, u_2, v_1, v_2 dentro de I , tal que $u_1 \leq u_2$ y $v_1 \leq v_2$ se cumple la desigualdad del rectángulo (por el lema 3). Esto es $\forall u_1, u_2, v_1, v_2 \in I$ tal que $u_1 \leq u_2$ y $v_1 \leq v_2$ se tiene que $C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0$.

Cotas de Frechet-Hoeffding para cópulas: Ya que se ha definido lo que es una cópula, se establecerán cotas para estas estableciendo el siguiente teorema: Sea C' una subcópula, entonces para toda dupla (u, v) en el dominio de la subcópula, $\forall (u, v) \in \text{Dom}C'$, se tiene que

$$W(u, v) = \max(u + v - 1, 0) \leq C'(u, v) \leq \min(u, v) = M(u, v). \quad (AI.3)$$

Esto se prueba suponiendo que (u, v) es un punto en el dominio de la subcópula C' , lo que, junto a los pisos y márgenes previamente establecidos, lleva a pensar que la subcópula será menor que cada una de sus marginales. Esto es $C'(u, v) \leq C'(u, 1) = u$ y que $C'(u, v) \leq C'(1, v) = v$ por lo que en consecuencia la cota inferior está dada por la más pequeña de las marginales, $\Rightarrow C'(u, v) \leq \min(u, v)$.

Por otro lado, el volúmen de la subcópula es positivo, lo que implica que la subcópula es mayor o igual a $u + v - 1$, por lo que la cota superior está dada por $\max(u + v - 1, 0)$, esto es: $V_{C'}([u, 1] \times [v, 1]) \geq 0$ implica $C'(u, v) \geq (u + v - 1) \Rightarrow C'(u, v) \geq \max(u + v - 1, 0)$. Se hace notar que ambas cotas son a su vez subcópulas. Hecho que mas tarde se usará para generar cópulas a travez del método de la mezcla.

Algunas propiedades de las cópulas: Ahora se detallarán algunos puntos sobre la continuidad, dependencia y formas de los límites de las cópulas. Sea C' una subcópula. Entonces para todo $u_1, u_2, v_1, v_2 \in \text{Dom}C'$ se tiene que $|C'(u_2, v_2) - C'(u_1, v_1)| \leq |u_2 - u_1| + |v_2 - v_1|$ por lo que se puede concluir que la subcópula C' es uniformemente continua en su dominio pues todos estos puntos se encuentran en el subplano I^2 el cual se definió continuo.

Siguiendo sobre la misma línea se sabe que el rango de la subcópula C' está sobre I , que esta es una función *2-creciente* además de ser uniformemente continua en el dominio con marginales no decrecientes (establecido anteriormente en los lemmas) por lo que las secciones horizontales, verticales y diagonales de la subcópula son no decrecientes y uniformemente continuas en I .

A partir de estas aseveraciones, se puede decir que la gráfica de una cópula es una superficie continua dentro del cubo I^3 cuyos límites están dados por:

$$(0, 0, 0), (1, 0, 0), (1, 1, 1) \text{ y } (0, 1, 0),$$

además de estar dentro de las superficies $M(u, v)$ y $W(u, v)$ establecidas por las cotas de Frechet-Hoeffding.

AI.2 Teorema de Sklar: Dentro de la literatura de cópulas, este teorema es la piedra angular sobre la cual se levanta toda la teoría. En este momento sólo se hará un esbozo del trabajo original de 1959. El teorema inicia definiendo a una función de distribución como una función F con dominio sobre la recta real extendida, $\text{Dom}F = \bar{\mathbb{R}}$, que además es no decreciente y cumple con los siguientes límites: $F(-1) = 0$ y $F(+1) = 1$.

También se define una función de distribución bivariada conjunta como una función H con dominio en $\bar{\mathbb{R}}^2$ que además es *2-creciente* y cumple con los siguientes límites: $H(x_1, -1) = H(-1, x_2) = 0$, y $H(+1, +1) = 1$. Además de tener marginales F y G dadas por $F(x_1) = H(x_1, 1)$ y $G(x_2) = H(1, x_2)$.

Sklar establece el siguiente teorema: Sea H una función conjunta de distribución con marginales F y G , entonces existe una cópula tal que $\forall x_1, x_2 \in \mathbb{R}$ se cumple que

$$H(x_1, x_2) = C(F(x_1), G(x_2)). \quad (\text{AI.4})$$

Si F y G son continuas entonces la cópula C es única, si las marginales son discontinuas la cópula no es única y sólo será únicamente determinada en el producto cruz de los rangos de las marginales, $\text{Ran}F \times \text{Ran}G$. Para probar este teorema se requieren de los dos siguientes lemas.

Lemma 4: Sea H una función de distribución conjunta con distribuciones marginales F y G , entonces existe una subcópula C' única, tal que el dominio de la subcópula sea el rango de las marginales, $\text{Dom}C' = \text{Ran}F \times \text{Ran}G$, además de que para toda dupla (x_1, x_2) en la recta real extendida, la función de distribución conjunta sea igual a la cópula, *i.e.* $\forall (x_1, x_2) \in \bar{\mathbb{R}}, H(x_1, x_2) = C'(F(x_1), G(x_2))$.

La prueba de este lemma se basa en el lemma 3 y suponer que $S_1 = S_2 = \bar{\mathbb{R}}$. Esto lleva a $|H(x_{1b}, x_{2b}) - H(x_{1a}, x_{2a})| \leq |F(x_{1b}) - F(x_{1a})| + |G(x_{2b}) - G(x_{2a})|$, pero al suponer $S_1 = S_2$ se tiene que $F(x_{1a}) = F(x_{1b})$ y $G(x_{2a}) = G(x_{2b})$, entonces $H(x_{1a}, x_{2a}) = H(x_{1b}, x_{2b})$.

El conjunto de pares ordenados $\{((F(x_1), G(x_2)), H(x_1, x_2)) | x_1, x_2 \in \mathbb{R}\}$ define una función real de dos variables C' (una subcópula) con dominio $\text{Dom}C' = \text{Ran}F \times \text{Ran}G$.

Lemma 5: Sea C' una subcópula que cumple con los requisitos de cópula, por lo que $C(u, v) = C'(u, v) \forall (u, v) \in \text{Dom}C'$. Esto es que cualquier subcópula puede ser extendida a una cópula, aunque la extensión es en generalmente no única. Para probarlo se inicia por el dominio de la subcópula, $\text{Dom}C' = S_1 \times S_2$, para después extender la función C'' al plano real extendido, con lo que $\text{Dom}C'' = \bar{S}_1 \times \bar{S}_2$ entonces C'' es una subcópula cuyo dominio puede ser llevado de nuevo a $[0, 1]^2$.

Para hacer esto es necesario recordar que las secciones horizontales, verticales y diagonales de la cópula son no decrecientes y continuas, además de que la extensión de subcópulas a cópulas no es única. Para ello, tomando en cuenta las cotas de la cópula, se define:

$$\lambda_1 = \begin{cases} \frac{a-a_1}{a_2-a_1} & \text{si } a_1 < a_2 \\ 1 & \text{si } a_1 = a_2 \end{cases},$$

$$\mu_1 = \begin{cases} \frac{b-b_1}{b_2-b_1} & \text{si } b_1 < b_2 \\ 1 & \text{si } b_1 = b_2 \end{cases}.$$

Ahora por interpolación lineal (usando como base el lemma 3) se tiene que la cópula puede ser expresada como la combinación lineal de la subcópula C'' valuada en distintos puntos, esto es: $C(a, b) = (1 - \lambda_1)(1 - \mu_1)C''(a_1, b_1) + (1 - \lambda_1)\mu_1C''(a_1, b_2) + \lambda_1(1 - \mu_1)C''(a_2, b_1) + \lambda_1\mu_1C''(a_2, b_2)$. El volúmen de la cópula sobre el plano antes especificado,

B , no es otra cosa sino: $V_C(B) = V_C([a, c] \times [b, d]) = (\lambda_2 - \lambda_1)(\mu_2 - \mu_1)V_C([a_1, a_2] \times [b_1, b_2])$, de donde se puede inferir que el volumen de la cópula es positivo, $V_C(B) \geq 0$, ya que $c \geq a$ y $d \geq b$. Esto implica que $\lambda_2 \geq \lambda_1$ y que $\mu_2 \geq \mu_1$. Revizando el volúmen para la cópula sobre el área B se tiene que:

$$\begin{aligned} V_C(B) = & (1 - \lambda_1)\mu_2 V_C([a_1, a_2] \times [d_1, d_2]) + \mu_2 V_C([a_1, c_1] \times [d_1, d_2]) + \\ & \lambda_2 \mu_2 V_C([c_1, c_2] \times [d_1, d_2]) + (1 - \lambda_1) V_C([a_1, a_2] \times [b_2, d_1]) + \\ & V_C([a_2, c_1] \times [b_2, d_1]) + \lambda_2 V_C([c_1, c_2] \times [b_2, d_1]) + \\ & (1 - \lambda_1)(1 - \mu_1) V_C([a_1, a_2] \times [b_1, b_2]) + (1 - \mu_1) V_C([a_2, c_1] \times [b_1, b_2]) + \\ & \lambda_2 (1 - \mu_1) V_C([c_1, c_2] \times [b_1, b_2]) \geq 0. \end{aligned}$$

Ya que se ha establecido que la cópula es una función siempre positiva y que puede ser formada a través de la combinación lineal de subcópulas, se procede a relacionarlas con las funciones bivariadas de distribución a través del siguiente teorema.

Teorema: Sean H una función de distribución conjunta con distribuciones de densidad de probabilidad marginales F y G . Entonces existe una cópula C tal que $\forall x, y \in \bar{I}\bar{R}$ se tiene que:

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

Si las distribuciones marginales F y G son continuas, entonces la cópula C es única, de otro modo está determinada únicamente en el dominio dado por $\text{Ran}(F) \times \text{Ran}(G)$. Siguiendo sobre la misma idea, si C es una cópula y F, G son funciones de distribución, entonces H es una función de distribución conjunta con marginales F y G .

Esto lleva al siguiente paso teórico que es relacionar a la cópula (función acumulada de distribución de probabilidad) con dos variables aleatorias. Para esto, Sklar establece el siguiente teorema:

Teorema: Sean X y Y variables aleatorias con funciones marginales de distribución F y G respectivamente, además de una función de distribución conjunta H .

Entonces existe una cópula C tal que:

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

De nuevo se establece que si F y G son continuas, entonces la cópula, C , es única. De otro modo, sólo está definida en el dominio dado por $\text{Ran}(F) \times \text{Ran}(G)$. Note que la cópula, C , de este teorema se llama “Cópula de X y Y ”, esto es C_{XY} .

Teorema: A continuación se discute sobre la unicidad de las cópulas. Sean X y Y variables aleatorias continuas con una cópula C_{XY} . Si α es una función estrictamente monótona cuyo dominio es $\text{Ran}(X)$ y β una función estrictamente monótona cuyo dominio es $\text{Ran}(Y)$, y ambas funciones son estrictamente crecientes en sus respectivos rangos, entonces la cópula C_{XY} es invariante ante transformaciones estrictamente crecientes de X y Y . Esto es $C_{\alpha(X)\beta(Y)} = C_{XY}$.

Para probar este teorema sean F_1, G_1, F_2 y G_2 las funciones de distribución de las variables aleatorias $X, Y, \alpha(X)$ y $\alpha(Y)$ respectivamente. Dado que α y β son funciones estrictamente crecientes, se tiene que $F_2(x) = P[\alpha(X) \leq x] = P[X \leq \alpha^{-1}(x)] = F_1(\alpha^{-1}(x))$, y $G_2(y) = G_1(\beta^{-1}(y))$.

Entonces, para toda realización x, y que se encuentre sobre los reales extendidos, $\forall x, y \in \bar{\mathbb{R}}$, se tiene que:

$$\begin{aligned} C_{\alpha(X)\beta(Y)}(F_2(x), G_2(y)) &= P[\alpha(X) \leq x, \beta(Y) \leq y] = P[X \leq \alpha^{-1}(x), Y \leq \beta^{-1}(y)] \\ &= C_{XY}(F_1(\alpha^{-1}(x)), G_1(\beta^{-1}(y))) \\ &= C_{XY}(F_2(x), G_2(y)). \end{aligned} \tag{AI.5}$$

Dado que X, Y son continuas, se tiene que $\text{Ran}(F_2) = \text{Ran}(G_2) = I$ entonces $C_{\alpha(X)\beta(Y)} = C_{XY}$ sobre I^2 .

1.3.3 Algunas propiedades notables y definiciones: Después de esta fundamentación técnica es posible referirse libremente al teorema de Sklar y a sus consecuencias. A continuación se darán algunas aclaraciones sobre este teorema y se mostrarán algunas de las más conocidas familias de cópulas.

Cómo se apuntó en la sección anterior, se usa una cópula cuando se tienen variables aleatorias dependientes, cuya dependencia va más allá de las medidas lineales como la correlación², de las cuales se tiene suficiente información, o se hacen supuestos sobre las distribuciones marginales de estas variables aleatorias³, pero se carece de información acerca de la distribución conjunta de estas.

Las cópulas permiten integrar la información sobre las marginales en una distribución conjunta después de conocer su interdependencia (más allá de la linealidad) a través de un parámetro, θ , o un vector de estos (en caso de cubrir a más de dos marginales). La gran ventaja de las cópulas es que las distribuciones marginales pueden venir de distintas familias de funciones de distribución, teniendo una única cópula que las agrupa una vez que se han especificado las marginales y la familia de cópula a usar. La elección de las funciones marginales y bivariadas es uno de los principales problemas al usar cópulas.

Antes de comenzar a enumerar algunas de las más conocidas familias de cópulas, se hará mención de algunos conceptos relacionados. Se conoce como *familia comprensiva de cópulas* a aquella familia que incluye las cotas de FrchetHoeffding (superior, C_U , e inferior, C_L) y a la *cópula producto*, C^\perp , la cual implica independencia de las variables aleatorias, esto es: $C(x_1, x_2) = F_1(x_1)F_2(x_2)$.

Por otro lado, se encuentra la perfecta dependencia positiva ó *comonotonicidad*⁴ cuando para cualquier par de realizaciones de la distribución conjunta, estas suben ó bajan al mismo tiempo, no necesariamente con la misma intensidad. Lo que puede escribirse como: $\forall(x_{1a}, x_{2a}), (x_{1b}, x_{2b})$, se tiene que $\{x_{1a} \leq x_{2a}, x_{1b} \leq x_{2b}\}$ ó $\{x_{1a} \geq x_{2a}, x_{1b} \geq x_{2b}\}$.

² Un ejemplo es el de variables aleatorias no correlacionadas pero dependientes. Suponga que se genera una corrida pseudo aleatoria de 100 números, a los cuales se les eleva al cuadrado. Si se toma la correlación entre estos, se obtendrá una $\rho=0$ aunque ambas series son claramente dependientes, sólo que su dependencia va más allá de una medida lineal.

³ Tal y como lo apunta Taleb en "The a priori problem of observed Probabilities", en la realidad el investigador se enfrenta a problemas que van más allá del control que se tiene sobre una urna ó un juego de azar. Lo que se hace en el mejor de los casos es recolectar información (muestreo) y realizar inferencias sobre esa muestra, suponiendo que convergerá a la distribución real y que existen los momentos muestrales necesarios para la estimación. La solución usual es suponer de entrada una distribución tratable para los datos ó truncar algunos de ellos. En realidad, se supuestos desde el momento de la inferencia, y estos supuestos la afectan a su vez.

⁴ *Comonotonicity*, en inglés.

En sentido contrario, se dice que un conjunto es contramonotónico ⁵ si se tiene que $\forall (x_{1a}, x_{2a}), (x_{1b}, x_{2b})$ se observa que $\{x_{1a} \leq x_{2a}, x_{1b} \geq x_{2b}\}$ ó $\{x_{1a} \leq x_{2a}, x_{1b} \geq x_{2b}\}$.

Haciendo uso de las definiciones anteriores, se puede definir a X_1 como una “función creciente” de X_2 si y sólo si su cópula es igual al mínimo de ambas marginales, *i.e.* la cota superior de FrchetHoeffding. Esto es: *sii* $C(\cdot) = C_u(\cdot)$. Lo que corresponde a comonotonicidad y perfecta correlación positiva.

Como se puede imaginar, existe la contraparte para la correspondencia contramonotónica y perfecta correlación negativa. En este caso, X_1 es llamada “función decreciente” de X_2 , *sii* $C(\cdot) = C_L(\cdot)$. Es decir si la cópula es igual a $\max \left[\sum_{j=1}^m F_j - m + 1, 0 \right]$.

Otra cualidad importante y llamativa de las cópulas es la *propiedad de invarianza*, la cual es demostrada en el párrafo correspondiente a la ecuación (1.3.5). La bondad de esta propiedad es ejemplificada cuando, por ejemplo, se conoce la distribución marginal de los activos y se conoce (o construye) la cópula correspondiente. Para trabajar con los logaritmos de estos, se usa la misma cópula, esto es: $C(x_1, x_2) = C(\ln(x_1), \ln(x_2))$.

También se hará mención del concepto de “cópula asociada”. Si $(U_1, U_2) \sim C$, entonces hay cópulas asociadas con los pares bivariados uniformes $(1 - U_1, 1 - U_2)$, $(U_1, 1 - U_2)$, $(1 - U_1, U_2)$, de estos el primer par corresponde a la cópula de supervivencia. Para considerarlas se inicia con los cuantiles de las distribuciones acumuladas marginales y sus respectivas funciones de cuantiles marginales de supervivencia, $F_1^{-1}(u_1) \sim F_1 \Rightarrow F_1^{-1}(1 - u_1) \sim \bar{F}_1$ y $F_2^{-1}(u_2) \sim F_2 \Rightarrow F_2^{-1}(1 - u_2) \sim \bar{F}_2$. Entonces para la cópula se tiene que $(1 - U_1, 1 - U_2) \sim \bar{C}$. En general,

$$\begin{aligned} \bar{F}(u) &= \bar{F}(F_1^{-1}(1 - u_1), \dots, F_m^{-1}(1 - u_m)) = \bar{F}(\bar{F}_1^{-1}(u_1), \dots, \bar{F}_m^{-1}(u_m)) \\ &= \bar{C}(u_1, \dots, u_m). \end{aligned} \tag{AI.6}$$

Para ejemplificar el manejo de cópulas, cuantiles y las funciones de distribución de probabilidad suponga que se desea escribir de forma extendida la probabilidad conjunta de supervivencia a dos tiempos de paro $(T_1 > t_1, T_2 > t_2)$. Esto es $S(t_1, t_2) = \Pr[T_1 > t_1, T_2 >$

⁵ *Countermonotonicity*, en inglés.

$t_2\} = (1 - F(t_1)) - F(t_2) + F(t_1, t_2) = S_1(t_1) - (1 - S_2(t_2)) + \bar{C}(1 - S_1(t_1), 1 - S_2(t_2)) = S_1(t_1) + S_2(t_2) - 1 + \bar{C}(1 - S_1(t_1), 1 - S_2(t_2))$ lo que en notación de cópula puede ser escrito como $\bar{C}(u_1, u_2) = u_1 + u_2 - 1 + C(1 - u_1, 1 - u_2) = \Pr[U_1 > u_1, U_2 > u_2]$.

Una vez que se han revizado las propiedades básicas que comparten todas las cópulas, así como el aparato técnico que dió lugar a estas funciones de distribución conjunta, se pueden revisar algunas de las familias más comunes de cópulas, sus límites y propiedades.

La siguiente sección no pretende ser exhaustiva en su examen, pues en la literatura se encuentran una gran cantidad, sino comentar las más usadas en aplicaciones y analizar cual será usada para el modelo.

1.3.4 Algunas familias de cópulas: Aunque el primer paso en la implementación de cópulas es la selección de las funciones que se ajustarán a las marginales, ese paso se dejará de lado hasta el momento en que se realice la aproximación de esta tesis. Por el momento se iniciará con la más sencilla de las cópulas, a saber: la cópula producto.

Cópula producto: Esta, la más simple de todas las cópulas, que refleja independencia entre las marginales y sirve como punto de comparación, tiene la forma:

$$C^\perp = C(u_1, u_2) = u_1 u_2. \quad (AI.7)$$

Cópula Farlie & Gumbel & Morgenstern (FGM): Propuesta por primera vez por Morgenstern en 1956, no es otra cosa que una generalización de la cópula producto cuando se introduce un parámetro de dependencia, θ . Empíricamente se ha visto que es útil cuando la dependencia entre las dos marginales es de magnitudes pequeñas⁶. La cópula FGM toma la forma:

$$C(u_1, u_2, \theta) = u_1 u_2 (1 + \theta(1 - u_1)(1 - u_2)). \quad (AI.8)$$

Por inspección se puede saber que cuando $\theta = 0$ que representa independencia, se llega a la cópula producto C^\perp .

⁶ Prieger, J. (2002), A flexible parametric selection model for non-normal data with application to health care usage. *Journal of Applied Econometrics* 17, 367392.

Cópula Gaussiana: Esta es una de las familias de cópulas más usadas en la literatura de riesgos por ser una extensión natural al supuesto de normalidad Gaussiana. La cópula parte de una distribución acumulada normal bivariada con un parámetro de dependencia $\theta \in [-1, 1]$ que “corta” parte del área central de la función, lo que ocasiona que la masa de probabilidad restante deba acumularse en las colas, generando con ello colas más anchas que la normal bivariada⁷.

Como se mencionó renglones arriba, el parámetro de dependencia esta confinado a la región entre -1 y 1. Estos valores corresponden a las cotas de Frchet (inferior y superior respectivamente). Como lo indican los posibles valores del parámetro de dependencia, la cópula normal admite dependencia positiva y negativa, aunque es incapaz de generar valores de colas pesadas. Empíricamente se ha visto que es confiable para representar valores cercanos a la media, tomando la siguiente forma:

$$\begin{aligned} C(u_1, u_2, \theta) &= \Phi_G(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \theta) \\ &= \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \left[\frac{1}{2\pi(1-\theta^2)^{1/2}} \right] \left[-\frac{(s^2 - 2\theta st + t^2)}{2(1-\theta^2)} \right] ds dt. \end{aligned} \quad (AI.9)$$

Cópula “t” de Student: Esta familia de cópulas es otra comunmente socorrida en la literatura financiera. Esta cópula tiene la peculiaridad de tener dos parámetros de dependencia θ_1 y θ_2 , Siendo θ_1 el parámetro que controla la anchura de las colas. Esta cópula es representada mediante:

$$C(u_1, u_2, \theta_1, \theta_2) = \int_{-\infty}^{t_{\theta_1}^{-1}(u_1)} \int_{-\infty}^{t_{\theta_2}^{-1}(u_2)} \left[\frac{1}{2\pi(1-\theta_2^2)^{1/2}} \right] \left[1 + \frac{(s^2 - 2\theta_2 st + t^2)}{\nu(1-\theta_2^2)} \right]^{-\theta_1+2/2} ds dt. \quad (AI.10)$$

Cómo características especiales de esta distribución, se puede remarcar que si $\theta_1 \rightarrow \infty$, la Cópula t tiende a la cópula Gaussiana, *i.e.* $\theta_1 \rightarrow \infty \Rightarrow C^t(u_1, u_2, \theta_1, \theta_2) \rightarrow \Phi_G(u_1, u_2, \theta)$

⁷ Existe una diferencia fundamental entre “colas pesadas” y “colas anchas”. Las primeras se refieren a funciones de distribución de las cuales es imposible determinar alguno de sus momentos muestrales (tiende a $\pm\infty$), *i.e.* $E[x^n]=\infty$ ó $E[x^n]=-\infty$, mientras que las colas anchas implican un exceso de kurtosis respecto a la normal (3).

Cópula de Clayton: Esta cópula, originalmente estudiada por Kimeldorf y Sampson en 1975⁸ es representada mediante:

$$C(u_1, u_2, \theta) = (u_1^{-1} + u_2^{-1} - 1)^{-1/\theta}. \quad (AI.11)$$

En este caso, el parámetro de dependencia está restringido a los reales positivos, $\theta \in (0, \infty)$, lo que implica que conforme el parámetro de dependencia se acerca a infinito, la cópula tiende a la cota superior de Frchet, mientras que si el parámetro de dependencia tiende a cero, la cópula se acerca a la cópula producto. Este punto será demostrado a continuación:

$$\begin{aligned} \lim_{\theta \rightarrow 0} [u_1^{-\theta} + u_2^{-\theta} - 1]^{-1/\theta} &= \lim_{\theta \rightarrow 0} \left[\frac{1}{\left(\frac{1}{u_1^\theta} + \frac{1}{u_2^\theta} - 1 \right)^{1/\theta}} \right] \\ &= \lim_{\theta \rightarrow 0} \left[\frac{u_1 u_2}{(u_1^\theta + u_2^\theta - (u_1^\theta u_2^\theta))^{1/\theta}} \right] = y. \end{aligned}$$

Para obtener el límite, se realiza una transformación lineal (logarítmica), la cual, tal y como se demostró anteriormente, no afecta a la cópula. Posteriormente se usan las propiedades de los límites para introducir el logaritmo dentro del límite y separar el numerador del resto del límite (no depende de θ , por lo que puede salir), para después se reacomodar y

⁸ Kimeldorf, G. and A. R. Sampson (1975), Uniform representations of bivariate distributions. *Communications in Statistics* 4, 617627.

usar la regla de L'Hopital, esto es:

$$\begin{aligned}
\ln(y) &= \ln \left(\lim_{\theta \rightarrow 0} \left[\frac{u_1 u_2}{(u_1^\theta + u_2^\theta - (u_1^\theta u_1^\theta))^{1/\theta}} \right] \right) \\
&= \lim_{\theta \rightarrow 0} \left[\ln \left(\frac{u_1 u_2}{(u_1^\theta + u_2^\theta - (u_1^\theta u_1^\theta))^{1/\theta}} \right) \right] \\
&= \ln(u_1 u_2) + \lim_{\theta \rightarrow 0} \left[\ln \left(\frac{1}{(u_1^\theta + u_2^\theta - (u_1^\theta u_1^\theta))^{1/\theta}} \right) \right] \\
&= \ln(u_1 u_2) + \lim_{\theta \rightarrow 0} \left[\frac{\ln((u_1^\theta + u_2^\theta - (u_1^\theta u_1^\theta)))}{-\theta} \right] \\
&= \ln(u_1 u_2) + \lim_{\theta \rightarrow 0} \left[\frac{\ln((+1 - 1 + u_1^\theta + u_2^\theta - (u_1^\theta u_1^\theta)))}{-\theta} \right] \\
&= \ln(u_1 u_2) + \lim_{\theta \rightarrow 0} \left[\frac{\ln((1 + (1 - u_1^\theta)(u_2^\theta - 1)))}{-\theta} \right],
\end{aligned}$$

aplicando la regla de L'Hopital se tiene que:

$$\ln(y) = \ln(u_1 u_2) + \lim \left[\frac{\theta u_2^{\theta-1} (1 - u_1^\theta) - \theta u_1^{\theta-1} (u_2^\theta - 1)}{(1 + (u_2^\theta - 1)(1 - u_1^\theta)) \frac{1}{1}} \right] = \ln(u_1 u_2).$$

Deshaciendo la transformación logarítmica al exponenciar, se llega al límite deseado: $y = u_1 u_2$. Que no es otra cosa que la cópula producto, C^\perp . Como puede verse por el dominio del parámetro de dependencia, esta cópula no puede lidiar con dependencias negativas, aunque presenta fuerte dependencia en la cola izquierda, mientras que la dependencia es débil del lado derecho.

Por estas propiedades, es que esta cópula puede ser usada para modelar la dependencia del riesgo incumplimiento con la tasa libre de riesgo (si la tasa aumenta, aumentan los incumplimientos), aunque no resulta útil para la dependencia del parámetro de prepago con la tasa libre, pues su relación es contramonotónica. Más tarde se regresará a esta cópula cuando sea integrada al modelo.

Cópula de Frank: Esta cópula, presentada en 1979, puede suponer parámetros de dependencia a lo largo de toda la recta real, $\theta \in \mathbb{R}$, correspondiendo los límites superior

e inferior a las cotas de Frchet- Hoeffding, mientras que el límite en cero corresponde nuevamente a la cópula producto. La forma de esta cópula es:

$$C(u_1 u_2, \theta) = -\theta^{-1} \ln \left[1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{(e^{-\theta} - 1)} \right]. \quad (AI.12)$$

Para mostrar el límite, primero se llevará la ecuación a una forma tal que sea posible usar el teorema de L'Hopital, lo que implica:

$$\begin{aligned} y &= \lim_{\theta \rightarrow 0} \left[-\theta^{-1} \ln \left[1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{(e^{-\theta} - 1)} \right] \right] \\ &= \lim_{\theta \rightarrow 0} \left[-\frac{1}{\theta} \ln \left[\frac{e^{-\theta} - 1 + e^{-\theta u_1 - \theta u_2} - e^{-\theta u_1} - e^{-\theta u_2} + 1}{(e^{-\theta} - 1)} \right] \right] \\ &= \lim_{\theta \rightarrow 0} \left[-\frac{1}{\theta} \ln \left[\frac{e^{\theta(u_1+u_2)} + e^{\theta} - e^{\theta(u_2+1)} - e^{\theta(u_1+1)}}{e^{\theta(u_1+u_2+1)} - 1} \right] \right] \\ &= \lim_{\theta \rightarrow 0} \left[-\frac{1}{\theta} \ln \left[\frac{e^{\theta(u_1+u_2+1)} + e^{2\theta} - e^{\theta(u_2+2)} - e^{\theta(u_1+2)}}{e^{\theta(u_1+u_2+1)} - e^{\theta(u_1+u_2+2)}} \right] \right] \\ &= \lim_{\theta \rightarrow 0} \left[-\frac{1}{\theta} \ln \left[\frac{e^{\theta(u_1+u_2+1)} + e^{2\theta} - e^{\theta(u_2+2)} - e^{\theta(u_1+2)}}{e^{\theta(u_1+u_2+1)} - e^{\theta(u_1+u_2+2)}} \right] \right] \\ &= \lim_{\theta \rightarrow 0} \left[\frac{\ln \left(\left[e^{\theta(u_1+u_2+1)} + e^{2\theta} - e^{\theta(u_2+2)} - e^{\theta(u_1+2)} \right]^{-1} \left[e^{\theta(u_1+u_2+1)} - e^{\theta(u_1+u_2+2)} \right] \right)}{\theta} \right]. \end{aligned}$$

Aplicando la regla dos veces, dado que en la primera sigue teniendo una forma indeterminada, se llega a:

$$\begin{aligned} y &= \lim_{\theta \rightarrow 0} \left[\frac{(u_1 + u_2 + 1) e^{\theta(u_1+u_2+1)} + 2e^{2\theta} - (u_2 + 2) e^{\theta(u_2+2)} - (u_1 + 2) e^{\theta(u_1+2)}}{e^{\theta(u_1+u_2+1)} + e^{2\theta} - e^{\theta(u_2+2)} - e^{\theta(u_1+2)}} \right] + \\ &\quad \lim_{\theta \rightarrow 0} \left[\frac{(u_1 + u_2 + 1) e^{\theta(u_1+u_2+1)} - (u_1 + u_2 + 2) e^{\theta(u_1+u_2+2)}}{e^{\theta(u_1+u_2+1)} - e^{\theta(u_1+u_2+2)}} \right] = \\ &\quad \lim_{\theta \rightarrow 0} \left[\frac{(u_1 + u_2 + 1)^2 e^{\theta(u_1+u_2+1)} + 4e^{2\theta} - (u_2 + 2)^2 e^{\theta(u_2+2)} - (u_1 + 2)^2 e^{\theta(u_1+2)}}{(u_1 + u_2 + 1) e^{\theta(u_1+u_2+1)} + 2e^{2\theta} - (u_2 + 2) e^{\theta(u_2+2)} - (u_1 + 2) e^{\theta(u_1+2)}} \right] + \\ &\quad \lim_{\theta \rightarrow 0} \left[\frac{(u_1 + u_2 + 1)^2 e^{\theta(u_1+u_2+1)} - (u_1 + u_2 + 2)^2 e^{\theta(u_1+u_2+2)}}{(u_1 + u_2 + 1) e^{\theta(u_1+u_2+1)} - (u_1 + u_2 + 2) e^{\theta(u_1+u_2+2)}} \right], \end{aligned}$$

aplicando los límites queda una fracción, la cual después de ser realizada nos conduce a $y = u_1 u_2$ que de nuevo es la cópula producto. Esta cópula, empíricamente, ha mostrado

ser útil para conjuntos de datos que no presentan colas pesadas pues su dependencia es débil y simétrica para las colas. Incluso ha mostrado tener una dependencia menor en las colas que la cópula Normal. Por esta razón es que este tipo de cópula no será usada en el trabajo.

Cópula Gumbel: Esta última familia de cópulas fue tratada por primera vez en 1960 y toma la forma:

$$C(u_1, u_2, \theta) = e^{-((-\ln(u_1))^\theta + (-\ln(u_2))^\theta)^{1/\theta}}. \quad (AI.13)$$

Para ésta cópula los valores de 1 e ∞ corresponden a la independencia y a la cota superior de Fréchet - Hoeffding. Esta cópula tampoco admite dependencias negativas pero muestra una fuerte dependencia en la cola derecha, siendo menor en la cola izquierda.

Apéndice II

Métodos para la generación de cópulas.

Método de inversión: Este es el método más rudimentario para generar cópulas, y su principal desventaja consiste en que es necesario conocer de antemano la distribución conjunta. Este método surge de forma directa del teorema de Sklar.

Dadas dos distribuciones marginales continuas, F_1 y F_2 , y una distribución conjunta continua, $F(x_1, x_2) = C(F_1(x_1), F_2(x_2))$. La cópula correspondiente es generada usando la función de cuantiles de las marginales¹ donde $x_1 = F_1^{-1}(u_1)$ y $x_2 = F_2^{-1}(u_2)$ y u_i son distribuciones marginales. Esto es

$$C(u_1, u_2) = F(F_1^{-1}(u_1), F_2^{-1}(u_2)). \quad (AII.1)$$

Un primer ejemplo² se puede encontrar dentro de la familia Arquimedean. Se parte de una función acumulada bivariada de la forma $F(x_1, x_2) = e^{-[e^{-x_1} + e^{-x_2} - (e^{-\theta x_1} + e^{-\theta x_2})^{-1/\theta}]}$

¹ El mismo procedimiento se sigue si se desea conocer la cópula asociada de supervivencia.

² Dado en Joe, H. (1993), "Parametric families of multivariate distributions with given margins". *Journal of Multivariate Analysis* 46, 262282.

la cual tiene dominio en el área real dada por $-\infty < x_1, x_2 < \infty$ con un parámetro de dependencia $\theta \geq 0$.

Para conocer sus distribuciones marginales, se procede de la misma forma que en cualquier otra función acumulada bivariada, esto es, se toma el límite cuando alguno de sus argumentos tiende a infinito. Lo que conduce a $F(x_1) = \lim_{x_2 \rightarrow \infty} [F(x_1, x_2)] = e^{-e^{-x_1}} = u_1$ y a $F(x_2) = \lim_{x_1 \rightarrow \infty} [F(x_1, x_2)] = e^{-e^{-x_2}} = u_2$.

Una vez conocidas las distribuciones acumuladas marginales, se procede a obtener la función de cuantiles respectiva, *i.e.* invertir la función, lo que lleva a: $x_1 = -\ln(-\ln(u_1))$ y $x_2 = -\ln(-\ln(u_2))$ respectivamente. Ahora sólo es cuestión de sustituir las funciones de cuantiles recién obtenidas como argumentos en la función acumulada bivariada. Esto lleva a:

$$\begin{aligned} C(x_1, x_2) &= e^{-\left[e^{-(-\ln(-\ln(u_1)))} + e^{-(-\ln(-\ln(u_2)))} - (e^{-\theta(-\ln(-\ln(u_1)))} + e^{-\theta(-\ln(-\ln(u_2)))})^{-1/\theta} \right]} \\ &= e^{\left[\ln(u_1) + \ln(u_2) + ((-\ln(u_1))^{-\theta} + (-\ln(u_2))^{-\theta})^{-1/\theta} \right]} \\ &= u_1 u_2 e^{((-\ln(u_1))^{-\theta} + (-\ln(u_2))^{-\theta})^{-1/\theta}}. \end{aligned} \tag{AIII.2}$$

Como puede imaginar, existen muchas cópulas que pueden formarse utilizando este método, tantas como funciones acumuladas de probabilidad multivariadas pueda imaginar. Con finalidad de hacer claro este método, usado en el desarrollo de la tesis, se realizará un segundo ejemplo. Se iniciará con la función acumulada de probabilidad bivariada $(1 + e^{-x_1} + e^{-x_2})^{-1}$.

Siguiendo el procedimiento anterior, se obtienen las funciones acumuladas marginales, a saber: $F(x_1) = (1 + e^{-x_1})^{-1}$ y $F(x_2) = (1 + e^{-x_2})^{-1}$. Ahora se invierten estas funciones, para obtener la función de cuantiles, lo que lleva a $x_1 = -\ln\left(\frac{u_1}{1-u_1}\right)$ y a $x_2 = -\ln\left(\frac{u_2}{1-u_2}\right)$. Para finalizar con el ejercicio, sólo se sustituyen estas funciones de

cuantiles en la función acumulada bivariada, lo que conduce a:

$$\begin{aligned}
 C &= \left(1 + e^{-\ln\left(\frac{u_1}{1-u_1}\right)} + e^{-\ln\left(\frac{u_2}{1-u_2}\right)} \right)^{-1} \\
 &= \left(\frac{u_1 u_2 + (1-u_1)u_2 + (1-u_2)u_1}{u_1 u_2} \right)^{-1} \\
 &= \frac{u_1 u_2}{u_1 + u_2 - u_1 u_2}.
 \end{aligned} \tag{AIII.3}$$

Tal y como se mencionó al principio del apéndice, este método requiere que el investigador conozca, o suponga, de antemano la forma de la distribución conjunta, lo que no siempre es posible. Por lo que el uso de este sencillo método es relativamente escaso. Para paliar este problema existen otros métodos para la formación de cópulas entre los cuales destacan el algebraico y el de mezcla, los cuales serán explicados a continuación.

Método algebraico: Este método también resulta muy simple, nace de la relación entre dos distribuciones marginales las cuales en un principio se suponen como independientes. A continuación se introduce un parámetro de dependencia θ y con ello se obtiene la cópula. Ejemplos conocidos de cópulas obtenidas usando este método son las cópulas de Plackett y la de Ali-Mikhail-Haq.

A continuación se ejemplifica este método replicando la obtención de la cópula de Ali-Mikhail-Haq, para ello se parte de las funciones marginales $F_1(x_1) = (1 + e^{-x_1})^{-1}$ y $F_2(x_2) = (1 + e^{-x_2})^{-1}$. El método requiere de buscar la función de distribución bivariada suponiendo independencia lo que lleva a pensar que $F_{12}(x_1, x_2) = F_1(x_1)F_2(x_2)$.

Recordando que las familias de cópulas incluyen a cópulas asociadas, *v.gr.* la cópula de supervivencia, se buscará la cópula de supervivencia, es decir $\frac{1-F(x_1, x_2)}{F(x_1, x_2)}$. Si se sustituye la función acumulada bivariada en esta se llega a:

$$\begin{aligned}
 \frac{1 - F(x_1, x_2)}{F(x_1, x_2)} &= \frac{1 - (1 + e^{-x_1})^{-1} (1 + e^{-x_2})^{-1}}{(1 + e^{-x_1})^{-1} (1 + e^{-x_2})^{-1}} = \frac{(1 + e^{-x_1})^{-1} (1 + e^{-x_2})^{-1} - 1}{(1 + e^{-x_1})^{-1} (1 + e^{-x_2})^{-1}} \\
 &= e^{-x_1} + e^{-x_2} + e^{-x_1 - x_2}.
 \end{aligned} \tag{AIII.4}$$

Si la última igualdad es expresada en términos de las distribuciones de supervivencia de las distribuciones originales, *i.e.* $\frac{1-F_i(x_i)}{F_i(x_i)} = \frac{1-(1+e^{-x_i})^{-1}}{(1+e^{-x_i})^{-1}} = (1 + e^{-x_i}) - 1 = e^{-x_i}$, se puede

expresar la ecuación AII.4 como:

$$\frac{1 - F(x_1, x_2)}{F(x_1, x_2)} = \frac{1 - F_1(x_1)}{F_1(x_1)} + \frac{1 - F_2(x_2)}{F_2(x_2)} + \frac{1 - F_1(x_1) - F_2(x_2)}{F_1(x_1) F_2(x_2)}.$$

Ahora basta con hacer el álgebra correspondiente en esa suma para llegar a la cópula. Por simplicidad se supone que $u_i = F_i(x_i)$, todo esto lleva a:

$$\begin{aligned} \frac{1 - C(u_1, u_2, \theta)}{C(u_1, u_2, \theta)} &= \frac{1 - u_1}{u_1} + \frac{1 - u_2}{u_2} + \frac{1 - u_1 - u_2}{u_1 u_2} \\ &= \frac{u_2 - u_1 u_2 + u_1 - u_1 u_2}{u_1 u_2} + (1 - \theta) \frac{1 - u_2 - u_1 + u_1 u_2}{u_1 u_2}. \end{aligned}$$

Reacomodando para despejar $C(u_1, u_2, \theta)$, se obtiene $C(u_1, u_2, \theta) = \frac{u_1 u_2}{1 - \theta((1 - u_1)(1 - u_2))}$. Finalmente se procede a describir el último método mencionado para formación de cópulas, el método de mezclado y sumas convexas.

Método de mezcla y sumas convexas: Durante la reseña hecha en el primer apéndice (AI) acerca de los límites de las cópulas, se habló de las cotas de Fréchet-Hoeffding y de la cópula producto, C_L, C_U, C^\perp respectivamente. Del mismo modo se mencionó el hecho de que estas cotas también eran cópulas bajo circunstancias especiales del parámetro de dependencia, θ . También se discutió el dominio de las subcópulas y de las cópulas como espacios en los reales extendidos, $\bar{\mathbb{R}}$, y de sus rangos sobre el intervalo $[0, 1]$.

Dados estos antecedentes, es que se puede hablar de sumas convexas entre cópulas, que mapean desde dominios similares y que por tanto pueden ser agregadas, lo que implica la existencia de cópulas que son producto de hacer sumas convexas entre funciones acumuladas multivariadas que ya han demostrado ser cópulas (Nelsen, 2006). Lo que implica que una suma convexa hecha mediante un ponderador $\pi \in [0, 1]$ también es una cópula, lo que se puede escribir como:

$$C^M = \pi_1 C_L + (1 - \pi_1) C_U. \quad (AII.5)$$

Esta misma idea puede ser usada para encontrar cópulas que son la esperanza de una colección infinita de cópulas indexadas por una variable continua η con una función de

distribución $\Lambda_\theta(\eta)$. Esto es:

$$C_\theta(u_1, u_2) = E_\eta[C_\eta(u_1, u_2)] = \int_{R(\eta)} C_\eta(u_1, u_2) d\Lambda_\theta(h). \quad (AIII.6)$$

De forma similar Marshall y Olkin, en su trabajo de 1988, consideraron la mezcla $H(x) = \int_{R(\eta)} [F(x)]^\eta d\Lambda(h)$, donde el índice $\eta > 0$. Ellos mostraron que para cualquier par de cópulas y distribuciones de cópulas, $H(x), \Lambda(\eta)$, existe una cópula $F(x)$ que la convierte en función de densidad. Aunque este tipo de ejemplos no serán usados durante el desarrollo de esta tesis.

Cópulas arquimedianas: Aunque no es un método de generación de cópulas, las cópulas arquimedianas son una clase especial de cópulas que se caracterizan por surgir de funciones pseudo invertibles. En general, son cópulas capaces de capturar amplios rangos de dependencia y de derivación más o menos sencilla.

Formalmente, se dice que una cópula es Arquimedea cuando existe una clase φ de funciones convexas decrecientes³ que mapean del intervalo $[0, 1]$ al intervalo $[0, \infty]$ con derivadas continuas en $(0, 1)$. Todas estas condiciones aseguran que exista la inversa, φ^{-1} , o pseudo inversa, $\varphi^{[-1]}$, de esta función, . Esto depende de que $\varphi(0) = \infty$ ó no⁴.

Si una función φ cumple con estas características entonces es capaz de generar funciones de distribución bivariadas y se le conoce como *función generadora*.⁵ Ejemplos de estas funciones son: $\varphi(x) = -\ln(x)$, $\varphi(x) = (1-x)^\theta$ y $\varphi(x) = x-\theta, \theta > 1$.

Siguiendo esta idea, Junker & May(2005)⁶ probaron que si la función φ es un generador, esta puede ser compuesta en otra función, $f \circ g$, que también será un generador, a condición de que la función tenga dominio y rango dentro del intervalo $[0, 1]$,

³ Esto significa que $\varphi(1)=0$ y $\varphi'(x)<0$, y convexas, i.e. $\varphi''(t)>0 \forall t \in (0,1)$.

⁴ La definición formal es : $\varphi^{[-1]}(x) = \begin{cases} \varphi^{-1}(x) & 0 \leq x \leq \varphi(0) \\ 0 & \varphi(x) \leq x \leq +\infty \end{cases}$.

⁵ Este concepto no tiene ninguna relación con el de función generadora de momentos.

⁶ Junker, M. and A. May (2005), Measurement of aggregate risk with copulas. *Econometrics Journal* 8, 428–454.

$g : [0, 1] \rightarrow [0, 1]$, y que sea una función cóncava estrictamente creciente, que además compla con $g(1) = 1$.

Para ejemplificar lo postulado por Junker & May, suponga que la función $f : [0, \infty] \rightarrow [0, \infty]$ es una función convexa estrictamente creciente que además cumple con $F(0) = 0$. Si esta función es compuesta con otra función que es un generador, entonces la nueva función, $f \circ g$, también lo es. Como parte de su trabajo, estos investigadores hacen notar que funciones que ya se usan en la literatura cumplen con estos requisitos. Por ejemplo

$$\begin{aligned} g(x) &= x^\nu & \nu &\in (0, 1), \\ g(x) &= \frac{\ln(ax+1)}{\ln(a+1)} & a &\in (0, 1), \\ g(x) &= \frac{e^{-\theta x} - 1}{e^{-\theta} - 1} & \theta &\in (-\infty, \infty), \\ f(\varphi) &= \varphi^\delta & \delta &\in (1, \infty), \\ f(\varphi) &= a^\varphi - 1 & a &\in (1, \infty), \\ f(\varphi) &= a^{-\varphi} - 1 & a &\in (0, 1), \end{aligned}$$

En general, las cópulas de la familia Arquimedeanas tienen el parámetro de dependencia dentro de la forma funcional φ y en general tienen la forma:

$$C(u_1, u_2, \theta) = \varphi^{-1}(\varphi(u_1) + \varphi(u_2)). \quad (AII.7)$$

Por último, se hace mención de la propiedad de simetría de esta clase de cópulas, *i.e.* $C(u_1, u_2) = C(u_2, u_1)$, y asociatividad, *i.e.* $C(C(u_1, u_2), u_3) = C(u_1, C(u_2, u_3))$. Estas dos propiedades son producto de la unicidad de las cópulas establecida en el teorema de Sklar.

Es importante señalar que independencia implica intercambiabilidad, pero la intercambiabilidad no implica la independencia de dos variables aleatorias (Nelsen 1995).

Apéndice III

Teoría de valores extremos

Introducción: Sería difícil leer este trabajo sin enmarcarlo en la crisis financiera de 2007 - 2008 causada originalmente por el mercado inmobiliario estadounidense. Los grandes auges o las grandes caídas en el valor de los instrumentos financieros no son eventos comunes, en realidad, se presentan con mayor frecuencia y severidad que la esperada en un mundo Gaussiano.

Antes de iniciar con la revisión de la teoría de valores extremos, es necesario conocer la razón primordial por la que a pesar de ser un número grande de créditos los involucrados en un bono hipotecario, estos no tienden a distribuirse como normales. Para ello, se analizará brevemente el teorema del límite central.

Este teorema es un caso particular de una transformación estabilizadora para que una variable aleatoria tienda a una función de distribución de probabilidad bien definida, *i.e.* $p \lim \hat{\theta}_n = \theta$, lo que implica que $z_n = \sqrt{n} (\hat{\theta}_n - \theta) \xrightarrow{d} f(z)$, donde $f(z)$ es una distribución bien definida con media y varianza positiva. Si un estimador cumple con esta propiedad, es llamado *n - raíz consistente*.

En particular, la versión de Lindberg - Feller del teorema del límite central (TLC) establece que: “La suma de variables aleatorias, sin importar su distribución original, tenderán a distribuirse normal”. Esto es, si se supone que la sucesión de variables aleatorias $\{x_i\}$, $i = 1 \cdots n$ es una secuencia de variables aleatorias con media μ_i y varianzas finitas σ_i^2 , donde las medias y varianzas muestrales están dadas por $\bar{\mu}_n = \frac{1}{n} (\mu_1 + \mu_2 + \cdots + \mu_n)$ y $\bar{\sigma}_n = \frac{1}{n} (\sigma_1 + \sigma_2 + \cdots + \sigma_n)$, además de que **ningún término domine la varianza**, *i.e.* $\lim_{n \rightarrow \infty} \frac{\max(\sigma_i)}{n\bar{\sigma}_n} = 0$ y que la varianza converja a una constante finita, *i.e.* $\bar{\sigma}^2 = \lim_{n \rightarrow \infty} \bar{\sigma}_n^2$, entonces las variables tenderán a distribuirse como normales con media μ y varianza finita, *i.e.* $\sqrt{n} (\bar{x}_n - \mu_n) \xrightarrow{d} N [0, \bar{\sigma}^2]$

El teorema no implica que las variables aleatorias vengan de la misma distribución, sólo pide que ninguna de las varianzas domine la muestra. Si se analiza esta última afirmación,

el lector encontrará que la velocidad de convergencia a la distribución normal dependerá de la velocidad con la que la máxima varianza dividida entre la varianza promedio tienda a cero. Mas aún, se puede inferir que dada la naturaleza de los valores extremos, siempre existirán valores que dominen la muestra y que entre mayor sea el parámetro de forma, γ , existirán valores que dominen más la muestra (más aberrantes), incumpliendo con ello los supuestos del Teorema del Límite Central.

Una vez establecida la razón por la cual no se puede cumplir con el TLC, se analizarán otros teoremas, también respecto a transformaciones estabilizadoras, que muestran la unicidad de la distribución límite para el máximo de un conjunto de variables aleatorias, así como las funciones de probabilidad hacia las cuales convergen. Estos teoremas son la distribución del máximo y el teorema del máximo de Fisher - Tippet respectivamente.

Distribución del máximo de un conjunto de variables aleato-

rias: Sean Y_1, Y_2, \dots, Y_N variables aleatorias independientes y no negativas con función de distribución común F , con $F(0) = 0$. Además, sea N una variable aleatoria Poisson con $\lambda > 0$, entonces:

$$\begin{aligned}
 P \left\{ \max_{0 \leq i \leq N} \{Y_i\} \leq x \right\} &= \sum_{i=0}^N P \left\{ \max_{0 \leq i \leq N} \{Y_i\} \leq x, N = n \right\} \\
 &= \sum_{i=0}^N P \left\{ \max_{0 \leq i \leq N} \{Y_i\} \leq y \right\} P \{N = n\} \\
 &= \sum_{i=0}^N P \{Y_0 \leq y, Y_1 \leq y, \dots, Y_n \leq y\} P \{N = n\} \\
 &= \sum_{i=0}^N P \{Y_0 \leq y\} P \{Y_1 \leq y\} \cdots P \{Y_n \leq y\} P \{N = n\} \\
 &= \sum_{i=0}^N [P \{Y \leq y\}]^n P \{N = n\} \\
 &= \sum_{i=0}^N F^n(y) P \{N = n\}.
 \end{aligned}$$

Que no es otra cosa mas que, tomar la probabilidad conjunta del máximo y de un número n de eventos, para después descomponerla suponiendo independencia entre los mismos y

hacer uso de que al tratarse x de un máximo, será mayor o igual que todas las variables aleatorias de la muestra.

Habiendo realizado este ejercicio mental, el lector debe hacer uso del hecho de que las variables tiene una distribución común, lo que permite factorizar todas sus funciones de distribución y demostrar con ello que la distribución del máximo converge a $F^n(y)$.

Estabilidad de las distribuciones y el Teorema del máximo (Fisher - Tippett): Una vez que se ha demostrado que existe una distribución para el máximo de una sucesión de variables aleatorias, es necesario mostrar la estabilidad de la distribución límite, para ello se hace uso del siguiente teorema de uso común en la literatura de valores extremos ¹.

Sean X_1, X_2, \dots, X_n variables aleatorias independientes e idénticamente distribuidas con una distribución F , entonces el máximo de estas variables, $\max_{1 \leq i \leq n} \{X_i\}$, tiene una distribución F^n . Se dice que una distribución F es estable si existen sucesiones de números reales $(a_n)_{n \in N}$ y $(b_n)_{n \in N}$ con $a_n > 0$ para toda $n \in N$ tales que:

$$P \left\{ \max_{1 \leq i \leq n} \{x_i\} \leq b_n + a_n x \right\} = P \left\{ \frac{\max_{1 \leq i \leq n} \{x_i\} - b_n}{a_n} \leq x \right\} \rightarrow F(x). \quad (AIII.1)$$

si y sólo si $n \rightarrow \infty$. En este caso, las variables aleatorias convergerán en probabilidad a alguna función de probabilidad de valores extremos. Por ejemplo, si se eligen $b_n = \ln n$ y

¹ Este apéndice está tomado en gran parte de Reiss, R.D. & Thomas, M.; *Statistical Analysis of Extreme Values with applications to Insurance, Finance, Hidrology and other fields*; Birkhuser Verlag, 2 Ed; Berlin; 2001. y de Venegas - Martinez, F.; *Riesgos Financieros y económicos, productos derivados y decisiones económicas bajo incertidumbre*; Cengage, 2 Ed, 2008.

$a_n = 1$, se obtiene la distribución Gumbell, a saber:

$$\begin{aligned}
 G_0^n(b_n + a_n x) &= G_0^n(\ln n + x) \\
 &= \left(e^{-e^{-(\ln n + x)}} \right)^n \\
 &= \left(e^{-e^{-\ln n} e^{-x}} \right)^n \\
 &= \left(e^{-\frac{1}{n} e^{-x}} \right)^n \\
 &= e^{-e^{-x}} \\
 &= G_0(x).
 \end{aligned}$$

En el caso de la distribución Frechet, se eligen $b_n = 0$ y $a_n = 1/n$, lo que llevará al lector a:

$$\begin{aligned}
 G_1^n(b_n + a_n x) &= G_1^n\left(0 + \frac{1}{n}x\right) \\
 &= \left(e^{-\left(\frac{1}{n}x\right)^{-\alpha}} \right)^n \\
 &= e^{-n\left(\frac{1}{n}x\right)^{-\alpha}} \\
 &= e^{-x^{-\alpha}}.
 \end{aligned} \tag{AIII.3}$$

Para el caso de la distribución Weibull, basta con tomar un $a_n = -1/n$ y seguir el proceso anterior.

Una vez establecida la estabilidad de las funciones de valores extremos, es posible recordar el teorema de Fisher - Tippet, el cual, después de definir un valor máximo estandarizado, $Y_n = \frac{M_n - \mu_n}{\sigma_n}$, dice que si el valor estandarizado converge a una función de distribución de probabilidad no degenerada, esta debe de ser una función de valores extremos, que en su parametrización gamma tienen la forma:

$$e^{-(1+\gamma y)^{\frac{-1}{\gamma}}}, \tag{AIII.4}$$

si $\gamma \neq 0, 1 + \gamma y > 0$ y

$$e^{-e^{-y}}, \tag{AIII.5}$$

si $\gamma = 0, -\infty \leq y \leq \infty$. Es importante recordar que $\gamma = 1/\alpha$. Habiendo tratado el tema de la estabilidad, es prudente recordar la condición de von Mises. Si una distribución de probabilidad cumple con ella, el máximo de un grupo de variables aleatorias será atraído al dominio de las funciones de valores extremos. La condición a cumplir es:

$$\lim_{x \rightarrow \infty} \frac{xf(x)}{\bar{F}(x)} = \alpha >, \quad (AIII.6)$$

donde $\bar{F}(x)$ es la distribución complementaria $1 - F(x)$. En el caso de la distribución Frechet, se tiene que: $\bar{F}(x) = 1 - e^{-x^{-\alpha}}$ y $f(x) = \alpha e^{-x^{-\alpha}} x^{-(1+\alpha)}$. Verificando la validez de la condición de Von Mises para esta distribución, se puede observar que:

$$\lim_{x \rightarrow \infty} \frac{xf(x)}{\bar{F}(x)} = \lim_{x \rightarrow \infty} \frac{x\alpha e^{-x^{-\alpha}} x^{-(1+\alpha)}}{1 - e^{-x^{-\alpha}}} = \frac{\alpha e^{-x^{-\alpha}} x^{-\alpha}}{1 - e^{-x^{-\alpha}}}$$

Este límite puede ser encontrado haciendo uso del teorema de L'Hopital, lo que conduce a:

$$\begin{aligned} \lim_{x \rightarrow \infty} \frac{xf(x)}{\bar{F}(x)} &= \lim_{x \rightarrow \infty} \frac{\alpha e^{-x^{-\alpha}} x^{-\alpha}}{1 - e^{-x^{-\alpha}}} \\ &= \lim_{x \rightarrow \infty} \frac{\alpha \left(e^{-x^{-\alpha}} (-\alpha) x^{-(1+\alpha)} + x^{-\alpha} e^{-x^{-\alpha}} \alpha x^{-(1+\alpha)} \right)}{e^{-x^{-\alpha}} (-\alpha) x^{-(1+\alpha)}} \\ &= \lim_{x \rightarrow \infty} \frac{-\alpha^2 + \alpha x^{-\alpha}}{-\alpha} \\ &= \alpha. \end{aligned}$$

Dada la similitud, haciendo uso de la transformación γ , se puede inferir que todas las distribuciones de valores extremos cumplen con la condición de Von Mises y por lo tanto se trata de distribuciones estables.

Finalmente, se desea recordar al lector que el j -ésimo momento, $E[X^j]$, de una función de valores extremos Frechét o Weibull puede ser escrito en términos de la función gamma, para ello se recuerda al lector que la función gamma está dada por:

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx,$$

lo que conlleva que $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$ y $\Gamma(1) = 1$.

Por lo tanto, es posible reescribir el j -ésimo momento de estas funciones de valores extremos como:

$$m_{j,G} = E[X^j] = \int x^j dG(x) = \int x^j g(x) dx. = \Gamma\left(1 - \frac{j}{\alpha}\right), \quad (AIII.7)$$

es importante que el lector recuerde esta característica de las funciones de valores extremos, pues será usada con frecuencia a lo largo del trabajo.

Apéndice IV

Máxima verosimilitud en dos etapas (2SML).

El objetivo de este apéndice es dar una explicación puntual de la metodología de Máxima Verosimilitud en dos etapas (2SML), así como de la consistencia y convergencia de los estimadores a una distribución normal, lo que hará posible la aplicación de pruebas de hipótesis sobre los parámetros y la estimación en general.

El tema de la convergencia será abordado siguiendo la línea de investigación planteada por László, G; Gyrgy, M & Zoltán (2002) y Xiaohong Chen & Yanqin Fan (2004) adaptando la nomenclatura de variables a la usada en este trabajo.

El análisis inicia suponiendo que se tiene un vector $\hat{\theta}$ de estimadores de los parámetros reales de las marginales, θ , y que el parámetro de dependencia $\alpha = (\alpha_n)$ consiste en una secuencia de variables aleatorias independientes e idénticamente distribuidas *i.i.d.* con media cero y varianza finita que siguen una distribución dependiente de los datos y los parámetros reales de las marginales, $f(x, \theta^*)$, donde el vector de parámetros θ^* es un dominio abierto en $F \subseteq R^d$. Se supone también que $\ln f$ es tres veces diferenciable y Lipschitz continuo. Estas condiciones aseguran que la solución sea localmente estable.

El problema atacado es la estimación de una función para los errores, e , de la predicción obtenida de la segunda estimación (la de α), que a su vez dependen del vector de parámetros

$\theta \in D_0$ y α , es decir, calcular un vector de parámetros de dependencia que a su vez dependen de una primer estimación realizada a través de ML de las marginales, $\hat{\theta}$, y un vector de dependencias, $\hat{\alpha}$, sea $\epsilon_n(\theta)$ el error estimado de predicción del proceso, el cual será el producto de la segunda estimación.

Suponga que se encuentra el vector de parámetros estimados, $\hat{\theta}$, y que estos son tomados como dados para la estimación de los errores a través del algoritmo de Máxima Verosimilitud usando el procedimiento tradicional por segunda ocasión, esto es:

$$V_N(\alpha) = - \sum_{n=1}^N \ln f(\alpha_n(\theta), \epsilon^*). \quad (AIV.1)$$

Esta segunda estimación arrojará un nuevo conjunto de parámetros asociados a la función de verosimilitud $f(\alpha_n(\theta), \epsilon^*)$. Dicho conjunto de parámetros requiere estar asociado a una función de distribución conocida de modo tal que sea posible realizar pruebas de hipótesis que los validen. Para llevar a cabo esta validación, es necesario conocer la matriz asintótica de covarianzas asociadas al primer proceso de verosimilitud, al cual se nombra como Σ_{ML} que no es otra cosa sino una matriz con errores al cuadrado, esto es:

$$\Sigma_{ML} = \lim_{N \rightarrow \infty} NE \left[\left(\hat{\theta}_N - \theta^* \right) \left(\hat{\theta}_N - \theta^* \right)^T \right]. \quad (AIV.2)$$

Suponiendo que se cumplen las condiciones de regularidad, se tiene que el error entre el parámetro real y el estimado, $\hat{\theta}_N - \theta^*$, está dado por el cambio de la verosimilitud cuando la muestra tiende a infinito más un infinitésimo de orden superior al primero, esto es:

$$\hat{\theta}_N - \theta^* = -^{-1}(\theta^*) \frac{1}{N} \sum_{n=1}^N \frac{\partial}{\partial \theta} \ln f(\epsilon_n(\theta), \alpha^*) \Big|_{\theta=\theta^*} + O_M(N^{-1}), \quad (A.IV.3)$$

donde $W(\theta) = \lim_{n \rightarrow \infty} E[-\ln f(\epsilon_n, \alpha^*)]$. A continuación se reescribe la matriz asintótica de covarianza como el límite de los cambios en las funciones de verosimilitud cuando cambian los parámetros que es:

$$\Sigma_{ML} = -^{-1}(\theta^*) P^* (-^T(\theta^*))^{-1}, \quad (AIV.4)$$

donde $P^* = \lim_{n \rightarrow \infty} E \left[\left(\frac{\partial}{\partial \theta} \ln f(\epsilon_n(\theta), \alpha^*) \right)^T \left(\frac{\partial}{\partial \theta} \ln f(\epsilon_n(\theta), \alpha^*) \right) \Big|_{\theta=\theta^*} \right]$ y $W_{\theta\theta}(\theta^*)$ es la matriz hessiana, la cual debe ser no singular.

Para la misma segunda estimación, se obtienen los mismos resultados si en vez maximizar la función de verosimilitud, se minimiza el error en la predicción (la base de MCO, básicamente se hace el problema dual), esto es: $V_N(\theta) = \sum_{n=1}^N \varepsilon_n^2(\theta)$. Para esta predicción en el error, se busca la matriz asintótica de covarianza, Σ_{PE} , que puede ser representada, haciendo uso del criterio de factorización de Fisher - Neyman¹, como $\Sigma_{PE} = (R^*)^{-1} (\sigma^*)^2$ con $R^* = \lim_{n \rightarrow \infty} E [\varepsilon_{\theta n}^T(\theta^*) \varepsilon_{\theta n}(\theta^*)]$.

Una vez expresado el problema recíproco, se procede a expresar la matriz de covarianzas de ML como una función del error de predicción, esto es:

$$\Sigma_{ML} = (R^*)^{-1} \mu^{-1}, \quad (AIV.5)$$

donde $\mu = \lim_{n \rightarrow \infty} E \left[\left(\frac{f'(e_n)}{f(e_n)} \right)^2 \right]$ es el parámetro de información de Fisher.² Este ejercicio se lleva a cabo dado que si dos conjuntos de datos tienen el mismo valor para un estadístico de Fisher, siempre llegarán a las mismas inferencias sobre el parámetro buscado, para mayor referencia, consúltese Stigler(1973)³. Lo que se pretende a través de este argumento es probar que se minimiza el error en las estimaciones al realizar el proceso de 2SML que busca maximizar la función de verosimilitud y que por lo tanto, los estimadores son consistentes y de mínima varianza.

Para probar la idea antes expresada, se define $h = \frac{f'}{f}$ y se hace uso de la P^* definida en (AIV.3), lo que conduce a

$$P^* = \lim_{m \rightarrow \infty} E [\varepsilon_{\theta m}^T(\theta^*) \varepsilon_{\theta m}(\theta^*)] E h^2(e_m) = R^* \mu, \quad (AIV.6)$$

para llevar a cabo esta sustitución, se hace uso de que el error de la primer estimación, que depende del vector de parámetros de las marginales, es una función de los errores de

¹ Se dice que el estadístico $T(X)$ es suficiente para el parámetro θ subyacente si la distribución condicional de probabilidad de la muestra, dado el estadístico $T(X)$ es independiente del parámetro θ , es decir $f_{\theta}(x) = h(x)g_{\theta}(T(x))$

² El parámetro de información de Fisher es una forma de medir el monto de información que una muestra de la variable aleatoria x posee sobre un parámetro θ asociado a la verosimilitud de la distribución de X , puede ser visto como una medida de la curvatura de la función de verosimilitud cuando cambian los parámetros

³ Stigler, Stephen (December 1973). "Studies in the History of Probability and Statistics. XXXII: Laplace, Fisher and the Discovery of the Concept of Sufficiency". *Biometrika* 60 (3): 439-445

la segunda estimación más un infinitésimo de algún orden, además de que los errores de ambas estimaciones son independientes entre si, esto es: $\varepsilon_n(\theta^*) = e_n + O_M(\alpha^n)$. De forma similar, la matriz hessiana $W_{\theta\theta}(\theta^*)$ es igual a:

$$(\theta^*) = - \lim_{n \rightarrow \infty} E \left[\begin{array}{c} f''(\cdot) \\ f(\cdot) \end{array} \varepsilon_{\theta,n}^T(\theta^*) \varepsilon_{\theta,n}(\theta^*) - \frac{f'(\cdot)}{f''(\cdot)} f'(\cdot) \varepsilon_{\theta,n}^T(\theta^*) \varepsilon_{\theta,n}(\theta^*) \right] - \lim_{n \rightarrow \infty} E \left[\begin{array}{c} f'(\cdot) \\ f(\cdot) \end{array} \varepsilon_{\theta\theta,n}(\theta^*) \right],$$

donde $f(\cdot) = f(\varepsilon_n(\theta^*))$. Dado que los errores de ambas estimaciones son independientes, el último término es igual a cero. Si se cumplen las condiciones de regularidad sobre la función f , se tiene que $Eh(e_n) = \int_{-\infty}^{\infty} \frac{f'(x)}{f(x)} f(x) dx = \int_{-\infty}^{\infty} f'(x) dx = 0$. Por lo que la matriz hessiana se reduce a:

$$\mathbf{W}_{\theta\theta}(\theta^*) = - \lim_{n \rightarrow \infty} E \left[\begin{array}{c} f''(\varepsilon_n(\theta^*)) \\ f(\varepsilon_n(\theta^*)) \end{array} \varepsilon_{\theta,n}^T(\theta^*) \varepsilon_{\theta,n}(\theta^*) \right] = R^* \int_{-\infty}^{\infty} \frac{f''(x)}{f(x)} f(x) dx = 0. \quad (AIV.7)$$

Por lo que se demuestra que la matriz Hessiana es la misma para ambas optimizaciones, esto es: $\mathbf{W}_{\theta\theta}(\theta^*) = Eh^2(e_n) R^* = R^* \mu$.

Ahora se verificará la validez de las estimaciones de los parámetros de dependencia de una forma similar a como se verificaron los primeros estimadores de las marginales. Inicie con un estimador para el parámetro de dependencia $\hat{\alpha}$ para el parámetro poblacional α^* , el cual satisface:

$$\hat{\eta}_N - \eta^* = O_M(N^{-1/2}), \quad (AIV.8)$$

se supone también que $\hat{\eta}_N \in F$, por lo que es posible definir un estimador parcialmente adaptativo $\bar{\theta}_N(\hat{\eta}_N)$ de θ^* como la solución de la segunda ecuación de verosimilitud, a saber:

$$V_N(\theta) = - \sum_{n=1}^N \frac{\partial}{\partial \theta} \ln f(\varepsilon_n(\theta), \hat{\eta}_N) = 0. \quad (AIV.9)$$

Para establecer la convergencia de esta segunda estimación, se establece en László, Gyorgy & Zoltán (2002) que:

$$\bar{\theta}_N(\hat{\alpha}_N) - \theta^* = -(R^*)^{-1} \mu^{-1} \frac{1}{N} \sum_{n=1}^N \frac{\partial}{\partial \theta} \ln f(\varepsilon_n(\theta), \alpha^*) \Big|_{\theta=\theta^*} + O_M(N^{-1}). \quad (AIV.10)$$

Para probar este teorema se hace un simil a la ecuación AIV.3, en donde se establece que la desviación en los parámetros corregidos está dada por el cambio en la función de verosimilitud por un cambio en los parámetros ponderado por la matriz Hessiana, esto es:

$$\bar{\theta}_N(\alpha_N) - \theta^* = - \mathbf{W}_{\theta\theta}^{-1}(\theta, \alpha) \Big|_{\theta=\theta^*} \frac{1}{N} \sum_{n=1}^N \frac{\partial}{\partial \theta} \ln f(\varepsilon_n(\theta), \alpha) \Big|_{\theta=\theta^*} + O_M(N^{-1}), \quad (\text{AIV.11})$$

de nuevo $W(v) = - \lim_{n \rightarrow \infty} E[\ln f(\varepsilon_n(\theta), \alpha)]$. Sea $F' \subseteq F_0$ un dominio compacto tal que el vector de parámetros α se encuentre en F' para cualquier conjunto de parámetros de dependencia, además de que la matriz Hessiana relacionada al proceso de verosimilitud este lejos de ser singular, es decir:

$$\sup_{\alpha \in F'} \|\mathbf{W}_{\theta\theta}^{-1}(\theta^*, \alpha)\| \leq K < \infty. \quad (\text{AIV.12})$$

La existencia de esta función F está justificada por ser la única solución para el Hessiano además de ser una matriz definida positiva que garantiza la continuidad de la función $\mathbf{W}_{\theta\theta}(\theta^*, \alpha)$ en α .

Dado que el proceso de cambio de la función de verosimilitud ante un cambio en los parámetros es L mixto,⁴ se introduce el cambio en el vector gradiente ante cambios en los parámetros de la primera estimación tal que: $\mathbf{g}_n(\theta, \alpha) = \frac{\partial}{\partial \theta} \ln(f(\varepsilon_n(\theta)), \alpha)$. Se hace notar al lector que el vector gradiente, $\mathbf{g}_n(\theta, \alpha)$, es un proceso L mixto uniforme en $\alpha \in F$, de lo que se sigue que:

$$\frac{1}{N} \sum_{n=1}^N \frac{\partial}{\partial \theta} \ln(f(\varepsilon_n(\theta)), \alpha) \Big|_{\theta=\theta^*} = O_M(N^{-1/2}), \quad (\text{AIV.13})$$

ya que para un par de procesos $(x_n(\theta))$ y $\frac{\Delta x}{\Delta \theta}$ que son L - mixtos uniformemente en una vecindad alrededor de θ que pertenece al espacio de parámetros de las marginales, $\theta, \theta + \mathbf{h} \in D$, se tiene que para el dominio compacto $D_0 \subseteq D$:

$$\sup_{\theta \in D_0} \left| \frac{1}{N} \sum_{n=1}^N x_n(\theta) \right| = O_M(N^{-1/2}), \quad (\text{AIV.14})$$

⁴ Sea (x_n) un proceso L mixto respecto a dos sigmas álgebra progresivamente medibles, $(\mathfrak{G}_n, \mathfrak{G}_n^+)$, y K acotadas, además sea $f(x)$ una función Lipschitz continua. Entonces el proceso $(y_n) = (f(x_n))$ es también L - mixto respecto a $(\mathfrak{G}_n, \mathfrak{G}_n^+)$.

por lo que se cumple la ecuación anterior, es decir, se cumple que el cambio en el vector gradiente sigue un infinitésimo de orden superior a un medio, por lo que se dice que converge.

Más aún, a partir de que la matriz hessiana está lejos de ser singular (ecuación AIV.12) y de que la diferencia entre los parámetros de dependencia estimados y poblacionales crece como un infinitésimo de orden superior (ecuación AIV.8), se tiene que el cambio en la matriz hessiana dado por un cambio en los parámetros también crece como un infinitésimo de orden superior, esto es:

$$\|\mathbf{W}_{\theta\theta}^{-1}(\theta^*, \hat{\alpha}_N) - \mathbf{W}_{\theta\theta}^{-1}(\theta^*, \alpha_N)\| = O_M(N^{-1/2}). \quad (\text{AIV.15})$$

A partir de las ecuaciones de variación en el vector gradiente y en el hessiano (AIV.13 y AIV.15 respectivamente) ocasionadas por ajustes en los parámetros de dependencia dados por la segunda estimación se puede ver que el comportamiento asintótico de el error de estimación $\bar{\theta}_N(\hat{\alpha}_N) - \theta^*$ es un error acotado por un infinitésimo de orden superior al primero $O_M(N^{-1})$ determinado por la ecuación AIV.11 que no es otra cosa sino el cambio en el vector gradiente ponderado por el hessiano más un infinitésimo de orden superior al primero, esto es:

$$\bar{\theta}_N(\hat{\alpha}_N) - \theta^* = -R^{-1}\mu^{-1} \frac{1}{N} \sum_{n=1}^N g_n(\theta^*, \hat{\alpha}_N) - O_M(N^{-1}). \quad (\text{AIV.16})$$

Para hacer manipulable la expresión, se hace una aproximación del vector gradiente por serie de Taylor alrededor del vector de parámetros de la segunda estimación, α^* , lo que conduce a:

$$g_n(\theta^*, \hat{\alpha}_N) = g_n(\theta^*, \hat{\alpha}) + \frac{\partial g_n}{\partial \alpha}(\theta^*, \alpha^*)(\hat{\alpha}_N - \alpha) + O_M(N^{-1}), \quad (\text{AIV.17})$$

que aun converge a un infinitésimo de orden superior al primero, si se recuerda que el proceso de cambio en el vector gradiente es un cociente de procesos *L - mixtos*, con lo cual se demuestra que la máxima verosimilitud en dos etapas es una metodología que muestra consistencia en los parámetros obtenidos en ambos momentos de la estimación.

Ahora se procederá a demostrar que los parámetros de estas estimaciones convergen en distribución a una distribución normal, lo que hará posible toda una batería de pruebas de hipótesis que validarán o rechazarán, siguiendo una metodología calara, el análisis econométrico. Para ello, se seguirá la línea de investigación planteada por Xiaohong Chen and Yanqin Fan (2004) de la universidad de Vanderbilt⁵ en la que se demuestra la convergencia de los parámetros de una estimación semiparamétrica de cópulas sin importar la distribución de las marginales.

La primer diferencia, afortunadamente simplificadora, con el trabajo de Chen y Fan (2004) es que sus estimaciones están hechas sobre datos que forman parte de una serie de tiempo con una estructura de dependencia compleja, mientras que los datos de este trabajo son agrupados para su análisis de manera tal que se forman varias cortes transversales (uno por periodo de análisis) analizados de forma independiente. Por lo que no es necesario pedir ergodicidad de ningún tipo a los datos del mismo.

Otra diferencia con el trabajo anteriormente citado es que, basado en la demostración de la irrelevancia de las marginales hecha a lo largo de la investigación de Chen y Fan, es posible suplir sus distribuciones empíricas o aproximaciones hechas con Kernels por las distribuciones marginales supuestas por el modelo sin hacer mella en las propiedades asintóticas de los estimadores. Todo esto a condición de que las funciones marginales sean procesos L - *Mixtos* con un decaimiento polinomial, condición que también fue establecida cuando se demostró la convergencia asintótica del modelo en el artículo de László, G; Gyrgy, M & Zoltán (2002).

De hecho, a grandes rasgos, sólo se establece una cota distribuida como normal para los estimadores de dependencia de la cópula, lo que garantiza que en muestras suficientemente grandes, los estimadores serán variables aleatorias distribuidas como normales.

El artículo busca expresar el valor esperado condicional de un proceso como función de la marginal del proceso y la cópula que lo une al valor del proceso en el periodo anterior,

⁵ Chen, X. & Fan, Y. Estimation of copula-based semiparametric time series models, Peer reviewed working paper No. 02-W26R, Department of Economics, Vanderbilt University, Nashville, TN 37235.

$E[y_t|y_{t-1}] = G * (.) C(y_{t-1}, \alpha^*)$. En el caso particular de este trabajo, se verá a una observación del corte transversal como función del resto de las observaciones, tomando como parámetro de dependencia óptimo como a aquel que maximiza la esperanza condicional a los datos de la muestra suponiendo que se distribuyen conjuntamente como la cópula que une a los parámetros independientes analizados.

Se supone también que la función de distribución marginal del parámetro de dependencia, $G^*(.)$, es la distribución marginal verdadera del parámetro de dependencia, la cual es absolutamente continua respecto a la medida de Lebesgue en la recta continua. así mismo, se le denominará $C(u_1, u_2; \alpha^*)$ a la cópula paramétrica verdadera que modela la relación entre los parámetros analizados. También se parte del supuesto de que el valor de dependencia α es absolutamente continuo respecto a la medida de Lebesgue en el rectángulo $[0, 1]^2$ y la cópula no converge a ninguna de las cotas de Fréchet-Hoeffding, pues de serlo casi seguramente (a.s.) se trataría de procesos determinísticos o por lo menos independientes.

En el artículo de Chen y Fan se hace una aproximación semiparamétrica (se usan kernels en lugar de funciones marginales paramétricas) como argumentos de la cópula y se llega a una densidad condicional que intenta modelar a una de las observaciones, y_t , tomando como argumento la observación pasada, y_{t-1} . Tratándose de un corte transversal, basta con suponer las observaciones, x_i , (el comportamiento de cada crédito) como independientes e idénticamente distribuidas, lo que deja la función de densidad de distribución condicional de cada observación como

$$h^*(x_i|X) = g^*(x_t) c(M^*(\theta_1), M^*(\theta_2); \alpha^*), \quad (AIV.18)$$

donde $h^*(x_i|X)$ es la distribución condicional verdadera de la i -ésima observación dadas en resto de las observaciones, $g^*(x_t)$ es la verdadera distribución marginal del parámetro de dependencia y $c(M^*(\theta_1), M^*(\theta_2); \alpha^*)$ es la cópula que rige la distribución conjunta de los parámetros de riesgo que dan lugar a la observación analizada. La principal diferencia entre el método semiparamétrico de Chen y Fan y el usado en esta investigación es la

determinación externa de la forma de las marginales, $M(\cdot)$, las cuales sólo se “pegan” en la cópula (sus estimaciones previamente realizadas). La validez de este procedimiento está respaldada por el concepto y formación de las cópulas (Explicado en AI) y la consistencia anteriormente demostrada.

Otra de las ventajas del uso de cópulas es su capacidad de separar la estructura de dependencia del parámetro de dependencia (lo que puede generar agrupamientos o *clusters*) del comportamiento de las marginales (que pueden presentar colas pesadas). De cumplirse con los supuestos de continuidad del parámetro y de la función de cópula en \mathbb{R} y \mathbb{R}^2 respectivamente además de que la cópula no alcance las cotas de Fréchet-Hoeffding(F-G), se puede transformar la distribución condicional de la i -ésima observación en la función de verosimilitud necesaria para realizar las estimaciones, a saber:

$$L(\alpha) = \frac{1}{n} \sum_{i=1}^n \ln g^*(x_i) + \frac{1}{n} \sum_{i=2}^n \ln c\left(\hat{M}(\lambda_1), \hat{M}(\lambda_2); \alpha\right), \quad (AIV.19)$$

donde $\hat{M}(\lambda_i)$ es la distribución marginal estimada del i -ésimo parámetro previamente estimada (el artículo original usa las funciones empíricas) que se “pega” a la cópula.

Esta transformación es posible si se supone que la cópula es positiva en el rectángulo $[0, 1]^2$, lo cual sucede *sii* es una función de densidad de probabilidad bivariada continua y bien definida. Esta condición permite que:

- i)* Si existen constantes $0 \leq \xi \leq 1$ y $0 \leq d \leq \infty$, una función similar a norma $\Lambda(\cdot) \geq 1$ y un conjunto \mathbf{K} tal que $\int_0^1 \Lambda(u) \times c(\lambda_1, \lambda_2, \alpha^*) du \leq \xi \Lambda(\lambda_i) + d1_{\mathbf{K}}(\lambda_i)$, entonces la sucesión de realizaciones $\{x_i\}$ es un proceso L - *mixto* con tasa de decaimiento exponencial $\beta_t \leq k \times e^{-an}$ para alguna $a > 0$. En este caso, $\Lambda(u)$ es una función creciente en el parámetro.
- ii)* Si existen constantes $\xi \in (0, 1]$, $0 < a$, $d < \infty$, una función similar a norma $\Lambda(\cdot) \geq 1$ y un conjunto \mathbf{K} tal que $\int_0^1 \Lambda(u) \times c(\lambda_1, \lambda_2, \alpha^*) du \leq \xi \Lambda(\lambda_i) - a [L(\lambda_i)]^\xi + d1_{\mathbf{K}}(\lambda_i)$, entonces la sucesión de realizaciones $\{x_i\}$ es un proceso L - *mixto* con tasa de decaimiento exponencial $\beta_t : \beta_t (1+n)^{\xi/1-\xi} \rightarrow 0$ conforme aumenta el tamaño de la muestra n .

Para probar estos dos postulados, Chen y Fan se basan en que la cópula sea positiva y no degenera en las distribuciones de las cotas de F-H, lo que garantiza que el proceso de serie de tiempo $\{y_t\}$ sea Markov, lo que lo hace geoméricamente ergódico. Para este trabajo, basta con suponer que las observaciones son independientes entre si, e idénticamente distribuidas. Esta condición hace que el proceso sea L-mixto con una tasa de decaimiento exponencial o polinomial, cumpliendo con los supuestos de los teoremas 5.2 de Down *et. al.* (1995) y 3.6 de Jarner y Roberts (2001) respectivamente.

Dado que la función de densidad del parámetro de dependencia, α , es continua, y por la definición de procesos L - mixtos, se puede concluir que todo el proceso (la cópula) es L - Mixto con alguna tasa de decaimiento sin importar sus distribuciones marginales, razón por la cual estas pueden simplemente “pegarse”.

Suponiendo que la función marginal de distribución del parámetro de dependencia está totalmente incorporada en la naturaleza de la cópula, esta puede ser ignorada, por lo que la función de verosimilitud de la distribución condicional del i-ésimo dato, tomando las marginales como dadas después de la primer estimación, está dado por la función de verosimilitud cuya única incógnita es el parámetro de dependencia, a saber:

$$\hat{\alpha} = \arg \max_{\alpha} \hat{L}(\alpha), \hat{L}(\alpha) = \frac{1}{n} \sum_{t=2}^n \ln c(\hat{M}(\lambda_1), \hat{M}(\lambda_2); \alpha). \quad (AIV.20)$$

Para determinar las propiedades de asintóticas del estimado del parámetro de dependencia, α , es necesario garantizar que la función de verosimilitud o sus derivadas, no exploten a infinito si las colas son demasiado pesadas. Para superar esta dificultad, Chen y Fan establecen la convergencia de la cópula bajo una métrica ponderada que es usada para establecer la convergencia asintótica del estimador a una normal.

Sea $W^*(\cdot)$ un proceso Gaussiano en $D[0, 1]$ tal que $W^*(0) = W^*(1) = 0$ y que

$$\begin{aligned} E\{W^*(v_1)W^*(v_2)\} &= \min\{v_1, v_2\} - v_1v_2 + \\ &\sum_{k=2}^{\infty} \{\text{cov}[I(U_1 \leq u_1), I(U_k \leq u_2)]\} + \\ &\sum_{k=2}^{\infty} \{\text{cov}[I(U_k \leq u_1), I(U_1 \leq u_2)]\}. \end{aligned} \quad (AIV.21)$$

Bajo estas definiciones Chen y Fan establecen el siguiente Lemma:

Suponga que el proceso de generación de datos $\{x_t\}$ es un proceso L-Mixto y que $w(\cdot)$ es una función continua sobre el intervalo $[0, 1]$ que es estrictamente positiva a lo largo del intervalo $(0, 1)$, que es simétrico en $u = 1/2$ y creciente en el intervalo $(0, 1/2)$, entonces:

- (1) Si la tasa de decaimiento, β_n , converge a un infinitésimo superior de algún tipo, $\beta_n = O(n^{-b})$, para alguna $b > 0$ y la medida $\int_0^1 \frac{1}{w(v)} \ln\left(1 + \frac{1}{w(v)}\right) dv < \infty$ es finita, se tiene que:

$$\sup_{u \in [0,1]} \left| \frac{\hat{U}_n(u) - u}{w(u)} \right| = o_{a.s.}(1), \sup_x \left| \frac{\hat{C}(x) - C^*(x)}{w(C^*(x))} \right| = o_{a.s.}(1), \quad (AIV.22)$$

que es algo similar a una demostración tipo “sandwich”, pues la cota gaussiana y la diferencia entre las funciones estimadas y real convergen. Por otro lado,

- (2) Si i) $\beta_n = O(n^{-b})$, para alguna $b > \gamma/(\gamma - 1)$ con $\gamma > 1$ y $\int_0^1 \left(\frac{1}{w(v)}\right)^{2\gamma} dv < \infty$ o ii) $\beta_n = O(n^{-b})$ para alguna $b > 1$ y $\int_0^1 \left(\frac{1}{w(v)}\right)^2 \ln\left(1 + \frac{1}{w(v)}\right) dv < \infty$, entonces:

$$\sqrt{n} \frac{(U_n(\cdot) - \cdot)}{w(\cdot)} \xrightarrow{dist} \frac{W^*(\cdot)}{w(\cdot)} \in D[0, 1], \sqrt{n} \sup_x \left| \frac{\hat{C}(x) - C^*(x)}{w(C^*(x))} \right| = O_p(1), \quad (AIV.23)$$

Lo que es la aplicación del mismo principio de convergencia anterior a las distintas tasas de decaimiento para los procesos *L-Mixtos*. Este resultado complementa el descrito por el trabajo de László, György & Zoltán(2002).

Para probar esta convergencia de las diferencias entre la cópula verdadera y la estimada a una distribución normal, se invita al lector a considerar la clase de funciones decrecientes, $\left\{ \frac{1}{w(v)} I(U_n \leq u) \right\} : u \in (0, 1/2]$, además de denotar al supremo del valor absoluto de la función decreciente, $F(U_n) \equiv \sup_{u \in (0, 1/2]} \left| \frac{1}{w(u)} I(U_n \leq u) \right|$, como la función envolvente. Ya que la función $\frac{1}{w(u)}$ es decreciente en el intervalo $(0, 1/2]$, la función del supremo será siempre mayor que la función anterior, esto es: $F(U_n) < \frac{1}{w(U_n)}$, por lo que el valor esperado de la medida que es función de la función envolvente es finito, $E[\{F(U_n)\} \ln[1 + F(U_n)]] < \infty$. Esto es debido a los supuestos hechos sobre $w(\cdot)$ y que la función $\{U_n\}$ (la de la marginal) está distribuida uniformemente sobre el intervalo abierto $(0, 1)$.

La demostración de Chen y Fan recurre a la aplicación 5 del teorema 1 de Rio (1995, página 924) que conduce a que la diferencia entre las distribuciones estimadas y real ponderada por la normal converge casi seguramente (*a.s.*), $\left| \frac{\{\hat{U}_n(u) - u\}}{w(u)} \right| = o_{a.s.}(1)$ para cualquier valor de u sobre el dominio de la función envolvente $(0, 1/2]$.

Ahora, para cualquier valor pequeño y positivo de ϵ , es posible formar una red de puntos $u_0 = 0 < u_1 < \dots < u_m = 1/2$ tal que la probabilidad de la envolvente sobre la función indicadora que toma valor de la red de puntos sea menor que ϵ , esto es: $\Pr \left\{ \frac{1}{w(u)} I(U_n \leq u) : u \in (u_i, u_{i+1}) \right\} < \epsilon$ para cada $i \in \{0, 1, \dots, m\}$; por lo tanto, el supremo del cociente de convergencia sobre la función envolvente es menor que el máximo del mismo cociente más un ϵ , *i.e.* $\sup_{u \in [0, 1/2]} \left| \frac{\hat{U}_n(u) - u}{w(u)} \right| \leq \max_i \left| \frac{\hat{U}_n(u) - u}{w(u)} \right| + \epsilon$.

Tomando secuencias de pequeñas variaciones $\epsilon_m \rightarrow 0$, es posible ver que el límite del supremo del cociente de convergencia es cero casi seguramente, *a.s.*, lo que se puede reescribir como:

$$\limsup_n \left\{ \sup_{u \in (0, 1/2]} \left| \frac{\hat{U}_n(u) - u}{w(u)} \right| \right\} = 0.$$

Por lo que la indicadora de la función envolvente, $\left\{ \frac{1}{w(u)} I(U_n \leq u) : u \in (0, 1/2] \right\}$, es de la clase de Glivenko-Cantnelli⁶. Para probarlo, basta con notar que la función $\frac{1}{w(u)}$ es simétrica cuando $u = 1/2$ y decreciente en el intervalo $(0, 1/2)$, además de tener un valor esperado finito. Al multiplicarlo por la indicadora, se obtiene una función empírica que convergerá la teórica.

Para el segundo resultado se sigue un razonamiento similar después de probar que la función envolvente es de la clase de Donsker, para ello es necesario tomar en cuenta los supuestos hechos sobre $w(\cdot)$, por lo que la función envolvente por puntos es menor o igual a la envolvente teórica. Si a eso se agrega el que el proceso $\{U_n\}$ es *L-mixto*, se tiene que las potencias pares de la media son finitas, $E[F(U_n)]^{2\gamma} < \infty$ cuando el decaimiento es

⁶ Se dice que una función es de la clase de Glivenko-Cantnelli si la función de distribución empírica $F_n(x)$ converge punto a punto a la función teórica $F(x)$ por la ley de los grandes números, esto es $\|F_n - F\|_\infty = \sup_{x \in \mathcal{R}} |F_n(x) - F(x)| \rightarrow 0_{a.s.}$

polinomial y $E \left\{ [F(U_n)]^2 \ln [1 + F(U_n)] \right\} < \infty$ para procesos con decaimiento exponencial. Aplicando el teorema 1 de Doukhan *et. al.* (1995), es posible concluir que la función envolvente por puntos, $\left\{ \frac{1}{w(u)} I(U_n \leq u) : u \in (0, 1/2] \right\}$, es de la clase de Donsker.

A continuación se resumen las condiciones de primer orden (Estipuladas por Chen y Fan) necesarias para que los valores del parámetro de dependencia (una variable aleatoria condicionada a los datos) converja en distribución a una normal:

- C1) El parámetro de dependencia verdadero es parte de un subconjunto compacto de un espacio real, $\alpha^* \in A, A \subset \mathbb{R}^d$, y la esperanza del cambio en el vector gradiente ante un cambio del parámetro de dependencia será cero *sii* el parámetro verdadero y el estimado son iguales. Lo que puede expresarse como: $E[l_\alpha(U_n, U); \alpha] = 0$ *sii* $\hat{\alpha} = \alpha^*$.
- C2) El cambio en el vector gradiente, $l_\alpha(u_1, u_2; \alpha)$, está bien definido en el espacio dado por el cuadrado unitario y el conjunto A , $(u_1, u_2; \alpha) \in (0, 1)^2 \times A$, además de que el vector gradiente $l_\alpha(u_1, u_2; \alpha)$ es Lipschitz continuo en $\hat{\alpha}$ con probabilidad uno, para cualquier valor del parámetro de dependencia estimado, $\alpha \in A$. Además de ser una función bien definida a lo largo del dominio $(u_1, u_2; \alpha) \in (0, 1)^2 \times A$. Esta, y la anterior son condiciones estándar.
- C3) El proceso de generación de datos $\{X_n\}$ es *L-Mixto* con una tasa de decaimiento polinomial $\beta_n = O(n^{-b})$ para alguna $b > 0$.

C4) El valor esperado de la función de puntaje es finito, esto es:

$E[\sup_{\hat{\alpha} \in A} \|l_{\alpha, j}(U_n, U_{n-1}; \hat{\alpha})\| \ln [1 + l_\alpha(U_{n-2}, U_n, \hat{\alpha})]] < \infty$, lo que no es otra cosa sino una condición de momento sobre la función de puntaje.

C5) Para $j = 1, 2$ se tiene que $E[\sup_{\hat{\alpha} \in A, j \in G^*} \|l_{\alpha, j}(G(X_n), G(X_{n-1}); \hat{\alpha})\| w(U_{n-2+j})] < \infty$ lo que implica que las derivadas de la función de puntaje (en la verosimilitud) será dominada por una función con primer momento finito al ser ponderada por la función $w(\cdot)$. Si las derivadas parciales están dominadas, se puede escoger a la función identidad como ponderadora y se cumple automáticamente la condición cinco, en caso de que las parciales no estén acotadas, será necesaria una función ponderadora del

tipo $w(u) = [u(1-u)]^{1-\xi}$, para alguna $\xi \in (0, 1)$.

Una vez establecidas las condiciones necesarias para un óptimo en la función de verosimilitud que arroja un parámetro distribuido como normal, es necesario establecer las condiciones suficientes que aseguren un máximo. Para ello se llamará $F_\delta = \{(\alpha, G) \in A \times g_\delta : \|\hat{\alpha} - \alpha^*\| \leq \delta\}$ a una función para valores pequeños de $\delta > 0$, también se denotará como una trayectoria suave perteneciente al espacio de las marginales a G_δ a $\{G_\eta : \eta \in [0, 1]\} \subset G_\delta$, con la particularidad de que $G_\eta|_{\eta=0} = G^*$, que en el caso de un óptimo para la marginal del parámetro de dependencia y para el parámetro, se tiene que: $\{(\alpha_\eta, G_\eta) : \eta \in [0, 1]\} \subset F_\delta$ es una trayectoria suave con $(\alpha_\eta, G_\eta)|_{\eta=0} = (\alpha^*, G^*)$. También es necesario definir:

$$A_n^* \equiv \frac{1}{n-1} \sum_{t=2}^n [l_\alpha(U_n, U_{n-1}; \hat{\alpha}) + W_1(U_{n-1}) + W_2(U_n)]. \quad (AIV.24)$$

Ahora es posible definir las condiciones suficientes para un máximo que garantiza la \sqrt{n} normalidad asintótica del parámetro de dependencia α , a saber:

S1) Satisfacer la condición C1 manteniendo el valor del parámetro dentro del espacio de los parámetros, $\alpha^* \in \text{int}(A)$.

Tener un hessiano, $\mathbf{B} \equiv E[l_{\alpha, \alpha}(U_{n-1}, U_n; \alpha^*)]$, definido negativo.

Mantener finita la varianza del vector gradiente conforme la muestra crece, $\Sigma \equiv \lim_{n \rightarrow \infty} \text{var}(\sqrt{n}A_n^*)$.

Asegurar la convergencia en probabilidad del parámetro estimado con el parámetro verdadero, $\|\hat{\alpha} - \alpha^*\| = o_p(1)$.

Garantizar la convergencia de las copulas a una tasa $n^{-1/2}$ bajo la métrica $w(\cdot)$, $\sup_x \left\{ \left| \frac{C_n(\cdot) - C(\cdot)}{w_2(C(\cdot))} \right| \right\} = O_p(n^{-1/2})$, donde la métrica es un proceso L-mixto con decaimiento exponencial.

S2) Las segundas derivadas, $l_{\alpha, \alpha}(u_1, u_2; \alpha^*)$, son continuas y están bien definidas en el espacio $(u_1, u_2; \alpha) \in (0, 1)^2 \times \text{int}(A)$.

S3) Dada la continuidad y regularidad de las segundas derivadas, se cumple el teorema de Young, por lo que es posible intercambiar el orden de integración y diferenciación

del vector gradiente, $l_\alpha(\lambda_{1,\eta}, \lambda_{2\eta}; \hat{\alpha}_\eta)$, con cambios pequeños cercanos al óptimo, $\eta \in (0, 1)$.

- S4) Los procesos que generan las variables aleatorias son *L-mixtos* con una tasa de decaimiento $\beta_n = O(n^{-b})$ para alguna $b > \gamma/(\gamma - 1)$ con $\gamma > 0$ si el decaimiento es polinomial y $\gamma > 1$ si el decaimiento es exponencial. También se pide que:

$$E \left\{ \sup_{(\hat{\alpha}, G) \in F_\delta} \|l_\alpha(G(x_n), G(x_{n-1}); \hat{\alpha})\| \right\}^{2\gamma} < \infty.$$

- S5) Finalmente se pide que el segundo momento del vector gradiente con la métrica (la cual fue anteriormente descrita) sea finito, esto es:

$$E \left\{ \sup_{(\hat{\alpha}, G) \in F_\delta} \|l_\alpha(G(x_n), G(x_{n-1}); \hat{\alpha})\| w(U_{n-2+j}) \right\}^2 < \infty,$$

para $j = 1, 2$. Todo ello implica que el vector gradiente es dominado por una métrica con segundo momento finito.

Bajo estos supuestos, es posible establecer (Chen y Fan lo hacen como Lemma) que la diferencia entre el parámetro estimado y el verdadero converge en probabilidad a un infinitésimo $n^{-1/2}$, $\hat{\alpha} - \alpha^* = B^{-1} + o_p(n^{-1/2})$, y que es distribuido asintóticamente como normal, $\sqrt{n}(\hat{\alpha} - \alpha^*) \rightarrow N(0, B^{-1}\Sigma B^{-1})$.

El lector interesado en el desarrollo puntual de este lemma puede verificar el trabajo de Chen y Fan marcado en la bibliografía. Más allá de esta sana revisión, es necesario hacer notar al lector que a lo largo de este apéndice se ha demostrado que los estimadores obtenidos usando 2SML son consistentes en sus dos etapas y que estos son distribuidos como normales, por lo que es posible aplicar las pruebas de hipótesis comunes establecidas por la mayor parte de los paquetes econométricos a condición de que el vector gradiente o la función de verosimilitud no exploten (por ello el cambio de métrica).

También es posible apreciar a lo largo de este apéndice que ambos trabajos (Zoltán *et. al.* y Chen *et. al.*) son complementarios, dado que llegan a conclusiones similares a través de caminos distintos, aunque siempre basándose en el hecho de que el proceso de generación de datos o el vector gradiente es *L-mixto*.

Apéndice V

Algoritmos usados para las estimaciones.

AV.1 Algoritmo de muestro inicial para realización de la estadística descriptiva.

```

. Sub muestreo()
    Worksheets("muestras").Range("A1:IR60000").ClearContents
    columnas = Worksheets("Datos mensuales").Cells(1, 2).Value
    largo = Worksheets("Datos mensuales").Cells(1, 4).Value
    Dim matrizmuestreo As Variant
    ReDim matrizmuestreo(largo, columnas)
    contador1 = 1
    For contador1 = 1 To columnas
        contador2 = 1
        obs = Worksheets("Datos mensuales").Cells(3, contador1).Value
        filaocupada = 0
        For contador2 = 1 To obs
            aleatorio = Rnd()
            If (aleatorio < (largo / obs * 1.3)) And (filaocupada < largo) Then
                filaocupada = filaocupada + 1
                matrizmuestreo(filaocupada, contador1) = Worksheets("Datos mensuales").
Cells ((contador2 + 5), contador1).Value
            Else
            End If
        Next contador2
    Next contador1
End Sub

```

```

        Next contador2
    Next contador1
Encabezados
For contador3 = 1 To largo
    For contador4 = 1 To columnas
        Worksheets("Muestras").Cells((contador3 + 3), contador4).Value =
matrizmuestreo (contador3, contador4)
    Next contador4
Next contador3
. End Sub

```

AV.2 Archivos “do” para la estimaciones de marginales por mes.

```

. log using {"Direccion del archivo para el visor \ nombre del archivo.smcl"}
, replace

. set memory 500M
. use {"Direccion del archivo de datos \ nombre del archivo.dta"}, clear

. rename var1 y1

. program define frechetdos

. version {numero de la version}

. args lnf gamma mu sigma

. quietly replace `lnf' = (-1 * ((1 + ('gamma' * (($ML_y1 - 'mu') / 'sigma')))) ^ (-1 / 'gamma')) +
(ln(1 / 'sigma')) + (((1 + 'gamma') / (-1 * 'gamma')) *
(ln(1 + ('gamma' * (($ML_y1 - 'mu') / 'sigma')))))

```

```

. end

. ml model lf frechetdos (y1=) (y1=) (y1=)

. ml check

. ml search -1 1

. ml maximize

. estimates store {nombre del vector de guardado}

. ml graph

. graph save Graph{"Direccion del archivo del grafico \ nombre del archivo.gph"}

. clear

. program drop frechetdos

. log close

```

AV.3 Archivos “do” para la estimaciones de cópulas FGM.

```

. use {"Direccion del archivo de los datos \ nombre del archivo.dta"}, clear

. generate y1= exp(-1*((1+('gamma1*( (var1+ 'media1) / 'desv.est.1^)))^
(-1/'gamma1^)))

. generate y2= exp(-1*((1+('gamma2*( (var2+ 'media2) / 'desv.est.2^)))^
(-1/'gamma2^)))

. replace y1= .999 if y1>.99

. replace y2= .999 if y2>.99

. replace y1= .001 if y1<.001

```

```

. replace y2= .001 if y2<.001

. 'Los reemplazos son para evitar errores al calcular logaritmos de cero, al tener colas
tan pesadas, para los mayores valores se asignan probabilidades de cero o uno.

. program define fgm

. version 1.0

. args lnf alpha

. quietly replace `lnf' = ln(((exp(-1*((1+(.0251871*(($ML_y1-879.9421)/29529.03)))^
(1/.0251871))))/29529.03)*(1+(.0251871*(($ML_y1-879.9421)/29529.03)))^
(-1/.0251871-1))*((exp(-1*((1+(.0817506*(($ML_y2-45.08434)/605.3342)))^
(1/.0817506))))/605.3342)*(1+(.0817506*(($ML_y2-45.08434)/605.3342)))^
(-1/.0817506-1))+
ln(1+'alpha'*((1-2*(exp(-1*((1+(.0817506*(($ML_y2-45.08434)/605.3342)))^
(1/.0817506))))*(1-2*(exp(-1*((1+(.0251871*(($ML_y1-879.9421)/29529.03)))^
(1/.0251871))))))))))

. end

. ml model lf fgm (y1=) (y2=)

. ml check

. ml maximize

. program drop fgm

```

AV.4 Archivos “do” para las estimaciones de las cópulas Ali-Mikhail-Haq

```

. use {"Direccion del archivo de los datos \ nombre del archivo.dta".dta", clear

```

```

. generate y1= exp(-1*((1+('gamma1'*( var1+ 'media1')/'desv.est.1')))^
(-1/'gamma1'))

. generate y2= exp(-1*((1+('gamma2'*( var2+ 'media2')/'desv.est.2')))^
(-1/'gamma2'))

. replace y1= .999 if y1>.99

. replace y2= .999 if y2>.99

. replace y1= .001 if y1<.001

. replace y2= .001 if y2<.001

. 'Los reemplazos son para evitar errores al calcular logaritmos de cero, al tener colas
tan pesadas, para los mayores valores se asignan probabilidades de cero o uno.

. program define amh

. version 1.0

. args lnf alpha

. quietly replace `lnf' = ln((-1*ln($ML_y1))^ (-1-'gamma1')/($ML_y1))
+ln((-1*ln($ML_y2))^ (-1-'gamma2')/($ML_y2))-4*
ln(1-'alpha'*(1-(((1*ln(($ML_y1)))^ (-'gamma1')-1)/'gamma1'))*
(1-(((1*ln(($ML_y2)))^ (-'gamma2')-1)/'gamma2'))
+ln(1-'alpha'*(1-2*(((1*ln(($ML_y2)))^ (-'gamma2')-1)/'gamma2'))
*(1-'alpha'*(1-(((1*ln(($ML_y1)))
^ (-'gamma1')-1)/'gamma1'))*(1-(((1*ln(($ML_y2)))
^ (-'gamma2')-1)/'gamma2'))^ 2-(2*(((1*ln(($ML_y2)))^
(-'gamma2')-1)/'gamma2'))
*(1-(((1*ln(($ML_y1)))^ (-'gamma1')-1)/'gamma1'))*(1-'alpha'*
(1-(((1*ln(($ML_y1)))^ (-'gamma1')-1)/'gamma1'))*(1-(((1*ln(($ML_y2)))^

```



```

(-'gamma2'-1)/'gamma2')))))))
. end
. ml model lf amh (y1=) (y2=)
. ml check
. ml maximize
. program drop amh

```

AV.5 Archivos “do” para la estimaciones de las cópulas Weibull.

```

. use {"Direccion del archivo de los datos \ nombre del archivo.dta"}, clear
. generate y1= exp(-1*((1+('gamma1'* (var1+ 'media1')/'desv.est.1'))^
(-1/'gamma1'))))
. generate y2= exp(-1*((1+('gamma2'* (var2+ 'media2')/'desv.est.2'))^
(-1/'gamma2'))))
. replace y1= .999 if y1>.99
. replace y2= .999 if y2>.99
. replace y1= .001 if y1<.001
. replace y2= .001 if y2<.001
. 'Los reemplazos son para evitar errores al calcular logaritmos de cero, al tener colas
tan pesadas, para los mayores valores se asignan probabilidades de cero o uno.
. program define wei
. version 1.0
. args lnf alpha

```

```

. quietly replace `lnf' = -1*((( -1*ln($ML_y1)) ^ (-`gamma1')-1/`gamma1')
^ (`gamma1')/`alpha'+((-1*ln($ML_y2)) ^ (-`gamma2')-1/`gamma2')
^ (`gamma2')/`alpha'))^`alpha'+(`gamma1'/`alpha'-1)*ln((-1*ln($ML_y1))
^ (-`gamma1')-1/`gamma1')+(`gamma2'/`alpha'-1)*ln((-1*ln($ML_y2))
^ (-`gamma2'-1)/`gamma2')+ln((-ln($ML_y1)) ^ (-`gamma1')-1/($ML_y1))
+ln((-ln($ML_y2)) ^ (-`gamma2')-1/($ML_y2)) +ln(((( -1*ln($ML_y1))
^ (-`gamma1')-1/`gamma1') ^ (`gamma1')/`alpha'+((-1*ln($ML_y2))
^ (-`gamma2')-1/`gamma2') ^ (`gamma2')/`alpha')) ^ (2*(`alpha')-1)
+1-`alpha'/`alpha'(((( -1*ln($ML_y1)) ^ (-`gamma1'-1/`gamma1') ^ (`gamma1')/`alpha'
+((-1*ln($ML_y2)) ^ (-`gamma2'-1)/`gamma2') ^ (`gamma2')/`alpha')) ^ (`alpha')-1)
. end

. ml model lf wei (y1=) (y2=)

. ml check
.
. ml maximize

. program drop wei

```

AV.6 Archivos en BVA para Excel para calcular el parámetro de dependencia usando Bootstrap

```

. Sub betaboot()

    repeticiones = Worksheets("Datos mensuales").Cells(1, 6).Value

    columnas = Worksheets("Datos mensuales").Cells(1, 2).Value

    largo = Worksheets("Datos mensuales").Cells(1, 4).Value

    Dim vectormedias As Variant

```

```

ReDim vectormedias(columnas)

Dim matrizbetas As Variant

ReDim matrizbetas(repeticiones, columnas)

For contador7 = 1 To repeticiones

    Dim muestrerapido As Variant

    ReDim muestrerapido(largo, columnas)

    contador5 = 1

    For contador5 = 1 To columnas

        contador6 = 1

        obs = Worksheets("Datos mensuales").Cells(3, contador5).Value

        filaocupada = 0

        acumuladomedia = 0

        For contador6 = 1 To obs

            aleatorio = Rnd()

            If (aleatorio < (largo / obs * 1.3)) And (filaocupada < largo) Then

                filaocupada = filaocupada + 1

                muestrerapido(filaocupada, contador5) =

Worksheets("Datos mensuales").Cells((contador6 + 5), contador5).Value

                acumuladomedia = acumuladomedia + Worksheets("Datos mensuales").

Cells((contador6 + 5), contador5).Value

            Else

                End If

```

```

Next contador6

vectormedias(contador5) = (acumuladomedia / largo)

Next contador5

'En este punto, todos tienen media correctas

'Aqui calcula las betas

contador8 = 3

contador9 = 1

For contador8 = 3 To columnas

    sumaactual31 = 0 'prepago y sobretasa

    sumaactual32 = 0 'sobretasa e incumplimiento

    sumaactual12 = 0 'prepago e incumplimienton

    For contador9 = 1 To largo

        If (((muestreorapido(contador9, contador8) > vectormedias(contador8)) And
            ((muestreorapido(contador9, (contador8 - 2)) > vectormedias((contador8 - 2)))) Or
            ((muestreorapido(contador9, contador8) < vectormedias(contador8)) And ((muestre-
orapido
(contador9, (contador8 - 2)) < vectormedias((contador8 - 2)))))) Then

            sumaactual31 = sumaactual31 + 1

        Else

            sumaactual31 = sumaactual31 - 1

        End If

        If (((muestreorapido(contador9, contador8) > vectormedias(contador8)) And
            ((muestreorapido(contador9, (contador8 - 1)) > vectormedias((contador8 - 1)))) Or

```

((muestreorapido(contador9, contador8) < vectormedias(contador8)) And ((muestreorapido

(contador9, (contador8 - 1)) < vectormedias((contador8 - 1)))) Then

sumaactual32 = sumaactual32 + 1

Else

sumaactual32 = sumaactual32 - 1

End If

If ((muestreorapido(contador9, (contador8 - 1)) > vectormedias((contador8 - 1)))

And ((muestreorapido(contador9, (contador8 - 2)) > vectormedias((contador8 - 2))))

Or

((muestreorapido(contador9, (contador8 - 1)) < vectormedias((contador8 - 1))) And

((muestreorapido(contador9, (contador8 - 2)) < vectormedias((contador8 - 2))))

Then

sumaactual12 = sumaactual12 + 1

Else

sumaactual12 = sumaactual12 - 1

End If

Next contador9

matrizbetas(contador7, contador8) = sumaactual31 / largo

matrizbetas(contador7, (contador8 - 1)) = sumaactual32 / largo

matrizbetas(contador7, (contador8 - 2)) = sumaactual12 / largo

contador8 = contador8 + 2

```

        Next contador8
    Next contador7
'Para pasar a pantalla las betas
    Sheets("Beta boot").Select
    Range("A11").Select
    Range(Selection, Selection.End(xlToRight)).Select
    Range(Selection, Selection.End(xlDown)).Select
    Selection.ClearContents
    Range("A3").Select
    For contador10 = 1 To columnas
        For contador11 = 1 To repeticiones
            Worksheets("Beta boot").Cells((contador11 + 10), (contador10 + 0)).Value
= matrizbetas(contador11, contador10)
        Next contador11
    Next contador10
. End Sub

```

AV.7 Archivos en BVA para Excel® para calcular las probabilidades de ocurrencia de marginales y cópulas sobre los valores promedio

```

. Sub valorcopula()
    repeticiones = Worksheets("Datos").Cells(1, 6).Value
    columnas = Worksheets("Datos").Cells(1, 2).Value
    muestras = Worksheets("Datos").Cells(1, 4).Value

```

```

'para subir parametros

Dim matrizparametros As Variant

ReDim matrizparametros(7, columnas)

For contador4 = 1 To columnas

For contador5 = 1 To 7

matrizparametros(contador5, contador4) = Worksheets("Parametros").
Cells((contador5 + 7), (contador4 + 1)).Value

Next contador5

Next contador4

' para verificar suba parametros

' For contador6 = 1 To columnas

' For contador7 = 1 To 7

' Worksheets("Parametros").Cells((contador7 + 20), (contador6 + 1)).Value = ma-
trizparametros(contador7, contador6)

' Next contador7

' Next contador6

'para muestrear probabilidades

Dim matrizpromedio As Variant

ReDim matrizpromedio(repeticiones, columnas)

Dim matrizmarginales As Variant

ReDim matrizmarginales(repeticiones, columnas)

contador1 = 1

```

```

For contador1 = 1 To repeticiones

' Dim matrizui As Variant

' ReDim matrizui(muestras, columnas)

contador2 = 3

For contador2 = 3 To columnas

obs = Worksheets("Datos").Cells(3, contador2).Value

filaocupada = 0

acumuladocpei = 0

acumuladocsei = 0

acumuladocsyp = 0

acumuladoprep = 0

acumuladoincu = 0

acumuladosobr = 0

contador3 = 1

For contador3 = 1 To obs

aleatorio = Rnd()

If (aleatorio < (muestras / obs * 1.3)) And (filaocupada < muestras) Then

filaocupada = filaocupada + 1

'prepago e incumplimiento

x1 = Worksheets("Datos").Cells((contador3 + 5), (contador2 - 2)).Value

x2 = Worksheets("Datos").Cells((contador3 + 5), (contador2 - 1)).Value

alpha = matrizparametros(1, (contador2 - 2))

```



```

gamma1 = matrizparametros(2, (contador2 - 2))
gamma2 = matrizparametros(3, (contador2 - 2))
mu1 = matrizparametros(4, (contador2 - 2))
mu2 = matrizparametros(5, (contador2 - 2))
sigma1 = matrizparametros(6, (contador2 - 2))
sigma2 = matrizparametros(7, (contador2 - 2))
u1 = (Exp(-1 * ((1 + (gamma1 * ((x1 - mu1) / sigma1))))(-1 / gamma1))))
u2 = (Exp(-1 * ((1 + (gamma2 * ((x2 - mu2) / sigma2))))(-1 / gamma2))))
cpei = (u1 * u2 * (1 + (alpha * (1 - u1) * (1 - u2))))
acumuladocpei = acumuladocpei + cpei
acumuladoprep = acumuladoprep + u1
'Sobretasa e incumplimiento
x1 = Worksheets("Datos").Cells((contador3 + 5), (contador2 - 0)).Value
x2 = Worksheets("Datos").Cells((contador3 + 5), (contador2 - 1)).Value
alpha = matrizparametros(1, (contador2 - 1))
gamma1 = matrizparametros(2, (contador2 - 1))
gamma2 = matrizparametros(3, (contador2 - 1))
mu1 = matrizparametros(4, (contador2 - 1))
mu2 = matrizparametros(5, (contador2 - 1))
sigma1 = matrizparametros(6, (contador2 - 1))
sigma2 = matrizparametros(7, (contador2 - 1))
u1 = (Exp(-1 * ((1 + (gamma1 * ((x1 - mu1) / sigma1))))(-1 / gamma1))))

```

```

u2 = (Exp(-1 * ((1 + (gamma2 * ((x2 - mu2) / sigma2)))^(1 / gamma2))))
csei = ((u1 * u2) / (1 - (alpha * (1 - u1) * (1 - u2))))
acumuladocsei = acumuladocsei + csei
acumuladosobr = acumuladosobr + u1
acumuladoincu = acumuladoincu + u2

'Sobretasa y prepago

x1 = Worksheets("Datos").Cells((contador3 + 5), (contador2 - 0)).Value
x2 = Worksheets("Datos").Cells((contador3 + 5), (contador2 - 2)).Value

alpha = matrizparametros(1, (contador2 - 0))
gamma1 = matrizparametros(2, (contador2 - 0))
gamma2 = matrizparametros(3, (contador2 - 0))
mu1 = matrizparametros(4, (contador2 - 0))
mu2 = matrizparametros(5, (contador2 - 0))
sigma1 = matrizparametros(6, (contador2 - 0))
sigma2 = matrizparametros(7, (contador2 - 0))

If ((Exp(-1 * ((1 + (gamma1 * ((x1 - mu1) / sigma1)))^(1 / gamma1)))) = 0) Then
u1 = 0.000001

Else

u1 = (Exp(-1 * ((1 + (gamma1 * ((x1 - mu1) / sigma1)))^(1 / gamma1))))

End If

If (Exp(-1 * ((1 + (gamma2 * ((x2 - mu2) / sigma2)))^(1 / gamma2)))) = 0) Then
u2 = 0.000001

```

```

Else
u2 = (Exp(-1 * ((1 + (gamma2 * ((x2 - mu2) / sigma2)))^(-1 / gamma2))))
End If
csyp = (Exp(-1 * (((u1^(gamma1 / alpha)) + (u2^(gamma2 / alpha)))^alpha)))
acumuladocsyp = acumuladocsyp + csyp
Else
End If
Next contador3
matrizpromedio(contador1, (contador2 - 2)) = (acumuladocpei / muestras)
matrizpromedio(contador1, (contador2 - 1)) = (acumuladocsei / muestras)
matrizpromedio(contador1, (contador2 - 0)) = (acumuladocsyp / muestras)
matrizmarginales(contador1, (contador2 - 2)) = (acumuladoprep / muestras)
matrizmarginales(contador1, (contador2 - 1)) = (acumuladoincu / muestras)
matrizmarginales(contador1, (contador2 - 0)) = (acumuladosobr / muestras)
contador2 = contador2 + 2
Next contador2
Next contador1
'para escribir promedios
contador8 = 1
contador9 = 1
For contador8 = 1 To repeticiones
For contador9 = 1 To columnas

```

```
Worksheets("Promedios").Cells((contador8 + 20), (contador9 + 1)).Value = ma-  
trizpromedio(contador8, contador9)
```

```
Next contador9
```

```
Next contador8
```

```
contador10 = 1
```

```
contador11 = 1
```

```
For contador10 = 1 To repeticiones
```

```
For contador11 = 1 To columnas
```

```
Worksheets("Marginales").Cells((contador10 + 20), (contador11 + 1)).Value = ma-  
trizmarginales(contador10, contador11)
```

```
Next contador11
```

```
Next contador10
```

```
. End Sub
```

Parte V

Bibliografía

- Nicholson, W; Microeconomic Theory, basic principles and extentions; South - Western College Publications: 9 Ed; 2004.
- Simon, C & Blume, L; Mathematics for Economists;— Norton & Company; New York; 1994; 930 pp
- Sullivan, William G., *et. al.*; Ingeniería económica de DeGarmo; Pearson Educación; 12 ed; México.
- Fabozzi, F; Fixed Income Mathematics; The McGraw-Hill Companies; 4th Ed; New York; 649 pp
- Kalotay A, Yang D & Fabozzi F; An Option-Theoretic Prepayment Model for Mortgages and Mortgage-Backed Securities; to appear in *International Journal of Theoretical and Applied Finance*; 2005
- Ross, S.M.; Introduction to Probability Models; Academic Press Inc.; 3rd Ed.; Orlando, Florida; USA; 1985
- Pravin K. Trivedi & David M. Zimmer; Copula modeling: An introduction for Practitioners; *Foundations and Trends in Econometrics*; Vol 1; No. 1; 2005; 1 - 111
- Edward W. Frees & Emiliano A. Valdez; Understanding relationships using copulas; *North American Actuarial Journal*; Vol 2; 1-25.
- Sklar, A. (1959), Fonctions de repartition 'a n dimensions et leurs marges, *Publ. Inst. Statist. Univ. Paris* 8, 229-231.
- Nelsen R.B; An Introduction to Copulas; Lecture Notes in Statistics;1999.
- Taleb, Nassim; The a Priori Problem of Observed Probabilities; not refereed.
- Prieger, J. (2002), A flexible parametric selection model for non-normal data with application to health care usage. *Journal of Applied Econometrics* 17, 367392.

- Kimeldorf, G. and A. R. Sampson; Uniform representations of bivariate distributions. *Communications in Statistics* 4; 1975 617627.
- Joe, H. (1993), "Parametric families of multivariate distributions with given margins". *Journal of Multivariate Analysis* 46, 262282.
- Junker, M. and A. May (2005), Measurement of aggregate risk with copulas. *Econometrics Journal* 8, 428454.
- Marshall, A. W. and I. Olkin (1988), Families of multivariate distributions. *Journal of the American Statistical Association* 83, 834841.
- Marshall, A. W. and I. Olkin (1990), Multivariate distributions generated from mixtures of convolution and product families. In: H. W. Block, A. R. Sampson, and T. H. Savits (eds.): *Topics in Statistical 104 References Dependence*, Vol. 16 of IMS Lecture Notes-Monograph Series. pp. 371393.
- Nelsen, R.; Copulas, Characterization, Correlation, and Counterexamples; *Mathematics Magazine*, Vol. 68, No. 3. (Jun., 1995), pp. 193-198.
- Hanagal, D; A Bivariate Weibull Regression Model; *Economic Quality Control*; Vol 20 (2005), No. 1, 1-8; Heldermann Verlag
- Freund, J.E. (1961): A bivariate extension of the exponential distribution. *Journal of Amer. Statist. Assoc.* 56, 971-77.
- Philip Hougaard; Survival Models for Heterogeneous Populations Derived from Stable Distributions; *Biometrika*; Vol.73, No. 2. (Aug., 1986), pp. 387-396.
- Nolan, J; Stable Distributions, Models for Heavy Tailed Data; Math/Stat Department American University; January , 2005; Working Paper
- Cheng K. Lee & Miin-Jye Wen; A Multivariate Weibull Disitribution; Working paper. Charlotte, North Carolina, USA & National Cheng Kung University, City Tainan, Taiwan, R.O.C.; 10 pp.
- Jenq-Tzong Shiau, Hsin-Yi Wang, and Chang-Tai Tsai; Bivariate frequency analysis of floods using copulas; *Journal of the American Water Resources Association, American Water Resources Association*; December 2006

- M.G. Kendall, "A new measure of rank correlation" *Biometrika*, 30 (1938) pp. 8193
- Hervé Abdi (2007); The Kendall Rank Correlation Coefficient; *Encyclopedia of Measurement and Statistics*; Neil Salkind (Ed.). Thousand Oaks (CA); 2007.
- Filip Lindskog, Alexander McNeil & Uwe Schmockkendall (2001); Kendall's Tau for elliptical distributions; Research supported by Credit Suisse, Swiss Re and UBS through RiskLab, Switzerland.
- Durbin, J. and A. S. Stuart (1951), Inversions and rank correlations. *Journal of Royal Statistical Society Series B* 2, 303309.
- Friedrich Schmid & Rafael Schmidt (2007); Nonparametric inference on multivariate versions of Blomqvists beta and related measures of tail dependence; *Metrika*; 66; 323354
- Reiss, R.D. & Thomas, M.; *Statistical Analysis of Extreme Values with applications to Insurance, Finance, Hidrology and other fields*; Birkhuser Verlag, 2 Ed; Berlin; 2001.
- Venegas - Martinez, F.; *Riesgos Financieros y económicos, productos derivados y decisiones económicas bajo incertidumbre*; Cengage, 2 Ed, 2008.
- Joe, H., and Xu, J. J. 1996. The estimation method of inference functions for margins for multivariate models. Technical Report 166. Vancouver: University of British Columbia, Department of Statistics.
- Patton, A. 2006a. Estimation of Multivariate Models for Time Series of Possibly Different Lengths. *Journal of Applied Econometrics* 21 (2): 147-173.
- László, G; Gyrgy, M & Zoltán, R. A two- step maximum- likelihood identification of non-Gaussian systems. Working paper; 2002.
- Chen, X. & Fan, Y. Estimation of copula-based semiparametric time series models, Peer reviewed working paper No. 02-W26R, Department of Economics, Vanderbilt University, Nashville, TN 37235.
- Stigler, Stephen (December 1973). "Studies in the History of Probability and Statistics. XXXII: Laplace, Fisher and the Discovery of the Concept of Sufficiency". *Biometrika* 60 (3): 439445
- Down, D., S. P. Meyn, and R. L. Tweedie (1995), Exponential and Uniform Ergodicity of

- Markov Processes, *The Annals of Probability* 23, 1671-1691.
- Jarner, S. and G. Robert (2001), Polynomial Convergence Rates of Markov Chains, Working paper, Lancaster University.
- Greene, H. W, *Econometric Analysis*, Prentice Hall, 5 ed, 2003, pp 260.
- Kalotay, A., Yang, D., Fabozzi, F.J. "An Option-Theoretic Prepayment Model for Mortgages and Mortgage-Backed Securities", *International Journal of Theoretical and Applied Finance*, 2003, in press.
- Curley, A. J. and Guttentag, J. M. (1974). *Explorations in Economic Research*, The Yield on Insured Residential Mortgages, pages 114161. National Bureau of Economic Research.
- Hayashi, F., *Econometrics*, Princeton University Press, 2000.
- FNMA, Single-Family REMIC Prospectus, Guaranteed Single-Family REMIC Pass - Through Certificates, August 1, 2007.
- Chung, K.L., "Teoría de la probabilidad y de los procesos estocásticos", Editorial Reverte, 3 Ed., 1983.
- Doob, J.L., Heuristic approach to the Kolmogorov - Smirnov Theorems, *The Annals of Mathematical Statistics*, 393 - 403.
- Rio, E. (1995), A Maximal Inequality and Dependent Marcinkiewicz-Zygmund Strong Laws, *Annals of Probability* 23, 918-937.