

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS MATEMÁTICAS
Departamento de Estadística e Investigación Operativa



**TEORÍA DE CÓPULAS Y CONTROL DE RIESGO
FINANCIERO**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR
PRESENTADA POR

Rosario Cintas del Río

Bajo la dirección del doctor:
Vicente Quesada Paloma

Madrid, 2007

ISBN: 978-84-669-3132-8

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS MATEMÁTICAS

DEPARTAMENTO DE ESTADÍSTICA E INVESTIGACIÓN OPERATIVA



**TEORÍA DE CÓPULAS
Y
CONTROL DE RIESGO FINANCIERO**

Rosario Cintas del Río

*Memoria para optar al grado de
Doctor en Ciencias Matemáticas
realizada bajo la dirección del
Dr.D. Vicente Quesada Paloma.
Diciembre 2006.*

Agradecimientos

En primer lugar quiero expresar mi agradecimiento al director de esta tesis y tutor de mis estudios de doctorado, el profesor Dr. D. Vicente Quesada Paloma, cuya dedicación y dirección ha sido fundamental en todo momento. Sus acertadas directrices, orientaciones y consejos, así como el excelente trato personal que me ha dispensado a lo largo de estos años resultan impagables por mi parte.

Tampoco puedo olvidar a mis compañeros Inés y José Luis. A la primera le agradezco el tiempo que ha dedicado a nuestras interesantes discusiones y al segundo sus excelentes sugerencias informáticas.

Por último, me gustaría mencionar a Benito, Alberto y Pablo, sin cuyo apoyo y cariño no hubiera sido posible concluir este trabajo.

Índice

Introducción	1
1 Cópulas y Variables Aleatorias	1
1.1. Conceptos Básicos	1
1.2. Teorema de Sklar. Interpretación probabilística de las cópulas	4
1.3. Cotas de Fréchet-Hoeffding	6
1.4. Gráficos y diagramas de contorno de una cópula	9
1.5. Propiedades de las cópulas	11
1.6. Cópula de supervivencia y función de supervivencia conjunta	16
1.7. Limitaciones del coeficiente de correlación lineal. Medidas de con- cordancia	19
1.8. Coeficiente de dependencia en colas (TDC)	23
1.9. Cópula empírica	26
1.10. Familias de cópulas más usuales en el entorno financiero	27
1.10.1. Cópulas elípticas	27
1.10.2. Cópulas Arquimedianas	30
1.10.3. Cópulas de valores extremos	36

2	Ajuste de Cópulas a Series Financieras Bivariantes	39
2.1.	Aproximación no paramétrica	40
2.2.	Aproximación semiparamétrica	44
2.3.	Aproximación paramétrica	48
2.4.	Bondad de ajuste	51
2.5.	Ajuste de cópulas a series diarias de retornos de Índices financieros	54
2.5.1.	Descripción de las series	54
2.5.2.	Ajuste no paramétrico de la cópula bivalente DowJones & Ibex35	61
2.5.3.	Ajuste paramétrico de la cópula bivalente DowJones & Ibex35	66
2.5.4.	Evaluación de la bondad de ajuste	68
2.5.5.	Modelos alternativos de cópulas mixtas	73
3	Filtrado de la Dependencia Temporal Marginal	79
3.1.	Modelos de volatilidad	83
3.1.1.	Características de las Series Financieras	84
3.1.2.	Formulación del Modelo ARCH	86
3.1.3.	Formulación del Modelo GARCH	89
3.1.4.	Extensiones del modelo GARCH	94
3.2.	Ajuste de procesos GARCH a los retornos diarios DowJones e Ibex35	97
3.2.1.	Selección y diagnóstico de los modelos	97
3.2.2.	Ajuste de cópulas bidimensionales a los retornos filtrados DowJones e Ibex35	108
4	Dependencia en Colas	111

4.1. Métodos gráficos para detectar la dependencia bivalente: Chi-plot y K-plot	113
4.1.1. Chi-plot	113
4.1.2. K-plot	118
4.2. Modelos de Extremos	121
4.2.1. Distribución de máximos	123
4.2.2. Distribución de excesos	126
4.2.3. Selección del umbral	132
4.2.4. Estimación de las colas de las distribuciones de pérdidas DowJones e Ibex35	133
4.2.5. Medidas de riesgo para las distribuciones de pérdidas univariantes DowJones e Ibex35	141
4.3. Medidas de Riesgo para una Cartera de Inversión	145
4.3.1. Medidas de riesgo calculadas sobre la cartera DowJones e Ibex35	148
5 Cópulas con Estructura Dinámica	159
5.1. Modelo de cópula dinámica DowJones&Ibex35	165
6 Conclusiones y Problemas abiertos	171
Bibliografía	175

Introducción

La gestión del riesgo operacional representa para las entidades financieras una de las tareas de mayor importancia en sus diferentes etapas de identificación, medida y control. Por otra parte, el nuevo Acuerdo de Capital de Basilea (Basilea II), proporciona un estándar internacional de supervisión que las instituciones financieras están obligadas a cumplir (BIS, 1996). Estas nuevas reglas marcan patrones a seguir en el análisis de las tres componentes esenciales del riesgo: riesgo de mercado, riesgo de crédito y riesgo operacional. De esta forma, se confirma la gestión del riesgo como una de las piezas clave para la determinación de las provisiones dinámicas de niveles mínimos de capital físico que las entidades deben mantener con el fin de cubrir las pérdidas esperadas. Por esta razón, resulta de interés establecer modelos de riesgo que sirvan a una doble finalidad: la de proponer controles internos que mejoren la propia gestión y la de presentar informes precisos a este respecto a las instituciones supervisoras externas.

Es en este entorno donde surge el problema de la modelización de series de retornos financieros, como ayuda para predecir la volatilidad en periodos futuros y probabilidades asociadas a ella. Una simple inspección gráfica de la distribución empírica de este tipo de series permite apreciar algunos rasgos comunes a todas ellas, como son la asimetría y las colas pesadas, es decir, la presencia de más realizaciones en las colas que las esperables en una distribución normal. Esta última observación resulta de vital importancia en el control de riesgos y, en particular, en el análisis y medida del riesgo potencial de pérdidas económicas, como explicaremos a continuación. Tradicionalmente las instituciones financieras cuantifican la posibilidad de que se produzca una pérdida más grande que una cierta cantidad fijada, sobre un horizonte temporal dado, mediante el llamado Valor en Riesgo (VaR) (Smith, 2002, Danielsson y otros 2000); por ejemplo, el VaR puede decirnos que más allá de 20 días se puede esperar una pérdida de al menos el

2%. De forma sencilla y en este contexto, una cartera puede considerarse como la combinación lineal de un conjunto de precios de activos individuales. Si la composición de la cartera se mantiene fija, es posible evaluar el riesgo mediante alguna técnica univariante (por ejemplo, mediante Teoría de Valores Extremos), tratando simplemente el precio de la cartera como variable de interés. Sin embargo, lo habitual es diseñar la cartera de forma que se maximice el rendimiento esperado sujeto a alguna restricción sobre el VaR de la misma. Para resolver un problema de esta naturaleza, en el que es necesario determinar los pesos de los diferentes activos, resulta esencial considerar la distribución conjunta de los precios. Las técnicas convencionales aplicadas en el cálculo del VaR se basan en la hipótesis de normalidad multivariante para la distribución conjunta de los retornos, pero esta hipótesis es cuestionable en la mayoría de los casos dado que, como ya hemos comentado, tratamos con series con colas más pesadas que las que proporcionaría la situación de normalidad, lo que complica la determinación de las probabilidades en las colas. Concretamente, en el caso de la dependencia entre mercados financieros, los trabajos de Ang y Chen (2002), Login y Solnik (2001) y Ang y Bekaert (2002), ponen en evidencia la hipótesis de normalidad aplicada a observaciones procedentes de mercados tanto domésticos como internacionales.

Los métodos tradicionales de simulación histórica (Crouhy, Galai y Mark 2001; Marshall 2001; Vilarino 2001), aunque pueden contemplar los fenómenos de asimetría y colas pesadas ya comentados, presentes en las distribuciones empíricas, tienen el inconveniente de no permitir predecir pérdidas fuera de la muestra y, por tanto, resultan pobres a la hora de estimar el VaR: las pérdidas pasadas resultan no ser buenos estimadores de las pérdidas futuras. Además las medidas de riesgo obtenidas a partir de las series históricas, con frecuencia son sensibles al tamaño de la ventana temporal empleada y dificultan los análisis de sensibilidad de determinados parámetros a pequeñas variaciones. De todo lo anterior, surge la necesidad de modelizar las distribuciones empíricas mediante modelos multivariantes suficientemente flexibles. Paralelamente a este desarrollo, surge el de los métodos inferenciales necesarios para determinar los modelos más apropiados a cada situación práctica.

Es en esta situación donde demuestran su utilidad las funciones cópulas, como una herramienta para describir la estructura de dependencia subyacente entre un grupo de variables aleatorias, de las cuales se conoce el comportamiento individual, y determinar el efecto de dicha estructura sobre el VaR bajo un modelo de probabilidad concreto, es decir, medimos y comparamos riesgos a partir de

los cuantiles de la distribución conjunta. Las funciones cópulas tienen una larga historia en Teoría de la Probabilidad aunque han sido estudiadas bajo diferentes nombres: t-normas, funciones de dependencia, operadores de Markov o medidas doblemente estocásticas. Para una introducción a la Teoría de cópulas pueden consultarse los textos de Joe (1997) y Nelsen (1999), o el trabajo de Frees y Valdez (1998). La idea de una función que caracterice la estructura de dependencia entre diferentes variables aleatorias proviene de los trabajos de Hoeffding (1940, 1948), aunque fue Sklar (1959) quién definió y estableció la denominación de *cópula* para una función que obtiene la distribución multivariante como función de las distribuciones marginales univariantes. Sin embargo, su aplicación en el marco financiero es muy reciente: la primera idea aparece en torno al estudio de las limitaciones del coeficiente de correlación lineal como medida de dependencia (Embretchs, McNeil y Straumann, 1999). El texto de Cherubini, Luciano y Vecchiato (2004) resulta bastante interesante en cuanto a aplicaciones en el campo financiero y actuarial y el trabajo de Embretchs, Lindskog y McNeil (2001) constituye un buen resumen desde el punto de vista teórico.

Las cópulas permiten representar funciones de distribución conjuntas así como distinguir el comportamiento de las marginales, a partir del patrón de dependencia capturado por la cópula misma. Esto resulta muy útil no sólo en la fase de modelización, sino también en las de estimación o simulación. Esencialmente, nos proporcionan una manera relativamente sencilla de extender los modelos financieros tradicionales, que partían de la hipótesis de normalidad conjunta, a situaciones más generales incluyendo aquellas en las que se supone normalidad únicamente en las marginales.

Actualmente, la Teoría de Cópulas está bastante consolidada, aunque las propuestas inferenciales aún resultan escasas. Las aportaciones de Genest y Rivest (1993), Genest, Ghoudi y Rivest (1995) y Joe (1997) resultan esenciales en este contexto, y Genest y Favre (2006) realizan una interesante recopilación de este tipo de técnicas. El interés, por tanto, se concentra en el desarrollo de modelos de cópulas multivariantes, capaces de capturar rasgos relevantes de series financieras y sobre los métodos inferenciales aplicables a dichos modelos, que resultan esenciales en la consecución de los siguientes fines: estimación de funciones de riesgo dependientes, predicción de series temporales financieras y evaluación de medidas de riesgo.

Objetivos y Resultados

Esta tesis tiene un doble objetivo. En primer lugar, introduce modelos tanto estáticos como dinámicos basados en funciones cópulas, capaces de capturar rasgos relevantes de series financieras bivariantes, con el fin de ser útiles en el control y valoración de riesgos potenciales de mercados financieros. En segundo lugar, analiza el comportamiento conjunto de dos mercados consolidados: el americano, representado por el índice Dow Jones, y el europeo, a través del índice Ibex35, razón por la cual, la teoría y la práctica van de la mano a lo largo de los diferentes desarrollos.

Fundamentalmente, nuestro análisis se ha referido al comportamiento de pérdidas extremas, reflejado en las colas inferiores, tanto de las distribuciones marginales como del modelo bivalente. La metodología que se presenta resulta aplicable a una gran variedad de situaciones reales de control de riesgo.

La tesis se ha estructurado a lo largo de cinco capítulos en los que se exponen los diferentes métodos y resultados, y un último capítulo dedicado a las conclusiones más relevantes, en el que además se describen los problemas abiertos a futuras investigaciones.

El **Capítulo 1** está dedicado a las nociones básicas, propiedades y herramientas utilizados en los capítulos posteriores. En él se establece la definición de cópula, relacionada con la de variable aleatoria, junto con las propiedades más relevantes de las mismas incluyendo el resultado fundamental, conocido como Teorema de Sklar (Sklar 1959), que permite la aplicación de la aproximación basada en cópulas en la modelización de estructuras de dependencia multivariantes. Incluimos un análisis detallado de las limitaciones del uso del coeficiente de correlación lineal como medida de dependencia, lo que justifica la necesidad de considerar otras medidas de concordancia, tales como la tau de Kendall o la rho de Spearman. A continuación, se plantea el concepto de dependencia en colas como una propiedad de la cópula, y no de las distribuciones marginales. Seguidamente se define la cópula empírica, para finalizar con una descripción de los modelos paramétricos de cópula más utilizados en el ámbito financiero, prestando especial atención a la familia de las arquimedianas y la de valores extremos, debido a las buenas propiedades que presentan.

El **Capítulo 2** aborda la cuestión relativa a la elección del modelo de cópula que representa de forma más adecuada la estructura de dependencia entre series

bivariantes de retornos de índices financieros. En este capítulo nos centramos exclusivamente en la elección del modelo estático de cópula, ya sea de tipo empírico o paramétrico, basado en la información proporcionada por las distribuciones empíricas continuas marginales; esto es, no se supone un modelo paramétrico para el comportamiento marginal, con el fin de evitar errores procedentes de malas especificaciones. Puesto que la función de distribución empírica es de naturaleza discreta, lo que puede originar problemas en un proceso posterior de ajuste de tipo paramétrico si optamos por el método de la máxima verosimilitud, ha sido necesario utilizar un procedimiento que dote a la función de distribución empírica de soporte continuo.

A continuación, en una primera aproximación, se obtiene la cópula empírica, que extrae de forma directa la estructura de dependencia de las series brutas de datos y permite estimar de forma consistente la cópula real que los representa. *Dicha cópula empírica se dedica a diferentes cometidos, como son la identificación del signo de la dependencia bivalente, la cuantificación de la dependencia conjunta y de la dependencia en colas*, esta última mediante un gráfico específico que permite detectar posibles convergencias en los valores de los coeficientes concretos. También servirá para calibrar diferentes familias de cópulas candidatas a utilizar en un ulterior ajuste paramétrico.

Si los modelos de cópulas candidatas al ajuste pertenecen a la familia de las arquimedianas, se utiliza un método alternativo no paramétrico (Genest y Rivest 1993) de estimación del parámetro $\theta \in \Theta$ identificativo de la cópula, basado en la τ de Kendall. Se calculan los correspondientes errores e intervalos de confianza. En este contexto, también es posible realizar una estimación de tipo paramétrico basada en el generador específico de cada modelo arquimediano, que se comparará con la anterior.

Posteriormente, en un planteamiento de naturaleza paramétrica (para el cual lo expuesto anteriormente no resulta excluyente sino complementario e incluso recomendable como tratamiento previo) supondremos que la cópula verdadera pertenece a una familia paramétrica $\{C_\theta, \theta \in \Theta\}$, con Θ el espacio de posibles valores del parámetro. Mediante un método de máxima pseudo-verosimilitud se obtendrán estimadores de $\theta \in \Theta$, consistentes y asintóticamente normales. En un paso siguiente, evaluamos la bondad del ajuste realizado, para lo que *se proponen métodos de tipo gráfico y analítico*.

Por último, se lleva a la práctica todo lo anteriormente expuesto, realizando el

análisis de la serie bivalente formada por los retornos logarítmicos diarios de los índices financieros Dow Jones e Ibex35. En esta implementación práctica, partimos de algunos de los modelos estáticos planteados inicialmente como candidatos al ajuste, para *crear dos modelos alternativos de cópulas que operan de una manera más eficiente, y que se obtienen como combinaciones lineales convexas de los primeros*. Estos modelos están identificados por un parámetro de dependencia tridimensional, cuya estimación se realiza por el método de máxima pseudo-verosimilitud antes indicado. *Para ellos se obtiene, como resultado teórico, la relación funcional entre los coeficientes de dependencia en colas y los parámetros de dependencia*. Esta situación ilustra de forma evidente la realimentación existente entre teoría y práctica, presente a lo largo de esta tesis, puesto que estos nuevos modelos de tipo mixto adquieren un carácter relevante en los desarrollos teóricos de capítulos posteriores, dado que capturan una amplia variedad de situaciones reales de dependencia bivalente. El capítulo, contemplado en su totalidad, constituye una *propuesta metodológica de inferencia para cópulas bivariantes estáticas*.

Dedicamos el **Capítulo 3** a la elección de los modelos marginales más adecuados a cada serie de retornos de los índices considerados. La realidad es que las observaciones de las series univariantes no son necesariamente independientes en el tiempo, por lo que el planteamiento básico de que partimos es que cada una de ellas sigue un modelo de serie temporal estacionario con estructura de volatilidad estocástica.

Abordamos, entonces, la representación del comportamiento de las distribuciones marginales de los retornos en dos fases: una primera en la que utilizaremos modelos tipo GARCH (modelos dinámicos de varianza condicionada; Engle 1982, Bollerslev, Chow y Kroner 1992) y máxima verosimilitud para obtener estimaciones de la volatilidad condicionada, para proceder, en una segunda fase, a confirmar mediante el análisis de datos exploratorio y los contrastes estadísticos pertinentes, que los residuos constituyen, al menos de forma aproximada, una serie i.i.d. con colas pesadas. Esta aproximación, que estima la distribución condicionada de los retornos a partir de la de los residuos y las estimaciones de la volatilidad y de la media condicionada, refleja dos rasgos presentes en la mayoría de las series de retornos financieros: la volatilidad estocástica y las colas pesadas de las distribuciones condicionadas de los retornos sobre horizontes temporales cortos. A lo largo del capítulo se establecen una serie de *procedimientos de selección y diagnóstico* de este tipo de modelos.

Los planteamientos teóricos previos se llevan a la práctica sobre cada una de las series de retornos logarítmicos diarios de los índices Dow Jones e Ibex35, para a continuación realizar la selección del modelo bidimensional de cópula que mejor se ajusta a la serie bivalente de residuos filtrados mediante los procesos tipo GARCH elegidos como modelos marginales. El modelo de cópula obtenido de esta forma es invariante en el tiempo, pero asume comportamientos marginales dinámicos.

Una vez resuelta la dificultad que suponía la presencia de dependencia temporal en las series marginales de retornos, mediante el filtrado de las mismas a través de procesos tipo GARCH, dedicamos el **Capítulo 4** al tratamiento detallado de la dependencia local en colas entre las series de retornos filtrados. En primer lugar, *adaptamos herramientas gráficas (Chi-plot, K-plot) para detectar la presencia de la dependencia, no solo de tipo general, sino también en las colas de la distribución bivalente*. A continuación, modelizamos las colas de las distribuciones marginales de los retornos filtrados. Para ello, nos servimos de métodos umbral (POT) que permiten el ajuste de distribuciones de Pareto generalizadas (GPD) a las muestras univariantes de residuos que superen un umbral dado. Se propone una *metodología de selección del umbral* más adecuado para el ajuste.

En un siguiente paso, se cuantifica el riesgo de la pérdida conjunta, concretada a través de una cartera de inversión con unos pesos de localización especificados. Se utilizan como modelos de pérdidas marginales los obtenidos mediante Distribuciones de Pareto Generalizadas (GPD) univariantes que representan los comportamientos de las colas de pérdidas de los índices financieros implicados. *Nuestra aportación consiste en ligar estas distribuciones marginales mediante los dos modelos mixtos de cópulas obtenidos en el capítulo 2, y utilizar esta estructura de dependencia conjunta para simular retornos de la cartera lineal*. Con este fin, se ha desarrollado un *algoritmo de simulación de muestras de retornos de la cartera basado en dichos modelos de cópula y en los modelos univariantes proporcionados por las GPD ajustadas*, en lugar de utilizar la aproximación cuantílica para cada distribución univariante. Al igual que en los Capítulos 2 y 3, las técnicas propuestas se aplican al caso Dow Jones e Ibex35.

Hasta el momento, se ha supuesto que el parámetro de dependencia del modelo o modelos de cópula seleccionados tiene un valor constante (o dos valores constantes en el caso de los modelos mixtos) lo que presupone que la dependencia bivalente permanece estática a lo largo del periodo muestral considerado.

Resulta más realista contemplar la posibilidad de que dicha dependencia posea una estructura que cambie a lo largo del tiempo, dado que en el caso concreto de los mercados financieros existe una clara componente dinámica que afecta a su evolución. El **Capítulo 5** plantea una *propuesta original que permite dotar al modelo de cópula de una estructura dinámica, basada en la evolución de la dependencia bivalente localizada en las colas, cuantificada a través de los coeficientes de dependencia (TDC)*. Una vez obtenida la serie de valores de dichos coeficientes a partir de la llamada cópula cola empírica, se propone para ella un modelo de serie temporal en el cual la influencia de valores cercanos al instante considerado sea mayor que la de los más alejados en el tiempo (alisado exponencial doble de Holt).

A partir de las predicciones de los TDC y de sus expresiones en función de los parámetros de dependencia de los modelos de cópula, es posible obtener la evolución dinámica de la dependencia ajustando el modelo de cópula bidimensional en cada instante. De esta forma, somos capaces de generar retornos de la cartera en base a la cópula dinámica y efectuar valoraciones del riesgo a lo largo de un horizonte temporal fijado previamente. El capítulo acaba con la obtención de los resultados correspondientes para el caso Dow Jones e Ibex35.

Esta tesis finaliza con un último capítulo, en el que se resumen las conclusiones más importantes de nuestro trabajo, así como las cuestiones abiertas que puedan considerarse como puntos de partida para futuras investigaciones.

Los resultados y metodologías expuestos a lo largo de los capítulos 2, 3, 4 y 5 detallados anteriormente, se han automatizado mediante diversos procedimientos informáticos realizados en lenguaje Splus 6.1 en su versión para Windows, de forma que permitan realizar un análisis similar al expuesto sobre cualquier muestra bivalente de retornos financieros. Al final de esta memoria se adjunta un anexo en el que se relacionan los diferentes procedimientos y funciones desarrollados para la implementación de las distintas técnicas introducidas.

Es nuestro deseo que este trabajo constituya un resultado útil, tanto desde el punto de vista teórico como práctico, y pueda ser considerado como una modesta aportación dentro del conjunto de metodologías estadísticas orientadas al control cuantitativo del riesgo operacional, inherente a los mercados financieros.

Cóputas y Variables Aleatorias

1.1. Conceptos Básicos

En términos generales, una cópula es una función de distribución multivariante definida sobre el cubo unidad $[0, 1]^n$ con marginales distribuidas uniformemente. Esta definición resulta natural si la cópula se obtiene a partir de una función de distribución multivariante continua, en cuyo caso la cópula es simplemente la función de distribución multivariante original con marginales univariantes transformadas; sin embargo, enmascara alguno de los problemas que aparecen cuando la cópula se construye utilizando otras técnicas, es decir, es necesario establecer el significado del término “función de distribución multivariante”. Por esta razón, debemos comenzar con una definición más abstracta y, posteriormente, establecer la definición operacional de cópula en términos de variables aleatorias y función de distribución.

Sean S_1, \dots, S_n subconjuntos no vacíos de $\overline{\mathbb{R}} = [-\infty, +\infty]$. Sea H una función real de n variables tales que $DomH = S_1 \times \dots \times S_n$ y para $a \leq b$ ($a_k \leq b_k \forall k$) sea $B = [a, b] = [a_1, b_1] \times \dots \times [a_n, b_n]$ cuyos vértices están en el $domH$.

Definición 1.1

Llamamos H -volumen de B a

$$V_H(B) = \sum sign(c)H(c)$$

donde la suma se toma sobre todos los vértices c de B y

$$\text{sign}(c) = \begin{cases} 1 & \text{si } c_k = a_k \text{ para un } n^\circ \text{ par de } k's \\ -1 & \text{si } c_k = a_k \text{ para un } n^\circ \text{ impar de } k's \end{cases}$$

Definición 1.2

Una función real H de n variables es n -creciente si $V_H(B) \geq 0 \forall B$ cuyos vértices están en el $\text{Dom}H$.

Sea H la función de variable real n -dimensional con $\text{Dom}H = S_1 \times \dots \times S_n$ donde cada S_k tiene un elemento más pequeño a_k . Supongamos que H verifica la siguiente propiedad:

Propiedad 1.1 $H(t) = 0 \forall t \in \text{dom}H$ tal que $t_k = a_k$ para al menos un k .

Si cada S_k es no vacío y tiene un elemento más grande b_k , entonces H tiene marginales, siendo las marginales unidimensionales de H las funciones H_k con $\text{Dom}H = S_k$ y con $H_k(x) = H(b_1, \dots, b_{k-1}, x, b_{k+1}, \dots, b_n) \forall x \in S_k$ que son las denominadas simplemente *marginales*. Las marginales de dimensión superior se definen de manera análoga.

Lema 1.1

Sean S_1, \dots, S_n subconjuntos no vacíos de $\overline{\mathbb{R}}$ y sea H la definida anteriormente, n -creciente y verificando la propiedad 1.1. con $\text{Dom}H = S_1 \times \dots \times S_n$. Entonces H es creciente en cada argumento, es decir, si $(t_1, \dots, t_{k-1}, x, t_{k+1}, \dots, t_n)$ y $(t_1, \dots, t_{k-1}, y, t_{k+1}, \dots, t_n)$ están en $\text{Dom}H$ y $x \leq y$, entonces

$$H(t_1, \dots, t_{k-1}, x, t_{k+1}, \dots, t_n) \leq H(t_1, \dots, t_{k-1}, y, t_{k+1}, \dots, t_n)$$

Lema 1.2

Sean S_1, \dots, S_n subconjuntos no vacíos de $\overline{\mathbb{R}}$ y sea H la definida anteriormente, n -creciente y verificando la propiedad 1.1., con marginales y $\text{Dom}H = S_1 \times \dots \times S_n$, entonces si $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$ son puntos en el $\text{Dom}H$ se verifica que

$$|H(x) - H(y)| \leq \sum_{k=1}^n |H_k(x) - H_k(y)|$$

(La demostración de ambos Lemas puede verse en Sklar y Schweizer, 1983).

Definición 1.3

Una función de distribución n -dimensional es una función H con dominio $\overline{\mathbb{R}}^n$ tal que H es n -creciente, verifica la propiedad 1.1. y $H(\infty, \dots, \infty) = 1$.

Del Lema 1.1 se sigue que las marginales de una función de distribución n -dimensional son funciones de distribución que denotaremos F_1, \dots, F_n .

Definición 1.4

Una Cópula n -dimensional es una función C con dominio $[0, 1]^n$ tal que:

- (i) C es n -creciente y verifica la propiedad 1.1.
- (ii) C tiene marginales C_k , $k = 1, 2, \dots, n$, que satisfacen $C_k(u) = u \forall u \in [0, 1]$.

Obsérvese que para cualquier cópula n -dimensional C , $n \geq 3$, las marginales k -dimensionales de C son cópulas k -dimensionales. De forma equivalente, una cópula n -dimensional C es una función con las siguientes propiedades:

1. Para cada $\mathbf{u} \in [0, 1]^n$, $C(\mathbf{u}) = 0$ si al menos una coordenada de \mathbf{u} es 0, y $C(\mathbf{u}) = u_k$ si todas las coordenadas de \mathbf{u} son iguales a 1 excepto u_k .
2. Para cada \mathbf{a} y \mathbf{b} en $[0, 1]^n$ tales que $a_i \leq b_i \forall i = 1 \dots n$ se verifica que $V_C[a, b] \geq 0$.

Puesto que las cópulas son funciones de distribución conjuntas (sobre $[0, 1]^n$), una cópula C induce una medida de probabilidad sobre $[0, 1]^n$ a través de

$$V_C([0, u_1] \times \dots \times [0, u_n]) = C(u_1, \dots, u_n)$$

y una extensión estándar a subconjuntos de Borel arbitrarios de $[0, 1]^n$. A partir de resultados conocidos de la teoría de la medida, podemos asegurar que existe una única medida de probabilidad sobre subconjuntos de Borel de $[0, 1]^n$ que coincide con V_C sobre el conjunto de elementos de $[0, 1]^n$. Denotaremos esta medida de probabilidad por V_C . A partir de la definición se sigue que una cópula C es una función de distribución sobre $[0, 1]^n$ con marginales distribuidas uniformemente en $[0, 1]$.

El siguiente Teorema se sigue directamente del Lema 1.2:

Teorema 1.1

Sea C una cópula n -dimensional. Entonces para cada $\mathbf{u}, \mathbf{v} \in [0, 1]^n$

$$|C(\mathbf{v}) - C(\mathbf{u})| \leq \sum_{k=1}^n |v_k - u_k|$$

Por tanto, C es uniformemente continua en $[0, 1]^n$.

1.2. Teorema de Sklar. Interpretación probabilística de las cópulas

El punto de partida para la aplicación financiera de las cópulas es su interpretación probabilística, es decir, la relación entre cópulas y funciones de distribución de variables aleatorias. Esta relación está esencialmente establecida en el *Teorema de Sklar* que asegura no solamente que las cópulas son funciones de distribución conjuntas, sino que el recíproco también es cierto: las funciones de distribución conjuntas se pueden reescribir en términos de las marginales y una única subcópula, que a su vez puede extenderse (en general, no de forma única) a una cópula. Es más, la mayoría del estudio de funciones de distribución conjuntas puede reducirse al estudio de cópulas.

Teorema 1.2 (*Sklar 1959*)

Sea H una función de distribución n -dimensional con marginales F_1, \dots, F_n . Entonces existe una cópula n -dimensional C tal que $\forall \mathbf{x} \in \overline{\mathbb{R}}^n$

$$H(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)) \quad (1.1)$$

Si F_1, \dots, F_n son todas continuas, entonces C es única, por tanto, está unívocamente determinada sobre $\text{Ran}F_1 \times \dots \times \text{Ran}F_n$.

Inversamente, si C es una cópula n -dimensional y F_1, \dots, F_n son funciones de distribución, entonces la función H definida anteriormente es una función de distribución n -dimensional con marginales F_1, \dots, F_n .

(La demostración puede verse en Sklar 1959).

De acuerdo con este resultado, cuando escribimos por ejemplo para el caso

bidimensional:

$$F(x, y) = C(F_1(x), F_2(y))$$

repartimos la probabilidad conjunta entre las marginales y una cópula, de forma que esta última solamente representa la asociación entre X e Y . Las cópulas separan el comportamiento marginal (representado por las F_i) del conjunto, en contra de lo que ocurre en la representación usual de probabilidades conjuntas vía función de distribución. Por esta razón, las cópulas son denominadas *funciones de dependencia* (Deheuvels 1978).

Observaciones: (Consecuencias del Teorema de Sklar)

$$P(X \leq x, Y > y) = F_1(x) - C(F_1(x), F_2(y))$$

$$P(X > x, Y \leq y) = F_2(y) - C(F_1(x), F_2(y))$$

$$P(X \leq x / Y \leq y) = \frac{C(F_1(x), F_2(y))}{F_2(y)}$$

$$P(X \leq x / Y > y) = \frac{F_1(x) - C(F_1(x), F_2(y))}{1 - F_2(y)}$$

$$P(X \leq x / Y = y) = C_{1/2}(F_1(x), F_2(y)) = \left[\frac{\partial C(u, v)}{\partial v} \right]_{u=F_1(x), v=F_2(y)}$$

Definición 1.5

Sea F una función de distribución univariante. Llamamos inversa generalizada de F a

$$F^{-1}(t) = \inf \{x \in \mathbb{R} : F(x) \geq t\} \quad \forall t \in [0, 1]$$

(Consideramos $\inf \emptyset = -\infty$).

Corolario 1.1

Sea H una función distribución n -dimensional con marginales continuas F_1, \dots, F_n y cópula C (C satisface (1.1)). Entonces para cualquier $\mathbf{u} \in [0, 1]^n$

$$C(u_1, \dots, u_n) = H(F_1(u_1), \dots, F_n(u_n))$$

La separación entre comportamientos marginales y dependencia conjunta explica la flexibilidad de las cópulas para establecer modelos de dependencia con numerosas aplicaciones prácticas y teóricas. La primera parte del teorema de Sklar permite construir distribuciones bivariantes de forma sencilla y flexible:

uniendo una pareja de distribuciones marginales univariantes en una función que satisface la definición de cópula. Esto difiere de la manera tradicional de construir distribuciones multivariantes, que tiene la restricción de que las marginales son usualmente del mismo tipo, es decir, las correspondientes variables aleatorias son transformaciones lineales afines la una de la otra mientras que las cópulas admiten marginales de diferentes tipos. Esta peculiaridad sugiere que cualquier problema de estimación debería descomponerse en dos etapas: la primera relativa a las marginales y la segunda a la cópula misma.

1.3. Cotas de Fréchet-Hoeffding

Consideremos las funciones

$$\begin{aligned}M^n(\mathbf{u}) &= \min(u_1, \dots, u_n) \\ \Pi^n(\mathbf{u}) &= u_1 \times \dots \times u_n \\ W^n(\mathbf{u}) &= \max(u_1 + \dots + u_n - n + 1, 0)\end{aligned}$$

Las funciones M^n y Π^n son n -cópulas $\forall n \geq 2$, mientras que W^n en general no es una cópula para cualquier $n \geq 3$.

Teorema 1.3

(Desigualdad de las cotas de Fréchet-Hoeffding)

Sea C una n -cópula, entonces para cada $\mathbf{u} \in [0, 1]^n$

$$W^n(\mathbf{u}) \leq C(\mathbf{u}) \leq M^n(\mathbf{u})$$

(La demostración puede verse en Fréchet 1957).

A la vista de esta desigualdad podemos preguntarnos qué representa el hecho de que la cópula que une las variables aleatorias coincida con alguna de las dos cotas (nos referiremos al caso bidimensional, ya que para $n > 2$, W^n no es una cópula). El hecho de que la distribución conjunta de las dos v.a. esté caracterizada por la cota superior de Fréchet, M^2 , indica una situación de dependencia perfecta positiva, también conocida como *comonotonicidad*. Por el

contrario, si la cópula que las liga es la cota inferior de Fréchet, W^2 , se dice que son *contramonotónicas*.

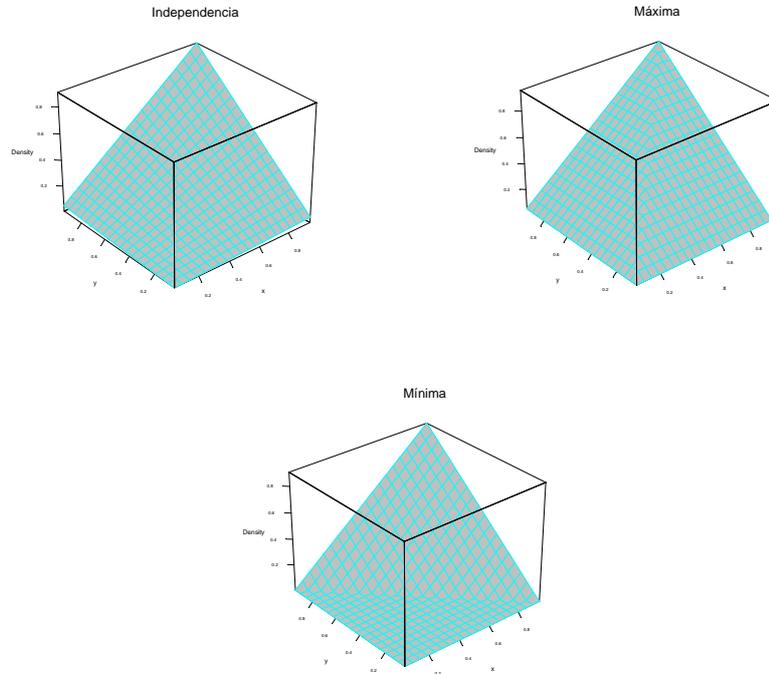


Figura 1.1: *Distribuciones de las Cópulas Independencia, Máxima y Mínima*

A pesar de que W^n no es una cópula para $n > 2$, es la mejor cota inferior posible en el sentido siguiente:

Teorema 1.4

Para cualquier $n > 2$ y $\mathbf{u} \in [0, 1]^n$, existe una n -cópula C (que depende de \mathbf{u}) tal que

$$C(\mathbf{u}) = W^n(\mathbf{u})$$

(La demostración puede verse en Nelsen 1999, pág.42)

Una consecuencia del Teorema de Sklar, es que la desigualdad que relaciona las cópulas máxima y mínima, W^2 y M^2 , puede escribirse como

$$\max(F_1(x) + F_2(y) - 1, 0) \leq F(x, y) \leq \min(F_1(x), F_2(y))$$

que es la conocida desigualdad de Fréchet-Hoeffding para funciones de distribución.

La existencia de cotas superior e inferior sugiere la posibilidad de establecer un orden entre las cópulas:

Definición 1.6

Sean C_1 y C_2 cópulas. Decimos que C_1 es más pequeña que C_2 (o C_2 más grande que C_1) y lo denotamos

$$C_1 \prec C_2 \quad (C_2 \succ C_1)$$

si

$$C_1(u, v) \leq C_2(u, v) \quad (\text{o } C_2(u, v) \geq C_1(u, v)) \quad \forall (u, v) \in [0, 1]^2$$

Este orden se denomina *orden de concordancia*.

Decimos que una familia paramétrica de cópulas $\{C_\theta, \theta \in \Theta\}$ está *ordenada positivamente* respecto al parámetro θ , si

$$\theta_1 \leq \theta_2 \Rightarrow C_{\theta_1} \prec C_{\theta_2} \quad \forall \theta_1, \theta_2 \in \Theta$$

Por el contrario, estará *ordenada negativamente* si

$$\theta_1 \leq \theta_2 \Rightarrow C_{\theta_1} \succ C_{\theta_2} \quad \forall \theta_1, \theta_2 \in \Theta$$

El orden así definido es solamente parcial, puesto que no todas las cópulas pueden compararse a través de él. Por ejemplo, si consideramos

$$C = \frac{1}{3}W + \frac{2}{3}M$$

(es inmediato comprobar que C es una cópula a partir de la definición 1.4 y del hecho de que W y M son cópulas) es posible encontrar puntos de $[0, 1]^2$ en los que $C \geq \Pi$ y otros en los que $C < \Pi$, en concreto

$$\begin{aligned} \frac{1}{3}\max(u+v-1, 0) + \frac{2}{3}\min(u, v) &> uv \quad \text{si } u = v = \frac{1}{2} \\ \frac{1}{3}\max(u+v-1, 0) + \frac{2}{3}\min(u, v) &< uv \quad \text{si } u = \frac{1}{4}, v = \frac{3}{4} \end{aligned}$$

con lo cual C y Π no son comparables.

1.4. Gráficos y diagramas de contorno de una cópula

El gráfico de una cópula, $z = C(u, v)$, es una superficie continua en el cubo unidad $[0, 1]^3$, acotada por el cuadrilátero de vértices $(0,0,0)$, $(1,0,0)$, $(1,1,1)$ y $(0,1,0)$.

Como sabemos, dada cualquier cópula C se verifica que:

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

por lo que el gráfico de C permanece entre los gráficos de las cotas de Fréchet-Hoeffding, es decir, las superficies $z = W(u, v)$ y $z = M(u, v)$.

Una manera sencilla de presentar y comparar el gráfico de una cópula, es a través de su *diagrama de contorno* que no es más que la representación gráfica de sus *curvas de nivel*. Éstas últimas son conjuntos en $[0, 1]^n$ dados por $C(u, v) = k$ con k constante. Obsérvese que los puntos $(k, 1)$ y $(1, k)$ pertenecen a la curva de nivel correspondiente a la constante k .

También se verifica que para cualquier cópula C y un $k \in [0, 1]$ dado, el gráfico de la curva de nivel $\{(u, v) \in [0, 1]^2 : C(u, v) = k\}$ debe permanecer dentro del triángulo determinado por $W(u, v) = k$ y $M(u, v) = k$.

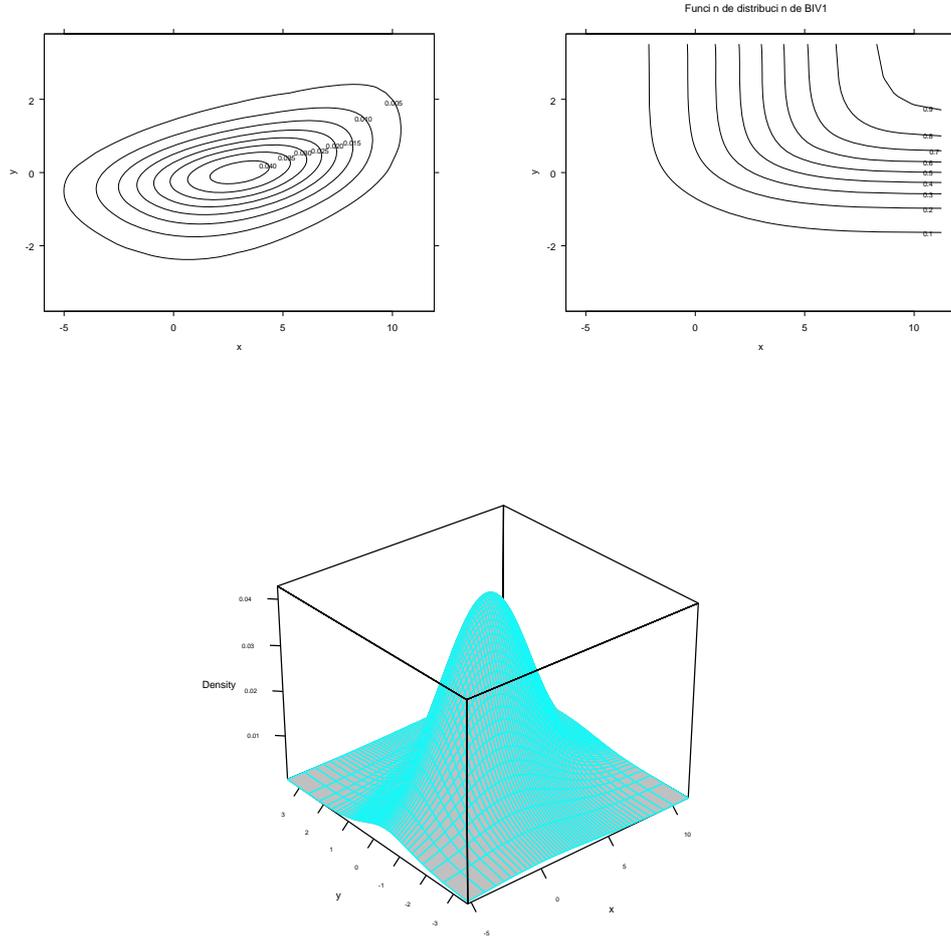


Figura 1.2: Densidad de la Cópula de Gumbel bivalente de parámetro 1.4 con marginales $N(3,4)$ y $t(3)$ (abajo) y diagramas de contorno de su densidad y su distribución (arriba)

1.5. Propiedades de las cópulas

Puesto que n v.a. X_1, \dots, X_n son independientes sii

$$H(x_1, \dots, x_n) = F_1(x_1) \times \dots \times F_n(x_n) \quad \forall x_1, \dots, x_n \in \overline{\mathbb{R}}$$

se puede establecer el siguiente resultado a partir del Teorema 1.2:

Teorema 1.5

Sea $(X_1, \dots, X_n)^T$ un vector de v.a. continuas con cópula C . Entonces X_1, \dots, X_n son independientes sii $C = \Pi^n$.

Una buena propiedad de las cópulas es que son invariantes bajo transformaciones estrictamente monótonas de las variables aleatorias, o en su caso se transforman de manera sencilla. Obsérvese que si la función de distribución de una v.a. X es continua y α es una función estrictamente monótona cuyo dominio contiene el $RanX$, entonces la función de distribución de la v.a. $\alpha(X)$ también es continua.

Proposición 1.1

Sea $(X_1, \dots, X_n)^T$ un vector de v.a. continuas con cópula C . Si $\alpha_1, \dots, \alpha_n$ son estrictamente crecientes sobre $RanX_1, \dots, RanX_n$, respectivamente, entonces $(\alpha_1(X_1), \dots, \alpha_n(X_n))^T$ también tiene cópula C .

Dem:

Sea F_i la función de distribución de X_i , $i = 1 \dots n$. Sea α_i^{-1} la inversa de α_i .

La variable transformada $\alpha_i(X_i)$ tiene función de distribución continua $\tilde{F} = F_i \circ \alpha_i^{-1}$. Recordemos que $F_i^{-1}(U) \sim F_i$ y $F_i^{-1}(u_i) = x_i$.

Por tanto

$$\begin{aligned} C(u_1, \dots, u_n) &= P[X_1 \leq F_1^{-1}(u_1), \dots, X_n \leq F_n^{-1}(u_n)] \\ &= P[\alpha_1(X_1) \leq \alpha_1 \circ F_1^{-1}(u_1), \dots, \alpha_n(X_n) \leq \alpha_n \circ F_n^{-1}(u_n)] \\ &= P[\alpha_1(X_1) \leq \tilde{F}_1^{-1}(u_1), \dots, \alpha_n(X_n) \leq \tilde{F}_n^{-1}(u_n)] \end{aligned}$$

□

De este resultado se siguen algunas consecuencias teóricas y prácticas de interés. Desde el punto de vista teórico, permite asegurar que cualquier funcional

o propiedad de la función de distribución conjunta que sea invariante bajo transformaciones estrictamente crecientes de las v.a., es un funcional o propiedad de su cópula (independientemente de sus funciones de distribución de probabilidad individuales), por tanto, resulta natural utilizar cualquier medida de dependencia para pares de v.a.: este es el núcleo de la definición de concordancia entre v.a. que se tratará con detalle en el punto 1.7.

Proposición 1.2

X e Y v.a. son comonotónicas sii

$$(X, Y) =_d (\alpha_1(Z), \alpha_2(Z))$$

para alguna variable Z y α_1 y α_2 funciones monótonas crecientes.

(Si X e Y son continuas se tiene el caso más simple de comonotonicidad $Y = \alpha(X)$ c.s. para alguna transformación creciente α).

Dem:

Sea $U \sim U(0, 1)$ y sea F la función de distribución conjunta de (X, Y) con marginales $F_1(x)$ y $F_2(y)$.

$$F(x, y) = P\{U \leq \min(F_1(x), F_2(y))\} = P\{U \leq F_1(x), U \leq F_2(y)\}$$

por tanto

$$(X, Y) =_d (F_1^{-1}(u), F_2^{-1}(u)) \tag{1.2}$$

F_1^{-1} y F_2^{-1} son transformaciones monótonas crecientes.

Inversamente, a partir de (1.2)

$$F(x, y) = P\{\alpha_1(Z) \leq x, \alpha_2(Z) \leq y\} = P\{Z \in A, z \in B\}$$

donde A y B son intervalos de la forma $(-\infty, k]$ o $(-\infty, k)$, por lo que $A \subset B$ o $B \subset A$. Es más

$$F(x, y) = \min\{P(Z \in A), P(Z \in B)\} = \min(F_1(x), F_2(y))$$

□

Proposición 1.3

X e Y v.a. son contramonotónicas sii

$$(X, Y) =_d (\alpha_1(Z), \alpha_2(Z))$$

para alguna variable Z y α_1 función monótona creciente y α_2 función monótona decreciente o viceversa.

Si X e Y son continuas se tiene el caso más simple de comonotonicidad $Y = \alpha(X)$ c.s. para alguna transformación decreciente α .

Demostración:

Sea $U \sim U(0, 1)$

$$(X, Y) =_d (F_1^{-1}(u), F_2^{-1}(1 - u)) = (X, Y) =_d (F_1^{-1}(u), F_2^{-1} \circ g(u))$$

donde $F_i^{-1}(q)$ es la función cuantílica de F_i y $g(x) = 1 - x$.

Entonces $u = F_1^{-1}$ es creciente y $v = F_2^{-1} \circ g$ es decreciente.

Para el inverso

$$(X, Y) = (u(z), v(z))$$

con u creciente y v decreciente.

Definimos

$$A = \{Z \in u^{-1}((-\infty, x])\}$$

$$B = \{Z \in v^{-1}((-\infty, y])\}$$

Si $A \cap B \neq \emptyset$, la monotonicidad de u y v implica que

$$P(A \cup B) = P(\Omega) = 1 = P(A) + P(B) - P(A \cap B)$$

y, por tanto

$$P(A \cup B) = P\{u(Z) \leq x, v(Z) \leq y\} = F_1(x) + F_2(y) - 1$$

Si $A \cap B = \emptyset$, entonces

$$F_1(x) + F_2(y) - 1 = 0$$

En cualquier caso

$$P(A \cup B) = P\{u(Z) \leq x, v(Z) \leq y\} = \max\{F_1(x) + F_2(y) - 1, 0\}$$

□

Desde el punto de vista aplicado, la propiedad de comonotonicidad junto con la invarianza, convierte a las cópulas en una potente herramienta para representar modelos de dependencia conjunta que permiten caracterizar de manera sencilla las situaciones extremas de independencia y dependencia perfecta (positiva o negativa).

Finalizaremos este punto con el siguiente resultado relativo a las derivadas parciales de una cópula:

Teorema 1.6

Sea C una cópula n -dimensional. Las derivadas parciales de orden k de C , $\frac{\partial^k}{\partial u_1 \dots \partial u_k} C(\mathbf{u})$, existen para casi todo $\mathbf{u} \in [0, 1]^n$.

Análogamente, para cada $\mathbf{u} \in [0, 1]^n$ $0 \leq \frac{\partial^k}{\partial u_1 \dots \partial u_k} C(\mathbf{u}) \leq 1$.

(El término *casi todo* se utiliza en el sentido de la medida de Lebesgue).

(Demostración y detalles en Nelsen 1999, pág.11)

Teniendo en cuenta lo anterior, dada una cópula :

$$C(u_1, \dots, u_n) = A_C(u_1, \dots, u_n) + S_C(u_1, \dots, u_n)$$

con

$$A_C(u_1, \dots, u_n) = \int_0^{u_1} \dots \int_0^{u_n} \frac{\partial^n}{\partial s_1 \dots \partial s_n} C(s_1, \dots, s_n) ds_1, \dots, ds_n$$

$$S_C(u_1, \dots, u_n) = C(u_1, \dots, u_n) - A_C(u_1, \dots, u_n)$$

Si $C = A_C$ en $[0, 1]^n$, se dice que C es *absolutamente continua*. En este caso C tiene *densidad*:

$$c(u_1, \dots, u_n) = \frac{\partial^n}{\partial u_1 \dots \partial u_n} C(u_1, \dots, u_n)$$

que por el teorema 1.1 sabemos que existe en el interior de $[0, 1]^n$ y es no negativa.

Si $C = S_C$ en $[0, 1]^n$, se dice que C es *singular* y

$$\frac{\partial^n}{\partial u_1 \dots \partial u_n} C(u_1, \dots, u_n) = 0 \quad \text{c.s. en } [0, 1]^n$$

El soporte de una cópula es el complementario de la unión de todos los subconjuntos abiertos $A \in [0, 1]^n$ con $V_C(A) = 0$. Cuando el soporte de C es $[0, 1]^n$, se dice que C tiene *soporte completo*. Si C es singular, su soporte tiene medida de Lebesgue cero y recíprocamente. Sin embargo, una cópula puede tener soporte completo con dos componentes, la absolutamente continua y la singular:

$$C = A_C + S_C$$

- La cópula producto Π es absolutamente continua:

$$\forall (u, v) \in [0, 1]^2 \quad A_C = \int_0^u \int_0^v \frac{\partial^2}{\partial u \partial t} utdudt = \int_0^u \int_0^v dudt = uv = \Pi$$

- La cota superior de Fréchet, W , es singular:

$$\forall (u, v) \in [0, 1]^2 \quad \frac{\partial^2}{\partial u \partial v} W(u, v) = 0$$

luego

$$A_C = \int_0^u \int_0^v \frac{\partial^2}{\partial u \partial v} W(u, v) dudt = 0 \neq W$$

- Análogamente para la cota inferior de Fréchet, M .

Para vectores aleatorios continuos, la densidad de la cópula se relaciona con la densidad f asociada a la función de distribución conjunta F . Concretamente, es igual a la razón entre la densidad conjunta y el producto de las densidades marginales:

$$c(F_1(x), F_2(y)) = \frac{f(x, y)}{f_1(x)f_2(y)}$$

ya que, por el Teorema de Sklar, podemos establecer la siguiente *representación canónica*:

De hecho, puesto que las transformaciones integrales son uniformes $U_1 = F_1(X)$ y $U_2 = F_2(X)$, se tiene que

$$X = F_1^{-1}(U_1) \quad Y = F_2^{-1}(U_2)$$

y dado que para variables aleatorias continuas estas transformaciones son estrictamente crecientes y continuas

$$c(u_1, u_2) = f(F_1^{-1}(u_1), F_2^{-1}(u_2)) \left| \begin{array}{cc} \frac{\partial X}{\partial u_1} & \frac{\partial X}{\partial u_2} \\ \frac{\partial Y}{\partial u_1} & \frac{\partial Y}{\partial u_2} \end{array} \right| = \frac{f(F_1^{-1}(u_1), F_2^{-1}(u_2))}{f_1(F_1^{-1}(u_1))f_2(F_2^{-1}(u_2))}$$

La representación canónica resulta útil cuando, para una distribución multivariante dada y marginales dadas, se quiere conocer la cópula que une esas marginales. También juega un papel fundamental en los procedimientos de estimación de cópulas.

1.6. Cópula de supervivencia y función de supervivencia conjunta

Un elemento clave en el control de riesgo de crédito es la llamada *Cópula de supervivencia*. Supongamos dos v.a. X e Y que representan, respectivamente, los tiempos de supervivencia de dos riesgos. Las funciones de supervivencia, que proporcionan la probabilidad de supervivencia individual más allá de un tiempo x e y respectivamente, vienen dadas por:

$$\begin{aligned} \bar{F}_1(x) &= P(X > x) = 1 - F_1(x) \\ \bar{F}_2(y) &= P(Y > y) = 1 - F_2(y) \end{aligned}$$

La función de supervivencia conjunta viene dada por:

$$\bar{H}(x, y) = P[X > x, Y > y]$$

cuyas marginales son $\bar{H}(x, -\infty)$ y $\bar{H}(-\infty, y)$, que son las funciones de supervivencia univariantes \bar{F}_1 y \bar{F}_2 .

Parece natural preguntarse si es posible establecer una relación análoga a la del Teorema de Sklar, pero entre funciones de supervivencia conjuntas y marginales. La respuesta viene dada a través de la cópula C de X e Y :

$$\overline{H}(x, y) = 1 - F_1(x) - F_2(y) + H(x, y) = \overline{F}_1(x) + \overline{F}_2(y) - 1 + C(F_1(x), F_2(y))$$

a partir de lo cual, podemos definir una función $\overline{C} : [0, 1]^2 \rightarrow [0, 1]$ tal que

$$\overline{C}(u, v) = u + v - 1 + C(1 - u, 1 - v)$$

con lo que tenemos

$$\overline{H}(x, y) = \overline{C}(\overline{F}_1(x), \overline{F}_2(y))$$

La función \overline{C} es una cópula y recibe el nombre de *Cópula de supervivencia* de X e Y . No debe confundirse con la función de supervivencia conjunta \widehat{C} de n v.a con función de distribución conjunta C , es decir, si $(u_1, \dots, u_n)^T$ tiene función de distribución C , entonces:

$$\widehat{C}(u_1, \dots, u_n) = P(U_1 > u_1, \dots, U_n > u_n) = \overline{C}(1 - u_1, \dots, 1 - u_n)$$

en el caso bidimensional

$$\widehat{C}(u, v) = 1 - u - v + C(u, v) = \overline{C}(1 - u, 1 - v)$$

Obsérvese que la función de distribución conjunta de dos transformaciones monótonas decrecientes de dos v.a. dadas viene representada a través de su cópula de supervivencia.

También es posible expresar, vía cópula de supervivencia, la probabilidad condicionada

$$P(U_1 > u / U_2 > v) = \frac{1 - u - v + C(u, v)}{1 - v} = \frac{\overline{C}(1 - u, 1 - v)}{1 - v}$$

y, por tanto

$$P(X > x / Y > y) = \frac{\overline{C}(\overline{F}_1(x), \overline{F}_2(y))}{\overline{F}_2(y)}$$

La *densidad de la cópula de supervivencia* se define como

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

que existe c.s. y es tal que $\bar{c}(u, v) = c(1 - u, 1 - v)$. Esto está relacionado con la densidad de la distribución de supervivencia \bar{F} , de la siguiente manera:

$$f(x, y) = \bar{c}(1 - F_1(x), 1 - F_2(y)) f_1(x) f_2(y)$$

así que

$$c(F_1(x), F_2(y)) = \bar{c}(1 - F_1(x), 1 - F_2(y))$$

Reviste un interés particular la función de supervivencia conjunta de dos uniformes. Sean U_1 y U_2 dos v.a $U(0, 1)$. con cópula C . Su función de supervivencia conjunta, denotada por \mathbb{C} , representa la probabilidad conjunta de que (U_1, U_2) sean mayores que u y v , respectivamente:

$$\mathbb{C}(u, v) = P(\mathbb{U}_1 > u, \mathbb{U}_2 > v)$$

A partir de lo anterior

$$\mathbb{C}(u, v) = 1 - u - v + C(u, v) = \bar{\mathbb{C}}(1 - u, 1 - v)$$

La probabilidad de supervivencia $\bar{F}(x, y)$ puede escribirse en términos de la función de supervivencia conjunta para uniformes como:

$$\bar{F}(x, y) = \mathbb{C}(F_1(x), F_2(y))$$

Hay diferentes conceptos que pueden expresarse en términos de cópulas, cópulas de supervivencia y funciones de supervivencia de $U(0, 1)$. Por ejemplo, la función de distribución de probabilidad para mínimos o máximos de dos v.a. se puede expresar de forma sencilla en términos de cópulas:

Sean $m = \min(X, Y)$, $M = \max(X, Y)$ y F_m y F_M las funciones de distribución del máximo y el mínimo, respectivamente.

Para el máximo, tenemos:

$$F_M(a) = P(M \leq a) = P(X \leq a, Y \leq a) = F(a, a) = C(F_1(a), F_2(a))$$

y para el mínimo:

$$\begin{aligned} F_m(a) &= P(m \leq a) = 1 - P(m > a) = 1 - P(X > a, Y > a) \\ &= 1 - \mathbb{C}(F_1(a), F_2(a)) = 1 - \overline{\mathbb{C}}(\overline{F}_1(a), \overline{F}_2(a)) \end{aligned}$$

donde, obviamente, el punto a debe ser elegido perteneciente al $RanF_1$ y $RanF_2$.

1.7. Limitaciones del coeficiente de correlación lineal. Medidas de concordancia

Comenzaremos estableciendo el significado de “independencia” en términos de cópulas:

X_1, \dots, X_n v.a son *independientes* si la cópula que une sus marginales es la cópula producto.

$$C(F_1(x_1), \dots, F_n(x_n)) = F_1(x_1)F_2(x_2)\dots F_n(x_n) = \Pi^n$$

Podemos, entonces, considerar medidas de dependencia que recojan la distancia entre la cópula empírica y la cópula producto. Veamos, en primer lugar, qué propiedades serían deseables para una medida de dependencia. Consideremos el caso bidimensional (X, Y) ; ∂ medida de dependencia debe cumplir:

1. Está definida para cada par (X, Y) .
2. $\partial(X, Y) = \partial(Y, X)$ (simetría)
3. $-1 \leq \partial(X, Y) \leq 1$
4. $\partial(X, Y) = 0$ sii X e Y son independientes
5. $\partial(X, Y) = 1$ sii X e Y son función estrictamente monótona no decreciente c.s la una de la otra (*comonotonicidad*).
6. $\partial(X, Y) = -1$ sii X e Y son función estrictamente monótona no creciente c.s la una de la otra (*contramonotonicidad*).
6. $\partial(X, Y) = \partial(T_1(X), T_2(Y))$ con T_1 y T_2 funciones estrictamente monótonas c.s.

7. Si la sucesión de v.a. $\{(X_n, Y_n)\}$ tiene cópula C_n con $C_n \xrightarrow{n \rightarrow \infty} C$ entonces $\partial_{C_n} \xrightarrow{n \rightarrow \infty} \partial_C$.

Recordemos la definición estándar del coeficiente de correlación de Pearson:

$$\rho = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sigma_X \sigma_Y}$$

ρ solamente cumple las propiedades 2 y 3 de entre las siete anteriores. El éxito de la correlación lineal se debe a sus buenas propiedades en el análisis lineal, especialmente para la combinación lineal de factores de riesgo F , donde

$$VaR(\alpha^T F) = \alpha^T Cov(F) \alpha$$

Sin embargo, su uso no es adecuado en situaciones en las que la dependencia multivariante es de tipo más general.

Alguno de los inconvenientes más notables que presenta son los siguientes:

- σ_X^2 y σ_Y^2 deben ser finitas para que ρ esté definida.
- La independencia siempre implica correlación cero, pero el recíproco sólo es cierto en el caso multivariante gaussiano.
- Correlaciones débiles no implican necesariamente baja dependencia.
- La correlación no es una medida invariante bajo transformaciones estrictamente monótonas (la cópula sí lo es).

Necesitamos, entonces, alguna otra medida de dependencia que no presente estos inconvenientes y, puesto que la cópula captura aquellas propiedades de la distribución conjunta que son invariantes bajo transformaciones estrictamente crecientes c.s., parece conveniente expresar dichas medidas en términos de la cópula.

Las medidas invariantes por cambios de escala utilizadas más frecuentemente son las llamadas *medidas de concordancia*: decimos que dos v.a son *concordantes* si valores grandes de una están asociados con valores grandes de la otra y lo mismo ocurre con los valores pequeños. En otro caso diremos que son *discordantes*. De

esta forma, la concordancia recoge aquellas relaciones no lineales que escapan a la correlación.

A continuación, recordamos las definiciones de las medidas de concordancia más utilizadas y proporcionamos nuevas expresiones para ellas en términos de cópulas:

- *Tau de Kendall* (se define como la probabilidad de concordancia menos la de discordancia)

Sean (X_1, Y_1) y (X_2, Y_2) dos vectores aleatorios iid con funciones de distribución conjuntas H_1 y H_2 y posibles cópulas C_1 y C_2 ,

$$\tau = [P(X_1 - X_2)(Y_1 - Y_2) > 0] - [P(X_1 - X_2)(Y_1 - Y_2) < 0]$$

Nelsen (1999) demuestra que puede expresarse en términos de las cópulas como:

$$\tau_C = 4 \int \int_{I^2} C_2(u, v) dC_1(u, v) - 1$$

- *Rho de Spearman*

Sea R_i el rango de x_i entre las x y S_i el rango de y_i entre las y . El coeficiente de correlación de rangos de Spearman es:

$$\rho_S = \frac{\sum_{i=1}^n (R_i - \bar{R})(S_i - \bar{S})}{\sqrt{\sum_{i=1}^n (R_i - \bar{R})^2} \sqrt{\sum_{i=1}^n (S_i - \bar{S})^2}}$$

que en términos de cópulas queda como:

$$\rho_C = 12 \int \int_{I^2} (C(u, v) - uv) dudv = 12 \int \int_{I^2} C(u, v) dudv - 3$$

(Obsérvese la comparación explícita con la cópula producto en este caso).

Esencialmente, estas dos medidas miden el grado de dependencia monotónica frente al coeficiente de correlación lineal, que mide simplemente el grado de dependencia lineal. Ambas medidas alcanzan el valor de la unidad para la cota

superior de Fréchet bivalente (una variable es transformación estrictamente creciente de la otra) y el valor -1 para la cota inferior de Fréchet (una variable es transformación estrictamente decreciente de la otra), lo cual no es cierto de forma general en el caso del coeficiente de correlación lineal.

Existen otras situaciones en las cuales el interés simplemente radica en cuando las variables son *concordantes* o *discordantes*, es decir, cuando la dependencia entre X e Y es positiva o negativa. En esas situaciones recurriremos a las siguientes definiciones.

Definición 1.7

Se dice que X e Y v.a. tienen dependencia positiva en cuadrante (PQD) si

$$P[X > x, Y > y] \geq P[X > x] P[Y > y] \quad \forall x, y \in \overline{\mathbb{R}}$$

La definición anterior es equivalente a

$$P[X \leq x, Y \leq y] \geq P[X \leq x] P[Y \leq y] \quad \forall x, y \in \overline{\mathbb{R}}$$

Definición 1.8

Se dice que X e Y v.a. están positivamente asociadas (PA) si

$$E[g_1(X, Y)g_2(X, Y)] \geq E[g_1(X, Y)] E[g_2(X, Y)]$$

siendo g_1 y g_2 funciones medibles de variable real, crecientes en ambas componentes y para las que las esperanzas anteriores existen.

(Para ampliar conceptos de dependencia ver Joe 1997).

PQD y PA son invariantes bajo transformaciones crecientes y verifican lo siguiente:

$$\text{Comonotonicidad} \Rightarrow \text{PA} \Rightarrow \text{PQD} \Rightarrow \rho_{X,Y} \geq 0, \tau_{X,Y} \geq 0, \rho_S(X, Y) \geq 0$$

Por tanto, la comonotonicidad es el tipo más fuerte de concordancia o dependencia positiva.

Para más detalles sobre tópicos relacionados con el uso del coeficiente de correlación lineal en el control de riesgos financieros, puede consultarse Embrechts, McNeil y Straumann (1999).

1.8. Coeficiente de dependencia en colas (TDC)

El concepto de dependencia en colas resulta relevante para el estudio de la dependencia entre valores extremos, lo cual es especialmente útil en la determinación de medidas de riesgo, tales como el VaR, que se concentran en el comportamiento de sucesos extremos. La dependencia en colas entre dos variables aleatorias continuas X e Y se puede plantear como una propiedad de la cópula y no de las distribuciones marginales, por lo que su cuantificación es invariante bajo transformaciones estrictamente crecientes de X e Y .

Con este fin, se define la siguiente medida empírica para extremos:

Definición 1.9

Sea $(X, Y)^T$ un vector de v.a. continuas con funciones de distribución marginales F_1 y F_2 . El coeficiente de dependencia en la cola superior de $(X, Y)^T$ es

$$\lim_{u \rightarrow 1^-} P \{Y > F_2^{-1}(u) / X > F_1^{-1}(u)\} = \lambda_u$$

siempre que $\lambda_u \in [0, 1]$ exista.

Si $\lambda_u \in (0, 1]$, decimos que X e Y son *dependientes asintóticamente en la cola derecha*.

Si $\lambda_u = 0$, decimos que X e Y son *asintóticamente independientes en la cola derecha*.

Puesto que

$$\begin{aligned} & P \{Y > F_2^{-1}(u) / X > F_1^{-1}(u)\} \\ = & \frac{1 - P \{X \leq F_1^{-1}(u)\} - P \{Y \leq F_2^{-1}(u)\} + P \{X \leq F_1^{-1}(u), Y \leq F_2^{-1}(u)\}}{1 - P \{X \leq F_1^{-1}(u)\}} \end{aligned}$$

Una definición equivalente (para v.a. continuas), que evidencia que el concepto de dependencia en colas es una propiedad de la cópula es la siguiente (Joe 1997, pág.33):

Definición 1.10

Sea C una cópula bivalente tal que $\lim_{u \rightarrow 1^-} \frac{(1-2u+C(u,u))}{(1-u)} = \lambda_u$ existe, entonces C tiene dependencia en la cola superior si $\lambda_u \in (0, 1]$ e independencia en la cola superior si $\lambda_u = 0$.

Podemos establecer estas definiciones en términos del VaR (Valor en Riesgo):

$$\begin{aligned}\lambda_u &= \lim_{\alpha \rightarrow 1^-} P \{Y > VaR_\alpha(Y) / X > VaR_\alpha(X)\} \\ &= \lim_{\alpha \rightarrow 1^-} \frac{1 - 2\alpha + C(\alpha, \alpha)}{1 - \alpha} = \lim_{\alpha \rightarrow 1^-} \frac{\widehat{C}(\alpha, \alpha)}{1 - \alpha}\end{aligned}\quad (1.3)$$

Para cópulas cuya expresión es complicada (por ejemplo la gaussiana), existe una expresión más manejable de λ_u (Embretchs, McNeil y Straumann 1999):

Sean (U, V) un par de v.a. $U(0, 1)$ con cópula C . Obsérvese que

$$\begin{aligned}P \{V \leq v / U = u\} &= \frac{\partial}{\partial u} C(u, v) \\ P \{V > v / U = u\} &= 1 - \frac{\partial}{\partial u} C(u, v)\end{aligned}$$

y análogamente si condicionamos por $V=v$.

Entonces:

$$\begin{aligned}\lambda_u &= \lim_{u \rightarrow 1^-} \frac{\widehat{C}(u, u)}{(1 - u)} = -\lim_{u \rightarrow 1^-} \frac{d}{du} \widehat{C}(u, u) \\ &= -\lim_{u \rightarrow 1^-} \left(-2 + \left[\frac{\partial}{\partial s} C(s, t) \right]_{s=t=u} + \left[\frac{\partial}{\partial t} C(s, t) \right]_{s=t=u} \right) \\ &= \lim_{u \rightarrow 1^-} (P \{V > u / U = u\} + P \{U > u / V = u\})\end{aligned}$$

Si además C es una cópula intercambiable, $C(u, v) = C(v, u)$, esta expresión se simplifica quedando como:

$$\lambda_u = 2 \lim_{u \rightarrow 1^-} P \{V > u / U = u\}$$

El concepto de dependencia en la cola inferior puede definirse de manera similar:

Si $\lim_{u \rightarrow 0^+} \frac{C(u,u)}{u} = \lambda_L$ existe, entonces C tiene *dependencia en la cola inferior* si $\lambda_L \in (0, 1]$ e *independencia en la cola inferior* si $\lambda_L = 0$.

De nuevo, existe una formulación alternativa útil en el caso de cópulas con expresiones complicadas:

$$\begin{aligned} \lambda_L &= \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u} = \lim_{u \rightarrow 0^+} \frac{dC(u, u)}{du} \\ &= \lim_{u \rightarrow 0^+} (P\{V < u/U = u\} + P\{U < u/V = u\}) \end{aligned} \quad (1.4)$$

Si además C es una cópula intercambiable:

$$\lambda_L = 2 \lim_{u \rightarrow 0^+} P\{V < u/U = u\}$$

Recordemos que :

$$\bar{C}(u, v) = 1 - u - v + C(u, v) = \hat{C}(1 - u, 1 - v)$$

por tanto:

$$\lim_{u \rightarrow 1^-} \frac{\hat{C}(u, u)}{(1 - u)} = \lim_{u \rightarrow 1^-} \frac{\bar{C}(1 - u, 1 - u)}{(1 - u)} = \lim_{u \rightarrow 0^+} \frac{\bar{C}(u, u)}{u}$$

por lo que el coeficiente de dependencia en la cola superior de C es el coeficiente de dependencia en la cola inferior de \bar{C} . Análogamente, el coeficiente de dependencia en la cola inferior de C es el coeficiente de dependencia en la cola superior de \bar{C} .

Proposición 1.4

Si X e Y tienen funciones de distribución continuas F_1 y F_2 respectivamente, entonces

$$\begin{aligned} \lambda_L &= \lim_{u \rightarrow 0^+} P(Y \leq F_2^{-1}(u) / X \leq F_1^{-1}(u)) \\ \lambda_U &= \lim_{u \rightarrow 1^-} P(Y > F_2^{-1}(u) / X > F_1^{-1}(u)) \end{aligned}$$

si existe el límite correspondiente.

Dem: Es inmediata al aplicar la transformación cuantílica a las definiciones de λ_L y λ_U .

1.9. Cópula empírica

Por último, trataremos una estimación empírica no paramétrica para cópulas introducida por Deheuvels (1979).

Definición 1.11

Sea $\{(x_k, y_k)\}_{k=1}^n$ una muestra de tamaño n obtenida a partir de una distribución bivalente. La Cópula empírica es la función C_n dada por

$$\begin{aligned} C_n\left(\frac{i}{n}, \frac{j}{n}\right) &= \frac{\text{n}^\circ \text{ de pares } (x, y) \text{ de la muestra tales que } x \leq x_{(i)} \text{ e } y \leq y_{(j)}}{n} \\ &= \frac{1}{n} \sum_{k=1}^n I(x_k \leq x_{(i)}, y_k \leq y_{(j)}) \quad i, j = 1 \dots n \end{aligned} \quad (1.5)$$

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

Definición 1.12

La frecuencia de la cópula empírica, c_n , viene dada por

$$c_n\left(\frac{i}{n}, \frac{j}{n}\right) = \begin{cases} \frac{1}{n} & \text{si } (x_{(i)}, y_{(j)}) \text{ es un elemento de la muestra} \\ 0 & \text{en otro caso} \end{cases}$$

La relación entre C_n y c_n es la siguiente:

$$C_n\left(\frac{i}{n}, \frac{j}{n}\right) = \sum_{p=1}^i \sum_{q=1}^j c_n\left(\frac{p}{n}, \frac{q}{n}\right)$$

e inversamente

$$c_n\left(\frac{i}{n}, \frac{j}{n}\right) = C_n\left(\frac{i}{n}, \frac{j}{n}\right) - C_n\left(\frac{i-1}{n}, \frac{j}{n}\right) - C_n\left(\frac{i}{n}, \frac{j-1}{n}\right) + C_n\left(\frac{i-1}{n}, \frac{j-1}{n}\right)$$

A partir de estas definiciones podemos obtener las versiones muestrales de las medidas de concordancia poblacionales, rho de Spearman (r_n) y tau de Kendall (t_n):

$$r_n = \frac{12}{n^2 - 1} \sum_{i=1}^n \sum_{j=1}^n \left[C_n \left(\frac{i}{n}, \frac{j}{n} \right) - \frac{i}{n} \cdot \frac{j}{n} \right]$$

$$t_n = \frac{2n}{n-1} \sum_{i=2}^n \sum_{j=2}^n \sum_{p=1}^{i-1} \sum_{q=1}^{j-1} \left[c_n \left(\frac{i}{n}, \frac{j}{n} \right) c_n \left(\frac{p}{n}, \frac{q}{n} \right) - c_n \left(\frac{i}{n}, \frac{q}{n} \right) c_n \left(\frac{p}{n}, \frac{j}{n} \right) \right]$$

(La comprobación puede verse en Nelsen 1999).

Deheuvels (1979) también demostró que cuando el tamaño muestral crece, la cópula empírica converge a la cópula verdadera.

Las cópulas empíricas permiten construir contrastes no paramétricos de independencia (Deheuvels 1981a y 81b, Genest y Rémillard 2004) además de ser un instrumento útil en el análisis exploratorio de los datos.

1.10. Familias de cópulas más usuales en el entorno financiero

A continuación, describimos las familias paramétricas de cópulas de uso más frecuente en el entorno financiero.

1.10.1. Cópulas elípticas

Las distribuciones elípticas juegan un papel importante en finanzas, debido a que constituyen una fuente de distribuciones multivariantes que poseen muchas de las buenas propiedades de la distribución normal multivariante y facilitan la obtención de modelos multivariantes para extremos y otras formas de dependencia no normales. Las cópulas elípticas no son más que las cópulas de distribuciones elípticas. La simulación a partir de este tipo de distribuciones es sencilla y, como consecuencia del Teorema de Sklar, también lo es a partir de cópulas elípticas. Además permiten calcular fácilmente los coeficientes de correlación de rangos y de dependencia en colas.

Definición 1.13

Sea X un vector aleatorio n -dimensional, $\mu \in \mathbb{R}$ y Σ una matriz simétrica de dimensión $n \times n$ y definida no negativa. Decimos que X tiene una distribución elíptica de parámetros μ , Σ y ϕ si dada la función característica de $X - \mu$ es de la forma

$$\varphi_{X-\mu}(t) = \phi(t^T \Sigma t)$$

Lo denotaremos $X \sim E_n(\mu, \Sigma, \phi)$ y ϕ recibe el nombre de *generador característico*.

Si $n = 1$, la clase de distribuciones elípticas coincide con la clase de distribuciones simétricas unidimensionales. Si $X \sim E_n(\mu, \Sigma, \phi)$, donde Σ es una matriz diagonal, entonces X tiene componentes incorreladas (si $0 < V(X_i) < +\infty$). Si X tiene componentes independientes, entonces $X \sim N_n(\mu, \Sigma)$. Obsérvese que la distribución normal multivariante es la única entre las distribuciones elípticas en la que componentes incorreladas implica componentes independientes. Un problema práctico en riesgo multivariante utilizando distribuciones elípticas es aquel en que todas las marginales son del mismo tipo.

Definición 1.14

Sea R una matriz simétrica, definida positiva con $\text{diag} R = 1$ y Φ_R^n la función de distribución normal n -variante estandarizada. La cópula gaussiana (MVN) se define como:

$$C_R(u_1, \dots, u_n) = \Phi_R^n(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n))$$

La función de densidad correspondiente es:

$$c_R(u_1, \dots, u_n) = \frac{1}{|R|^{1/2}} \exp\left(-\frac{1}{2} \varsigma^T (R^{-1} - I) \varsigma\right)$$

con $\varsigma_n = \Phi^{-1}(u_n)$.

En el caso bivalente

$$C_R(u, v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi(1-R_{12}^2)^{1/2}} \exp\left\{-\frac{s^2 - 2R_{12}st + t^2}{2(1-R_{12}^2)}\right\} ds dt$$

R_{12} es sencillamente el coeficiente de correlación lineal de la correspondiente distribución normal bivalente.

Las cópulas gaussianas no tienen dependencia en colas (ni superior ni inferior), es decir, $\lambda_u = \lambda_L = 0$ para $R_{12} < 1$. Esta cópula ha sido la única utilizable durante mucho tiempo; de hecho la metodología RiskMetrics de J.P. Morgan la usa hace varios años en gestión de riesgos de carteras mediante simulaciones Monte Carlo.

Existe otra cópula dentro de la familia elíptica que reviste interés, es la llamada *t-cópula*.

Si X tiene la representación estocástica

$$X =_d \mu + \frac{\sqrt{v}}{\sqrt{S}}Z$$

donde $\mu \in \mathbb{R}, S \sim \chi_v^2$ y $Z \sim N_n(\mathbf{0}, \Sigma)$ son independientes, entonces X tiene una distribución *t de Student n-variante* de media μ (para $v > 1$) y v grados de libertad.

Su matriz de covarianzas es $\frac{v}{v-2}\Sigma$ (para $v > 2$). Si $v \leq 2$ entonces $Cov(X)$ no está definida. En este caso interpretamos Σ como el parámetro de forma de la distribución de X .

Definición 1.15

Sea R una matriz simétrica, definida positiva con $diagR = 1$ y X el vector aleatorio definido anteriormente. La cópula de Student multivariante (MVT) se define como:

$$C_{v,R}^t(u_1, \dots, u_n) = t_{v,R}^n(t_v^{-1}(u_1), \dots, t_v^{-1}(u_n))$$

donde

$$R_{ij} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}} \quad \forall i, j \in \{1, \dots, n\}$$

y $t_{v,R}^n$ denota la función de distribución de $\frac{\sqrt{v}}{\sqrt{S}}Y$ con $S \sim \chi_v^2$ e $Y \sim N_n(\mathbf{0}, R)$ independientes. t_v denota las marginales (iguales) de $t_{v,R}^n$, es decir, la función de distribución de $\frac{\sqrt{v}}{\sqrt{S}}Y_1$.

En el caso bivalente la expresión de la cópula es:

$$C_{v,R}^t(u,v) = \int_{-\infty}^{t_v^{-1}(u)} \int_{-\infty}^{t_v^{-1}(v)} \frac{1}{2\pi(1-R_{12}^2)^{1/2}} \left\{ 1 + \frac{s^2 - 2R_{12}st + t^2}{v(1-R_{12}^2)} \right\}^{-(v+2)/2} ds dt$$

Obsérvese que R_{12} es sencillamente el coeficiente de correlación lineal usual de la correspondiente distribución t_v bivalente si $v > 2$. En este caso, el coeficiente de dependencia en la cola superior (y debido a la simetría radial también el de la cola inferior) viene dado por:

$$\lambda_u = 2 \lim_{x \rightarrow \infty} P(X_2 > x / X_1 = x) = 2\bar{t}_{v+1} \left(\sqrt{v+1} \sqrt{1-R_{12}} / \sqrt{1+R_{12}} \right)$$

Este coeficiente es creciente en R_{12} y decreciente en v , como era de esperar. Es más, tiende a cero cuando $v \rightarrow \infty$ para $R_{12} < 1$ e implica extremos asintóticamente dependientes para $R_{12} = -1$.

1.10.2. Cópulas Arquimedianas

Muchas de las familias paramétricas de cópulas interesantes pertenecen a la llamada clase de cópulas Arquimedianas que capturan una gran variedad de estructuras de dependencia. Como veremos, la representación arquimediana de cópulas permite reducir el estudio de una cópula multivalente a una única función univalente.

Definición 1.16

Se dice que una cópula C es Arquimediana si es de la forma:

$$C(u_1, \dots, u_n) = \begin{cases} \varphi^{-1}(\varphi(u_1) + \dots + \varphi(u_n)) & \text{si } \sum_{i=1}^n \varphi(u_i) \leq \varphi(0) \\ 0 & \text{en otro caso} \end{cases}$$

con $\varphi : [0, 1] \rightarrow [0, \infty]$ una función continua, estrictamente decreciente en $[0, \varphi(0)]$ con $\varphi(1) = 0$, $\varphi'(u) < 0$ y $\varphi''(u) > 0 \forall 0 \leq u \leq 1$.

$\varphi(u)$ se denomina *generador de la cópula*. Si $\varphi(0) = \infty$, decimos que φ es un *generador estricto* (puesto que un gran número de resultados requieren que φ tenga inversa en $(0, \infty)$, solamente consideraremos generadores estrictos).

Estas cópulas poseen algunas propiedades algebraicas elementales, como son las siguientes:

Sea C una cópula arquimediana con generador φ , entonces

- (i) C es simétrica, es decir, $C(u, v) = C(v, u) \forall u, v \in [0, 1]$.
- (ii) C es asociativa, es decir, $C(C(u, v), w) = C(u, C(v, w)) \forall u, v, w \in [0, 1]$.
- (iii) $k\varphi$ es también un generador de C , con k cualquier constante positiva
- (iv) Las distribuciones marginales de U y V son $U(0, 1)$

$$\begin{aligned} C(u, 1) &= u \quad \forall u \in [0, 1] \\ C(1, v) &= v \quad \forall v \in [0, 1] \end{aligned}$$

- (v) Sus curvas de nivel son convexas.

(La demostración puede verse en Nelsen 1999).

Adicionalmente, tenemos que el uso de cópulas arquimediana simplifica mucho los cálculos, por ejemplo, la tau de Kendall viene dada por:

$$\tau_C = 1 + 4 \int_0^1 \frac{\varphi(u)}{\varphi'(u)} du$$

y la dependencia en colas se puede expresar en términos de los generadores.

Teorema 1.7

Sea φ un generador estricto tal que φ^{-1} pertenece a la clase de transformadas de Laplace de variables aleatorias estrictamente positivas. Si $\varphi^{-1}(0)$ es finita, entonces $C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v))$ no tiene dependencia en la cola superior. Si C tiene dependencia en la cola superior, entonces $\varphi^{-1}(0) = -\infty$ y el coeficiente de dependencia en la cola superior viene dado por:

$$\lambda_u = 2 - 2 \lim_{s \rightarrow 0^+} \left[\frac{\varphi^{-1}(2s)}{\varphi^{-1}(s)} \right]$$

(La demostración puede verse en Joe 1997, pág.103)

Los dos resultados siguientes resultan necesarios para la aplicación del método de calibración de cópulas basado en medidas de dependencia muestrales propuesto en el punto 3.11.

Teorema 1.8

Sea una cópula arquimediana con generador φ en Ω . Sea $K_C(t)$ la C -medida del conjunto

$$\{(u, v) \in [0, 1]^2 : C(u, v) \leq t\}$$

o, equivalentemente, del conjunto

$$\{(u, v) \in [0, 1]^2 : \varphi(u) + \varphi(v) \geq \varphi(t)\}$$

Entonces, para cualquier $t \in [0, 1]$

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

(Demostración en Nelsen 1999, pág. 102).

Este resultado permite interpretar, de forma natural, la C -medida de una región por encima o por debajo de una curva de nivel como una función de distribución. La siguiente proposición presenta una interpretación probabilística del Teorema 1.8.

Proposición 1.5

Sean U y V v.a. cuya función de distribución conjunta es la cópula arquimediana C generada por φ en Ω . Entonces la función dada por el Teorema 1.8 es la función de distribución de la v.a. $C(u, v)$.

Algunos ejemplos de familias de cópulas arquimedianas más utilizadas en el entorno financiero son las siguientes:

Familia Gumbel

Sea $\varphi(t) = (-\ln t)^\theta, \theta \geq 1$.

Claramente $\varphi(t)$ es continua, $\varphi(1) = 0$, $\varphi'(t) = -\theta(-\ln t)^{\theta-1} \frac{1}{t}$, por tanto φ es estrictamente decreciente en $[0, \infty]$.

$\varphi''(t) \geq 0$ en $[0, 1]$, por lo que φ es convexa. Como $\varphi(0) = \infty$, φ es un generador estricto.

La expresión general en el caso bidimensional de esta familia es

$$C_\theta(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v)) = \exp\left(-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{1/\theta}\right) \quad (1.6)$$

con función de densidad

$$\begin{aligned} c_\theta(u, v) &= (-\ln u)^{\theta-1} \left[-1 + \theta + (-\ln u)^\theta + (-\ln v)^\theta\right]^{\frac{1}{\theta}} \\ &\quad \left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{\left(\frac{1}{\theta}-2\right)} (-\ln v)^{(\theta-1)} \\ &\quad / \exp\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{1/\theta} uv \end{aligned}$$

Además $C_1 = \Pi$ y $\lim_{\theta \rightarrow \infty} C_\theta = M$ (recordemos que $\Pi(u, v) = uv$ y $M(u, v) = \min(u, v)$).

En cuanto a la dependencia en colas, puesto que $\varphi^{-1}(s) = \exp(-s^{1/\theta})$ y

$$\varphi^{-1'}(s) = -\exp(-s^{1/\theta})/\theta$$

según el Teorema 1.7 el coeficiente de dependencia en la cola superior viene dado por:

$$\begin{aligned} \lambda_u &= 2 - 2 \lim_{s \rightarrow 0^+} \left[\frac{\varphi^{-1'}(2s)}{\varphi^{-1'}(s)} \right] \\ &= 2 - 2^{1/\theta} \lim_{s \rightarrow 0^+} \left[\frac{\exp(-(2s)^{1/\theta})}{\exp(-s^{1/\theta})} \right] = 2 - 2^{1/\theta} \end{aligned}$$

La expresión de la tau de Kendall para esta familia es:

$$\tau_\theta = 1 + 4 \int_0^1 \frac{t \ln t}{\theta} dt = 1 - \frac{1}{\theta}$$

Familia Clayton

Sea $\varphi(t) = \frac{(t^{-\theta}-1)}{\theta}$, $\theta \in [-1, \infty) \setminus \{0\}$.

Las cópulas bidimensionales pertenecientes a la familia de Clayton tienen la expresión:

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

con función de densidad

$$c_{\theta}(u, v) = (1 + \theta) u^{(-1-\theta)} v^{(-1-\theta)} \left(-1 + \frac{1}{u^{\theta}} + \frac{1}{v^{\theta}} \right)^{(-2-\frac{1}{\theta})}$$

Para $\theta > 0$ las cópulas son estrictas y su expresión queda simplificada:

$$C_{\theta}(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$$

La familia de Clayton tiene dependencia en la cola inferior para $\theta > 0$ y como $\varphi^{-1}(s) = (1 + \theta s)^{-1/\theta}$, a partir del Teorema 1.7 tenemos que $\lambda_u = 0$ y el coeficiente de dependencia en la cola inferior viene dado por:

$$\begin{aligned} \lambda_L &= 2 \lim_{s \rightarrow \infty} [\varphi^{-1}(2s)/\varphi^{-1}(s)] \\ &= 2 \lim_{s \rightarrow \infty} \left[\frac{(1 + 2\theta s)^{-\frac{1}{\theta}-1}}{(1 + \theta s)^{-\frac{1}{\theta}-1}} \right] \\ &= 2 \cdot 2^{-\frac{1}{\theta}-1} = 2^{-\frac{1}{\theta}} \end{aligned}$$

La expresión de la tau de Kendall es:

$$\tau_{\theta} = 1 + 4 \int_0^1 \frac{t^{\theta+1} - t}{\theta} dt = 1 + \frac{4}{\theta} \left(\frac{1}{\theta+2} - \frac{1}{2} \right) = \frac{\theta}{\theta+2}$$

además, $C_{-1} = W$, $\lim_{\theta \rightarrow 0} C_{\theta} = \Pi$ y $\lim_{\theta \rightarrow \infty} C_{\theta} = M$.

Familia de Frank

Sea $\varphi(t) = -\ln \frac{e^{-\theta t} - 1}{e^{-\theta} - 1}$, $\theta \in (-\infty, +\infty) \setminus \{0\}$.

La familia de Frank viene dada por:

$$C_{\theta}(u, v) = -\frac{1}{\theta} \ln \left(1 + \frac{(\exp(-\theta u) - 1)(\exp(-\theta v) - 1)}{e^{-\theta} - 1} \right)$$

con función de densidad

$$c_{\theta}(u, v) = \frac{\theta \exp\{\theta(1 + u + v)\}(-1 + \exp \theta)}{(\exp \theta - \exp(\theta + \theta u) + \exp(\theta(u + v)) - \exp(\theta + \theta v))^2}$$

Las cópulas de esta familia son arquimedianas estrictas.

Además:

$$\begin{aligned} \lim_{\theta \rightarrow \infty} C_\theta &= M \\ \lim_{\theta \rightarrow -\infty} C_\theta &= W \\ \lim_{\theta \rightarrow 0} C_\theta &= \Pi \end{aligned}$$

Los miembros de esta familia son las únicas cópulas arquimedianas que satisfacen $C(u, v) = \overline{C}(u, v)$ (simetría radial), por lo tanto representan patrones de dependencia simétrica (la dependencia es la misma entre retornos positivos que entre retornos negativos).

Como

$$\varphi^{-1}(s) = -\frac{1}{\theta} \ln \left(1 - (1 - e^{-\theta}) e^{-s} \right)$$

y

$$\varphi^{-1'}(s) = \frac{-\frac{1}{\theta} (1 - e^{-\theta}) e^{-s}}{1 - (1 - e^{-\theta}) e^{-s}}$$

y si $\varphi^{-1'}(0) = -\frac{e^\theta - 1}{\theta}$ es finita, la familia de Frank no tiene dependencia en la cola superior y, además, debido a la simetría radial ($C = \overline{C}$) tampoco existiría dependencia en la cola inferior.

La expresión de la tau de Kendall para esta familia es

$$\tau_\theta = 1 - 4 \frac{(1 - D_1(\theta))}{\theta}$$

con $D_k(x) = \frac{k}{x^k} \int_0^x \frac{t^k}{e^t - 1} dt$ (función de Debye) para cualquier entero positivo.

1.10.3. Cópulas de valores extremos

Definición 1.17

Se dice que una cópula C es de valores extremos (EV) si satisface la siguiente relación

$$C(u_1^t, \dots, u_n^t) = C^t(u_1, \dots, u_n) \quad \forall t > 0$$

La cópula de Gumbel es de valores extremos:

$$\begin{aligned} C(u^t, v^t) &= \exp\left(-\left[(-\ln u^t)^\alpha + (-\ln v^t)^\alpha\right]^{1/\alpha}\right) \\ &= \exp\left(-\left(t^\alpha\left[(-\ln u)^\alpha + (-\ln v)^\alpha\right]\right)^{1/\alpha}\right) \\ &= \left[\exp\left(-\left[(-\ln u)^\alpha + (-\ln v)^\alpha\right]^{1/\alpha}\right)\right]^t \\ &= C^t(u, v) \end{aligned}$$

Si nos preguntamos por la relación existente entre las cópulas de valores extremos y la teoría de extremos multivariantes, la respuesta es la siguiente:

Sea $X_{n,m}^+ = \text{máx}(X_{n,1}, \dots, X_{n,k}, \dots, X_{n,m})$ con $\{X_{n,k}\}$ k v.a. iid.

Sea G_n la distribución marginal del extremo univariante $X_{n,m}^+$. Entonces, la distribución límite G de $(X_{n,1}^+, \dots, X_{n,k}^+, \dots, X_{n,m}^+)$ es tal que:

$$G(X_1^+, \dots, X_k^+, \dots, X_m^+) = C(G_1(X_1^+), \dots, G_k(X_k^+), \dots, G_n(X_n^+))$$

con C una cópula EV y G_k $k = 1 \dots n$ una distribución de valores extremos univariante no degenerada.

Esta relación proporciona un método para construir distribuciones de valores extremos multivariantes. Veamos el caso bidimensional con algo más de detalle:

Dada la cópula C de valores extremos, es decir

$$C(u^t, v^t) = C^t(u, v) \quad \forall (u, v) \in [0, 1]^2 \quad t > 0$$

C es *máx-estable*, es decir, si $(X_1, Y_1), \dots, (X_n, Y_n)$ son pares aleatorios i.i.d a partir de C , cópula de valores extremos, y si

$$\begin{aligned} M_n &= \text{máx}\{X_1, \dots, X_n\} \\ N_n &= \text{máx}\{Y_1, \dots, Y_n\} \end{aligned}$$

C también es una cópula asociada con (M_n, N_n) .

Joe (1997) demuestra que las cópulas EV se pueden representar de la forma:

$$C(u, v) = \exp \left\{ \ln(uv) A \left(\frac{\ln u}{\ln(uv)} \right) \right\}$$

con $A : [0, 1] \rightarrow [\frac{1}{2}, 1]$ función convexa tal que $\max(t, 1-t) \leq A(t) \leq 1 \forall t \in [0, 1]$. La función $A(t)$ se denomina *función de dependencia*.

Como en el caso de v.a. univariantes, es posible demostrar que la cópula límite para la sucesión $\{(a_n + b_n M_n, c_n + d_n N_n)\}_{n \in \mathbb{N}}$ es una cópula EV, si la sucesión converge débilmente en distribución para algunas sucesiones de números a_n, b_n, c_n y d_n , bajo ciertas condiciones de regularidad.

En el caso de la cópula de Gumbel, con una forma sencilla que permite generar muestras aleatorias fácilmente, la función de dependencia es:

$$A(t) = (t^\alpha + (1-t)^\alpha)^{1/\alpha}$$

Ajuste de Cópulas a Series Financieras Bivariantes

En los puntos anteriores se ha establecido la definición de función cópula en términos de variables aleatorias y sus propiedades, así como las ventajas que aporta su utilización como funciones de dependencia. A continuación, abordaremos la cuestión relativa a la elección de la cópula que represente de forma más adecuada la relación de dependencia entre series de retornos de índices financieros.

Recordemos que este proceso de estimación implica, en primer lugar, la correspondiente a las funciones de distribución marginales que mejor se ajusten a las observaciones muestrales de los retornos de cada índice considerado para, a continuación, ligarlas en una función cópula que capture la dependencia conjunta.

La discusión sobre cuáles son las distribuciones marginales de naturaleza paramétrica más adecuadas a este tipo de datos la postergaremos al capítulo siguiente, donde se tendrá en cuenta la naturaleza específica y la dependencia temporal inherente a los mismos. En este capítulo nos centraremos exclusivamente en la elección de la cópula estática, ya sea de tipo empírico o paramétrico, basada en la información proporcionada por las distribuciones empíricas continuas marginales.

2.1. Aproximación no paramétrica

Una primera aproximación **no paramétrica** consiste en obtener la información extraída de las series brutas de datos y proporcionada por la cópula empírica, que recoge la estructura de dependencia subyacente en la muestra bidimensional y permite estimar de forma consistente la cópula real que los representa. Para este cometido, emplearemos la definida en el punto 1.9 (Deheuvels 1979).

En términos generales, nuestro interés se centra en el análisis de dos series temporales correspondientes a sendos índices de mercado, con colas pesadas y alta curtosis. Sean X e Y las variables aleatorias que representan los logaritmos de los retornos diarios de cada índice. Se trata de estimar la distribución de probabilidad conjunta de X e Y , es decir, queremos encontrar probabilidades de la forma $H(x, y) = P(X \leq x, Y \leq y)$.

Denotamos por $F_1(x)$ y $F_2(y)$ las distribuciones marginales de $H(x, y)$, que supondremos continuas. Entonces, por el Teorema de Sklar, existe una única cópula C , asociada a la distribución conjunta H , tal que

$$C(u, v) = H(F_1^{-1}(u), F_2^{-1}(v)) \quad 0 \leq u \leq 1, 0 \leq v \leq 1$$

El procedimiento a seguir, en esta primera aproximación, será el siguiente:

1. Obtener las funciones de distribución empíricas marginales, $F_{1emp}(x)$ y $F_{2emp}(y)$, procedentes de las series brutas de retornos.
2. A partir de ellas, definir las transformaciones integrales a uniformes estandarizadas de las variables aleatorias implicadas X e Y :

$$U = F_{1emp}(X) \quad V = F_{2emp}(Y)$$

3. Obtener la cópula empírica que une U y V , $C_{emp}(u, v)$.

Obsérvese que el punto 1 puede plantear problemas si optamos por el método de máxima verosimilitud en un proceso posterior de ajuste de tipo paramétrico, puesto que para el desarrollo de la función de verosimilitud necesitamos distribuciones marginales continuas, lo cual entra en conflicto con el hecho de que

la función de distribución empírica es de naturaleza discreta. Esta necesidad proviene del hecho de que la derivada de algunas cópulas usadas en la verosimilitud no están definidas si $u, v \in \{1\}$. Para salvar este inconveniente, podemos modificar la función de distribución empírica con el fin de convertirla en absolutamente continua (De Matteis 2001, Dudewicz y Mishra 1988).

Definición 2.1

Sean n observaciones x_1, \dots, x_n iid. La **función de distribución empírica** de la muestra $\forall x \in \mathbb{R}$ viene dada por

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I_{(-\infty, x]}(x_i)$$

siendo I la función indicador usual.

Definición 2.2

Sean $a, b \in \mathbb{R}$ tales que $a \leq \min(x_1, \dots, x_n) \leq \max(x_1, \dots, x_n) \leq b$.

Sean z_1, \dots, z_n los valores de x_1, \dots, x_n ordenados de forma creciente. Definimos $z_0 = a$ y $z_{n+1} = b$.

Llamamos **función de distribución empírica continua** a

$$F_{emp}(x; a, z_1, \dots, z_n, b)$$

que toma el valor 0 si $x \leq a$, el valor 1 si $x \geq b$ y, para puntos intermedios, $a < x < b$, toma el valor asignado por el segmento lineal que une los puntos medios de las barras de los intervalos adyacentes $[z_i, z_{i+1}]$. El punto medio de la barra más a la izquierda $[z_1, z_2]$ se une al punto $(0, 0)$ y el de la barra más a la derecha $[z_{n-1}, z_n]$ se une al punto $(b, 1)$.

Haciendo uso de esta definición, obtenemos las funciones de distribución marginales empíricas continuas, $F_{1emp}(x)$ y $F_{2emp}(y)$, que tienen como funciones de densidad $f_1(x)$ y $f_2(y)$.

Proposición 2.1

Sea X una v.a. con función de distribución F . Sea F^{-1} la función cuantílica de F

$$F^{-1}(\alpha) = \inf\{x/F(x) \geq \alpha\}$$

Entonces

(i) Para cualquier U v.a. distribuida como una uniforme estándar se verifica que $F^{-1}(u) \sim F$.

(ii) Si F es continua, la v. a. $F(x)$ se distribuye como una $U(0, 1)$.

(Demostración en Feller 1966)

La proposición anterior garantiza que las funciones de distribución marginales F_1 y F_2 de $(x_1, y_1), \dots, (x_n, y_n)$ son v.a. distribuidas como $U(0, 1)$.

Definimos

$$u_i = F_{1emp}(x_i) \quad , \quad v_i = F_{2emp}(y_i) \quad i = 1 \dots n$$

con lo que $(u_i)_{i=1 \dots n}$ y $(v_i)_{i=1 \dots n}$ se distribuyen aproximadamente como uniformes estándar. Podemos, entonces, calcular la cópula empírica de $(u_1, v_1) \dots (u_n, v_n)$.

La cópula empírica así determinada será muy útil empleada como herramienta en el análisis exploratorio de los datos, ya que podemos utilizarla para los siguientes cometidos:

1. Identificar el signo de la dependencia bivalente

La existencia de un orden parcial entre las cópulas garantiza que cualquier cópula, incluida la empírica, está acotada por las cópulas máxima y mínima

$$W \leq C_{emp} \leq M$$

por tanto, podemos representar los diagramas de contorno de la cópula empírica obtenida y compararlos con los de las cotas de Fréchet y los de la cópula independencia que, como sabemos, caracterizan las situaciones de dependencia extrema. La proximidad a cualquiera de ellas indicará el signo de la dependencia bivalente o la posibilidad de independencia entre las variables X e Y .

Asimismo, la inspección gráfica de estos diagramas de contorno, junto con la de la representación de la densidad empírica, permitirá efectuar comparaciones entre pérdidas conjuntas extremas y ganancias conjuntas extremas, además de detectar la presencia de asimetría en colas.

2. Cuantificar la dependencia conjunta

El cálculo de las versiones muestrales de las medidas de concordancia τ y r_s , cuyas expresiones en términos de la cópula empírica vienen dadas por la

ecuaciones recogidas en el punto 1.9, también informarán sobre el signo y la cantidad de dependencia entre X e Y recogida en la cópula empírica.

3. Evaluar las posibles familias paramétricas candidatas para el ajuste

Las estimaciones muestrales de las medidas de concordancia citadas anteriormente pueden servir para calibrar diferentes familias paramétricas: después de estimar los parámetros para una familia paramétrica concreta mediante, por ejemplo, máxima verosimilitud, se puede realizar una comparación del valor de τ y r_s de la cópula estimada con las muestrales. También es posible, como se detallará más adelante al tratar los ajustes de tipo paramétrico para cópulas de la familia arquimediana, plantear un algoritmo de calibración basado en la estimación del parámetro θ de la cópula obtenida a partir de la relación existente entre dicho parámetro y una medida de dependencia muestral concreta, por ejemplo la τ de Kendall.

4. Cuantificar la dependencia en colas

Las estimaciones empíricas de los índices de dependencia en la cola superior e inferior, además de informar sobre la dependencia extrema conjunta, pueden utilizarse para la comparación entre cópulas. Recordemos la expresión del índice de dependencia en la cola superior:

$$\lambda_u = \lim_{\alpha \rightarrow 1^-} \frac{1 - 2\alpha + C(\alpha, \alpha)}{1 - \alpha}$$

Se puede construir un gráfico de la función $\frac{1-2\alpha+C(\alpha,\alpha)}{1-\alpha}$ con $C = C_{emp}$, y haciendo tender $\alpha \rightarrow 1^-$ observar posibles convergencias. Análogamente se puede proceder para el índice de dependencia en la cola inferior.

2.2. Aproximación semiparamétrica

En cuanto a la elección de la familia paramétrica $\{C_\theta, \theta \in \Theta\}$, y dado que en la mayoría de los estudios empíricos que implican datos de series financieras aparece de forma evidente la dependencia entre valores extremos (dependencia en colas), resulta aconsejable encontrar especificaciones de modelos que reflejen el fenómeno de las colas pesadas. Desde el punto de vista empírico, la clase de *cópulas arquimedianas* aparece como una elección adecuada para reflejar esta situación, además de poseer buenas propiedades que las hacen muy manejables (véase el punto 1.10.2).

Si nos decidimos por una cópula perteneciente a esta última familia, es posible utilizar una idea apuntada por Genest y McKey (1986), con el fin de elegir la mejor cópula. Es un método simple basado en medidas de dependencia muestrales, que puede considerarse como una adaptación no paramétrica del conocido método de los momentos: a partir de la relación explícita entre el parámetro θ de la cópula y la τ de Kendall se infiere una estimación para θ . Este método permite también contrastar para diferentes cópulas arquimedianas cuál de ellas representa mejor la estructura de dependencia comparándolas con la cópula empírica. Para ello utilizaremos el algoritmo elaborado por Genest y Rivest (1993) que propone determinar dos estimaciones para la función de distribución K de una cópula arquimediana desconocida C : una de ellas de naturaleza no paramétrica, por tanto independiente de las marginales y del parámetro de la cópula, y otra de tipo paramétrico $K_C(t)$, vista en el Teorema 1.8. De esta forma, para la estimación paramétrica necesitamos el parámetro θ estimado previamente. Aunque la estimación paramétrica puede utilizarse directamente para construir la cópula arquimediana bivalente, resulta más conveniente proponer diferentes elecciones de cópulas arquimedianas y seleccionar a través de la estimación de θ aquella cuya estimación paramétrica esté más próxima a la estimación no paramétrica.

El algoritmo a seguir es el siguiente:

(i) Estimación no paramétrica

Sea $(X_1, Y_1) \dots (X_n, Y_n)$ una muestra aleatoria con función de distribución bivalente $H(X, Y)$, distribuciones marginales $F_1(X)$ y $F_2(Y)$ y cópula arquimediana $C(U, V)$ con $U = F_1(X)$ y $V = F_2(Y)$ variables aleatorias distribuidas $U(0, 1)$.

Buscamos una estimación de la función de distribución univariante

$$K(w) = P[C(U, V) \leq w] = P[H(X, Y) \leq w]$$

para ello, definimos la v.a. $W = \hat{H}(X, Y)$, con \hat{H} la estimación empírica de la distribución bivalente H .

A continuación, determinamos la estimación empírica de la distribución de W . Sea

$$W_i = \hat{H}(x_i, y_i) = \frac{\text{Card}\{(x_j, y_j) \mid x_j < x_i, y_j < y_i\}}{(n-1)} \quad (2.1)$$

\hat{H} es la estimación empírica buscada. Entonces, un estimador no paramétrico de $K(w)$ viene dado por:

$$K_n(w) = \frac{1}{n} \sum_{i=1}^n \{i : 1 \leq i \leq n, w_i \leq w\} \quad (2.2)$$

Proposición 2.2 (*Genest y Rivest*)

Sea $H(x, y)$ una función de dependencia absolutamente continua respecto a la medida de Lebesgue y sea $(x_1, y_1) \dots (x_n, y_n)$ una m.a.s. obtenida a partir de H . Bajo condiciones de regularidad débiles, la distribución de los W_i converge a $K(w) = P\{H(x, y) \leq w\}$, y la función de distribución empírica $K_n(w)$ de los w_i es un estimador \sqrt{n} -consistente de $K(w)$.

(Demostración en Genest y Rivest 1993, Pág 1042)

Es importante establecer la relación existente entre los W_i definidos por (2.1) y el valor muestral de la τ de Kendall. Recordemos que esta última se define como la diferencia entre el número de pares concordantes y el número de pares discordantes, dividida por el número total de pares $\binom{n}{2}$. Sean

$$I_{ij} = \begin{cases} 1 & \text{si } X_j \leq X_i \text{ e } Y_j < Y_i \\ 0 & \text{en otro caso} \end{cases}$$

$$I_{ji} = \begin{cases} 1 & \text{si } X_i \leq X_j \text{ e } Y_i < Y_j \\ 0 & \text{en otro caso} \end{cases}$$

Con esta notación $I_{ij} + I_{ji}$ puede tomar los valores 0, 1 y 2, pero valdrá 1 si y sólo si las observaciones i y j son concordantes. Entonces

$$W_i = \frac{\sum_{j=1}^n I_{ij}}{n-1}$$

por lo que el número de pares concordantes c se obtiene como

$$W_i = \frac{\sum_{j=1}^n I_{ij}}{n-1} \iff \sum_{j=1}^n I_{ij} = (n-1)W_i \Rightarrow \sum_{i,j=1}^n I_{ij} = (n-1) \sum_{i=1}^n W_i$$

El número de pares discordantes d es

$$d = \binom{n}{2} - c = \binom{n}{2} - (n-1) \sum_{i=1}^n W_i$$

por lo que el cálculo de τ es

$$\hat{\tau} = \frac{c-d}{\binom{n}{2}} = \frac{2(n-1) \sum_{i=1}^n W_i - \frac{n(n-1)}{2}}{\frac{n(n-1)}{2}} = \frac{4(n-1) \sum_{i=1}^n W_i - n(n-1)}{n(n-1)} = 4\overline{W} - 1 \quad (2.3)$$

La conexión de lo anterior con la cópula empírica es la siguiente:

Sean R_i y S_i , rangos de x_i e y_i respectivamente. Después de ajustar la escala de los ejes mediante el factor $\frac{1}{n+1}$ (n tamaño de la muestra) con el fin de que todos los puntos pertenezcan al cuadrado unidad $[0, 1]^2$, se obtiene

$$W_i = C_n \left(\frac{R_i}{n+1}, \frac{S_i}{n+1} \right)$$

con C_n la cópula empírica definida en el Capítulo 1 (expresión 1.5), que en términos de rangos puede escribirse como:

$$C_n(u, v) = \frac{1}{n} \sum_{i=1}^n I \left(\frac{R_i}{n+1} \leq u, \frac{S_i}{n+1} \leq v \right)$$

Entonces

$$\overline{W} = \int \int_{[0,1]^2} C_n(u, v) dC_n(u, v)$$

A partir de (2.3) y del hecho de que bajo condiciones de regularidad adecuadas $C_n \rightarrow_{n \rightarrow \infty} C$, con C la cópula verdadera, se puede concluir (como hizo Hoeffding

1948) que τ_n es un estimador asintóticamente insesgado de la versión poblacional de la τ de Kendall, dada en términos de la cópula por:

$$\tau = 4 \int \int_{[0,1]^2} C(u, v) dC(u, v) - 1$$

De forma general, si $\theta = h(\tau)$ para alguna función h , nos referiremos a

$\hat{\theta}_n = h(\tau_n)$ como el estimador de θ basado en la τ de Kendall.

Según la Proposición 3.1 de Genest y Rivest (1993)

$$\sqrt{n} \frac{\tau_n - \tau}{4S} \approx N(0, 1)$$

donde

$$S^2 = \frac{1}{n} \sum_{i=1}^n (W_i + \widetilde{W}_i - 2\overline{W})^2$$

y

$$\widetilde{W}_i = \frac{1}{n-1} \sum_{j=1}^n I_{ji} = \frac{1}{n-1} \sum \{j : X_i \leq X_j, Y_i \leq Y_j\}$$

Por aplicación del Teorema de Slutsky (también conocido como método Delta) se obtiene que, cuando $n \rightarrow \infty$:

$$\hat{\theta}_n \approx N\left(\theta, \frac{1}{n} (4Sh'(\tau_n))^2\right)$$

De acuerdo con lo anterior, un intervalo de confianza para θ al nivel

$100(1 - \alpha)\%$ viene dado por

$$\hat{\theta}_n \pm z_{\frac{\alpha}{2}} \frac{1}{\sqrt{n}} 4S |h'(\tau_n)|$$

Para un estimador consistente alternativo de la varianza asintótica de la varianza asintótica de τ_n puede consultarse Samara y Randless (1988).

(ii) Estimación paramétrica

Consideramos la dada por

$$K_c(w) = w - \frac{\varphi(w)}{\varphi'(w^+)}$$

y denominamos a este estimador $K_\varphi(w)$.

(iii) Comparación entre las dos estimaciones

La idea central consiste en que el estimador $\hat{\theta}$ es adecuado si la distribución paramétrica $K_\varphi(w)$ es similar a la no paramétrica $K_n(w)$. Esta comparación puede hacerse de forma gráfica o analítica.

Gráficamente, a través de un QQ-Plot de las dos distribuciones. Si la elección de $\hat{\theta}$ es buena, este gráfico proporcionará una línea recta como resultado y la cópula elegida ajustará bien a la muestra.

Analíticamente, Frees y Valdez (1998) sugieren minimizar la siguiente distancia

$$\int (K_\varphi(w) - K_n(w))^2 dK_n(w)$$

2.3. Aproximación paramétrica

En un **planteamiento de naturaleza paramétrica** (para el cual lo expuesto anteriormente no resulta excluyente sino complementario e incluso recomendable como tratamiento previo) supondremos que la cópula verdadera pertenece a una familia paramétrica $\{C_\theta, \theta \in \Theta\}$, con Θ el espacio de posibles valores del parámetro. Mediante el método de máxima verosimilitud se obtendrán estimadores de $\theta \in \Theta$ consistentes y asintóticamente normales.

Entre las técnicas estadísticas clásicas, la estimación de máxima verosimilitud es una alternativa al método de los momentos, que es usualmente más eficiente, particularmente cuando el parámetro θ es multidimensional. En el contexto en que estamos situados, se requiere una adaptación a la inferencia realizada sobre el parámetro de dependencia, basada exclusivamente en rangos, adaptación que fue descrita en términos generales por Oakes (1994) y más tarde quedó formalizada por Genest, Ghoudi y Rivest (1995) y por Shih y Louis (1995). Este método, denominado de máxima pseudo-verosimilitud o de máxima verosimilitud canónica (CML), requiere que C_θ sea absolutamente continua con densidad c_θ y maximiza el logaritmo de la verosimilitud de la forma

$$l(\theta) = \sum_{i=1}^n \log \left(c_\theta \left(\frac{R_i}{n+1}, \frac{S_i}{n+1} \right) \right)$$

expresión que se obtiene sin más que sustituir las marginales desconocidas F_1 y F_2 en la verosimilitud clásica por sus contrapartidas empíricas $F_{1emp}(x)$ y $F_{2emp}(y)$.

Este método puede parecer menos atractivo que el basado en la inversión de la τ de Kendall, ya que implica una mayor cantidad de cálculos y requiere la expresión explícita de la densidad c_θ . sin embargo, puede aplicarse de forma más general, concretamente si θ es multidimensional.

A partir de las distribuciones empíricas marginales continuas, F_{1emp} y F_{2emp} , y de la representación canónica de la cópula:

$$h(x, y) = f_1(x)f_2(y)c(F_{1emp}(x_i), F_{2emp}(y_i))$$

con $c(u, v) = \frac{\partial^2}{\partial u \partial v} C(u, v)$ función de densidad de la cópula C , la ecuación de verosimilitud será

$$\mathcal{L}(\theta; u, v) = \prod_{i=1}^n c(u_i, v_i)$$

con $u_i = F_{1emp}(x_i)$ y $v_i = F_{2emp}(y_i)$ y la estimación de θ se determina en la forma usual como:

$$\hat{\theta} = \arg \max_{\theta \in \Theta} \sum_{i=1}^n \log \mathcal{L}(\theta; u_i, v_i) \tag{2.4}$$

o equivalentemente, en términos de rangos, como raíz de la ecuación

$$\frac{\partial}{\partial \theta} l(\theta) = \sum_{i=1}^n \left[\frac{\partial}{\partial \theta} c_\theta \left(\frac{R_i}{n+1}, \frac{S_i}{n+1} \right) / c_\theta \left(\frac{R_i}{n+1}, \frac{S_i}{n+1} \right) \right] = 0 \tag{2.5}$$

Genest, Ghoudi y Rivest (1995) demuestran que, bajo ciertas condiciones de regularidad, el estimador $\hat{\theta}$ obtenido a partir de (2.4) o de (2.5) es consistente y asintóticamente tiene distribución normal bajo condiciones de regularidad similares a las impuestas en el método de máxima verosimilitud.

Proposición 2.3

Bajo condiciones de regularidad adecuadas, el estimador $\hat{\theta}$ obtenido a partir de (2.4) es consistente y $n^{1/2}(\hat{\theta} - \theta)$ es asintóticamente normal con varianza $\nu^2 = \sigma^2/\beta^2$, donde ν^2 depende exclusivamente de la cópula verdadera C_θ .

Un estimador consistente de ν^2 viene dado por

$$\hat{\nu}_n^2 = \frac{\hat{\sigma}_n^2}{\hat{\beta}_n^2}$$

con

$$\begin{aligned}\widehat{\sigma}_n^2 &= \frac{1}{n} \sum_{i=1}^n (A_i - \bar{A})^2 \\ \widehat{\beta}_n^2 &= \frac{1}{n} \sum_{i=1}^n (B_i - \bar{B})^2\end{aligned}$$

con \bar{A} y \bar{B} las medias aritméticas de A_i y B_i respectivamente. El cálculo de las pseudo-observaciones A_i y B_i puede hacerse como sigue:

Si reordenamos $(X_i, Y_i)_{i=1\dots n}$ de forma que $X_1 < \dots < X_n$ (es decir, $R_1 = 1, \dots, R_n = n$) y utilizamos la siguiente notación

$$\begin{aligned}L(\theta, u, v) &= \log c_\theta(u, v) \\ L_\theta(\theta, u, v) &= \frac{\partial}{\partial \theta} \log c_\theta(u, v) \\ L_u(\theta, u, v) &= \frac{\partial}{\partial u} \log c_\theta(u, v) \\ L_v(\theta, u, v) &= \frac{\partial}{\partial v} \log c_\theta(u, v)\end{aligned}$$

entonces

$$B_i = L_\theta \left(\widehat{\theta}_n, \frac{i}{n+1}, \frac{S_i}{n+1} \right) \quad i = 1, \dots, n$$

y

$$\begin{aligned}A_i &= B_i - \frac{1}{n} \sum_{j=1}^n L_\theta \left(\widehat{\theta}_n, \frac{j}{n+1}, \frac{S_j}{n+1} \right) \times L_u \left(\widehat{\theta}_n, \frac{j}{n+1}, \frac{S_j}{n+1} \right) \\ &\quad - \frac{1}{n} \sum_{S_j \geq S_i} L_\theta \left(\widehat{\theta}_n, \frac{j}{n+1}, \frac{S_j}{n+1} \right) \times L_v \left(\widehat{\theta}_n, \frac{j}{n+1}, \frac{S_j}{n+1} \right)\end{aligned}$$

Un intervalo de confianza para θ al nivel $100(1 - \alpha)\%$ vendrá dado por

$$\widehat{\theta}_n \pm z_{\frac{\alpha}{2}} \frac{\widehat{v}_n}{\sqrt{n}}$$

2.4. Bondad de ajuste

Por último, necesitamos confirmar el ajuste realizado a través de algún **contraste de bondad de ajuste (GOF)**. Proponemos el conocido de Kolmogorov-Smirnov que permite detectar, incluso con muestras pequeñas, discrepancias entre la distribución teórica y la empírica. El estadístico utilizado para el contraste es la máxima diferencia entre la distribución empírica y la teórica

$$T = \max \left\{ \left| \widehat{F}(x) - F(x) \right| \right\}$$

Utilizaremos el planteamiento de una muestra para contrastar que

$$K_c(t) = t - \frac{\varphi(t)}{\varphi'(t^+)} = P[C(U, V) \leq t] \quad (2.6)$$

aplicado a $C(F_{1emp}(x), F_{2emp}(y))$ es una $U(0, 1)$.

A partir del estimador $\widehat{\theta}$, obtenido por ejemplo por máxima verosimilitud, también podemos realizar un QQ-Plot de $K_{C(F_{1emp}(x), F_{2emp}(y))}(\cdot)$ frente a cuantiles de la $U(0, 1)$. Si el ajuste es bueno debemos obtener una línea recta.

En el caso de comparar $K_n(w)$ y $K_C(t)$ es necesario utilizar la versión del contraste de Kolmogorov-Smirnov para dos muestras, en la que la hipótesis nula asume la igualdad de las dos distribuciones citadas.

Si se utiliza la aproximación paramétrica de máxima verosimilitud, un criterio suplementario para examinar la bondad del ajuste de nuestros modelos es el que se basa en la comparación de los valores negativos de las funciones de verosimilitud. Akaike (1973,1977) desarrolló una estrategia de decisión basada en la medida de información de Leibler (1951), dado que esta medida proporciona un criterio de ordenación de modelos estadísticos alternativos. Recordemos el estimador de máxima verosimilitud para el parámetro θ :

$$\widehat{\theta} = \arg \max_{\theta \in \Theta} \sum_{i=1}^n \log \mathcal{L}(\theta; u_i, v_i)$$

Definición 2.3

El criterio de información de Akaike se define como

$$AIC = 2(-\log \mathcal{L}(\theta; u_i, v_i)) + 2q$$

donde q es el número de parámetros estimados en el modelo.

El valor de AIC informa sobre qué modelo proporciona un mejor ajuste: cuanto menor sea este valor mejor es el modelo considerado.

La figura 2.1 resume el procedimiento propuesto para el ajuste.

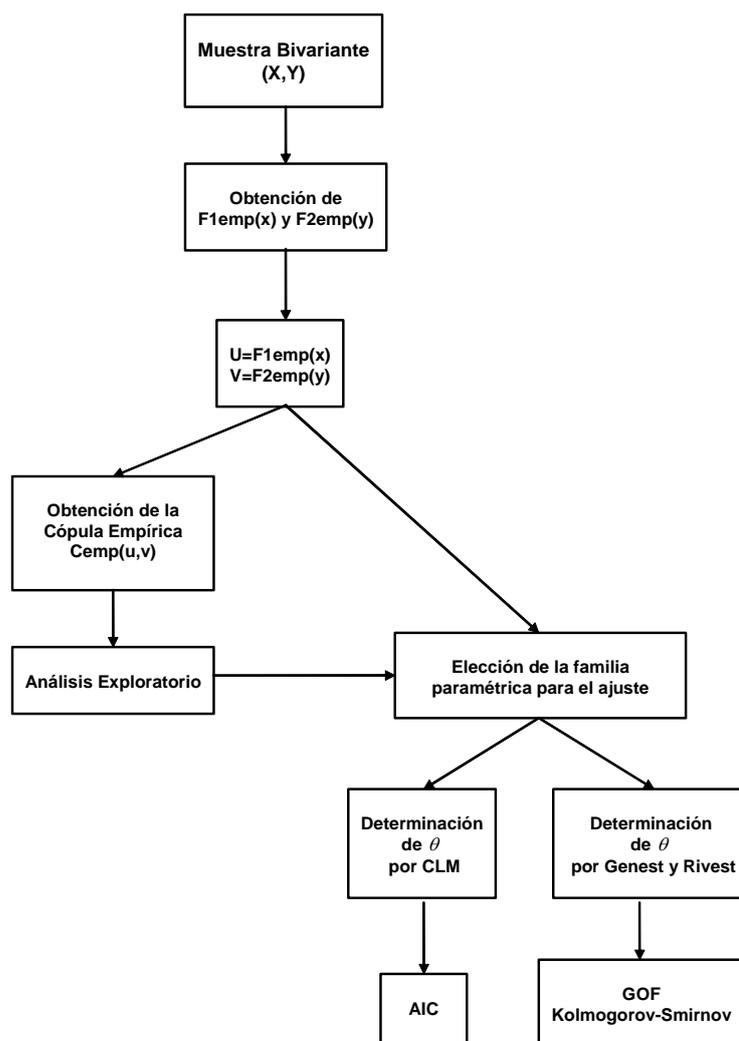


Figura 2.1: *Procedimiento de selección y ajuste de la Cópula bivalente*

2.5. Ajuste de cópulas a series diarias de retornos de Índices financieros

2.5.1. Descripción de las series

A continuación, aplicamos los procedimientos propuestos en los puntos anteriores para determinar la cópula que mejor representa la relación de dependencia entre retornos diarios de índices financieros. Para ello, hemos seleccionado dos índices financieros ampliamente conocidos: DowJones e Ibex35. El primero como representante del comportamiento del mercado americano y el segundo del europeo, más concretamente del español. Las dos series temporales brutas correspondientes, contienen información del valor diario al cierre de ambos índices desde el 3 de Enero de 1994 al 28 de Marzo de 2006. Los datos del índice Ibex35 provienen de los proveedores del mercado madrileño Telekurs Financial, y los del Dow Jones directamente del mercado americano (American Stock Exchange).

La primera cuestión abordada ha sido el tratamiento de los valores missing correspondientes a periodos no operativos en uno u otro mercado, afectados por la diferencia horaria (fin de semana, fiestas, etc.). Al objeto de no perder información, se han alineado las series aplicando el siguiente procedimiento a cada una de ellas:

- (i) localización del periodo t en el que existe un valor missing en la serie.
- (ii) sustitución del mismo por la media \bar{x}_T (con el fin de conservar la tendencia), siendo T el número total de observaciones de la serie.
- (iii) obtención del valor final a utilizar en ese periodo a través de una media móvil asimétrica de cuatro puntos, calculada como:

$$x_t = \frac{x_{t-2} + x_{t-1} + \bar{x}_T + x_{t+1}}{4}$$

La incorporación del valor x_{t+1} tiene la finalidad de reflejar en el valor resultante el impacto, al inicio de la siguiente sesión, de las noticias acumuladas en el periodo inactivo. La figura 2.2 muestra los gráficos de ambas series convenientemente alineadas.

A partir de estas series se obtienen las correspondientes a los retornos logarítmicos, que constituyen el objeto de nuestro análisis. Llamamos retorno logarítmico

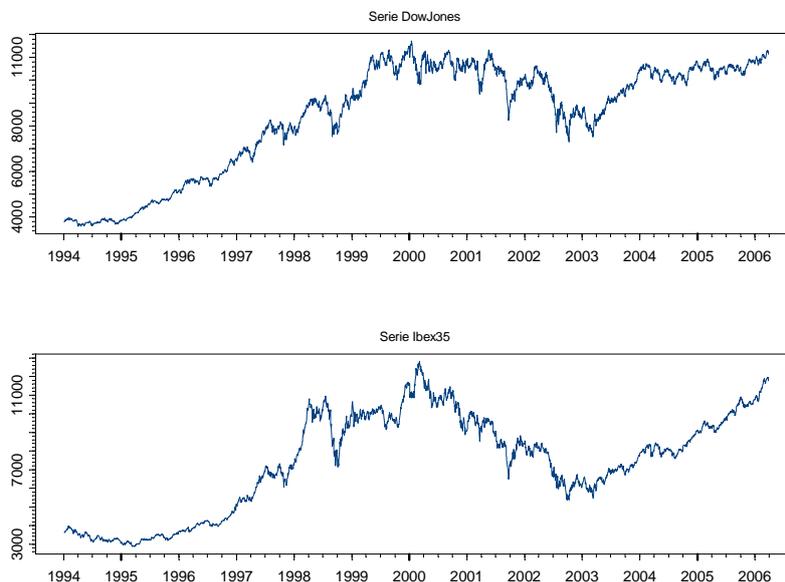


Figura 2.2: Valores diarios al cierre de los índices DowJones e Iboex35

en un periodo t a

$$R_t = \log(P_t/P_{t-1})$$

donde P_t y P_{t-1} son los valores al cierre del índice correspondiente en los periodos t y $t-1$, respectivamente. La representación de las series de retornos resultantes se muestra en la figura 2.3.

Contrastamos la normalidad de los retornos mediante el estadístico de Jarque-Bera (1980)

$$JB = \frac{T}{6} \left(\widehat{skew}^2 + \frac{(\widehat{kurt} - 3)^2}{4} \right)$$

con \widehat{skew} la asimetría muestral y \widehat{kurt} la curtosis muestral. Bajo la hipótesis nula de que los datos se distribuyen normalmente, JB se distribuye según una χ_2^2 . Los resultados de aplicar el contraste a las dos series consideradas aparecen en la siguiente tabla

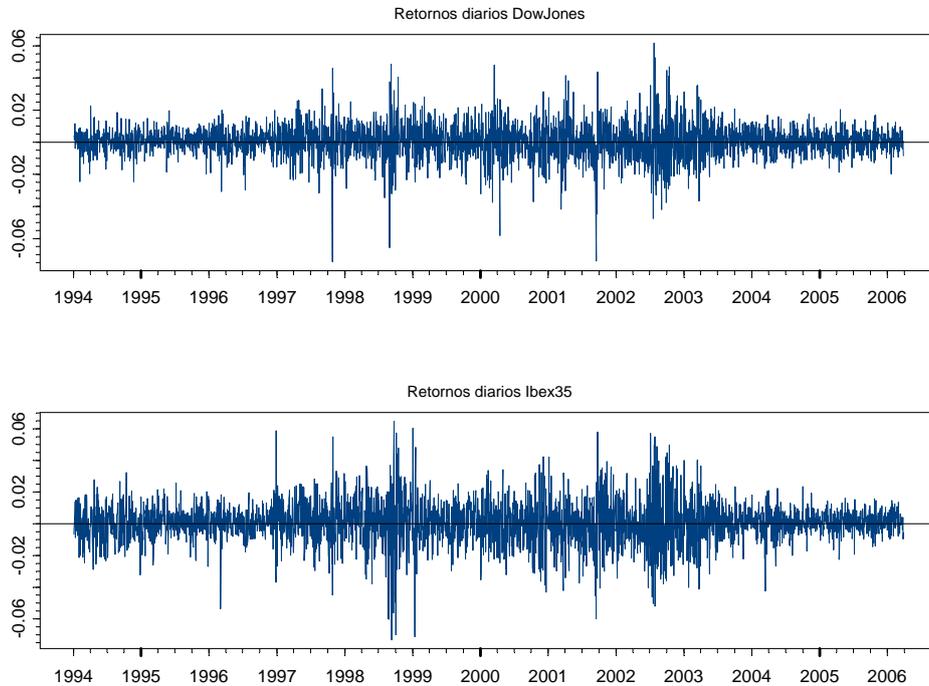


Figura 2.3: *Retornos logarítmicos diarios DowJones e Ibex35*

	Estadístico JB	p-valor
DowJones	2619.619	0
Ibex35	1271.999	0

por lo que en ambos casos se rechaza la hipótesis de normalidad.

En la figura 2.4 se muestra una descripción gráfica de la serie de retornos logarítmicos correspondientes al índice DowJones. En la esquina superior izquierda aparece el histograma, que refleja una distribución leptocúrtica (obsérvese la densidad Kernel ajustada en la imagen inferior), y a su derecha el correspondiente box-plot presenta una serie de alta frecuencia en colas, lo que coincide con la información aportada por el QQ-plot que enfrenta los cuantiles empíricos con los de una normal, en el que se pueden apreciar ambas colas de la distribución empírica más pesadas que las que corresponderían a una distribución normal.

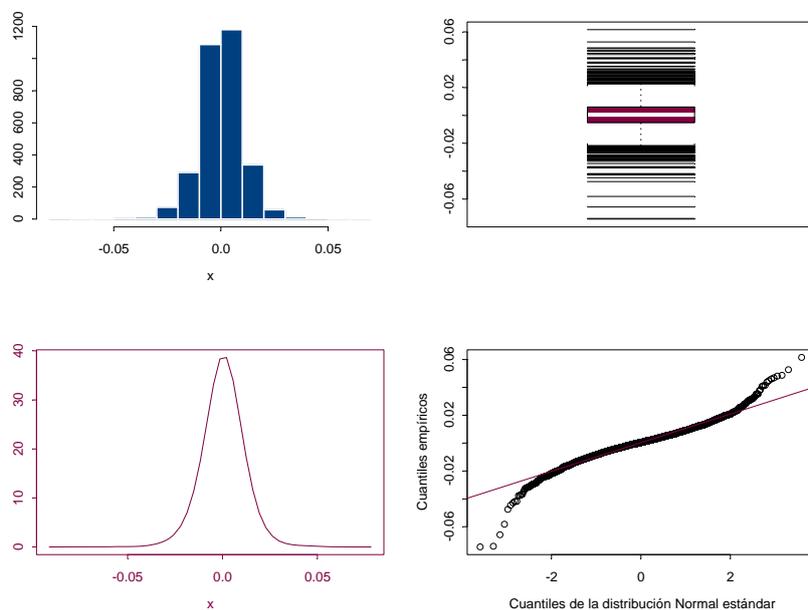


Figura 2.4: Descripción gráfica de los retornos logarítmicos diarios DowJones

Un análisis similar se deduce de la inspección de la figura 2.5 en la que aparece la descripción gráfica de los retornos logarítmicos del índice Ibex35, que vuelve a presentar una serie de alta frecuencia con colas pesadas.

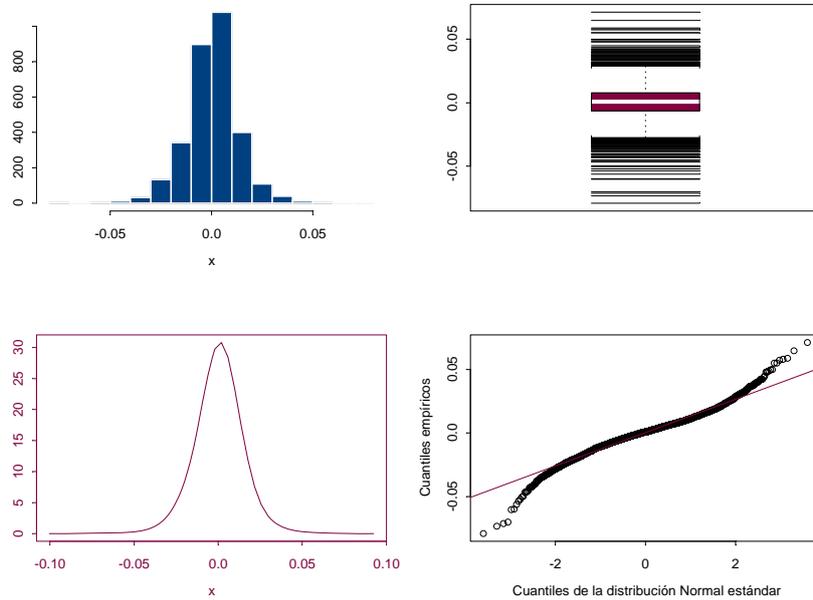


Figura 2.5: Descripción gráfica de los retornos *Ibox35*

Las características numéricas correspondientes a cada serie de retornos aparecen en la tabla 2.1

	Nº obs.	Q_1	Mediana	Q_3	Media	Asimetr.	Curt.
DowJones	3080	-0.00511	0.00045	0.0060	0.00035	-0.2485	7.49
Ibox35	3080	-0.00652	0.00082	0.0076	0.00038	-0.2032	6.122

Tabla 2.1: Características numéricas de los retornos logarítmicos diarios *DowJones* e *Ibox35*

El aspecto de las series (figura 2.3) lleva a sospechar la existencia de estacionariedad. Para comprobarlo, utilizamos el contraste KPSS (Kwiatkowski y otros, 1992). Se obtiene un valor de 0.2812 para el estadístico del contraste en la serie de retornos *DowJones*, y un valor de 0.1748 para el *Ibox35*. Dado que el cuantil 90% del estadístico KPSS es 0.349 y que se trata de un contraste unilateral de cola derecha, no es posible rechazar la hipótesis de estacionariedad

para ninguna de las dos series. Se puede consultar una simulación del algoritmo para obtener la distribución del estadístico del contraste KPSS en Zivot y Wang (2003, página 124). Por otra parte, podemos valorar la dependencia temporal inspeccionando los correspondientes correlogramas representados en la figura 2.6.

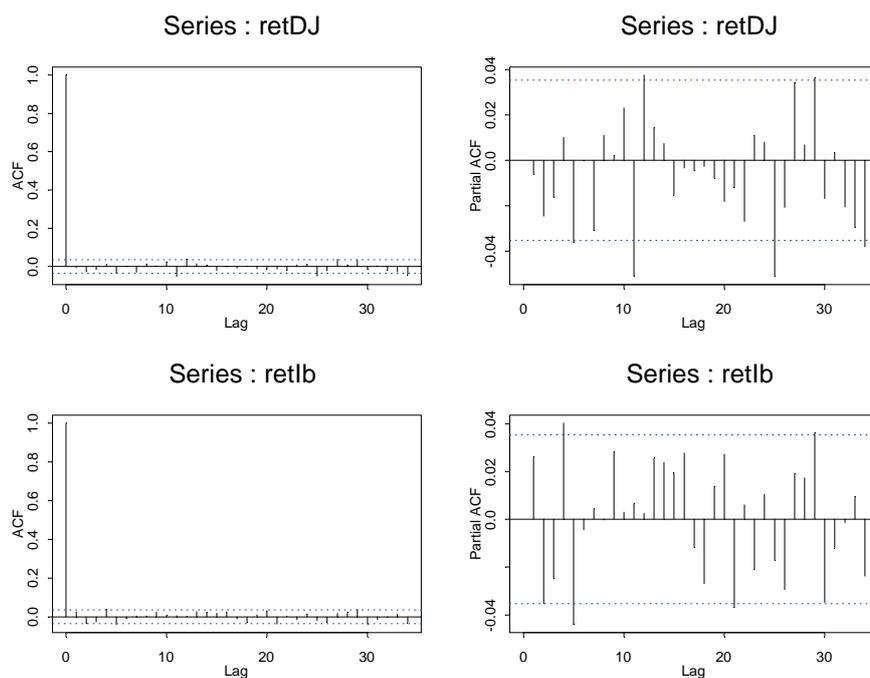


Figura 2.6: *Funciones de autocorrelación simple y parcial de los retornos DowJones (superior) e Ibx35 (inferior)*

Los autocorrelogramas representados no aportan evidencias contra la independencia temporal en los retornos diarios; sin embargo, ambas series presentan heterocedasticidad condicionada (el llamado efecto ARCH), una versión en tiempo discreto de volatilidad estocástica, como se puede observar en la representación de los autocorrelogramas de los retornos al cuadrado (figura 2.7). En el Capítulo 3, nos ocuparemos de analizar y contrastar formalmente este efecto sobre las series consideradas, y de ajustar convenientemente los modelos que lo representan a las distribuciones marginales empíricas.

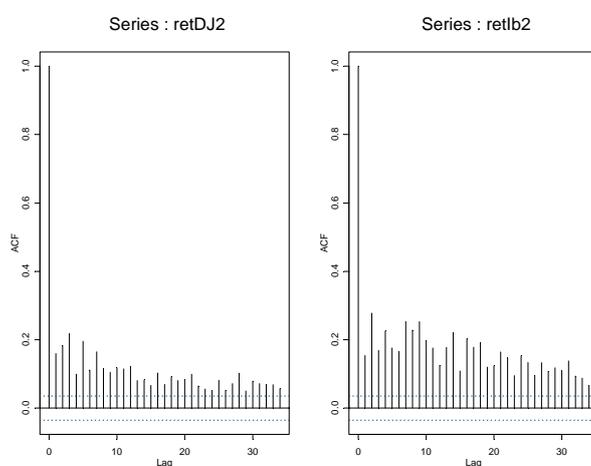


Figura 2.7: Autocorrelogramas de los retornos al cuadrado DowJones e Ibex35

2.5.2. Ajuste no paramétrico de la cópula bivalente DowJones & Ibex35

Comenzamos el procedimiento de ajuste según lo propuesto en el punto 2.1. En primer lugar, se obtienen las distribuciones empíricas continuas a partir de las dos series de retornos y, a continuación, se determina la cópula empírica que las liga. En la figura 2.8 aparecen representados el diagrama de dispersión de los retornos logarítmicos diarios que constituyen la serie bivalente (DowJones, Ibex35) junto con el correspondiente al soporte de la cópula empírica. Este último permite apreciar la dependencia presente en ambas colas de la distribución bivalente.

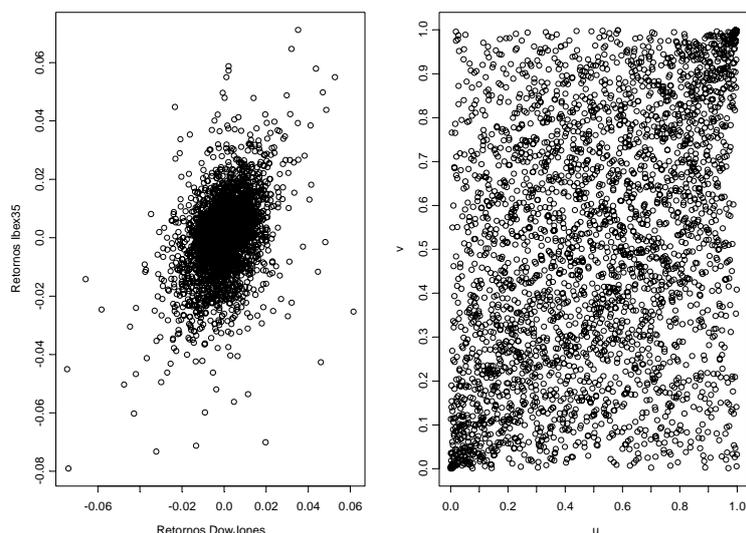


Figura 2.8: *Diagrama de dispersión de la serie bivalente de retornos logarítmicos (DowJones, Ibex35) (izquierda) y soporte de la cópula empírica (derecha)*

Como ya se ha explicado, no supondremos ningún modelo paramétrico para las funciones de distribución marginales, al objeto de evitar los errores procedentes de una mala especificación. En la figura 2.9 podemos observar la densidad correspondiente a la cópula empírica obtenida (esquina superior izquierda), su función de distribución de probabilidad (esquina superior derecha) y en la parte inferior aparecen representados los correspondientes diagramas de contorno.

Si superponemos los diagramas de contorno de la función de distribución asociada a la cópula empírica y el correspondiente a la cópula independencia, es posible confirmar la existencia de dependencia bivalente (parte superior de la figura 2.10). Es más, el diagrama de contorno de la función de densidad de probabilidad de la cópula empírica confirma la existencia de dependencia en ambas colas de la distribución bivalente (figura 2.9). Por otra parte, si superponemos los diagramas de contorno de la función de distribución asociada a la cópula empírica y los correspondientes a las cópulas máxima y mínima, observamos que el signo de esta dependencia es positivo (parte inferior de la figura 2.10) .

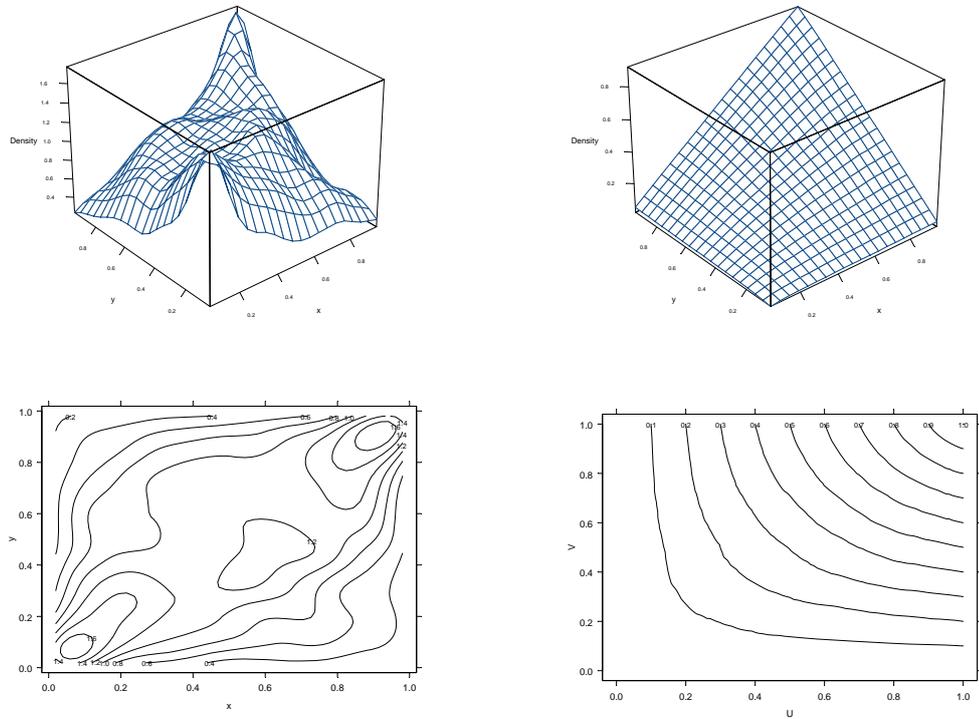


Figura 2.9: *Función de densidad, de distribución y diagramas de contorno de la Cópula empírica*

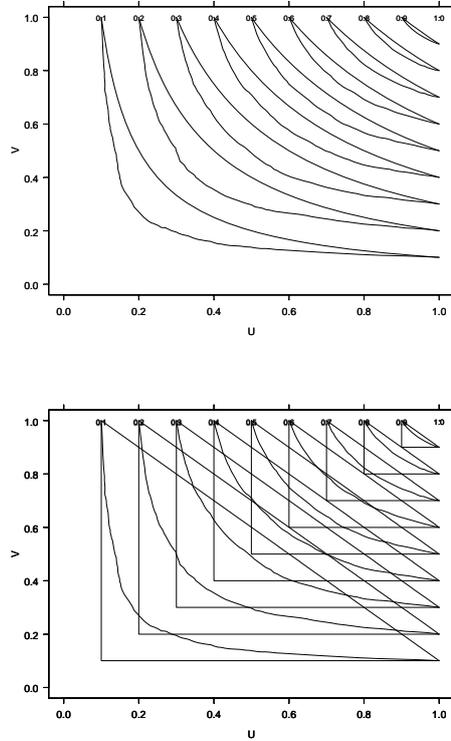


Figura 2.10: Comparación diagramas de contorno de la cópula empírica con la cópula independencia (arriba) y las cópulas máxima y mínima (abajo)

El cálculo de las medidas de dependencia muestrales confirma el diagnóstico anterior. Si utilizamos como indicadores exploratorios del grado de dependencia el coeficiente de correlación de Pearson, la r_s de Spearman y la τ de Kendall, sus valores muestrales son los siguientes:

$\hat{\rho}$	\hat{r}_s	$\hat{\tau}$
0.4103	0.3593	0.2515

Para el análisis de la dependencia en colas a través de la cópula empírica, es necesario recurrir a una representación gráfica de los valores de los coeficientes de dependencia en las colas superior e inferior, puesto que no existe una expresión concreta para dichos coeficientes en este caso.

Representamos la serie de valores de la función

$$\frac{1 - 2\alpha + C_{emp}(\alpha, \alpha)}{(1 - \alpha)}$$

para valores crecientes de α hacia uno en el caso de la cola superior, y análogamente, representamos la serie de valores de la función

$$\frac{C_{emp}(\alpha, \alpha)}{\alpha}$$

para valores decrecientes de α hacia cero en el caso de la cola inferior.

El análisis del gráfico resultante permite detectar la presencia de dependencia en colas en la cópula empírica y estudiar posibles convergencias en el valor de ambos coeficientes. Como se aprecia en la figura 2.11, el valor del índice en la cola inferior crece de manera lineal para valores de α inferiores a 0.8, a medida que el valor de α se acerca a cero. En el índice superior observamos el mismo comportamiento pero cuando α tiende a la unidad, para valores de α superiores a 0.2. Por tanto, existe evidencia clara de dependencia, tanto en la cola derecha como en la izquierda, de la distribución empírica bivariante.

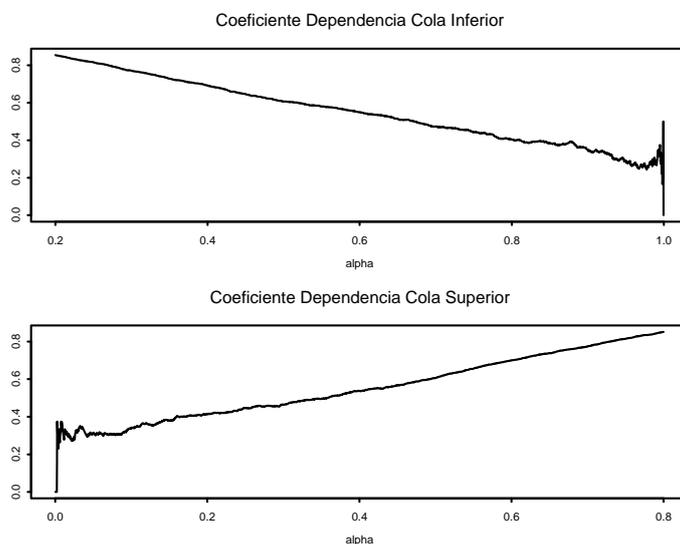


Figura 2.11: *Coeficiente de dependencia en las colas inferior y superior de la cópula empírica DowJones e Ibx35*

2.5.3. Ajuste paramétrico de la cópula bivalente DowJones & Ibex35

Basándonos en los resultados del análisis exploratorio realizado hasta el momento y en el conocimiento de las peculiaridades propias de las series financieras, parece que la elección más acertada debería recaer en alguna cópula perteneciente al grupo de las arquimedianas que presentase dependencia en colas. Proponemos las siguientes familias uniparamétricas arquimedianas para realizar el ajuste: Clayton, Frank y Gumbel (veáse el punto 1.10.2 para un mayor detalle en la descripción de estas familias). Recordemos que la familia de Clayton presenta dependencia en la cola inferior, mientras que la de Gumbel lo hace en la cola superior. La familia de Frank no tiene dependencia en colas, pero parece interesante incluirla en el análisis comparativo como elemento de contraste con el fin de valorar la eficacia de los procedimientos propuestos. Por razones análogas, también está justificada la presencia de la familia gaussiana.

Estimación del parámetro θ mediante el procedimiento de Genest y Rivest (1993)(GR)

Comenzaremos estimando el parámetro θ para las familias arquimedianas consideradas, a partir de la relación existente entre el valor muestral de la τ de Kendall y el generador asociado a la familia de cópulas, que recordemos era la siguiente:

$$\hat{\tau} = 1 + 4 \int_0^1 \frac{\varphi(t)}{\varphi'(t)} dt$$

Para nuestras tres familias esta expresión se concreta en las siguientes igualdades:

Clayton	$\tau = \frac{\theta}{\theta+2}$	$\tau \in [-1, 1] \setminus \{0\}$
Frank	$\tau = 1 - \frac{4}{\theta} [D_1(-\theta) - 1]$	$\tau \in [-1, 1] \setminus \{0\}$
Gumbel	$\tau = 1 - \theta^{-1}$	$\tau \in [0, 1]$

con $D_1(\theta) = \frac{1}{\theta} \int_0^\theta \frac{t}{e^t-1} dt$ (integral de Debye) y $D_1(-\theta) = D_1(\theta) + \frac{\theta}{2}$.

En el caso de la familia de Frank, se ha utilizado para obtener el valor de θ la aproximación de la integral de Debye que aparece en Abramowitz y Stegun (1954).

El valor $\hat{\tau} = 0.2515$ proporciona las siguientes estimaciones de θ :

	$\hat{\theta}$	<i>Err.St</i>	$IC_{0.95}(\theta)$
<i>Clayton</i>	0.6721	0.0444	[0.5851; 0.7591]
<i>Frank</i>	2.34	0.0115	[2.3285; 2.3625]
<i>Gumbel</i>	1.336	0.0222	[1.2925; 1.3795]

Tabla 2.2: *Estimaciones de θ por el procedimiento de Genest y Rivest para los tres modelos arquimedianos junto con sus intervalos de confianza al 95%*

Estimación del parámetro θ mediante máxima verosimilitud canónica (CML)

La estimación del parámetro θ se efectúa ahora para todos los modelos considerados, incluyendo la familia gaussiana, y viene dada por la ecuación (2.4). Los valores proporcionados para θ por la estimación de máxima verosimilitud canónica son:

	$\hat{\theta}$	<i>Err.St</i>	$IC_{0.95}(\theta)$
<i>Clayton</i>	0.5422	0.0267	[0.4899; 0.5945]
<i>Frank</i>	2.4666	0.0203	[2.3766; 2.5063]
<i>Gumbel</i>	1.3354	0.0312	[1.2743; 1.3965]
<i>Gaussiana</i>	0.3925	0.0193	[0.3547; 0.4303]

Tabla 2.3: *Estimaciones de θ por el procedimiento CML para los cuatro modelos considerados junto con sus intervalos de confianza al 95%*

Un resumen de las estimaciones obtenidas por ambos procedimientos, junto con su ordenación utilizando el criterio de Akaike aparece en la tabla 2.4.

Familia de cópulas	Genest y Rivest	CML	AIC
<i>Gumbel</i>	1.336	1.3354	-561.5811
<i>Gaussiana</i>	-	0.3925	-512.2128
<i>Clayton</i>	0.6721	0.5422	-465.6281
<i>Frank</i>	2.34	2.4666	-452.0353

Tabla 2.4: *Ordenación de las estimaciones de θ por ambos procedimientos según AIC*

A partir de estas estimaciones, cuantificamos de nuevo la dependencia muestral recogida por cada modelo ajustado e incorporamos, como un elemento más

de ayuda a la elección de la cópula más adecuada, la información proporcionada por los índices cola correspondientes. Los resultados obtenidos para las estimaciones de θ realizadas mediante el procedimiento de Genest y Rivest se recogen en la tabla 2.5 , junto con los proporcionados por la cópula empírica.

Familia de Cópulas	τ	r_s	λ_l	λ_u
<i>Gumbel</i>	0.2515	0.3657	0	0.3199
<i>Clayton</i>	0.2515	0.3675	0.3565	0
<i>Frank</i>	0.2469	0.3462	0	0
Cópula empírica	0.2515	0.3593	-	-

Tabla 2.5: *Medidas de dependencia muestrales asociadas a las cópulas uniparamétricas estimadas mediante Genest y Rivest*

Procedemos de manera análoga para las estimaciones realizadas mediante máxima verosimilitud (tabla 2.6).

Familia de Cópulas	τ	r_s	λ_L	λ_U
<i>Gumbel</i>	0.2511	0.3652	0	0.3195
<i>Gaussiana</i>	0.2567	0.3925	0	0
<i>Clayton</i>	0.2133	0.3138	0.2784	0
<i>Frank</i>	0.2589	0.3812	0	0
Cópula empírica	0.2515	0.3593	-	-

Tabla 2.6: *Medidas de dependencia muestrales asociadas a las cópulas uniparamétricas estimadas mediante CML*

2.5.4. Evaluación de la bondad de ajuste

Comenzaremos la evaluación de la bondad de ajuste de los modelos uniparamétricos propuestos utilizando **métodos de tipo gráfico**.

1. En primer lugar, comprobamos si $K_C(\cdot)$, dada por la ecuación

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

aplicada a $C(F_{1emp}(x), F_{2emp}(y))$ sigue una $U(0, 1)$.

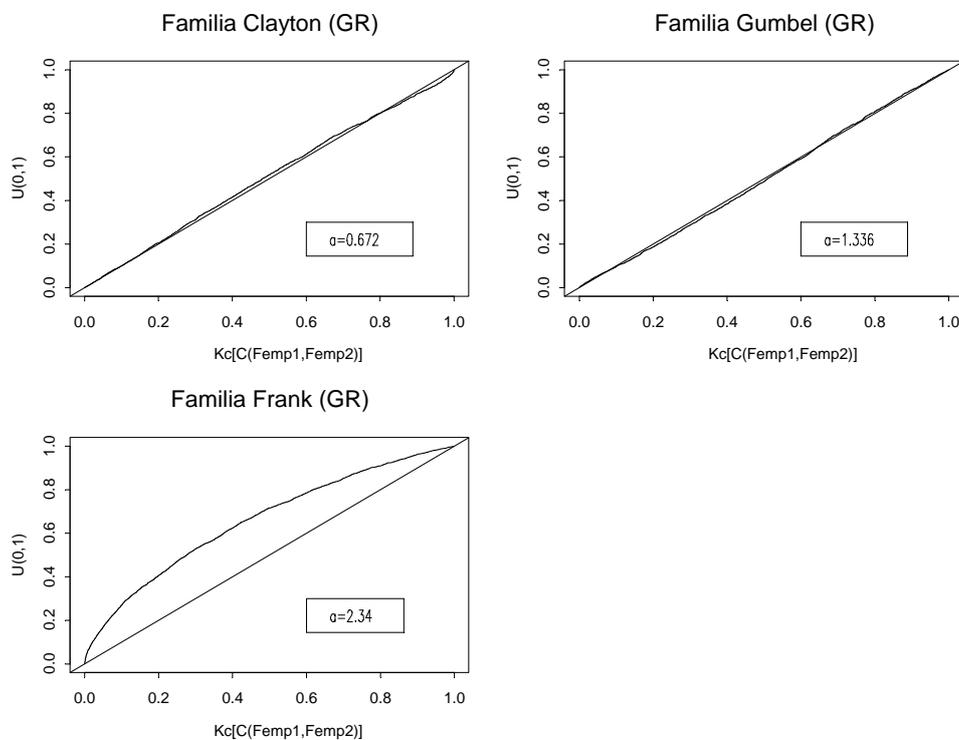


Figura 2.12: Comparación $K_C(C_{\hat{\theta}}(F_{1emp}(x), F_{2emp}(y)))$ con $U(0, 1)$ (Genest y Rivest)

La figura 2.12 muestra los QQ-plots que enfrentan las diferentes estimaciones realizadas mediante el procedimiento de Genest y Rivest a la $U(0, 1)$. Se aprecia como la cópula de Frank resulta la peor elección, mientras que las estimaciones obtenidas para las familias Gumbel y Clayton parecen buenas candidatas para el ajuste del modelo uniparamétrico.

Si procedemos de manera análoga para las estimaciones obtenidas mediante CML, observamos (figura 2.13) resultados muy similares: la familia de Frank queda descartada y la de Clayton y Gumbel ofrecen buenos resultados, apreciándose en esta última un mejor ajuste en la cola derecha.

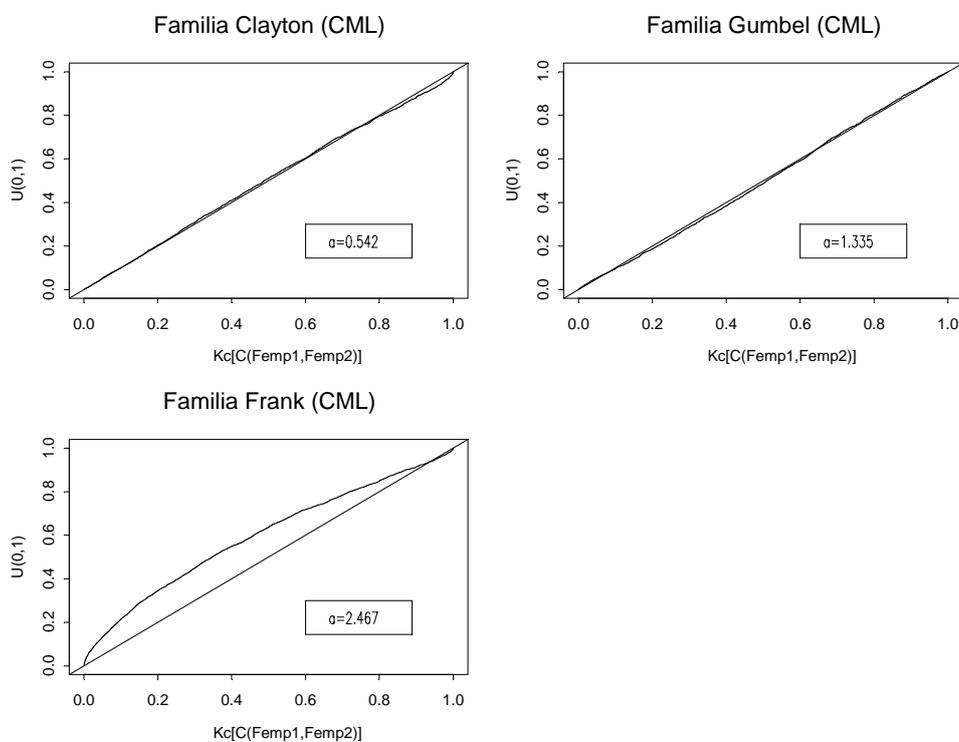


Figura 2.13: Comparación $K_C(C_{\hat{\theta}}(F_{1emp}(x), F_{2emp}(y)))$ con $U(0, 1)$ (CML)

2. A continuación, comparamos la estimación no paramétrica proporcionada por (2.2) con la paramétrica dada por (2.6).

La figura 2.14 muestra los QQ-plots correspondientes a las dos funciones de distribución estimadas mediante el procedimiento de Genest y Rivest. Los resultados indican que, de nuevo, la familia de Frank proporciona una estimación paramétrica más alejada de la empírica que en las otras dos familias. Entre estas últimas, la de Gumbel también presenta un mejor ajuste en la cola derecha.

Encontramos resultados similares al aplicar este procedimiento gráfico a las estimaciones obtenidas mediante máxima verosimilitud. Si observamos la figura 2.15, las familias de Clayton y Gumbel proporcionan buenos resultados mientras que la de Frank se aleja bastante de la distribución empírica.

Por último, completamos el estudio anterior mediante **métodos analíticos**. En realidad, ya hemos usado uno de ellos (criterio de Akaike) para ordenar los

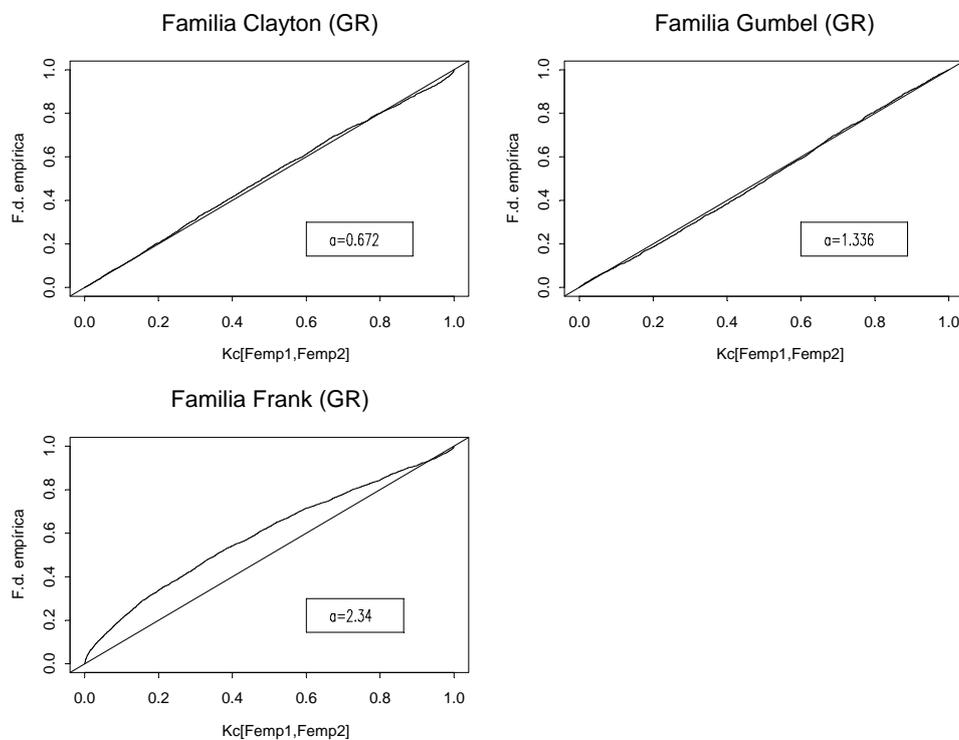


Figura 2.14: Comparación $K_C(C_{\hat{\theta}}(F_{1emp}(x), F_{2emp}(y)))$ y $K_n(w)$ (Genest y Rivest)

modelos ajustados mediante máxima verosimilitud. Ahora proponemos emplear el contraste de Kolmogorov-Smirnov, para una muestra, en el que la hipótesis nula consiste en que $K_C(C_{\hat{\theta}}(F_{1emp}(x), F_{2emp}(y)))$ sigue una distribución $U(0, 1)$. Los p-valores suministrados por el contraste aplicado a los resultados obtenidos mediante los dos procedimientos de estimación, aparecen en la tabla siguiente:

Familia de Cópulas	p-valor GR	p-valor CML
<i>Gumbel</i>	0.2817	0.2822
<i>Clayton</i>	0.0908	0.0434
<i>Frank</i>	0.0	0.0

Si utilizamos la versión del contraste de Kolmogorv-Smirnov para dos muestras con el fin de comparar las distribuciones $K_C(t)$ y $K_n(w)$, en el que la hipótesis nula asume la igualdad de las dos distribuciones, se obtienen en todos los casos

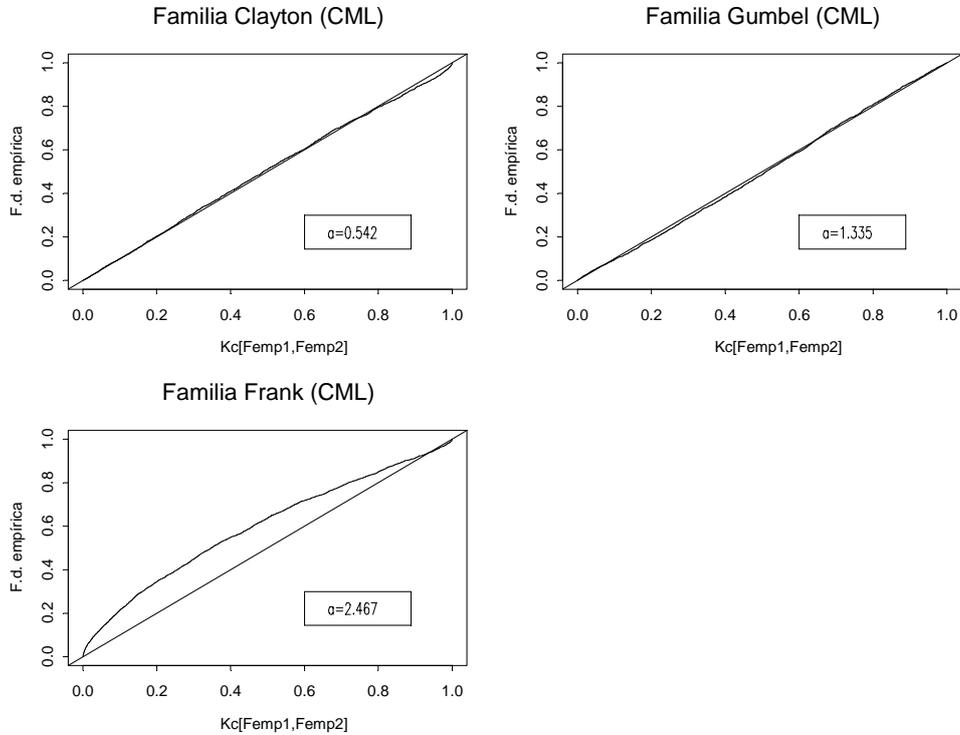


Figura 2.15: Comparación $K_C(C_{\hat{\theta}}(F_{1emp}(x), F_{2emp}(y)))$ y $K_n(w)$ (CML)

p-valores nulos debido a que se trata de un contraste bastante conservador, por lo que en esta situación concreta no aporta información que pueda ayudar en el proceso de selección del modelo.

Por último, podemos preguntarnos si la diferencia entre las estimaciones de θ obtenidas para cada familia mediante los dos procedimientos empleados, tiene influencia sobre los valores suministrados por los modelos ajustados. Una posible medida de la distancia entre las dos estimaciones puede ser

$$\sup_{x,y} \left| C_{\hat{\theta}_1}(u, v) - C_{\hat{\theta}_2}(u, v) \right|$$

Aplicando esta distancia a los diversos modelos ajustados, se obtiene

Familia de Cópulas	$\sup_{x,y} C_{\hat{\theta}_1}(u, v) - C_{\hat{\theta}_2}(u, v) $
<i>Gumbel</i>	0.000094
<i>Clayton</i>	0.010358
<i>Frank</i>	0.003341

A la vista de todo lo anterior, el modelo uniparamétrico de cópula proporcionado por la familia Gumbel se confirma como la mejor elección para representar la relación de dependencia existente entre las series de retornos logarítmicos diarios de los índices DowJones e Ibex35. El modelo de Clayton también proporciona buenos resultados, por lo que debería ser considerado como una posible segunda elección. En cuanto a la cópula gaussiana, aunque no proporciona un mal ajuste en términos de AIC (tabla 2.4), es un modelo que no contempla la existencia de dependencia en colas, por lo cual no parece muy apropiado dada la naturaleza de nuestros datos. No existe duda en cuanto a los malos resultados que ofrece el modelo perteneciente a la familia de Frank, por lo que queda claramente desechado para nuestros propósitos.

2.5.5. Modelos alternativos de cópulas mixtas

Teniendo en cuenta las conclusiones a las que hemos llegado en el ajuste de cópulas uniparamétricas, así como la información proporcionada por el modelo de cópula empírica extraído de la muestra bidimensional de retornos, podemos mejorar los resultados obtenidos hasta el momento proponiendo algún modelo de cópula mixta que capture con más precisión los rasgos peculiares del comportamiento conjunto de los retornos de los dos índices considerados. En concreto, nos referimos a la dependencia en ambas colas de la distribución bivalente empírica. El modelo perteneciente a la familia Gumbel, que ha resultado preferido en el análisis previo, contempla la dependencia en la cola superior. La situación inversa aparece al considerar como elección el modelo de Clayton, en el que la dependencia reflejada es la tiene lugar en la cola inferior. Parece, por tanto, que una mezcla de ambos modelos operaría de manera más eficiente sobre nuestra muestra bivalente de retornos. Por otra parte, puesto que el modelo uniparamétrico que ha ofrecido los mejores resultados ha sido el perteneciente a la familia Gumbel, otra opción interesante la constituye la mezcla de un elemento

de la familia Gumbel con un modelo de cópula de supervivencia (veáse el punto 2.6) asociado a dicha familia.

Si denotamos por $C(u, v; \boldsymbol{\theta})$ el modelo de cópula mixta, consideramos combinaciones lineales convexas de la forma:

$$C(u, v; \boldsymbol{\theta}) = \theta_3 C_{Fam1}(u, v; \theta_1) + (1 - \theta_3) C_{Fam2}(u, v; \theta_2)$$

que se concretarán en dos modelos mixtos:

	<i>Fam1</i>	<i>Fam2</i>
<i>Modelo 1</i>	Gumbel	Gumbel supervivencia
<i>Modelo 2</i>	Gumbel	Clayton

Específicamente, el modelo 1 tiene como expresión

$$C_{Mod1}(u, v; \boldsymbol{\theta}) = \theta_3 C_{Gu}(u, v; \theta_1) + (1 - \theta_3)(u + v - 1 + C_{Gu}(1 - u, 1 - v; \theta_2))$$

donde C_{Gu} es la cópula de Gumbel $(u, v) \in [0, 1]^2$ y $\theta_1, \theta_2 \in [1, +\infty)$.

Para el modelo 2 la formulación es

$$C_{Mod2}(u, v; \boldsymbol{\theta}') = \theta'_3 C_{Gu}(u, v; \theta'_1) + (1 - \theta'_3) C_{Cl}(u, v; \theta'_2)$$

con C_{Cl} la cópula Clayton $(u, v) \in [0, 1]^2$, $\theta'_1 \in [1, +\infty)$ y $\theta'_2 \in [-1, +\infty) \setminus \{0\}$.

Proposición 2.4

$C_{Mod1}(u, v; \boldsymbol{\theta})$ y $C_{Mod2}(u, v; \boldsymbol{\theta}')$ son cópulas.

Demostración:

Es inmediata a partir de la definición 1.4 y del hecho de que C_{Gu} y C_{Cl} son cópulas.

□

Los modelos así construidos vuelven a ser cópulas y, por tanto, disfrutan de todas las propiedades de las mismas. Los parámetros θ_1 y θ_2 (θ'_1 y θ'_2) administran el grado de dependencia en la cola superior e inferior, respectivamente, y el parámetro θ_3 (θ'_3) asigna el peso del primer elemento de la combinación lineal convexa y fija la estructura conjunta.

El ajuste se realiza mediante máxima verosimilitud, en la forma usual

$$\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta} \in \Theta} \sum_{i=1}^n \log \mathcal{L}(\boldsymbol{\theta}; u_i, v_i)$$

con $u_i = F_{1emp}(x_i)$ y $v_i = F_{2emp}(y_i)$.

Los modelos mixtos ajustados se comparan con los uniparamétricos mediante el criterio AIC. También se utilizan para calcular los coeficientes de dependencia en colas (recordemos que estos coeficientes son una característica de la cópula, por lo que exclusivamente es necesaria la función cópula correspondiente a cada modelo para calcularlos).

Proposición 2.5

Existen los coeficientes de dependencia en ambas colas (superior e inferior) para $C_{Mod1}(u, v; \theta)$ y $C_{Mod2}(u, v; \theta')$, y sus expresiones son, para $C_{Mod1}(u, v; \theta)$:

$$\begin{aligned}\lambda_{L,1} &= (1 - \theta_3)(2 - 2^{1/\theta_2}) \\ \lambda_{U,1} &= \theta_3(2 - 2^{1/\theta_1})\end{aligned}$$

y para $C_{Mod2}(u, v; \theta')$:

$$\begin{aligned}\lambda_{L,2} &= (1 - \theta'_3)2^{-1/\theta'_2} \\ \lambda_{U,2} &= \theta'_3(2 - 2^{1/\theta'_1})\end{aligned}$$

Demostración:

A partir de la definición de la cópula de Gumbel, dada por la expresión (1.6), se tiene que

$$C_{Mod1}(u, u; \theta) = \theta_3 u^{2^{\frac{1}{\theta_1}}} + (1 - \theta_3) \left(2u - 1 + (1 - u)^{2^{\frac{1}{\theta_2}}} \right) \quad (2.7)$$

Sustituyendo (2.7) en la expresión (1.4), obtenemos el coeficiente de dependencia en la cola inferior para el modelo 1:

$$\begin{aligned}\lambda_{L,1} &= \lim_{u \rightarrow 0^+} \left[\frac{\theta_3 u^{2^{\frac{1}{\theta_1}}} + (1 - \theta_3) \left(2u - 1 + (1 - u)^{2^{\frac{1}{\theta_2}}} \right)}{u} \right] \\ &= \theta_3 \lim_{u \rightarrow 0^+} \left[\frac{u^{2^{\frac{1}{\theta_1}}}}{u} \right] + (1 - \theta_3) \lim_{u \rightarrow 0^+} \left[\frac{2u - 1 + (1 - u)^{2^{\frac{1}{\theta_2}}}}{u} \right] \\ &= (1 - \theta_3) \lim_{u \rightarrow 0^+} \left[2 - 2^{\frac{1}{\theta_2}} (1 - u)^{2^{\frac{1}{\theta_2}} - 1} \right] \\ &= (1 - \theta_3)(2 - 2^{1/\theta_2})\end{aligned}$$

Análogamente, sustituyendo (2.7) en la expresión (1.3), obtenemos el coeficiente de dependencia en la cola superior para el modelo 1:

$$\begin{aligned}
\lambda_{U,1} &= \lim_{u \rightarrow 1^-} \left[\frac{1 - 2u + \theta_3 u^{2^{\frac{1}{\theta_1}}} + (1 - \theta_3) \left(2u - 1 + (1 - u)^{2^{\frac{1}{\theta_2}}} \right)}{1 - u} \right] \\
&= \lim_{u \rightarrow 1^-} \left(\frac{1 - 2u}{1 - u} \right) + \theta_3 \lim_{u \rightarrow 1^-} \left(\frac{u^{2^{\frac{1}{\theta_1}}}}{1 - u} \right) + (1 - \theta_3) \lim_{u \rightarrow 1^-} \left[\frac{2u - 1 + (1 - u)^{2^{\frac{1}{\theta_2}}}}{1 - u} \right] \\
&= 2 - \theta_3 \lim_{u \rightarrow 1^-} 2^{\frac{1}{\theta_1}} u^{2^{\frac{1}{\theta_1}} - 1} + (1 - \theta_3) \lim_{u \rightarrow 1^-} \left[2 - 2^{\frac{1}{\theta_2}} (1 - u)^{2^{\frac{1}{\theta_2}} - 1} \right] \\
&= 2 - \theta_3 2^{\frac{1}{\theta_1}} - 2(1 - \theta_3) \\
&= \theta_3 \left(2 - 2^{\frac{1}{\theta_1}} \right)
\end{aligned}$$

Procediendo de manera similar, se obtiene para el modelo 2

$$C_{Mod1}(u, u; \theta') = \theta'_3 u^{2^{\frac{1}{\theta'_1}}} + (1 - \theta'_3) \left(2u^{-\theta'_2} - 1 \right)^{-\frac{1}{\theta'_2}} \quad (2.8)$$

Sustituyendo (2.8) en la expresión (1.4), obtenemos el coeficiente de dependencia en la cola inferior para el modelo 2:

$$\begin{aligned}
\lambda_{L,2} &= \lim_{u \rightarrow 0^+} \left[\frac{\theta'_3 u^{2^{\frac{1}{\theta'_1}}} + (1 - \theta'_3) \left(2u^{-\theta'_2} - 1 \right)^{-\frac{1}{\theta'_2}}}{u} \right] \\
&= \theta'_3 \lim_{u \rightarrow 0^+} \left(\frac{u^{2^{\frac{1}{\theta'_1}}}}{u} \right) + (1 - \theta'_3) \lim_{u \rightarrow 0^+} \left[\frac{\left(2u^{-\theta'_2} - 1 \right)^{-\frac{1}{\theta'_2}}}{u} \right]
\end{aligned}$$

El límite incluido en el primer sumando es cero y el correspondiente al segundo sumando coincide con el coeficiente de dependencia en la cola inferior de la cópula Clayton, por lo que:

$$\lambda_{L,2} = (1 - \theta'_3) 2^{-\frac{1}{\theta'_2}}$$

Finalmente, sustituyendo (2.8) en la expresión (1.3), obtenemos el coeficiente

de dependencia en la cola superior para el modelo 2:

$$\begin{aligned}
 \lambda_{U,2} &= \lim_{u \rightarrow 1^-} \left[\frac{1 - 2u - \theta'_3 u^{2^{\frac{1}{\theta'_1}}} + (1 - \theta'_3) (2u^{-\theta'_2} - 1)^{-\frac{1}{\theta'_2}}}{1 - u} \right] \\
 &= \lim_{u \rightarrow 1^-} \left(\frac{1 - 2u}{1 - u} \right) + \theta'_3 \lim_{u \rightarrow 1^-} \left(\frac{u^{2^{\frac{1}{\theta'_1}}}}{1 - u} \right) + (1 - \theta'_3) \lim_{u \rightarrow 1^-} \left[\frac{(2u^{-\theta'_2} - 1)^{-\frac{1}{\theta'_2}}}{1 - u} \right] \\
 &= 2 + \theta'_3 \lim_{u \rightarrow 1^-} \left[-2^{\frac{1}{\theta'_1}} u^{2^{\frac{1}{\theta'_1}} - 1} \right] + \\
 &\quad + (1 - \theta'_3) \lim_{u \rightarrow 1^-} (-1) \left(-\frac{1}{\theta'_2} \right) (2u^{-\theta'_2} - 1)^{-\frac{1}{\theta'_2} - 1} 2 (-\theta'_2) u^{-\theta'_2 - 1} \\
 &= 2 - \theta'_3 2^{\frac{1}{\theta'_1}} + (1 - \theta'_3) \frac{1}{\theta'_2} 2 (-\theta'_2) = 2 - \theta'_3 2^{\frac{1}{\theta'_1}} - 2(1 - \theta'_3) \\
 &= \theta'_3 \left(2 - 2^{\frac{1}{\theta'_1}} \right)
 \end{aligned}$$

□

En la tabla 2.7 aparecen las estimaciones de máxima verosimilitud para los parámetros de los dos modelos de cópulas mixtas, así como su ordenación según el criterio de Akaike junto con el valor de los coeficientes de dependencia en colas inferior y superior.

Podemos apreciar que el modelo 1 es el que proporciona mejores resultados, y si comparamos los modelos mixtos con los uniparamétricos ajustados, ambos operan mejor que cualquiera de los considerados.

Familia de cópulas	CML	AIC	λ_L	λ_U
<i>Modelo 1</i>	1.251 ; 1.252 ; 0.497	-652.0370	0.1310	0.1293
<i>Modelo 2</i>	1.171 ; 1.021 ; 0.624	-624.7183	0.1906	0.1205
<i>Gumbel</i>	1.3354	-561.5811	0	0.3199
<i>Gaussiana</i>	0.3925	-512.2128	0	0
<i>Clayton</i>	0.5422	-465.6281	0.2784	0
<i>Frank</i>	2.4666	-452.0353	0	0

Tabla 2.7: Ordenación de todos los modelos estimados mediante CML según AIC

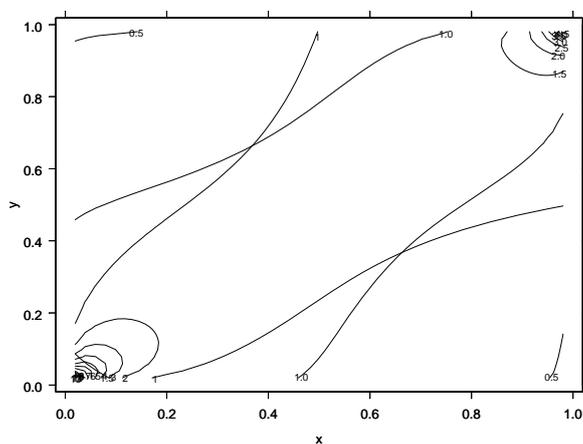


Figura 2.16: Diagrama de contorno de la cópula mixta Gumbel & Clayton en el que se aprecia la dependencia en ambas colas

Filtrado de la Dependencia Temporal Marginal

Uno de los objetivos fundamentales de nuestro estudio es analizar el comportamiento del riesgo inherente a los mercados financieros. Con este fin, abordamos en el presente capítulo la cuestión relativa a la elección de los *modelos marginales* más adecuados a cada índice considerado. El planteamiento básico del que partimos es que cada serie de retornos sigue un modelo de serie temporal estacionario con estructura de volatilidad estocástica. La presencia de volatilidad estocástica implica, como ya es conocido, que los retornos de índices financieros no son necesariamente independientes en el tiempo. Por tanto, trabajando con estos modelos existen dos tipos de distribuciones de retornos a considerar: la *distribución condicionada*, donde la condición la establece la volatilidad actual, y la *distribución marginal* o estacionaria del proceso. Ambas distribuciones son relevantes para el control del riesgo. Las colas de la distribución de retornos condicionada son esencialmente objeto de interés en el cálculo de medidas de riesgo de mercado y será donde se concentre nuestro trabajo en el capítulo siguiente, una vez filtrada la dependencia temporal marginal. La clave es la posible extensión de una pérdida originada por un movimiento de mercado adverso sobre el día (días) siguiente, dada la volatilidad registrada hasta el momento actual. La estimación de colas no condicionadas proporciona información diferente, pero complementaria, sobre el riesgo.

Esquemáticamente, las aproximaciones existentes para estimar la distribución de beneficio-pérdida (P&L) de una cartera se pueden dividir en tres grupos:

1. Método de simulación histórica no paramétrica (HS).
2. Métodos paramétricos basados en un modelo econométrico para volatilidades y en la suposición de normalidad condicionada (Riskmetrics de J.P. Morgan 1995, y la mayoría de modelos de la familia ARCH/GARCH).
3. Métodos basados en Teoría de valores extremos (EVT).

En la *aproximación HS* (Crouhy, Galai y Mark 2001; Marshall 2001; Vilarinho 2001), la distribución P&L estimada de una cartera viene dada simplemente por la distribución empírica de pérdidas y ganancias pasadas. El método es fácil de implementar y evita suposiciones sobre la forma de la distribución. Sin embargo, este método dificulta notoriamente la estimación de los cuantiles extremos, así como la extrapolación más allá de las observaciones pasadas. Es más, las estimaciones de cuantiles obtenidas por HS tienden a ser muy volátiles siempre que aparece una observación grande en la muestra. Finalmente, el método es incapaz de distinguir entre periodos de alta y baja volatilidad, en particular si se utiliza una muestra grande para mitigar la influencia de los dos primeros problemas en la calidad de las estimaciones en colas.

Los modelos más refinados que incluyen la aproximación de normalidad condicionada, tales como los de la *familia GARCH* (Engle 1982, 2000, Bollerslev, Chow y Kroner 1992, Bollerslev, Engle y Nelson, 1994, Gouriéroux 1997) con modelos dinámicos de varianza condicionada de los retornos, proporcionan estimaciones del VaR que reflejan la volatilidad registrada hasta el momento actual. La debilidad principal de esta aproximación está en suponer la normalidad condicionada, lo cual no parece ajustarse a la mayoría de los datos reales.

Los *métodos basados en EVT* (Embretchs y otros, 1997) tienen dos características que los hacen atractivos para la estimación de las colas: están basados en una teoría estadística sólida y proporcionan una forma paramétrica para la cola de la distribución. Por tanto, estos métodos permiten extrapolaciones más allá del rango de los datos. Sin embargo, ninguno de los métodos basados en EVT para estimación de cuantiles proporciona estimaciones VaR que reflejen la historia de la volatilidad actual. Dada la heterocedasticidad condicionada presente en la mayoría de las series financieras, bien documentada por el considerable éxito de los modelos de la familia ARCH-GARCH, creemos que esta será una mejor elección para cualquier tipo de estimación del VaR. No obstante, no desechamos estos métodos sino que serán utilizados en una fase posterior, detallada en el

capítulo 4, donde emplearemos los llamados *métodos umbral* (POT) a partir de EVT para estimar la distribución de las colas de los términos de error correspondientes a las distribuciones marginales filtradas mediante procesos tipo GARCH (McNeil y Frey, 2000).

Por todo lo anteriormente expuesto, nuestra propuesta para representar el comportamiento de las distribuciones marginales de los retornos incluye dos fases: en primer lugar utilizaremos modelos tipo GARCH y máxima verosimilitud para obtener estimaciones de la volatilidad condicionada, para proceder, a continuación, a confirmar a través de análisis de datos exploratorio y los contrastes estadísticos pertinentes, que los términos de error (residuos) constituyen, al menos aproximadamente, una serie i.i.d. con colas pesadas.

Los resultados obtenidos permitirán construir una estimación de la distribución condicionada de los retornos a partir de la distribución de los residuos y las estimaciones de la volatilidad y de la media condicionada. Esta aproximación refleja dos rasgos presentes en la mayoría de las series de retornos financieros: la volatilidad estocástica y las colas pesadas de las distribuciones condicionadas de los retornos sobre horizontes temporales cortos.

De manera formal, sea $(X_t)_{t \in N}$ una serie temporal estrictamente estacionaria que representa las observaciones diarias de retornos generadas por precios de activos financieros.

Suponemos que esta serie viene generada por

$$X_t = \mu_t + \sigma_t Z_t \tag{3.1}$$

donde las innovaciones Z_t son un proceso estricto de ruido blanco (iid) de media cero, varianza 1, función de distribución marginal $F_Z(z)$ y con μ_t y σ_t medibles respecto a G_{t-1} (información sobre el proceso de retornos hasta el instante (t-1)).

Sea $F_X(x)$ la distribución marginal de cada X_t y $F_{X_{t+1}+\dots+X_{t+k}/G_t}(x)$ la distribución predictiva de los retornos de los k días siguientes, suponiendo conocidos los retornos hasta el día t.

Nuestro interés consiste en estimar cuantiles de las colas de este tipo de distribuciones, lo cual permitirá determinar las medidas de riesgo financiero usuales, tales como el VaR o la Pérdida esperada, tal y como se detalla a continuación.

Sabemos que, para $0 < q < 1$, se define el *cuantil no condicionado* de F como

$$x_q = \inf \{x \in \mathbb{R} : F_X(x) \geq q\}$$

por tanto, el *cuantil condicionado* será un cuantil de la distribución predictiva de retornos sobre los siguientes k días:

$$x_q^t(k) = \inf \{x \in \mathbb{R} : F_{X_{t+1}+\dots+X_{t+k}/G_t}(x) \geq q\}$$

Puesto que una medida alternativa de riesgo viene dada por la llamada *pérdida esperada* (ES), si consideramos la *pérdida esperada no condicionada*, tendremos:

$$S_q = E[X/X > x_q]$$

y la *pérdida esperada condicionada* será:

$$S_q^t(k) = E \left[\sum_{j=1}^k X_{t+j} / \sum_{j=1}^k X_{t+j} > x_q^t(k), G_t \right]$$

Para simplificar la exposición, podemos limitarnos al cálculo del VaR y la pérdida esperada en distribuciones predictivas que contemplan un solo instante, para las que utilizaremos la notación x_q^t y S_q^t .

Puesto que

$$F_{X_{t+1}/G_t}(x) = P \{ \mu_{t+1} + \sigma_{t+1} Z_{t+1} \leq x / G_t \} = F_Z \left(\frac{x - \mu_{t+1}}{\sigma_{t+1}} \right)$$

estas medidas se simplifican y quedan como:

$$\begin{aligned} x_q^t &= \mu_{t+1} + \sigma_{t+1} z_q \\ S_q^t &= \mu_{t+1} + \sigma_{t+1} E[Z/Z > z_q] \end{aligned}$$

ambas dependiendo de z_q , q -cuantil superior de la distribución marginal de Z_t que por hipótesis no depende de t . Por tanto, nuestra tarea consistirá en determinar z_q , valor que permitirá el cálculo de las medidas de riesgo anteriormente citadas.

El proceso de filtrado de los retornos marginales se puede esquematizar en el desarrollo de las siguientes fases:

1) En primer lugar elegimos un proceso específico entre los que verifican (3.1), es decir, un modelo particular para la varianza y media condicionadas. Podemos decantarnos por un tipo GARCH para la varianza y un ARMA, AR o incluso un valor constante para la media. Procedemos al ajuste del modelo sin establecer ninguna suposición sobre $F_Z(z)$, distribución de las innovaciones.

2) Estimamos μ_{t+1} y σ_{t+1} a partir del modelo ajustado y obtenemos los residuos, considerándolos como una realización de un ruido blanco estricto.

3) Comprobamos la no normalidad de los residuos (mediante los contrastes indicados para esta situación y, por ejemplo, gráficamente a través de un QQ-plot de los mismos frente a la Normal) y la existencia de colas pesadas y curtosis elevada.

4) Postulamos alguna distribución conocida para los residuos, de manera que responda al patrón de colas pesadas y curtosis elevada. En concreto, proponemos la distribución t-Student transformada para tener varianza 1. Los modelos tipo GARCH con innovaciones siguiendo una t de Student pueden ajustarse mediante máxima verosimilitud. Los resultados serán más satisfactorios cuanto más iguales sean en cuanto a longitud las colas positiva y negativa de la distribución de los retornos.

5) Comparamos de los resultados obtenidos con el caso de innovaciones distribuidas normalmente, que es de esperar proporcionen un peor ajuste, dada la presencia de colas pesadas.

6) Aplicamos los procedimientos inferenciales propuestos en el Capítulo 2, con el fin de estimar el modelo de cópula que mejor capture la relación de dependencia presente en la muestra bivalente de retornos filtrados según los procesos tipo GARCH. Para finalizar, comparamos las conclusiones obtenidas con las procedentes del tratamiento de la muestra bivalente de retornos brutos.

3.1. Modelos de volatilidad

Hasta hace una década, la mayoría de los modelos relativos a series temporales financieras y macroeconómicas se centraban en los primeros momentos condicionados, y cualquier dependencia temporal en momentos de orden superior se trataba como una perturbación. Sin embargo, la importancia creciente del

riesgo y las consideraciones de incertidumbre en la teoría económica moderna, han llevado a desarrollar nuevas técnicas de series temporales en econometría que permiten establecer modelos para varianzas y covarianzas que varían en el tiempo. Un instrumento importante en este desarrollo lo constituyen la clase de modelos para heterocedasticidad condicionada autorregresiva (ARCH) introducidos por Engle (1982). Paralelamente al éxito de los modelos de series temporales clásicos, en los que aparece el uso de la media condicionada frente a la no condicionada, la clave de los modelos ARCH está en la incorporación de los momentos de segundo orden condicionados frente a los no condicionados. Mientras que en aquellos modelos la matriz de covarianzas no condicionadas para las variables de interés puede ser invariante en el tiempo, las varianzas y covarianzas condicionadas a menudo dependen, y no trivialmente, de los estados pasados del universo. La comprensión de la naturaleza exacta de esta dependencia temporal es crucial para muchas series financieras y macroeconómicas tales como las relativas a inversiones, valoración de opciones, la estructura de los índices de interés y la dinámica general de las relaciones entre valoraciones de activos financieros. Por tanto, desde la perspectiva de la inferencia econométrica, la pérdida de eficiencia asintótica por el descuido de la heterocedasticidad puede ser arbitrariamente grande y, cuando se efectúan las predicciones económicas, sería deseable obtener estimaciones más precisas del error de predicción condicionando por el conjunto de información que se tiene hasta el momento. Muchos procedimientos estadísticos han sido diseñados con la intención de que fueran robustos frente a grandes errores, pero ninguno había utilizado la existencia de agrupamientos temporales de datos atípicos para predecir su ocurrencia y minimizar sus efectos. Esta es exactamente la aproximación que hacen los modelos tipo ARCH.

3.1.1. Características de las Series Financieras

La predicción de rentabilidades de activos financieros, tales como stocks e índices de cambio, es de interés obvio en la práctica financiera (Alexander 2001). La realidad es que rara vez se generan predicciones seguras, en muchas ocasiones ni tan siquiera es posible predecir el signo de la rentabilidad del día siguiente. La relación entre rentabilidad y riesgo juega un papel prominente en muchas teorías y modelos financieros, como por ejemplo la valoración de opciones. Dado que la volatilidad es vista a menudo como una medida de este riesgo (cuando se aceptan distribuciones normales, condicionadas o no, el riesgo es una función de la volati-

lidad), se puede estar interesado no sólo en obtener predicciones de rentabilidades sobre activos financieros, sino también predicciones de la volatilidad asociada. La evidencia reciente muestra que la volatilidad de activos financieros no es constante, sin embargo sí se manifiesta una cierta regularidad, en el sentido de que periodos de alta volatilidad alternan con otros más tranquilos, por lo que parece posible predecir este riesgo que varía a lo largo del tiempo. La mayoría de los modelos comúnmente utilizados en la práctica financiera para describir rentabilidades y volatilidades son lineales. Existen, sin embargo, distintos indicios de que pudieran ser más apropiados modelos no lineales (Pagan 1996, Frey 1997). Uno de ellos es el hecho de que los datos de series temporales financieras presentan un *comportamiento asimétrico* (por ejemplo, rentabilidades negativas grandes aparecen más frecuentemente que rentabilidades positivas grandes). Otro sería que la aparición de rentabilidades negativas grandes es, a menudo, prelude de periodos de volatilidad sustancial, mientras que la aparición de rentabilidades positivas grandes genera volatilidades menos significativas. Una cuestión importante es la de discernir cuando estas rentabilidades negativas grandes son simplemente sucesos atípicos, o cuando son consecuencia natural de un proceso subyacente que sigue un modelo no lineal que debe ser identificado, modelo que se puede utilizar no sólo para describir el patrón observado en la serie financiera, sino más importante aún, con el fin de obtener las mejores predicciones. Huelga decir que tales peculiaridades deberían incorporarse en el modelo de serie temporal utilizado para la descripción y predicción; en otro caso, nos arriesgamos a obtener predicciones demasiado por debajo o demasiado por encima. Incluso en el caso univariante, el conjunto de formas funcionales permitidas es amplio, y se puede acomodar a una gran familia paramétrica de modelos ARCH. Por tanto, es conveniente tener conocimiento de las regularidades empíricas que cada modelo recoge para seleccionar el ARCH apropiado.

A partir de una sencilla inspección gráfica y descriptiva de las observaciones temporales, podemos resumir estos rasgos característicos de las series financieras de la siguiente forma:

- (i) *Los valores atípicos aparecen más veces de lo esperado.*
- (ii) *Las puntuaciones atípicas generalmente son negativas.*
- (iii) *Los valores atípicos tienden a presentarse en clusters.*
- (iv) *Valores atípicos negativos provocan inestabilidades significativas.*

(Rentabilidades del mercado de stock negativas grandes provocan volatilidades grandes)

(v) Los periodos inactivos incrementan la volatilidad en la reapertura.

(vi) Movimientos conjuntos de volatilidades.

Para finalizar conviene comentar la relación existente entre variables macroeconómicas y volatilidad. Puesto que los valores del stock están fuertemente ligados a la salud de la economía, es natural esperar que medidas de la incertidumbre macroeconómica tales como las varianzas condicionadas de la producción industrial, índices de interés, crecimiento monetario, etc. deberían ayudar a explicar los cambios en la volatilidad del mercado de stocks. Esta idea refuerza el interés en la obtención de modelos tipo ARCH, que representan el comportamiento de estos momentos de segundo orden condicionados, bajo esta nueva óptica macroeconómica.

3.1.2. Formulación del Modelo ARCH

Si una variable aleatoria y_t se observa a partir de la función de densidad condicionada $f\left(\frac{y_t}{y_{t-1}}\right)$, la predicción de valores actuales basada en la información pasada, bajo suposiciones clásicas, es simplemente $E\left(\frac{y_t}{y_{t-1}}\right)$, que depende del valor de la variable que condiciona, y_{t-1} . La varianza de la predicción en ese periodo es $V\left(\frac{y_t}{y_{t-1}}\right)$. Tal expresión reconoce que la predicción de la varianza condicionada depende de la información pasada y, es más, puede ser una variable aleatoria.

Consideremos inicialmente la autorregresión de primer orden:

$$y_t = \gamma y_{t-1} + \varepsilon_t$$

donde $\{\varepsilon_t\}$ es un ruido blanco con $E(\varepsilon_t) = \sigma^2$.

La media condicionada de y_t es γy_{t-1} , y la media no condicionada es cero. La varianza condicionada de y_t es σ^2 , mientras que la varianza no condicionada es $\frac{\sigma^2}{1 - \gamma^2}$.

La aproximación clásica en la representación de la heterocedasticidad consiste

en introducir una variable exógena x_t que predice la varianza. Con media cero conocida, el modelo será:

$$y_t = \varepsilon_t x_{t-1}$$

donde, de nuevo $V(\varepsilon_t) = \sigma^2$ y $V(y_t) = \sigma^2 x_{t-1}^2$.

Esta solución del problema no parece satisfactoria, ya que lo que se requiere es una especificación de las causas de los cambios de la varianza, más que el reconocimiento de que la media y la varianza condicionadas pueden variar en el tiempo. Quizás a causa de esta dificultad, rara vez aparecen correcciones de la heterocedasticidad en modelos de series temporales.

Un modelo preferible es:

$$\begin{aligned} y_t &= \varepsilon_t h_t^{\frac{1}{2}} \\ h_t &= \alpha_0 + \alpha_1 y_{t-1}^2 \end{aligned}$$

con $V(\varepsilon_{t-1}) = 1$.

Esto es un ejemplo de lo que será llamado un *Modelo Autorregresivo de Heterocedasticidad Condicionada* (ARCH). No es exactamente un modelo bilineal, pero se acerca mucho a uno. Añadiendo la suposición de normalidad, se puede expresar más directamente en términos de el *conjunto de información pasada disponible hasta el instante t*, que denotaremos por G_t . Usando densidades condicionadas:

$$\begin{aligned} \frac{y_t}{G_{t-1}} &\sim N(0, h_t) \\ h_t &= \alpha_0 + \alpha_1 y_{t-1}^2 \end{aligned}$$

La función de la varianza se puede expresar de forma más general como:

$$h_t = h(y_{t-1}, y_{t-2}, \dots, y_{t-p}, \alpha)$$

donde p es el *orden del proceso* ARCH y α es un vector de parámetros desconocidos.

El *modelo de regresión* ARCH se obtiene suponiendo que la media de y_t viene dada por $x_t \beta$, una combinación lineal de variables endógenas y exógenas retardadas incluidas en el conjunto de información G_{t-1} , con β un vector de parámetros desconocidos.

Formalmente:

$$\begin{aligned}y_t/G_{t-1} &\sim N(x_t\beta, h_t) \\h_t &= h(\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-p}, \alpha) \\ \varepsilon_t &= y_t - x_t\beta\end{aligned}\tag{3.2}$$

La función de la varianza se puede generalizar para que incluya x actuales y retardadas como parte del conjunto de información. La función h queda, entonces, como:

$$h_t = h(\varepsilon_{t-1}, \dots, \varepsilon_{t-p}, x_t, x_{t-1}, \dots, x_{t-p}, \alpha)$$

o, sencillamente:

$$h_t = h(G_{t-1}, \alpha)$$

Esta generalización representa una extensión sencilla de los resultados.

En particular, si la función h se factoriza de la forma:

$$h_t = h_\varepsilon(\varepsilon_{t-1}, \dots, \varepsilon_{t-p}, \alpha)h_x(x_t, \dots, x_{t-p})$$

los dos tipos de heterocedasticidad pueden ser tratados secuencialmente, primero corrigiendo la componente x y, a continuación, ajustando el modelo ARCH sobre los datos transformados.

El modelo formulado en (3.2) tiene una serie de características que lo hacen atractivo para aplicaciones econométricas, puesto que es útil para captar los rasgos peculiares de las series financieras, ya comentados en el punto anterior, y permite predecir la varianza cambiante en el tiempo basándose en errores pasados. A partir de suposiciones sencillas, los beneficios de carteras de activos financieros se toman como funciones de las medias y varianzas esperadas de las rentabilidades. Cualquier cambio en la demanda de activos debe asociarse con cambios en estas medias y varianzas esperadas. Si se asume que la media sigue una regresión o un modelo de serie temporal clásicos, inmediatamente la varianza se ve restringida a ser constante en el tiempo. Como ya se ha comentado, la alternativa de utilizar variables exógenas para explicar cambios en la varianza, no es apropiada en la mayoría de los casos, lo que hace del modelo ARCH el más indicado en estas situaciones.

Dada $(X_t)_{t \in \mathbb{N}}$ sucesión de variables aleatorias, la formulación más utilizada para un modelo ARCH de orden p , $p \in \mathbb{N}$, es la siguiente:

$$\begin{aligned} X_t &= \mu + \epsilon_t \\ \epsilon_t &= \sigma_t Z_t \\ h_t &= \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \dots + \alpha_p \epsilon_{t-p}^2 \end{aligned} \tag{3.3}$$

donde $(Z_t)_{t \in \mathbb{Z}}$ es una sucesión de variables aleatorias i.i.d de media cero y varianza unidad, $\alpha_0 > 0$ y $\alpha_i \geq 0$ para $i = 1, 2, \dots, p$.

3.1.3. Formulación del Modelo GARCH

En aplicaciones empíricas del modelo ARCH a menudo se requiere un retardo relativamente largo en la ecuación de la varianza condicionada y, para evitar problemas con la varianza de las estimaciones de los parámetros, generalmente se impone una estructura de retardos fija. Resulta, entonces, de inmediato interés práctico extender la clase de modelos ARCH de forma que permita simultáneamente una memoria más larga y una estructura de retardos más flexible. Esto se consigue con la clase de modelos GARCH(p,q) (Modelo Generalizado de Heterocedasticidad Condicionada Autorregresiva).

Sean:

- $\varepsilon_t \equiv$ proceso estocástico en tiempo discreto con valores reales
- $\psi_t \equiv$ Conjunto (σ -álgebra) de toda la información hasta el instante t

El modelo GARCH(p,q) viene definido por:

$$\varepsilon_t / \psi_{t-1} \sim N(0, h_t) \tag{3.4}$$

$$h_t = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^q \beta_i h_{t-i} = \alpha_0 + A(L)\varepsilon_t^2 + B(L)h_t \tag{3.5}$$

donde $p > 0$, $q \geq 0$, $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1..p$, $\beta_i \geq 0$ con $i = 1..q$.

Seguimos la línea de Engle (1982) asumiendo que la distribución condicionada es Normal, pero naturalmente también se pueden considerar otras distribuciones;

por ejemplo, en lugar de ε_{t-i}^2 en la ecuación (3.5), puede ser más apropiado en algunas aplicaciones utilizar $|\varepsilon_{t-i}|$.

Para $p=0$ el modelo se reduce a un ARCH(q), y para $p=q=0$, ε_t es sencillamente un ruido blanco.

En el proceso ARCH(q) la varianza condicionada se especifica como una función lineal únicamente de las varianzas muestrales pasadas, mientras que el proceso GARCH(p,q) permite introducir también varianzas condicionadas retardadas, lo que se corresponde con alguna clase de mecanismo de aprendizaje adaptativo.

El modelo de regresión GARCH(p,q) se obtiene tomando los ε_t como innovaciones de una regresión lineal:

$$\varepsilon_t = y_t - x_t' b$$

donde y_t es la variable dependiente, x_t' es un vector de variables explicativas y b es un vector de parámetros desconocidos.

Si todas las raíces de $1-B(z)=0$ están fuera del círculo unidad, (3.5) se puede escribir como:

$$\begin{aligned} h_t &= \alpha_0 (1 - B(1))^{-1} + A(L)(1 - B(L))^{-1} \varepsilon_t^2 \\ &= \alpha_0 \left(1 - \sum_{i=1}^q \beta_i \right)^{-1} + \sum_{i=1}^{\infty} \delta_i \varepsilon_{t-i}^2 \end{aligned}$$

que junto con (3.4) puede interpretarse como un ARCH(∞).

Los δ_i se obtienen a partir del desarrollo en serie de potencias de:

$$\begin{aligned} D(L) &= A(L)(1 - B(L))^{-1} \\ \delta_i &= \begin{cases} \alpha_i + \sum_{j=1}^n \beta_j \delta_{i-j}, & i = 1 \dots p \\ \sum_{j=1}^n \beta_j \delta_{i-j}, & i = p + 1, \dots \end{cases} \end{aligned}$$

con $n = \text{mín}(q, i - 1)$.

Entonces, si $B(1) < 1$, δ_i será decreciente para i más grande que $m = \text{máx}\{p, q\}$.

Así, si $D(1) < 1$, el proceso GARCH(p,q) se puede aproximar, con un grado determinado de precisión, por un ARCH(p) estacionario para un valor suficientemente grande de q . Pero como en el caso análogo de un ARMA, el proceso GARCH puede justificarse, posiblemente, a través de una descomposición de Wald de los argumentos, lo que proporcionaría una descripción más escueta.

A partir de la teoría elaborada sobre procesos ARCH(q) finito dimensionales que hemos visto en la sección anterior, es de esperar que la condición $D(1) < 1$, o equivalentemente $A(1) + B(1) < 1$, sea suficiente para garantizar la estacionariedad en sentido amplio (Milhoj 1985).

Una representación equivalente del proceso GARCH(p,q) viene dada por

$$\varepsilon_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^q \beta_i \varepsilon_{t-i}^2 - \sum_{i=1}^q \beta_i \nu_{t-i} + \nu_t \quad (3.6)$$

con

$$\nu_t = \varepsilon_t^2 - h_t = (\eta_t^2 - 1) h_t$$

donde $\eta_t \sim N(0, 1)$ i.i.d.

Obsérvese que, por definición, ν_t es serialmente incorrelada de media cero. Además, el proceso GARCH(p,q) se puede interpretar como un proceso ARMA en ε_t^2 de órdenes $m = \max(p, q)$ y q , respectivamente. Sin embargo, aunque (3.6) puede ser más significativa desde un punto de vista teórico, la formulación dada por (3.4) y (3.5) resulta más sencilla para trabajar de forma práctica.

Estructura de las funciones de autocorrelación y autocorrelación parcial

Puesto que es conocido el uso de las funciones de autocorrelación y autocorrelación parcial para identificar y chequear el comportamiento de series temporales de la forma ARMA (Box-Jenkins 1976), a continuación veremos como estas dos funciones del proceso al cuadrado son útiles para identificar y chequear el comportamiento de series temporales cuya varianza condicionada sigue un proceso de la forma GARCH. (La idea de utilizar el proceso al cuadrado para chequear la adecuación del modelo no es nueva; Granger y Anderson (1978) encontraron que algunas de las series modelizadas por Box-Jenkins presentaban residuos al cuadrado autocorrelados, incluso si los residuos mismos parecían no estar correlados a lo largo del tiempo).

Consideremos el proceso general GARCH(p,q):

$$\varepsilon_t/\psi_{t-1} \sim N(0, h_t)$$

$$h_t = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^q \beta_i h_{t-i} = \alpha_0 + A(L)\varepsilon_t^2 + B(L)h_t$$

y supongamos que el proceso tiene momento de cuarto orden finito.

Hay que tener en cuenta que las condiciones para la existencia del momento de cuarto orden son desconocidas; sin embargo, dado un orden específico del modelo estas condiciones se pueden obtener. Por ejemplo, la condición necesaria y suficiente para el proceso GARCH(1,2) obtenida de esta manera es:

$$\alpha_2 + 3\alpha_1^2 + 3\alpha_2^2 + \beta_1^2 + 2\alpha_1\beta_1 - 3\alpha_2^3 - 3\alpha_1^2\alpha_2 + 6\alpha_1\alpha_2\beta_1 + \alpha_2\beta_1^2 < 1$$

Sea la función de autocovarianza para ε_t^2 denotada por:

$$\gamma_n = \gamma_{-n} = Cov(\varepsilon_t^2, \varepsilon_{t-n}^2)$$

De (3.6) se sigue inmediatamente que:

$$\begin{aligned} \gamma_n &= \sum_{i=1}^p \alpha_i \gamma_{n-i} + \sum_{i=1}^q \beta_i \gamma_{n-i} \\ &= \sum_{i=1}^m \varphi_i \gamma_{n-i} \begin{cases} n \geq q+1 \\ m = \max(p, q) \\ \varphi_i = \alpha_i + \beta_i, i = 1 \dots p \end{cases} \end{aligned} \quad (3.7)$$

$\alpha_i \equiv 0$ para $i > p$ y $\beta_i \equiv 0$ para $i > q$.

De donde se obtienen las siguientes ecuaciones, análogas a las de Yule-Walker:

$$\rho_n = \gamma_n \gamma_0^{-1} = \sum_{i=1}^m \varphi_i \rho_{n-i}, n \geq q + 1 \tag{3.8}$$

Por tanto, las q primeras autocorrelaciones para ε_t^2 dependen directamente de los parámetros $\alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q$ pero, dados $\rho_q, \dots, \rho_{q+1-m}$ la anterior ecuación en diferencias únicamente determina las autocorrelaciones en retardos más altos. Esto es similar a lo que ocurre con las autocorrelaciones de un proceso ARMA(m,p).

Obsérvese también, que (3.7) depende de los parámetros $\alpha_1, \dots, \alpha_p$ y β_1, \dots, β_q solamente a través de $\varphi_1, \dots, \varphi_m$.

Sea ϕ_{kk} la autocorrelación parcial k para ε_t^2 que se obtiene resolviendo el conjunto de k ecuaciones con k incógnitas $\phi_{k1}, \dots, \phi_{kk}$:

$$\rho_n = \sum_{i=1}^k \phi_{ki} \rho_{n-i} \quad n = 1 \dots k$$

Por (3.8) ϕ_{kk} se corta después del retardo q para un proceso ARCH(p):

$$\phi_{kk} \begin{cases} \neq 0, k \leq p \\ = 0, k > p \end{cases}$$

Este comportamiento es idéntico al de la función de autocorrelación parcial de un AR(p).

También, por (3.8) y otros resultados ya conocidos en series temporales, la función de autocorrelación parcial de ε_t^2 para un proceso $GARCH(p, q)$ es, en general, no nula pero tiende a desaparecer.

En la práctica, naturalmente, los ρ_n y las ϕ_{kk} serán desconocidas. Si embargo, los análogos muestrales $\hat{\rho}_n$ proporcionan una estimación consistente para ρ_n , y ϕ_{kk} se estima consistentemente por el coeficiente k -ésimo, $\hat{\phi}_{kk}$, de una autorregresión de orden k para ε_t^2 (Granger y Newbold 1977). Estas estimaciones, junto con su varianza asintótica bajo la hipótesis nula de no GARCH, se pueden utilizar

en la etapa de identificación preliminar, resultando también útiles para chequear el diagnóstico del modelo.

3.1.4. Extensiones del modelo GARCH

En la mayoría de los casos, el modelo básico GARCH resulta una elección razonablemente adecuada para representar y estimar la volatilidad condicionada presente en las series financieras. Sin embargo, dado que en la formulación del modelo intervienen únicamente los residuos al cuadrado, ε_{t-i}^2 , la volatilidad condicionada no refleja la influencia del signo de los residuos. Sin embargo, es un hecho conocido en las series financieras que las “malas noticias” (shocks negativos) tienen una mayor influencia sobre la volatilidad que las “buenas noticias” (shocks positivos). Esto es debido a que las malas noticias tienden a bajar los precios, lo que produce un efecto de palanca y genera una mayor volatilidad. Por esto, nos referimos al fenómeno del impacto asimétrico de las noticias como *efecto de apalancamiento financiero*. A continuación, presentamos algunas variantes del modelo GARCH básico que incorporan el efecto de apalancamiento financiero en su formulación.

Modelo TGARCH

El modelo GARCH umbral (TGARCH) (Zakoian, 1991) es de la forma siguiente:

$$h_t = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \gamma_i S_{t-i} \varepsilon_{t-i}^2 + \sum_{i=1}^q \beta_i h_{t-i}$$

donde

$$S_{t-i} = \begin{cases} 1 & \text{si } \varepsilon_{t-i} > 0 \\ 0 & \text{si } \varepsilon_{t-i} \leq 0 \end{cases}$$

esto es, dependiendo de cuando ε_{t-i} está por encima o por debajo del umbral del 0, ε_{t-i}^2 tiene un efecto diferente sobre la varianza h_t : si ε_{t-i} es positivo, el efecto total viene dado por $\alpha_i \varepsilon_{t-i}^2$, mientras que si ε_{t-i} es negativo, el efecto total viene dado por $(\alpha_i + \gamma_i) \varepsilon_{t-i}^2$. Por tanto, valores positivos de γ_i producen mayor influencia de las llamadas malas noticias.

Este modelo también es conocido como modelo GJR, debido a que fue propuesto de manera muy similar por Glosten, Jagannathan y Runkle (1993).

Modelo EGARCH

Nelson (1991) propuso el siguiente GARCH exponencial:

$$\log h_t = \alpha_0 + \sum_{i=1}^p \alpha_i \frac{|\varepsilon_{t-i}| + \gamma_i \varepsilon_{t-i}}{h_{t-i}^{1/2}} + \sum_{i=1}^p \beta_i \log h_{t-i} \quad (3.9)$$

Obsérvese que cuando ε_{t-i} es positivo, “buenas noticias”, el efecto total es $(1 + \gamma_i) |\varepsilon_{t-i}|$, mientras que si ε_{t-i} es negativo o hay “malas noticias”, el efecto total es $(1 - \gamma_i) |\varepsilon_{t-i}|$. Las malas noticias pueden tener un gran impacto sobre la volatilidad, si el valor de γ_i es negativo. Una ventaja adicional del modelo EGARCH sobre el GARCH básico es que la varianza condicionada h_t siempre es positiva sin importar los valores de los coeficientes de (3.9), ya que el modelo se refiere a $\log h_t$ en lugar de h_t .

Modelo PGARCH

El modelo GARCH básico se puede tratar como un caso particular del modelo GARCH potencia (PGARCH) propuesto por Ding, Granger y Engle (1993):

$$h_t^{d/2} = \alpha_0 + \sum_{i=1}^p \alpha_i (|\varepsilon_{t-i}| + \gamma_i \varepsilon_{t-i})^d + \sum_{i=1}^q \beta_i h_{t-i}^{d/2} \quad (3.10)$$

donde d es un exponente positivo. Si $d = 2$, (3.10) se reduce al modelo GARCH básico con efecto apalancamiento. Ding, Granger y Engle (1993) demuestran que el modelo PGARCH también incluye casos especiales de muchas otras variantes GARCH.

Curvas de impacto de noticias

Todos los modelos planteados en los puntos anteriores son capaces de recoger el efecto de apalancamiento financiero, por lo que la elección particular de uno de ellos implica el uso de un criterio de selección. Podemos emplear de nuevo el criterio de Akaike (definición 2.3), junto con las llamadas *curvas de impacto de noticias*, propuestas por Engle y Ng (1993), que representan gráficamente la relación funcional entre la varianza condicionada en el instante t y el término de error en el instante $(t - 1)$, manteniendo constante la información en el instante $(t - 2)$ y anteriores.

La tabla 3.1 recoge las expresiones de la varianza no condicionada, $\bar{\sigma}^2$, de diferentes modelos GARCH, y la tabla 3.2 contiene las expresiones de las curvas de impacto de noticias para los mismos modelos con órdenes $p = 1$ y $q = 1$.

$$\begin{array}{l}
 \text{GARCH}(p,q) \\
 \text{TGARCH}(p,q) \\
 \text{PGARCH}(p,q) \\
 \text{EGARCH}(p,q)
 \end{array}
 \left|
 \begin{array}{l}
 \bar{\sigma}^2 = \alpha_0 \left/ \left[1 - \sum_{i=1}^p \alpha_i (1 + \gamma_i^2) - \sum_{i=1}^q \beta_i \right] \right. \\
 \bar{\sigma}^2 = \alpha_0 \left/ \left[1 - \sum_{i=1}^p (\alpha_i + \gamma_i / 2) - \sum_{i=1}^q \beta_i \right] \right. \\
 \bar{\sigma}^2 = \alpha_0^2 \left/ \left[1 - \sum_{i=1}^p \alpha_i \sqrt{2/\pi} - \sum_{i=1}^q \beta_i \right]^2 \right. \\
 \bar{\sigma}^2 = \exp \left\{ \left(\alpha_0 + \sum_{i=1}^p \alpha_i \sqrt{2/\pi} \right) \left/ \left(1 - \sum_{i=1}^q \beta_i \right) \right. \right\}
 \end{array}
 \right.$$

Tabla 3.1: Varianza no condicionada de los diferentes procesos GARCH

GARCH(1,1)	$\sigma_t^2 = A + \alpha_1 (\epsilon_{t-1} + \gamma_1 \epsilon_{t-1})^2$ $A = \alpha_0 + \beta_1 \bar{\sigma}^2$
TGARCH(1,1)	$\sigma_t^2 = A + (\alpha_1 + \gamma_1 S_{t-1}) \epsilon_{t-1}^2$ $A = \alpha_0 + \beta_1 \bar{\sigma}^2$
PGARCH(1,1)	$\sigma_t^2 = A + 2\sqrt{A}\alpha_1 (\epsilon_{t-1} + \gamma_1 \epsilon_{t-1}) + \alpha_1^2 (\epsilon_{t-1} + \gamma_1 \epsilon_{t-1})^2$ $A = (\alpha_0 + \beta_1 \bar{\sigma})^2$
EGARCH(1,1)	$\sigma_t^2 = A \exp \{ \alpha_1 (\epsilon_{t-1} + \gamma_1 \epsilon_{t-1}) / \bar{\sigma} \}$ $A = \bar{\sigma}^{2\beta_1} \exp \{ \alpha_0 \}$

Tabla 3.2: Curvas de impacto de noticias de los diferentes procesos GARCH

3.2. Ajuste de procesos GARCH a los retornos diarios DowJones e Ibex35

3.2.1. Selección y diagnóstico de los modelos

En el capítulo anterior, se había constatado la no normalidad de los retornos mediante el estadístico de Jarque-Bera y la observación de los gráficos correspondientes a las funciones de autocorrelación de los retornos al cuadrado de los dos índices considerados (figura 3.1), hacía sospechar la presencia de efecto ARCH (heterocedasticidad condicionada) en ambas series.

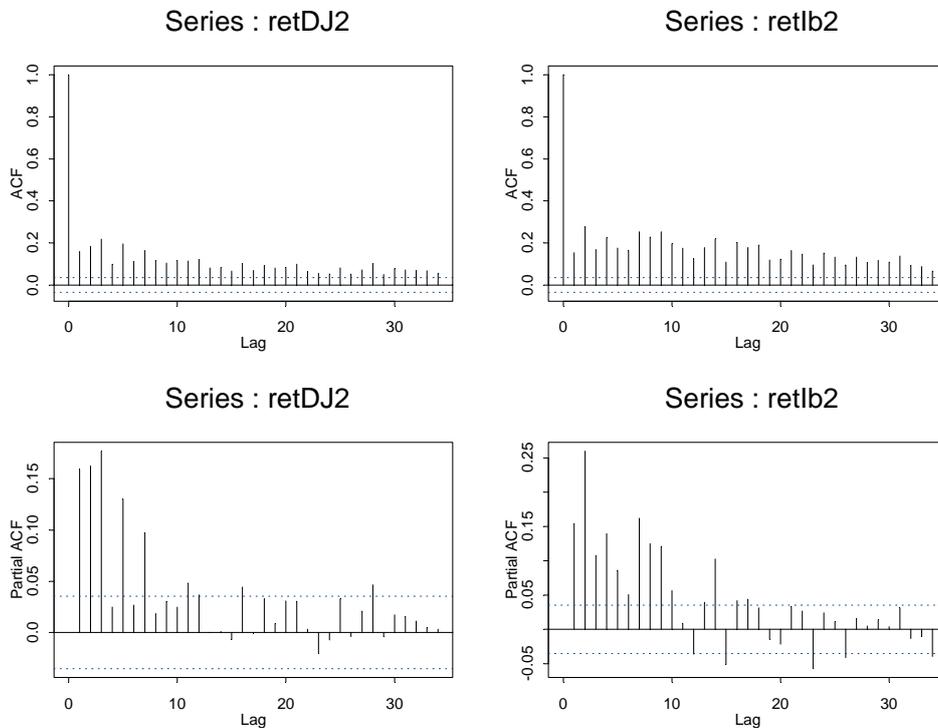


Figura 3.1: Funciones de autocorrelación simple y parcial de los retornos logarítmicos diarios al cuadrado DowJones e Ibex35

A partir de la formulación del modelo ARCH(p) dada por (3.3) y, puesto que ϵ_t tiene media cero $\forall t$, suponiendo que existe el momento de segundo orden de

X_t

$$\text{Var}(\epsilon_t) = E(\epsilon_t^2) = h_t$$

donde Var y E denotan, respectivamente, la varianza y la esperanza condicionadas por la información hasta el instante $(t - 1)$. Entonces, (3.3) puede escribirse como

$$\epsilon_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \dots + \alpha_p \epsilon_{t-p}^2 + U_t \quad (3.11)$$

donde $U_t = \epsilon_t^2 - E(\epsilon_t^2)$ es un ruido blanco de media cero. La ecuación (3.11) permite afirmar que los residuos al cuadrado de un proceso ARCH(p) siguen un proceso autorregresivo de orden p . Engle (1982), propone un contraste para detectar el efecto ARCH basado en la ecuación (3.11):

Se contrasta la hipótesis nula $\alpha_1 = \alpha_2 = \dots = \alpha_p = 0$ mediante el estadístico $LM = TR^2$. Bajo la hipótesis nula, LM sigue una distribución χ^2 con p grados de libertad. T es el tamaño muestral y R^2 es el coeficiente de determinación de la regresión de ϵ_t^2 sobre p valores retardados de ϵ_t^2 , obtenido a partir de (3.11) después de ajustar el modelo correspondiente según (3.3).

Los resultados del contraste anterior aplicado a las dos series de retornos aparecen en la tabla 3.3. La hipótesis nula de ausencia de efecto ARCH es rechazada en ambos casos.

Retornos	LM	P-valor
DowJones	340.08	0.0
Ibex35	547.89	0.0

Tabla 3.3: Contraste LM de la presencia de efecto ARCH en las series de retornos logarítmicos diarios DowJones e Ibex35

Formalmente, consideraremos $(Z_t)_{t \in \mathbb{Z}}$ sucesión de variables aleatorias iid de media cero y varianza unidad. El proceso $(X_t)_{t \in \mathbb{Z}}$ es un ARMA(p_2, q_2) para la media condicionada y un GARCH(p_2, q_2) para la varianza condicionada si

satisface

$$\begin{aligned}
 X_t &= \mu_t + \epsilon_t \\
 \mu_t &= \mu + \sum_{i=1}^{p_1} \phi_i (X_{t-1} - \mu) + \sum_{j=1}^{q_1} \theta_j \epsilon_{t-j} \\
 \epsilon_t &= \sigma_t Z_t \\
 \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{p_2} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q_2} \beta_j \sigma_{t-j}^2
 \end{aligned} \tag{3.12}$$

donde $\alpha_0 > 0$, $\alpha_i \geq 0$ para $i = 1, 2, \dots, p_2$, $\beta_j \geq 0$ para $j = 1, 2, \dots, q_2$ y Z_t es independiente de $(X_s)_{s \leq t}$. Los polinomios $\phi(z) = 1 - \phi_1 z - \dots - \phi_{p_1} z^{p_1}$ y $\theta(z) = 1 - \theta_1 z - \dots - \theta_{q_1} z^{q_1}$ no tienen raíces comunes ni dentro del círculo unidad.

A continuación, ajustamos modelos univariantes ARMA(1, 1) para la media condicionada y del tipo GARCH(1, 1) para la varianza condicionada, mediante máxima verosimilitud, a cada una de las series de retornos marginales. Supondremos que las innovaciones Z_t siguen una distribución gaussiana o una distribución t de Student, y compararemos los resultados. Los diferentes procesos tipo GARCH(1, 1) que consideramos incorporan, en todos los casos, un término que contempla la presencia del efecto de apalancamiento financiero. En concreto utilizamos los procesos TGARCH(1, 1), EGARCH(1, 1) y PGARCH(1, 1). En este último caso, la estimación de máxima verosimilitud de la potencia produjo un valor de 0.855, por lo que el proceso se reajustó con un valor fijo para la potencia de 1.

Los resultados obtenidos, según los distintos procesos considerados, para la serie de retornos logarítmicos diarios DowJones con innovaciones gaussianas son los siguientes:

$\mu - \text{TGARCH}(1, 1)$	Coefficientes	Error std.	t	$\Pr(> t)$
μ	3.325e-004	1.52e-004	2.180	1.46e-002
α_0	1.392e-006	2.13e-007	6.538	3.63e-011
$ARCH(1)$	1.320e-002	7.90e-003	1.670	4.75e-002
$GARCH(1)$	9.188e-001	7.03e-003	130.608	0.0
Coef. apalanc.	1.082e-001	9.69e-003	11.167	0.0

ARMA(1, 1) – EGARCH(1, 1)	Coeficientes	Error std.	t	$\Pr(> t)$
μ	0.00041	0.0002	1.40	7.97e-002
$AR(1)$	-0.9065	0.1585	-5.71	5.93e-009
$MA(1)$	0.8982	0.1667	5.38	3.82e-008
α_0	-0.4649	0.0458	-10.13	0.0
$ARCH(1)$	0.1571	0.0152	10.27	0.0
$GARCH(1)$	0.9633	0.0041	231.19	0.0
Coef. apalanc.	-0.6990	0.0895	-7.80	3.99e-015

μ – PGARCH(1, 1, 1)	Coeficientes	Error std.	t	$\Pr(> t)$
μ	0.0002	0.0001	1.71	4.29e-002
α_0	0.0001	2e-005	7.27	2.21e-013
$ARCH(1)$	0.0687	0.0067	10.19	0.0
$GARCH(1)$	0.9288	0.0064	143.13	0.0
Coef. apalanc.	-0.7531	0.09477	-7.94	1.33e-015

En todos los casos se utiliza el estadístico t usual, $\hat{\alpha}/\hat{\sigma}_{\hat{\alpha}}$, para contrastar cuando el parámetro correspondiente es cero.

Para comprobar la no autocorrelación de los residuos utilizamos el contraste de Ljung-Box, que plantea la hipótesis nula de aleatoriedad y utiliza el estadístico

$$Q_{LB} = n(n+2) \sum_{j=1}^h \frac{\rho_j^2}{n-j}$$

donde n es el tamaño muestral, ρ_j es la autocorrelación en el retardo j y h es el número de retardos considerado. La hipótesis nula se rechaza si $Q_{LB} > \chi_{1-\alpha; h}^2$, con α el nivel de significación fijado previamente. Los resultados de aplicar el contraste anterior tanto a los residuos estandarizados como a su cuadrado en los tres modelos ajustados no permiten rechazar la hipótesis nula de no autocorrelación (con 12 retardos considerados). También se efectuado el contraste basado en el estadístico LM a los residuos para constatar la ausencia de efecto ARCH en los mismos. Los resultados de ambos contrastes se recogen en las tablas 3.4 y 3.5.

	$Q_{LB} \text{ res}_{std}$	p-valor	$Q_{LB} \text{ res}_{std}^2$	p-valor
$\mu - PGARCH(1, 1, 1)$	15.89	0.19	7.49	0.82
$ARMA(1, 1) - EGARCH(1, 1)$	13.93	0.30	10.03	0.61
$\mu - TGARCH(1, 1)$	15.95	0.19	8.83	0.71

Tabla 3.4: *Contraste Q_{LB} aplicado a los residuos DowJones en los tres procesos GARCH ajustados*

	LM	p-valor
$\mu - PGARCH(1, 1, 1)$	7.38	0.87
$ARMA(1, 1) - EGARCH(1, 1)$	9.784	0.63
$\mu - TGARCH(1, 1)$	8.69	0.75

Tabla 3.5: *Contraste LM aplicado a los residuos DowJones en los tres procesos GARCH ajustados*

Para terminar, quedaría por analizar la normalidad de los residuos. Efectuado el contraste de Jarque-Bera sobre los residuos de los tres procesos ajustados, se obtienen p-valores nulos, con lo cual es posible rechazar la hipótesis de normalidad en los tres supuestos. La figura 3.2 muestra los QQ-plots de los residuos de los tres modelos ajustados frente a la normal tipificada. Se aprecia de forma evidente, que la cola izquierda en los tres casos se aleja bastante del supuesto de normalidad.

Una vez realizado el diagnóstico de los procesos ajustados, pasamos a decidir cual es el más adecuado para representar el comportamiento marginal de la serie de retornos logarítmicos diarios DowJones. La tabla 3.6 muestra la ordenación, según el criterio de Akaike, de los tres modelos:

	AIC	Valor verosimilitud
$\mu - PGARCH(1, 1, 1)$	-20166.67	10088
$ARMA(1, 1) - EGARCH(1, 1)$	-20140.11	10073
$\mu - TGARCH(1, 1)$	-20136.64	10077

Tabla 3.6: *Ordenación según AIC de los procesos ajustados a los retornos diarios DowJones con innovaciones gaussianas*

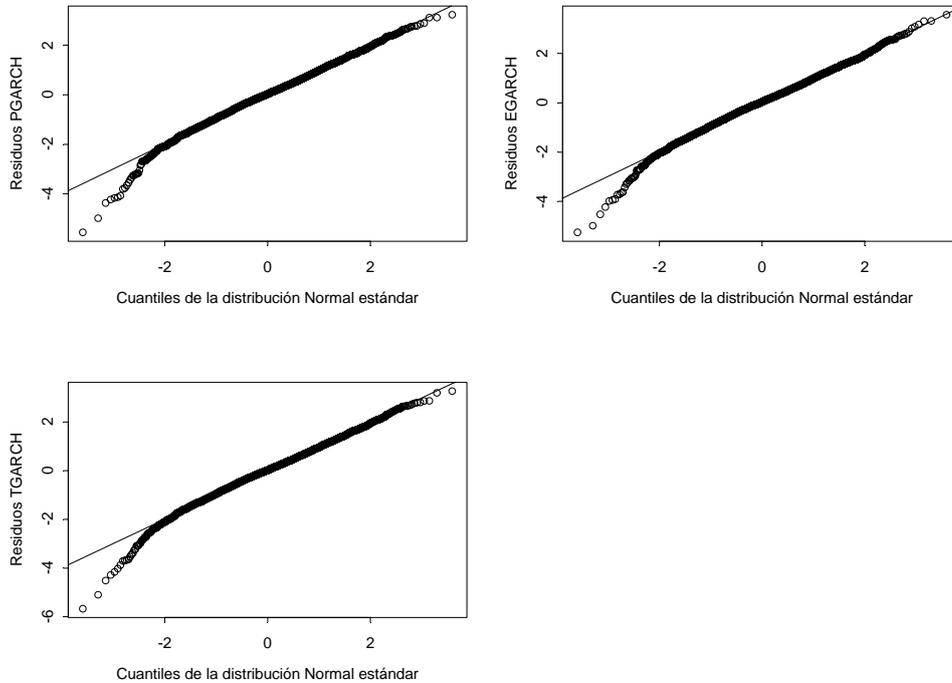


Figura 3.2: *QQ-plots de los residuos estandarizados DowJones frente a la normal estándar*

En la figura 3.3 están representadas las curvas de impacto de noticias generadas por los tres procesos ajustados a la serie de retornos logarítmicos diarios DowJones con innovaciones gaussianas. La inspección de esta representación gráfica permite constatar la presencia del efecto de apalancamiento financiero en todos ellos (ya confirmado por el hecho de que el coeficiente correspondiente al mismo es significativo en todos los ajustes) que se traduce en la contribución asimétrica que tienen las innovaciones sobre la volatilidad. Podemos observar como valores negativos de los residuos, malas noticias, disparan el valor de la volatilidad condicionada. También se observa que el modelo más robusto frente al impacto de valores negativos grandes es el correspondiente al PGARCH(1,1,1) ajustado. Esta información, junto con la proporcionada por la tabla 3.6, hace que el proceso PGARCH(1,1,1), con media constante, sea el preferido para representar el comportamiento marginal de la serie de retornos DowJones.

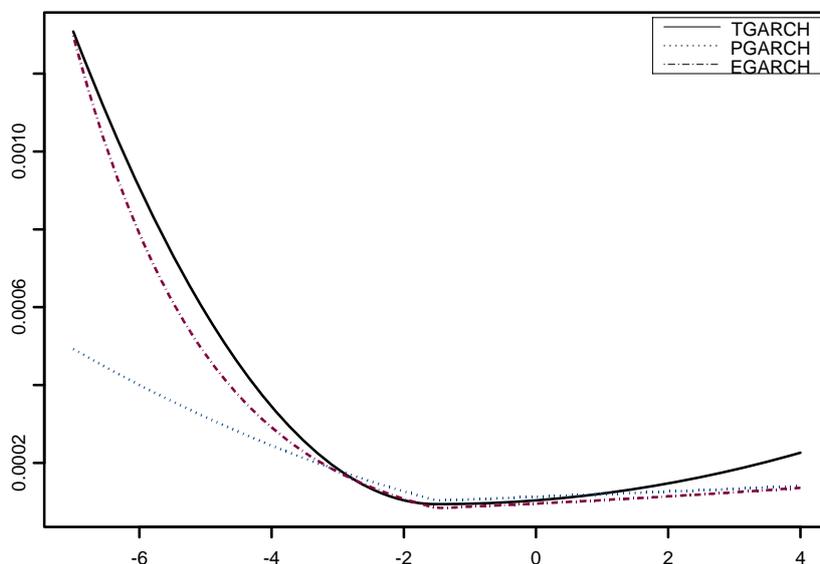


Figura 3.3: *Curvas de impacto de noticias de los tres procesos GARCH ajustados a la serie de retornos logarítmicos diarios DowJones con innovaciones gaussianas*

Como ya habíamos comentado, la hipótesis de normalidad no resulta apropiada para los residuos de nuestro modelo. Por ello, una vez seleccionado el proceso PGARCH(1,1,1), con media constante, como el más adecuado a nuestros propósitos, podemos efectuar el reajuste del mismo, suponiendo que las innovaciones siguen una distribución t de Student. Los coeficientes estimados para el proceso PGARCH(1,1,1) con media constante e innovaciones distribuidas según una t de Student, con una estimación de 10.863 grados de libertad y desviación típica estimada 1.6755, son los siguientes:

$\mu - \text{PGARCH}(1, 1, 1)$	Coeficientes	Error std.	t	$\text{Pr}(> t)$
μ	4.09e-004	1.45e-004	2.807	2.51e-003
α_0	1.44e-004	0.29e-004	4.926	4.41e-007
$ARCH(1)$	0.0617	8.74e-003	7.063	9.99e-013
$GARCH(1)$	0.9366	8.15e-003	114.857	0.0
Coef. apalanc.	-0.7738	0.1336	-5.79	3.86e-009

El contraste de Jarque -Bera permite rechazar la hipótesis de normalidad en los residuos y, en cuanto a la autocorrelación, los valores que toma el estadístico Q_{LB} de Ljung-Box son 16.11 (p-valor de 0.186) para los residuos y 8.41 (p-valor de 0.752) para los residuos al cuadrado, por lo que no es posible rechazar la hipótesis nula de aleatoriedad en los 12 retardos considerados. El estadístico LM toma un valor de 8.28 con un p-valor de 0.796, por lo que aceptamos la ausencia de efecto ARCH en los residuos. El valor AIC para este modelo es de -20221.34, inferior al -20166.67 obtenido para el modelo con innovaciones gaussianas, y con un valor para la verosimilitud de 10117.

Si realizamos un análisis análogo para la serie de retornos logarítmicos diarios del índice Ibex35, obtenemos los siguientes coeficientes estimados en los diversos procesos ajustados:

$\mu - \mathbf{TGARCH}(1, 1)$	Coefficientes	Error std.	t	$\Pr(> t)$
μ	5.96e-004	1.86e-004	3.21	6.70e-004
α_0	2.76e-006	4.22e-007	6.53	3.71e-011
$ARCH(1)$	5.06e-002	9.75e-003	5.19	1.10e-007
$GARCH(1)$	8.90e-001	9.75e-003	92.91	0.0
Coef. apalanc.	8.59e-002	1.20e-002	7.12	6.37e-013

$\mu - \mathbf{EGARCH}(1, 1)$	Coefficientes	Error std.	t	$\Pr(> t)$
μ	4.91e-004	1.77e-004	2.78	2.73e-003
α_0	-0.4486	0.0464	-9.66	0.0
$ARCH(1)$	0.1944	0.0167	11.62	0.0
$GARCH(1)$	0.9666	0.0042	225.70	0.0
Coef. apalanc.	-0.3371	0.0586	-5.74	5.08e-009

$\mu - \mathbf{PGARCH}(1, 1, 1)$	Coefficientes	Error std.	t	$\Pr(> t)$
μ	5.60e-004	1.67e-004	3.35	4.006e-004
α_0	2.26e-004	0.32e-004	6.87	3.70e-012
$ARCH(1)$	0.0895	6.98e-003	12.81	0.0
$GARCH(1)$	0.9106	7.03e-003	129.42	0.0
Coef. apalanc.	-0.3611	0.0562	-6.41	8.12e-011

El análisis de la autocorrelación de los residuos mediante el contraste de Ljung-Box proporciona los resultados que aparecen en la tabla 3.7. Podemos observar que mientras que la ausencia de correlación es aceptable al analizar los residuos al cuadrado, en el caso de la serie simple de residuos no es posible aceptar la

hipótesis de aleatoriedad en los casos PGARCH y TGARCH.

	$Q_{LB} res_{std}$	p-valor	$Q_{LB} res_{std}^2$	p-valor
$\mu - PGARCH(1, 1, 1)$	22.12	0.03	7.91	0.791
$\mu - EGARCH(1, 1)$	20.35	0.06	7.38	0.830
$\mu - TGARCH(1, 1)$	21.94	0.03	8.12	0.775

Tabla 3.7: *Contraste Q_{LB} aplicado a los residuos Ibe35 en los tres procesos GARCH ajustados*

En la tabla 3.8 se recogen los valores del estadístico LM , que permiten confirmar la ausencia de efecto ARCH en las tres series de residuos Ibe35.

	LM	p-valor
$\mu - PGARCH(1, 1, 1)$	7.78	0.80
$\mu - EGARCH(1, 1)$	7.31	0.83
$\mu - TGARCH(1, 1)$	7.96	0.78

Tabla 3.8: *Contraste LM aplicado a los residuos Ibe35 en los tres procesos GARCH ajustados*

De nuevo recurrimos al criterio de Akaike y a las curvas de impacto de noticias, como herramientas de ayuda en la selección del mejor modelo para filtrar la dependencia temporal de la serie de retornos Ibe35. La ordenación según AIC de los tres procesos ajustados aparece en la tabla 3.9.

	AIC	Valor verosimilitud
$\mu - PGARCH(1, 1, 1)$	-18754.89	9382
$\mu - TGARCH(1, 1)$	-18748.77	9379
$\mu - EGARCH(1, 1)$	-18745.43	9378

Tabla 3.9: *Ordenación según AIC de los procesos ajustados a los retornos diarios Ibe35 con innovaciones gaussianas*

La figura 3.4, en la que están representadas las curvas de impacto de noticias asociadas a cada proceso ajustado, permite apreciar un comportamiento intermedio del proceso EGARCH, que recordemos era el que proporcionaba mejores resultados en términos de autocorrelación de los residuos estandarizados, en cuanto a estabilidad frente a la aparición de malas noticias (valores negativos de los errores).

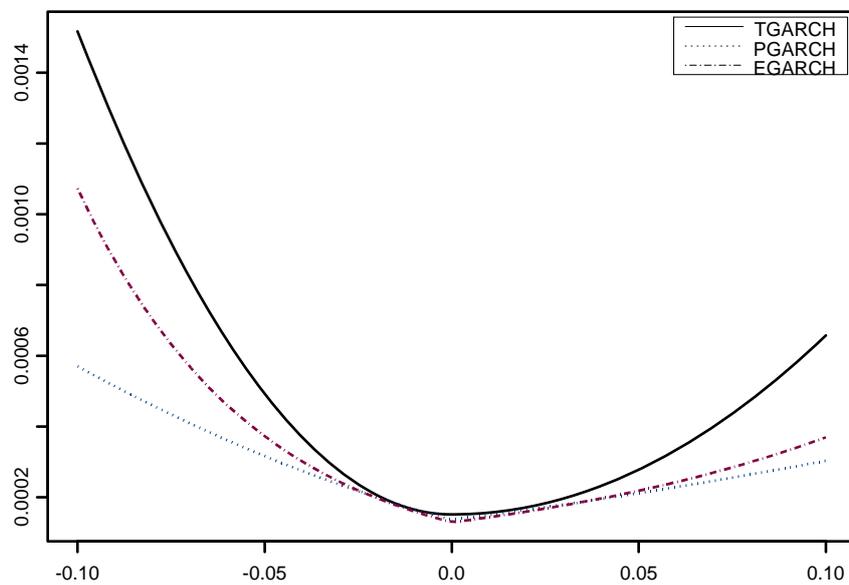


Figura 3.4: *Curvas de impacto de noticias de los tres procesos GARCH ajustados a la serie de retornos logarítmicos diarios Ibex35 con innovaciones gaussianas*

En lo que respecta a la normalidad de los residuos, el contraste de Jarque-Bera proporciona p-valores nulos en los tres casos, por lo que es posible rechazar la hipótesis de residuos gaussianos en todos los ajustes realizados. En la figura 3.5 aparecen los QQ-plots de los residuos Ibex35 frente la normal estándar. En las tres series de residuos se observa la presencia de colas pesadas, es decir, ambas colas reflejan frecuencias más altas que las esperadas en el caso gaussiano, lo que induce a pensar que el ajuste mejoraría considerando innovaciones con distribución *t* de Student.

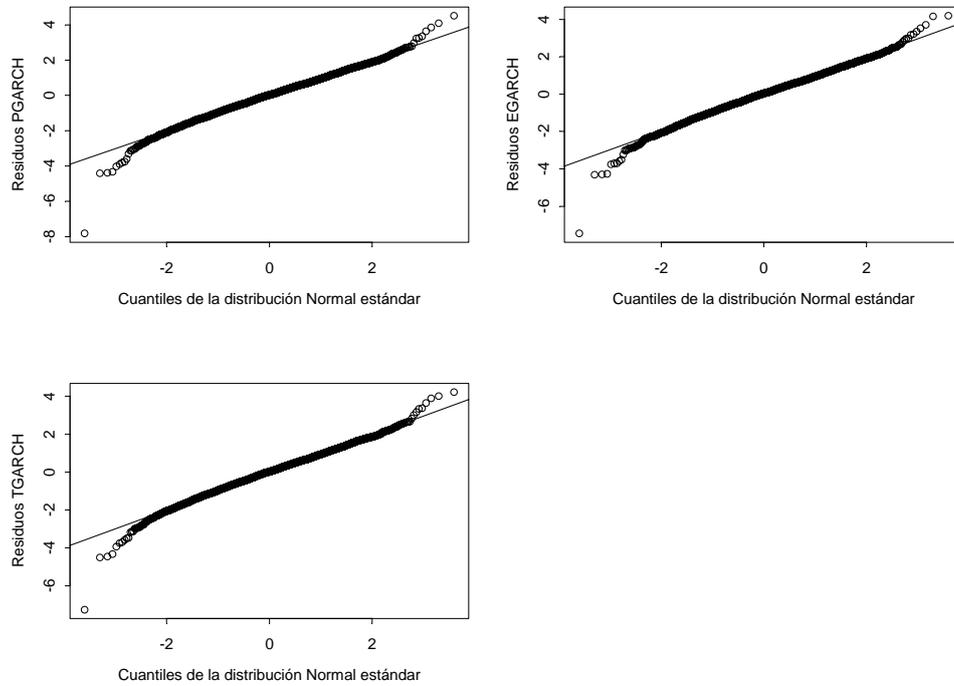


Figura 3.5: *QQ-plots de los residuos estandarizados Ibox35 frente a la normal estándar*

Realizado este ajuste, los coeficientes estimados para el proceso EGARCH(1,1) con media constante e innovaciones distribuidas según una t de Student, con una estimación de 10.617 grados de libertad y desviación típica estimada 1.4158, son los siguientes:

$\mu - \mathbf{EGARCH}(1, 1)$	Coefficientes	Error std.	t	$\Pr(> t)$
μ	7.02e-004	1.7e-004	3.93	4.29e-005
α_0	-0.3203	0.0450	-7.10	7.27e-013
$ARCH(1)$	0.1649	0.0193	8.53	0.0
$GARCH(1)$	0.9785	0.0041	236.15	0.0
Coef. apalanc.	-0.4383	0.0819	-5.34	4.73e-008

El contraste de Jarque -Bera permite rechazar la hipótesis de normalidad en los residuos y, en cuanto a la autocorrelación, los valores que toma el estadístico Q_{LB} de Ljung-Box son 20.88 (p-valor de 0.053) para los residuos estandarizados y 7.69 (p-valor de 0.8083) para los residuos estandarizados al cuadrado, por lo que no es posible rechazar la hipótesis nula de aleatoriedad en los 12 retardos considerados. El estadístico LM toma un valor de 7.55 con un p-valor de 0.819, por lo que aceptamos la ausencia de efecto ARCH en los residuos. El valor AIC para este modelo es de -18824.18, inferior a cualquiera de los obtenidos para los modelos ajustados con innovaciones gaussianas, y con un valor para la verosimilitud de 9418.

3.2.2. Ajuste de cópulas bidimensionales a los retornos filtrados DowJones e Ibex35

A partir de los modelos GARCH ajustados en el punto anterior, se obtienen los residuos o *retornos filtrados*, \hat{z}_t , de cada una de las series de retornos logarítmicos diarios univariantes. Los residuos se calculan a partir de la muestra (x_1, \dots, x_T) y de los parámetros estimados, como

$$\hat{z}_t = \frac{(x_t - \hat{\mu}_t)}{\hat{\sigma}_t} \quad (3.13)$$

para $t = 1, 2, \dots, T$.

De esta forma, se obtiene una muestra bivalente de retornos filtrados que no presentan evidencias en contra de la hipótesis de i.i.d. en las marginales, pero que conservan la dependencia contemporánea, como se aprecia en la figura 3.6 en la que aparece el diagrama de dispersión correspondiente a la muestra bivalente de retornos DowJones e Ibex35, filtrados a través de un proceso PGARCH(1,1,1) con media constante para los retornos DowJones y de un EGARCH(1,1) con media constante para los retornos Ibex35, en ambos casos con innovaciones según una t de Student. En la parte derecha de este mismo gráfico aparece el soporte de la cópula empírica de las innovaciones, en el que es posible observar la dependencia en las colas de la distribución bivalente

Sean Z_1 y Z_2 las variables aleatorias que representan los retornos filtrados DowJones e Ibex35, respectivamente. Sea F la función de distribución de probabilidad bivalente de (Z_1, Z_2) con funciones de distribución marginales continuas F_1 y F_2 . Con el fin de analizar la dependencia entre los residuos, ajustamos

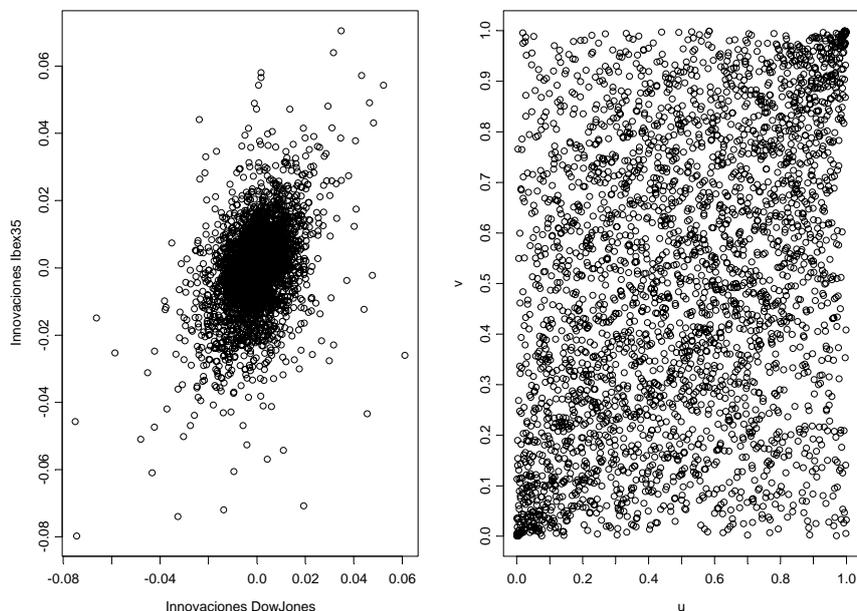


Figura 3.6: Diagrama de dispersión (izquierda) y soporte de la cópula empírica (derecha) de los retornos filtrados DowJones e Ibx35

modelos de cópulas del tipo

$$F(z_1, z_2) = C_\theta(F_1(z_1), F_2(z_2))$$

Dada la muestra de retornos filtrados $(\hat{\mathbf{z}}_1, \hat{\mathbf{z}}_2)$, estimamos el parámetro θ de la cópula C mediante los procedimientos de Genest y Rivest y máxima verosimilitud canónica (CML), ya descritos y utilizados en el Capítulo 2. En el caso CML, el estimador vendrá dado en esta ocasión por

$$\hat{\theta} = \arg \max_{\theta \in \Theta} \sum_{i=1}^n \log \mathcal{L}(\theta; F_{1emp}(z_{1i}), F_{2emp}(z_{2i}))$$

Los resultados obtenidos en el ajuste del modelo de cópula más adecuado para representar la relación de dependencia entre los retornos filtrados DowJones e Ibx35 aparecen en la tabla 3.10 ordenados según AIC. Las conclusiones son

semejantes a las obtenidas en el Capítulo 2 al tratar las series brutas de retornos, puesto que la estructura de dependencia contemporánea entre los dos índices se ha conservado a través del filtrado marginal. Los dos modelos mixtos encabezan la clasificación, seguidos del perteneciente a la familia de Gumbel, todos ellos con estimaciones de los parámetros muy similares a las obtenidas para las series brutas. En el caso de los modelos mixtos, aparecen tres valores estimados: los dos primeros corresponden a cada modelo de los que integran la mixtura y el tercero concreta la combinación lineal convexa.

Familia de cópulas	$\hat{\theta}$	AIC	λ_L	λ_U
<i>Modelo 1</i>	1.252 ; 1.253 ; 0.498	-652.03	0.1312	0.1297
<i>Modelo 2</i>	1.172 ; 1.022 ; 0.625	-624.72	0.1903	0.1209
<i>Gumbel</i>	1.3355	-561.58	0	0.3195
<i>Gaussiana</i>	0.3926	-512.23	0	0
<i>Clayton</i>	0.5423	-465.63	0.2785	0
<i>Frank</i>	2.4667	-452.04	0	0

Tabla 3.10: Ordenación según AIC de todos los modelos de cópula estimados para los retornos filtrados DowJones e Ibex35

Dependencia en Colas

Es un hecho conocido y ya comentado que la ausencia de normalidad en términos univariantes va asociada a los fenómenos de colas pesadas y leptocurtosis. En un planteamiento multivariante, estos rasgos característicos pueden referirse tanto a las distribuciones marginales como a la distribución que representa los movimientos conjuntos de los diferentes mercados. Anteriormente ya nos hemos referido a este concepto como *dependencia en colas*. Intuitivamente, podemos imaginar mercados en los que las distribuciones marginales presenten colas pesadas, mientras que los movimientos extremos de los mercados implicados sean ortogonales, o casos en los que los retornos de cada mercado se distribuyan normalmente pero los grandes movimientos se presentan conjuntamente bajo otro patrón. El uso de funciones cópula permite representar estas dos peculiaridades, colas pesadas y dependencia en colas, de forma separada. Recordemos que para representar la dependencia en colas considerábamos la probabilidad de que un suceso con probabilidad menor o igual que α ocurra en la primera variable, dado que un suceso con probabilidad menor o igual que α ha ocurrido en la segunda (podríamos estar hablando, por ejemplo, de la probabilidad de observar un *crash* de mercado con probabilidad menor o igual que 0,01 en el índice Ibex35, dado que ha ocurrido un *crash* con probabilidad menor o igual que 0,01 en el índice Dow Jones):

$$\lambda_\alpha = P(F_1(x) \leq \alpha / F_2(y) \leq \alpha) = \frac{P(F_1(x) \leq \alpha, F_2(y) \leq \alpha)}{P(F_2(y) \leq \alpha)} = \frac{C(\alpha, \alpha)}{\alpha}$$

Si calculamos esta medida a lo largo de la cola inferior, es decir, para valores muy pequeños de α , obtenemos el índice de dependencia en la cola inferior

definido en el punto 1.8 del Capítulo 1. Una interpretación análoga es posible para el índice de dependencia en la cola superior, que representa la probabilidad de que los precios se disparen simultáneamente en los dos mercados (en nuestro ejemplo, en los mercados español y estadounidense).

El cálculo de estos índices constituye una primera aproximación en la cuantificación de la dependencia en colas de una muestra bivalente, a través de la cópula que liga las funciones de distribución marginales correspondientes. Parece que una solución más ajustada a la naturaleza del problema, debe establecerse en términos de las colas de las distribuciones univariantes marginales, dado que nuestro objetivo es el cálculo de cuantiles altos que permitan determinar el valor de las distintas medidas de riesgo financiero. Es en este punto donde intervienen métodos estadísticos tradicionales de tratamiento de valores extremos, que combinados con la aportación que hacen las funciones cópula en la representación de estructuras de dependencia conjunta, permitirán enfocar el problema en la aparición contemporánea de sucesos extremos a través de sus distribuciones de probabilidad. En el Capítulo 3 hemos salvado la dificultad que suponía, a la hora de aplicar estos métodos, la presencia de dependencia temporal en las series marginales de retornos, mediante el filtrado de las mismas a través de procesos tipo GARCH. Nos disponemos a lo largo del presente capítulo, a resolver las siguientes cuestiones:

1. Comprobación mediante métodos gráficos (Chi-plots, K-plots) de la existencia de dependencia local en colas entre las series de retornos filtrados.
2. Modelización de las colas de las distribuciones marginales de las series de retornos filtrados previamente a través de procesos tipo GARCH. Para ello utilizaremos *Métodos Umbral* (POT) que permitirán ajustar *Distribuciones de Pareto Generalizadas* (GPD) a las muestras univariantes de residuos que superen un umbral dado. (Gilli y Këllezi 2005, Tajvidi 2004)
3. Cálculo de las medidas de riesgo financiero usuales VaR (Valor en Riesgo) y ES (Pérdida esperada) para cada índice considerado, a través de las GPD ajustadas y para diferentes niveles de confianza. Los resultados obtenidos se compararán con los procedentes de un supuesto gaussiano.
4. Selección del modelo de cópula más efectivo para capturar la dependencia conjunta de las distribuciones GPD que representan los comportamientos de los índices financieros en las colas marginales.

5. Cálculo de las medidas de riesgo financiero usuales VaR (Valor en Riesgo) y ES (Pérdida esperada), para diferentes niveles de confianza, pero en este caso, a través del modelo de cópula elegido, al objeto de cuantificar el riesgo financiero conjunto.

4.1. Métodos gráficos para detectar la dependencia bivalente: Chi-plot y K-plot

4.1.1. Chi-plot

El Chi-plot definido por Fisher y Switzer (1985, 2001) está diseñado para ser un instrumento gráfico sencillo, rápido y útil en la tarea de explorar estructuras de dependencia bivalente, dependiendo de los datos únicamente a través de los valores de sus rangos. La información que proporciona complementa la obtenida de un diagrama de dispersión, en el cual puede resultar difícil valorar la presencia de algún patrón diferente a una relación monótona. Su construcción es la siguiente:

Sea $(X_1, Y_1), \dots, (X_n, Y_n)$ una muestra aleatoria simple de H , función de distribución conjunta continua de la v.a. (X, Y) y sea $I(A)$ la función indicatriz usual del suceso A . Para cada observación (x_i, y_i) , sean:

$$H_i = \sum_{j \neq i} I(X_j \leq X_i, Y_j \leq Y_i) / (n - 1) \quad (4.1)$$

$$F_i = \sum_{j \neq i} I(X_j \leq X_i) / (n - 1)$$

$$G_i = \sum_{j \neq i} I(Y_j \leq Y_i) / (n - 1)$$

$$S_i = \text{sign} \left\{ \left(F_i - \frac{1}{2} \right) \left(G_i - \frac{1}{2} \right) \right\}$$

$$\chi_i = (H_i - F_i G_i) / \{ F_i (1 - F_i) G_i (1 - G_i) \}^{\frac{1}{2}}$$

$$\lambda_i = 4S_i \max \left\{ \left(F_i - \frac{1}{2} \right)^2, \left(G_i - \frac{1}{2} \right)^2 \right\}$$

El Chi-plot es un diagrama de dispersión de los pares $(\lambda_i, \chi_i)_{i=1\dots n}$.

El valor de λ_i es una medida de la distancia de la observación (X_i, Y_i) al centro de los datos. Adicionalmente, se pueden trazar las líneas horizontales en $\chi = -c_p/n^{\frac{1}{2}}$ y $\chi = c_p/n^{\frac{1}{2}}$, donde c_p se selecciona de modo que aproximadamente el 100p% de los pares (λ_i, χ_i) se sitúen entre estas dos líneas de control. Para $p = 0, 90, 0.95, 0.99$ los valores de c_p son 1.54, 1.78, 2.18, respectivamente, (Fisher y Switzer, 2001). Se pueden calcular otros valores mediante métodos Monte Carlo. Además, los autores recomiendan representar solamente los pares para los que $|\lambda_i| < 4 \left(\frac{1}{n-1} - \frac{1}{2} \right)^2$, con el fin de evitar observaciones engañosas. Todos los valores de λ_i deben estar en el intervalo $[-1, 1]$. En caso de que los datos constituyan una muestra bivalente con marginales continuas independientes, los valores de λ_i se distribuyen uniformemente. Sin embargo, si X e Y están asociadas, los valores de λ_i se presentarán formando agrupaciones; en particular, valores positivos de λ_i indican que X_i e Y_i son relativamente grandes (a la vez) o relativamente pequeñas (a la vez) respecto a sus medianas, mientras que λ_i negativos corresponden a X_i e Y_i situados en lados opuestos respecto a sus medianas.

En cada observación muestral, χ_i se puede interpretar como el coeficiente de correlación asociado con $(n-1)$ pares (X_{ij}, Y_{ij}) , ambos dicotomizados de la siguiente manera:

$$\begin{aligned} X_{ij} &= \begin{cases} 1 & \text{si } X_j \leq X_i \\ 0 & \text{en otro caso} \end{cases} \\ Y_{ij} &= \begin{cases} 1 & \text{si } Y_j \leq Y_i \\ 0 & \text{en otro caso} \end{cases} \end{aligned}$$

para todo $j \neq i$. Por tanto, $-1 \leq \chi_i \leq 1, \forall i = 1\dots n$. Además, $\sqrt{n}\chi_i$ es la raíz cuadrada del estadístico chi-cuadrado utilizado tradicionalmente para contrastar la independencia en la tabla de contingencia generada por los puntos de corte (X_i, Y_i) .

Si Y es función estrictamente creciente de X , entonces $\chi_i = 1$ para todos los puntos de corte muestrales, por el contrario, si Y es función estrictamente decreciente de X , $\chi_i = -1$ en esos mismos puntos. χ_i también puede considerarse como una medida empírica de la *dependencia en cuadrante positiva* (PQD),

definida como (Joe, 1997):

$$P(X \leq a_1, Y \leq a_2) \geq P(X \leq a_1) P(Y \leq a_2) \quad \forall a_1, a_2 \in \mathbb{R}$$

es decir, PQD se define globalmente para todo $a_1, a_2 \in \mathbb{R}$, mientras que χ_i mide esta dependencia localmente y la representa frente a la distancia del punto al centro de los datos.

Por otra parte, Fisher y Switzer (1985, 2001) sugieren que los patrones de dependencia observados mediante los chi-plots pueden ser útiles para identificar el modelo de cópula subyacente. Por ejemplo, datos normales bivariantes con correlación de Pearson positiva, ρ , producen pares (λ_i, χ_i) con $\chi_i \approx \rho_n$ cuando $\lambda_i \approx 0$.

El Chi-plot también es eficaz para detectar de la presencia de dependencia en colas, la cual habíamos cuantificado mediante los coeficientes de dependencia en la cola superior e inferior, λ_L y λ_U , definidos en términos de la cópula bivalente correspondiente. Teóricamente, el valor de χ_i es cero si

$$P(X \leq a_1, Y \leq a_2) = P(X \leq a_1) P(Y \leq a_2)$$

que es lo mismo que decir

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

Si la cópula tiene independencia en la cola inferior, sabemos que $\lambda_L = 0$ y todos los χ_i también son cero. por tanto, la inspección gráfica de los Chi-plots construidos sobre las observaciones situadas en el primer y tercer cuadrante permitirá explorar la dependencia en las colas superior e inferior, respectivamente, de la distribución bivalente, así como detectar la posible presencia de outliers.

Chi-plot de la muestra bivalente de retornos filtrados Dow Jones e Ibex35

Comenzamos la exploración de la dependencia local existente en la muestra bivalente, obtenida mediante el filtrado de los retornos logarítmicos diarios del

índice Dow Jones a través de un proceso PGARCH(1,1,1) con media constante y de los retornos logarítmicos diarios del índice Ibex35, en este caso a través de un proceso EGARCH(1,1) y media constante, construyendo el Chi-plot correspondiente. En la figura 4.1 aparece el diagrama de dispersión de los retornos filtrados junto con la representación del Chi-plot de los mismos. Se aprecia claramente la ausencia de independencia entre ambas series univariantes, ya que la mayoría de los pares representados se sitúan fuera de las líneas de control representadas por las líneas horizontales discontinuas. los valores de $\chi_i > 0$ son los que predominan, por lo que el signo de la dependencia puede considerarse positivo.

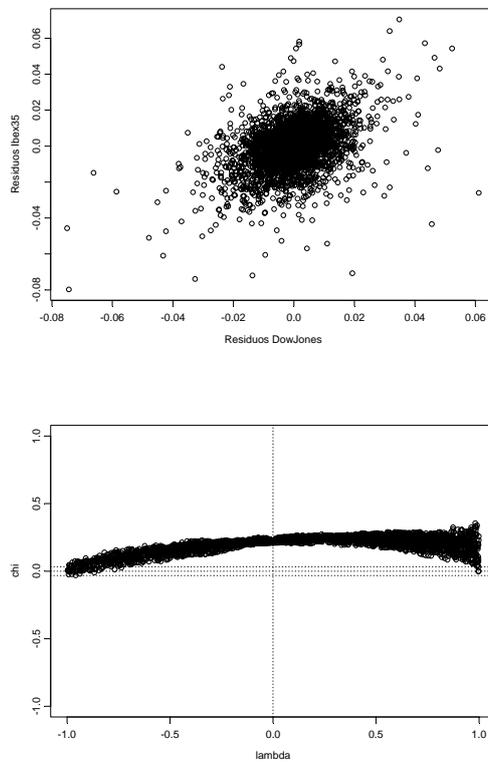


Figura 4.1: *Diagrama de dispersión (arriba) y Chi-plot (abajo) de la muestra bivalente de retornos filtrados DowJones e Ibex35*

La representación (figura 4.2) de los chi-plot correspondientes a los valores de la muestra bivalente que se sitúan en el primer y tercer cuadrante, permite confirmar la existencia de dependencia de signo positivo en ambas colas. En ambos casos, la práctica totalidad de los puntos se sitúan fuera de las bandas de control, con valores positivos para χ_i .

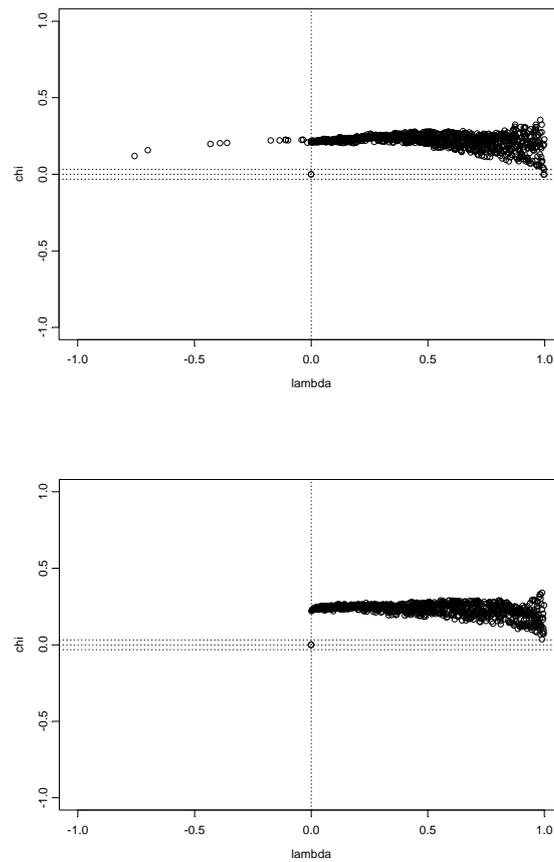


Figura 4.2: Chi-plot de los retornos filtrados DowJones e Ibox35 del primer cuadrante (arriba) y del tercer cuadrante (abajo)

4.1.2. K-plot

Al igual que los chi-plot, los K-plot (forma abreviada de Kendall-plot), desarrollados por Genest y Boies (2003), también se construyen sobre los rangos de las observaciones pero, así como los primeros se inspiran en el estadístico chi-cuadrado de independencia, los K-plots utilizan la transformación integral de probabilidades multivariantes, produciendo un gráfico similar al QQ-plot convencional.

Sea $(X_1, Y_1), \dots, (X_n, Y_n)$ una muestra aleatoria simple de H , función de distribución conjunta continua de la v.a. (X, Y) . Para construir el K-plot debe procederse de la siguiente manera:

(a) Para cada $1 \leq i \leq n$, se calcula H_i como en (4.1).

(b) Se ordenan los valores H_i de forma que $H_{(1)} \leq \dots \leq H_{(n)}$.

(c) Se dibujan los pares $(W_{i:n}, H_i)$, $1 \leq i \leq n$, donde $W_{i:n}$ representa la esperanza del estadístico de orden i -ésimo en una muestra de tamaño n , calculada a partir de la distribución K_0 de los H_i bajo la hipótesis nula de independencia.

Veamos, a continuación, como se determina la forma de K_0 bajo la hipótesis de independencia. A partir de la definición de la densidad de un estadístico de orden, se tiene que

$$W_{i:n} = n \binom{n-1}{i-1} \int_0^1 w [K_0(w)]^{i-1} [1 - K_0(w)]^{n-i} dK_0(w) \quad (4.2)$$

para todo $1 \leq i \leq n$. Bajo ciertas condiciones de regularidad (Genest y Rivest, 1993, Proposición 2.1), la función de distribución empírica K_n de las pseudo-observaciones H_1, \dots, H_n es un estimador \sqrt{n} -consistente y normal de

$$K(w) = P \{H(X, Y) \leq w\} \text{ con } 0 \leq w \leq 1$$

conclusión que se podía intuir dado que $H_i = \widehat{H}_n(X_i, Y_i) \rightarrow H$ cuando $n \rightarrow \infty$, con \widehat{H}_n la función de distribución empírica basada en los (X_j, Y_j) para $j \neq i$. A pesar de la no independencia estocástica de los H_i , Genest y Rivest (1993) demuestran que cuando $H = F \times G$

$$K(w) = K_0(w) = P \{UV \leq w\} = w - w \log(w) \quad 0 \leq w \leq 1$$

con U y V variables aleatorias independientes con distribución uniforme en $[0, 1]$.

Por tanto, esta será la elección de K que utilizaremos en (4.2) para calcular los $W_{i:n}$. Obsérvese que K depende únicamente de la cópula asociada con H :

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

$$\begin{aligned} K(w) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} I(H(x, y) \leq w) dH(x, y) \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} I(C(F(x), G(y)) \leq w) dC(F(x), G(y)) \\ &= \int_0^1 \int_0^1 I(C(u, v) \leq w) dC(u, v) \\ &= P\{C(U, V) \leq w\} \end{aligned}$$

K proporciona, por tanto, un resumen univariante de la dependencia implícita en la cópula C .

Genest y Rivest (1993, proposición 2.1) demuestran que

$$K_n(w) \rightarrow^P K(w) \quad \text{cuando } n \rightarrow \infty \quad \forall 0 \leq w \leq 1$$

y, por tanto también

$$K_n^{-1}(p) \rightarrow^P K^{-1}(p) \quad \text{cuando } n \rightarrow \infty \quad \forall 0 \leq p \leq 1$$

Puesto que K_n y su inversa están acotadas, esta convergencia se extiende naturalmente a sus esperanzas, con las consecuencias siguientes:

(i) Para un entero arbitrario $n \geq 1$ y $0 \leq p \leq 1$, sea $\lceil np \rceil$ el entero más pequeño mayor o igual que np . Entonces

$$H_{(\lceil np \rceil)} = K_n^{-1}(p) \rightarrow K^{-1}(p)$$

y, por tanto

$$\lim_{n \rightarrow \infty} E[H_{(\lceil np \rceil)}] = \lim_{n \rightarrow \infty} W_{\lceil np \rceil:n} = K_0^{-1}(p)$$

bajo la hipótesis nula de independencia.

(ii) Para un n suficientemente grande, los pares $(W_{i:n}, H_{(i)})$ tienden a concentrarse a lo largo de la curva $p \rightarrow (K_0^{-1}(p), K^{-1}(p))$, es decir, los puntos del k-plot parecerán una representación de $w \rightarrow K^{-1}(K_0(w))$.

(iii) El gráfico tenderá a ser lineal cuando $K = K_0$, es decir, bajo la hipótesis de independencia.

(iv) Todos los puntos del gráfico caerán sobre el eje horizontal ($p \equiv 0$) cuando X e Y sean contramonotónicas ($\tau(X, Y) = -1$), ya que $K^{-1}(p) = 0$ para todos los valores posibles de $0 \leq p \leq 1$.

(v) Todos los puntos del gráfico caerán sobre la curva $K_0(p)$ cuando X e Y sean comonotónicas ($\tau(X, Y) = 1$), puesto que $K^{-1}(p) = p$ sobre el intervalo $[0, 1]$.

K-plot de la muestra bivalente de retornos filtrados Dow Jones e Ibex35

La inspección gráfica de los K-plots construidos sobre la muestra bivalente de retornos filtrados (figura 4.3) conduce a las mismas conclusiones que habíamos establecido a partir de los Chi-plot.

Existe dependencia positiva entre los retornos filtrados ya que los puntos del gráfico se sitúan en una curva por encima de la diagonal principal. Además, los K-plots correspondientes a valores de la muestra que se sitúan en el primer y tercer cuadrante, confirman la presencia de dependencia positiva en ambas colas, derecha e izquierda, y en los dos casos la curvatura de los gráficos correspondientes es superior a la obtenida para el total de pares integrantes de la muestra bivalente.

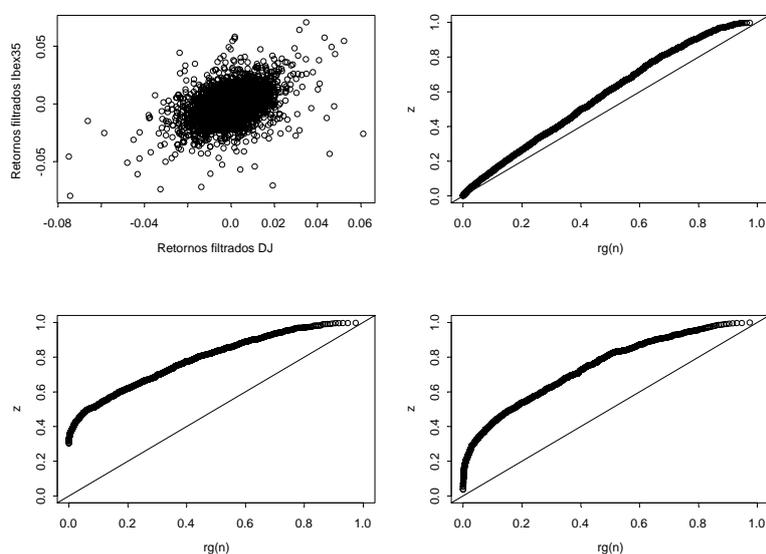


Figura 4.3: *Diagrama de dispersión y K-plot de la muestra bivalente de retornos filtrados DowJones e Ibe35(arriba). K-plots de los retornos bivalentes del primer cuadrante (abajo izquierda) y del tercer cuadrante (abajo derecha)*

4.2. Modelos de Extremos

La teoría de valores extremos (EVT) proporciona un fundamento teórico firme sobre el que construir modelos estadísticos que describan el comportamiento de sucesos extremos (McNeil 1998, McNeil y Frey 2000, Embrechts, Resnick y Samorodnitsky 1998 y 1999, Longin 1977a y 1977b), lo que en nuestro caso se concreta en la obtención de modelos relativos a colas de series financieras, que permitirán determinar medidas de riesgo y establecer los intervalos de confianza relacionados.

Los riesgos financieros pueden considerarse como variables aleatorias que, a partir de estados futuros del universo, toman valores representando beneficios y pérdidas. Estos riesgos pueden tratarse individualmente, o como parte de un proceso estocástico en el que los riesgos presentes dependen de los riesgos previos. Los valores potenciales de un riesgo tienen una distribución de probabilidad que, evidentemente, no es observable de manera exacta, pero que es posible estimar

a partir de información parcial procedente de pérdidas pasadas debidas a riesgos similares.

Los sucesos extremos ocurren cuando un riesgo toma valores en la cola de su distribución. Puesto que en lo que sigue trataremos con dos medidas concretas de riesgo, VaR y Pérdida media esperada (ES), convendremos en que una pérdida es un número positivo y un beneficio es un número negativo. Lógicamente, en el ámbito financiero, EVT está más desarrollada como una teoría de grandes pérdidas más que una teoría de beneficios pequeños.

En términos generales, existen dos métodos para modelizar los valores extremos:

- Método de máximos por bloques

Proporciona modelos para las observaciones más grandes de muestras grandes de observaciones idénticamente distribuidas (por ejemplo, si se registran pérdidas y beneficios diarios, o por hora, a partir de un instrumento o grupo de instrumentos financieros, este método proporciona un modelo apropiado para el máximo anual o cuatrimestral de tales valores). El punto delicado de este método lo constituye la elección apropiada de los periodos que definen los bloques.

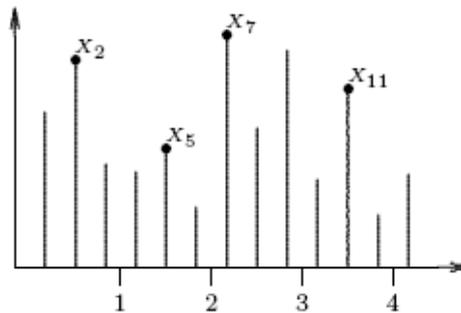
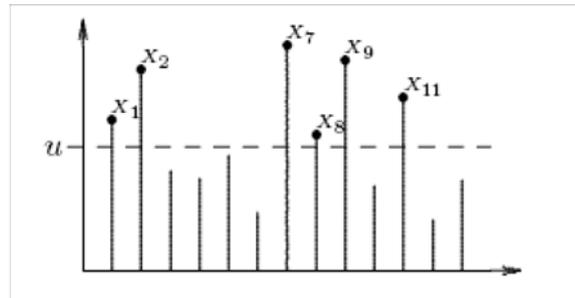


Figura 4.4: *Máximos por bloques*

- Métodos umbral (POT o Peaks Over Threshold)

Proporcionan modelos para las observaciones grandes que exceden un umbral alto. Los modelos POT generalmente resultan más útiles en aplicaciones prácticas, debido a su mayor eficiencia al tratar con un número limitado de datos, como es el caso de los valores extremos. Dentro de esta clase de modelos exis-

ten dos tipos de análisis: los modelos semiparamétricos construidos en torno al *estimador de Hill* (Danielsson y de Vries, 1997) y los modelos paramétricos basados en la *distribución de Pareto generalizada (GPD)*. Nos inclinamos por estos últimos por razones de simplicidad, tanto en la exposición como en la ejecución, ya que proporcionan expresiones paramétricas simples de las medidas de riesgo extremo, para las que es relativamente fácil dar estimaciones de error estadístico utilizando técnicas inferenciales de máxima verosimilitud.



Excesos sobre un umbral u

4.2.1. Distribución de máximos

A continuación, presentamos los resultados teóricos fundamentales que se utilizan tanto en el método de máximos por bloques como en el método de umbrales.

Cuando modelizamos el máximo de una v.a., EVT proporciona un resultado similar al del Teorema central del límite para sumas de v.a. pero aplicado sobre máximos de los bloques definidos. Es el llamado teorema de Fisher -Tippett.

Teorema 4.1 (*Fisher-Tippett*)

Sea $\{X_n\}$ un a sucesión de variables aleatorias i.i.d. con M_n el máximo del bloque de tamaño n . Si existen las constantes $c_n > 0$, $d_n \in \mathbb{R}$ y H función de distribución no degenerada tal que

$$\frac{M_n - d_n}{c_n} \rightarrow_d H$$

entonces H pertenece a una de las tres distribuciones de valores extremos siguientes:

$$\begin{aligned} \Phi_\alpha(x) &= \begin{cases} 0 & \text{si } x \leq 0 \\ e^{-x^{-\alpha}} & \text{si } x > 0 \end{cases} & \alpha > 0 & \text{(Fréchet)} \\ \Psi_\alpha(x) &= \begin{cases} e^{-(-x)^\alpha} & \text{si } x \leq 0 \\ 1 & \text{si } x > 0 \end{cases} & \alpha > 0 & \text{(Weibull)} \\ \Lambda(x) &= e^{-e^{-x}}, \quad x \in \mathbb{R} & & \text{(Gumbel)} \end{aligned}$$

La prueba de este resultado puede verse en Fisher-Tippett, 1928; Gnedenko, 1943.

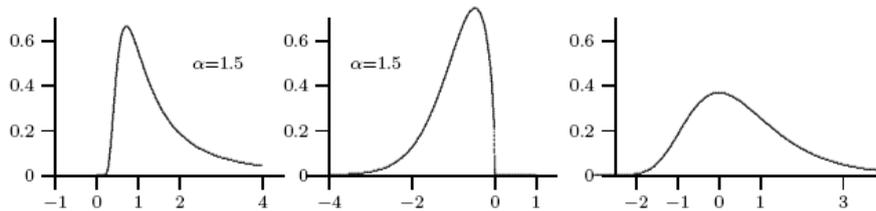


Figura 4.5: *Funciones de densidad Fréchet, Weibull y Gumbel*

Obsérvese que la distribución de Fréchet tiene una cola que decae polinomialmente y se ajusta bien a distribuciones de colas pesadas, mientras que la de la distribución Gumbel decae exponencialmente por lo que caracteriza distribuciones de colas delgadas. Finalmente, la distribución Weibull es la distribución asintótica de distribuciones con punto final finito.

Jenkinson (1955) y Von Misses (1954) propusieron la siguiente representación uniparamétrica:

Definición 4.1 (*GEV*)

Se denomina *distribución de valores extremos generalizada (GEV)* a

$$H_\xi(x) = \begin{cases} e^{-(1+\xi x)^{-\frac{1}{\xi}}} & \text{si } \xi \neq 0 \\ e^{-e^{-x}} & \text{si } \xi = 0 \end{cases} \quad \text{con } x \text{ tal que } 1 + \xi x > 0$$

Esta generalización se obtiene tomando $\xi = \alpha^{-1}$ para la distribución de Fréchet, $\xi = -\alpha^{-1}$ para la Weibull e interpretando la de Gumbel como el caso límite para $\xi = 0$.

Dado que, en general, se desconoce el tipo de distribución límite del máximo muestral, la representación generalizada es particularmente útil cuando se emplea en la estimación el método de máxima verosimilitud. La distribución GEV, definida anteriormente, es la distribución límite de extremos normalizados. Puesto que en la práctica desconocemos la distribución verdadera de los retornos y, por tanto, las constantes de normalización c_n y d_n también son desconocidas, utilizaremos la especificación triparamétrica siguiente de la GEV, que constituye la distribución límite de los máximos no normalizados:

$$H_{\xi, \mu, \sigma}(x) = H_{\xi}\left(\frac{x - \mu}{\sigma}\right) \quad x \in \mathcal{D}, \quad \text{con} \quad \mathcal{D} \in \begin{cases} \left(-\infty, \mu - \frac{\sigma}{\xi}\right) & \text{si } \xi < 0 \\ (-\infty, \infty) & \text{si } \xi = 0 \\ \left(\mu - \frac{\sigma}{\xi}, \infty\right) & \text{si } \xi > 0 \end{cases}$$

Los dos parámetros adicionales μ y σ son los parámetros de localización y escala, respectivamente, y representan las constantes de normalización desconocidas.

Las cantidades de interés no son los parámetros en sí mismos, sino los cuantiles (también llamados *niveles de retorno*) de la GEV estimada:

$$R^k = \left\{ H_{\xi, \mu, \sigma}^{-1}\left(1 - \frac{1}{k}\right)\right.$$

Sustituyendo ξ , μ y σ por sus estimaciones $\hat{\xi}$, $\hat{\mu}$ y $\hat{\sigma}$ obtenemos

$$\hat{R}^k = \begin{cases} \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}} \left(1 - \left(-\log\left(1 - \frac{1}{k}\right)\right)^{-\hat{\xi}}\right) & \text{si } \xi \neq 0 \\ \hat{\mu} - \hat{\sigma} \log\left(-\log\left(1 - \frac{1}{k}\right)\right) & \text{si } \xi = 0 \end{cases}$$

(Un valor de 7 para \hat{R}^k indica que la pérdida máxima observada durante el periodo de un año excederá el 7% una vez en 10 años, en media).

4.2.2. Distribución de excesos

Una aproximación alternativa consiste en considerar la distribución de excesos sobre cierto umbral (POT). Se trata de estimar, a partir de la función de distribución desconocida F , la función de distribución F_u de valores de x sobre el umbral u :

Definición 4.2

Sea X una v.a. con función de distribución F y punto final por la derecha $x_F \leq \infty$. Para un $u < x_F$ fijo

$$F_u(y) = P\{X - u \leq y / X > u\} \quad \text{con } 0 \leq y \leq x_F - u$$

es la función de distribución de los excesos de la v.a X (con función de distribución F) sobre el umbral u .

La distribución de excesos representa la probabilidad de que una pérdida exceda el umbral u como máximo un cantidad y , dado que ha superado el umbral u . De forma general, asumiremos que F tiene punto final derecho infinito, es decir, se permite la posibilidad de pérdidas arbitrariamente grandes, aunque puede ocurrir en ciertas aplicaciones que sea finito, por ejemplo, si F representa pérdidas de crédito expresadas como una proporción de exposición, puede ser una distribución *Beta* en el intervalo $[0, 1]$ en la que los resultados mayores que 1 tienen probabilidad nula.

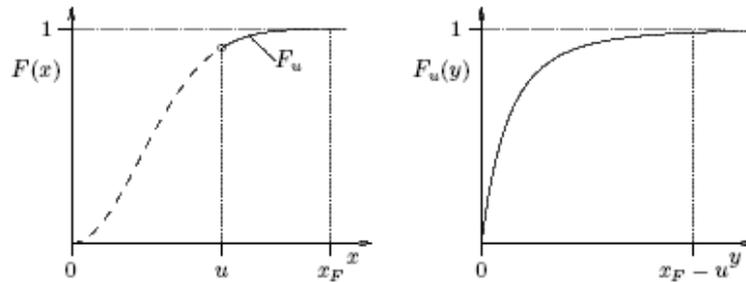


Figura 4.6: Función de distribución F y función de distribución de excesos F_u

Definición 4.3

Se denomina función media de excesos a

$$e(u) = E[X - u | X > u]$$

F_u se puede escribir en términos de F como:

$$F_u(y) = \frac{F(u+y) - F(u)}{1 - F(u)} = \frac{F(x) - F(u)}{1 - F(u)} \quad (4.3)$$

Las realizaciones de X permanecen en su mayor parte entre 0 y u , y la estimación de F en ese intervalo generalmente no plantea problemas. Sin embargo, la de F_u puede complicarse debido a que, en general, existe un número pequeño de observaciones en este área.

Definición 4.4 (GPD)

Sea G_ξ dada por

$$G_\xi(x) = \begin{cases} 1 - (1 + \xi x)^{-\frac{1}{\xi}} & \text{si } \xi \neq 0 \\ 1 - e^{-x} & \text{si } \xi = 0 \end{cases}$$

donde

$$\begin{aligned} x &\geq 0 & \text{si } \xi &\geq 0 \\ 0 &\leq x \leq -\frac{1}{\xi} & \text{si } \xi &< 0 \end{aligned}$$

G_ξ es la llamada **distribución de Pareto generalizada (GPD)**.

El caso G_0 puede interpretarse como el límite de G_ξ cuando $\xi \rightarrow 0$.

Se pueden introducir en la definición dos parámetros, uno de localización y otro de escala, con lo cual tenemos la familia $G_{\xi,\mu,\sigma}$ sustituyendo el argumento x por $\left(\frac{x-\mu}{\sigma}\right)$ para $\mu \in \mathbb{R}$ y $\sigma > 0$ y reajustando el soporte convenientemente. También nos referiremos a $G_{\xi,\mu,\sigma}$ como GPD.

La distribución GPD juega un papel importante a la hora de determinar la distribución de los excesos sobre un determinado umbral, como queda patente en el siguiente resultado.

Teorema 4.2 (Pickands, 1975)

Para una gran clase de funciones de distribución F , la función de distribución de excesos F_u , para un valor de u grande, puede aproximarse por:

$$F_u(y) \cong G_{\xi, \sigma}(y) \quad u \rightarrow \infty$$

donde

$$G_{\xi, \sigma}(y) = \begin{cases} 1 - \left(1 + \frac{\xi}{\sigma} y\right)^{-\frac{1}{\xi}} & \text{si } \xi \neq 0 \\ 1 - e^{-\frac{y}{\sigma}} & \text{si } \xi = 0 \end{cases}$$

para

$$\begin{aligned} y &\in [0, (x_f - u)] && \text{si } \xi \geq 0 \\ y &\in \left[0, -\frac{\sigma}{\xi}\right] && \text{si } \xi < 0 \end{aligned}$$

Este resultado permite afirmar que el modelo GPD es el natural para la distribución de excesos desconocida sobre umbrales suficientemente altos, por tanto, para un riesgo individual X_i con función de distribución F supondremos que, para un cierto umbral u , la distribución de excesos sobre u es una GPD para algún ξ y σ . ξ es el parámetro de forma o “índice cola” (que puede ser negativo, positivo o cero) y σ es el parámetro de escala. El índice cola ξ da una indicación de la pesadez de la cola: cuanto más grande sea, más gruesa es la cola. Como, en general, no es posible fijar una cota superior para pérdidas financieras, las distribuciones que mejor se ajustan a series de retornos financieros son aquellas en las que $\xi > 0$.

Suponiendo que la distribución de la cola es una GPD, podemos obtener expresiones analíticas para el VaR y ES como funciones de los parámetros de la GPD. En primer lugar, despejamos $F(x)$ de (4.3):

$$F(x) = (1 - F(u)) F_u(y) + F(u)$$

y sustituyendo F_u por la GPD y $F(u)$ por la estimación $\frac{n - N_u}{n}$, donde n es el número total de observaciones y N_u es el número de observaciones sobre el umbral u , se obtiene el estimador de la cola derecha:

$$\widehat{F}(x) = \frac{N_u}{n} \left(1 - \left(1 + \frac{\widehat{\xi}}{\widehat{\sigma}} (x - u) \right)^{-\frac{1}{\widehat{\xi}}} \right) + \left(1 - \frac{N_u}{n} \right)$$

simplificando

$$\widehat{F}(x) = 1 - \frac{N_u}{n} \left(1 + \frac{\widehat{\xi}}{\widehat{\sigma}} (x - u) \right)^{-\frac{1}{\widehat{\xi}}} \quad (4.4)$$

que es válida para $x > u$.

Invirtiéndolo la expresión anterior para una probabilidad dada p :

$$\widehat{VaR}_p = u + \frac{\widehat{\sigma}}{\widehat{\xi}} \left(\left(\frac{n}{N_u} p \right)^{-\widehat{\xi}} - 1 \right) \quad (4.5)$$

que no es más que la estimación de un cuantil, cuando el cuantil es un parámetro desconocido de una distribución subyacente desconocida.

En cuanto a la pérdida esperada ES para una probabilidad dada p , podemos reescribir su expresión como:

$$\widehat{ES}_p = \widehat{VaR}_p + E \left[X - \widehat{VaR}_p / X > \widehat{VaR}_p \right] \quad (4.6)$$

donde el segundo término de la suma es el valor esperado de los excesos sobre el umbral \widehat{VaR}_p .

La función media de excesos para una GPD con $\xi < 1$ es

$$e(u) = E [X - u / X > u] = \frac{\sigma + \xi u}{1 - \xi} \quad \text{con } \sigma + \xi u > 0 \quad (4.7)$$

Esta función proporciona la media de los excesos de X sobre diferentes valores del umbral u .

Por otra parte, el modelo GPD para la distribución de excesos sobre el umbral u tiene una buena propiedad de estabilidad: si tomamos cualquier umbral más alto, tal como VaR_p para $p > F(u)$, la distribución de excesos sobre VaR_p es también GPD con el mismo parámetro de forma, pero diferente parámetro de escala. (Embretchs Klüpelberg y Mikosch 1997).

Proposición 4.1

si X es una v.a que sigue una GPD, entonces, para todos los enteros r tales que $r < \frac{1}{\xi}$, existen los r primeros momentos.

(La demostración puede verse en Embretchs Klüpelberg y Mikosch 1997, Pág. 165)

Análogamente, a partir de la definición de la pérdida esperada que recordemos era:

$$ES_p = E[X / X > VaR_p] \quad (4.8)$$

y a partir de (4.6), para $Z = VaR_p - u$ y X la v.a. que representa los excesos y sobre u , se tiene que:

$$\widehat{ES}_p = \widehat{VaR}_p + \frac{\widehat{\sigma} + \widehat{\xi} (\widehat{VaR}_p - u)}{1 - \widehat{\xi}} = \frac{\widehat{VaR}_p}{1 - \widehat{\xi}} + \frac{\widehat{\sigma} - \widehat{\xi}u}{1 - \widehat{\xi}} \quad (4.9)$$

Estimación de máxima verosimilitud

Dada una muestra $y = (y_1, \dots, y_n)$, el logaritmo de la función de verosimilitud para la GPD es

$$L(\xi, \sigma; y) = \begin{cases} -n \log \sigma - \left(\frac{1}{\xi} + 1\right) \sum_{i=1}^n \log \left(1 + \frac{\xi}{\sigma} y_i\right) & \text{si } \xi \neq 0 \\ -n \log \sigma - \frac{1}{\sigma} \sum_{i=1}^n y_i & \text{si } \xi = 0 \end{cases} \quad (4.10)$$

Calcularemos los valores $\widehat{\xi}$ y $\widehat{\sigma}$ que maximizan esta función, para la muestra determinada por las observaciones que exceden el umbral u . Los cuantiles altos y la pérdida esperada pueden calcularse a partir de las ecuaciones (4.5) y (4.9) sustituyendo los parámetros por sus estimaciones.

Intervalos de confianza

Si admitimos que se mantiene la teoría de grandes muestras, podemos construir intervalos de confianza para los parámetros ξ y σ mediante inversión del contraste de razón de verosimilitudes. Dado que en el procedimiento de estimación únicamente intervienen los excesos, puede ocurrir que las estimaciones se obtengan a partir de conjuntos de datos pequeños, lo que hace que no siempre se pueda confiar en las propiedades de optimalidad asintótica de los estimadores de máxima verosimilitud. Para tamaños muestrales pequeños, sugerimos utilizar la corrección basada en la metodología de Lawley (1956).

Además de intervalos de confianza simples para cada parámetro, también es posible construir intervalos conjuntos. Por ejemplo, un intervalo conjunto para los parámetros ξ y σ viene dado por el contorno al nivel $\chi_{\alpha,2}^2$ del logaritmo de la función de verosimilitud relativa, definida como $L(\xi, \sigma) - L(\hat{\xi}, \hat{\sigma})$.

Para construir intervalos de confianza simples, por ejemplo para ξ , calculamos

$$L^*(\xi) = \max_{\sigma} L(\xi, \sigma)$$

y el intervalo de confianza al nivel $(1 - \alpha)$ viene dado por los valores de ξ que satisfacen

$$L^*(\xi) > L(\hat{\xi}, \hat{\sigma}) - \frac{1}{2}\chi_{\alpha,1}^2$$

donde $\chi_{\alpha,1}^2$ es el cuantil $(1 - \alpha)$ de la distribución χ^2 con 1 grado de libertad. La curva $(\xi, L^*(\xi))$ se denomina *perfil del logaritmo de la verosimilitud* y, en general, no es simétrica respecto a su máximo.

Los intervalos de confianza basados en máxima verosimilitud para el VaR , se pueden obtener a partir de una versión reparametrizada de la GPD, definida como una función de ξ y VaR_p :

$$G_{\xi, VaR_p}(y) = \begin{cases} 1 - \left(1 + \frac{\left(\frac{n}{N_u}p\right)^{-\xi} - 1}{VaR_p - u} y\right)^{-\frac{1}{\xi}} & \text{si } \xi \neq 0 \\ 1 - \frac{n}{N_u} p e^{\left(\frac{y}{VaR_p - u}\right)} & \text{si } \xi = 0 \end{cases}$$

La función de densidad correspondiente es

$$g_{\xi, VaR_p}(y) = \begin{cases} \frac{\left(\frac{n}{N_u}p\right)^{-\xi} - 1}{\xi(VaR_p - u)} \left(1 + \frac{\left(\frac{n}{N_u}p\right)^{-\xi} - 1}{VaR_p - u} y\right)^{-\frac{1}{\xi} - 1} & \text{si } \xi \neq 0 \\ -\frac{\frac{n}{N_u} p e^{\left(\frac{y}{VaR_p - u}\right)}}{VaR_p - u} & \text{si } \xi = 0 \end{cases}$$

Análogamente, usando la siguiente reparametrización para $\xi \neq 0$:

$$G_{\xi, ES_p}(y) = \left\{ 1 - \left(1 + \frac{\xi + \left(\frac{n}{N_u}p\right)^{-\xi} - 1}{(ES_p - u)(1 - \xi)} y\right)^{-\frac{1}{\xi}} \right.$$

$$g_{\xi, ES_p}(y) = \left\{ \frac{\xi + \left(\frac{n}{N_u} p\right)^{-\xi} - 1}{\xi(1-\xi)(ES_p - u)} \left(1 + \frac{\xi + \left(\frac{n}{N_u} p\right)^{-\xi} - 1}{(1-\xi)(ES_p - u)} \right)^{-\frac{1}{\xi} - 1} \right.$$

es posible calcular un intervalo de confianza basado en máxima verosimilitud para la pérdida esperada.

4.2.3. Selección del umbral

El principal problema del método POT está en la selección del umbral u . La elección del umbral u es, básicamente, un compromiso entre un valor suficientemente alto que garantice la aplicación del resultado asintótico del Teorema 4.2, y un valor suficientemente bajo que permita seleccionar un número suficiente de valores muestrales que permita realizar la estimación de la cola de la distribución.

Se puede utilizar una herramienta gráfica, basada en la representación de la *función muestral media de excesos*

Definición 4.5

Llamamos *función muestral media de excesos* a

$$e_n(u) = \frac{\sum_{i=k}^n (x_i^n - u)}{n - k + 1} \quad k = \text{mín} \{i / x_i^n > u\}$$

con $n - k + 1$ el número de observaciones que exceden el umbral u .

La función muestral media de excesos es una estimación de la función media de excesos, $e(u)$, que para la función GPD se ha establecido en la definición 4.3 y es lineal para $\xi < 1$, por tanto, la representación de los puntos

$$(u, e_n(u)) \quad \text{con } x_1^n < u < x_n^n$$

informará gráficamente sobre la linealidad para diferentes valores de u , es decir, se elegirá $u > 0$ tal que $e_n(u)$ sea aproximadamente lineal para $x \geq u$. Una vez elegido un valor para el umbral u , se puede proceder a realizar un QQ-plot que enfrente los cuantiles muestrales de los excesos a ese umbral y los de la GPD correspondiente. Si se sitúan sobre una línea recta, el hecho de que los datos provienen de una GPD es aceptable. Para datos financieros, $\alpha = \frac{1}{\xi}$ en la mayoría de los casos, toma valores en el rango (3,4) (Embretchs Klüpelberg y Mikosch 1997, Pág. 291).

Por otra parte, según el teorema 4.2, si una distribución GPD resulta un modelo razonable para los excesos sobre un umbral u_0 , entonces los excesos sobre un umbral más alto u también deben seguir una distribución GPD con el mismo parámetro de forma. Por tanto, las estimaciones de dicho parámetro ξ deben ser constantes sobre u_0 . En la práctica, las estimaciones de estas cantidades no resultarán exactamente constantes, pero deben estabilizarse en torno a un valor. Por tanto, la representación gráfica de las estimaciones $\hat{\xi}$ frente a distintos valores de u , junto con los intervalos de confianza para estas cantidades, facilitará la elección de u_0 como el valor más pequeño del umbral para el que las estimaciones se vuelven aproximadamente constantes.

4.2.4. Estimación de las colas de las distribuciones de pérdidas DowJones e Ibex35

Anteriormente se ha justificado la elección de métodos umbral (POT) para obtener modelos univariantes de las colas de las distribuciones marginales, debido a la posibilidad que ofrecen de estimar por separado las colas de cada distribución univariante y por razones de simplicidad, ya que proporcionan expresiones paramétricas simples de las medidas de riesgo extremo. Serán, por tanto, distribuciones GPD las elegidas para este cometido.

Dado que nuestro objetivo es el cálculo de medidas de riesgo financiero, en lo que sigue convendremos que una pérdida es un número positivo y un beneficio un número negativo, de manera que limitaremos el análisis a la cola derecha de las distribuciones marginales.

En primer lugar, exploramos el comportamiento de las dos colas de pérdidas, mediante un QQ-plot en el que la distribución de referencia es la exponencial. Recordemos que si la distribución de excesos sobre un umbral es de colas finas, entonces la GPD es exponencial con $\xi = 0$ y el QQ-plot debe ser lineal.

La figura 4.7 representa estos QQ-plots para los retornos negativos de los dos índices. En ambos casos, observamos un distanciamiento de la linealidad, más acusado en las pérdidas correspondientes al índice Dow Jones. Por tanto, el siguiente paso consistirá en determinar el valor más adecuado para el umbral sobre el que se ajustarán las GPD correspondientes. Para ello, calculamos y representamos la función media de excesos para las dos series de retornos filtrados (figura 4.8). Se aprecia una situación similar en los dos gráficos: la función es

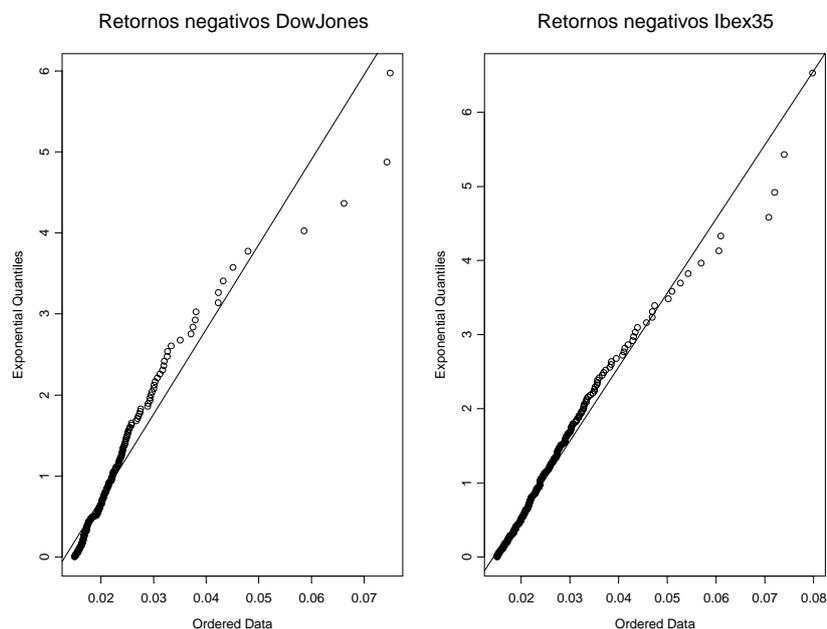


Figura 4.7: *QQ-plots de los retornos negativos DowJones e Ibex35 sobre el umbral $u=0.015$ frente a la exponencial*

lineal para $u < -0.02$ y $u > 0.02$, por lo que una primera aproximación sobre el valor del umbral en ambos casos podría ser $u = 0.02$.

El gráfico 4.9 muestra la variación de $\hat{\xi}$ según distintos valores del umbral u . Se aprecia que en el caso DowJones (parte superior del gráfico) el valor de ξ se estabiliza para valores del umbral inferiores o iguales a 0.02 y en el caso Ibex35 (parte inferior del gráfico) lo hace para valores de hasta 0.03, lo cual confirma la aproximación obtenida para los posibles valores de u mediante la función media de excesos (recordemos que la elección del umbral constituye un compromiso entre valores altos del mismo para asegurar el resultado asintótico establecido en el teorema 4.2, y valores suficientemente bajos que permitan seleccionar un número suficiente de observaciones para realizar las estimaciones de la cola de la distribución).

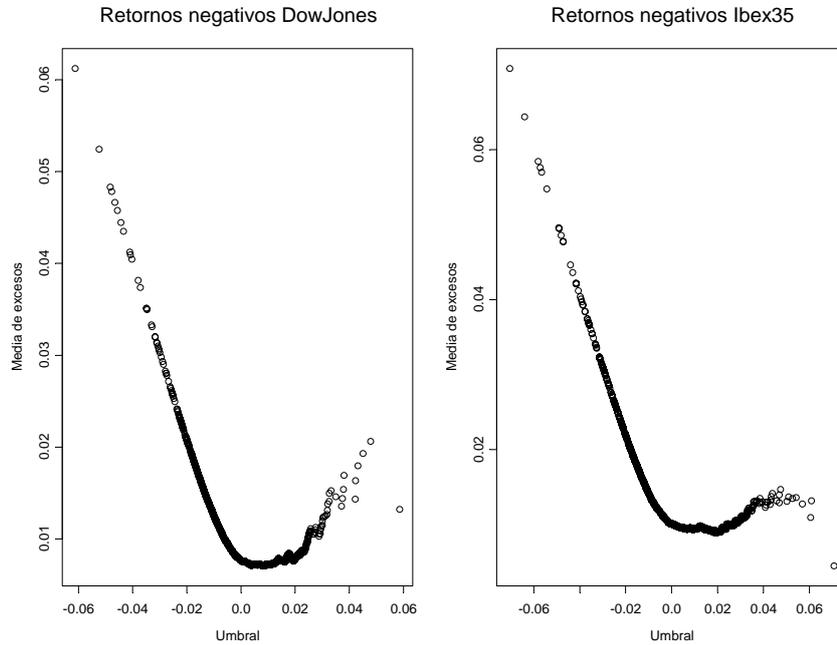


Figura 4.8: *Función media de excesos para los retornos negativos DowJones e Ibex35*

Una vez determinados los valores de los umbrales, procedemos a obtener las estimaciones de máxima verosimilitud de los parámetros ξ y σ de las distribuciones GPD correspondientes a las colas de pérdidas en cada uno de los índices, a partir de (4.10). Los resultados obtenidos aparecen en la tabla 4.1. Los valores entre paréntesis a la derecha de las estimaciones de los parámetros corresponden a los errores estándar de estimación.

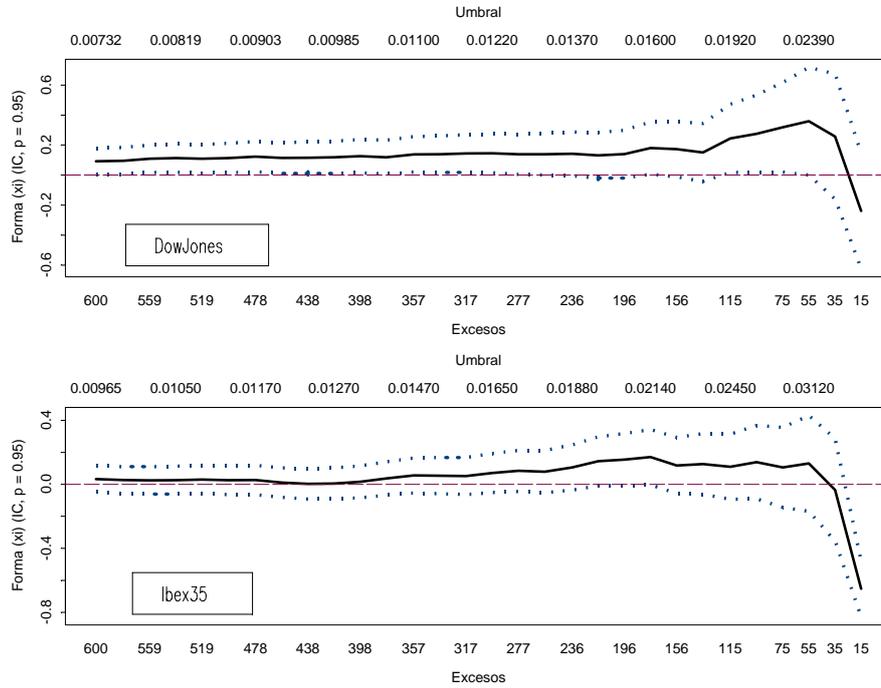


Figura 4.9: Estimaciones del valor de ξ como función del umbral u para los retornos negativos DowJones e Ibex35

	DowJones	Ibex35
u	0.02	0.02
n	103	208
$\hat{\xi} (\hat{se})$	0.2859 (0.1283)	0.1491 (0.0828)
$\hat{\sigma} (\hat{se})$	0.0055 (0.0088)	0.0076 (0.0008)
% datos $> u$	96.65	93.24
$-max(logveros.)$	-402.0554	-775.5655

Tabla 4.1: Resultados de la estimación de las colas de pérdidas DowJones e Ibex35 para $u = 0.02$

Los valores de las estimaciones $\widehat{\xi}_{DJ} = 0.2859$ y $\widehat{\xi}_{IB} = 0.1491$ se alejan de cero, por lo que nos encontramos con distribuciones de colas pesadas. Dichos valores, también garantizan la existencia de momentos de primer orden en ambas distribuciones (recordemos que existen los momentos de orden $r < 1/\xi$, para r entero), por lo que es posible calcular la pérdida esperada (ES) definida en (4.9).

Para realizar el diagnóstico de los modelos ajustados utilizamos métodos de tipo gráfico. En la figura 4.10 aparecen cuatro gráficos, construidos sobre la GPD estimada a partir de las pérdidas de los retornos filtrados del índice Dow Jones sobre el umbral $u = 0.02$.

En la esquina superior izquierda aparece la GPD estimada de la distribución de excesos y en la parte superior derecha se muestra la estimación de la cola (4.4). El gráfico de la esquina inferior izquierda se corresponde con un gráfico de dispersión de los datos transformados a residuos exponenciales de parámetro unidad $W_i = \left(1 + \widehat{\xi} \frac{x_i - \widehat{\mu}}{\widehat{\sigma}}\right)^{-\frac{1}{\xi}}$ y, por último, en la parte derecha aparece un QQ-plot de los cuantiles de la distribución empírica frente a los de la GPD ajustada.

A la vista de todos ellos, podemos concluir que la distribución GPD de parámetros $\widehat{\xi}_{DJ} = 0.2859$ y $\widehat{\sigma}_{DJ} = 0.0055$ se ajusta razonablemente a la cola derecha (pérdidas) de los retornos filtrados Dow Jones.

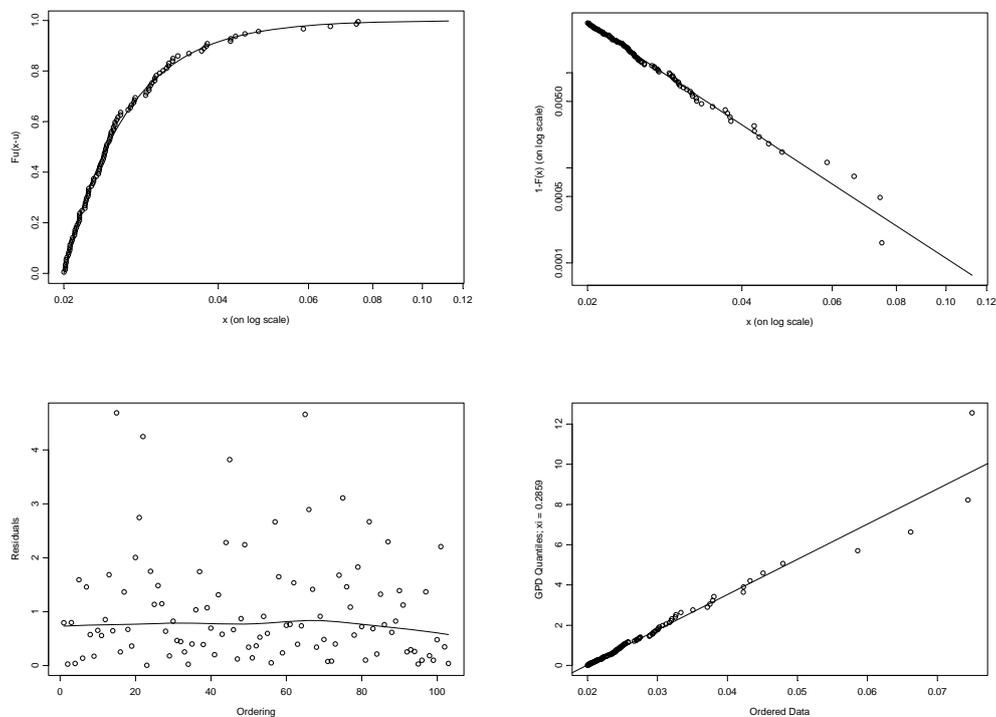


Figura 4.10: Gráficos de diagnóstico para la GPD ajustada a la distribución de pérdidas DowJones sobre el umbral $u = 0.02$

Procediendo de igual modo para efectuar el diagnóstico correspondiente a la GPD estimada, en este caso, a partir de las pérdidas de los retornos filtrados del índice Ibex35 sobre el umbral $u = 0.02$, recurrimos a la figura 4.11. Una observación detallada de los cuatro gráficos en ella representados y ya descritos anteriormente, confirma un buen ajuste de la distribución GPD de parámetros $\hat{\xi}_{IB} = 0.1491$ y $\hat{\sigma}_{IB} = 0.0076$ a la cola derecha (pérdidas) de los retornos filtrados Ibex35.

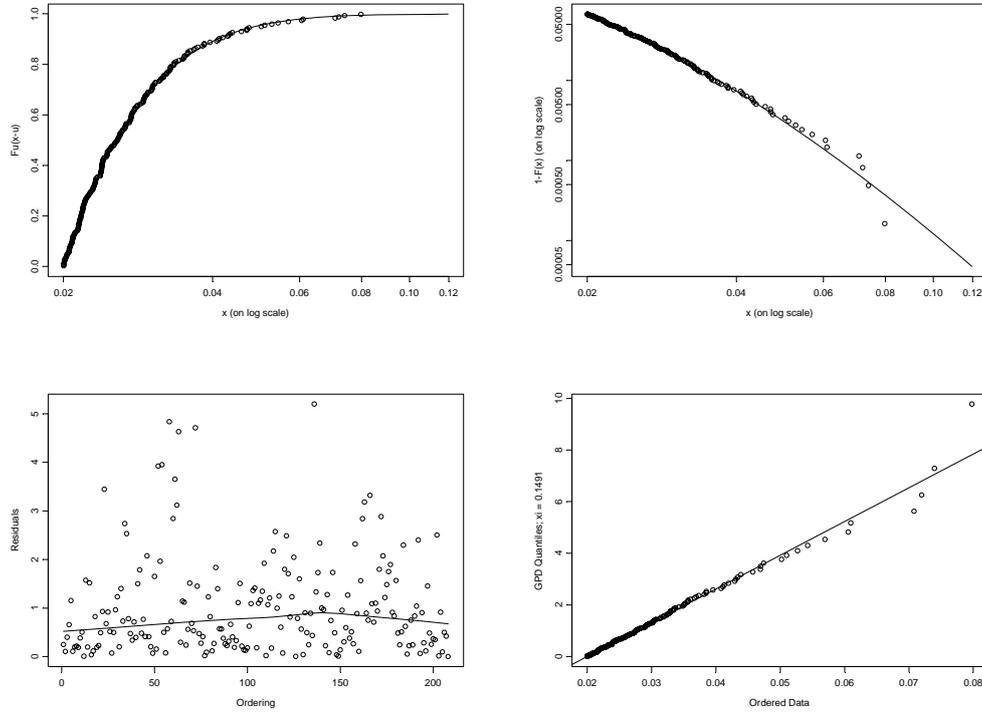


Figura 4.11: Gráficos de diagnóstico para la GPD ajustada a la distribución de pérdidas Ibo35 sobre el umbral $u = 0.02$

Por último, podemos comparar los resultados obtenidos mediante simulación de muestras a partir de las distribuciones GPD estimadas, del mismo tamaño que las originales de residuos filtrados. En la figura 4.13 aparecen representadas las observaciones muestrales originales en la columna de la izquierda y las simuladas en la columna de la derecha. Los cuatro gráficos superiores corresponden al índice Dow Jones y los cuatro inferiores al índice Ibo35. En las filas primera y tercera observamos los diagramas de dispersión correspondientes a la totalidad de las observaciones, mientras que los de las filas segunda y cuarta corresponden a los excesos sobre el umbral. En ambos índices se observa bastante similitud entre la disposición de las observaciones originales y las simuladas.

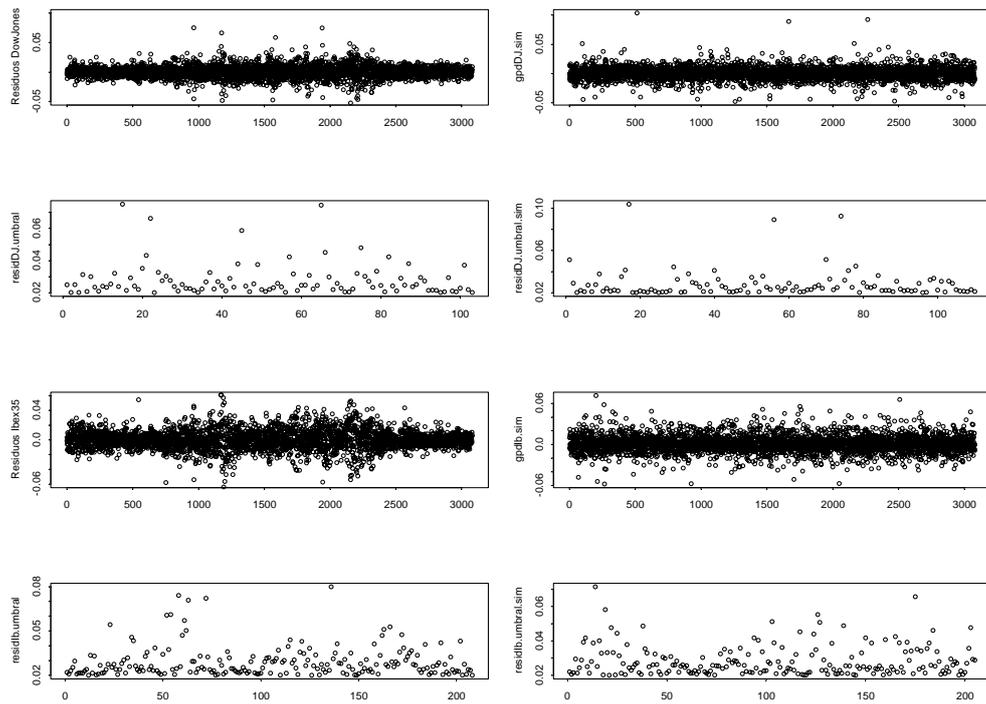


Figura 4.12: *Datos originales (izquierda) y muestras simuladas (derecha) a partir de las distribuciones GPD ajustadas a los residuos DowJones (parte superior) e Ibox35 (parte inferior) sobre un umbral $u = 0.02$. En las filas segunda y cuarta se representan los excesos al umbral en los datos originales (izquierda) y en las muestras simuladas (derecha).*

4.2.5. Medidas de riesgo para las distribuciones de pérdidas univariantes DowJones e Ibex35

A partir de las distribuciones GPD ajustadas a las pérdidas univariantes de los dos índices considerados, es posible calcular las medidas de riesgo financiero usuales, VaR y ES, según las expresiones (4.5) y (4.9), respectivamente y para distintos niveles de probabilidad fijados.

Las estimaciones de ambas medidas para la cola de pérdidas estimada (según una GPD de parámetros $\hat{\xi} = 0.2859$ y $\hat{\sigma} = 0.0055$) de los retornos filtrados Dow Jones y para distintos niveles de probabilidad p , junto con sus intervalos de confianza al 95%, aparecen en la tabla 4.2.

Dado que las distribuciones GPD trabajan mejor para cuantiles altos y puesto que el porcentaje de observaciones que caen por debajo del umbral elegido, $u = 0.02$, es de 96.65%, se han seleccionado como niveles de probabilidad para el cálculo, tanto del VaR como de ES, $p = (0.98, 0.99, 0.995, 0.999)$. La tabla 4.3 recoge los valores de dichas medidas calculados en el caso normal. Si consideramos, por ejemplo, $p = 0.99$, la interpretación de los valores obtenidos es la siguiente: la pérdida diaria del índice Dow Jones será superior al 2.8% con una probabilidad del 1%, y si la pérdida diaria es superior al 2.8%, la media de las pérdidas esperadas será del 3,91% con probabilidad 1%. Por otra parte, si comparamos las dos tablas, se observa que para ambas medidas y en todos los niveles de probabilidad p considerados, se obtienen valores superiores en el caso de colas estimadas según distribuciones GPD frente al supuesto gaussiano, es decir, el hecho de no contemplar el fenómeno de colas pesadas conlleva una subestimación de la pérdida futura. Obsérvese, también, que esta diferencia aumenta cuanto más alto es el cuantil considerado.

La última columna de ambas tablas proporciona el valor de la razón de la pérdida esperada al cuantil correspondiente, según el nivel de probabilidad fijado. De nuevo observamos valores inferiores en todos los casos cuando se considera la distribución normal. En este último caso, dicha razón tiende a la unidad a medida que $p \rightarrow 1$, ya que

$$ES_p = E[X / X > VaR_p] = k(VaR_p)$$

donde $K(x) = \phi(x)/(1 - \Phi(x))$ es el recíproco de la razón de Mill, con $\phi(x)$ y $\Phi(x)$ la función de densidad y la función de distribución de la normal estándar (la

razón de Mill está incorporada en la mayoría de los paquetes estadísticos usuales); para más detalles consultar Johnson y Kotz (1970). Se puede utilizar la expresión asintótica siguiente

$$k(x) = x(1 + x^{-2} + o(x^{-2}))$$

cuando $x \rightarrow \infty$, a partir de la cual es claro que la razón

$$ES_p/VaR_p \rightarrow_{p \rightarrow 1} 1$$

En el supuesto GPD, para $\xi = 0.2859 > 0$ esta razón converge a

$$(1 - \xi)^{-1} \rightarrow_{p \rightarrow 1} 1.40$$

Por tanto, la pérdida esperada, una vez superado el VaR, tenderá a ser 1.4 veces superior a dicho valor en el caso GPD, mientras que ambas medidas tienden a igualarse en un supuesto gaussiano.

p	VaR_p	$IC_{VaR_p;0.95}$	ES_p	$IC_{ES_p;0.95}$	$\frac{ES_p}{VaR_p} (\rightarrow_{p \rightarrow 1} 1.40)$
0.98	0.0231	[0.0223; 0.0240]	0.0321	[0.0294; 0.0388]	1.3896
0.99	0.0280	[0.0263; 0.0303]	0.0391	[0.0347; 0.0522]	1.3964
0.995	0.0340	[0.0311; 0.0384]	0.0475	[0.0400; 0.0723]	1.3970
0.999	0.0536	[0.0445; 0.0770]	0.0749	[0.0546; 0.1124]	1.3973

Tabla 4.2: Estimaciones puntuales e intervalos de confianza al 95% para las medidas de riesgo DowJones según una GPD($\hat{\xi} = 0.2859; \hat{\sigma} = 0.0055$)

p	VaR_p	ES_p	$\frac{ES_p}{VaR_p} (\rightarrow_{p \rightarrow 1} 1)$
0.98	0.0216	0.0255	1.1805
0.99	0.0244	0.0280	1.1475
0.995	0.0271	0.0304	1.1217
0.999	0.0325	0.0354	1.0892

Tabla 4.3: Cuantiles y ES de la $N(5.82e-05;0.0105)$ calculados sobre los retornos filtrados DowJones

Los resultados de un análisis similar al anterior, pero realizado en esta ocasión sobre los retornos filtrados Ibex35 aparecen en las tablas 4.4 (estimación según una GPD de parámetros $\hat{\xi} = 0.1491$ y $\hat{\sigma} = 0.0076$) y 4.5 (según la normal), para los mismos niveles de probabilidad. De nuevo, todos los valores de las medidas de riesgo y para todos los niveles de probabilidad considerados, son superiores en el caso de estimar la cola de pérdidas mediante una distribución GPD. Para $p = 0.99$, la pérdida diaria del índice Ibex35 será superior al 3.68% con probabilidad 1%, y si la pérdida diaria supera el 3.68%, la media de las pérdidas esperadas será del 4.87% con igual probabilidad. También en este caso, la suposición de normalidad en la cola de pérdidas produce una subestimación de la pérdida futura, ya sea en valor (VaR) o en media (ES).

En cuanto a la razón $\frac{ES_p}{VaR_p}$, ésta toma valores superiores en el caso GPD, en el que cuando $p \rightarrow 1$ tiende al valor 1.17, mientras que en el caso gaussiano su valor sigue siendo 1. En este caso los valores del límite en ambas razones están más próximos que los correspondientes al índice Dow Jones, debido a que el valor estimado del parámetro $\hat{\xi} = 0.1491$ de la distribución GPD está más cerca de cero (recordemos que en el caso Dow Jones $\hat{\xi} = 0.2859$) y, por tanto, las colas de la distribución son menos pesadas y más parecidas a las de la distribución normal que las del caso anterior.

p	VaR_p	$IC_{VaR_p;0.95}$	ES_p	$IC_{ES_p;0.95}$	$\frac{ES_p}{VaR_p} (\rightarrow_{p \rightarrow 1} 1.17)$
0.98	0.0301	[0.0286; 0.0318]	0.0408	[0.0377; 0.0463]	1.3554
0.99	0.0368	[0.0344; 0.0397]	0.0487	[0.0438; 0.0581]	1.3233
0.995	0.0442	[0.0406; 0.0495]	0.0574	[0.0500; 0.0729]	1.2986
0.999	0.0646	[0.0555; 0.0833]	0.0814	[0.0650; 0.1196]	1.2600

Tabla 4.4: Estimaciones puntuales e intervalos de confianza al 95% para las medidas de riesgo Ibex35 según una GPD($\hat{\xi} = 0.1491$; $\hat{\sigma} = 0.0076$)

p	VaR_p	ES_p	$\frac{\text{ES}_p}{\text{VaR}_p} (\rightarrow_{p \rightarrow 1} 1)$
0.98	0.0277	0.0326	1.1768
0.99	0.0314	0.0359	1.1433
0.995	0.0347	0.0389	1.1210
0.999	0.0416	0.0453	1.0889

Tabla 4.5: Cuantiles y ES de la $N(3.20e-04;0.0133)$ calculados sobre los retornos filtrados *Ibex35*

Para finalizar, en la figura 4.14 aparecen representadas las estimaciones de las colas de pérdidas según la distribución GPD correspondiente a cada índice considerado, junto con las curvas correspondientes al perfil del logaritmo de la verosimilitud, sobre las que se marcan los intervalos de confianza al 95% para el $\widehat{\text{VaR}}_{0.99}$ y el $\widehat{\text{ES}}_{0.99}$. Se aprecia la asimetría de dicha curva respecto a su máximo en el caso de la estimación $\widehat{\text{ES}}_{0.99}$, como era de esperar dada la presencia de sucesos extremos.

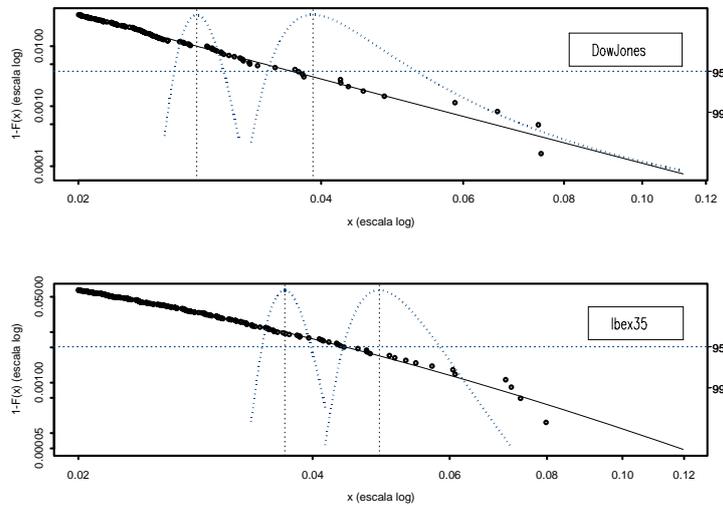


Figura 4.13: Estimación de la cola de pérdidas según una GPD. $\text{VaR}_{0.99}$ y $\text{ES}_{0.99}$ con sus correspondientes intervalos de confianza al 95% para los retornos filtrados *DowJones* (arriba) e *Ibex35* (abajo)

4.3. Medidas de Riesgo para una Cartera de Inversión

Una consecuencia directa del Teorema de Sklar es el uso de las funciones cópulas como instrumento para medir el riesgo de mercado, en concreto para calcular el *valor en riesgo* (VaR) de una cartera de inversión construida sobre índices de mercado.

Si denotamos por R los retornos de la cartera sobre un horizonte dado T , VaR es el valor tal que, para una probabilidad α fijada

$$P(R \leq VaR_\alpha) = \alpha$$

De forma equivalente, utilizando la función de distribución de R , $F(r)$, VaR se puede definir como la solución z^* de la ecuación $F(z^*) = \alpha$. A su vez, la función de distribución $F(r)$ se puede escribir vía cópulas como sigue:

Consideremos una cartera con dos activos, sean X e Y sus retornos continuos sobre un horizonte común T y sea $\beta \in [0, 1]$ el peso de X . El retorno de la cartera es

$$R = \beta X + (1 - \beta)Y$$

con función de distribución

$$\begin{aligned} F(r) &= P\{R \leq r\} = P\{\beta X + (1 - \beta)Y \leq r\} \\ &= \int_{-\infty}^{+\infty} P\left\{X \leq \frac{1}{\beta}r - \frac{1 - \beta}{\beta}y, Y = y\right\} f_2(y) dy \\ &= \int_{-\infty}^{+\infty} C_{1/2}\left(F_1\left(\frac{1}{\beta}r - \frac{1 - \beta}{\beta}y\right), F_2(y)\right) f_2(y) dy \end{aligned}$$

con la probabilidad condicionada obtenida a través de:

$$C_{1/2}(u, v) = P[U \leq u/V \leq v] = \lim_{\Delta v \rightarrow 0^+} \frac{C(u, v + \Delta v) - C(u, v)}{\Delta v} = \frac{\partial C(u, v)}{\partial v}$$

A partir de esta expresión es posible calcular el VaR de la cartera para distintos niveles de confianza (probabilidades que determinan el cuantil) y horizontes fijados, considerando distribuciones conjuntas no normales, a partir de marginales de los retornos también no normales y estudiar su sensibilidad respecto a la combinación de activos establecida, variando los pesos de localización de cada uno de ellos (en nuestro caso, variando el valor de β). Es más, la aproximación mediante

la cópula permite cambiar las marginales mientras se mantiene fija la cópula (la familia elegida) o utilizar diferentes familias de cópulas manteniendo fijas las marginales. Esto permite jugar con situaciones tan diferentes como aquellas en las que se eliminan los efectos de colas pesadas en las distribuciones marginales considerando retornos distribuidos normalmente (obtendremos valores del VaR más pequeños en valor absoluto) o aquellas en las que se mantienen estos efectos considerando, por ejemplo, retornos distribuidos según una *t* de Student. También es posible suponer independencia entre retornos y utilizar la cópula independencia, con lo que de nuevo queda de manifiesto la gran flexibilidad aportada por la representación vía cópulas.

Nuestra propuesta consiste en cuantificar el riesgo de la pérdida conjunta, concretada a través de una cartera de inversión con unos pesos de localización especificados, utilizando como modelos de pérdidas marginales los obtenidos mediante el ajuste de distribuciones GPD univariantes a las colas de pérdidas de los índices implicados, ligados por el modelo de cópula que mejor represente la dependencia presente en las colas de las distribuciones univariantes.

La distribución de los retornos de la cartera debe aproximarse por simulación, dado que solo se conoce una expresión paramétrica exacta para ella en algunos modelos particulares. Para ello, utilizaremos la idea propuesta por Embrechts, McNeil y Straumann (1999) que permite, mediante distribuciones condicionadas, simular observaciones de modelos basados en cópulas.

Sean F_1, \dots, F_n distribuciones marginales continuas de forma que la cópula $C(u_1, \dots, u_n)$ especifica una única distribución multivariante

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n))$$

El problema teórico de simulación se reduce al que plantea la existencia de una solución de la ecuación anterior. Nuestro objetivo es simular un vector aleatorio $(U_1, \dots, U_n)^t$ a partir de C , basándonos en que $(F_1^{-1}(U_1), \dots, F_n^{-1}(U_n))^t$ tiene distribución F .

Consideramos el caso general $n > 2$ e introducimos la notación

$$C_i(u_1, \dots, u_n) = C(u_1, \dots, u_i, 1, \dots, 1) \quad i = 2, \dots, n - 1$$

para representar las distribuciones marginales i -dimensionales de $C(u_1, \dots, u_n)$. Bajo esta notación, $C_1(u_1) = u_1$ y $C_n(u_1, \dots, u_n) = C(u_1, \dots, u_n)$.

La distribución condicionada de u_i dados los valores de las $(i - 1)$ primeras componentes de $(u_1, \dots, u_n)^t$, se puede escribir en términos de derivadas y densidades de las marginales i -dimensionales

$$\begin{aligned} C_i(u_i/u_1, \dots, u_{i-1}) &= P[U_i \leq u_i / U_1 = u_1, \dots, U_{i-1} = u_{i-1}] \\ &= \frac{\delta^{i-1} C_i(u_1, \dots, u_i)}{\delta u_1 \dots \delta u_{i-1}} \bigg/ \frac{\delta^{i-1} C_{i-1}(u_1, \dots, u_{i-1})}{\delta u_1 \dots \delta u_{i-1}} \end{aligned}$$

suponiendo que existen tanto el numerador como el denominador. En ese caso, seguiremos el siguiente algoritmo:

- Simulamos un valor u_1 a partir de una $U(0, 1)$
- Simulamos un valor u_2 a partir de $C_2(u_2/u_1)$
- Simulamos un valor u_3 a partir de $C_3(u_3/u_1, u_2)$
- \vdots
- Simulamos un valor u_n a partir de $C_n(u_n/u_1, \dots, u_{n-1})$

En nuestro caso ($n = 2$) una vez generado el vector aleatorio (U, V) , según el procedimiento anterior y a partir del modelo de cópula seleccionado según se indicó en el Capítulo 3, supondremos que las distribuciones marginales correspondientes son las GPD ajustadas a las pérdidas de los índices considerados:

$$\begin{aligned} U &\sim GPD(\hat{\xi}_{DJ}, \hat{\sigma}_{DJ}) \\ V &\sim GPD(\hat{\xi}_{IB}, \hat{\sigma}_{IB}) \end{aligned}$$

a partir de las cuales simularemos, mediante el método Montecarlo, los valores X_{sim} e Y_{sim} que producirán la muestra aleatoria bidimensional final $(x_{sim}^i, y_{sim}^i)_{i=1 \dots n}$.

Utilizando esta muestra bidimensional, generamos la muestra de retornos logarítmicos de la cartera según la expresión siguiente

$$r_i^* = \log(\lambda_1 \exp(x_{sim}^i) + \lambda_2 \exp(y_{sim}^i)) \quad (4.11)$$

La muestra así generada, permitirá calcular las medidas de riesgo en el periodo siguiente, VaR_{t+1} , al nivel de probabilidad α deseado de la siguiente forma

$$VaR_{t+1} = -r_{t+1, \alpha}^* \quad (4.12)$$

con $r_{t+1,\alpha}$ el cuantil de orden α de la distribución de retornos de la cartera. La pérdida esperada, ES_{t+1} , se calculará como

$$ES_{t+1} = -\frac{1}{N_{t+1}} \sum_{i=1}^N r_{i,t+1}^* I \{x \leq -VaR_{t+1}, x \in \mathbb{R}\} (r_{i,t+1}^*) \quad (4.13)$$

donde N_{t+1} es el número de retornos simulados $r_{i,t+1}^*$ con valor menor o igual que $-VaR_{t+1}$, es decir

$$N_{t+1} = \sum_{i=1}^N I \{x \leq -VaR_{t+1}, x \in \mathbb{R}\} (r_{i,t+1}^*)$$

4.3.1. Medidas de riesgo calculadas sobre la cartera DowJones e Ibex35

En lo que sigue, consideraremos una cartera de inversión integrada en un 50% por el índice Dow Jones y en un 50% por el índice Ibex35. En la figura 4.14 observamos el diagrama de dispersión de las pérdidas correspondientes a los retornos filtrados de ambos índices situados por debajo del umbral $u = -0.02$.

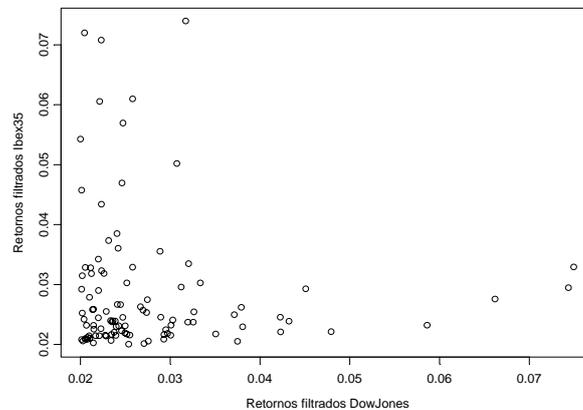


Figura 4.14: Diagrama de dispersión de los retornos filtrados DowJones e Ibex35 por debajo del umbral $u = -0.02$

Se han simulado 3000 retornos de la cartera especificada anteriormente para cada uno de los modelos mixtos de cópula con los que estamos trabajando. El

procedimiento de simulación se ajusta al algoritmo detallado en el punto anterior. Para cada modelo mixto de cópula se han seguido los siguientes pasos:

- En primer lugar, se obtiene la expresión de la función de densidad conjunta $c(u_1, u_2)$.
- Simulamos un valor u_1 a partir de una $U(0, 1)$
- A continuación, se determina la función de densidad condicionada $c(u_2/u_1)$.
- Se obtiene la función de distribución empírica $F^*(u_2/u_1)$ asociada a $c(u_2/u_1)$.
- Simulamos un valor u a partir de una $U(0, 1)$.
- Se obtiene el valor u_2 tal que $F^*(u_2/u_1) = u$.

El par (u_1, u_2) corresponde a un valor muestral del modelo de cópula en cuestión.

Las expresiones correspondientes a las funciones de densidad conjuntas de los modelos mixtos son, para el Modelo 1:

$$c_{Mod1}(u_1, u_2, \theta_1, \theta_2, \theta_3) = \theta_3 \times c_{G_u}(u_1, u_2, \theta_1) + (1 - \theta_3) \times c_{G_uS}(u_1, u_2, \theta_2)$$

con

$$\begin{aligned} c_{G_u}(u_1, u_2, \theta_1) &= (-\log u_1)^{(-1+\theta_1)} \left[-1 + \theta_1 + \left((-\log u_1)^{\theta_1} + (-\log u_2)^{\theta_1} \right) \right]^{\frac{1}{\theta_1}} \times \\ &\quad \left[(-\log u_1)^{\theta_1} + (-\log u_2)^{\theta_1} \right]^{\left(-2+\frac{1}{\theta_1}\right)} (-\log u_2)^{(-1+\theta_1)} \times \\ &\quad \left[\exp \left((-\log u_1)^{\theta_1} + (-\log u_2)^{\theta_1} \right)^{\frac{1}{\theta_1}} \times u_1 \times u_2 \right]^{-1} \\ c_{G_uS}(u_1, u_2, \theta_2) &= c_{G_u}((1 - u_1), (1 - u_2), \theta_2) \end{aligned}$$

y

$$\theta_1 = 1.252 \quad , \quad \theta_2 = 1.253 \quad , \quad \theta_3 = 0.498$$

Análogamente, la función de densidad conjunta para el Modelo 2 viene dada por

$$c_{Mod2}(u_1, u_2, \theta'_1, \theta'_2, \theta'_3) = \theta'_3 \times c_{G_u}(u_1, u_2, \theta'_1) + (1 - \theta'_3) \times c_{CI}(u_1, u_2, \theta'_2)$$

con

$$c_{CI}(u_1, u_2, \theta'_2) = (1 + \theta'_2) u_1^{(-1-\theta'_2)} u_2^{(-1-\theta'_2)} \left(-1 + \frac{1}{u_1^{\theta'_2}} + \frac{1}{u_2^{\theta'_2}} \right)^{\left(-2-\frac{1}{\theta'_2}\right)}$$

y

$$\theta'_1 = 1.172 \quad , \quad \theta'_2 = 1.022 \quad , \quad \theta'_3 = 0.625$$

En el gráfico 4.15 aparecen representadas las funciones de densidad condicionadas para diferentes valores de u_1 (0.1, 0.3, 0.5, 0.7 y 0.9) y sus correspondientes funciones de distribución empíricas para el Modelo 1 (parte superior) y el Modelo 2 (parte inferior).

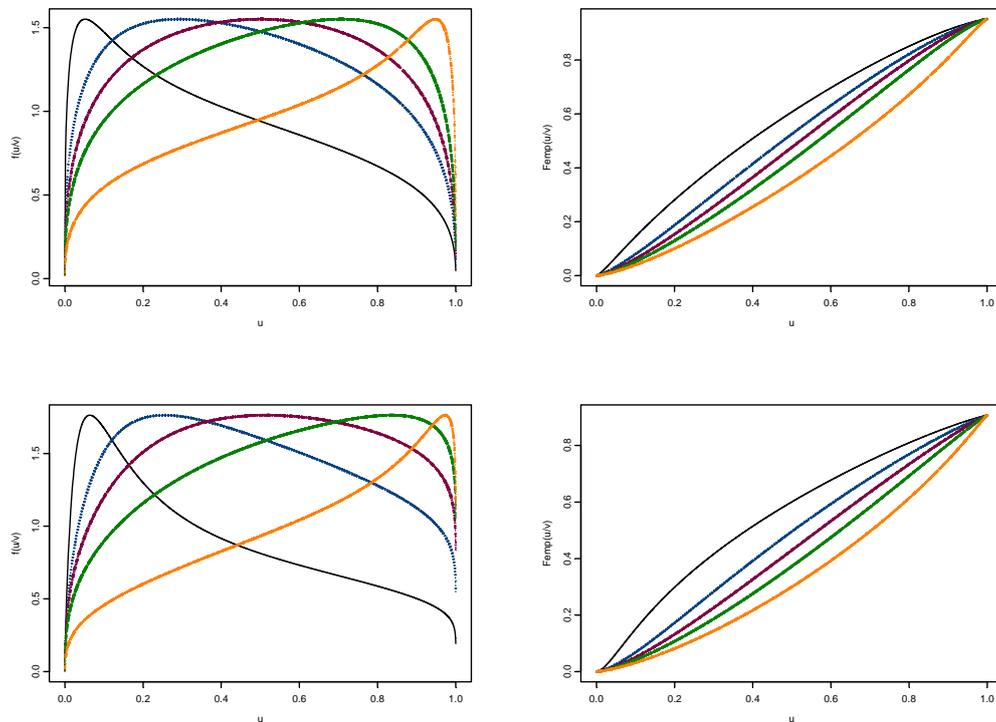


Figura 4.15: *Funciones de densidad condicionadas para distintos valores de v (izquierda) y sus correspondientes funciones de distribución empíricas(derecha). La parte superior corresponde al Modelo 1 y la inferior al Modelo 2*

Mediante el método Montecarlo, simulamos los valores X_{sim} e Y_{sim} , que producirán la muestra aleatoria bidimensional final (x_{sim}^i, y_{sim}^i) , a partir de las GPD ajustadas anteriormente a las distribuciones de pérdidas marginales DowJones e

Ibex35. Esta muestra bidimensional permite obtener los retornos logarítmicos de la cartera según (4.11) y las medidas de riesgo asociadas a las pérdidas correspondientes según (4.12) y (4.13). Los resultados obtenidos para los dos modelos mixtos, junto con los producidos por la cópula normal de parámetro 0.3926 aparecen en la tabla 4.6.

p	VaR_p			ES_p		
	<i>Modelo 1</i>	<i>Modelo 2</i>	<i>Normal</i>	<i>Modelo 1</i>	<i>Modelo 2</i>	<i>Normal</i>
0.98	0.0203	0.0127	0.0215	0.0259	0.0158	0.0286
0.99	0.0252	0.0149	0.0256	0.0294	0.0180	0.0339
0.995	0.0273	0.0170	0.0296	0.0326	0.0198	0.0400
0.999	0.0349	0.0210	0.0451	0.0417	0.0243	0.0582

Tabla 4.6: *Medidas de riesgo para los retornos logarítmicos de la cartera DowJones e Ibex35 simulados según los modelos mixtos y la cópula normal de parámetro 0.3926*

Para obtener los retornos generados a partir de la cópula normal únicamente se necesita generar una muestra a partir de una normal bivariante y, a continuación hacer que las marginales sean uniformes. Una vez generado el vector aleatorio (U, V) que proviene de la cópula normal de parámetro θ , supondremos que cada una de sus componentes procede a su vez de la GPD correspondiente. El algoritmo es el siguiente.

- Simulamos un valor u_1 a partir de una $N(0, 1)$.
- Obtenemos u_2 a partir de una $N(\theta u_1, 1 - \theta^2)$, con θ el parámetro de la cópula normal.
- Entonces

$$U = \Phi(u_1)$$

$$V = \Phi(u_2)$$

con Φ la función de distribución de la normal.

- A continuación suponemos que

$$U \sim GPD\left(\hat{\xi}_{DJ}, \hat{\sigma}_{DJ}\right)$$

$$V \sim GPD\left(\hat{\xi}_{IB}, \hat{\sigma}_{IB}\right)$$

a partir de las cuales simularemos, mediante el método Montecarlo, los valores X_{sim} e Y_{sim} que producirán la muestra aleatoria bidimensional final $(x_{sim}^i, y_{sim}^i)_{i=1\dots n}$.

Utilizando esta muestra bidimensional, generamos la muestra de retornos logarítmicos de la cartera según la expresión (4.11) y calculamos las medidas de riesgo según (4.12) y (4.13).

En la figura 4.16 se pueden apreciar las violaciones de los retornos de la cartera sobre el VaR y el ES al nivel 0.99, simulados a partir de cada uno de los modelos mixtos.

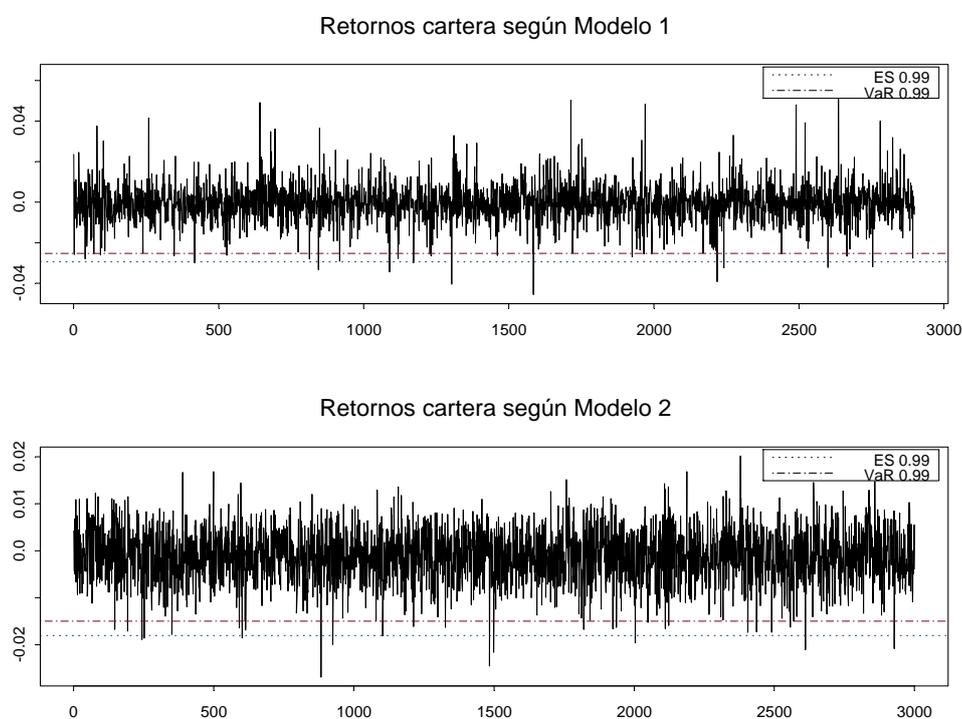


Figura 4.16: *Violaciones al $VaR_{0.99}$ de los retornos simulados según los dos modelos mixtos.*

Si consideramos que los retornos de la cartera proceden de una distribución GPD, podemos realizar la estimación de la misma con los resultados que aparecen en la tabla 4.7. Se ha establecido un valor para el umbral de $u = 0.01$, a la vista de la forma de la función de excesos medios (figura 4.17) y de la sensibilidad de la estimación $\hat{\xi}$ al valor del umbral u (figura 4.18).

	Modelo 1	Modelo 2
u	0.01	0.01
n	323	147
$\hat{\xi} (\hat{se})$	0.01085 (0.06318)	0.00152 (0.0943)
$\hat{\sigma} (\hat{se})$	0.00568 (0.000478)	0.00304 (37×10^{-4})
% datos $> u$	88.85	95.1
$-\max(\log veros.)$	-1343.558	-704.7064

Tabla 4.7: Resultados de la estimación GPD de las colas de pérdidas de los retornos de la cartera DowJones&Ibex35 según los modelos mixtos para $u = 0.01$

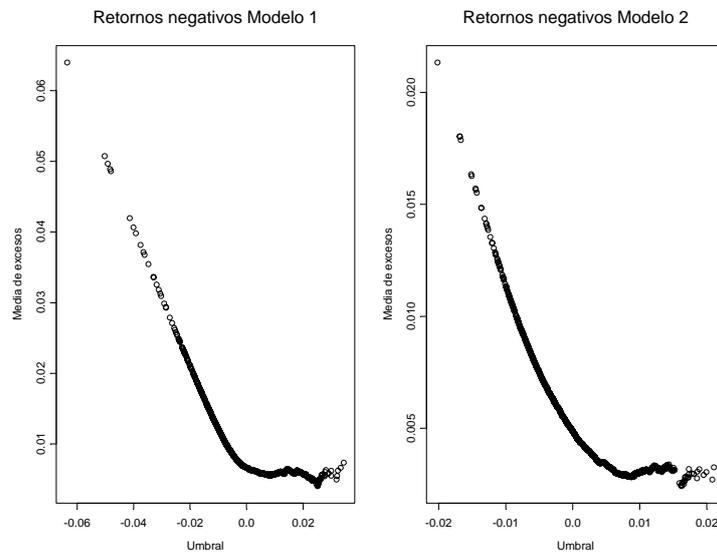


Figura 4.17: Función de excesos medios para las pérdidas de la cartera DowJones&Ibex35 según los dos modelos mixtos

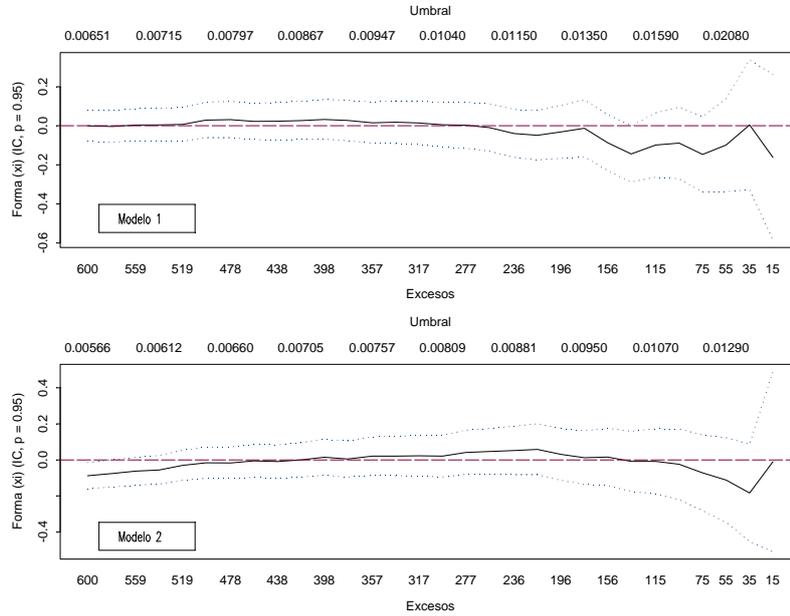


Figura 4.18: *Estimaciones del valor de ξ como función del umbral u para los retornos negativos de la cartera DowJones&Ibex35 generados según los modelos mixtos de cópulas*

Los gráficos 4.19 y 4.20 permiten afirmar que las dos distribuciones se ajustan razonablemente a las colas de pérdidas de los retornos de la cartera generados a partir de los dos modelos mixtos.

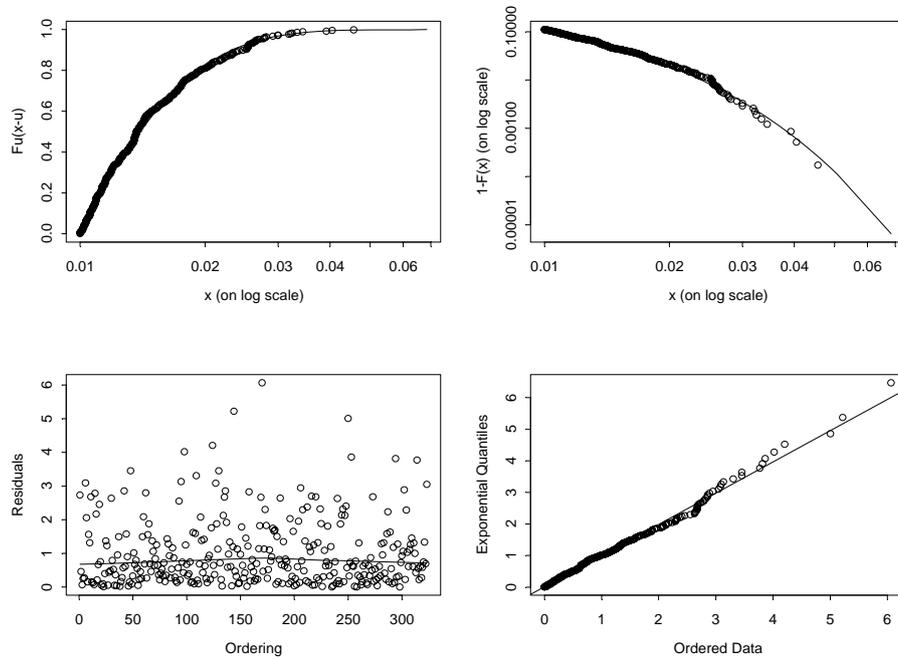


Figura 4.19: Gráficos de diagnóstico para la GPD ajustada a la distribución de pérdidas de la cartera DowJonesIber35 sobre el umbral $u = 0.01$ (Modelo 1)

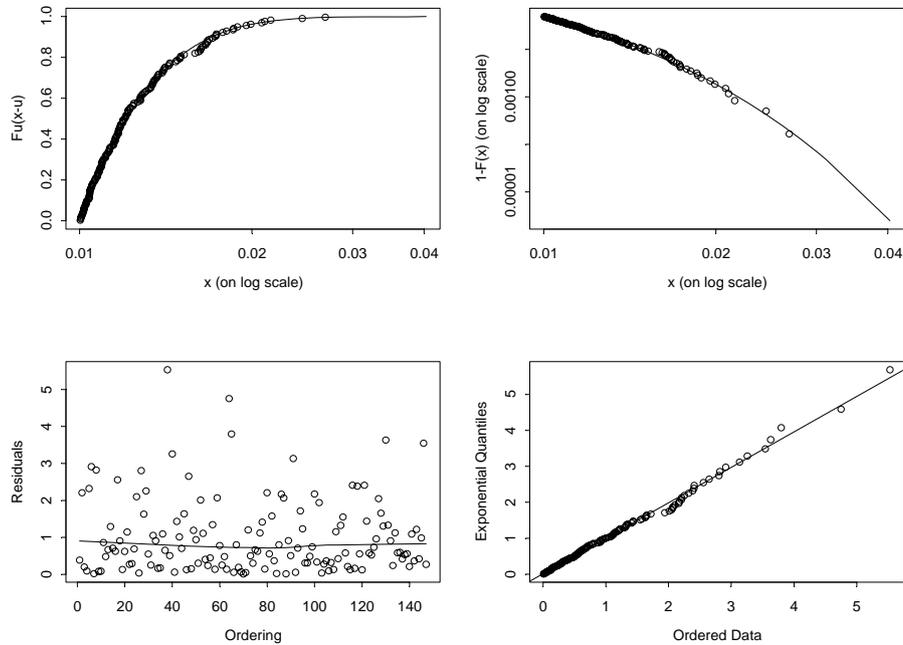


Figura 4.20: Gráficos de diagnóstico para la GPD ajustada a la distribución de pérdidas de la cartera DowJones e Ibex35 sobre el umbral $u = 0.01$ (Modelo 2)

A partir de estas distribuciones GPD ajustadas para cada modelo, se han obtenido las estimaciones puntuales de las medidas de riesgo junto con sus intervalos de confianza al 95%. En la tabla 4.8 aparecen los correspondientes a las pérdidas de los retornos de la cartera simulados según el Modelo 1 y en la 4.9 los referentes a las pérdidas simuladas según el Modelo 2. En ambos casos se observan valores muy próximos a los obtenidos directamente a partir de la serie de retornos de la cartera (tabla 4.6).

p	\mathbf{VaR}_p	$\mathbf{IC}_{\mathbf{VaR}_p;0.95}$	\mathbf{ES}_p	$\mathbf{IC}_{\mathbf{ES}_p;0.95}$	$\frac{\mathbf{ES}_p}{\mathbf{VaR}_p} \rightarrow 1,01$
0.98	0.0198	[0.0188; 0.0210]	0.025	[0.0240; 0.0283]	1.2626
0.99	0.0238	[0.0224; 0.0256]	0.0297	[0.0273; 0.0339]	1.2478
0.995	0.0279	[0.0260; 0.0308]	0.0338	[0.0306; 0.0402]	1.2114
0.999	0.0374	[0.0334; 0.0454]	0.0435	[0.0372; 0.0574]	1.1631

Tabla 4.8: *Estimaciones puntuales e intervalos de confianza al 95% para las medidas de riesgo de la cartera según una GPD* ($\hat{\xi} = 0.01085$; $\hat{\sigma} = 0.00568$) (Modelo 1)

p	\mathbf{VaR}_p	$\mathbf{IC}_{\mathbf{VaR}_p;0.95}$	\mathbf{ES}_p	$\mathbf{IC}_{\mathbf{ES}_p;0.95}$	$\frac{\mathbf{ES}_p}{\mathbf{VaR}_p} \rightarrow 1,001$
0.98	0.0127	[0.0122; 0.0132]	0.0157	[0.0149; 0.0169]	1.2362
0.99	0.0148	[0.0140; 0.0157]	0.0178	[0.0167; 0.0199]	1.2027
0.995	0.0169	[0.0159; 0.0183]	0.0200	[0.0184; 0.0233]	1.1834
0.999	0.0218	[0.0198; 0.0261]	0.0249	[0.0217; 0.0334]	1.1422

Tabla 4.9: *Estimaciones puntuales e intervalos de confianza al 95% para las medidas de riesgo de la cartera según una GPD* ($\hat{\xi} = 0.00152$; $\hat{\sigma} = 0.00304$) (Modelo 2)

Se puede observar que la razón $\frac{\mathbf{ES}_p}{\mathbf{VaR}_p}$ converge a valores próximos a 1 en los dos modelos, lo cual era de esperar dado que las estimaciones del parámetro de forma en la distribuciones GPD ajustadas toma valores próximos a cero en ambos casos ($\hat{\xi}_{Mod1} = 0.01085$, $\hat{\xi}_{Mod2} = 0.00568$). Recordemos también, que la cópula gaussiana no era una mala opción en términos de AIC (tabla 2.4) aunque no contemplaba la dependencia en colas presente en nuestra muestra bivalente. De ahí, que si comparamos las estimaciones de las medidas de riesgo calculadas para los retornos procedentes de los modelos mixtos (tablas 4.8 y 4.9) con las obtenidas a partir de la cópula gaussiana (tabla 4.6), en todos los casos esta última proporciona valores superiores para todos los niveles de probabilidad considerados.

Cóputas con Estructura Dinámica

Parece lógico que en esa búsqueda del patrón de dependencia subyacente entre las series de retornos analizadas, aparezca como elemento sustancial la información proporcionada por observaciones pasadas, incorporada no únicamente en los comportamientos marginales sino también en el conjunto. La influencia de esta información pasada, en general, no tiene por qué ser del mismo tipo e intensidad en los modelos marginales que en el considerado para capturar la dependencia multivariante. Por tanto, la determinación de modelos de cóputas con estructura dinámica resulta relevante en nuestro estudio, y a este fin vamos a dedicar el presente capítulo.

Las formulaciones tipo GARCH, empleadas en el Capítulo 3 para filtrar la dependencia temporal marginal y obtener así una muestra bivariante en la que se mantiene la dependencia contemporánea, proponen modelos de retornos futuros de un determinado índice financiero conociendo su historia, pero no resuelven el problema de predecir ese índice conociendo no sólo sus valores anteriores sino también los de otros que puedan estar relacionados con él.

Hasta ahora hemos supuesto que el parámetro de dependencia de la cóputa que modeliza el comportamiento conjunto posee un valor constante, o dos valores constantes en el caso de los modelos mixtos, lo que presupone que la dependencia entre los dos procesos marginales permanece estática a lo largo del periodo muestral considerado. Sin embargo, resulta más realista contemplar la posibilidad

de que esta relación de dependencia conjunta, expresada a través del modelo de cópula seleccionado como el más adecuado, posea una estructura que cambie a lo largo del tiempo, dado que en el caso de los mercados financieros existe una clara componente dinámica que afecta a su evolución. Podemos encontrar diversos autores que han desarrollado sus trabajos en esta línea de investigación: Patton (2001a, 2001b, 2003) introdujo el concepto de cópula condicional, extendido recientemente por Fermanian y Wegkamp (2004), mientras que la aproximación de Van der Goorberg, Genest y Werker (2005) consiste en proponer un modelo de evolución temporal para el parámetro de dependencia de la cópula a través de la medida de asociación τ de Kendall, relacionada con aquel por la expresión ya conocida

$$\tau(\theta) = 4EC_{\theta}(U, V) - 1$$

en la que (U, V) se distribuye según C_{θ} , y la esperanza se toma respecto a U y V .

Puesto que a lo largo de este estudio hemos prestado una especial atención a la dependencia bivalente localizada en colas, trataremos de dotar al modelo paramétrico de cópula elegido de una estructura dinámica, relacionando su parámetro identificativo con alguna medida de dependencia local, que establezca dicha dependencia como una propiedad de la cópula misma. Un candidato idóneo para esta tarea resulta ser el *coeficiente de dependencia en colas (TDC)*, definido en la Sección 1.8 del Capítulo 1, ya utilizado anteriormente a lo largo de diversos desarrollos, para el que se propondrá un modelo de evolución temporal adecuado.

Sean $r_{i,t+1}$ los retornos correspondientes al índice i en los periodos $t + 1$ y t , respectivamente, y sea

$$G_t = \sigma((r_{1,s}, r_{2,s}) : s \leq t)$$

toda la información disponible sobre los retornos de ambos índices hasta el instante t .

Suponemos que las distribuciones marginales de los retornos de cada índice vienen dadas por procesos GARCH del tipo de los ya especificados en el Capítulo 3, y que ambas distribuciones marginales están ligadas por el modelo paramétrico mixto $C_{\theta}(U, V)$, con $\theta = \{\theta_1, \theta_2, \theta_3\}$.

Recordemos las expresiones del coeficiente de dependencia en colas, en función de los parámetros de dependencia de la cópula, obtenidas para el modelo paramétrico

mixto dado por miembros de la familia Gumbel y Gumbel supervivencia:

$$\lambda_L = (1 - \theta_3)(2 - 2^{1/\theta_2}) \tag{5.1}$$

$$\lambda_U = \theta_3(2 - 2^{1/\theta_1}) \tag{5.2}$$

Los parámetros θ_1 y θ_2 reflejan el grado de dependencia del modelo mixto y θ_3 podría considerarse un parámetro de forma, que regula la estructura de dependencia mediante la asignación de pesos a los modelos integrantes de la composición.

La primera cuestión a resolver es la obtención de las series de valores que reflejen la evolución temporal de la dependencia en ambas colas. Sean λ_U y λ_L , los coeficientes de dependencia en la cola superior e inferior respectivamente, asociados a nuestro modelo. Las series de valores $\{\widehat{\lambda}_{U,n}^t\}_{t \in T}$ y $\{\widehat{\lambda}_{L,n}^t\}_{t \in T}$ que recogen la evolución temporal de ambos coeficientes, se obtienen a partir de los estimadores no paramétricos propuestos por Schmidt y Stadtmüller (2003), basados en la definición de *cópula cola* y cuya construcción se detalla a continuación.

Definición 5.1 (*Cópula cola*)

Sea F una función de distribución n -dimensional. Si para los subconjuntos $I, J \in \{1, \dots, n\}, I \cap J = \emptyset$, existe el siguiente límite en $\overline{\mathbb{R}}_+^n$

$$\Lambda_U^{I,J}(x) = \lim_{t \rightarrow \infty} P \left\{ X_i > F_i^{-1} \left(1 - \frac{x_i}{t} \right), \forall i \in I \mid X_j > F_j^{-1} \left(1 - \frac{x_j}{t} \right) \forall j \in J \right\}$$

entonces la función $\Lambda_U^{I,J} : \overline{\mathbb{R}}_+^n \rightarrow \mathbb{R}$ es la llamada *cópula cola superior* asociada a F para $I; J$. La correspondiente *cópula cola inferior* se define como

$$\Lambda_L^{I,J}(x) = \lim_{t \rightarrow \infty} P \left\{ X_i \leq F_i^{-1} \left(\frac{x_i}{t} \right), \forall i \in I \mid X_j \leq F_j^{-1} \left(\frac{x_j}{t} \right) \forall j \in J \right\}$$

siempre que dicho límite exista.

$F_1^{-1}, \dots, F_n^{-1}$ denotan las correspondientes inversas generalizadas de las funciones de distribución marginales F_1, \dots, F_n y $\overline{\mathbb{R}}_+^n := [0, \infty]^n \setminus \{(\infty, \dots, \infty)\}$.

Mientras que las funciones *cópula* representan la estructura de dependencia de un vector aleatorio multivariante, las funciones *cópula cola* describen la estructura de dependencia de distribuciones multi-dimensionales en las colas.

En un entorno bivariante, consideraremos un vector aleatorio $(X, Y)'$ con función de distribución F y funciones de distribución marginales continuas G y

H. La estimación resulta más fácil si se utiliza la siguiente modificación de la cópula cola:

$$\begin{aligned}\Lambda_U(x, y) &= y \cdot \Lambda_U^{\{1\}, \{2\}}(x, y) \\ \Lambda_L(x, y) &= y \cdot \Lambda_L^{\{1\}, \{2\}}(x, y) \quad x \in \overline{\mathbb{R}}_+, y \in \mathbb{R}_+\end{aligned}$$

donde los índices $\{1\}$ y $\{2\}$ pueden obviarse. Tendremos que $\Lambda_U(x, \infty) = x$ y $\Lambda_L(x, \infty) = x$ para todo $x \in \mathbb{R}_+$.

La siguiente definición establece el ya conocido concepto de dependencia en colas en términos de la cópula cola.

Definición 5.2 (*Dependencia en colas*)

El vector aleatorio bivalente $(X, Y)'$ se considera dependiente en la cola superior si existe $\Lambda_U(1, 1)$ y

$$\lambda_U = \Lambda_U(1, 1) = \lim_{v \rightarrow 1^-} P \{X > G^{-1}(v) / Y > H^{-1}(v)\} > 0$$

Si $\lambda_U = 0$ se dice que $(X, Y)'$ es independiente en la cola superior. λ_U es el coeficiente de dependencia en la cola superior. Análogamente, se define el coeficiente de dependencia en la cola inferior como

$$\lambda_L = \Lambda_L(1, 1)$$

si el límite existe. Diremos que existe dependencia en la cola inferior si $\lambda_L > 0$.

La cópula cola y los coeficientes de dependencia en colas no dependen de las marginales sino únicamente de la cópula, por ejemplo

$$\lambda_U = \lim_{v \rightarrow 1^-} \frac{1 - 2v + C(v, v)}{1 - v} \quad \text{y} \quad \Lambda_U(x, y) = x + y - \lim_{t \rightarrow \infty} t \left(1 - C\left(1 - \frac{x}{t}, 1 - \frac{y}{t}\right)\right)$$

Sea C_n la cópula empírica definida por

$$C_n(u, v) = F_n(G_n^{-1}(u), H_n^{-1}(v)) \quad (u, v)' \in [0, 1]^2$$

con F_n , G_n y H_n las funciones de distribución empíricas de F , G y H y n el tamaño de la muestra bivalente considerada. $C_n((a, b] \times (c, d])$ es la medida sobre el intervalo $(a, b] \times (c, d] \subset [-\infty, \infty]^2$ dada por

$$C_n((a, b] \times (c, d]) = C_n\left(\left((a, b] \times (c, d]) \cap [0, 1]^2\right)\right)$$

Como ya se ha comentado con anterioridad, la elección de las funciones de distribución empíricas para modelizar las marginales evita errores debidos a especificaciones equivocadas en los ajustes paramétricos de dichos modelos. Un estudio empírico de las consecuencias producidas por malas especificaciones e interpretaciones erróneas de la correspondiente estructura de dependencia en valores extremos puede consultarse en Frahm, Junker y Schmidt (2006).

Definición 5.3 (*Cópula cola empírica*)

Sean $R_{1n}^{(j)}$ y $R_{2n}^{(j)}$ los rangos de $X^{(j)}$ e $Y^{(j)}$ $j = 1, \dots, n$, respectivamente. Se define la cópula cola empírica superior como

$$\begin{aligned} \widehat{\Lambda}_{U,n}(x, y) &= \frac{n}{k} C_n \left(\left(1 - \frac{kx}{n}, 1 \right] \times \left(1 - \frac{ky}{n}, 1 \right] \right) \\ &\approx \frac{1}{k} \sum_{j=1}^n I_{\{R_{1n}^{(j)} > n-kx \text{ y } R_{2n}^{(j)} > n-ky\}} \end{aligned} \tag{5.3}$$

Análogamente, se define la cópula cola empírica inferior como

$$\begin{aligned} \widehat{\Lambda}_{L,n}(x, y) &= \frac{n}{k} C_n \left(\frac{kx}{n}, \frac{ky}{n} \right) \\ &\approx \frac{1}{k} \sum_{j=1}^n I_{\{R_{1n}^{(j)} \leq kx \text{ y } R_{2n}^{(j)} \leq ky\}} \end{aligned} \tag{5.4}$$

con $k \in \{1, \dots, n\}$, $k = k(n) \rightarrow \infty$ y $k/n \rightarrow 0$ cuando $n \rightarrow \infty$.

Las definiciones anteriores están basadas en la siguiente representación de la cópula empírica propuesta por Genest, Ghoudi y Rivest (1995):

$$\widetilde{C}_n(u, v) = \frac{1}{n} \sum_{j=1}^n I_{\{G_n(X^{(j)}) \leq u \text{ y } H_n(Y^{(j)}) \leq v\}} \quad (u, v) \in [0, 1]^2$$

Schmidt y Stadtmüller (2003) proponen los siguientes estimadores no paramétricos del coeficiente de dependencia en colas, basados en las definiciones anteriores de las cópulas cola empíricas:

$$\widehat{\lambda}_{U,n} = \widehat{\Lambda}_{U,n}(1, 1) \tag{5.5}$$

$$\widehat{\lambda}_{L,n} = \widehat{\Lambda}_{L,n}(1, 1) \tag{5.6}$$

Ambos autores demuestran que, bajo ciertas condiciones de regularidad,

$$\begin{aligned}\sqrt{k} \left\{ \widehat{\lambda}_{U,n} - \lambda_U \right\} &\rightarrow dN(0, \sigma_U^2) \\ \sqrt{k} \left\{ \widehat{\lambda}_{L,n} - \lambda_L \right\} &\rightarrow dN(0, \sigma_L^2)\end{aligned}$$

con $N(0, \sigma_U^2)$ y $N(0, \sigma_L^2)$ variables aleatorias centradas distribuidas normalmente con varianzas

$$\begin{aligned}\sigma_U^2 &= \lambda_U + \left(\frac{\partial}{\partial x} \Lambda_U(1,1) \right)^2 + \left(\frac{\partial}{\partial y} \Lambda_U(1,1) \right)^2 \\ &\quad + 2\lambda_U \left(\left(\frac{\partial}{\partial x} \Lambda_U(1,1) - 1 \right) \left(\frac{\partial}{\partial y} \Lambda_U(1,1) - 1 \right) - 1 \right)\end{aligned}$$

y

$$\begin{aligned}\sigma_L^2 &= \lambda_L + \left(\frac{\partial}{\partial x} \Lambda_L(1,1) \right)^2 + \left(\frac{\partial}{\partial y} \Lambda_L(1,1) \right)^2 \\ &\quad + 2\lambda_L \left(\left(\frac{\partial}{\partial x} \Lambda_L(1,1) - 1 \right) \left(\frac{\partial}{\partial y} \Lambda_L(1,1) - 1 \right) - 1 \right)\end{aligned}$$

Para más detalles sobre propiedades de convergencia, normalidad y consistencia se puede consultar Schmidt y Stadtmüller (2003).

Consideraremos ventanas temporales de un cuatrimestre de negocio (periodos de 81 días) cada uno de las cuales comprende los 40 días anteriores al día t considerado, los 40 siguientes y el propio día t . Para cada uno de esos periodos, se calcula la estimación empírica $\widehat{\lambda}_{L,n}^t$, $n = 81$, según la expresión (5.6).

Puesto que las estimaciones del VaR son altamente sensibles al comportamiento de las colas y, en concreto, a la dependencia en colas presente en la distribución de retornos de la cartera, la elección del umbral k vendrá determinada por el nivel p del VaR correspondiente. Recordemos que el cálculo del coeficiente de dependencia en la cola inferior es equivalente a:

$$\lambda_L = \lim_{p \rightarrow 1^-} P \{ Y < VaR_{1-p}(Y) / X < VaR_{1-p}(X) \} > 0$$

Mediante un procedimiento análogo efectuado sobre el coeficiente de dependencia en la cola superior

$$\lambda_U = \lim_{p \rightarrow 0^+} P \{ Y > VaR_{1-p}(Y) / X > VaR_{1-p}(X) \} > 0$$

se obtienen las estimaciones empíricas $\widehat{\lambda}_{U,n}^t$, $n = 81$, según la expresión (5.5).

Una vez obtenidas las series de estimaciones empíricas de los coeficientes de dependencia, en la cola superior e inferior, en cada periodo temporal, se propone un modelo de serie temporal para cada una de ellas en el cual la influencia de valores cercanos al instante t considerado sea superior a la de los más alejados en el tiempo. La especificación se corresponde con la de un *modelo de alisado exponencial doble de Holt* (Gardner y Dannenbring 1980, Gardner 1985), adecuado para series que presentan una tendencia de tipo lineal en su evolución, cuya especificación es la siguiente:

$$\begin{aligned}\widehat{\lambda}_{t+k} &= a_t + b_t k + \epsilon_t \\ a_t &= \alpha \widehat{\lambda}_t + (1 - \alpha)(a_{t-1} + b_{t-1}) \quad 0 < \alpha \leq 1\end{aligned}\quad (5.7)$$

$$b_t = \beta(a_t - a_{t-1}) + (1 - \beta)b_{t-1} \quad 0 < \beta \leq 1 \quad (5.8)$$

La expresión (5.7) corresponde a la ecuación de nivel y la (5.8) a la de la pendiente.

Finalmente, es posible incorporar una estructura dinámica al modelo de cópula a través de las relaciones establecidas en (5.1) y (5.2), que en cada instante t permiten determinar, a partir de los valores $\lambda_{L,t}$ y $\lambda_{U,t}$, los parámetros θ_1 y θ_2 . El parámetro θ_3 permanece constante, según el valor obtenido en la estimación de máxima verosimilitud realizada en el Capítulo 2, y la evolución temporal afectará a los valores de los parámetros θ_1 y θ_2 que son los que gobiernan el grado de dependencia del modelo mixto.

5.1. Modelo de cópula dinámica DowJones&Ibex35

Las series de estimaciones empíricas de los coeficientes de dependencia en cada cuatrimestre considerado, obtenidas según las expresiones (5.5) y (5.6), respectivamente, aparecen representadas en la figura 5.1. El umbral k utilizado se ha determinado en función del nivel de probabilidad asociado al VaR, considerando un valor que garantice la inclusión de un 15% de observaciones por la izquierda (TDC inferior) o por la derecha (TDC superior). A continuación se ha procedido a realizar un doble suavizado exponencial de Holt sobre cada serie. En cada una de ellas el valor seleccionado del parámetro es $\alpha = 0,1$, resultante al minimizar el error cuadrático medio del ajuste (figura 5.2).

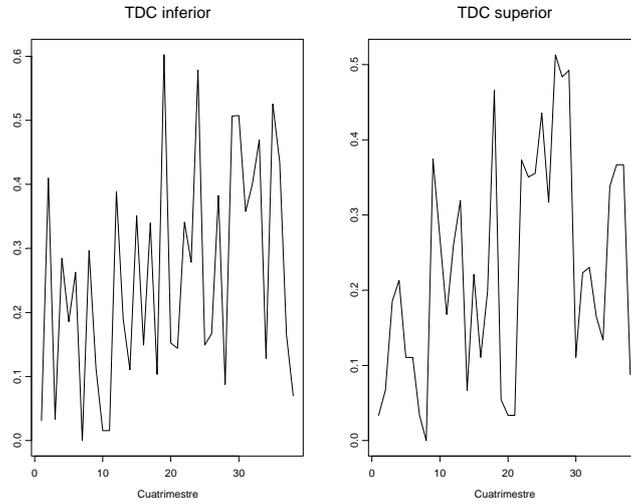


Figura 5.1: Estimaciones empíricas de los TDC en cada cuatrimestre obtenidas a partir de la muestra bivalente filtrada (DowJones,Ibex35)

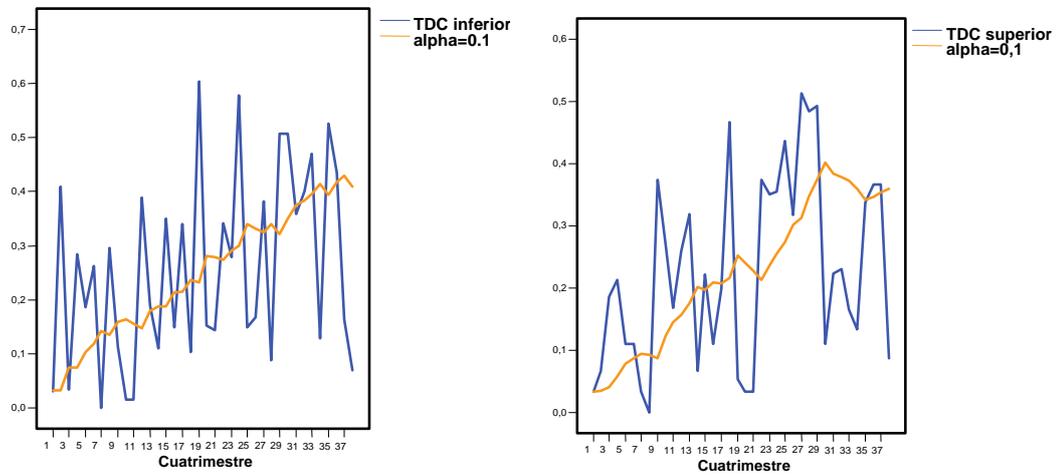


Figura 5.2: Serie de estimaciones empíricas de los TDC y los correspondientes suavizados de Holt para $\alpha = 0,1$.

En base a los modelos ajustados se han realizado predicciones de ambos coeficientes para los próximos 4 años (12 cuatrimestres), que han permitido determinar el valor de los parámetros de dependencia de cada modelo mixto de cópula, en base a las expresiones siguientes, obtenidas a partir de (5.1) y (5.2):

$$\begin{aligned}\widehat{\theta}_{1,t} &= \log(2) \Big/ \log\left(2 - \frac{\widehat{\lambda}_{U,t}}{\widehat{\theta}_3}\right) \\ \widehat{\theta}_{2,t} &= \log(2) \Big/ \log\left(2 - \frac{\widehat{\lambda}_{L,t}}{(1 - \widehat{\theta}_3)}\right) \\ \widehat{\theta}_3 &= 0.497\end{aligned}$$

La evolución temporal de los parámetros de dependencia del modelo mixto puede observarse en la figura 5.3. En ella se aprecia que el parámetro θ_2 tiene un crecimiento más acusado. Recordemos que este parámetro determina el grado de dependencia de la cópula Gumbel supervivencia, que regula la dependencia entre las pérdidas (cola inferior) de la muestra bivalente.

Se han efectuado simulaciones de 12 muestras de tamaño 2000 para cada par (θ_1, θ_2) del Modelo 1 con $\theta_3 = 0.497$, que corresponden a cuatrimestres consecutivos de los cuatro años siguientes al 28 de Marzo de 2006. Para cada una de ellas se han generado los retornos de la cartera integrada por los índices Dow Jones e Ibex35, con un peso de 0.5 para ambos índices, considerando cada marginal una GPD de parámetros los especificados para cada índice en el Capítulo 4 (tabla 4.1).

A partir de estos retornos se han determinado las medidas de riesgo correspondientes, VaR y ES, para niveles de probabilidad $p = (0.98, 0.99, 0.995, 0.999)$. La evolución temporal del VaR, calculado para cada cuatrimestre considerado, puede apreciarse en el gráfico 5.4. En él se observa que a medida que aumenta el cuantil calculado (la pérdida asociada al valor en riesgo es mayor y, por tanto más peligrosa) el Var obtenido para el modelo de cópula estático está por debajo de los valores en riesgo generados por el modelo dinámico, lo que conduce a pensar que el riesgo estaría siendo subvalorado a través del modelo estático. Por otra parte, la posibilidad de calcular valores en riesgo para periodos temporales consecutivos (en nuestro caso cuatrimestres), permite prevenir situaciones de riesgo extremo a lo largo de un horizonte temporal fijado previamente: la aparición de picos o cumbres en la serie de predicciones señala con antelación periodos especialmente

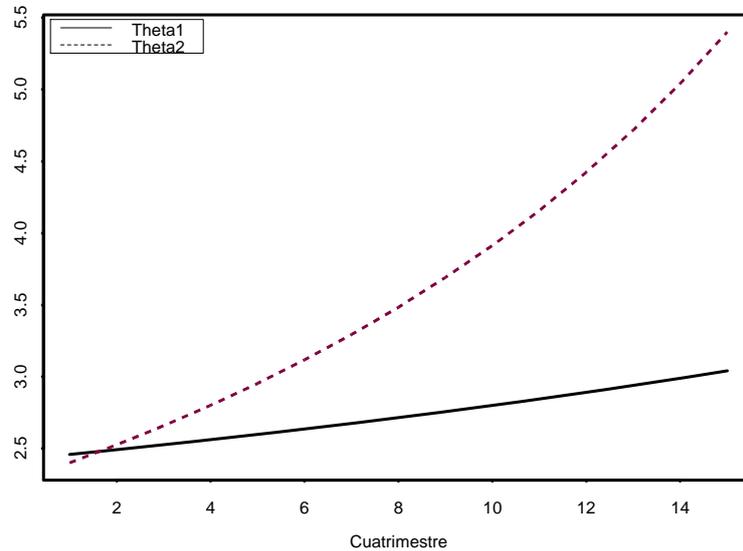


Figura 5.3: *Evolución temporal de los parámetros θ_1 y θ_2 del modelo mixto Gumbel-Gumbel supervivencia*

peligrosos, lo que facilita la aplicación de medidas preventivas, en el caso de que sea posible, o cuanto menos la posibilidad de estar sobre aviso. El gráfico 5.4 señala como especialmente peligroso el cuarto cuatrimestre a partir de la fecha considerada, en todos los niveles de probabilidad.

Finalmente, en cuanto al valor del ES (la otra medida de riesgo en consideración) el gráfico 5.5 revela un comportamiento similar, e incluso más acusado, que en el caso VaR. Se puede apreciar que en todos los niveles de probabilidad considerados, la pérdida esperada obtenida sobre el modelo estático de cópula infravalora, en la mayoría de los cuatrimestres considerados, las indicadas por el modelo dinámico. De nuevo, los cuatrimestres 4 y 5 aparecen como especialmente peligrosos.

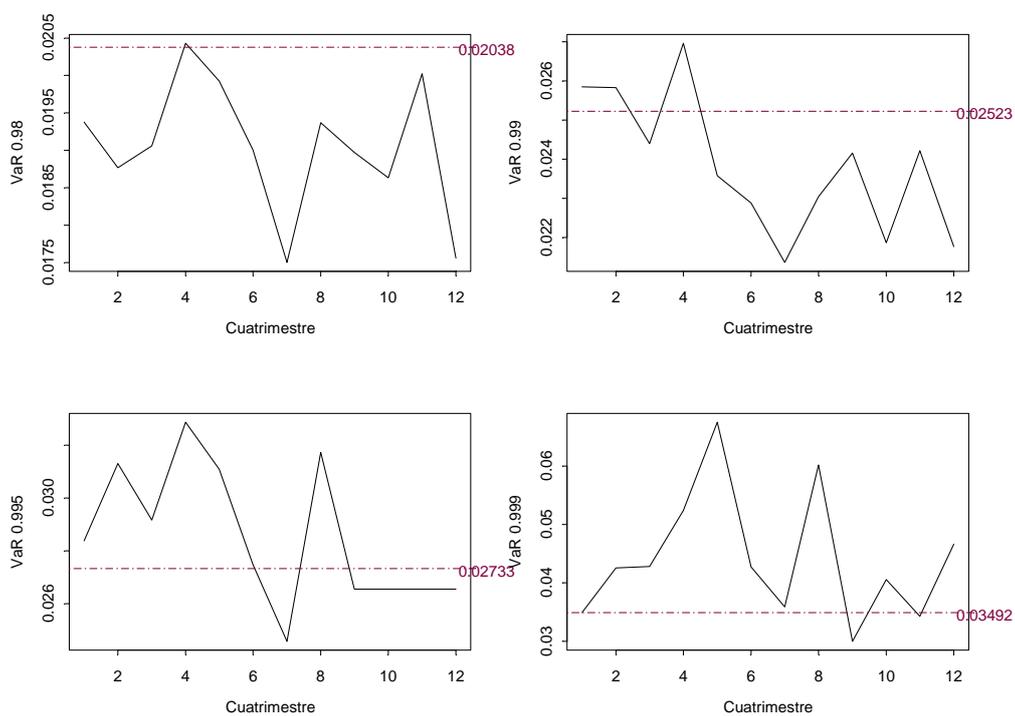


Figura 5.4: Predicciones del VaR a un horizonte de 4 años para diferentes niveles de probabilidad, según el Modelo 1 con estructura dinámica (la línea discontinua corresponde al VaR obtenido a partir del modelo estático)

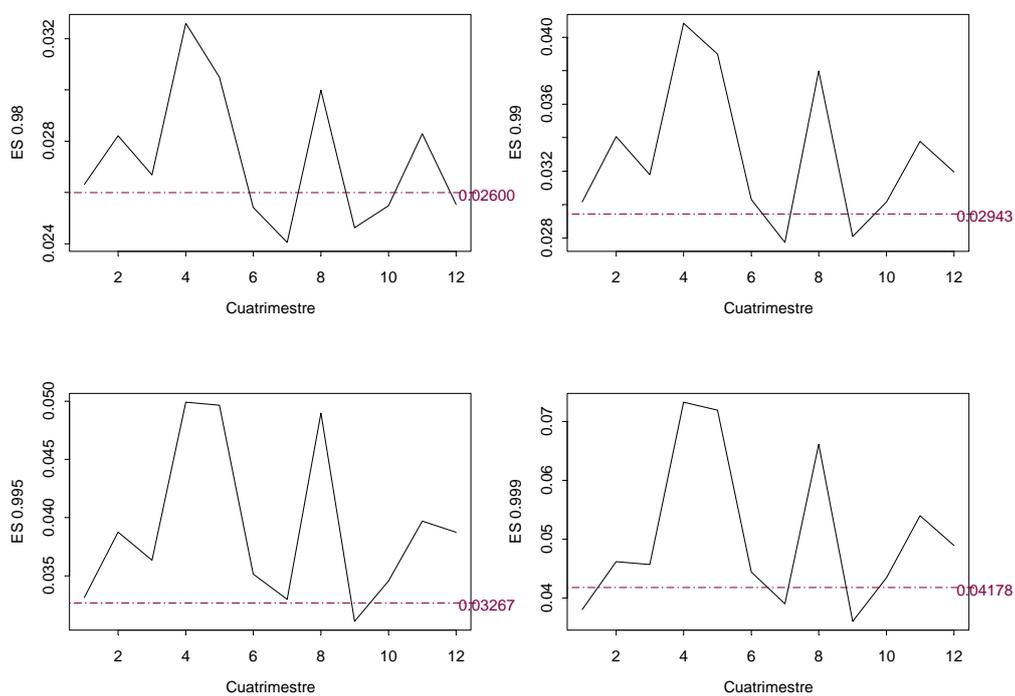


Figura 5.5: Predicciones del ES a un horizonte de 4 años según el Modelo 1 para diferentes niveles de probabilidad, con estructura dinámica (la línea discontinua corresponde al ES obtenido a partir del modelo estático)

6

Conclusiones y Problemas abiertos

Considerando nuestro trabajo en su totalidad y con carácter general, creemos que es posible afirmar que la obtención de modelos basados en cópulas constituye una aproximación interesante, flexible y eficiente en la representación de estructuras de dependencia de series financieras bivariantes, que supera las carencias y malas especificaciones inherentes a aquellas otras que parten de supuestos gaussianos y tratan únicamente el aspecto lineal de la correlación.

En cada uno de los capítulos se han tratado y resuelto de forma individualizada las diferentes cuestiones que surgen en un planteamiento general del desafío que constituye la elaboración de una metodología estadística sólida que permita abordar con garantías de éxito la valoración y control de riesgos de mercados financieros.

Dado que un modelo multivariante presenta dos componentes: la univariante o marginal, que caracteriza cada una de las variables, y la estructura de dependencia existente entre esas variables marginales, resulta posible plasmar de forma separada estos dos aspectos del problema en términos matemáticos con el fin de obtener los modelos apropiados. Esto es lo que se ha hecho en el capítulo 1, de forma genérica, conectando las funciones cópulas con la noción de variable aleatoria. Los conceptos, propiedades y resultados establecidos en él constituyen el fundamento teórico de partida que nos ha permitido acometer el resto de cuestiones.

Hemos iniciado nuestro estudio con una propuesta de métodos inferenciales que faciliten la selección del modelo de cópula más adecuado a nuestra muestra bivalente, prestando especial atención a la información bruta extraída de la muestra, proporcionada por las funciones de distribución empíricas univariantes, tratadas convenientemente, ligadas por su cópula empírica. Posteriormente se utilizan métodos paramétricos para realizar el ajuste de diferentes modelos basados en cópulas, y se indica como evaluar la bondad del ajuste a través de métodos gráficos y analíticos. Es aquí donde se perfilan como modelos más eficientes aquellos que combinan diferentes familias de cópulas, puesto que son capaces de cubrir de una forma más precisa y flexible diferentes grados y estructuras de dependencia.

También se ha resuelto la dificultad que presenta la presencia de dependencia temporal en las muestras univariantes. Después de evaluar diferentes alternativas, hemos optado por realizar un filtrado de dicha dependencia marginal a través de procesos tipo GARCH con el fin de obtener marginales aproximadamente i.i.d con colas pesadas, manteniendo la relación de dependencia contemporánea. En el capítulo 3 se propone una metodología de ajuste para series financieras basada en este tipo de procesos.

Una vez superadas las dificultades anteriores estamos ya en disposición de abordar el núcleo de nuestro problema: la representación de la relación existente entre pérdidas de índices financieros, con el fin de valorar el riesgo potencial de carteras de inversión construidas sobre ellos. A lo largo del capítulo 4 se trata exhaustivamente esta cuestión, proporcionando herramientas que permitan tratar el problema en todas sus facetas. Las distribuciones GPD resultan de gran utilidad en la modelización de las pérdidas univariantes, proporcionando buenos resultados en la obtención de cuantiles altos.

Por último, nos planteamos la posibilidad de dotar de una estructura dinámica al modelo de cópula seleccionado. Puesto que nuestro interés fundamental se centra sobre el fenómeno de dependencia en colas, incorporamos la información proporcionada por observaciones pasadas a este respecto no sólo en las distribuciones marginales (procesos GARCH) sino también en la conjunta. Los candidatos ideales para esta tarea resultan ser los coeficientes de dependencia en colas (TDC), superior e inferior.

En el camino recorrido desde el inicio de nuestro estudio hasta su elaboración final, nos hemos encontrado con nuevos retos que han llamado nuestra atención

y a los que esperamos poder dedicar tiempo y esfuerzo en el futuro tales como:

- La utilización de modelos mixtos construidos como combinaciones lineales convexas de familias de cópulas se perfila como una herramienta potente para capturar diferentes estructuras y grados de dependencia. A la flexibilidad característica de los modelos basados en cópulas, que permiten ligar diferentes modelos marginales con diferentes estructuras bivariantes en un número infinito de formas, se añade la proporcionada por los pesos de la combinación lineal que permiten dar mayor importancia a la dependencia de uno u otro signo o a la ausencia de dependencia en colas. Resultaría interesante seguir analizando estructuras distintas a las ya utilizadas en esta tesis.

- La posibilidad de dotar a los modelos de cópula de una estructura que refleje la evolución temporal conjunta de observaciones pasadas, también resulta una línea de trabajo interesante, en su doble faceta de obtener aproximaciones alternativas a la propuesta en esta tesis, que permitan obtener modelos dinámicos de cópulas, y en la de explorar diferentes usos de los mismos.

- Por último, quedaría pendiente la tarea de generalizar los resultados obtenidos al caso multidimensional.

Para finalizar, únicamente señalar que para la realización de los diferentes procedimientos empleados a lo largo de los desarrollos se ha utilizado el paquete estadístico Splus 6.1 en su versión para Windows, especialmente el módulo S+FinMetrics en su versión 2.0.1 que incorpora funciones específicas de tratamiento de series financieras y la librería MASS realizada por Venables, W. y Ripley, B. versión 7.0.

Bibliografía

- [1] Alexander, C. (2001): *Market Models: A Guide to Financial Data Analysis*. John Wiley and Sons.
- [2] Artzner, P.; Delbaen, F.; Eber, J. y Heath, D. (1999): *Coherent Measures of Risk*. *Mathematical Finance* 9, 203-288.
- [3] Abramowitz, M. y Stegun, I. (1970): *Handbook of Mathematical Functions*. Dover.
- [4] Akaike, H. (1973): *Information Theory and An Extension of the Maximum likelihood Principle*. Petrov, B. N. y Csaki, F. (Eds.). The second International Symposium on Information Theory. Budapest: Akadémiai Kiado.
- [5] Akaike, H. (1977): *On Entropy Maximization Principle*. Krishnaia, P. R. (Ed.). *Applications of Statistics*. Holland: North-Holland. Publishing Co.
- [6] Ang, A. y Bekaert, G. (2002): *International Asset Allocation with Regime Shifts*. *Review of Financial Studies*, 15: 1137-1187.
- [7] Ang, A. y Chen, J. (2002): *Asymmetric Correlations of Equity Portfolios*. *Review of Financial Studies*, 63: 443-494.
- [8] Bank of International Settlements (1996): *Amendment to the Capital Accord to Incorporate Market Risks*. Basle Committee on Banking Supervision.
- [9] Bollerslev, T., Chou, R. y Kroner, K. (1992): *ARCH Modeling in Finance*. *Journal of Econometrics*, 52:5-59.
- [10] Bollerslev, T., Engle, R.F. y Nelson, D. B. (1994): *ARCH Models*. In *Handbooks of Econometrics*. (Eds. Engle, R.F. y McFadden, D. L.). Vol 4, Elsevier Science B.V.

- [11] Box, G. E. y Jenkins, G. M. (1976): *Time Series Analysis Forecasting and Control*. Holden-Day Inc. San Francisco.
- [12] Carmona, R. A. (2004): *Statistical Analysis of Financial Data in S-Plus*. Springer Texts in Statistics. Springer. New York.
- [13] Cherubini, U., Luciano, E. y Vecchiato, W. (2004): *Copula Methods in Finance*. Wiley, New York.
- [14] Crouhy, M., Galai, D. y Mark, R. (2001): *Risk Management*. McGraw-Hill. New York.
- [15] Danielsson, J., De Vries, C. (1997): *Value-at-Risk and extreme Returns*. School of Economics and Political Science. Paper N° 273: 1-33.
- [16] Danielsson, J., Jorgensen, N., De Vries, C. y Phornchanok, J. (2000): *On the (Ir)Relevancy of Value-at-Risk Regulation*. Incluido en *Measuring Risk in Complex Stochastic Systems*. (Frankel, J., Härdle, W., Stahl, G.) Kaiser-slautern and Berlin.
- [17] Deheuvels, P. (1979): *La Fonction de Dependence Empirique et ses Propriétés. Un test non parametrique d'indépendance*. Acad. Roy. Belg. Bull. Cl. Sci (5) 65: 274-292.
- [18] Deheuvels, P. (1981a): *A Kolmogorov-Smirnov type test for Independence and Multivariate Samples*. Rev. Roumaine Math. Pures Appl. 26: 213-226.
- [19] Deheuvels, P. (1981b): *A Non Parametric Test for Independence*. Publ. Inst. Statist. Univ. Paris 26: 29-50.
- [20] Ding, Z., Granger, C. W. J. y Engle, R.F. (1993): *A Long Memory Property of Stock Market of Returns and A New Model*. Journal of Empirical Finance, 1: 83-106.
- [21] Dudewicz, E. J. y Mishra, S. N. (1988): *Modern Mathematical Statistics*. John Wiley & Sons.
- [22] Engle, R. F. (1982): *Autorregressive Conditional Heteroskedasticity With Estimates of the Variance of United Kingdom Inflation*. Econometrica, 50 (4): 987-1007.
- [23] Engle, R. F. y Ng, V. (1993): *Measuring and Testing the Impact of News on Volatility*. Journal of Finance, 48 (5): 1749-1778.

- [24] Engle, R. F.(2000): *ARCH Selected Readings*. Oxford University Press.
- [25] Embrechts, P. ; McNeil, A. ; Straumann, D. (1999): *Correlation and Dependence in Risk Management: Properties and Pitfalls*. Cambridge University Press, 176-223.
- [26] Embrechts, P.; Klüppelberg, C. y Mikosch, T. (1997): *Modelling Extremal Events for Insurance and Finance*. Springer. New York.
- [27] Embrechts, P., Resnick, S. y Samorodnitsky, G. (1998): *Living In the Edge* RISK Magazine, 11 (1): 96-100.
- [28] Embrechts, P., Resnick, S. y Samorodnitsky, G. (1999): *Extreme Value Theory As a Risk Management Tool*. North American Actuarial Journal, 3 (2): 30-41.
- [29] Embrechts, P. ; Lindskog, F. ; McNeil, A. (2001). *Modelling Dependence with Copulas and Applications to Risk Management*. Department of Mathematics ETHZ. Zurich.
- [30] Feller, W. (1966): *An Introduction to Probability Theory and Its Applications*. Vol 2. John Wiley & Sons.
- [31] Fermanian, J. D. y Wegkamp, M. (2004): *Time Dependent Copulas*. Preprint INSEE.
- [32] Fisher, N. I.y Switzer, P. (1985): *Chi-plot for Assessing Dependence*. Biometrika 72 (2):253-265
- [33] Fisher, N. I.y Switzer, P. (2001): *Graphical Assessment of Dependence: Is a Picture Worth 100 Tests?*. the American Statistician, 55: 233-239.
- [34] Fisher, R. A., Tippett, L. H. C. (1928): *Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample*. Proceedings of the Cambridge Philosophical Society. Vol. 24: 180-190.
- [35] Frahm, G., Junker, M., Schmidt, R. (2006): *Estimating of Tail-Dependence Coefficient : Propertie and Pitfalls*. Insurance: Mathematics and Economics 37: 80-100.
- [36] Fréchet, M (1957): *Les Tableaux de Corrélacion Dont Les Marges Sont Données*.Annals Univ. Lyon, Sci., 20:13-31.

- [37] Frees, E.W. y Valdez, E.A. (1998): *Understanding Relationships Using Copulas*. North Amer. Act. J. 2:1-25.
- [38] Frey, R. (1997): *Derivative Asset Analysis in Models With Level-Dependent and Stochastic Volatility*. CWI Quaterly. Amsterdam, 10: 1-34.
- [39] Gardner, E. y Dannenbring, D. G. (1980): *Forecasting With Exponential Smoothing: Some Guideliness for Model selection*. Decision Sciencies. Vol. 11, N° 2: 370-383.
- [40] Gardner, E. (1985): *Exponential Smoothing: The State of Art*. Journal of Forecasting. Vol.4 N° 1: 1-38.
- [41] Genest, C., Mckey, J. (1986): *The Joy of Copulas: Bivariate Distributions With Uniform Marginals*. American Statistics, 40: 280-285.
- [42] Genest, C., Rivest, L.-P. (1993): *Statistical Inference Procedures for Bivariate Archimedian Copulas*. Journal of the American Statistical Association, 88 (423): 1034-1043.
- [43] Genest, C., Ghoudi, K. y Rivest, L.-P. (1995): *A Semiparametric Estimation Procedure of Dependence Parameters in Multivariate Families of Distribution*. Biometrika, 40: 543-552.
- [44] Genest, C., Boies, J.-C. (2003): *Detecting dependence with Kendall plots*. The American Statistician, 57, 275-284.
- [45] Genest, C. y Rémillard, B. (2004): *Tests of Independence and Randomness Based on the Empirical Copula Process*. Sociedad de Estadística e I.O. Test Vol. 13 No. 2: 335-369.
- [46] Genest, C. y Favre, A.-C. (2006): *Everything You Always Wanted Know About Copula Modelling But Were Afraid to Ask*. Journal of Hydrologic Engineering, 12. Próxima publicación.
- [47] Gilli, M. y Këllezzi, E. (2005): *An Application of Extreme Value Theory for Measuring Financial Risk*. Elsevier Science. Preprint.
- [48] Glosten, L. R., Jagannathan, R. y Runkle, D. E. (1993): *On the Relation Between the Expected Value and the Volatility of The Nominal Excess Return on Stocks*. Journal of Finance, 48 (5): 1779-1801.

-
- [49] Gnedenko, B. V. (1943): *Sur la Distribution Limite du Terme Maximum D'une Série Aléatoire*. Annals of Math. 44:423-453.
- [50] Gouriéroux, C. (1997): *ARCH-Models and Financial Applications*. Springer Series in Statistics. Springer. New York.
- [51] Granger, C.W.J. y Newbold, P. (1977): *Forecasting Economic Time Series* Academic Press. New York.
- [52] Granger, C. W. J. y Andersen, A. P.: 1978: *Introduction to Bilinear Time Series Models*. Vandenhoeck and Ruprecht, Göttingen
- [53] Hoeffding, W. (1940): *Maszstabinvariante Korrelationstheorie*. Schr. Math. Inst. und Inst. Angewandte Math. Universität Berlín, 5:181-233.
- [54] Hoeffding, W. (1948): *A Class of Statistics With Asymptotically Normal Distribution*. Annals Math. Statistics, 19: 293-325.
- [55] Jarque, C. M. y Bera, A. (1980): *Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals*. Economic Letters, 6 (3): 255-259.
- [56] Jenkinson, A. F. (1955): *The Frequency Distribution of The Annual Maximum (or Minimum) Values of Meteorological Elements*. Quarterly Journal of the Royal Meteorological Society, 81: 158-171.
- [57] Joe, H. (1997): *Multivariate Models and Dependence Concepts*. Chapman&Hall.
- [58] Johnson, N. L., y Kotz, S. (1970). *Distributions in Statistics. Continuous Univariate Distributions - 2*. New York: Wiley.
- [59] Krause, A. y Olson, M. (2002): *The Basics of S-Plus*. Springer. New York.
- [60] Kwiatkowski, D., Phillips, P., Schmidt, P. y Shin, Y. (1992): *Testing the Null Hypothesis of Stationary Against The Alternative of A Unit Root. How Sure Are We That Economic Series have A Unit Root?*. Journal of Econometrics, 54: 159-178.
- [61] Lawley, D.N. (1956): *A General Method for Approximating to the Distribution of Likelihood Ratio Criteria*. Biometrika, 43: 295-303.

- [62] Leibler, R. A. y Kullback, S. (1951): *On Information and Sufficiency*. Annals of Mathematical Statistics, 22:79-86.
- [63] Longin, F. (1977a): *From Value at Risk To Stress Testing, The Extreme Value Approach*. Discussion Paper 97-004. CERESSEC.
- [64] Longin, F. (1977b): *Beyond The VaR*. Discussion Paper 97-011. CERESSEC.
- [65] Longin, F. y Solnik, B. (2001):. *Extreme Correlation of International Equity Markets*. Journal of Finance, 56(2): 649-676.
- [66] Marshall, C. (2001): *Measuring and Managing Operational Risk in Financial Institutions: Tools, Techniques and Other Resources*. John Wiley & Sons. Singapore.
- [67] De Matteis, R. (2001): *Fitting Copulas to Data*. Diploma thesis Institute of Mathematics of the University of Zurich.
- [68] McNeil, A. (1998): *Calculating Quantile Risk Measures for Financial Return Series Using Extreme Value Theory*. Preprint. ETH Zürich.
- [69] McNeil, A. y Frey, R. (2000): *Estimation of Tail Related Risk Measures for Heteroscedastic Financial Time Series: An Extreme Value Approach*. Journal of Empirical Finance, 7: 271-300.
- [70] Milhoj, A. (1985). *The Moment Structure of ARCH Processes*. Scandinavian Journal of Statistics 12: 281-292.
- [71] Nelsen, R. (1999): *An Introduction to Copulas*. Springer-Verlag.
- [72] Nelson, D. B. (1991): *Conditional Heteroskedasticity in Asset Returns: A New Approach*. Econometrica, 59 (2): 347-370.
- [73] Oakes, D. (1994): *Multivariate Survival Distributions*. J. Nonparam. Statist. 3 (3-4): 343-354.
- [74] Pagan, A. (1996): *The Econometrics of Financial Markets*. Journal of Empirical Finance, 3: 15-102.
- [75] Patton, A. J. (2001a): *Modelling Time-Varying Exchange Rate Dependence Using the Conditional Copula*. Economics Working Paper Series 2001-09, Department of Economics, UC San Diego.

- [76] Patton, A. J. (2001b): *Estimation of Copula Models for Time Series of Possibly Different Lengths*. Economics Working Paper Series 2001-17, Department of Economics, UC San Diego
- [77] Pickands, J. (1975): *Statistical Inference Using Extreme Order Statistics*. The Annals of Statistics 3: 199-131.
- [78] Riskmetrics (1995): *Riskmetrics Thecnical Document*. 3^a ed. J. P. Morgan.
- [79] Samara, B. y Randles, R. H. (1988). *Kendall's Tau*. Comm. Statist. Theory Methods, 17: 3191-3205.
- [80] Shih, J.H. y Louis, T.A. (1995): *Inferences on the Association Parameter in Copula Models for Bivariate Survival Data*. Biometrics, 51: 1384-1399.
- [81] Sklar, A (1959): *Fonctions de Repartition à n-dimensions et Leur Marges*. Publ. Inst. Statist. Univ. París.8: 229-231.
- [82] Sklar, A (1996): *Random Variables, Distribution functions and Copulas- a personal look backward and forward in Distributions with fixed Marginals and Related Topics*. Rüschendorf, L., Schweizer, B. y Taylor, M. Institute of Mathematical Statistics. Hayward, CA:1-14.
- [83] Sklar, A. y Schweizer, B.(1983): *Probabilistic metric Spaces*. North-Holland. New York.
- [84] Schmidt, R. y Stadmüller, U. (2003): *Nonparametric Estimation of Tail Dependence*. Working Paper. London Bussiness School. London. U.K.
- [85] Smith, R. (2002): *Risk management: Value at Risk and Beyond*. Cambridge U.P.
- [86] Van der Goorberg, R. W., Genest, C. y Werker, J. M. (2005): *Bivariate Option Pricing Using Dynamic Copula Models*. Insurance: Mathematics and Economics, 37: 101-114.
- [87] Tajvidi, N. (2004): *Confidence Intervals and Accuracy Estimation for Heavy-Tailed Generalized Pareto Distributions*. Extremes 6: 111-123. Kluwer Academic Publishers. Netherlands.
- [88] Venables, W. N. y Ripley, B. D. (2002): *Modern Applied Statistics With S*. Springer. New York.

- [89] Vilarino, A. (2001): *Turbulencias Financieras y Riesgos de Mercado*. Prentice Hall. Madrid.
- [90] Von Mises, R. (1954): *La Distribution De La Plus Grande De n Valeurs*. American Mathematical Society. Reprinted in Selected Papers. Vol II: 271-294. Providence. R. I.
- [91] Zakoian, J. (1991): *Threshold Heteroskedasticity Model*. Unpublished manuscript. INSEE.
- [92] Zivot, E. y Wang, J. (2003) *Modeling Financial Time Series With S-Plus*. Springer-Verlag.

Relación de Procedimientos

A continuación, se relacionan los procedimientos y funciones desarrollados para la implementación de las diferentes técnicas propuestas a lo largo de esta tesis. Para su realización se ha utilizado el paquete estadístico Splus 6.1 en su versión para Windows, especialmente el módulo S+FinMetrics en su versión 2.0.1 que incorpora funciones específicas de tratamiento de series financieras y la librería MASS realizada por Venables, W. y Ripley, B. versión 7.0.

También señalar, que han sido de gran ayuda los textos de Carmona (2004), Krause y Olson (2002), Venables y Ripley (2002) y Zivot y Wang (2003).

Capítulo 2:

- Alinear y describir series.scc

Incluye las funciones

```
miqqnorm <- function(x)
```

```
forma <- function(x)
```

- Retornos.scc
- Cópula empírica.scc

Incluye las funciones

```
empiric.df <-function(data,x)
```

```
tail.index.empirical.copula <-function(copula)
```

- Estimación Genest-Rivest.scc

Incluye las funciones

```
calc.theta <- function(ktau)
kfi.1 <- function(t,a)
kfi.2 <- function(t,a)
kfi.3 <- function(t,a)
varW <- function(y)
varW.2 <-function(y)
sd.clayton <- function(copula)
sd.gumbel <- function(copula)
```

- Estimación CML.scc

Incluye las funciones

```
neg.log.lik <- function(param,copula,data)
AIC <- function (copula)
L.theta.clayton <- function(a,u,v)
L.u.clayton <- function(a,u,v)
L.v.clayton <- function(a,u,v)
L.theta.frank <- function(a,u,v)
L.u.frank <- function(a,u,v)
L.v.frank <- function(a,u,v)
N.cl
M.cl
sd.cl
N.frank
M.frank
sd.frank
```

```
L.theta.gumbel <- function(a,u,v)
```

```
L.u.gumbel <- function(a,u,v)
```

```
L.v.gumbel <- function(a,u,v)
```

```
N.gumbel
```

```
M.gumbel
```

```
sd.gumbel
```

```
L.theta.normal <- function(a,u,v)
```

```
L.u.normal <- function(a,u,v)
```

```
L.v.normal <- function(a,u,v)
```

```
N.normal
```

```
M.normal
```

```
sd.normal
```

```
sum1
```

- Estimación modelos mixtos.scc

Incluye las funciones

```
duv.mix <-function(u,v,par1)
```

```
negloglik <- function(par1)
```

```
ti.low.mix
```

```
ti.up.mix1
```

```
du.g
```

```
du.gs
```

```
duv.mix2 <- function(u,v,param2)
```

```
negloglik2 <- function(param2)
```

```
du.g2
```

du.c2

Capítulo 3:

- Filtrado GARCH-DowJones.scc
- Filtrado GARCH-Ibex35.scc

Capítulo 4:

- Dependencia en colas.scc

Incluye las funciones

chi.plot.ur_function(datos)

chi.plot.ll_function(datos)

gpdDJ.sim

gpdIb.sim

residDJ.umbral

residDJ.umbral.est

residIb.umbral

residIb.umbral.est

- Densidades.scc

Incluye las funciones

fx.mod1<-function(x,y,a,b,c)

fx.mod2<-function(x,y,a2,b2,c2)

- Simulación modelos mixtos.scc

Incluye las funciones

rfib <- function(n)

generación.mod1

generación.mod2

muestra.mod1.3000

muestra.mod2.3000

- Simulación retornos cartera.scc

Incluye las funciones

```
retornos.sim<- function (Q, sim.u, sim.v, x.est, y.est, lambda1, lambda2)
```

```
vi.m1
```

```
vi.m2
```

```
Risk.portf.normal
```

```
icm1.var
```

```
icm1.es
```

```
icm2.var
```

```
icm2.es
```

Capítulo 5:

- Cópula dinámica.scc

Incluye las funciones

```
lambda.inf <- function(x.rank,y.rank)
```

```
lambda.sup<-function(x.rank,y.rank)
```

```
muestra.mod1.2000
```

```
risk.m1.2000
```