

WORKING PAPER



PRUEBAS DE ESTRÉS EN LAS INSTITUCIONES FINANCIERAS: ENFOQUE DE LA TEORÍA DE VALORES EXTREMOS CON COPULAS EXTREMAS

Por: Eduardo Court M.

eduardocourt@yahoo.es

Profesor de la escuela de Post Grado de la USMP

Los reguladores establecen que las empresas financieras deben de hacer simulaciones de crisis con el fin de conocer el monto de pérdidas potenciales en el caso de fluctuaciones fuertes de los mercados. Sin embargo, a pesar del carácter obligatorio de las mismas, estas simulaciones de crisis no son usadas para calcular la exigencia de los fondos propios. Sin embargo, el uso de simulaciones de crisis es uno de los cuatro principios de base definidos por Basilea para que un modelo interno sea válido. El objetivo de estas simulaciones es complementar los mecanismos de medición de riesgos y especialmente las medidas del VaR.

Estas simulaciones de crisis, llamados escenarios de estrés, (stress testing), deben ser las herramientas para comprender la exposición del establecimiento financiero a una crisis grave. Contrariamente al VaR, estas deben de responder a la siguiente pregunta:

“cuál es el monto de pérdida anual al que la institución financiera deberá hacer frente luego de la próxima crisis si el portafolio de negociación no cambia?”

La noción de probabilidad desaparece. Estadísticamente los test de estrés hacen referencia a un máximo y no a un cuantil. Es una visión de riesgo extremo. El punto de vista es totalmente diferente del que nos da el VaR, ya que el regulador no está totalmente enterado de lo que concierne a esta medida. Las medidas que nos da el test de estrés son antes que todo destinadas al riesgo de gerencia y a la dirección general. Estas deben de estar bastante detalladas para poder comprender la exposición de la entidad financiera y también para verificar que este riesgo extremo está en un nivel soportable.

Existen dos grandes familias de test de estrés: Los métodos objetivos y los llamados métodos subjetivos. El primer tipo usa los hechos históricos (hechos que hayan ocurrido en la realidad) para construir escenarios, mientras que el segundo se basa en hipótesis de trabajo. Para ser más preciso, podemos clasificar los test de estrés según tres enfoques:

1.- Usamos la data de crisis pasadas y empleamos un método de simulación histórica sobre los periodos con problemas para calcular una pérdida potencial máxima. (y no un VaR). La entidad financiera tendrá entonces una estimación de que podría ocasionar una crisis similar con el portafolio actual.

2.- Hacemos un análisis estructurado de escenarios. Los acontecimientos históricos se usan aplicándolos a las condiciones actuales del mercado. Por ejemplo, podríamos analizar un alza de las tasas de la FED como si ocurriera el día de hoy. Contrariamente al primer método, este usará cifras simuladas a partir de cifras actuales y de acontecimientos pasados.

3.- El tercer enfoque es un análisis de escenarios específicos. Para esto debemos de identificar una situación de riesgo que amenace a la entidad financiera. Este método no usa acontecimientos pasados, usa hipótesis sobre crisis potenciales futuras.

Los test de estrés surgen entonces de las preguntas que debe de plantearse la entidad financiera sobre su riesgo, por ejemplo:

- ¿Cuál es el impacto de un alza de tasas de 100 puntos básicos?
- ¿Qué pasaría si la curva de tasas se invierte?
- ¿Cuáles son las repercusiones de una devaluación?
- ¿Qué pasaría en ausencia de liquidez?
- ¿Cuál es el impacto de un alza de precios inmobiliarios?
- Etc.

Es necesario saber que en la actualidad no existe un estándar preciso para la elaboración de escenarios de crisis, por lo que estaremos bien si estamos en la capacidad de responder a estas tres preguntas:

- ¿Cuáles serían las pérdidas si se produce el escenario X ?
- ¿Cuáles son los peores escenarios para la institución?
- ¿Qué podemos hacer para limitar las pérdidas en este caso?

Sabemos por otro lado que los reguladores buscan influenciar la elección de la dirección en materia de gestión del riesgo, por lo que debemos tener en mente dos exigencias:

- La credibilidad del escenario a los ojos del CEO
- Una buena lectura de la exposición del portafolio a los diferentes factores de riesgo definidos.

1.1 Metodologías escogidas por las instituciones financieras

Una encuesta reciente de Deloitte muestra estas metodologías:

Evaluación del test de estrés	
(1) Escenarios apoyados en política económica	40%
(2) Escenarios con shock de mercados	70%
(3) Shocks históricos	65%
(4) Test de estrés global	30%
(5) Test de estrés micro	44%
(6) Teoría de valores extremos	5%

Los métodos 1,2,4 y 5 son métodos subjetivos; los métodos 3 y 6 corresponden a un enfoque histórico y a un enfoque sistemático usando los extremos.

1.2 El enfoque histórico

Este es un enfoque pedido por los reguladores. La idea es simple: nos concentramos sobre la evolución de nuestros factores de riesgo en un periodo dado (típicamente del 1ro de enero al 31 de diciembre) y deducimos de ahí el peor periodo, el cual constituirá nuestro escenario de crisis. Indudablemente, la calidad del análisis dependerá de la elección de nuestros peores periodos.

De esta forma, el enfoque más simple consiste en listar para cada factor de riesgo la variación al alza o baja más importante sobre un periodo de tiempo

elegido. A partir de ahí podemos simular escenarios de crisis seleccionando algunos de estos valores, los otros factores de riesgo hallados quedan idénticos a su valor actual. Automáticamente vemos que la ausencia de correlación entre los diferentes factores es un obstáculo a la credibilidad de los escenarios, de ahí la idea de elegir como escenario de crisis una configuración histórica incluyendo todos los factores. Lo peor sería definido entonces sobre la base del portafolio actual: ya sea calculando los valores potenciales del portafolio para los periodos pasados, o ya sea calculando la media ponderada de las variaciones de los factores para cada periodo; en los dos casos lo peor es definido como el periodo donde se obtuvo el mínimo. Este tipo de enfoque nos da entonces un escenario de crisis cuya credibilidad está asegurada debido a que ya se produjo.

La ventaja de este método reposa en su credibilidad y en su facilidad de cálculo. Él también nos da un marco cuantitativo: la pérdida potencial máxima calculada en el pasado en base al portafolio actual nos permite tener un orden de magnitud de las pérdidas acumuladas para todo escenario creado. Sin embargo, este tipo de enfoque presenta el inconveniente de ser poco comprensible: ya que obtenemos una pérdida agregada y dependiente del peor periodo, lo que hace difícil para la dirección, la localización objetiva de las zonas de riesgo de su portafolio.

1.3 El test de estrés macro

Este es un tipo de enfoque subjetivo, ya que se trata de una metodología que no se basa en el análisis de los datos pasados sino más bien se basa en un acontecimiento económico o político no resuelto o atendido, y a partir de esto trata de recrear un encadenamiento de acontecimientos los cuales posteriormente calibra de manera cuantitativa para crear diversos escenarios de crisis.

Tenemos de esta forma un método de dos etapas:

- 1.- Se elabora un escenario dinámico a partir de un dato (por ejemplo, el Dow Jones cae en 8%)
- 2.- Construcción de un balance final basado en términos de factores de riesgo del portafolio.

Se le llama test de estrés macro ya que se deduce la dinámica de la crisis y los balances de resultados a partir de modelos macro económicos.

Este tipo de escenarios son indispensables ya que nos permiten crear escenarios de crisis que no pueden ser inferidos de los datos del pasado, y son asimismo posibles de ocurrir. Sin embargo, para que estos escenarios sean eficaces deben ser creíbles. Por lo tanto, debemos tener en cuenta:

- 1.- De usar diferentes escenarios de profundidad de la crisis:
- 2.- Debemos emplear en su construcción opiniones de muchas personas con credibilidad y de variadas competencias, y también a personas con responsabilidades muy altas;

3.- Debemos crear de manera regular nuevos escenarios adaptados a los cambios políticos y económicos.

Como hemos mencionado, estos programas de test de estrés macro corresponden a simulaciones de crisis. Su objetivo es calcular la pérdida máxima a la que deberá hacer frente la banca si se dan los escenarios. Este tipo de análisis forman parte de los procesos de validación del VaR exigidos por los reguladores.

La teoría de valores extremos, no es la teoría más robusta para crear las simulaciones de crisis, sin embargo, nos permite cuantificar escenarios, es decir, obtener la probabilidad de ocurrencia y de dar un tiempo de retorno implícito. En otras palabras, nos permite asociar un grado de severidad/gravedad a cada escenario y compararlo en términos de probabilidad.

1.3.1 El caso unidimensional

Asumimos un tiempo de retorno \mathcal{T} y buscamos el acontecimiento implícito a ese tiempo de retorno. Sea F la distribución del rendimiento diario de un activo. Tendremos:

$$r = F^{-1}\left(1 - \frac{1}{\mathcal{T}}\right) \quad (1.1)$$

Generalmente \mathcal{T} toma valores muy grandes (5 años, 10 años, etc.), lo que implica umbrales de confianza extremadamente débiles para los cuantiles. Por ejemplo, si $\mathcal{T} = 5$ años, y si asumimos 250 días de trading por año, el umbral de confianza es igual a:

$$\alpha = 1 - \frac{1}{5 \times 250} = 99.92\% \quad (1.2)$$

Para un $\mathcal{T} = 50$ años, tendremos un $\alpha = 99.992\%$. Como estos umbrales de confianza se encuentran muy elevados, el uso de la teoría de valores extremos aparece como pertinente. Sea φ el tamaño de los bloques. Tendremos:

$$\alpha_{GEV} = 1 - \frac{\varphi}{\mathcal{T}} \quad (1.3)$$

GEV=Generalized Extreme Value distribution

Por ejemplo, para un $\mathcal{T} = 5$ años, α_{GEV} toma el valor de 98%.

1.3.2 El caso multidimensional

En el caso multidimensional la construcción de escenarios de crisis a partir de la teoría multivariada de valores extremos es compleja. Por este motivo usaremos la teoría de extremos solo para calcular los tiempos de retorno implícitos.

Los escenarios bivariados de crisis pueden ser vistos como la representación de una zona de fracaso. Sean \mathcal{X}_1^+ y \mathcal{X}_2^+ los estadísticos del máximo de las variables aleatorias X_1 y X_2 . Tendremos:

$$\begin{aligned}\Pr(\mathcal{X}_1^+ > \mathcal{X}_1, \mathcal{X}_2^+ > \mathcal{X}_2) &= 1 - \Pr(\mathcal{X}_1^+ \leq \mathcal{X}_1) - \Pr(\mathcal{X}_2^+ \leq \mathcal{X}_2) + \Pr(\mathcal{X}_1^+ \leq \mathcal{X}_1, \mathcal{X}_2^+ \leq \mathcal{X}_2) \\ &= 1 - \mathbf{F}_1(\mathcal{X}_1) - \mathbf{F}_2(\mathcal{X}_2) + \mathbf{C}(\mathbf{F}_1(\mathcal{X}_1), \mathbf{F}_2(\mathcal{X}_2)) \\ &= \overline{\mathbf{C}}(\mathbf{F}_1(\mathcal{X}_1), \mathbf{F}_2(\mathcal{X}_2))\end{aligned}\quad (1.4)$$

Sea \mathcal{T} el tiempo de retorno. La zona de fracaso se representa por:

$$\left\{ (\mathcal{X}_1, \mathcal{X}_2) \in \mathbb{R}^2 \mid \overline{\mathbf{C}}(\mathbf{F}_1(\mathcal{X}_1), \mathbf{F}_2(\mathcal{X}_2)) < \frac{1}{\mathcal{T}} \right\} \quad (1.5)$$

Observemos que, para un escenario de crisis dado, podemos calcular el tiempo de retorno implícito usando la siguiente ecuación:

$$\mathcal{T} \frac{1}{\Pr(\mathcal{X}_1^+ > \mathcal{X}_1, \mathcal{X}_2^+ > \mathcal{X}_2)} = \frac{1}{\overline{\mathbf{C}}(\mathbf{F}_1(\mathcal{X}_1), \mathbf{F}_2(\mathcal{X}_2))} \quad (1.6)$$

La estructura de dependencia juega un rol importante para el cálculo de \mathcal{T} .

Si la dependencia no es perfecta, obtendremos un tiempo de retorno más alto.

Un programa de test de estrés debe ser un escenario extremo con posibilidad de ocurrencia. Es absurdo hacer escenarios de crisis con tiempos de retorno muy largos (muchos años). En general lo ideal es trabajar con retornos a 5 o 10 años, excepcionalmente 25 años; es por este motivo que es importante tener una estimación de la función de dependencia.

1.4 El método WCS (Worst-case scenario)

Este método consiste en dividir un periodo de referencia T en N sub intervalos del mismo tamaño. Sea X_n la variable aleatoria que describe la posición neta del portafolio para el n -ésimo sub intervalo. En este contexto, podemos definir la medida del VaR a partir del número de sub intervalos que presenten una pérdida mayor a un valor dado x^- :

$$\frac{\sum_{n=1}^N \mathbf{1}_{|X_n \leq x^-|}}{N} = 1 - \alpha \quad (1.7)$$

donde α es el umbral de confianza. Presentado de esta forma, x^- no es otra cosa que la medida del VaR. Contrariamente, el análisis WCS considera la distribución de la máxima pérdida (o la función de la distribución de lo peor) entre los sub intervalos. Lo que buscamos entonces es caracterizar la función de distribución \mathbf{G}_N de $X_N^- = \min(X_1, \dots, X_n, \dots, X_N)$. Si asumimos que las variables

aleatorias X_n son independientes y tienen la misma ley de distribución F , tendremos:

$$\begin{aligned} \mathbf{G}_N(x) &= \Pr(\mathcal{X}_N^- \leq x) \\ &= 1 - \Pr(X_1 > x, \dots, X_n > x, \dots, X_N > x) \\ &= 1 - [1 - F(x)]^N \end{aligned} \quad (1.8)$$

Denominemos a g_N y f como las funciones de densidad correspondientes, entonces podemos mostrar que:

$$g_N(x) = Nx [1 - F(x)]^{N-1} f(x) \quad (1.9)$$

$$\mathbf{G}_N^{-1}(\alpha) = \mathbf{F}^{-1}\left(1 - (1 - \alpha)^{\frac{1}{N}}\right) \quad (1.10)$$

donde $\mathbf{G}_N^{-1}(\alpha)$ es el cuantil de la peor función de distribución.

Observación

Cuando $N=1$, es decir que estamos considerando solo un día de trading, tendremos $\mathbf{G}_1(x) = F(x)$, $g_1(x) = f(x)$ y $\mathbf{G}_1^{-1}(\alpha) = F^{-1}(\alpha)$. En este caso, $|\mathbf{G}_1^{-1}(\alpha)|$ es exactamente el valor en riesgo.

$|\mathbf{G}_1^{-1}(\alpha)|$ aparece como una extensión de la medida del VaR. Observemos que \mathfrak{R}_{WCS} . Esta medida de riesgo depende entonces de dos parámetros, N el número de días de trading y α la probabilidad ($\alpha_{WCS} = 1 - \alpha_{VaR}$) que la pérdida sobrepase a \mathfrak{R}_{WCS} . Hagamos un ejemplo, supongamos que la posición neta de un portafolio X sigue una ley gaussiana estándar $\mathcal{N}(0,1)$. Tomamos los valores para $|\mathbf{G}_1^{-1}(\alpha)|$ y los organizamos en una tabla:

$N - \alpha$	0.5%	1%	5%	10%
1	2.5758	2.3263	1.6449	1.2816
5	3.0896	2.8769	2.3187	2.0365
10	3.2899	3.0889	2.5679	2.3087
100	3.8900	3.7178	3.2834	3.0748
250	4.1069	3.9432	3.5334	3.3384

Cuando el horizonte considerado es un día de trading y para un α igual a 1%, la medida de riesgo es igual a 2.33, el valor en riesgo. Si tomamos 5 días de trading, el riesgo pasa a 2.88. Podemos apreciar que la medida del riesgo depende del número de días de trading. La curvatura de la función $|\mathbf{G}_1^{-1}(\alpha)|$ en función de N nos da una cierta visión dinámica del riesgo. A partir de la peor

función de distribución, podemos construir otras medidas de riesgo. Por ejemplo, podríamos considerar la peor esperanza $\eta(N)$ definida por:

$$\begin{aligned}\eta(N) &= -\mathbb{E}_{[x_N]} \\ &= -N \int [1 - \mathbf{F}(x)]^{N-1} \int(x) x dx\end{aligned}\quad (1.11)$$

Esta peor esperanza depende fuertemente de la función de densidad y especialmente de la cola izquierda de la misma. También podemos apreciar que en el caso que la distribución de pérdidas potenciales sea leptokurtica, la diferencia de riesgo entre esta distribución y la distribución gaussiana estimada puede crecer fuertemente con N .

Generalmente el análisis WCS está asociado a las simulaciones de crisis, debido a que ambas están relacionadas con el tiempo de retorno. En general, una metodología VaR no capta las crisis debido a que el tiempo de retorno basado en VaR tiene poco realismo. El análisis WCS o la teoría de los extremos son las teorías más adaptadas a este tema.

1.5 La teoría de los valores extremos

1.5.1 La heurística de los extremos

Una distribución gaussiana está perfectamente caracterizada por sus dos primeros momentos. El conocimiento de la mediana y de la varianza nos permite entonces calcular todos los cuantiles. Bajo la hipótesis de normalidad, el análisis estadístico trata sobre las características más visibles y más fáciles de interpretar. Sin embargo, la gestión del riesgo es la gestión de eventos inesperados, de acontecimientos raros. Con la distribución gaussiana, estos acontecimientos tienen un carácter previsible.

Contrariamente a la inferencia estadística gaussiana, la teoría de valores extremos consiste en analizar las ocurrencias que presentan las frecuencias más débiles. El análisis gaussiano consiste en focalizarse sobre los acontecimientos que se concentran alrededor del modo de la función de densidad, y esto hace que, bajo este análisis, no se pueda distinguir las partes accidentales de la forma general de esta función de densidad.

La teoría de los extremos nos permite una aproximación a las partes extremas de la función de densidad, esto significa que dejamos de lado una gran cantidad de datos y nos concentramos en aquellos extremos que vamos a modelar.

Hay que entender que la teoría de valores extremos no es una herramienta de modelamiento del carácter leptokurtico de una distribución. Sin embargo, nos permite caracterizar la ley de los extremos. La teoría de valores extremos nos permitirá estimar directamente la distribución a partir de los datos. La idea es la siguiente:

- Que podemos decir de la distribución límite $\lim_{N \rightarrow \infty} \mathbf{G}_N(x)$
- Sin hacer hipótesis sobre la distribución \mathbf{F} ?

Podríamos afirmar que la teoría de valores extremos nos permite predecir el futuro con cierta certeza. Pero el pasado contiene mucha información, que no necesariamente es observable, sobre los extremos como para poder modelar $G_\infty(x)$, sin querer conocer la función F . Y esto será posible bajo ciertas condiciones, $G_\infty(x)$ no puede ser declinante muchas veces, de hecho, solo un número limitado de veces, de las leyes de distribución.

1.5.2 La teoría clásica

1.5.2.1 Las tres leyes de extremos

Consideremos N variables aleatorias $X_1, \dots, X_n, \dots, X_N$ independientes y con la misma ley de distribución F en las cuales buscamos estudiar los comportamientos extremos $\mathcal{X}_N^- = \min(X_1, \dots, X_n, \dots, X_N)$ y $\mathcal{X}_N^+ = \max(X_1, \dots, X_n, \dots, X_N)$. Para esto podemos usar la estadística de orden Y sobre X , es decir, tendremos:

$$\mathcal{X}_N^- = Y_1 \leq \dots \leq Y_n \leq \dots \leq Y_N = \mathcal{X}_N^+ \quad (1.12)$$

Consideremos ahora la variable aleatoria $\wedge = \sum_{n=1}^{k(N)} h(Y_n)$ esta \wedge tiene varios usos en finanzas:

- Si $h(y) = \frac{1}{k(N)} x y$, \wedge corresponde a la media de los $k(N)$ -ésima pérdidas más importantes;
- Si $h(y) = (y^- y)_+$ y $k(N) = N$, \wedge representa la suma de pérdidas más grandes que un umbral;
- Si $h(y) = y$ y $k(N) = \alpha N$, \wedge es el cuantil α de X

Con la teoría de los extremos nos interesamos directamente en $\wedge = Y_1$ y $\wedge = Y_N$. Para esto requerimos del teorema de Fisher-Tippett que nos permite caracterizar la ley de distribuciones extremas:

Teorema 1: (Embrechts, Kluppelberg y Mikosh, [1997])

Supongamos N variables aleatorias $X_1, \dots, X_n, \dots, X_N$ independientes y con la misma ley de distribución F . Si existen constantes a_N y b_N y una distribución límite no degenerada G tal que:

$$\lim_{N \rightarrow \infty} \Pr \left(\frac{\mathcal{X}_N^+ - b_N}{a_N} \leq x \right) = G(x) \quad \forall x \in \mathbb{R} \quad (1.13)$$

Entonces G pertenece a uno de los tres tipos de distribución.

Uno de los resultados más importantes de la teoría de extremos es: bajo ciertas condiciones, la ley de distribución de extremos es una de las tres leyes paramétricas que se muestran en el cuadro anterior. Podríamos también caracterizar los tres tipos de distribución por una distribución única:

$$\mathbf{G}(x) = \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{\frac{1}{\xi}} \right\} \quad (1.14)$$

definida en el soporte $\Delta = \left\{ x : 1 + \xi \left(\frac{x - \mu}{\sigma} \right) > 0 \right\}$ y que denominaremos $\mathcal{G}(\mu, \sigma, \xi)$

Esta función de distribución corresponde a la ley de probabilidad de Von Mises, y es conocida con el nombre de GEV (Generalized Extreme Value distribution). Tendremos entonces las correspondencias siguientes:

Resaltemos que los parámetros μ y σ son los límites de b_N y a_N .

La función de densidad correspondiente es:

$$g(x) = \frac{1}{\sigma} \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{\left(\frac{1 + \xi}{\xi} \right)} \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{\frac{1}{\xi}} \right\} \quad (1.15)$$

Si graficamos la función, cada vez que μ cambia, veremos una traslación de funciones. μ es por lo tanto un parámetro de localización. Mientras que σ tiene el comportamiento de una varianza, es por eso que se le considera como un parámetro de dispersión. Por último, el parámetro ξ está relacionado con el carácter leptokurtico de la función de distribución \mathbf{F} . Es por esto que normalmente se le llama índice de cola o índice de valor extremo.

1.5.2.2 Observaciones a la ley de extremos.

La ley de factores extremos es un campo prolífico de investigación, veamos algunos resultados importantes.

Observación: Si retomamos el análisis WCS y aplicamos este al máximo \mathcal{X}_N^+ tendremos:

$$\mathbf{G}_N(x) \mathbf{F}(x)^N$$

$$g_N(x) = N \mathbf{F}(x)^{N-1} f(x) \quad (1.16)$$

y

$$\lim_{N \rightarrow \infty} \mathbf{G}_N(x) = \mathbf{G}(x) \quad (1.17)$$

Es entonces interesante estudiar la velocidad de convergencia que permite darnos una idea del periodo de referencia del análisis WCS para que este sea equivalente a la teoría de extremos.

Para medir la convergencia, Coles y Tawn [1999] proponen estudiar gráficamente el comportamiento de $-\ln\left(-\ln\left(\mathbf{G}\left(\frac{x-b_N}{a_N}\right)\right)\right)$ en función de N y compararlo con aquel de la ley límite. Para las leyes gaussianas y exponenciales, la ley límite es la distribución de Gumbel. Para las otras dos es la distribución de Frechet. El problema es construir los siguientes (a_N) y (b_N) . Para la ley exponencial, podemos demostrar que si $a_N = 1$ y $b_N = -\ln(N)$, $\frac{\mathcal{X}_N^+ - b_N}{a_N}$ converge bien en ley hacia la familia de tipo III. Para las otras leyes, el conocimiento de las continuaciones de (a_N) y (b_N) no está verificado; esto debido a que numéricamente no son adecuadas para valores débiles de N . En este caso podemos construir las continuaciones de forma numérica adoptando un buen criterio de adecuación entre $\mathbf{G}_N\left(\frac{x-b_N}{a_N}\right)$ y la ley límite $\mathbf{G}(x)$. En nuestro caso hemos usado la norma ℓ_2 para el criterio de adecuación. Podremos ver entonces que la convergencia de la ley exponencial es mucho más rápida que la de la ley gaussiana. Y las distribuciones que tienen colas leptokúrticas presentan velocidades de convergencia más débiles.

Observación: el cuantil $\mathbf{G}^{-1}(\alpha)$ de orden α para la distribución GEV está dado por la fórmula siguiente:

$$\mathbf{G}^{-1}(\alpha) = \mu - \frac{\sigma}{\xi} \left[1 - (-\ln \alpha)^{-\xi} \right] \quad (1.18)$$

Este cuantil es fuertemente influenciado por los parámetros σ y ξ ya que tenemos

$$\partial_\alpha \mathbf{G}^{-1}(\alpha) = \frac{\sigma}{\alpha (-\ln \alpha)^{1+\xi}} \quad (1.19)$$

Si graficamos esta función (diferentes curvas de cuantiles en función de ξ) veremos que mientras más varianza tenga la función \mathbf{F} , esta es más leptokúrtica, y el cuantil será más elevado.

Observación: Podemos demostrar que el modo de distribución GEV es alcanzado por el valor $\mu + \sigma(1-\xi)^\xi$. En cuanto a la media $\mathbb{E}[\mathcal{X}_N^+]$, podemos calcularla fácilmente usando un procedimiento de integración numérica del tipo Gauss-Legendre o Gauss-Hermite.

Ya casi tenemos todos los elementos para aplicar la teoría de valores extremos a la gestión de riesgos. Sea P la pérdida potencial de un portafolio. Medir el riesgo se convierte entonces en analizar la ley del máximo de $-P$, es decir, la peor distribución. A partir de esta podremos calcular los peores cuantiles, la

esperanza o el modo de los peores, el tiempo de retorno, etc. Solo nos falta desarrollar la inferencia estadística y la metodología de estimación de una distribución GEV.

1.5.2.3 Como se estiman los parámetros de la distribución GEV

Existen varios métodos para estimar los parámetros de la distribución GEV podríamos mencionar el método de estimación del índice de cola del tipo Hill o Pickands, el método de los momentos o el método del umbral basado en la distribución de Pareto generalizada. El más popular, sin embargo, es el de máxima verosimilitud.

Sea θ el vector de parámetros, tendremos:

$$\theta \begin{bmatrix} \mu \\ \sigma \\ \xi \end{bmatrix}$$

Consideremos una muestra de datos $\{X_t\}$ de dimensión $T = \varphi N$ con $\varphi \in \mathbb{R}$. Dividimos esta muestra en N bloques y definimos \mathcal{X}_n^+ de la siguiente manera:

$$\mathcal{X}_n^+ = \max \left(\left\{ X_{1+n(t-1)}, t = 1, \dots, \varphi \right\} \right) \quad (1.20)$$

La expresión de verosimilitud de la observación n es entonces:

$$L(\mathcal{X}_n^+; \theta) = \frac{1}{\sigma} \left[1 + \xi \left(\frac{\mathcal{X}_n^+ - \mu}{\sigma} \right) \right]^{-\frac{1+\xi}{\xi}} \exp \left\{ - \left[1 + \xi \left(\frac{\mathcal{X}_n^+ - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right\} \quad (1.21)$$

De acá deducimos la siguiente expresión para la logo similitud:

$$\ell(\mathcal{X}_n^+; \theta) = -\ln \sigma - \left(\frac{1+\xi}{\xi} \right) \ln \left(1 + \xi \left(\frac{\mathcal{X}_n^+ - \mu}{\sigma} \right) \right) - \left[1 + \xi \left(\frac{\mathcal{X}_n^+ - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \quad (1.22)$$

El estimador de máxima verosimilitud corresponderá entonces a:

$$\hat{\theta}_{ML} = \arg \max_{\theta \in \Theta} \sum_{n=1}^N \ell(\mathcal{X}_n^+; \theta) \quad (1.23)$$

con Θ como espacio de los parámetros. El cálculo de la expresión 1.23 es relativamente simple siempre y cuando se tomen algunas precauciones, especialmente para la especificación de Θ a causa de la singularidad al punto $\xi = 0$. Ver más abajo en el punto 1.5.4.1 la forma de uso del algoritmo de Bryoden et al. en el que se aplica un jacobiano analítico.

1.5.3 Concepto del tiempo de retorno

Para poder usar la teoría de extremos en finanzas es necesario introducir el concepto de tiempo de retorno. Consideremos un acontecimiento E que tiene una probabilidad de ocurrencia igual a α . La ley de aparición de este elemento es entonces una variable aleatoria geométrica de probabilidad α . Su esperanza matemática es:

$$\tau = \frac{1}{\alpha} \quad (1.24)$$

τ es el tiempo de retorno de E .

Veamos un ejemplo, consideremos un VaR de 99%. El tiempo de retorno asociado al acontecimiento de que el VaR sea sobrepasado es:

$$\tau = \frac{1}{1-0.99} = 100$$

Si el periodo de tenencia es de un día, el tiempo de retorno del VaR 99% es de 100 días. Esto quiere decir que se sobrepasará el VaR diario en promedio cada 100 días.

Definición: Sea $\alpha = \Pr\{\mathcal{X}_n^+ \leq x\}$. El tiempo de retorno del acontecimiento $\Pr\{\mathcal{X}_n^+ > x\}$ es igual a $(1-\alpha)^{-1}$ periodos de referencia. El tiempo de retorno en días es entonces:

$$\tau = \varphi \frac{1}{1-\alpha} \quad (1.25)$$

Observación Podemos comparar los resultados entre un VaR tradicional y un VaR GEV a partir de la siguiente relación:

$$\alpha_{GEV} = 1 - \varphi(1 - \alpha_{VaR}) \quad (1.26)$$

1.5.4 Forma general de la ley de extremos

Podemos mostrar que la forma general de la ley de extremos es $\exp(h(x))$.

Tenemos $\mathbf{G}_N(x) = \mathbf{F}(x)^N$ y:

$$d\mathbf{G}_N(x) = \mathbf{F}(x)^N \ln \mathbf{F}(x) dN + N\mathbf{F}(x)^{N-1} f(x) dx \quad (1.27)$$

Deducimos que:

$$d \ln \mathbf{G}_N(x) = \ln \mathbf{F}(x) dN + N \frac{f(x)}{\mathbf{F}(x)} dx := h_N(x) \quad (1.28)$$

Ahora si es fácil mostrar que, bajo ciertas condiciones de regularidad, $h_N(x) = h(x)$ debido a que $N \rightarrow \infty$ y $\mathbf{F}(x) \rightarrow 1^-$. Si introducimos una restricción de estabilidad (la distribución de los extremos de una muestra debe ser la misma

que la de la muestra original) se demuestra que h es la solución de una distribución funcional que solo tiene tres soluciones.

1.5.4.1 Jacobiano analítico de la función de loge verosimilitud

Tenemos:

$$\begin{aligned}\partial_{\mu}\ell(\mathcal{X}_n^+; \theta) &= \frac{1 + \xi - \omega^{-\frac{1}{\xi}}}{\sigma\omega} \\ \partial_{\sigma}\ell(\mathcal{X}_n^+; \theta) &= \frac{\left((1 + \xi) - \omega^{-\frac{1}{\xi}} \right) (\mathcal{X}_n^+ - \mu) - \sigma\omega}{\sigma^2\omega} \\ \partial_{\xi}\ell(\mathcal{X}_n^+; \theta) &= \left(1 - \omega^{-\frac{1}{\xi}} \right) \left(\frac{1}{\xi^2} \ln \omega - \frac{(\mathcal{X}_n^+ - \mu)}{\xi\sigma\omega} \right) - \frac{(\mathcal{X}_n^+ - \mu)}{\sigma\omega}\end{aligned}\quad (1.29)$$

con

$$\omega = 1 + \xi \left(\frac{(\mathcal{X}_n^+ - \mu)}{\sigma} \right) \quad (1.30)$$

1.6 Las cópulas

1.6.1 Representación canónica de una distribución bivariada

Una cópula es una herramienta estadística introducida por Sklar en 1959.

Definición: una cópula bidimensional o 2-cópula es una función C que tiene las siguientes propiedades:

- I. $DomC = [0,1] \times [0,1]$;
- II. $C(0,u) = C(u,0)$ y $C(u,1) = C(1,u) = u$ para todo u en $[0,1]$;
- III. C es 2-creciente :

$$C(v_1, v_2) - C(v_1, v_2) - C(u_1, v_2) + C(u_1, u_2) \geq 0 \quad (1.31)$$

Para todo $(u_1, u_2) \in [0,1]^2, (v_1, v_2) \in [0,1]^2$ tal que $0 \leq u_1 \leq v_1 \leq 1$ y $0 \leq u_2 \leq v_2 \leq 1$

Esta definición quiere decir que C es una distribución con márgenes uniformes. Entonces diremos que C es una distribución uniforme multidimensional.

Sean F_1 y F_2 dos distribuciones de probabilidad. Es evidente que $C(F_1(x_1), F_2(x_2))$ define una distribución de probabilidad bidimensional $F(x_1, x_2)$ donde los márgenes están dados.

En efecto, nosotros sabemos que $U_i = F_i(X_i)$ define una transformación uniforme (Probability integral transformation o PIT).

Sea $U = F(X)$. El soporte de U es evidentemente $[0,1]$. Y verificamos que la distribución U es uniforme:

$$\begin{aligned}
 \Pr\{U \leq u\} &= \Pr\{F(X) \leq u\} \\
 &= \Pr\{X \leq F^{-1}(u)\} \text{ (debido a que } F \text{ es una función creciente)} \\
 &= F(F^{-1}(u)) \\
 &= u
 \end{aligned} \tag{1.32}$$

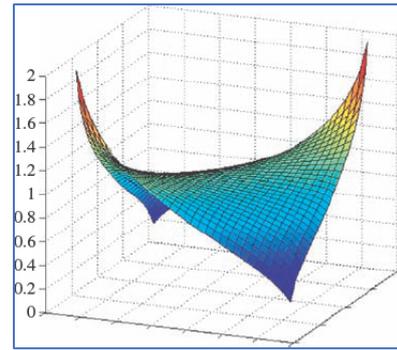
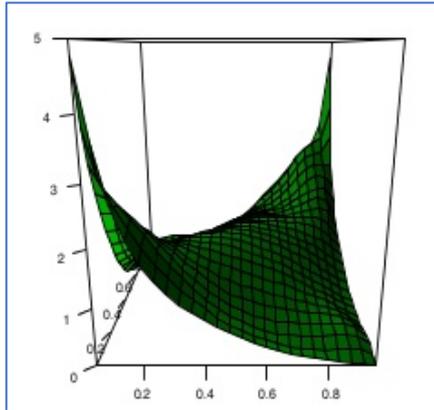
Además, verificamos que:

$$\begin{aligned}
 C(F_1(x_1), F_2(\infty)) &= C(F_1(x_1), 1) \\
 &= F_1(x_1)
 \end{aligned} \tag{1.33}$$

Las copulas son una herramienta poderosa para construir distribuciones multidimensionales donde los márgenes están dados (“Distributions with Given Marginals”)

Observación: Las copulas nos permiten construir modelos multidimensionales compatibles con modelos unidimensionales. En el sistema financiero no es extraño encontrar que una institución financiera use un modelo sofisticado de una dimensión (por ejemplo, puede usar un modelo de volatilidad estocástica para valorar las opciones), y un modelo más simple para valorizar estructuras multidimensionales (por ejemplo, usarán Black & Scholes para valorar productos estructurados). El solo hecho de que el modelo multidimensional no sea compatible con el unidimensional los puede llevar a correr graves riesgos (por ejemplo, que un trader arbitre los dos modelos).

Observación: La construcción de modelos multidimensionales no gaussianos es difícil. Por ejemplo, en las distribuciones multivariadas continuas tenemos pocas familias de probabilidades: exponenciales, gamma, Dirichlet, Liouville, logística, Pareto, exponencial natural; además, estas distintas distribuciones de probabilidad tienen el inconveniente que los márgenes son del mismo tipo. No es posible tener una distribución con un margen gaussiano, un margen uniforme, un margen de Pareto, etc. Las cópulas permiten resolver este problema. Ver gráfico:



Ejemplos de copulas de distribución bivariada

Hay un resultado aún más interesante demostrado por Sklar:

Teorema: Sea una función F de distribución bidimensional cuyos márgenes son F_1 y F_2 . Entonces F admite una representación copula:

$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2)) \quad (1.34)$$

La copula C es única si los márgenes son continuos. Si no, solo la sub copula está determinada de manera única sobre $RanF_1 \times RanF_2$.

Este teorema es importante porque nos permite asociar a cada distribución bidimensional una copula (que puede ser única). Tendremos entonces una representación canónica de la distribución: de un lado, tenemos los márgenes F_1 y F_2 , es decir, las direcciones unidimensionales; del otro lado, tenemos la copula que permite relacionar esos márgenes, esta nos explica la dependencia entre las direcciones unidimensionales.

Este teorema se puede usar para construir funciones de copulas a partir de distribuciones bidimensionales. Por ejemplo, consideremos la distribución logística bivariada de Gumbel

$$F(x_1, x_2) = (1 - e^{-x_1} + e^{-x_2})^{-1}$$

definida sobre \mathbb{R}^2 . Podemos mostrar que los márgenes son $F_1(x_1) = F(x_1, \infty) = (1 + e^{-x_1})^{-1}$ y $F_2(x_2) = (1 + e^{-x_2})^{-1}$. Las funciones inversas son entonces $F_1^{-1}(u_1) = -\ln(1 - u_1)$ y $F_2^{-1}(u_2) = -\ln(1 - u_2)$. De donde deducimos que la función copula es:

$$\begin{aligned}
\mathbf{C}(u_1, u_2) &= \mathbf{F}(\mathbf{F}_1^{-1}(u_1), \mathbf{F}_2^{-1}(u_2)) \\
&= \left(1 + \frac{1-u_1}{u_1} + \frac{1-u_2}{u_2}\right)^{-1} \\
&= \frac{u_1 u_2}{u_1 + u_2 - u_1 u_2} \tag{1.35}
\end{aligned}$$

Esta copula se llama copula logística de Gumbel (Gumbel logistic copula). En algunos casos no podremos obtener una versión analítica de la copula. Este es el caso de la copula normal:

$$\mathbf{C}(u_1, u_2) = \Phi(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \rho)$$

Si la distribución bivariada es absolutamente continua, entonces esta admite una densidad y tendremos:

$$f(x_1, x_2) = c(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2)) \times f_1(x_1) \times f_2(x_2) \tag{1.36}$$

con $c(u_1, u_2)$ como la densidad de la copula \mathbf{C} . Observemos que la condición

$$\mathbf{C}(v_1, v_2) - \mathbf{C}(v_1, u_2) - \mathbf{C}(u_1, v_2) + \mathbf{C}(u_1, u_2) \geq 0 \tag{1.37}$$

es equivalente a la positividad de la densidad:

$$\mathbf{C}(u_1, u_2) = \partial_{1,2}^2 \mathbf{C}(u_1, u_2) \geq 0 \tag{1.38}$$

cuando esta existe. Podemos usar la ecuación 1.36 para caracterizar la densidad de una cúpula normal. Tendremos

$$\begin{aligned}
&\frac{1}{2\pi\sqrt{1-\rho^2}} \left[x_1^2 + x_2^2 - 2\rho x_1 x_2 \right] \\
&= c(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2)) \times \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x_1^2\right) \times \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x_2^2\right) \tag{1.39}
\end{aligned}$$

de donde:

$$c(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2)) = \frac{1}{\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)} \left[x_1^2 + x_2^2 - 2\rho x_1 x_2 \right] + \frac{1}{2} \left[x_1^2 + x_2^2 \right]\right) \tag{1.40}$$

1.6.2 Clase Frechet, distribución mínima y máxima y orden de concordancia

Definición: \mathbf{F} corresponde a la clase de Frechet $(\mathbf{F}_1, \mathbf{F}_2)$ y observamos que $\mathbf{F} \in \mathcal{F}(\mathbf{F}_1, \mathbf{F}_2)$ sí y solo si los márgenes de \mathbf{F} son \mathbf{F}_1 y \mathbf{F}_2 :

$$\mathbf{F}(x_1, \infty) = \mathbf{F}_1(x_1) \tag{1.41}$$

y

$$\mathbf{F}(\infty, x_2) = \mathbf{F}_2(x_2) \quad (1.42)$$

Caracterizar la clase de Frechet $\mathcal{F}(\mathbf{F}_1, \mathbf{F}_2)$ es encontrar el conjunto de funciones de copulas \mathcal{C} porque tenemos:

$$\mathcal{F}(\mathbf{F}_1, \mathbf{F}_2) = \{ \mathbf{F} : \mathbf{F}(x_1, x_2) = \mathbf{C}(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2)), \mathbf{C} \in \mathcal{C} \} \quad (1.43)$$

El problema de la caracterización de la clase de Frechet $\mathcal{F}(\mathbf{F}_1, \mathbf{F}_2)$ es así independiente de los márgenes \mathbf{F}_1 y \mathbf{F}_2 .

Las funciones extremas de \mathbf{F}^- y de \mathbf{F}^+ pertenecientes a $\mathcal{F}(\mathbf{F}_1, \mathbf{F}_2)$ son:

$$\mathbf{F}^-(x_1, x_2) = \max(\mathbf{F}_1(x_1) + \mathbf{F}_2(x_2) - 1, 0) \quad (1.44)$$

y

$$\mathbf{F}^+(x_1, x_2) = \min(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2)) \quad (1.45)$$

\mathbf{F}^- y \mathbf{F}^+ son llamados los bornes de base y de altura de Frechet respectivamente. Las copulas asociadas son:

$$\mathbf{C}^-(u_1, u_2) = \max(u_1 + u_2 - 1, 0) \quad (1.46)$$

y

$$\mathbf{C}^+(u_1, u_2) = \min(u_1, u_2) \quad (1.47)$$

Estas son las copulas extremas de \mathcal{C} . Tenemos:

$$\mathbf{C}^-(u_1, u_2) \leq \mathbf{C}(u_1, u_2) \leq \mathbf{C}^+(u_1, u_2) \quad (1.48)$$

para todo $(u_1, u_2) \in \mathbb{R}^2$ y $\mathbf{C} \in \mathcal{C}$

Observación: Consideremos la clase de Frechet $\mathcal{F}(\mathbf{F}_1, \mathbf{F}_2)$ con $\mathbf{F}_1 = \mathbf{F}_2 = \mathcal{N}(0, 1)$. Así, la distribución normal con dos dimensiones de correlación ρ pertenece a $\mathcal{F}(\mathbf{F}_1, \mathbf{F}_2)$. Sin embargo, muchas otras distribuciones no normales pertenecen a esta clase de Frechet. Este es, por ejemplo, el caso de la siguiente distribución:

$$\mathbf{F}(x_1, x_2) = \frac{\Phi(x_1)\Phi(x_2)}{\Phi(x_1) + \Phi(x_2) - \Phi(x_1)\Phi(x_2)} \quad (1.49)$$

Consideremos la distribución normal

$$\mathbf{F}(x_1, x_2; \rho) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}[x_1^2 + x_2^2 - 2\rho x_1 x_2]\right) dx_1 dx_2 \quad (1.50)$$

Tendremos:

$$\mathbf{F}^-(x_1, x_2) := \mathbf{F}(x_1, x_2; -1) = \max(\Phi(x_1) + \Phi(x_2) - 1, 0) \quad (1.51)$$

y

$$\mathbf{F}^+(x_1, x_2) := \mathbf{F}(x_1, x_2; +1) = \min(\Phi(x_1), \Phi(x_2)) \quad (1.52)$$

El caso $\rho = 0$ corresponde a la independencia de las variables aleatorias y tendremos:

$$\mathbf{F}^\perp(x_1, x_2) := \mathbf{F}(x_1, x_2; 0) = \Phi(x_1)\Phi(x_2) \quad (1.53)$$

De acá deducimos la expresión que produce la copula:

$$\mathbf{C}^\perp(u_1, u_2) = u_1 u_2 \quad (1.54)$$

Introduzcamos ahora un orden parcial sobre las copulas.

Definición: Sean \mathbf{C}_1 y \mathbf{C}_2 dos copulas. Decimos que \mathbf{C}_1 es más pequeño que \mathbf{C}_2 y observamos que $\mathbf{C}_1 \prec \mathbf{C}_2$ sí y solo si $\mathbf{C}_1(u_1, u_2) \leq \mathbf{C}_2(u_1, u_2)$ para todo $(u_1, u_2) \in [0, 1]^2$.

Esta relación de orden es parcial ya que no podemos comparar todas las copulas entre ellas. Sin embargo, siempre tendremos

$$\mathbf{C}^- \prec \mathbf{C} \prec \mathbf{C}^+ \quad (1.55)$$

Podemos deducir que una estructura de dependencia positiva es una función de copula \mathbf{C} que verifica la siguiente desigualdad:

$$\mathbf{C}^\perp \prec \mathbf{C} \prec \mathbf{C}^\perp \quad (1.56)$$

De la misma forma, una estructura de dependencia negativa es una función de copula \mathbf{C} que verifica la desigualdad siguiente:

$$\mathbf{C}^- \prec \mathbf{C} \prec \mathbf{C}^\perp \quad (1.57)$$

Sin embargo, como esta relación es de orden parcial, existen las funciones de copulas \mathbf{C} tal que $\mathbf{C} \not\prec \mathbf{C}^\perp$ y $\mathbf{C} \not\prec \mathbf{C}^-$. Existen entonces funciones copulas que no son ni estructuras de dependencia positiva, ni estructura de dependencia negativa.

Observación: El orden \prec es llamado orden de concordancia. El corresponde al dominio estocástico de primer orden en las funciones de distribución.

Observación: Consideremos la copula $\mathbf{C}(u_1, u_2; \theta)$ con un parámetro θ . Observamos $\mathbf{C}_\theta(u_1, u_2) = \mathbf{C}(u_1, u_2; \theta)$. La familia de copulas $\{\mathbf{C}_\theta\}$ se dice que está completamente ordenada si para todo $\theta_2 \geq \theta_1$ podemos verificar que:

$$\mathbf{C}_{\theta_2} \succ \mathbf{C}_{\theta_1} \quad (1.58)$$

(familia positivamente ordenada) o

$$\mathbf{C}_{\theta_2} \prec \mathbf{C}_{\theta_1} \quad (1.59)$$

(familia negativamente ordenada). Por ejemplo, la familia de copulas normales $\{\mathbf{C}_\rho\}$ es positivamente ordenada.

1.6.3 Copulas y variables aleatorias

Consideremos dos variables aleatorias X_1 y X_2 de distribuciones \mathbf{F}_1 y \mathbf{F}_2 . Denominamos \mathbf{F} a la distribución del vector aleatorio $\mathbf{X} = (X_1, X_2)$. Definimos entonces la copula del vector aleatorio (X_1, X_2) por la copula de \mathbf{F} :

$$\mathbf{F}(x_1, x_2) = \mathbf{C}_{\langle X_1, X_2 \rangle}(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2)) \quad (1.60)$$

Con esta definición podemos dar una nueva interpretación a las copulas \mathbf{C}^- , \mathbf{C}^\perp y \mathbf{C}^+ :

- Dos variables aleatorias X_1 y X_2 son contra-monotónicas (countermonotonic) o $\mathbf{C}\langle X_1, X_2 \rangle = \mathbf{C}^-$ - si existe una variable aleatoria X tal que $X_1 = f_1(X)$ y $X_2 = f_2(X)$ con f_1 como una función no creciente y f_2 una función no decreciente;
- Dos variables aleatorias X_1 y X_2 son independientes si la estructura de dependencia es la copula producto \mathbf{C}^\perp ;
- Dos variables aleatorias X_1 y X_2 son co-monótonas (comonotonic) -o $\mathbf{C}\langle X_1, X_2 \rangle = \mathbf{C}^+$ - si es que existe una variable aleatoria X tal que $X_1 = f_1(X)$ y $X_2 = f_2(X)$ con f_1 y f_2 como funciones no decrecientes.

Consideremos un vector aleatorio gaussiano $\mathbf{X} = (X_1, X_2)$ de correlación ρ .

$\mathbf{C}\langle X_1, X_2 \rangle = \mathbf{C}^-$ corresponde al caso:

$$X_2 = -X_1 \quad (1.61)$$

entonces $\mathbf{C}\langle X_1, X_2 \rangle = \mathbf{C}^+$ corresponde al caso:

$$X_2 = X_1 \quad (1.62)$$

Si $\mathbf{X} = (X_1, X_2)$ es un vector aleatorio cuyos márgenes son log normal. Entonces tenemos:

$$\mathbf{C}\langle X_1, X_2 \rangle = \mathbf{C}^- \Leftrightarrow X_2 = \exp(-\ln X_1) = \frac{1}{X_1} \quad (1.63)$$

y

$$\mathbf{C}\langle X_1, X_2 \rangle = \mathbf{C}^- \Leftrightarrow X_2 = X_1 \quad (1.64)$$

Podemos ahora desarrollar uno de los teoremas más importantes de las copulas:

Teorema: sean dos variables aleatorias continuas X_1 y X_2 de margen \mathbf{F}_1 y \mathbf{F}_2 y de copula $\mathbf{C}\langle X_1, X_2 \rangle$. Si h_1 y h_2 son dos funciones estrictamente crecientes sobre $\text{Im}(X_1)$ e $\text{Im}(X_2)$ respectivamente, entonces

$$\mathbf{C}\langle h_1(X_1), h_2(X_2) \rangle = \mathbf{C}\langle X_1, X_2 \rangle \quad (1.65)$$

Por lo tanto, la copula es invariante por las transformaciones estrictamente crecientes de las variables aleatorias.

Podemos también demostrar este teorema usando las propiedades de las distribuciones. Denominemos \mathbf{F} y \mathbf{G} a las distribuciones de los vectores aleatorios (X_1, X_2) y $(h_1(X_1), h_2(X_2))$. Observamos que los márgenes de \mathbf{G} son:

$$\begin{aligned} \mathbf{G}_1(x_1) &= \Pr\{h_1(X_1) \leq x_1\} \\ &= \Pr\{X_1 \leq h_1^{-1}(x_1)\} \text{ (ya que } h_1 \text{ es una función creciente)} \\ &= \mathbf{F}_1(h_1^{-1}(x_1)) \end{aligned} \quad (1.66)$$

y $\mathbf{G}_2(x_2) = \mathbf{F}_2(h_2^{-1}(x_2))$. Entonces tenemos $\mathbf{G}_1^{-1}(u_1) = h_1(\mathbf{F}_1^{-1}(u_1))$ y $\mathbf{G}_2^{-1}(u_2) = h_2(\mathbf{F}_2^{-1}(u_2))$. Deducimos el resultado:

$$\begin{aligned} \mathbf{C}\langle h_1(X_1), h_2(X_2) \rangle(u_1, u_2) &= \mathbf{G}(\mathbf{G}_1^{-1}(u_1), \mathbf{G}_2^{-1}(u_2)) \\ &= \Pr\{h_1(X_1) \leq \mathbf{G}_1^{-1}(u_1), h_2(X_2) \leq \mathbf{G}_2^{-1}(u_2)\} \\ &= \Pr\{X_1 \leq h_1^{-1}(\mathbf{G}_1^{-1}(u_1)), X_2 \leq h_2^{-1}(\mathbf{G}_2^{-1}(u_2))\} \\ &= \Pr\{X_1 \leq \mathbf{F}_1^{-1}(u_1), X_2 \leq \mathbf{F}_2^{-1}(u_2)\} \\ &= \mathbf{C}\langle X_1, X_2 \rangle(u_1, u_2) \end{aligned} \quad (1.67)$$

Por ejemplo, tendremos

$$\begin{aligned} \mathbf{C}\langle X_1, X_2 \rangle &= \mathbf{C}\langle \ln X_1, X_2 \rangle \\ &= \mathbf{C}\langle \ln X_1, \ln X_2 \rangle \\ &= \mathbf{C}\langle X_1, \exp X_2 \rangle \\ &= \mathbf{C}\langle \sqrt{X_1}, \exp X_2 \rangle \end{aligned} \quad (1.68)$$

La aplicación de transformaciones crecientes no modifica la copula, solamente los márgenes. Así la copula de la distribución log normal bivariada es la misma que aquella de la distribución normal bivariada.

1.6.4 Dependencia y correlación

Podríamos interpretar la copula del vector aleatorio $\mathbf{X} = (X_1, X_2)$ como una re-parametrización o una normalización de la distribución conjunta F luego de eliminar los efectos de los márgenes F_1 y F_2 . Cuando las copulas fueron descubiertas se les dio el nombre de función de dependencia. La copula es de hecho la función de dependencia de \mathbf{X} . Además, es una estadística exhaustiva de esta dependencia. Consideremos una estadística no exhaustiva de la dependencia, es decir una medida de dependencia. Esta será una medida de dependencia si se explica exclusivamente a partir de la copula.

Entre estas estadísticas, podemos considerar de manera general las medidas de concordancia.

Definición: Una medida numérica κ de asociación entre dos variables aleatorias continuas X_1 y X_2 cuya copula es C , es una medida de concordancia si esta satisface las siguientes propiedades:

1. κ es definida para cada par de variables aleatorias continuas (X_1, X_2)
2. $-1 = \kappa_{X, -X} \leq \kappa_C \leq \kappa_{X, X} = 1$;
3. $\kappa_{X_1, X_2} = \kappa_{X_2, X_1}$;
4. Si X_1 y X_2 son independientes, entonces $\kappa_{X_1, X_2} = \kappa_{C^\perp} = 0$;
5. $\kappa_{-X_1, X_2} = \kappa_{X_1, -X_2} = -\kappa_{X_1, X_2}$;
6. Si $C_1 \prec C_2$, entonces $\kappa_{C_1} \leq \kappa_{C_2}$;
7. Si $\{(X_{1,n}, X_{2,n})\}$ es una secuencia de variables aleatorias continuas cuya copula es C_n y si $\{C_n\}$ converge hacia C (pointwise convergence), entonces $\lim_{n \rightarrow \infty} \kappa_{C_n} = \kappa_C$.

Entre todas las medidas de convergencia, hay tres medidas muy populares que tienen un rol importante en la estadística no paramétrica: El tau de Kendall, la rho de Spearman y el índice de Gini. Las tres se pueden explicar por medio de copulas:

$$\begin{aligned} \tau &= 4 \iint_{I^2} C(u_1, u_2) dC(u_1, u_2) - 1 \\ \rho &= 12 \iint_{I^2} C u_1, u_2 dC(u_1, u_2) - 3 \\ \gamma &= 2 \iint_{I^2} (|u_1 + u_2 - 1| - |u_1 - u_2|) dC(u_1, u_2) \end{aligned} \quad (1.69)$$

Observación: Consideremos una muestra de n observaciones $\{(x_1, y_1), \dots, (x_n, y_n)\}$ del vector aleatorio (X, Y) . La tau de Kendall es la probabilidad de concordancia $(X_i - X_j)(Y_i - Y_j) > 0$ menos la probabilidad de

no concordancia de pares $(X_i - X_j)(Y_i - Y_j) < 0$. La rho de Spearman es la correlación de las estadísticas de rango R_X y R_Y :

$$\rho = \text{cor}(\mathbf{F}_X(X), \mathbf{F}_Y(Y)) \quad (1.70)$$

Consideremos el coeficiente de correlación lineal de Pearson. ¿Es una medida de concordancia?; recordemos que la expresión de este coeficiente es:

$$\rho\langle X_1, X_2 \rangle = \frac{\mathbb{E}[X_1 X_2] - \mathbb{E}[X_1]\mathbb{E}[X_2]}{\sigma[X_1]\sigma[X_2]} \quad (1.71)$$

Usando las funciones super-aditivas podemos mostrar los resultados siguientes:

- Si la copula de (X_1, X_2) es \mathbf{C}^\perp , entonces $\rho\langle X_1, X_2 \rangle = 0$;
- ρ es una medida creciente con relación al orden de concordancia
 $\mathbf{C}_1 \succ \mathbf{C}_2 \Rightarrow \rho_1\langle X_1, X_2 \rangle \geq \rho_2\langle X_1, X_2 \rangle \quad (1.72)$
- $\rho\langle X_1, X_2 \rangle$ está delimitada
 $\rho^-\langle X_1, X_2 \rangle \leq \rho\langle X_1, X_2 \rangle \leq \rho^+\langle X_1, X_2 \rangle \quad (1.73)$
 y los bornes alcanzados por las copulas de Frechet son \mathbf{C}^- y \mathbf{C}^+ .

Es evidente que las propiedades (1), (3), (5) y (7) han quedado verificadas. Para la propiedad (5) tenemos:

$$\begin{aligned} \rho\langle -X_1, X_2 \rangle &= \frac{\mathbb{E}[-X_1 X_2] - \mathbb{E}[-X_1]\mathbb{E}[X_2]}{\sigma[-X_1]\sigma[X_2]} \\ &= \frac{\mathbb{E}[-X_1 X_2] - \mathbb{E}[X_1]\mathbb{E}[-X_2]}{\sigma[X_1]\sigma[-X_2]} \\ &= \rho\langle X_1, -X_2 \rangle \\ &= \frac{\mathbb{E}[X_1 X_2] - \mathbb{E}[X_1]\mathbb{E}[X_2]}{\sigma[X_1]\sigma[X_2]} \\ &= -\rho\langle X_1, X_2 \rangle \end{aligned} \quad (1.74)$$

Sin embargo, si las propiedades (1) - (7) son verificadas, tendremos:

$$\kappa_{\mathbf{C}^-} = -1 \quad (1.75)$$

y

$$\kappa_{\mathbf{C}^+} = -1 \quad (1.76)$$

Si usamos la interpretación de límites de Frechet en términos de variables aleatorias, tendremos:

$$\rho^-(X_1, X_2) = \rho^+(X_1, X_2) = \frac{\mathbb{E}[f_1(X)f_2(X)] - \mathbb{E}[f_1(X)]\mathbb{E}[f_2(X)]}{\sigma[f_1(X)]\sigma[f_2(X)]} \quad (1.77)$$

La solución a la ecuación $\rho^-(X_1, X_2) = -1$ es:

$$f_1(x) = a_1x + b \quad (1.78)$$

y

$$f_2(x) = a_2x + b \quad (1.79)$$

con $a_1a_2 < 0$. Para la ecuación $\rho^+(X_1, X_2) = 1$, la condición es $a_1a_2 > 0$. Aparte de las variables aleatorias gaussianas, hay pocas variables aleatorias con capacidad de satisfacer estas condiciones. Observemos también que el coeficiente de correlación de Pearson es una medida de concordancia, entonces este no debe variar por transformaciones crecientes:

$$\rho(X_1, X_2) = \rho\langle f_1(X_1), f_2(X_2) \rangle \quad (1.80)$$

Una vez más, la solución de esta ecuación es $f_1(x) = a_1x + b$ y $f_2(x) = a_2x + b$ con $a_1 > 0$ y $a_2 > 0$. Ahora podemos entender mejor porque a esta medida se le llama lineal.

A modo de ilustración, consideremos el modelo de Black & Scholes con dos activos. Los precios de los activos $S_1(t)$ y $S_2(t)$ son procesos de difusión cuya representación es:

$$\begin{cases} dS_1(t) = u_1S_1(t)dt + \sigma_1S_1(t)dW_1(t) \\ dS_2(t) = u_2S_2(t)dt + \sigma_2S_2(t)dW_2(t) \end{cases} \quad (1.81)$$

con $W_1(t)$ y $W_2(t)$ dos movimientos brownianos- \mathcal{F}_t :

$$\mathbb{E}[W_1(t)W_2(t)|\mathcal{F}_{t_0}] = \rho(t-t_0) \quad (1.82)$$

Podemos mostrar que

$$\rho^- \leq \rho\langle S_1(t), S_2(t) \rangle \leq \rho^+ \quad (1.83)$$

con

$$\rho^\pm = \frac{\exp(\pm\sigma_1\sigma_2(t-t_0)) - 1}{\sqrt{\exp(\sigma_1^2(t-t_0)) - 1} \cdot \sqrt{\exp(\sigma_2^2(t-t_0)) - 1}} \quad (1.84)$$

$\rho(S_1(t), S_2(t)) = \rho^-$ (resp. ρ^+) sí y solo si $C\langle S_1(t), S_2(t) \rangle = C^-$ o todavía $\rho\langle W_1(t), W_2(t) \rangle = -1$ (resp. $C\langle S_1(t), S_2(t) \rangle = C^+$ o todavía $\rho\langle W_1(t), W_2(t) \rangle = +1$).

Definición:

Una medida numérica δ de asociación entre dos variables aleatorias continuas X_1 y X_2 cuya copula es C , es una medida de dependencia si satisface las siguientes propiedades:

1. δ se define para todo par (X_1, X_2) de variables aleatorias continuas;
2. $0 = \delta_{C^\perp} \leq \delta_C \leq \delta_{C^+} = 1$;
3. $\delta_{X_1, X_2} = \delta_{X_2, X_1}$;
4. $\delta_{X_1, X_2} = \delta_{C^\perp} = 0$ sí y solo si X_1 y X_2 son independientes;
5. $\delta_{X_1, X_2} = \delta_{C^+} = \delta_C = 1$ sí y solo si X_1 es una función estrictamente monótona de X_2 ;
6. Si h_1 y h_2 son funciones estrictamente monótonas respectivamente sobre $\text{Im } X_1$ y $\text{Im } X_2$ entonces $\delta_{h_1(X_1), h_2(X_2)} = \delta_{X_1, X_2}$;
7. Si $\{(X_{1,n}, X_{2,n})\}$ es una secuencia de variables aleatorias continuas cuya copula es C_n , y si $\{C_n\}$ converge hacia C (pointwise convergence), entonces $\lim_{n \rightarrow \infty} \delta_{C_n} = \delta_C$

Podemos verificar estas medidas a partir de los trabajos de Schweizer y Wolff que verifican esta definición:

$$\begin{aligned} \sigma &= 12 \iint_{I^2} |C(u_1, u_2) - C^\perp(u_1, u_2)| du_1 du_2 \\ \Phi^2 &= 90 \iint_{I^2} |C(u_1, u_2) - C^\perp(u_1, u_2)|^2 du_1 du_2 \end{aligned} \quad (1.85)$$

Observación:

Podemos interpretar las medidas de concordancia como una generalización de la correlación lineal en el caso que la distribución bivariada no sea gaussiana. Estas toman el valor de -1 para la estructura de dependencia más negativa C^- (la misma que corresponde a una correlación lineal igual a -1 en el caso gaussiano) y $+1$ para la estructura de dependencia más positiva C^+ (la misma que corresponde a una correlación lineal igual a $+1$ en el caso gaussiano). Estas dos copulas representan la forma de dependencia más extrema, es decir que las variables aleatorias X_1 y X_2 son perfectamente dependientes debido a que

$$X_2 = f(X_1) \quad (1.86)$$

con una función determinista. Sin embargo, estas medidas de dependencia solo tomarán el valor de 1 si la dependencia es monótona. Entonces, existen numerosas copulas que corresponden a dependencias perfectas y cuyo δ es estrictamente inferior a 1 (las copulas más conocidas son las “shuffles of min”).

Observación:

Para la familia de copulas Normales, verificamos que

$$\delta\langle C_\rho \rangle = \delta\langle C_{-\rho} \rangle \quad (1.87)$$

Observación:

$\delta\langle C \rangle = 0$ sí y solo sí $C = C^\perp$. Esta medida no está verificada por las medidas de concordancia y la correlación lineal.

1.6.5 El caso multidimensional

El caso multidimensional presenta algunas dificultades con relación al caso bidimensional. La principal dificultad es la noción de dependencia negativa. Por ejemplo, ¿cómo podemos caracterizar la dependencia negativa extrema? Tomemos por ejemplo el caso de un vector gaussiano (X_1, X_2, X_3) . Supongamos que X_1 está perfectamente anti correlacionado con X_2 y que X_2 esté perfectamente anti correlacionado con X_3 . Entonces X_1 está perfectamente correlacionado con X_3 . La expresión de la matriz de correlación será:

$$\rho = \begin{pmatrix} 1 & -1 & 1 \\ & 1 & -1 \\ & & 1 \end{pmatrix} \quad (1.88)$$

Esto significa que

$$\begin{pmatrix} 1 & -1 & -1 \\ & 1 & -1 \\ & & 1 \end{pmatrix} \notin \rho\langle X_1, X_2, X_3 \rangle \quad (1.89)$$

Esto explica que el límite mínimo de Frechet no existe (en el caso general). Igual si tenemos

$$C^- \prec C \prec C^+ \quad (1.90)$$

con

$$C^-(u_1, \dots, u_n) = \max\left(\sum_{i=1}^n u_i - n + 1, 0\right) \quad (1.91)$$

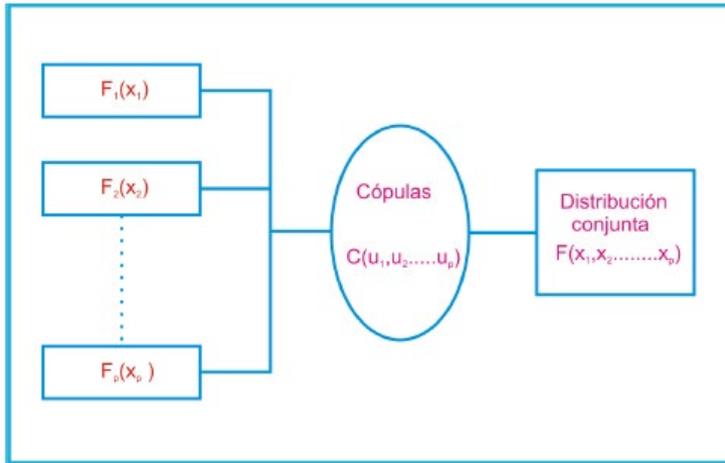
y

$$C^+(u_1, \dots, u_n) = \min(u_1, \dots, u_n) \quad (1.92)$$

C^- no es una copula.

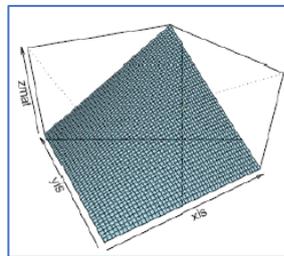
1.6.6 Familias de copulas

Enlaces entre leyes marginales, copulas y distribución conjunta

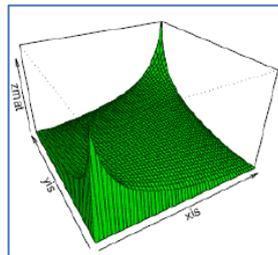


Ya hemos visto que la copula Normal se define de la manera siguiente:

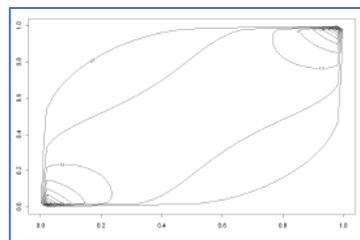
Función de distribución de la copula Normal



Función de densidad de la copula Normal



Contorno de la copula Normal



$$C(u_1, u_2; \rho) = \Phi(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \rho) \quad (1.93)$$

La copula asociada a la distribución bivariada de student es:

$$C_v(u_1, u_2; \rho) = t_v(t_v^{-1}(u_1), t_v^{-1}(u_2); \rho) \quad (1.94)$$

La copula Frank corresponde a:

$$C(u_1, u_2; \theta) = \frac{1}{\theta} \ln \left(1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1} \right) \quad (1.95)$$

con $\theta \in \mathbb{R}$

Observación:

Estas tres familias de copulas son llamadas familias comprensivas (comprehensive family) ya que ellas contienen las tres copulas C^- , C^\perp y C^+ .

Algunas familias de copulas presentan estructuras de dependencia más limitadas. Este es el caso de la copula Gumbel $\theta \geq 0$ definida por:

$$C(u_1, u_2; \theta) = \exp \left(- \left[(-\ln u_1)^\theta + (-\ln u_2)^\theta \right]^{\frac{1}{\theta}} \right) \quad (1.96)$$

que verifica la siguiente desigualdad

$$C^\perp \prec C \prec C^+ \quad (1.97)$$

La copula FGM (Farlie, Gumbel, Morgenstern) tiene la siguiente expresión

$$C(u_1, u_2; \theta) = u_1 u_2 + \theta u_1 u_2 (1 - u_1)(1 - u_2) \quad (1.98)$$

con $\theta \in [-1, 1]$. Esta copula es una perturbación de la copula producto y verificamos que

$$-\frac{1}{3} \leq \rho \langle C \rangle \leq \frac{1}{3} \quad (1.99)$$

La rho de Spearman toma valores en un intervalo bastante restringido. Por lo que no podemos usar esta copula para modelar variables aleatorias que tengan una fuerte dependencia.

También podemos construir distribuciones bivariadas que son aún menos comparables. Igual si las copulas Normales, de Frank o de Gumbel son diferentes (en términos de comportamientos extremos), desde ciertos puntos de vista son similares. De hecho, el conjunto de estas tres copulas es muy pequeño con respecto al conjunto de copulas \mathcal{C} . Existen entonces estructuras de dependencia que no se parecen mucho a las tres copulas precedentes. Para explicar esto consideremos la región alcanzable $\mathcal{B}(\tau, \rho)$ por los estadísticos

(τ, ρ) que es definida por

$$(\tau, \rho) \in \mathcal{B}(\tau, \rho) \Leftrightarrow \begin{cases} \frac{3\tau - 1}{2} \leq \rho \leq \frac{1 + 2\tau - \tau^2}{2} & \tau \geq 0 \\ \frac{\tau^2 + 2\tau - 1}{2} \leq \rho \leq \frac{1 + 3\tau}{2} & \tau \leq 0 \end{cases} \quad (1.100)$$

Podemos ver que los límites no pueden ser mejorados y que siempre hay una copula que corresponde a un punto de la frontera $\mathcal{B}(\tau, \varrho)$.

1.7 Copulas de valores extremos

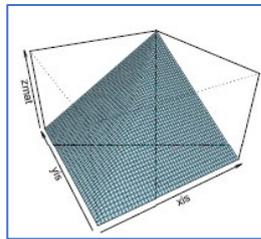
Definición: Una copula de valores extremos (extreme value copula) verifica la relación siguiente:

$$C(u_1^t, \dots, u_n^t) = C^t(u_1, \dots, u_n) \quad (1.101)$$

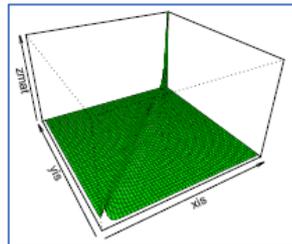
para todo real positivo.

Por ejemplo, la copula Gumbel es una copula de valores extremos:

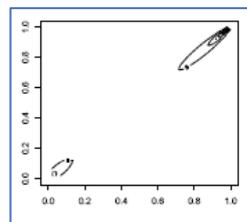
Función de distribución de la copula de Gumbel



Función de densidad de la copula de Gumbel



Líneas de nivel de la copula de Gumbel



$$\begin{aligned}
\mathbf{C}(u_1^t, u_2^t; \theta) &= \exp\left(-\left[(-\ln u_1^t)^\theta + (-\ln u_2^t)^\theta\right]^{\frac{1}{\theta}}\right) \\
&= \exp\left(-\left(t^\theta \left[(-\ln u_1)^\theta + (-\ln u_2)^\theta\right]\right)^{\frac{1}{\theta}}\right) \\
&= \left[\exp\left(-\left[(-\ln u_1)^\theta + (-\ln u_2)^\theta\right]^{\frac{1}{\theta}}\right)\right]^t \\
&= \mathbf{C}^t(u_1, u_2; \theta)
\end{aligned} \tag{1.102}$$

que no es el caso de la copula FMG:

$$\begin{aligned}
\mathbf{C}(u_1^t, u_2^t; \theta) &= u_1^t u_2^t + \theta u_1^t u_2^t (1 - u_1^t)(1 - u_2^t) \\
&= u_1^t u_2^t (1 + \theta - \theta u_1^t - \theta u_2^t + \theta u_1^t u_2^t) \\
&\neq u_1^t u_2^t (1 + \theta - \theta u_1 - \theta u_2 + \theta u_1 u_2)^t \\
&\neq \mathbf{C}^t(u_1, u_2; \theta)
\end{aligned} \tag{1.103}$$

El nombre “extreme value copula” sugiere una relación entre la teoría de valores extremos y estas copulas. Consideremos el vector \mathcal{X}_n^+ de las estadísticas de orden máximo. Tendremos

$$\mathcal{X}_n^+ = \begin{bmatrix} X_{n:n}^{(1)} \\ \vdots \\ X_{n:n}^{(m)} \end{bmatrix} \tag{1.104}$$

Denominemos $\mathbf{F}_{n:n}$ a la distribución de \mathcal{X}_n^+ definida por:

$$\begin{aligned}
\mathbf{F}_{n:n}(\mathbf{x}) &= \Pr\{\mathcal{X}_n^+ \leq \mathbf{x}\} \\
&= \Pr\{X_{n:n}^{(1)} \leq x^{(1)}, \dots, X_{n:n}^{(m)} \leq x^{(m)}\}
\end{aligned} \tag{1.105}$$

La teoría multidimensional de valores extremos se interesa en la ley límite

$$\begin{aligned}
\lim_{n \rightarrow \infty} \Pr\left(\frac{X_{n:n}^{(1)} - b_n^{(1)}}{a_n^{(1)}} \leq x^{(1)}, \dots, \frac{X_{n:n}^{(m)} - b_n^{(m)}}{a_n^{(m)}} \leq x^{(m)}\right) &= \mathbf{G}(x^{(1)}, \dots, x^{(m)}) \\
&= \mathbf{G}(\mathbf{x})
\end{aligned} \tag{1.106}$$

Luego, la representación canónica de la distribución \mathbf{G} :

$$\mathbf{G}(x^{(1)}, \dots, x^{(m)}) = \mathbf{C}\langle \mathbf{G} \rangle(\mathbf{G}_1(x^{(1)}), \dots, \mathbf{G}_m(x^{(m)})) \tag{1.107}$$

es evidente que los márgenes de $\mathbf{G}(\mathbf{x})$ verifican el teorema de Fisher-Tippet. Entonces, los márgenes de una distribución multidimensional de valores

extremos son de Gumbel, y/o de Frechet, y/o de Weibull. En lo que concierne a la copula $C\langle G \rangle$ tenemos el siguiente resultado:

Teorema: $C\langle G \rangle$ es una copula de valores extremos.

La representación copula permite caracterizar fácilmente las distribuciones multidimensionales de valores extremos. Tomemos el ejemplo del vector aleatorio $(X^{(1)}, X^{(2)})$ cuya distribución es

$$F(x_1, x_2) = \exp\left(-\left[(-\ln \Phi(x_1))^\theta + (-\ln x_2)^\theta\right]^{\frac{1}{\theta}}\right) \quad (1.108)$$

$X^{(1)}$ es entonces una variable aleatoria gaussiana y $X^{(2)}$ es una variable aleatoria uniforme. La distribución G entonces es:

$$G(x^{(1)}, x^{(2)}) = \exp\left(-\left[(-\ln \wedge(x^{(1)}))^\theta + (-\ln \psi(x^{(2)}))^\theta\right]^{\frac{1}{\theta}}\right) \quad (1.109)$$

Para construir una distribución multidimensional de valores extremos, es suficiente con enganchar los márgenes GEV con una copula de valores extremos. Por ejemplo, con una distribución como la siguiente:

$$G(x^{(1)}, x^{(2)}) = \exp\left(-\left[(-\ln \mathcal{G}\mathcal{E}\mathcal{V}(0,1,1))^\theta + (-\ln \mathcal{G}\mathcal{E}\mathcal{V}(0,1,1.2))^\theta\right]^{\frac{1}{\theta}}\right) \quad (1.110)$$

En el caso bivariado, podemos demostrar que la expresión de la copula $C\langle G \rangle$ es:

$$C\langle G \rangle(u_1, u_2) = \exp\left[\ln u_1 u_2 A\left(\frac{\ln u_1}{\ln(u_1 u_2)}\right)\right] \quad (1.111)$$

con A como una función convexa que verifica $A(0) = A(1) = 1$ y $\max(w, 1-w) \leq A(w) \leq 1$. Ahora es fácil de mostrar que

$$\begin{aligned}
& \max(w, 1-w) \leq A(w) \leq 1 \\
& \Leftrightarrow \max\left(\frac{\ln u_1}{\ln(u_1 u_2)}, \frac{\ln u_2}{\ln(u_1 u_2)}\right) \leq A\left(\frac{\ln u_1}{\ln(u_1 u_2)}\right) \leq 1 \\
& \Leftrightarrow \min(\ln u_1, \ln u_2) \geq \ln(u_1 u_2) A\left(\frac{\ln u_1}{\ln(u_1 u_2)}\right) \geq \ln(u_1 u_2) \\
& \Leftrightarrow \min(u_1, u_2) \geq \exp\left[\ln(u_1 u_2) A\left(\frac{\ln u_1}{\ln(u_1 u_2)}\right)\right] \geq u_1 u_2 \\
& \Leftrightarrow \mathbf{C}^+ \succ \mathbf{C}\langle \mathbf{G} \rangle \succ \mathbf{C}^\perp \tag{1.112}
\end{aligned}$$

ya que $\ln(u_1 u_2) \leq 0$

por tanto

$$\mathbf{C}^\perp \prec \mathbf{C}\langle \mathbf{G} \rangle \prec \mathbf{C}^+ \tag{1.113}$$

Los extremos son siempre positivamente dependientes. Y esta propiedad se generaliza fácilmente al caso multidimensional.

Podemos distinguir tres situaciones:

1. $\mathbf{C}\langle \mathbf{G} \rangle = \mathbf{C}^\perp$;
2. $\mathbf{C}\langle \mathbf{G} \rangle = \mathbf{C}^+$; y
3. $\mathbf{C}\langle \mathbf{G} \rangle$ es una copula comprendida entre la copula producto y la copula de los límites altos de Fréchet.

La forma cómo podemos determinar a priori que caso verifica la copula $\mathbf{C}\langle \mathbf{G} \rangle$ es a través de la dependencia de la cola.

Definición:

Si una copula bivariada \mathbf{C} es tal que el límite

$$\lim_{u \rightarrow 1} \frac{\bar{\mathbf{C}}(u, u)}{1-u} = \lim_{u \rightarrow 1} \frac{1-2u + \mathbf{C}(u, u)}{1-u} = \lambda \tag{1.114}$$

existe, entonces \mathbf{C} tiene una dependencia de cola si $\lambda \in (0, 1]$ y no tiene dependencia de cola si $\lambda = 0$.

Por la copula producto \mathbf{C}^\perp , tenemos:

$$\begin{aligned}
\lim_{u \rightarrow 1} \frac{\bar{\mathbf{C}}(u, u)}{1-u} &= \lim_{u \rightarrow 1} \frac{1-2u + u^2}{1-u} \\
&= \lim_{u \rightarrow 1} (1-u) \\
&= 0 \tag{1.115}
\end{aligned}$$

Por la copula C^+ , tenemos:

$$\begin{aligned}\lim_{u \rightarrow 1} \frac{\bar{C}(u, u)}{1-u} &= \lim_{u \rightarrow 1} \frac{1-2u + \min(u, u)}{1-u} \\ &= \lim_{u \rightarrow 1} \frac{1-2u+u}{1-u} \\ &= 1\end{aligned}\quad (1.116)$$

Por la copula C^- , la dependencia de cola es nula:

$$\begin{aligned}\lim_{u \rightarrow 1} \frac{\bar{C}(u, u)}{1-u} &= \lim_{u \rightarrow 1} \frac{1-2u + \max(2u-1, 0)}{1-u} \\ &= 0\end{aligned}\quad (1.117)$$

Tenemos una relación muy interesante entre la dependencia de cola de C y aquella de la copula de valores extremos $C\langle G \rangle$:

Teorema: La dependencia de cola de $C\langle G \rangle$ es igual a la dependencia de cola de C .

Observación: Los extremos son entonces independientes si la copula no tiene dependencia de cola.

Observación: En el caso unidimensional, la caracterización de los extremos depende de la cola de distribución. En el caso multidimensional, ella depende de la dependencia de las colas de distribución.

La medida λ se puede interpretar en el marco del valor en riesgo. Consideremos la probabilidad condicional que sobrepasamos el valor en riesgo en una dirección sabiendo que ya lo hemos pasado en la otra dirección:

$$\begin{aligned}\lambda(\alpha) &= \Pr\{X_2 > VaR_2(\alpha) | X_1 > VaR_1(\alpha)\} \\ &= \Pr\{X_2 > F_2^{-1}(\alpha) | X_1 > F_1^{-1}(\alpha)\} \\ &= \frac{\Pr\{F_2(X_2) > \alpha, F_1(X_1) > \alpha\}}{\Pr\{F_1(X_1) > \alpha\}} \\ &= \frac{1-2\alpha + F(F_1^{-1}(\alpha)), F_2^{-1}(\alpha)}{1-\alpha} \\ &= \frac{1-2\alpha + C(\alpha, \alpha)}{1-\alpha}\end{aligned}\quad (1.118)$$

Verificamos que:

$$\lambda = \lim_{\alpha \rightarrow 1^+} \lambda(\alpha)\quad (1.119)$$

Dicho de otra forma, λ es la probabilidad de observar un extremo en una dirección sabiendo que ya tenemos un extremo en otra dirección.

Para demostrar la importancia de la dependencia de cola, consideramos los valores tomados por $\lambda(\alpha)$ para las copulas Normal y de Student. Para la copula Normal, tenemos $\lambda = 0$ si $\rho < 1$, y para la de Student, $\lambda \neq 0$ si $\rho > -1$.

1.8 Modelación de los dominios de atracción

Sea F la distribución multidimensional cuyos márgenes son F_1, \dots, F_m y cuya copula es C . Sea G la distribución multidimensional de los valores extremos. Observamos que G_1, \dots, G_m son los márgenes de G , y $C\langle G \rangle$ su correspondiente copula. Tenemos entonces el teorema siguiente:

Teorema: $F \in MDA(G)$ sí y solo si (MDA=máximo dominio de atracción)

1. $F_i \in MDA(G_i)$ para todo $i = 1, \dots, m$;
2. $C \in MDA(C\langle G \rangle)$

Ya vimos como modelar el máximo dominio de atracción en el caso unidimensional y también vimos como calcular los coeficientes de normalización. Estos coeficientes de normalización son válidos en el caso multidimensional. Entonces, la única dificultad es la determinación de $C\langle G \rangle$:

$$\begin{aligned} \lim_{n \rightarrow \infty} \Pr \left(\frac{X_{n:n}^{(1)} - b_n^{(1)}}{a_n^{(1)}} \leq x^{(1)}, \dots, \frac{X_{n:n}^{(m)} - b_n^{(m)}}{a_n^{(m)}} \leq x^{(m)} \right) \\ = C\langle G \rangle(G_1(x^{(1)}), \dots, G_m(x^{(m)})) \end{aligned} \quad (1.120)$$

Teorema: $C \in MDA(C\langle G \rangle)$ si C satisface la siguiente relación:

$$\lim_{t \rightarrow \infty} C^t \left(u_1^t, \dots, u_m^t \right) = C\langle G \rangle(u_1, \dots, u_m) \quad (1.121)$$

Observación: si C es una copula de valores extremos, $C \in MDA(C)$. En efecto, tenemos:

$$\begin{aligned} \lim_{t \rightarrow \infty} C^t \left(u_1^t, \dots, u_m^t \right) &= \lim_{t \rightarrow \infty} C \left(\left(u_1^t \right)^t, \dots, \left(u_m^t \right)^t \right) \\ &= \lim_{t \rightarrow \infty} C(u_1, \dots, u_m) \\ &= C(u_1, \dots, u_m) \end{aligned} \quad (1.122)$$

En el caso bidimensional, este teorema se puede escribir con la función convexa A :

Teorema: $C \in MDA(C\langle G \rangle)$ con

$$\mathbf{C}\langle \mathbf{G} \rangle(u_1, u_2) = \exp \left[\ln(u_1, u_2) A \left(\frac{\ln u_1}{\ln(u_1, u_2)} \right) \right] \quad (1.123)$$

si \mathbf{C} satisface la siguiente relación:

$$\lim_{u \rightarrow 0} \frac{1 - \mathbf{C}((1-u)^{1-t}, (1-u)^t)}{u} = A(t) \quad (1.124)$$

para todo $t \in [0, 1]$.

Consideremos un ejemplo con la distribución de $(X^{(1)}, X^{(2)})$ definida por:

$$\mathbf{F}(x_1, x_2) = \left((1 - e^{-x_1})^{-\theta} + x_2^{-\theta} - 1 \right)^{-\frac{1}{\theta}} \quad (1.125)$$

sobre $[0, \infty] \times [0, 1]$. Los márgenes de $\mathbf{F}(x_1, x_2)$ entonces son:

$$\begin{aligned} \mathbf{F}_1(x_1) &= \mathbf{F}(x_1, 1) \\ &= 1 - e^{-x_1} \end{aligned} \quad (1.126)$$

y

$$\begin{aligned} \mathbf{F}_2(x_2) &= \mathbf{F}(\infty, x_2) \\ &= x_2 \end{aligned} \quad (1.127)$$

$X^{(1)}$ es entonces una variable aleatoria exponencial y $X^{(2)}$ es una variable aleatoria uniforme. Busquemos ahora modelar la distribución bidimensional de los extremos:

$$\lim_{n \rightarrow \infty} \Pr \left(\frac{X_{n:n}^{(1)} - b_n^{(1)}}{a_n^{(1)}} \leq x^{(1)}, \frac{X_{n:n}^{(2)} - b_n^{(2)}}{a_n^{(2)}} \leq x^{(2)} \right) = \mathbf{C}\langle \mathbf{G} \rangle(\mathbf{G}_1(x^{(1)}), \mathbf{G}_2(x^{(2)})) \quad (1.128)$$

Sabemos que

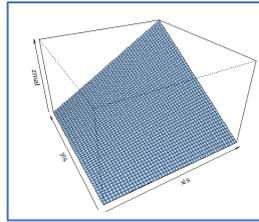
$$\lim_{n \rightarrow \infty} \Pr \left(\frac{X_{n:n}^{(1)} - \ln n}{1} \leq x^{(1)} \right) = \wedge(x^{(1)}) \quad (1.129)$$

y

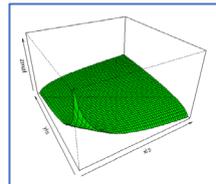
$$\lim_{n \rightarrow \infty} \Pr \left(\frac{X_{n:n}^{(2)} - 1}{n^{-1}} \leq x^{(2)} \right) = \psi_1(x^{(2)}) \quad (1.130)$$

La copula de \mathbf{F} es la copula de Clayton:

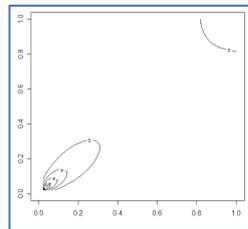
Función de distribución de la copula de Clayton



Función de densidad de la copula de Clayton



Líneas de nivel de la copula de Clayton



$$C(u_1, u_2) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-\frac{1}{\theta}} \quad (1.131)$$

Tenemos

$$\begin{aligned} \lim_{u \rightarrow 0} \frac{1 - C((1-u)^t, (1-u)^{1-t})}{u} &= \lim_{u \rightarrow 0} \frac{1 - \left((1-u)^{-\theta t} + (1-u)^{-\theta(1-t)} - 1 \right)^{-\frac{1}{\theta}}}{u} \\ &= \lim_{u \rightarrow 0} \frac{1 - (1 + \theta u + o(u))^{-\frac{1}{\theta}}}{u} \\ &= \lim_{u \rightarrow 0} \frac{u + o(u)}{u} \\ &= 1 \end{aligned} \quad (1.132)$$

$A(t)$ es entonces igual a 1, y deducimos que:

$$C\langle \mathbf{G} \rangle = C^\perp \quad (1.133)$$

Finalmente, obtenemos

$$\begin{aligned}
\mathbf{G}(x^{(1)}, x^{(2)}) &= \lim_{n \rightarrow \infty} \Pr\left(X_{n:n}^{(1)} - \ln n \leq x^{(1)}, n\left(X_{n:n}^{(2)} - 1\right) \leq x^{(2)}\right) \\
&= \wedge(x^{(1)}) \times \psi_1(x^{(2)}) \\
&= \exp\left(-e^{-x^{(1)}}\right) \exp\left(x^{(2)}\right)
\end{aligned} \tag{1.134}$$

Si cambiamos la copula de \mathbf{F} , solo se modificará la copula $\mathbf{C}\langle\mathbf{G}\rangle$. Por ejemplo, si \mathbf{C} es la copula Normal de parámetro $\rho < 1$, entonces

$$\mathbf{G}(x^{(1)}, x^{(2)}) = \exp\left(-e^{-x^{(1)}}\right) \exp\left(x^{(2)}\right) \tag{1.135}$$

Si \mathbf{C} es la copula Normal de parámetro $\rho = 1$, tendremos

$$\mathbf{G}(x^{(1)}, x^{(2)}) = \min\left(\exp\left(-e^{-x^{(1)}}\right), \exp\left(x^{(2)}\right)\right) \tag{1.136}$$

Si \mathbf{C} es la copula de Gumbel, tendremos

$$\mathbf{G}(x^{(1)}, x^{(2)}) = \exp\left(-\left[\exp\left(-\theta x^{(1)}\right) + \left(-x^{(2)}\right)^\theta\right]^{\frac{1}{\theta}}\right) \tag{1.137}$$

1.9 Copulas Archimax

Consideremos una nueva familia de copulas que aglutina la mayor parte de las familias conocidas de copulas, especialmente las copulas Arquimedianas y todas las copulas de valores extremos. Esta nueva familia nos proporciona una mayor flexibilidad para su modelamiento. A priori podemos conocer los diferentes dominios máximos de atracción.

Definición: Una función bivariada es una copula Archimax sí y solo si ella es de la forma

$$\mathbf{C}_{\phi, A}(x, y) = \phi^{-1}\left[\left(\phi(x) + \phi(y)\right) A\left(\frac{\phi(x)}{\phi(x) + \phi(y)}\right)\right] \tag{1.138}$$

para todo $0 \leq x, y \leq 1$, con:

1. $A: [0, 1] \rightarrow \left[\frac{1}{2}, 1\right]$ tal que $\max(t, 1-t) \leq A(t) \leq 1$ para todo $0 \leq t \leq 1$;
2. $\phi: (0, 1] \rightarrow [0, \infty)$ es una función convexa, decreciente que verifica $\phi(1) = 0$

Adoptaremos la convención $\phi(0) = \lim_{t \rightarrow 0} \phi(t)$ y $\phi^{-1}(s) = 0$ para $s \geq \phi(0)$

Podemos observar rápidamente que esta nueva familia de copulas contiene todas las copulas Arquimedianas y todas las copulas de valores extremos.

Efectivamente, si $\phi(t) = \ln\left(\frac{1}{t}\right)$, obtenemos:

$$\mathbf{C}_{\phi,A}(x,y) = \mathbf{C}_A(x,y) = \exp \left[\ln(xy) A \left(\frac{\ln(y)}{\ln(xy)} \right) \right] \quad (1.139)$$

para todo $0 \leq x, y \leq 1$. Reconocemos acá la forma general de las copulas bivariadas de valores extremos. Si ahora elegimos tomar $A(t) \equiv 1$, encontramos la forma general de las copulas Arquimedianas:

$$\mathbf{C}_{\phi,A}(x,y) = \mathbf{C}_\phi(x,y) = \phi^{-1}(\phi(x) + \phi(y)) \quad (1.140)$$

El interés del modelaje de la fórmula 1.138 está en el hecho de que con una elección cuidadosa de ϕ y A , obtenemos una copula que estará dentro del máximo dominio de atracción de cualquier "atractor" \mathbf{C}_A dado.

Proposición: Sea $\mathbf{C}_{\phi,A}$ una copula Archimax tal que $\phi\left(1 - \frac{1}{t}\right) \in \text{RV}_{-m}$ para un cierto $m \geq 1$. Entonces $\mathbf{C}_{\phi,A}$ pertenece a un dominio-max de atracción de la copula \mathbf{C}_{A^*} donde la función A^* está dada por:

$$A^*(t) = \left[t^m + (1-t)^m \right]^{1/m} A^{1/m} \left(\frac{t^m}{t^m + (1-t)^m} \right) \quad (1.141)$$

para todo $t \in [0,1]$. Este atractor puede ser considerado desde dos puntos de vista: \mathbf{C}_{A^*} es una copula Archimax de generador A^* y $\phi(t) = \ln\left(\frac{1}{t}\right)$, o, \mathbf{C}_{A^*} es una copula Archimax de generadores A y $\phi^*(t) = \ln^m\left(\frac{1}{t}\right)$. Además, A^* y A coinciden si y solo si $m = 1$.

Observación: Podemos ahora demostrar que la única copula Arquimediana y de valores extremos es la copula de Gumbel.

Razonando en sentido inverso, vemos que $\mathbf{C}_{\phi,A}$ pertenece al dominio-max de atracción de \mathbf{C}_{A^*} si $\phi\left(1 - \frac{1}{t}\right) \in \text{RV}_{-m}$ y

$$A(t) = \left[t^{1/m} + (1-t)^{1/m} \right]^m \left\{ A^* \left(\frac{t^{1/m}}{t^{1/m} + (1-t)^{1/m}} \right) \right\}^m \quad (1.142)$$

Ahora nos falta saber si la relación 1.142 define bien una función de dependencia. Al poner $D(t) = \partial_t \ln(A^*(t))$, $\mathbf{H}^*(t) = t + (1-t)D(t)$ y $h^*(t) = \partial_t \mathbf{H}^*(t)$ encontraremos que la relación 1.142 es verdadera si y solo si

$$\frac{h^*(t)}{\mathbf{H}^*(t)(1-\mathbf{H}^*(t))} \geq \frac{m}{t(1-t)} \quad (1.143)$$

para todo $t \in [0,1]$. De pasada podemos resaltar que esta relación es válida para todo $1 \leq m' \leq m$. Así, si $A(t)$ es definida por la ecuación 1.142, $\mathbf{C}_{\phi,A}$ pertenece al dominio-max de atracción de \mathbf{C}_{A^*} previendo que $\phi\left(1-\frac{1}{t}\right) \in \text{RV}_{-m'}$ para $1 \leq m' \leq m$. En fin, podemos también resaltar que para los generadores A y ϕ cualesquiera tales que $\phi\left(1-\frac{1}{t}\right) \in \text{RV}_{-m}$ tenemos

$$\lambda = \lim_{u \rightarrow 1^-} \frac{1-2u + \mathbf{C}_{\phi,A}(u,u)}{1-u} = 2 - \left\{ 2A\left(\frac{1}{2}\right) \right\}^{\frac{1}{m}} \leq 1 \quad (1.144)$$

De esta manera, las copulas Archimax presentan una dependencia a nivel superior de la cola de distribución salvo si $m=1$ y $A \equiv 1$.

Como ejemplo de copula Archimax podríamos considerar el generador de Tawn que se define por $0 \leq \theta \leq 1$ para

$$A_\theta(t) = \theta t^2 - \theta t + 1 \quad (1.145)$$

para $t \in [0,1]$ y combinarlo al generador de la familia de Clayton ($\eta > 0$)

$$\phi_{1,\eta}(t) = \frac{t^{-\eta} - 1}{\eta} \quad (1.146)$$

O con el generador de copulas de Genest & Ghoudi ($0 < \eta \leq 1$)

$$\phi_{2,\eta}(t) = (1-t^\eta)^{\frac{1}{\eta}} \quad (1.147)$$

Podemos subrayar que $\phi_{2,\eta}(t) \in \text{RV}_{-\frac{1}{\eta}}$. Así, para $\eta \neq 1$, $\mathbf{C}_{\phi_{2,\eta},A}(x,y)$ no se encuentra en el dominio de atracción de \mathbf{C}_A . En fin, supongamos que queremos construir una copula en el dominio-max de atracción de

$$A_{\alpha,\beta}^*(t) = \left[\alpha^k t^k + \beta^k (1-t)^k \right]^{\frac{1}{k}} + \left[(1-\alpha)^k t^k + (1-\beta)^k (1-t)^k \right]^{\frac{1}{k}} \quad (1.148)$$

para $t \in [0,1]$ con $0 \leq \alpha, \beta \leq 1$ y $k \geq 1$. Para los valores $\alpha = \frac{1}{10}$, $\beta = \frac{3}{10}$, $m = \frac{3}{2}$ y $k = 1.6$ se verifica la ecuación 1.143. La copula original A es entonces aquella definida por la ecuación 1.142.

1.10 Modelación financiera con copulas

La modelación financiera consiste en:

- En especificar un modelo para cada factor, y

- En correlacionar los diferentes modelos para obtener una modelación que una los diferentes factores.

Normalmente hacemos estas dos etapas en simultaneo, considerando un modelo multidimensional, que generalmente es gaussiano.

Lo interesante de las copulas es justamente el poder especificar el modelo multidimensional respetando las dos etapas. Hay muchas ventajas cuando procedemos de esta manera:

1. La extensión de modelos unidimensionales es más fácil;
2. La elección de la estructura de dependencia (la copula) puede hacerse independientemente de la elección de los modelos unidimensionales (los márgenes);
3. La representación como copula permite reducir la complejidad de la calibración del modelo multidimensional, ya que los parámetros de la copula y los parámetros de los márgenes pueden estimarse separadamente;
4. En un enfoque directo de modelación del modelo multidimensional, una mala especificación de los márgenes introduce una distorsión sistemática en la estimación de los parámetros de la copula, e inversamente;
5. Por último, queda claro que una representación por copula nos permite comprender mejor cuales son las hipótesis subyacentes del modelo, sobre todo desde la óptica de la gestión de riesgos.

Veamos ahora diferentes ejemplos que nos muestran las dificultades o los errores que podemos encontrar cuando especificamos directamente un modelo multidimensional.

1.- Supongamos que queremos modelar la dependencia conjunta de tres índices CAC40, Dow Jones, y Nikkei. Para un periodo que va del 1/6/2002 al 17/11/2017, la matriz de correlación de rendimientos es

$$\hat{\rho} = \begin{pmatrix} 1 & 0.158 & 0.175 \\ & 1 & 0.0589 \\ & & 1 \end{pmatrix} \quad (1.149)$$

Si suponemos que el vector aleatorio de rendimientos es gaussiano, tendremos

$$C\langle \text{CAC40, NIKKEI} \rangle \succ C\langle \text{CAC40, Dow Jones} \rangle \succ C\langle \text{NIKKEI, Dow Jones} \rangle \quad (1.150)$$

Ahora supongamos que la dependencia es Normal (sin hacer hipótesis sobre los márgenes), y estimamos la matriz de parámetros ρ de la copula normal tridimensional y obtenemos

$$\rho = \begin{pmatrix} 1 & 0.207 & 0.157 \\ & 1 & 0.0962 \\ & & 1 \end{pmatrix} \quad (1.151)$$

Fácilmente podemos verificar que

$$\mathbf{C}\langle \text{CAC40}, \text{Dow Jones} \rangle \succ \mathbf{C}\langle \text{CAC40}, \text{NIKKEI} \rangle \succ \mathbf{C}\langle \text{NIKKEI}, \text{Dow Jones} \rangle \quad (1.152)$$

El orden de dependencia ha cambiado.

2.- Consideremos una distribución hiperbólica generalizada GH. La función de densidad que le corresponde es

$$f(x) = a(\lambda, \alpha, \beta, \delta) \left(\delta^2 + (x - \mu)^2 \right)^{\frac{1}{4}(2\lambda-1)} \\ \times \mathbf{K}_{\lambda-\frac{1}{2}} \left(\alpha \sqrt{\delta^2 + (x - \mu)^2} \right) \exp(\beta(x - \mu)) \quad (1.153)$$

donde \mathbf{K} es la función de Bessel modificada de tercer tipo y

$$f(x) = A(\lambda, \alpha, \beta, \delta, \rho) \left(\delta^2 + (x - \mu)^\top \rho^{-1} (x - \mu) \right)^{\frac{1}{4}(2\lambda-N)} \\ \times \mathbf{K}_{\lambda-\frac{N}{2}} \left(\alpha \sqrt{\delta^2 + (x - \mu)^\top \rho^{-1} (x - \mu)} \right) \exp(\beta^\top (x - \mu)) \quad (1.154)$$

con

$$A(\lambda, \alpha, \beta, \delta, \rho) = \frac{(\alpha^2 - \beta^\top \rho \beta)^{\frac{1}{2}\lambda}}{(2\pi)^{\frac{N}{2}} \alpha^{\lambda-\frac{N}{2}} \delta^\lambda \mathbf{K}_\lambda \left(\delta \sqrt{\alpha^2 - \beta^\top \rho \beta} \right)} \quad (1.155)$$

Esta distribución es poco manejable desde el punto de vista de su implementación, ya que es difícil de simular y la estimación de parámetros es relativamente complicada. Estos problemas se pueden evitar si construimos una distribución hiperbólica generalizada multidimensional a partir de los márgenes hiperbólicos generalizados y de una copula.