

Modelo de cálculo de capital económico por riesgo de crédito para portafolios de créditos a personas físicas

Adán Díaz-Hernández¹
José C. Ramírez-Sánchez²

Recibido 26 de noviembre 2007, Aceptado 4 de enero 2008

Resumen

Este artículo discute una nueva metodología para estimar el capital económico por riesgo de crédito para un portafolio de créditos al menudeo -o de personas físicas-, utilizando los conceptos generales de cópula y medidas de dependencia, así como algunos de los resultados centrales de Teoría de Valores Extremos (TVE). La superioridad del modelo propuesto sobre las técnicas tradicionales de estimación se demuestra al aplicar cópulas Elípticas Generalizadas y cópulas Agrupadas del tipo t de Student para modelar la estructura de dependencia de los parámetros de riesgo PD , EAD y LGD . Además, se utiliza el método POT para analizar el comportamiento de las pérdidas extremas del portafolio.

Abstract

This paper discusses a new methodology to estimate the economic capital by credit risk for a retail portfolio based on the general concepts of copula and dependence measures as well as some core results of the Extreme Value Theory (EVT). The superiority of the proposed approach over the traditional estimation techniques is demonstrated in the application of Elliptical Generalized copulas and Grouped copulas of the t Student type to model the dependence structure of the risk parameters PD , EAD and LGD . Furthermore, the POT method is used to analyze the extreme losses behavior.

Clasificación JEL: C14, C15, C16 y G32

Palabras clave: Capital Económico, Riesgo de Crédito, Cópulas, Valores Extremos

1. Introduction

El capital económico (CE) es el capital en riesgo que los accionistas invierten para limitar, con cierto nivel estadístico de confianza, la probabilidad de incumplimiento de una posición en un horizonte de tiempo dado. A diferencia del capital contable o del regulatorio, el CE provee una estimación del riesgo basada en medidas objetivas, transparentes (o expresadas en unidades monetarias específicas) y comprensivas (al incluir a diferentes clases de riesgos) que permiten utilizar más eficientemente el capital.

¹ Alumno del doctorado en Ciencias Financieras, ITESM-CCM.

² Director Académico de la División de Negocios, ITESM-CCM, josecarlos.ramirez@itesm

En la literatura existe una amplia gama de trabajos que permiten modelar el riesgo de crédito de portafolios de préstamos comerciales -empresas o personas morales- de acuerdo con los lineamientos propuestos por CreditMetrics (CM), KMV PortfolioManager (KMV), CreditRisk+, o CreditPortfolioView (para una revisión de estos modelos ver Crouhy 2000 y Crosbie 2002). Pero esta no parece ser la misma situación para el caso de los préstamos a personas físicas. La escasez de trabajos en este rubro ha obligado a los practicantes a adaptar o modificar los modelos de préstamos comerciales para medir el riesgo de crédito en portafolios de préstamos personales, según sean las necesidades o las disponibilidades de información de las instituciones financieras (ver De Andrade 2004 y Perli 2004). De hecho, gran parte de las reglas que en materia de riesgo de crédito establece Basilea II en su enfoque más avanzado de calificaciones internas (IRB), tienen sus bases sobre ciertos resultados asintóticos para portafolios grandes bajo versiones bastante simplificadas de modelos del tipo KMV/CM.

El objetivo principal de este documento es proponer e implementar una metodología para el cálculo del CE de un portafolio de créditos al menudeo -o de personas físicas-, utilizando como herramientas los conceptos generales de cópula y medidas de dependencia, así como algunos de los resultados centrales de Teoría de Valores Extremos (TVE). La metodología propuesta considera, al igual que los modelos tradicionales y el enfoque regulatorio, que las pérdidas del portafolio atribuidas al riesgo de crédito son determinadas por los parámetros de riesgo de probabilidad de incumplimiento (PD), exposición al incumplimiento (EAD) y pérdida dado el incumplimiento (LGD). Sin embargo, a diferencia de la práctica habitual, el presente trabajo de investigación supone que la relación de dependencia entre los parámetros de riesgo es establecida mediante cópulas multivariadas. En concreto la metodología busca: 1) evaluar el impacto que tienen las reglas de capitalización propuestas en el enfoque más avanzado de IRB sobre los requerimientos de capital del portafolio; 2) medir la efectividad de un grupo alternativo de cópulas (las Elípticas simétricas, las Elípticas Generalizadas, así como las de Mezcla de Normales Agrupadas) frente a otro grupo que ha recibido mayor uso y atención en la literatura como modelos de estructuras de dependencia (la cópula Gaussiana y la de Independencia); y 3) hacer comparaciones con otras metodologías de cálculo de CE consagradas en la práctica. Para analizar adecuadamente el comportamiento de las pérdidas extremas del portafolio, es decir, el área de la cola de la distribución de pérdidas simuladas, se emplea uno de los métodos de la TVE, denominado Peaks Over Threshold (POT), en la estimación del Valor en Riesgo (Var_q), el Déficit Esperado (ES_q) y, posteriormente, el CE del portafolio.

El resto del documento está organizado como sigue. En sección 2 se aborda el concepto de cópula, medidas de dependencia, así como las familias de cópulas específicas que se aplicarán para modelar la estructura de dependencia de los parámetros de riesgo. La estimación de los mismos, así como la composición del portafolio de créditos se presentan en la sección 3. En la sección 4 se presentan los resultados obtenidos de los ajustes de las cópulas utilizadas: Gaussiana, t de Student simétrica, t de Student Generalizada y t de Student Agrupada, las dos últimas, como posibles alternativas a las dos primeras, se implementa un análisis de bondad de ajuste para determinar estadísticamente al mejor de los modelos. En la sección 5, se realiza un ejercicio para comparar la robustez de los resultados de la metodología aquí propuesta contra algunos modelos

alternativos, tales como el de incumplimiento de un sólo factor tipo KMV/CM, un modelo basado en matrices de transición y el enfoque más avanzado de IRB propuesto por Basilea II. Finalmente, en la sección 6 se exponen las conclusiones relativas a los resultados obtenidos. Los algoritmos utilizados en la estimación de los modelos de cópulas son descritos en un Apéndice al final del documento.

2. Cópulas Elípticas Generalizadas y cópulas Agrupadas

La función de distribución (f.d.) conjunta de un vector aleatorio de factores de riesgo contiene tanto la descripción del comportamiento marginal de los componentes individuales como una descripción de su estructura de dependencia. El concepto de cópula³ permite describir dicha estructura de dependencia sin especificar la forma de las f.d. marginales. El Teorema Sklar (1959) que indica que cualquier distribución multivariada se puede descomponer en una cópula y sus marginales; si las distribuciones marginales son continuas la cópula es única. De acuerdo con este resultado, dada una f.d. F con marginales F_1, \dots, F_d continuas, existe una única cópula C tal que

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)), \quad (1)$$

para todo $x_1, \dots, x_d \in \mathfrak{R} \equiv [-\infty, \infty]$. La expresión (1) muestra explícitamente cómo se pueden construir distribuciones conjuntas F combinando distribuciones marginales con cópulas C . Dado un vector aleatorio $\mathbf{X} = (X_1, \dots, X_d)'$ que tiene f.d. F_X con distribuciones marginales continuas F_1, \dots, F_d , se dice que la cópula C asociada a \mathbf{X} (o a F_X) es la f.d. del vector aleatorio $(F_1(x_1), \dots, F_d(x_d))'^4$.

Uno de los ejemplos de cópulas más simples es la cópula de independencia, que está dada por

$$\Pi(u_1, \dots, u_d) = \prod_{i=1}^d u_i. \quad (2)$$

Entre las cópulas más ampliamente utilizadas en riesgos, y en particular en riesgo de crédito, se encuentra la cópula Gaussiana, la cual se define como la cópula asociada a algún vector aleatorio gaussiano $\mathbf{Y} \prec N_d(\mu, \Sigma)$ y está dada por

$$C_P^{Ga}(u_1, \dots, u_d) = \Phi_P(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d)) \quad (3)$$

donde $P = \Delta^{-1}\Sigma\Delta^{-1}$ es la matriz de correlaciones, con $\Delta = \text{diag}(\sqrt{\text{Var}(Y_1)}, \dots, \sqrt{\text{Var}(Y_d)})$. Otra cópula bastante popular en finanzas es la t de Student (simétrica), que se define como la cópula asociada al vector aleatorio $\mathbf{X} \prec t_d(\alpha, 0, P)$ con distribución t de Student multivariada. De acuerdo con esta definición, la cópula t de Student tiene la forma

$$C_{\alpha, P}^t(u_1, \dots, u_d) = t_{\alpha, P}(t_{\alpha}^{-1}(u_1), \dots, t_{\alpha}^{-1}(u_d)) \quad (4)$$

³ Se define a una cópula d -dimensional como una f.d. sobre $[0,1]^d$ con distribuciones marginales uniformes estándar.

⁴ Más precisamente, se puede verificar fácilmente que en este caso la cópula asociada a X está dada por $C(u_1, \dots, u_d) = F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))$.

Las dos cópulas anteriores pertenecen a una familia de cópulas más generales llamadas cópulas Elípticas Generalizadas (incluidas las simétricas), cuya fácil manipulación (y estimación) para dimensiones grandes la hacen una alternativa más atractiva frente a otras familias multivariadas. Se dice que un vector aleatorio \mathbf{Y} tiene distribución Elíptica Generalizada si se le puede representar como

$$\mathbf{Y}^d = \mu + R A U^k, \quad (5)$$

donde U^k es un vector aleatorio uniformemente distribuido sobre la esfera unitaria $S^{k-1} = \{x \in \mathbb{R}^k | x'x = 1\}$, R es una v.a., $\mu \in \mathbb{R}^d$, y $A \in M_{d \times k}(\mathbb{R})$. Esencialmente, la estructura de dependencia entre R y U^k determinan la f.d. de \mathbf{Y} . En el caso particular en que $R \geq 0$ es una v.a. independiente de U^k , se tiene que \mathbf{Y} tiene distribución Elíptica (Simétrica)⁵.

Frahm et al. 2003 propone modelar la asimetría mediante una distribución condicional en escala para R , es decir, considera $F_{R|U^{(d)}=u}(r) = P\{\gamma(u)R^* \leq r\}$, donde la función de escala γ está dada por

$$\gamma(u) = \gamma_0 + \sum_{i=1}^d \alpha_i \left(\delta \left(\frac{Au}{\|Au\|}, \nu_i^+ \right) \right)^{\eta_i} + \sum_{i=1}^d \beta_i \left(\delta \left(\frac{Au}{\|Au\|}, \nu_i^- \right) \right)^{\theta_i}, \quad (6)$$

con $\gamma_0 > 0$, $\alpha_1, \dots, \alpha_d, \beta_1, \dots, \beta_d \geq 0$, $\eta_1, \dots, \eta_d, \theta_1, \dots, \theta_d \geq 0$. Además, R^* es una v.a. positiva independiente de U^d , $\delta(u, \nu) = \frac{1}{2} - \frac{1}{2} \text{signo}(\langle u, \nu \rangle)$. $FB_{\left(\frac{1}{2}, \frac{d-1}{2}\right)}(\langle u, \nu \rangle^2)$ definida para $u, \nu \in S^{d-1}$ denotando por $FB_{(\alpha, \beta)}$ a la f.d. $Beta(\alpha, \beta)$. Por otro lado, los vectores de referencia ν_1^+, \dots, ν_d^+ son los vectores propios de Σ y ν_1^-, \dots, ν_d^- sus conjugados negativos, los cuáles existen ya que se supone que $\Sigma = AA'$ es positiva definida.

Cuando se define $F_{R|U^{(d)}=u}(r) = F_{(d, \gamma(u))}\left(\frac{r^2}{d}\right)$, donde $F_{(m, n)}$ es la f.d. F de Fisher con grados de libertad (m, n) , a la distribución resultante para \mathbf{Y} se conoce como t de Student Generalizada y su densidad está dada por

$$f(x) = \frac{\Gamma((w(x) + 2)/2)}{\Gamma(w(x)/2)(\pi w(x))^{d/2} |\Sigma|^{1/2}} \left(1 + \frac{(x - \mu)' \Sigma^{-1} (x - \mu)}{w(x)} \right)^{-(w(x)+d)/2}, \quad (7)$$

y $x \neq \mu$, donde $w(\mathbf{x}) = \gamma(\mathbf{x} - \mu)$. A la cópula asociada al vector aleatorio \mathbf{Y} se llama cópula Elíptica Generalizada.

Existe un grupo de cópulas llamadas cópulas de Mezcla de Normal Agrupadas. Uno de los ejemplos más comunes es la cópula t de Student Agrupada, definida por Demarta et al. 2004, en donde se construye una cópula para un vector aleatorio \mathbf{X} de manera que ciertos subvectores de \mathbf{X} tengan cópulas t de Student pero con diferentes niveles de dependencia en la cola entre sí. Sean las v.a. $\mathbf{Z} \sim N_d(\mathbf{0}, P)$ y $U \sim U(0, 1)$ independientes entre sí. Considérese una partición del conjunto $\{1, \dots, d\}$ en m subconjuntos de tamaos $\{s_1, \dots, s_m\}$ y para

⁵ Las propiedades y resultados más importantes de las distribuciones Elípticas Generalizadas se pueden revisar en Frahm et al. 2003, mientras que para el caso simétrico se recomienda Cambanis et al. 1981.

$k = 1, \dots, m$ sea ν_k el parámetro de grados de libertad asociado al grupo k . Denotando por G_α a la f.d. univariada Gamma Inversa $GI(\frac{\alpha}{2}, \frac{\alpha}{2})$, sea $W_k = G_{\alpha_k}^{-1}$ (esto implica que W_1, \dots, W_m son comonóticos o perfectamente dependientes⁶). Entonces la cópula t de Student Agrupada está dada por la f.d. del vector aleatorio

$$(t_{\alpha_1}(X_1), \dots, t_{\alpha_1}(X_{s_1}), t_{\alpha_2}(X_{s_1+1}), \dots, t_{\alpha_2}(X_{s_1+s_2}), \dots, t_{\alpha_m}(X_d))', \quad (8)$$

donde

$$\mathbf{X} = \left(\sqrt{W_1}Z_1, \dots, \sqrt{W_1}Z_{s_1}, \sqrt{W_2}Z_{s_1+1}, \dots, \sqrt{W_2}Z_{s_1+s_2}, \dots, \sqrt{W_m}Z_d \right)'$$

es un vector aleatorio con distribución t de Student Agrupada.

Asociadas intrínsecamente al concepto de cópula, se encuentran las medidas de dependencia. La más conocida y popular es correlación lineal de Pearson que, aunque juega un papel central en la teoría financiera, funciona bien solamente en el contexto Normal multivariado o más generalmente, para distribuciones Elípticas simétricas (ver Embrechts 2002). En contraste con la correlación lineal, las llamadas correlaciones de rango resultan ser funciones de la cópula (suponiendo marginales continuas), y por tanto, pueden ser utilizadas en su parametrización sin importar la familia de cópulas de que se trate. Entre las más utilizadas están la correlación de tau de Kendall y la rho de Spearman. Para dos v.a. X_1 y X_2 se definen la tau de Kendall y la rho de Spearman, respectivamente, como

$$\rho_\tau(X_1, X_2) = E \left[\text{signo}((X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2)) \right] \quad (9)$$

y

$$\rho_S(X_1, X_2) = \rho(F_1(X_1), F_2(X_2)),$$

donde $(\tilde{X}_1, \tilde{X}_2)$ es una copia independiente de (X_1, X_2) cuyas marginales correspondientes son F_1 y F_2 . Aquí $\rho(X_1, X_2)$ denota la correlación lineal de Pearson de X_1 y X_2 .

Otras medidas de dependencia para parejas de v.a. que también dependen solamente de la cópula de las v.a. X_1 y X_2 con marginales continuas, son los coeficientes de dependencia de la cola. Estos coeficientes proporcionan medidas de dependencia extrema, es decir, permiten cuantificar la magnitud de la dependencia en las colas de una distribución divariada. Se define el coeficiente de dependencia de X_1 y X_2 de la cola superior como

$$\lambda_u \equiv \lambda_u(X_1, X_2) = \lim_{q \rightarrow 1^-} P \{ X_2 > F_2^{-1}(q) | X_1 > F_1^{-1}(q) \}, \quad (10)$$

siempre y cuando el límite $\lambda_u \in [0, 1]$ exista. Análogamente, el coeficiente de dependencia de la cola inferior está dado por

$$\lambda_1 \equiv \lambda_1(X_1, X_2) = \lim_{q \rightarrow 1^+} P \{ X_2 \leq F_2^{-1}(q) | X_1 \leq F_1^{-1}(q) \}.$$

⁶ Es decir, su cópula asociada es la f.d. del vector aleatorio (U, \dots, U) , que se denota como cópula de comonotividad y está dada por $M(u_1, \dots, u_d) = \min\{u_1, \dots, u_d\}$.

Existe una relación entre la tau de Kendall y la correlación lineal para vectores elípticamente distribuidos con marginales continuas (para una prueba se recomienda ver Lindskog et al. 2003), la cuál establece que

$$\rho(X_1, X_2) = \text{sen} \left(\frac{\pi}{2} \rho_\tau(X_1, X_2) \right). \quad (11)$$

Para el caso de v.a. continuas cuya cópula asociada es una t de Student $C_{\alpha, P}^t(u_1, u_2)$, con $P = \rho I_2$, los coeficientes de dependencia están dados por $\lambda = 2t_{\alpha+1} \left(-\sqrt{(\alpha+1)\frac{1-\rho}{1+\rho}} \right)$. Sin embargo, existe una expresión general para los coeficientes de dependencia de la cola de vectores con f.d. Elíptica simétrica (en este caso $\lambda_l = \lambda_u$) cuando la v.a. R de la representación (4) es de variación regular con índice α . Dicha expresión, obtenida por Hult and Lindskog 2001, coincide, para $x = \frac{\pi}{4} - \frac{1}{2} \arcsin \rho$, con el valor de la función

$$\lambda(\alpha, x) = \frac{\int_x^{\frac{\pi}{2}} \cos^\alpha t dt}{\int_0^{\frac{\pi}{2}} \cos^\alpha t dt} \quad (12)$$

Entre los numerosos procedimientos estadísticos de estimación para las distribuciones Elípticas, son pocos los que están diseñados para trabajar sobre las cópulas Elípticas independientemente de sus marginales. Kostadinov 2005 propone un método de estimación no paramétrico, el cuál se describe resumidamente en el Algoritmo 1 del Apéndice y que básicamente se enfoca en estimar robustamente las medidas de dependencia ρ_τ y λ , así como el índice de la cola α . éste mismo algoritmo se utiliza para estimar los parámetros de la cópula t de Student Agrupada. Para el caso de las cópulas Elípticas Generalizadas se estiman μ y Σ utilizando estimadores propuestos por Frahm 2003.

3. Descripción del Portafolio y estimación de los parámetros de riesgo

3.1 Descripción del Portafolio

El portafolio de créditos a utilizar para la implementación de la metodología propuesta incluye a las tres subclases de activos al menudeo que las instituciones bancarias en México identifican: las exposiciones aseguradas por la vivienda del acreditado (como el portafolio hipotecario), las que tienen características revolventes (como el portafolio de tarjetas de crédito) y todas las demás exposiciones (como el portafolio de créditos al consumo). Los datos utilizados para la composición del portafolio son reales pero, como es común esperar en estas situaciones, se ha decidido mantener en secreto el origen de la fuente de información por razones de confidencialidad. Esta limitación no afecta, sin embargo, el alcance de los resultados ya que el método propuesto puede implementarse en casi cualquier institución financiera sin ninguna pérdida de generalidad.

Debido a que los productos que componen el portafolio están formados por un número grande de exposiciones, la modelación de los incumplimientos individualmente por cuenta es poco efectiva, por lo que es conveniente segmentar cada producto en grupos de créditos homogéneos en su perfil de riesgo. El

portafolio ha sido segmentado mediante una técnica clasificatoria basada en árboles de decisión llamada Detección de Interacción Automática Ji-Cuadrada (CHAID) que permite identificar divisiones óptimas cuando las variables son continuas, nominales e incluso categóricas (ver Biggs 1991). El procedimiento incluye una serie de reglas que son establecidas por las trayectorias de variables independientes a lo largo del árbol con el fin de obtener una partición de la población en grupos significativamente diferentes en relación con alguna variable clave. Tras utilizar como variable clave al parámetro PD y como variables independientes el producto, número de pagos vencidos, tiempo en libros, uso de línea de crédito y morosidad observada; el método CHAID arrojó una segmentación de 62 grupos (cubetas) repartidas en 10 productos. Cada producto tiene asociada una cubeta predeterminada de créditos incumplidos, mientras que el resto de éstas incluye alguna de las variables independientes.

Tabla 1
Composición del portafolio de créditos

Producto	Subgrupos	Volumen	Exposición
A	4	2,077	\$ 52,369,300
B	6	455,902	\$ 1,913,675,323
C	9	25,465	\$ 1,343,202,349
D	8	265,835	\$16,304,975,098
E	5	90,373	\$1,015,491,658
F	7	92,361	\$ 387,029,421
G	12	2,206,829	\$ 33,518,889,702
H	4	9,451	\$ 5,845,773,697
I	3	20,642	\$ 10,976,303,005
J	4	7,897	\$ 796,397,845
Total	62	3,176,832	\$ 72,154,107,399

La tabla 1 resume los resultados de la segmentación, mostrando para cada producto el número de subgrupos, el volumen en número de cuentas y la exposición en saldo (en el producto revolving G ésta se expresa como el límite de línea de crédito).

3.2 Estimación de los parámetros de riesgo

Una vez segmentado el portafolio, es necesario estimar los parámetros de riesgo. Para tal efecto conviene introducir, primero, la notación a utilizar. Sean $i = 1, \dots, M$ los grupos o cubetas en los que se divide el portafolio (en nuestro caso $M = 62$); $t = 1, \dots, n$ las fechas de observación; $j = 1, \dots, N_{it}$ los créditos y $e_{it}^{(j)}$ su exposición; $Y_{it}^{(j)}$ la indicadora del incumplimiento para cada crédito j dentro de una ventana anual a partir de la fecha de observación y, finalmente, $ead_{it}^{(j)}$ el saldo de dicho crédito al momento de su incumplimiento.

a) *Probabilidad de incumplimiento*. Para calcular la probabilidad de incumplimiento PD_i del grupo i se utiliza como estimador a la tasa de incumplimiento

observada, TI_{it} correspondiente a la ventana anual t (vista a partir de la fecha de observación t) en el grupo i ; esto es:

$$\widehat{TI}_{it} = \frac{n_{it}}{N_{it}}, \quad (13)$$

donde $n_{it} = \sum_{j=1}^{N_{it}} Y_{it}^{(j)}$ es el número de incumplimientos observados en la correspondiente ventana de observación y TI_{i1}, \dots, TI_{in} las realizaciones del parámetro de riesgo PD_i para cada cubeta $i = 1, \dots, M$.

b) *Exposición al momento del incumplimiento.* Asociada al parámetro EAD_i de cada grupo i se emplea la variable CCF_i o factor de conversión del crédito, definida como la proporción pérdida de la exposición inicial de los créditos al momento del incumplimiento. Para cada ventana de observación $t = 1, \dots, n$ se calcula CCF_i mediante el estimador insesgado

$$\widehat{CCF}_{it} = \frac{\sum_{j=1}^{N_{it}} ead_{it}^{(j)} Y_{it}^{(j)}}{\sum_{j=1}^{N_{it}} e_{it}^{(j)} Y_{it}^{(j)}} \quad (14)$$

c) *Pérdida dado el incumplimiento.* En el cálculo de LGD es primordial conocer, primero, el comportamiento de las recuperaciones de los créditos que incumplen exactamente en cada fecha de observación. Por eso es importante definir para cada producto p (formado a su vez por un conjunto de cubetas) el porcentaje que se recupera del saldo ead_{pt} de todos los créditos incumplidos en t durante los primeros k meses posteriores (ventana de recuperación); esto es de interés para definir la correspondiente tasa de recuperación τ_{pt} . En caso que se consideren los costos fijos c_p en el proceso de recuperación de cada producto p , el estimador de τ_{pt} está dado por

$$\widehat{\tau}_{pt} = \sum_{\tau=1}^k \frac{1 - c_p}{(1 + r)^\tau} \quad (15)$$

donde $R_{pt}^{(\tau)}$ es la recuperación del mes τ posterior correspondiente a los créditos del producto $p = 1, \dots, K$ que incumplieron en t ; $c_{ik}^{(j)}$ es el factor de costos y r es la tasa de interés adecuada para descontar los flujos. Con el cálculo de τ_{pt} se puede obtener, para cada fecha de observación $t = 1, \dots, n$, el estimador

$$\widehat{LGD}_{pt} = 1 - \widehat{\tau}_{pt} \quad (16)$$

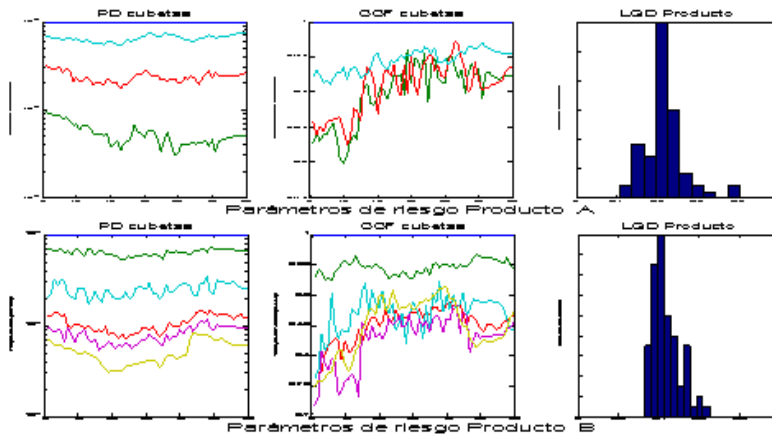
como una realización de LGD_p o de la pérdida dado el incumplimiento para el producto p en una ventana de recuperación de k meses. Se decidió considerar $k = 24$ meses para el caso de los portafolios de Consumo y Tarjeta de Crédito, y $k = 36$ meses para el de portafolio Hipotecario. Los tamaos de las ventanas de recuperación se eligieron de manera que las estimaciones $\widehat{\tau}_{pt}$ se estabilizaran a partir de k o más observaciones.

Los resultados de la consolidación de los parámetros de riesgo por producto son mostrados en las Figura 1. En concreto, las figuras muestran las estimaciones históricas de PD (a la izquierda) y CCF (al centro) de todas las cubetas

que componen al portafolio y adicionalmente se muestra el histograma de las estimaciones históricas de *LGD* que se tienen por producto (a la derecha). En todos los casos se observan niveles diferenciados del parámetro *PD* para cada cubeta en que se segmenta el producto, así como un comportamiento de *LGD* diferente para cada producto. Se presentan solamente las gráficas correspondientes a los productos *A* y *B*.

Figura 1

Estimaciones históricas de PD (izquierda), CCF (al centro) y LGD (derecha) por producto.

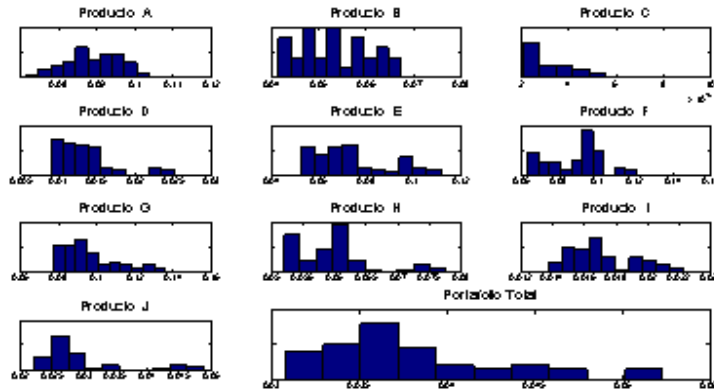


De la agregación de las estimaciones históricas de los parámetros de riesgo considerando la posición actual del portafolio, se obtienen las pérdidas históricas (como porcentaje de saldo), con las que se estiman empíricamente la pérdida esperada (*PE*) y el Valor en Riesgo (*VaR_q*) para ciertos niveles de confianza, por ejemplo, $q = .99, .995, .999, .999$. Los resultados de la simulación histórica de las pérdidas se muestran en la Tabla 2 y Figura 2.

Tabla 2
Estimaciones empíricas de PE y VaR_q de las pérdidas históricas por producto y portafolio total.

Producto	PE	VaR _q Emp			
		99%	99.5%	99.9%	99.95%
A	8.87%	10.35%	10.37%	10.37%	10.37%
B	5.31%	6.72%	6.73%	6.73%	6.73%
C	0.34%	0.89%	0.90%	0.90%	0.90%
D	1.34%	2.50%	2.51%	2.51%	2.51%
E	7.35%	11.15%	11.15%	11.15%	11.15%
F	8.95%	12.05%	12.10%	12.10%	12.10%
G	9.56%	12.92%	12.93%	12.93%	12.93%
H	5.99%	7.76%	7.76%	7.76%	7.76%
I	1.71%	2.23%	2.23%	2.23%	2.23%
J	2.88%	4.88%	4.88%	4.88%	4.88%
TOTAL	3.79%	5.21%	5.22%	5.22%	5.22%

Figura 2
Histogramas de las pérdidas históricas por producto y portafolio total.



De acuerdo con la Tabla 2, para niveles de confianza q cercanos a 1, el estimador empírico de VaR_q (columna VaR_qEmp) resulta poco apropiado debido al reducido número de 60 observaciones. En contraste, las estimaciones empíricas de la pérdida esperada (PE) pueden considerarse como estimaciones aceptables (columna PE). Los histogramas de la Figura 2 muestran comportamientos en los que no se pueden observar realizaciones de pérdidas suficientemente extremas. Para tener un buen ajuste al comportamiento distribucional marginal

se ajustan densidades kernel de suavización de Epanechnikov⁷.

4. Implementación del modelo de cálculo de CE usando cópulas

Para modelar la estructura de dependencia, es decir, la cópula asociada del vector aleatorio $(PD_i, EAD_i)_{i=1}^M$, se supone que $(TI_{i1}, CCF_{i1})_{i=1}^M, \dots, (TI_{in}, CCF_{in})_{i=1}^M$ las estimaciones conjuntas asociadas a dichos los parámetros de riesgo son observaciones independientes e idénticamente distribuidas (i.i.d.). Se considerará despreciable la posible correlación serial ocasionada por el uso de ventanas de observación que se traslapan, no obstante el cambio mes a mes que sufre la composición del portafolio por los créditos originados y los liquidados. Por otro lado, se supone que LGD es una v.a independiente de los parámetros de riesgo PD y EAD . Para la estimación de la cópula asociada al vector aleatorio $(LGD_p)_{p=1}^K$ también se supone que las estimaciones conjuntas $(\widehat{LGD}_{p1})_{p=1}^K, \dots, (\widehat{LGD}_{pm})_{p=1}^K$ son observaciones i.i.d. de éste.

En la agrupación de las M cubetas en sus correspondientes productos, se considera que las correspondientes v.a. PD y CCF son dependientes entre sí y entre cubetas, y que la estructura de dependencia está dada por alguna cópula. Más precisamente, se supone que el vector aleatorio $2k$ -dimensional

$$PDEAD_p = (CCF_{i1}, TI_{i1}, \dots, CCF_{ik}, TI_{ik})' \quad (17)$$

tiene marginales absolutamente continuas y una cópula C^{PDEAD} asociada, donde $\{i_1, \dots, i_k\} \subset \{1, \dots, M\}$ es un conjunto de cubetas que forman alguno de los $p = 1, \dots, K$ productos ajenos entre sí. Para todo el portafolio, se supone también que el vector aleatorio K -dimensional

$$LGD = (LGD_1, \dots, LGD_K)' \quad (18)$$

posee marginales absolutamente continuas y alguna cópula asociada C^{LGD} .

Estructuras de dependencia multidimensionales comúnmente utilizadas, como la cópula de Independencia (2) y la cópula Gaussiana (3), son ajustadas a los datos de los parámetros de riesgo PD y CCF como ejercicios iniciales. Por el momento, el parámetro LGD se considera constante e igual a su media histórica. Para el ajuste Gaussiano, la matriz de correlaciones se estima mediante el estimador usual

$$\hat{\Sigma} = \left(\sum_{k=1}^n y_k y_k' \right) \quad (19)$$

donde $y_k = \left(\Phi^{-1}(\hat{F}_1(x_{ik})), \dots, \Phi^{-1}(\hat{F}_d(x_{dk})) \right)'$ y \hat{F}_i es la f.d. empírica del vector de observaciones $(x_{i1}, \dots, x_{in})'$ de la v.a. X_i , para cada $i = 1, \dots, d$.

⁷ Para una muestra de observaciones x_1, \dots, x_n el estimador de densidad kernel está dado por la función $\hat{f}(x) = \frac{\sum_{j=1}^n K_h(x-x_j)}{n \int K_h(t-u) dt}$ donde $K_h(x) = \left(\frac{1}{h}\right) K\left(\frac{x}{h}\right)$, $K(u) = \left(\frac{3}{4}\right) (1-u^2) 1_{\{|u|<1\}}$ es la llamada función kernel de Epanechnikov y el parámetro h se llama ancho de banda.

Para efectos de comparación, en todos los ajustes marginales el ancho de banda será el dado por la regla de Silverman⁸.

Los resultados obtenidos en cada ejercicio se muestran en las Tablas 5 y 6 abajo, mismas en las que por producto y portafolio total se estiman la *PE* y el *VaR_q*; éste último mediante los métodos *POT*, Empírico y el ajuste de una distribución Beta.

Tabla 3

Estimaciones de PE y VaR_q ajustando la cópula de Independencia a PD y CCF de todo el portafolio.

Producto	PE	VaRqEmp			
		99%	99.5%	99.9%	99.95%
A	8.91%	10.61%	10.79%	11.20%	11.32%
B	5.37%	6.58%	6.67%	6.96%	7.01%
C	0.38%	1.05%	1.15%	1.31%	1.35%
D	1.36%	1.83%	1.88%	2.00%	2.04%
E	7.36%	11.48%	11.87%	12.65%	12.80%
F	8.88%	10.76%	10.99%	11.37%	11.61%
G	9.73%	11.72%	11.93%	12.42%	12.65%
H	6.00%	7.65%	7.74%	7.89%	7.96%
I	1.73%	2.30%	2.35%	2.46%	2.51%
J	2.87%	4.39%	4.52%	4.84%	4.93%
TO TA L	3.83%	4.27%	4.33%	4.43%	4.47%

Tabla 4

Estimaciones PE y VaR_q ajustando una cópula Gaussiana a PD y CCF de todo el portafolio.

Producto	PE	VaRqEmp			
		99%	99.5%	99.9%	99.95%
A	8.92%	10.33%	10.52%	10.85%	10.93%
B	5.36%	6.82%	6.96%	7.17%	7.21%
C	0.37%	0.94%	1.02%	1.14%	1.21%
D	1.37%	2.41%	2.51%	2.68%	2.72%
E	7.41%	12.25%	12.66%	13.55%	14.01%
F	8.94%	12.07%	12.41%	13.11%	13.19%
G	9.74%	13.52%	13.83%	14.55%	14.73%
H	6.00%	7.43%	7.51%	7.72%	7.74%
I	1.72%	2.19%	2.24%	2.34%	2.36%
J	2.88%	4.60%	4.69%	4.86%	4.93%
TO TA L	3.84%	5.04%	5.17%	5.39%	5.42%

⁸ Para la elección del parámetro de ancho de banda de la densidad kernel se puede utilizar la regla de Silverman, que permite determinar un ancho de banda *h* óptimo (para ajustar densidades normales) el cuál está dado por $(40\sqrt{\pi}/n)^{1/5} \hat{\sigma}$, donde *n* es el número de observaciones y $\hat{\sigma}$ la volatilidad estimada de los datos.

Las estimaciones de PE (columna PE), comparadas con sus estimadores históricos, parecen ser poco sensibles a la manera en que se modela la estructura de dependencia de los parámetros de riesgo. Sin embargo, las estimaciones empíricas de VaR_q (columnas $VaRqEmp$), en ambos ajustes, son relativamente inferiores respecto del comportamiento histórico de las pérdidas.

Una estructura de dependencia que le asigna probabilidad positiva a los eventos de observas conjuntas extremas es la cópula t de Student. El siguiente ejercicio consiste en ajustar esta cópula bajo las mismas consideraciones de LGD y densidades kernel. La matriz de correlaciones se estima también con la ecuación (19) tomando $y_k = \left(t_\alpha^{-1}(\hat{F}_1(x_{ik})), \dots, t_\alpha^{-1}(\hat{F}_d(x_{dk})) \right)'$, donde t_α denota la f.d. t de Student univariada cuyo parámetro de grados de libertad α es estimado aplicando el Algoritmo 1 del Apéndice.

Tabla 5

Estimaciones de PE y VaR_q ajustando una cópula t de Student Simétrica PD y CCF de todo el portafolio.

		VaRqEmp			
Producto	PE	99%	99.5%	99.9%	99.95%
A	8.90%	10.94%	11.12%	11.49%	11.73%
B	5.36%	7.19%	7.32%	7.61%	7.67%
C	0.37%	0.98%	1.08%	1.21%	1.25%
B	1.36%	2.56%	2.69%	2.84%	2.87%
C	7.37%	12.77%	13.35%	14.07%	14.42%
B	8.90%	12.62%	13.05%	13.66%	13.93%
C	9.70%	14.18%	14.49%	15.30%	15.47%
B	6.00%	7.80%	7.91%	8.13%	8.17%
C	1.72%	2.27%	2.34%	2.46%	2.48%
B	2.87%	4.79%	4.91%	5.08%	5.14%
TOTA L	3.82%	5.33%	5.46%	5.68%	5.78%

La Tabla 5 revela que el ajuste de la cópula t de Student Simétrica al portafolio total es más adecuado que los ajustes de las cópulas de Independencia y Gaussiana ya que las estimaciones de VaR_q obtenidas son consistentes con las estimaciones históricas.

4.1 Ajuste de cópulas Elípticas Simétricas y Generalizadas al portafolio de créditos

En esta sección se modela la estructura de dependencia al interior de cada producto utilizando alguna cópula Elíptica Simétrica y Generalizada. ésta última cópula se emplea para introducir y, por tanto, evaluar el efecto de asimetría en la estructura de dependencia de los parámetros de riesgo.

Aplicando el Algoritmo 2 se ajusta para cada producto p una cópula C^pPDEAD para modelar la estructura de dependencia del vector aleatorio $pPDEAD$ dado por la ecuación (17), al tiempo que al vector aleatorio LGD

se le ajusta de manera independiente una cópula C^{CLGD} . La distribución marginal de los parámetros de riesgo contenidos en los vectores anteriores se modela, como se ha venido haciendo en los ejercicios anteriores, ajustando densidades kernel de suavización a las observaciones históricas de PD , EAD y LGD .

Para cada cubeta i la pérdida (como porcentaje de su saldo) se define como la v.a.

$$L_i = TI_i * CCF_i * LGD_i \tag{20}$$

Entonces con las observaciones simuladas $p\widehat{PDEAD}_1, \dots, p\widehat{PDEAD}_N$, $p = 1, \dots, K$ y $\widehat{LGD}_1, \dots, \widehat{LGD}_N$ de los parámetros de riesgo y aplicando (20) se construyen, para cada cubeta $i \in \{i_1, \dots, i_k\}$ del producto p en cuestión, las pérdidas simuladas L_{i1}, \dots, L_{iN} . Para hacer la agregación de pérdidas para cada producto (como porcentaje de su saldo) es necesario conocer la composición de un portafolio de créditos actual, es decir, el saldo actual e_i de cada cubeta i . Entonces, simplemente se pondera por el saldo de cada producto p para obtener las pérdidas simuladas por producto, es decir,

$$L_{pj} = \sum_{i \in \{i_1, \dots, i_k\}} \frac{e_i}{\sum_{j \in \{i_1, \dots, i_k\}} e_j} L_{ij}, \quad j = 1, \dots, N \tag{21}$$

son las pérdidas simuladas del producto p formado por las cubetas $i \in \{i_1, \dots, i_k\}$.

La tabla 8 muestra los estimadores del índice de la cola α^9 que resultan de ajustar las cópulas Elípticas a cada producto y al portafolio total. Como se puede apreciar, los comportamientos a observaciones extremas conjuntas para los parámetros PD y CCF son bastante diferentes entre productos.

Tabla 6

Grados de libertad para (PD, CCF) por producto y portafolio total.

(PD, CCF)	
Producto	Grados de libertad α_p (Depend. Cola)
A	20.5
B	38.8
C	11.8
D	14.2
E	67.9
F	9.7
G	10.8
H	2.0
I	22.1
J	3.2
TOTAL	276.8

⁹ La v.a. espectral R de la representación (5) se supone de variación regular en infinito con índice $\alpha > 0$. Dicho índice mide la ocurrencia de observaciones conjuntas extremas, a menor valor de α mayor dependencia entre las observaciones extremas conjuntas. En el caso de la cópula t de Student coincide con los grados de libertad.

Los resultados de las simulaciones de las cópulas ajustadas *t* Simétrica y Generalizada están contenidos en las tablas 7 y 8, respectivamente. Dichas tablas incluyen las estimaciones del VaR_q por los tres métodos de estimación de altos cuantiles propuestos: POT^{10} (Columna VaRqPOT), Empírico (columna VaRqEmp) y Ajuste de una distribución Beta (VaRqBeta).

Tabla 7
Estimaciones de PE y VaR_q por producto ajustando una cópula *t* de Student Simétrica a los parámetros de riesgo PD y CCF.

Producto	PE	VaRqPOT				VaRqEmp				VaRqBeta			
		99%	99.5%	99.9%	99.95%	99%	99.5%	99.9%	99.95%	99%	99.5%	99.9%	99.95%
A	8.88%	10.93%	11.15%	11.57%	11.72%	10.91%	11.19%	11.58%	11.80%	10.85%	11.08%	11.56%	11.75%
B	5.33%	7.27%	7.41%	7.83%	7.70%	7.28%	7.40%	7.59%	7.64%	7.40%	7.65%	8.18%	8.39%
C	0.38%	1.01%	1.12%	1.36%	1.46%	1.03%	1.13%	1.33%	1.38%	0.79%	0.85%	0.96%	1.03%
D	1.36%	2.55%	2.68%	2.89%	2.95%	2.56%	2.67%	2.90%	2.96%	2.30%	2.43%	2.70%	2.81%
E	7.39%	12.98%	13.58%	14.62%	14.95%	13.00%	13.49%	14.56%	14.82%	12.00%	12.59%	13.85%	14.36%
F	8.97%	12.73%	13.14%	13.92%	14.19%	12.78%	13.19%	13.88%	14.06%	12.77%	13.24%	14.22%	14.61%
G	9.59%	13.90%	14.36%	15.17%	15.43%	13.86%	14.37%	15.31%	15.57%	13.22%	13.66%	14.58%	14.95%
H	6.00%	7.91%	8.03%	8.20%	8.23%	7.98%	8.00%	8.20%	8.26%	7.51%	7.69%	8.06%	8.20%
I	1.72%	2.29%	2.35%	2.45%	2.49%	2.28%	2.33%	2.45%	2.49%	2.28%	2.35%	2.49%	2.55%
J	2.88%	4.93%	5.06%	5.24%	5.27%	4.89%	5.02%	5.22%	5.28%	4.30%	4.48%	4.86%	5.02%

Tabla 8
Estimaciones de PE y VaR_q por producto ajustando una cópula *t* de Student Generalizada a los parámetros de riesgo PD y CCF.

Producto	PE	VaRqPOT				VaRqEmp				VaRqBeta			
		99%	99.5%	99.9%	99.95%	99%	99.5%	99.9%	99.95%	99%	99.5%	99.9%	99.95%
A	8.88%	11.05%	11.35%	12.00%	12.26%	11.05%	11.39%	12.08%	12.25%	10.83%	11.05%	11.53%	11.71%
B	5.33%	7.32%	7.50%	7.83%	7.94%	7.27%	7.48%	7.84%	7.94%	7.39%	7.64%	8.17%	8.38%
C	0.38%	1.04%	1.18%	1.55%	1.71%	1.06%	1.17%	1.43%	1.64%	0.79%	0.85%	0.96%	1.04%
D	1.36%	2.57%	2.70%	2.92%	2.99%	2.60%	2.72%	2.93%	2.99%	2.30%	2.43%	2.70%	2.81%
E	7.39%	13.14%	13.84%	15.14%	15.58%	13.19%	13.91%	15.18%	15.68%	11.96%	12.54%	13.78%	14.29%
F	8.97%	12.79%	13.24%	14.10%	14.40%	12.83%	13.27%	14.03%	14.34%	12.74%	13.19%	14.17%	14.56%
G	9.59%	13.88%	14.36%	15.24%	15.54%	13.83%	14.43%	15.35%	15.70%	13.21%	13.65%	14.57%	14.94%
H	6.00%	7.92%	8.04%	8.20%	8.24%	7.86%	7.98%	8.18%	8.27%	7.51%	7.68%	8.05%	8.20%
I	1.72%	2.38%	2.47%	2.66%	2.72%	2.42%	2.48%	2.62%	2.69%	2.28%	2.35%	2.49%	2.55%
J	2.88%	4.94%	5.08%	5.25%	5.29%	4.90%	5.03%	5.21%	5.29%	4.30%	4.48%	4.86%	5.02%

Según los resultados obtenidos, aparentemente no hay ventaja significativa de modelar la asimetría en la estructura de dependencia con el uso de las cópulas elípticas generalizadas, sin embargo, se debe utilizar un criterio estadístico que permita evaluar la adecuación de los diferentes ajustes de las cópulas a los datos observados.

4.2 Ajuste de la cópula *t* Agrupada al portafolio de créditos

La cópula *t* de Student Agrupada permite modelar la estructura de dependencia entre grupos de v.a. que a su vez tienen estructura de dependencia dada por alguna cópula *t* de Student pero cuyo comportamiento de dependencia a eventos extremos es diferenciado entre grupos. Ésta cópula se ajusta a los parámetros de riesgo PD y CCF del portafolio total, utilizando los estimadores obtenidos en el paso (i) del Algoritmo 2 para los parámetros de grados de libertad α_p de

¹⁰ Para revisar los resultados más importantes de este método derivado de la Teoría de Valores Extremos se recomienda ver Embrechts 1997.

cada producto p y la matriz de correlaciones P que resulta para el portafolio total. A partir de la expresión (8) y utilizando los estimadores de los parámetros anteriores se simulan observaciones de $(pPDEAD)_{p=1}^K$, que al ser agregados por producto y portafolio Total se muestran en la Figura 3.

Según la Tabla 9, las estimaciones por producto de PE y VaR_q mediante éste enfoque son consistentes con los ajustes de las cópulas t de Student Simétrica y Generalizada. La estimación de VaR_q del portafolio total permite cuantificar el efecto de diversificación cuando se le compara con la suma de los VaR_q 's de todos los productos ponderados por el saldo total (renglón inferior Total pond $\rho = 1$ de la tabla), así como con la raíz cuadrada de suma de sus cuadrados (Total pond $\rho = 0$).

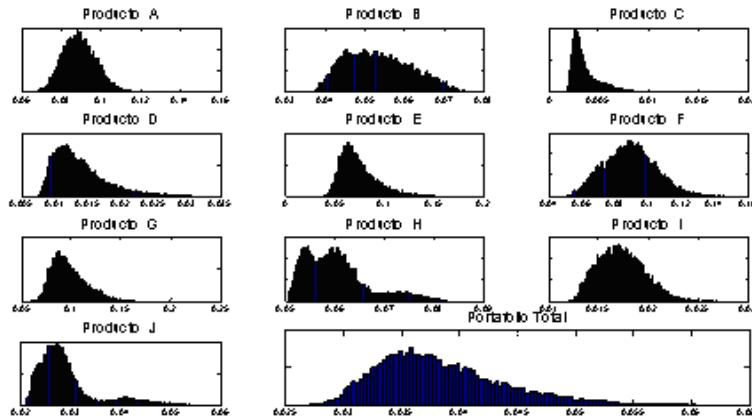
Tabla 9

Estimaciones de PE y VaR_q por producto y portafolio total ajustando una cópula t de Student Agrupada a los parámetros de riesgo PD y CCF.

Producto	PE	VaRqPOT				VaRqEmp				VaRqBeta			
		99%	99.5%	99.9%	99.95%	99%	99.5%	99.9%	99.95%	99%	99.5%	99.9%	99.95%
A	8.93%	10.99%	11.21%	11.63%	11.78%	11.00%	11.23%	11.54%	11.71%	10.96%	11.19%	11.69%	11.88%
B	5.35%	7.23%	7.37%	7.66%	7.65%	7.23%	7.36%	7.57%	7.62%	7.42%	7.67%	8.20%	8.42%
C	0.38%	1.01%	1.12%	1.36%	1.46%	1.03%	1.11%	1.28%	1.38%	0.80%	0.86%	1.00%	1.05%
D	1.37%	2.96%	2.69%	2.89%	2.96%	2.88%	2.68%	2.89%	2.95%	2.34%	2.47%	2.75%	2.87%
E	7.37%	13.01%	13.57%	14.49%	14.76%	13.07%	13.51%	14.61%	14.91%	11.97%	12.56%	13.82%	14.32%
F	8.82%	12.76%	13.17%	13.94%	14.20%	12.83%	13.21%	13.85%	14.15%	12.86%	13.35%	14.37%	14.78%
G	9.73%	14.43%	14.89%	15.69%	15.94%	14.44%	14.88%	15.70%	16.02%	13.85%	14.36%	15.42%	15.84%
H	6.01%	7.88%	8.00%	8.16%	8.19%	7.82%	7.98%	8.16%	8.21%	7.52%	7.69%	8.06%	8.21%
I	1.73%	2.30%	2.37%	2.60%	2.55%	2.31%	2.37%	2.60%	2.57%	2.30%	2.37%	2.52%	2.57%
J	2.89%	4.96%	5.01%	5.22%	5.27%	4.85%	4.98%	5.24%	5.29%	4.30%	4.48%	4.86%	5.01%
TOTAL	3.84%	5.36%	5.52%	5.80%	5.89%	5.36%	5.53%	5.83%	5.90%	5.15%	5.31%	5.64%	5.78%
Total pond rho=1	3.84%	5.71%	5.90%	6.20%	6.30%	5.72%	5.89%	6.20%	6.32%	5.46%	5.67%	6.10%	6.27%
Total pond rho=0		2.90%	2.99%	3.15%	3.20%	2.90%	2.99%	3.15%	3.21%	2.78%	2.88%	3.09%	3.18%

Figura 3

Histogramas de las pérdidas simuladas por producto y portafolio total ajustando una cópula t de Student Agrupada a los parámetros PD y CCF.



Para establecer estadísticamente la bondad de ajuste de cada modelo de cópulas ajustado, se emplea un estadístico basado en formas bilineales propuesto por Panchenko 2005. En la Tabla 9 se presentan los valores promedio de las distan-

cias Q basadas en formas bilineales que se obtienen al implementar, por cada producto y al portafolio total.

Tabla 10

Distancias Q (estadística de prueba) para cada tipo de cópulas ajustado.

Producto	Cópula Independencia portafolio total	Cópula Gaussiana portafolio total	Cópula t Simétrica portafolio total	Cópula t Simétrica por producto	Cópula t Generalizada por producto	Cópula t Agrupada
A	0.10415	0.07264	0.04061	0.04016	0.04069	0.03980
B	0.09646	0.06839	0.03817	0.03786	0.03793	0.03670
C	0.09191	0.06615	0.03667	0.03662	0.03669	0.03600
D	0.09699	0.06858	0.03841	0.03791	0.03807	0.03700
E	0.10004	0.07006	0.03916	0.03866	0.03895	0.03810
F	0.09482	0.06797	0.03774	0.03752	0.03754	0.03670
G	0.09040	0.06519	0.03615	0.03612	0.03612	0.03580
H	0.10676	0.07413	0.04152	0.04152	0.04134	0.03970
I	0.11059	0.07632	0.04255	0.04275	0.04463	0.04080
J	0.10658	0.07577	0.04206	0.04213	0.04223	0.04110
TOTAL	0.08873	0.06400	0.03549	0.03549	0.03549	0.03460

Los valores de la estadística de prueba Q muestran que la cópula t de Student Agrupada es, de entre las cópulas utilizadas, la que presenta un mejor ajuste a los datos (a la cópula empírica de los parámetros de riesgo). No obstante que el reducido número de observaciones históricas de los parámetros de riesgo no permite establecer la significancia de la prueba, la Tabla 10 indica que el ajuste de la cópula t de Student Agrupada es ligeramente superior a los ajustes de las cópulas t de Student Simétrica y Generalizada, en tanto que los supuestos de independencia y gaussiano no son objetivos con lo observado en los datos.

Con fundamento en el criterio anterior, se decide modelar la estructura de dependencia de los parámetros de riesgo del portafolio PD y CCF con la cópula t de Student Agrupada. Por otro lado, la estructura de dependencia del parámetro LGD entre productos del portafolio se modela con una cópula t de Student simétrica. Con el fin de estresar las distribuciones marginales del modelo, todas las densidades kernel se ajustan con el doble del ancho de banda empleado en los ejercicios anteriores. Los resultados obtenidos se presentan en las Tablas 11 y 12, donde esta última contiene además el cálculo de la medida de riesgo Déficit Esperado (ES_q), también estimada por los tres métodos implementados para VaR_q .

Tabla 11

Estimaciones de PE y VaR_q ajustando una cópula t de Student Agrupada a los parámetros PD y CCF , y una cópula t de Student a la LGD .

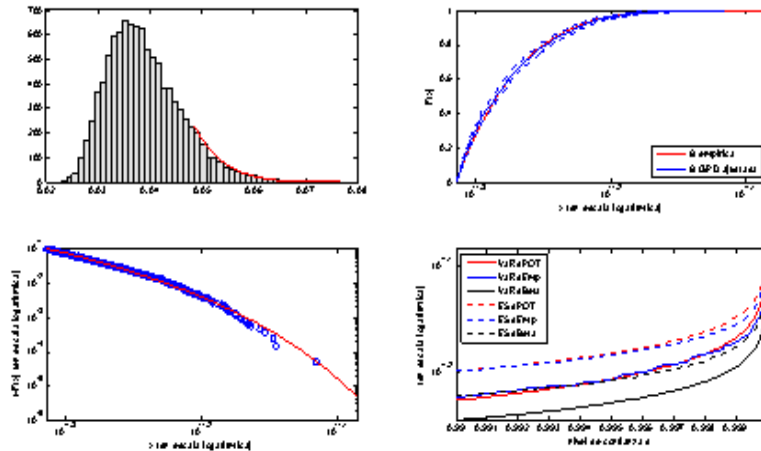
Producto	PE	VaRqPOT					VaRqEmp				VaRqBeta			
		99%	99.5%	99.9%	99.95%	99%	99.5%	99.9%	99.95%	99%	99.5%	99.9%	99.95%	
A	9.01%	14.03%	14.79%	16.46%	17.13%	13.99%	14.86%	16.79%	17.76%	13.46%	14.01%	15.19%	15.66%	
B	5.48%	9.73%	10.20%	11.07%	11.37%	9.08%	10.17%	11.03%	11.46%	9.45%	9.97%	11.09%	11.55%	
C	0.43%	2.00%	2.29%	2.85%	3.06%	2.09%	2.36%	2.73%	2.85%	1.46%	1.63%	2.03%	2.21%	
D	1.41%	3.25%	3.52%	4.10%	4.32%	3.25%	3.57%	4.10%	4.26%	2.98%	3.20%	3.70%	3.90%	
E	7.61%	17.55%	19.00%	21.89%	22.96%	17.44%	19.16%	22.11%	22.77%	15.74%	16.86%	19.28%	20.26%	
F	9.02%	15.90%	16.66%	18.09%	18.57%	15.81%	16.73%	18.24%	18.73%	15.63%	16.48%	18.32%	19.06%	
G	9.83%	15.51%	16.09%	17.13%	17.46%	15.53%	16.12%	17.28%	17.56%	14.86%	15.48%	16.80%	17.33%	
H	6.09%	11.83%	12.48%	13.68%	14.12%	11.77%	12.37%	13.80%	14.40%	11.63%	12.38%	14.00%	14.66%	
I	1.78%	3.49%	3.66%	3.99%	4.10%	3.40%	3.64%	4.04%	4.16%	3.79%	4.08%	4.71%	4.97%	
J	3.03%	7.79%	8.58%	10.32%	11.03%	7.67%	8.51%	10.53%	11.15%	6.96%	7.54%	8.80%	9.33%	
TOTAL	3.91%	5.96%	6.21%	6.74%	6.93%	5.95%	6.19%	6.72%	7.03%	5.76%	5.99%	6.49%	6.69%	
Total pond rho=1	3.90%	7.23%	7.62%	8.39%	8.66%	7.22%	7.63%	8.46%	8.70%	6.98%	7.40%	8.30%	8.66%	
Total pond rho=0		3.41%	3.57%	3.88%	3.99%	3.41%	3.57%	3.91%	4.01%	3.29%	3.47%	3.84%	3.99%	

Tabla 12
 Estimaciones de ES_q ajustando una cópula t de Student Agrupada a los parámetros PD y CCF, y una cópula t de Student a la LGD.

Producto	ESqPOT				ESqEmp				ESqBeta			
	99%	99.5%	99.9%	99.95%	99%	99.5%	99.9%	99.95%	99%	99.5%	99.9%	99.95%
A	15.09%	15.81%	17.40%	18.04%	15.09%	15.81%	17.67%	18.30%	14.23%	14.74%	15.85%	16.34%
B	10.34%	10.73%	11.45%	11.69%	10.32%	10.74%	11.50%	11.77%	10.17%	10.66%	11.73%	12.22%
C	2.38%	2.63%	3.13%	3.31%	2.41%	2.60%	2.90%	3.01%	1.71%	1.88%	2.31%	2.48%
D	3.62%	3.88%	4.40%	4.61%	3.63%	3.86%	4.37%	4.59%	3.29%	3.51%	4.01%	4.21%
E	19.48%	20.76%	23.31%	24.25%	19.57%	20.81%	23.09%	23.78%	17.30%	18.36%	20.66%	21.66%
F	16.88%	17.52%	18.71%	19.11%	16.92%	17.61%	18.74%	19.01%	16.82%	17.62%	19.36%	20.12%
G	16.25%	16.72%	17.55%	17.82%	16.30%	16.81%	17.57%	17.76%	15.72%	16.30%	17.55%	18.10%
H	12.66%	13.20%	14.25%	14.63%	12.64%	13.26%	14.43%	14.75%	12.87%	13.38%	14.93%	15.62%
I	3.72%	3.86%	4.14%	4.24%	3.71%	3.86%	4.17%	4.26%	4.19%	4.47%	5.11%	5.37%
J	8.90%	9.65%	11.33%	12.02%	8.94%	9.79%	11.43%	12.05%	7.77%	8.32%	9.59%	10.11%
TOTAL	6.30%	6.53%	7.00%	7.17%	6.29%	6.50%	7.08%	7.34%	6.08%	6.30%	6.79%	6.99%
Total pond rho=1	7.75%	8.09%	8.75%	8.99%	7.76%	8.11%	8.77%	8.98%	7.56%	7.95%	8.83%	9.21%
Total pond rho=0	3.62%	3.76%	4.02%	4.11%	3.63%	3.77%	4.03%	4.11%	3.53%	3.70%	4.06%	4.22%

Los estimaciones obtenidas por el método POT son bastante consistentes con los estimadores empíricos, al tiempo que permiten obtener estimaciones más confiables para niveles de q bastante cercanos a 1, esto debido a su base sobre resultados asintóticos de la TVE (ver Embrechts et al. 1997 y Díaz 2003). En la Figura 4 se muestra el desempeño de los estimadores de VaR_q y ES_q por los tres métodos POT , Empírico y el ajuste de una Beta. Se presentan los resultados para el caso del portafolio total

Figura 4
 Estimaciones de VaR_q y ES_q (métodos POT , Empírico y Beta)
 Portafolio Total.



La primera de las subgráficas (de izquierda a derecha y de arriba a abajo) contiene el histograma de las pérdidas simuladas y la parte de la cola a la que se ajusta una DPG , mientras que en la segunda aparece su ajuste con la distribución empírica. En la tercera se presentan el ajuste entre la Distribución de Excesos F_u y la empírica de la cola de la distribución de pérdidas. Final-

mente, en la cuarta se subgráfica se presenta el comportamiento de los diferentes estimadores de las medidas de riesgo VaR_q y ES_q , para $q \in (.99, 1)$.

Además de la cópula t de Student Agrupada, existen otras cópulas relacionadas a la cópula t de Student como es el caso de la cópula t de Valores Extremos y la cópula t de Cola Inferior, sin embargo, su manipulación y estimación para dimensiones grandes es bastante complicada (ver Demarta et al. 2004).

5. Estimaciones de CE bajo modelos tradicionales y el enfoque IRB

Con el fin de comparar el modelo propuesto con otros enfoques, se implementan un modelo de incumplimiento de un sólo factor, un modelo basado en matrices de transición y el cálculo del requerimiento de capital regulatorio según las reglas de Basilea II (ver BCBS 2004) en su enfoque avanzado de calificaciones internas (IRB).

El modelo de un factor tipo KM/CM ha sido ampliamente utilizado para la estimación de capital económico en portafolios de créditos revolventes y de consumo (ver RMA 2003). Sus parámetros importantes, la correlación de activos (AVC) y la correlación entre incumplimientos (LC) se calibran con información histórica de las tasas de incumplimiento. Concretamente, el CE_i (como porcentaje del saldo) de la cubeta i se estima como

$$K_i = LGD_i \cdot CCF_i \cdot \Phi \left(\frac{\Phi^{-1}(PD_i) + \sqrt{1 - AVC_i} \Phi^{-1}(q)}{\sqrt{1 - AVC_i}} \right), \quad (22)$$

donde q es el nivel de confianza y PD_i se estima como el valor promedio de las correspondientes tasas de incumplimiento observadas.

Con las mismas bases teóricas que el modelo de un factor del tipo KMV/CM , el enfoque avanzado de calificaciones internas (IRB) de Basilea II, utiliza resultados asintóticos sobre los cuantiles de la función de pérdidas del portafolio cuando el número de exposiciones es suficientemente grande (ver Frey et al. 2003). Más precisamente, este enfoque regulatorio propone estimar el requerimiento de capital RC_i , para cada cubeta i de créditos homogéneos en perfil de riesgo, como

$$RC_i = LGD_i \cdot CCF_i \cdot \Phi \left(\frac{\Phi^{-1}(PD_i) + \sqrt{\rho} \Phi^{-1}(.999)}{\sqrt{1 - \rho}} \right)$$

donde el nivel de confianza se fija en $q = .999$ y la correlación de activos ρ está preestablecida según el tipo de portafolio al que pertenezca la cubeta (0.04 para tarjeta de crédito, 0.15 para hipotecario y $.03 \left(\frac{1 - e^{-35PD_i}}{1 - e^{-35}} \right) + .16 \left(\frac{1 - e^{-35PD_i}}{1 - e^{-35}} \right)$ para el portafolio de consumo).

En ambos enfoques los parámetros CCF y LGD se consideran constates, por lo que para fines de la aplicación serán iguales a su valor promedio histórico.

En el siguiente cuadro resumen (Tabla 13) se muestran los resultados de CE y capital regulatorio obtenidos bajo el modelo de un factor y las reglas de IRB de Basilea II. Se agregan también las estimaciones de VaR_q por el método POT que se obtuvieron en el ajuste de la cópula t de Student Agrupada, así como las que resultan de ajustar densidades kernel que dan más peso a

observaciones recientes de los parámetros de riesgo mediante el uso de un factor de decaimiento, por ejemplo, $\lambda = .97$. Pensando en niveles de confianza q realistas, se consideran 99.5% y 99.9% (si por ejemplo, la institución aspira la calificación $mxAAA$, que en escala internacional corresponde a BBB , que indica una probabilidad de incumplimiento del 0.5%, el nivel de confianza a considerar es $q = 99.5\%$).

Tabla 13

Estimaciones de CE por el modelo de un factor KMV/CM, el capital regulatorio bajo el enfoque IRB de Basilea II y el CE del modelo Cópula t Student agrupada.

Producto	Modelo de un factor KMV/CM		Capital Regulatorio IRB Basilea II	Cópula t Agrupada		Cópula t Agrupada $\lambda=.97$	
	99.5%	99.9%		99.5%	99.9%	99.5%	99.9%
A	12.93%	14.07%	14.51%	14.79%	16.45%	14.76%	16.62%
B	9.65%	10.92%	10.64%	10.20%	11.07%	10.48%	11.42%
C	1.00%	1.24%	0.63%	2.29%	2.85%	2.31%	2.71%
D	2.16%	2.41%	2.59%	3.52%	4.10%	3.30%	3.85%
E	10.43%	11.17%	13.80%	19.00%	21.89%	17.62%	20.80%
F	12.71%	13.92%	14.88%	16.66%	18.09%	16.52%	18.22%
G	14.92%	16.65%	15.52%	16.09%	17.13%	16.06%	16.20%
H	7.17%	7.47%	9.57%	12.46%	13.68%	12.15%	13.17%
I	2.03%	2.11%	2.91%	3.66%	3.99%	3.63%	4.02%
J	4.57%	5.09%	7.31%	8.58%	10.32%	7.83%	9.81%
TOTAL	5.55%	6.08%	6.47%	6.21%	6.74%	6.09%	6.41%

En la mayoría de los casos el modelo de un factor indica un CE menor que el requerimiento de capital regulatorio bajo Basilea II, lo cual se debe en esencia a que, no obstante que la construcción en ambos enfoques es casi la misma, la correlación de activos que el primero utiliza es calculada con información histórica, mientras que en el enfoque *IRB* las correlaciones de activos están predeterminadas y el nivel de confianza que usa es del .999. Por su parte, el modelo propuesto basado en la cópula *t* de Student Agrupada refleja niveles de *CE* superiores al capital regulatorio y al estimado con el modelo de un factor tipo *KMV/CM*. Una ventaja del modelo propuesto es que permite modelar la dependencia existente entre los factores de riesgo y entre todos los grupos de créditos y productos que forman al portafolio; en consecuencia este modelo proporciona un efecto de diversificación del portafolio que no es posible apreciar en el modelo de un factor tipo *KMV/CM* ni en las reglas del enfoque más avanzado de *IRB* propuestas en Basilea II.

Finalmente, en la Tabla 13 se presentan los resultados obtenidos de implementar un modelo basado en matrices de transición de pagos vencidos. El supuesto de que el proceso de migración de pagos vencidos (incluido el estado absorbente de incumplimiento) sigue una cadena de Markov, en la práctica no se satisface (ver Lando et al. 2002), siendo éste una de las grandes deficiencias de este enfoque. Sin embargo, igual que en los dos últimos modelos presentados en esta sección, ni la estructura de dependencia de los parámetros de riesgo ni la dependencia entre incumplimientos son considerados bajo esta metodología que, no obstante, es ampliamente utilizada en la práctica.

Tabla 14
Comparación del modelo propuesto con un enfoque de matrices de transición.

Producto	Matrices Transición		Cópula t agrupada	
	PE	VaRqEmp 99.5%	PE	VaRqEmp 99.95%
A	9.40%	32.74%	9.01%	14.79%
B	4.11%	17.34%	5.48%	10.20%
C	0.36%	5.58%	0.43%	2.29%
D	0.91%	7.90%	1.41%	3.52%
E	6.21%	30.00%	7.61%	19.00%
F	6.57%	21.96%	9.02%	16.66%
G	8.23%	68.77%	9.83%	16.09%
H, I, J	3.04%	12.94%	3.20%	4.36%

Como la información disponible para su implementación contiene agregado al portafolio hipotecario, el número de productos en que se divide al portafolio total se reduce. Las estimaciones de PE difieren considerablemente en algunos productos, esto se debe muy probablemente a que el supuesto de transiciones no ocurre de manera Markoviana. Además, el CE estimado por el modelo de matrices de transición es muy superior al que se estima con la metodología de cópulas propuesta, lo cual se debe principalmente al hecho de que no es muy realista suponer que cada suma de exposición total por pagos vencidos es una v.a. Bernoulli con probabilidad de éxito la probabilidad de incumplimiento estimada.

6. Conclusiones

La metodología de cálculo del CE por riesgo de crédito propuesta en este trabajo de investigación muestra las ventajas de modelar la estructura de dependencia de los parámetros de riesgo mediante el uso de cópulas Elípticas Generalizadas y de Mezclas de Normales Agrupadas, para el caso concreto de familias del tipo t de Student. Este tipo de cópulas arrojó mejores resultados en la modelación del comportamiento de las pérdidas extremas del portafolio con respecto a las estructuras de dependencia comúnmente utilizadas en finanzas, como son la cópula Gaussiana y la cópula de Independencia. El ejercicio de bondad de ajuste reveló que el mejor ajuste lo proporciona la cópula t de Student Agrupada, al tiempo que no se encontraron diferencias significativas entre los ajustes de la cópula t de Student Simétrica y la cópula t de Student Generalizada en cada producto, no obstante que ésta última sí modela la asimetría en la estructura de dependencia. Una de las principales razones por las que se utiliza la cópula t de Student Agrupada obedece a que permite modelar la dependencia entre productos con diferentes índices de dependencia extrema, lo que hace posible la agregación de las pérdidas del portafolio total, resaltando el beneficio del efecto de diversificación.

Al comparar la metodología propuesta con otros modelos, se observó que el modelo de incumplimiento de un sólo factor tiende a subestimar el CE , mientras

que el enfoque de matrices de transición de pagos vencidos tiende a sobreestimarlos. Para la mayoría de los productos, el modelo de un sólo factor arroja un CE menor que el requerimiento de capital regulatorio bajo Basilea II, lo cual se debe al hecho de que la correlación de activos que el primero utiliza es calculada con información histórica, mientras que en el enfoque IRB las correlaciones de activos están predeterminadas, así como su nivel de confianza (99.9%). Las estimaciones de CE en el modelo propuesto parecen ser, en general, superiores al requerimiento regulatorio que establecen las reglas de capitalización del enfoque más avanzado de IRB propuestas en Basilea II. Estos resultados obtenidos ponen en duda la aplicabilidad de tales reglas en el mercado mexicano, por lo que el capital regulatorio bajo Basilea II puede ser insuficiente para cubrir el capital en riesgo que enfrentan las instituciones financieras que otorgan créditos al menudeo. Una ventaja de la metodología expuesta, es su gran flexibilidad para modelar no sólo la dependencia existente entre los factores de riesgo sino, también, entre todos los grupos de créditos y productos que forman al portafolio; lo que hace posible que se muestren las bondades de la diversificación del portafolio que no es posible apreciar en el modelo de un factor tipo KMV/CM ni en las reglas del enfoque más avanzado de IRB propuestas en Basilea II.

Apéndice

Denótese las realizaciones de n copias independientes del vector aleatorio \mathbf{X} (el conjunto de los datos) por

$$\{(x_1, \dots, x_{dt})'\}_{t=1}^n \equiv \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ x_{d1} & \dots & \dots & x_{dn} \end{pmatrix} = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \dots \quad \mathbf{x}_n]$$

Algoritmo 1

- (i) Determinar las correlaciones tau de Kendall de cada pareja (X_i, X_j) mediante el estimador $\hat{\rho}_\tau^{(n)}(X_i, X_j) = \binom{n}{2}^{-1} \sum_{k>l} \text{sign}[(x_{ik} - x_{il})(x_{jk} - x_{jl})]$;
- (ii) Usar la ecuación (11) para calcular $\hat{\rho}_{ij}$ y estimar así la matriz de correlación ρ_{ij} ;
- (iii) Estimar los coeficientes de dependencia de la cola inferior $\hat{\Lambda} = (\lambda_{ij}^n)_{i,j=1}^d$ como

$$\lambda_{ij}^n \equiv \hat{\lambda}_l^{(n,r)}(X_i, X_j) = \frac{1}{n} \sum_{k=l}^n \frac{\sqrt{2}}{r} \mathbf{1}_{\{Q_k < r\}} \text{sen} 2\phi_k,$$

donde para $k = 1, \dots, n$ las cantidades (Q_k, ϕ_k) satisfacen las ecuaciones $\hat{F}_i(x_{ik}) = Q_k \text{sen} \phi_k$, $\hat{F}_j(x_{jk}) = Q_k \text{cos} \phi_k$, donde cada \hat{F}_i es la f. d. empírica de la componente X_i ;

(iv) Estimar el índice de la cola α de la v.a. espectral R como

$$\hat{\alpha}^n = \arg \min \left\| \lambda\left(\alpha, \frac{\pi}{4}(1 - \hat{\rho}_{ij})\right)_{ij} - \hat{\Lambda}^{(n)} \right\|_{\times},$$

donde $L(\alpha, \hat{\rho}_{\tau})^{(n)}$ es la matriz $\lambda(\alpha, \frac{\pi}{4}(1 - \hat{\rho}_{ij}))_{ij}$, la función $\lambda(\alpha, x)$ está definida como en (12) y $\|A\|_{\times} = \sum_{i,j=l}^d A_{ij}^2$ define la norma en L^2 definida sobre $M_{d \times d}(\mathfrak{R})$.

Algoritmo 2

(i) Estimar los parámetros ρ_{ij} y α de una distribución Elíptica Generalizada sobre el vector \mathbf{X} usando el Algoritmo 1;

(ii) Ajustar a cada componente de \mathbf{X} alguna f.d. marginal F_j^* continua. Si por ejemplo, el número de observaciones n no es muy grande, se puede utilizar un kernel de suavización;

(iii) Simular observaciones y_1, \dots, y_N de un vector \mathbf{Y} cuya representación está dada por la ecuación (5), utilizando las estimaciones de los parámetros anteriores y estructura de dependencia de dicha representación (bajo la especificación de la f.d. de la v.a. espectral R);

(iv) Construir observaciones $\hat{u}_1, \dots, \hat{u}_N$ de la cópula empírica con las observaciones simuladas y_1, \dots, y_N , a saber,

$$\hat{u}_i = (U_{i1}, \dots, U_{id})' \left(\hat{F}_1(y_{i1}), \dots, \hat{F}_1(y_{id}) \right)',$$

donde \hat{F}_j es un estimador de la f.d. de Y_j , por ejemplo, se puede usar una variante de la f.d. empírica $\hat{F}_j(x) = \frac{1}{n+1} \sum_{i=1}^n \mathbf{1}_{\{y_{ij} < x\}}$.

(v) Con las observaciones simuladas $\hat{u}_1, \dots, \hat{u}_N$, las cuales incorporan la estructura de dependencia de la cópula C , construir $\hat{x}_j = F_j^{*-1}(\hat{u}_j)$, $j = 1, \dots, N$. Los vectores $\hat{x}_1, \dots, \hat{x}_N$ son observaciones simuladas del vector original \mathbf{X} .

Bibliografía

- Basel Committee on Banking Supervision (BCBS) (2004). International Convergence of Capital Measurement and Capital Standards. A Revised Framework.
- Biggs, D., B. De Ville y E. Suen (1991). A method of choosing multiway partitions for classification and decision trees. *Journal of Applied Statistics*, 62, pp. 18-49.
- Crosbie, P. y J. Bohn (2002), Modelling Default Risk. *KMV working paper*.
- Crouhy, M., D. Galai y R. Mark (2000). A comparative analysis of current credit risk models. *Journal of Banking and Finance*, 24, 59-117.
- De Andrade, F. y L. Thomas (2004). Structural Models In Consumer Credit. Risk and Insurance 0407001, *EconWPA*.
- Díaz, A. (2003). Teoría de Valores Extremos para sucesiones de variables aleatorias dependientes, Tesis de licenciatura, UNAM.

- Embrechts, P., C. Klppelberg y T. Mikosch (1997). Modelling Extremal Events for Insurance and Finance. Springer Verlag, Berlin.
- Embrechts, P., A. McNeil y D. Straumann (2002). Correlation and dependency in risk management: properties and pitfalls. *Risk Management for Central Bank Foreign Reserves*, 243-261, Frankfurt: European Central Bank.
- Frahm, G. y M. Junker (2003). Generalized elliptical distributions: models and estimation. *Research Center Caesar Financial Engineering*.
- Frey, R. y A. McNeil (2003). Dependent defaults in models of portfolio credit risk. *Journal of Risk*, 6(1), 59-92.
- Hult, H. y F. Lindskog (2001). Multivariate extremes, aggregation and dependence in elliptical distributions. Research paper, *RiskLab*.
- Joos, P., K. Vanhoof, H. Ooghe y N. Sierens (1998). Credit classification: A comparison of logit models and decision trees. *10th European Conference on Machine Learning*, April 24, Chemnitz (Germany), p. 59-72.
- Lando, T. y T. Skodeberg (2002). Analyzing rating transitions and rating drift with continuous observations. *Journal of Banking and Finance*, 26, 423-444.
- Lindskog, F., A. McNeil y U. Schmock (2003). Credit Risk - Measurement, Evaluation and Management. Physica-Verlag, Heidelberg.
- Panchenko, V. (2005). Goodness of fit test for copulas. *Physica A*, 355, 176-182.
- Perli, R. y W. Nayda (2004). Economic and Regulatory Capital Allocation for Revolving Retail Exposures. *Journal of Banking & Finance*, 28(4), 789-809.
- RMA (2003). Retail Credit Economic Capital Estimation-Best Practices. *The Risk Management Association (RMA)*.