

# Pérdidas inesperadas por riesgo operativo en una entidad financiera con Teoría de Cópulas

## Unexpected Losses for Operational Risk in a Financial Institution with Copula Theory

---

Gloria Inés Macías Villalba\*

(Fecha de recepción: 13 de febrero de 2017. Fecha de aceptación: 26 de junio de 2017)

### RESUMEN

En la búsqueda por el equilibrio y estabilidad del sistema financiero; Basilea II propone métodos avanzados para la medición del riesgo operativo que pueden ser desarrollados por las entidades involucradas y ajustados a los requerimientos del supervisor. Estas metodologías permiten a las entidades estimar las pérdidas por riesgo operativo de manera más objetiva, pero los modelos de cálculo de pérdida, generalmente se centran en los eventos de mayor frecuencia pero de menor cuantía en el cuerpo de la distribución, sin embargo, también es importante considerar eventos menos frecuentes pero de alto impacto, esto es, eventos extremos cuya distribución debe ser analizada de manera separada y con modelos específicos.

Entre los métodos utilizados para calcular las pérdidas inesperadas por riesgo operativo, están la teoría de valores extremos TVE, y la teoría de cópulas. Tanto la teoría de valores extremos como la de cópulas buscan distribuciones de ajuste que simulen el comportamiento de los datos extremos, sin embargo, la teoría de cópulas añade un elemento fundamental que hace referencia a la dependencia entre las variables, que para el caso del riesgo operativo sería importante contemplar debido a que se analizan pérdidas por tipos de eventos en cada una de las líneas de negocio.

Clasificación JEL: G210, G220 y G170.

**Palabras clave:** Riesgo operativo, distribución de pérdidas agregadas, pérdidas inesperadas, cópulas, eventos extremos.

---

\* Grupo de Investigación en Ingeniería Financiera, Universidad Autónoma de Bucaramanga, Bucaramanga, Santander, Colombia  
gmacias@unab.edu.co.

## ABSTRACT

*In the search for balance and stability in the financial system, Basel II proposes advanced methods to measure operational risk that can be developed by the entities involved and adjusted to the requirements of the regulator. These methodologies allow such entities to estimate losses due to operating risk more objectively, however loss calculation models are generally focus on higher frequency but less value events from the body of the distribution. Nevertheless, it is also important to consider less frequent but high impact events, this is extreme events whose distribution must be analyzed separately with specific models.*

*Among the methods used to calculate unexpected losses due to operational risk are TVE Extreme Value Theory and Copula Theory. Both, Extreme Value Theory and Copula Theory model distributions simulate the behavior of extreme data, however, copulas theory adds a fundamental element that makes reference to the dependence between variables, which in the case of the operational risk would be important to consider since losses are analyzed by types of events in each business line.*

*JEL clasification: G210, G220 y G170*

**Keywords:** *Operational Risk, Distribution of Aggregate Losses, Expected Losses, Copula, Extreme Events.*

## Introducción

Las empresas se enfrentan a diferentes situaciones de riesgo, entendidas como la posibilidad de sufrir una pérdida o daño a un activo. El riesgo es inevitable si forma parte fundamental de la toma de decisiones sobre la actividad principal de la empresa; de ahí que hay riesgos inherentes que se deben gestionar para disminuir el impacto o la posible pérdida, y según el tipo de actividad que desarrolle la empresa o negocio.

El sector financiero es el más regulado en materia de gestión de riesgos, con lineamientos de orden internacional desde el comité de Basilea, y en el caso de Colombia con la Superfinanciera. Sin embargo, la norma sugiere a las entidades que desarrollen modelos propios o internos que se ajusten a las condiciones exigidas, es por esto que los analistas necesitan conocer diferentes metodologías para medir los riesgos financieros que se adapten a las características de las variables, y enmarcarlos dentro del proceso de gestión de riesgo.

Por otra parte, con la llegada de Basilea II, se incluyó el riesgo operativo dentro de las diferentes líneas de negocio y tipos de eventos, para el cálculo del requerimiento de capital mínimo. Con este nuevo reto, las entidades del sector financiero iniciaron su tarea en la identificación de los factores de riesgo operativo; y con análisis de tipo cualitativo o semicuantitativo desarrollaron estrategias de mitigación adecuadas según la exposición frente al riesgo. Este análisis no había avanzado hacia la medición cuantitativa, debido a que las entidades no contaban con registros de riesgo operativo que les permitieran tener suficientes datos para aplicar metodologías más objetivas y robustas que se ajustaran a las características de las distribuciones de pérdidas por riesgo operativo.

Sin embargo, en Basilea II (Comité de Basilea, 2004), se dieron pautas para el cálculo del requerimiento de capital por riesgo operativo con tres métodos cuantitativos: el Indicador Básico, el Método Estándar y los Métodos de Medición Avanzada (MMA). Mediante el enfoque MMA los mismos bancos definen su requerimiento de capital según los modelos internos de su propio desarrollo, siempre y cuando cumplan con el aval de la entidad supervisora. Una de las metodologías propuestas por Basilea en MMA, es el enfoque de distribución de pérdidas (Loss Distribution Approach, LDA).

El modelo LDA ha sido ampliamente utilizado en el entorno internacional, en el sector financiero, ya que provee una técnica adecuada para satisfacer la necesidad de presentar el requerimiento del capital económico sugerido por Basilea y solicitado por los entes supervisores de los países miembros. En diferentes trabajos se ha utilizado el modelo LDA para el cálculo del requerimiento de capital por la exposición a riesgo operacional, como (Frachot, Georges, y Roncalli, 2001), (Jöhnemark, 2012), (Shevchenko y Peters, 2013), (Chernoval, Svetlozar T., y Fabozzi, 2007), entre otros.

En Colombia, el modelo LDA se ha aplicado durante menos tiempo, sin embargo, existen trabajos realizados por varios autores que aportan en su avance, para citar algunos: (Mora A. a., 2010), calcula las pérdidas por riesgo operativo utilizando dos metodologías: la primera con la aplicación del modelo propuesto por (Böcker y Klüppelberg, 2005), con ajuste de distribuciones subexponenciales, y un segundo método con la teoría de valores extremos; (Franco y Murillo, 2008) aplica el modelo LDA para los siete tipos de eventos por riesgo operativo a una entidad comisionista de bolsa para estimar la carga de capital; (Bernal, 2013) utiliza la literatura y normativa del sector financiero del modelo LDA y lo contextualiza a una Pyme; por su

parte (Arias, 2010), utiliza el proceso del LDA con los datos de la cola de la distribución y desarrolla la teoría de valores extremos.

En la aplicación de los métodos de medición avanzada, en los análisis estadísticos de los datos, se encuentra que en general la distribución de pérdidas<sup>1</sup> por riesgo operativo presenta problemas de elongación y asimetría, es decir, alta curtosis y sesgada a la derecha. Lo anterior implica un análisis más detallado de la distribución de pérdidas, y no tomarla como un todo; de ahí la necesidad de utilizar modelos de medición diferentes. Por una parte usar métodos tradicionales para modelar el cuerpo de la distribución, y por otra, hacer uso de modelos de medición especializados para las pérdidas inesperadas ubicados en la cola de la distribución.

Las distribuciones asimétricas y leptocúrticas se dan porque los eventos por riesgo operativo tienen eventos que presentan alta frecuencia pero con impactos o severidades bajas y eventos que pocas veces se dan, pero con un impacto muy alto. Sin embargo, estos últimos, por el hecho de tener baja frecuencia no se pueden ignorar, hay que modelarlos y calcular las pérdidas inesperadas, de esta forma, se establecen acciones para que se controlen las causas que generan el riesgo.

El trabajo de investigación pretende proporcionar una visión del modelo de cópulas para ajustarla al comportamiento de las pérdidas por riesgo operativo y enfocado al estudio de las distribuciones de las colas. Analizando en cada una, las características, propiedades, y la aplicación con datos de una entidad del sector financiero colombiano para calcular la pérdida inesperada.

Para las distribuciones de las colas, hay dos teorías que han tenido gran relevancia en la gestión de otros riesgos; corresponde a la teoría de valores extremos TVE, y la teoría de cópulas de extremos. Tanto la teoría de valores extremos como la de cópulas buscan distribuciones de ajuste que simulen el comportamiento de los datos extremos, sin embargo la teoría de cópulas añade un elemento fundamental que hace referencia a la dependencia no lineal entre los eventos de riesgo operativo.

Con esta investigación se busca presentar una metodología de medición de riesgo operativo que se ajuste a las características de las distribuciones de severidades para el cálculo de las pérdidas no esperadas por eventos de riesgo operativo, y además que sirvan de apoyo en el sector financiero colombiano.

---

<sup>1</sup> También llamada Distribución de Severidad.

## 1. Marco conceptual

El desarrollo de la investigación requiere el fundamento teórico en temas relacionados con los modelos a utilizar en la medición de las pérdidas por riesgo operativo. Teniendo en cuenta que la medición de las pérdidas se centra en las colas de la distribución, se abordarán temas como la teoría asintótica, la teoría de valores extremos y la teoría de cópulas.

Las distribuciones de pérdidas por eventos de riesgo operativo presentan características como asimetría y alta curtosis, por tanto, se pueden catalogar como asintóticas. Las distribuciones asintóticas se encuentran ubicadas en la estadística de los valores extremos, que describe eventos no comunes. Es una teoría que tiene gran utilidad en diferentes campos, y ahora en el campo de las finanzas cobra importancia para el análisis de riesgos financieros, como es el caso de las pérdidas por eventos de riesgo operativo; teniendo en cuenta que la principal debilidad de la distribución normal es su incapacidad para modelar datos de alta asimetría y curtosis.

Para iniciar, es importante mencionar algunos estadísticos de orden. Las variables  $(x_1, x_2, x_3, x_4, \dots, x_n)$ , forman una muestra aleatoria simple, son independientes e idénticamente distribuidas (i.i.d). Los estadísticos de orden mínimo y máximo son los que más se destacan de la muestra, el  $x_{1:n} = \text{Min}(x_1, x_2, x_3, x_4, \dots, x_n)$ , y el  $x_{n:n} = \text{Max}(x_1, x_2, x_3, x_4, \dots, x_n)$ . El rango también es considerado un estadístico de orden, indica en forma general, el grado de dispersión de los datos contenidos en la muestra. Se determina como  $R = x_{n:n} - x_{1:n}$ . Otro estadístico de orden es la mediana, que separa la distribución en la mitad, es decir, establece el valor a partir del cual 50% de la distribución es mayor que él. Los estadísticos de orden máximo y mínimo son los que determinan la teoría de valores extremos, y de acuerdo con el tipo distribución se pueden hacer análisis útiles en situaciones prácticas.

La teoría asintótica surge cuando se analizan las distribuciones para las siguientes situaciones: a)  $n$  tiende a infinito; b) Se desconoce la función de distribución  $F(x)$  de la población, o c) El tamaño de la muestra es desconocido. Según (Castillo, 1988), si se aplica la primera situación, a) cuando  $n$  tiende a infinito, se pueden determinar las distribuciones límites tanto para el máximo como para el mínimo.

En cuanto a distribuciones asintóticas de máximos y mínimos, los estudios realizados por Fisher y Tippett en 1928, demostraron que existen tres tipos de distribuciones límites no degeneradas para máximos y mínimos. Estas distribuciones son Frechét, Gumbel y Weibull que presentan caracte-

rísticas asintóticas en términos de alta curtosis y asimetría. Sin embargo, Fréchet (1927) realizó los primeros análisis, obteniendo una distribución asintótica para los valores grandes.

La teoría de valores extremos está relacionada con los eventos que ocurren rara vez, es decir, los eventos que ocurren con poca probabilidad, pero que en términos de impacto o pérdida pueden ser catastróficos, de ahí la importancia de analizar esta información y determinar el tipo de distribución que permita modelar y realizar estimaciones. Los eventos que presentan estas características se encuentran en la cola de la distribución, por lo tanto otro aspecto a tener en cuenta es a partir de qué valor se pueden considerar datos extremos.

Una de las primeras aplicaciones de la teoría de valores extremos la presenta Bernoulli en 1709 en el campo actuarial, más adelante se destaca el trabajo realizado por von Bortkiewicz en 1922 con el análisis de los grandes valores (máximos y mínimos), donde resalta la importancia de realizar la aproximación de Poisson a la distribución Binomial. Sin embargo, en 1921 Fisher analizó la situación de los valores atípicos (outliers) de una distribución y la necesidad de ser modelados en distribuciones diferentes a la normal.

Se puede decir que el teorema de la teoría de valores extremos fue propuesto por Fisher y Tippett en 1928, cuando demostraron cuáles eran las distribuciones límites de extremos no degeneradas que satisfacen las condiciones de máximos y mínimos, y lo presentan a través de la Distribución Generalizada de Valores Extremos.

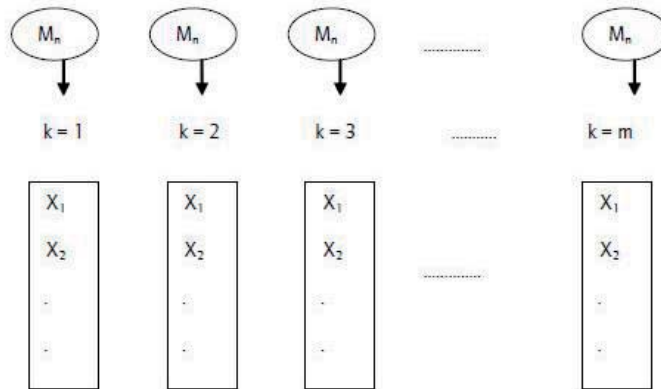
## 2. Marco metodológico

En la teoría de valores extremos es importante realizar una etapa inicial para determinar cuáles son las observaciones a tener en cuenta para la distribución de valores objeto de estudio que permita realizar los cálculos necesarios. Por lo anterior la pregunta es ¿Cómo obtener los valores extremos?, para ello existen especialmente dos técnicas, los bloques máximos y picos sobre el umbral (POT).

En el método de Máximo por Bloques, se divide la muestra que se ha tomado durante un periodo de tiempo en un número determinado de bloques, y de cada bloque se selecciona el evento de mayor valor. La serie de valores extremos se obtiene con el máximo de la serie de cada bloque, y a partir de esta información se realizan los análisis de la teoría de valor ex-

tremo. Los parámetros se estiman con la máxima verosimilitud asumiendo independencia entre los máximos de los bloques. En el gráfico se observa la metodología, donde la muestra se divide en  $m$  bloques y de cada bloque se identifica el  $M_n$  (valor máximo), ver Figura 1.

Figura 1. Método bloques máximos



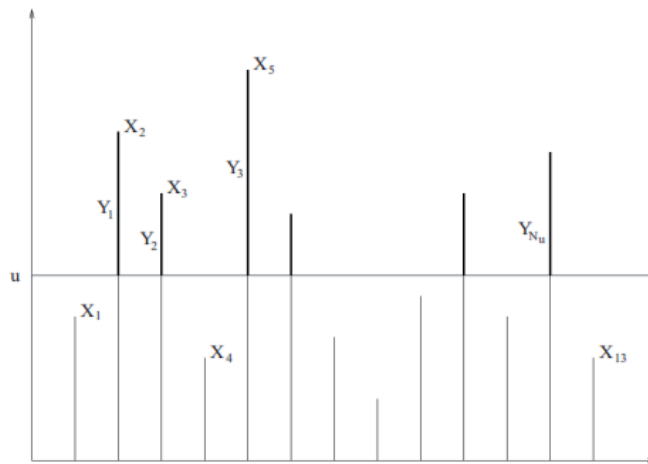
Fuente: (Mora A. b., 2010).

Una de las ventajas de esta metodología es poder tener observaciones a lo largo de toda la muestra, y esto reduce la dependencia entre las observaciones extremas. La dificultad que puede presentar esta metodología es la selección del tamaño del bloque, porque se puede incurrir en el error de bloques pequeños que generan sesgo y los bloques grandes generan varianzas altas para la estimación de los parámetros.

La técnica Picos sobre el Umbral (POT), es considerada más práctica que el máximo por bloques, especialmente cuando las muestras son muy grandes. Se obtiene la distribución de valores extremos identificando las observaciones que excedan un umbral dado  $u$ . En esta metodología el número de valores extremos no depende del número de datos de la muestra, depende del umbral seleccionado.

La Figura 2 muestra las observaciones  $X_1, X_2, X_3, \dots, X_n$ , independientes e idénticamente distribuidas, y  $Y_i = X_i - u$ , para todos los  $X_i > u$ . Con los  $Y_{N_u}$ , se forma la distribución de excesos sobre el umbral.

Figura 2. Método POT



Fuente: (Embrechts, Kluppelberg y Mikosch, 1997).

Con los datos extremos  $X_i > u$  se pueden calcular los excesos sobre el umbral  $u$  llamados  $Y_{N_u}$  de tal forma que se obtiene una distribución especial. La selección del umbral conlleva problemas de sesgo y varianza, si el umbral es muy bajo se presenta sesgo en la estimación de los parámetros porque serían muchos datos que no cumpliría la teoría asintótica, y si el umbral es muy alto se obtienen pocos datos lo que genera una alta varianza en la estimación de los parámetros.

Para determinar el umbral adecuado existen varios métodos, uno es el estimador de Hill. Propuesto por Hill Bruce M. en 1975. Es la técnica más popular para determinar el índice de cola, o parámetro de forma  $\xi$  para las distribuciones de valores extremos. Su aplicación está condicionada sólo para distribuciones con dominio máximo de atracción de Fréchet, es decir cuando  $\xi > 0$  lo que hace su aplicación limitada.

Para la aplicación del estimador de Hill, las observaciones deben estar ordenadas de mayor a menor. Dada una muestra  $(x_1, x_2, x_3, x_4, \dots, x_n)$  que procede de una población, si los valores de secuencia  $x_1, x_2, x_3, x_4, \dots, x_n$ , se ordenan en orden decreciente,  $X_1 \geq X_2 \geq X_3 \geq X_4 \geq X_n$ . El estimador de Hill se calcula con la siguiente fórmula:



$$\xi_{n,k}^{(H)} = \left( \frac{1}{k} \sum_{i=1}^k \ln X_i - \ln X_{k+1} \right)^{-1}, 2 \leq k \leq n$$

Este estimador presenta las propiedades asintóticas de consistencia y normalidad. En el estimador de Hill  $k$  es el umbral en el extremo de la distribución. En la anterior ecuación se puede observar que el estimador de Hill es el promedio de las diferencias entre las observaciones más extremas menos la  $k + 1$  observación más cercana al extremo.

El problema en la aplicación del estimador de Hill es la elección del  $k$ , es decir el umbral o dato a partir del cual se obtienen las diferencias. Una forma de calcular el umbral para determinar  $k$  es por medio del gráfico de Hill, donde se grafican los estimadores de Hill para diferentes valores de  $k$  que están asociados a diferentes umbrales. Se determina cuál es el valor de  $k$  apropiado (ni muy grande ni muy pequeño) cuando en el gráfico de Hill se observa estable.

Otra técnica para determinar el umbral es el gráfico de exceso medio. Se toma el valor de umbral ( $u$ ) de tal forma que la función de exceso medio  $e(x)$  sea lineal con pendiente positiva, donde,  $e(x)$  = Media aritmética de los valores que exceden  $u$ , y la fórmula se muestra en la ecuación (1).

$$e(x) = \frac{\sum_{i=1}^n (x_{i,n} - u)^+}{\sum_{i=1}^n 1_{(x_{i,n} > u)}} \quad (1)$$

Con los datos de la muestra se obtienen las medias de los excesos para diferentes valores de umbrales ( $u$ ), y se identifica el umbral óptimo donde la función inicia con pendiente positiva.

Por otra parte, la teoría de cópulas se utilizará como el método principal de esta investigación. En el análisis multivariado es muy probable encontrar relaciones entre las variables individuales, y de ahí la necesidad de determinar en qué forma el movimiento de una variable depende de otra, tomada como una dependencia transversal. Es muy común utilizar el coeficiente de correlación lineal de Pearson para encontrar el grado de relación de dos variables, pero tiene el defecto de asumir normalidad de las distribuciones individuales, y se sabe que la mayoría de las series financieras no presentan esta característica por condiciones de colas pesadas o asimetría en las distribuciones. Por otro lado, las relaciones entre las variables por lo general no son lineales.

La teoría de cópulas propuesta por (Sklar A., 1959), corrige los problemas anteriores debido a que determina la dependencia entre las distribuciones como una función conjunta a través del comportamiento individual de las variables sin importar el tipo de distribución, es decir, no requiere el supuesto de normalidad. Además, con la teoría de cópulas se puede describir en forma completa cómo es la dependencia entre las variables, o la estructura de dependencia sin resumirla en un solo número; y lo más importante es que permite separar la función de distribución conjunta en dos elementos de análisis: *a)* las distribuciones marginales de las variables individuales o unidimensionales y, *b)* la interdependencia de las probabilidades.

Una cópula se puede definir como una función de distribución multivariada; de manera más concreta, es una función de distribución conjunta cuyas distribuciones marginales univariadas se distribuyen uniformemente entre  $[0,1]$ , y relaciona un grupo de variables. Entonces, la cópula es una función que combina las distribuciones marginales de las variables en una distribución multivariante específica.

El Teorema de Sklar se considera como uno de los aportes fundamentales en la teoría de cópulas, porque establece la relación que existe entre la función multivariada y las distribuciones marginales. De acuerdo con el trabajo realizado por Abe Sklar en 1959 (Sklar A., 1959), se presentan las siguientes definiciones y colorario.

Si  $G$  es una función de distribución conjunta  $n$ -dimensional con marginales 1-dimensional  $F_1, \dots, F_n$  entonces existe una función  $C$  (llamada “ $n$ -cópula”) desde el  $n$ -cubo unidad al intervalo unidad tal que, ecuación (2):

$$G(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)) \quad (2)$$

Sí  $F_1, \dots, F_n$  son continuas,  $C$  es única y su rango está determinado en:

$$\text{ran}(F_1) \times \dots \times \text{ran}(F_n)$$

## Funciones de Distribución

Sea  $n$  un positivo entero. Un  $n$ -intervalo

$$N = [(x_1, \dots, x_n), (y_1, \dots, y_n)], \quad (x_m \leq y_m \text{ para todo } m \leq n), \quad (3)$$

Es el producto cartesiano de los intervalos cerrados  $[x_1, y_1], \dots, [x_n, y_n]$ . En particular, el  $n$ -cubo unidad  $I^n$  es la  $n$ -ésima potencia del intervalo unidad cerrado  $I = [0,1]$ , y el  $n$ -espacio continuo  $E^n$  es la  $n$ -ésima potencia de la recta real  $E = [-\infty, +\infty]$ .

Una función real  $G$  en la posición  $n$  es una función cuyo dominio,  $Dom G$ , es un subconjunto no vacío de  $E^n$ , y cuyo rango,  $Ran G$ , es un subconjunto de la línea recta no continua.

**Definición 2.** Una función de distribución  $n$ -dimensional es una función real  $G$  con  $n$ -posiciones cuyo dominio es todo el conjunto  $E^n$  y cuyo rango es un subconjunto de  $I$ . Satisface las siguientes condiciones:

- a)  $G(+\infty, \dots, +\infty) = 1$
- b)  $G(x_1, x_2, x_3, \dots, x_n) = 0$  si  $x_m = -\infty$  para cualquier  $m \leq n$
- c)  $G$  es  $n$ -creciente
- d)  $F_m(x_m) = G(+\infty, \dots, +\infty, \dots, x_m, +\infty, \dots, +\infty)$

Una consecuencia de b) y c) es que una función de distribución  $n$ -dimensional no decreciente en cada una de sus variables separadamente. Inmediatamente, cada marginal de una función de distribución  $n$ -dimensional es en sí misma una función de distribución 1-dimensional (univariada).

**Definición 3.** Una cópula  $n$ -dimensional (brevemente,  $n$ -cópula) es una función real  $C$  con  $n$  posiciones con  $Dom C = I^n$ ,  $Ran C = I$ , la cual satisface las siguientes condiciones:

- a)  $C(1, \dots, 1, x_m, 1, \dots, 1) = x_m$  para cada  $m \leq n$  y todos los  $x_m$  en  $I$
- b)  $C(x_1, \dots, x_n) = 0$  si  $x_m = 0$  para cualquier  $m \leq n$
- c)  $C$  es  $n$ -creciente

Según las condiciones a) - c) existe una cópula única: la función idéntica de en  $I$ ; según esto, toda  $n$ -cópula es no decreciente en cada variable separada y conjuntamente es continua.

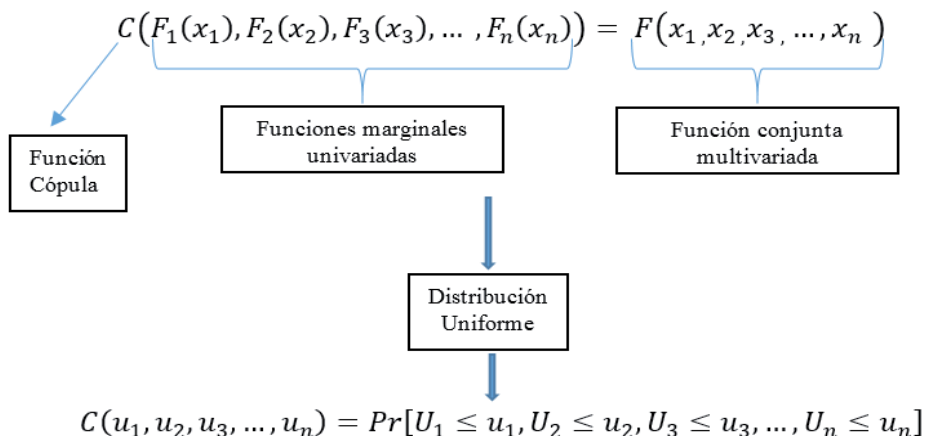
**Corolario.** Sean  $X_1, \dots, X_n$  variables aleatorias definidas en el mismo espacio probabilístico. Las distribuciones marginales de las variables  $X_1, \dots, X_n$  están denotadas por  $F_1, \dots, F_n$  respectivamente, y la función de distribución conjunta de  $X_1, \dots, X_n$  denotada por  $G$ . Entonces  $F_m$  es la  $m$ -ésima marginal de  $G$  para todo  $m \leq n$  entonces existe una  $n$ -cópula  $C$  para todo. Sean  $X_1, \dots, X_n$  tal que:  $G(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n))$ .

Si cada  $F_m$  es continua, entonces la cópula es única.

Por otro lado como notación, una cópula  $C : [0,1]^n \rightarrow [0,1]$ , es una función de distribución cuyas distribuciones marginales univariadas son uniformes con valores entre  $[0,1]$ . (Giacomini, 2005), afirma que una cópula  $n$ -dimensional  $C : [0,1]^n \rightarrow [0,1]$  satisface las propiedades propuestas por (Sklar A., 1959). Según las propiedades, las cópulas son distribuciones uniformes multivariadas. La primera y tercera propiedades afirman que las cópulas son funciones conectadas y que todas las cajas -dimensionales con vértices en  $[0,1]$  tienen un  $C$ -volumen no negativo. Ambas garantizan que las cópulas son funciones de distribución, mientras que la segunda propiedad dice que las cópulas tienen distribuciones marginales uniformes.

(Salinas, Maldonado, y Díaz, 2010), explican el teorema de Sklar de la siguiente forma:

Figura 3: Esquema de la Cópula



Significa que la cópula es una función conjunta multivariada que captura la información sobre la dependencia de un grupo de variables aleatorias  $x_1, x_2, x_3, \dots, x_n$ , y está conformada por funciones marginales univariadas o individuales que se distribuyen uniformemente, sin importar el tipo de distribución que tenga cada una de las variables.

Como consecuencia del teorema de (Sklar A., 1959), Fréchet-Hoeffding establecen los límites superior e inferior para las cópulas, y según (Melo & Becerra, 2008) se pueden incluir como otra propiedad,

- 1) Para cada cópula  $C(u_1, u_2, u_3, \dots, u_n) = C(u)$  se establecen los siguientes límites:

$$L_F(u_1, u_2, u_3, \dots, u_n) \leq C(u_1, u_2, u_3, \dots, u_n) \leq U_F(u_1, u_2, u_3, \dots, u_n) \quad (4)$$

Donde  $L_F = \max\{\sum_{i=1}^n u_i + 1 - n, 0\}$ ; y  $U_F = \min\{u_1, u_2, u_3, \dots, u_n\}$

Entonces,

$$\max\{\sum_{i=1}^n u_i + 1 - n, 0\} \leq C(u) \leq \min\{u_1, u_2, u_3, \dots, u_n\} \quad (5)$$

Si se analiza el significado literal de cópula, se relaciona con términos como atadura, enlace, ligar algo con otra cosa, entonces tiene que ver con el concepto de dependencia. Por lo tanto, es importante determinar cuáles son las diferentes formas en que las variables aleatorias se asocian a través de las cópulas. Se trata de encontrar medidas de dependencia que no tengan las limitaciones del coeficiente de correlación de *Pearson* por ser un único valor que analiza la fuerza y dirección entre dos variables aleatorias, pero como una relación lineal. Se proponen dos medidas de dependencia o asociación: Las medidas de correlación de rango, y los coeficientes de dependencia en las colas.

Las medidas de correlación de rango permiten identificar la correlación entre variables aleatorias a través de un rango, sin suponer que esta relación es lineal, y pueden determinar el grado de asociación de valores grandes de una variable con los valores grandes de la otra variable, de la misma forma con los valores pequeños. Es decir, estas medidas se asocian con el concepto de concordancia entre variables aleatorias. Además, según (Melo y Becerra, 2008), pueden ser obtenidas en función de la cópula a la cual se ajusta la función de distribución conjunta de las variables aleatorias y no de las distribuciones marginales, lo que implica que las medidas de correlación

de rango son invariantes para transformaciones monótonas crecientes de las variables aleatorias involucradas.

Lo anterior, teniendo en cuenta que la cópula muestra todas las posibles dependencias entre las variables aleatorias, si es perfecta positiva es una dependencia comonotónica, y si es perfecta negativa es contramonotónica. Por otro lado, la cópula es invariante ante transformaciones monótonas crecientes, es decir, que la variable aleatoria tiene la misma cópula que la transformación monótona de dicha variable.

Se presenta como medida de correlación de rango, la Tau de Kendall ( $\tau$ ), que mide el grado de dependencia entre dos variables aleatorias y se define como la diferencia entre las probabilidades de concordancia y discordancia entre pares independientes. La concordancia se da, si el cambio de una variable se asocia con el cambio de la otra variable en el mismo sentido; por ejemplo, un cambio positivo de una variable se asocia con un cambio positivo de la otra variable. Por otra parte, la discordancia se da, si un cambio positivo de una variable tiende a asociarse a un cambio negativo de la otra variable, y viceversa. La  $\tau$  de Kendall está determinada por la cópula de la distribución conjunta, y no por las distribuciones marginales.

Existen diferentes tipos de cópulas agrupadas en familias de cópulas dependiendo del criterio utilizado, y de la aplicación que se le quiera dar. Según (Melo y Becerra, 2008) una clasificación es por los tipos extremos de relación de dependencia: comonotonicidad, contramonotonicidad e independencia; otra, por la forma funcional de la cópula: implícitas y explícitas; y por último una tipología muy usada es por las características particulares de las cópulas: Elípticas, Arquimedianas, de Valor Extremo, y las Arquimex. Para el desarrollo de esta investigación se utiliza la última clasificación, es decir, se establecen las familias de cópulas según las características específicas, pero con tres primeras familias, que son las más comunes para análisis de riesgos financieros.

Las cópulas elípticas, son muy utilizadas en los mercados financieros, y se asocian a las distribuciones elípticas. Su característica más importante son que se asocian a variables aleatorias, con función de distribución multivariada simétrica y representan relaciones de dependencia simétricas ya sean de la cola izquierda o derecha de las distribuciones implicadas. Además, permiten calcular de manera sencilla los coeficientes de correlación de rango y los coeficientes de dependencia en las colas. Las más conocidas en esta familia de cópulas son la Cópula normal o Gaussiana y la Cópula  $t$  de student.

Por su parte, las cópulas Arquimedianas se caracterizan porque capturan más información sobre la estructura de dependencia, por tanto son las más estudiadas y utilizadas en la práctica; al recoger diferentes tipos de estructuras de dependencia, adicionales a las que capturan las elípticas. Además, se pueden construir con mayor facilidad. Gran variedad de cópulas paramétricas pertenecen a la familia Arquímedes, su representación permite trabajar cópulas con función de distribución multivariada y reducirla a una función única univariada. Son varias las cópulas que pertenecen a este grupo de Arquimedianas. Las principales son: las cópulas de Frank, las de Clayton y la de Gumbel.

Como la cópula de Gumbel es la utilizada en la investigación, se amplía un poco más su descripción.

En la cópula de Gumbel, el generador de esta cópula, según Melo y Becerra (2008), está dado por,  $\varphi(t) = (-\log(t))^\theta$ , con  $\theta > 1$ .

La cópula de Gumbel tiene la forma,

$$C^{Gu}(u_1, u_2, u_3, \dots, u_n) = \exp\left(-(\tilde{u}_1^\theta + \tilde{u}_2^\theta + \dots + \tilde{u}_n^\theta)^{\frac{1}{\theta}}\right) \quad (6)$$

Donde  $\tilde{u}_i \equiv -\log u_i$ .

La función de distribución bivariada se expresa como,

$$C_\theta(u, v) = \exp\left\{-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{\frac{1}{\theta}}\right\} \quad (7)$$

Y la función de densidad es,

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

Por último están las cópulas de Valores Extremos, que se derivan de la estructura de dependencia de la distribución generalizada de valor extremo multivariada. En esta familia de cópulas, se da mayor énfasis a los eventos de las colas (raros o extremos) de las distribuciones marginales. Se trata

de encontrar los límites de cópulas asociadas a los valores máximos de las variables aleatorias independientes e idénticamente distribuidas. Es de uso común en la medición de riesgos financieros, por la relevancia que se le da a la estructura de la cópula, los valores extremos de las distribuciones marginales, y la forma como se asocian. En las cópulas de valores extremos también se encuentra la Gumbel de independencia y comonotividad.

### 3. Hipótesis

Teniendo en cuenta que las pérdidas por eventos de riesgo operativo en el sector financiero son variables aleatorias que modelan distribuciones leptocúrticas y sesgadas, se pretende mostrar que: La medición de la pérdida inesperada con influencia de factores de dependencia de los eventos de riesgo operativo en una entidad financiera, se pueden modelar con teoría de cópulas a través de parámetros de dependencia en las colas de la distribución.

### 4. Metodología

El trabajo se enmarca en una investigación empírica, y el propósito es utilizar modelos de medición que se ajusten a las pérdidas por riesgo operativo en las colas de las distribuciones; para determinar las pérdidas inesperadas, es decir, en situaciones de turbulencia, cuánto podría perder la entidad por la materialización de los eventos de riesgo operativo, y así evaluar qué tan necesarias son el cálculo del capital económico de la entidad financiera.

Para la aplicación del modelo de medición al contexto colombiano, es necesario contar con bases de datos de pérdidas detalladas por factores de riesgo operativo, por líneas de negocio y tipos de eventos. La superintendencia financiera de Colombia (Superfinanciera, 2007) en el marco de aplicación del Sistema de Administración de Riesgo Operativo (SARO), establece los criterios que las entidades deben tener en cuenta para hacer el registro de los eventos por riesgo operativo, y algunas han iniciado con la recolección de información desde el 2008 aproximadamente.

Los primeros análisis se realizan para establecer la forma de las distribuciones de pérdidas por riesgo operativo con datos internos de una entidad financiera que facilita la información y convenio de confidencialidad; se analizan las distribuciones de pérdidas para la línea de negocio de mayor participación y dos de sus eventos. Para el procesamiento de la información se utilizan paquetes estadísticos como el SPSS, software de simulación ma-



temática Matlab y herramientas que son complementos de Excel como Risk Simulator.

Una vez identificadas las características de las distribuciones de pérdidas por riesgo operativo, y reconociendo que estas distribuciones presentan alta curtosis y sesgo positivo, se procede a aplicar modelos que permitan determinar las pérdidas no esperadas, es decir, nos centramos en la cola de la distribución.

Para aplicar la teoría de cópulas en el cálculo de pérdidas inesperadas, se hace necesario determinar a partir de qué valor se considera la cola de la distribución, es un umbral mínimo de pérdidas que define una gran pérdida mínima, porque la cola es a la derecha y se consideran como distribuciones asintóticas (Castillo, 1988). Dado que el umbral es fundamental, debido a que es el punto de transición de las pequeñas pérdidas y las pérdidas de gran magnitud se debe tener cuidado para establecer el dato del umbral correctamente; en esta parte se revisan métodos gráficos y la técnica de medias de excesos para encontrar el umbral adecuado y en la medición se usan modelos paramétricos, entre ellos los de bloques máximos, el método POT o picos sobre el umbral de la teoría de valores extremos (Fisher y Tippet, 1928).

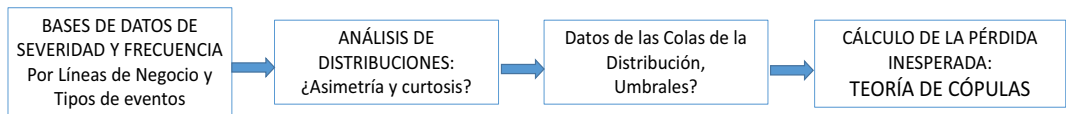
De esta forma, la aplicación de la teoría de cópulas en los datos extremos para medir la pérdida inesperada por eventos de riesgo operativo en el sector financiero incorpora el análisis de la influencia de factores de correlación a través de parámetros de dependencia no lineal en las colas. Es decir, que una vez obtenida la distribución de la cola teniendo en cuenta el método de selección del umbral aplicado en la teoría de valor extremo, se realiza un análisis conjunto para establecer los parámetros de dependencia, y las distribuciones marginales de dos eventos de una línea de negocio y finalmente calcular la pérdida inesperada contemplando la dependencia entre dichos eventos.

La teoría de cópulas ha sido aplicada para la medición del riesgo operativo, entre otras, se cuenta con algunas referencias como, (Gallardo, 2009) afirma que para determinar la forma paramétrica de la cópula se debe conocer la dependencia entre las variables y utiliza para medir la dependencia no lineal el coeficiente de correlación de Spearman y la Tau de Kendall. El documento de (Chavez Gudiño, 2009), utiliza la distribución generalizada de Pareto multivariada de tipo logístico anidado para algunas variables, teniendo en cuenta que es importante para situaciones donde exista dependencia en la ocurrencia de eventos extremos, usa la densidad angular para estimar

los parámetros de dependencia, y propone un método para determinar el orden jerárquico de las variables. También está el trabajo desarrollado por (Fantazzini, Valle, y Giudici, 2008), inicia con una descripción del modelo, luego presenta las distribuciones marginales utilizados para modelar la frecuencia y la severidad de las pérdidas y se revisan los principales puntos de la teoría de la cópula con una muestra empírica, y el cálculo del valor en riesgo (OPVaR).

La metodología propuesta se puede resumir en la Figura 4.

Figura 4. Metodología Propuesta



Fuente: Elaboración propia.

## 5. Resultados y discusión

Los datos de pérdidas por riesgo operativo proceden de una entidad financiera con un nivel de activos que no supera los 900 mil millones de pesos. Las operaciones que realiza se concentran en la línea banca personal y minorista; por tanto, es allí donde se va a observar una mayor concentración de riesgos operativos. Un resumen estadístico de los datos de pérdida para la línea de negocio se presenta en la Tabla 1.

Tabla 1. Datos estadísticos línea de negocio

MIN	Q1	Q2	Q3	MÁX
0%	25%	50%	75%	100%
\$ 10,000	\$ 98,725	\$ 467,801	\$ 3,957,758	\$58,555,329
		\$ 58,545,329		
		\$ 3,859,033		
\$ 88,725		\$ 54,597,571		

Forma Distribución	
Curtosis	13.87
Coefficiente de Asimetría	3.54
Prueba K - S	0.000%

Fuente: Elaboración propia

Los datos estadísticos muestran el intervalo de movimiento de las pérdidas es muy amplio. Al observar la diferencia entre el primer cuartil con el dato menor, y el tercer cuartil con el dato mayor, se evidencia un profundo sesgo a la derecha de la distribución.

Para los dos eventos EAP y FE<sup>2</sup> estudiados, el comportamiento de los datos se presenta en la Tabla 2,

Tabla 2. Datos estadísticos eventos EAP y FE

EVENTO: EAP

MIN	Q1	Q2	Q3	MÁX
0%	25%	50%	75%	100%
\$ 10,000	\$ 29,478	\$ 69,477	\$ 229,902	\$ 57,581,498
		\$ 57,571,498		
		\$ 200,424		
\$ 19,478		\$ 57,351,596		

EVENTO: FE

MIN	Q1	Q2	Q3	MÁX
0%	25%	50%	75%	100%
\$ 95,385	\$ 434,968	\$ 1,920,299	\$ 14,929,146	\$ 58,555,329
		\$ 58,459,944		
		\$ 14,494,178		
\$ 339,583		\$ 43,626,183		

Forma Distribución	EAP	FE
Curtosis	187.78	3.21
Coefficiente de Asimetría	12.02	1.97
Prueba K - S	0.0000%	0.0001%

Fuente: Elaboración propia.

Del análisis estadístico se puede concluir,

- Los diferentes grupos de pérdidas por riesgo operativo presentan distribuciones con alta dispersión, rango amplio y desviación estándar muy superior a la media.
- La forma de la distribución para las pérdidas para la línea de negocio y tipo de evento muestran alta curtosis y asimetría a la derecha; por tanto presentan colas pesadas.
- El exceso de curtosis en las diferentes distribuciones analizadas demuestra que presentan colas pesadas
- Ninguna de las distribuciones analizadas se comporta como la distribución normal.

<sup>2</sup> EAP: Ejecución y Administración de Procesos; FE: Fraude Externo.

En el análisis de los datos se pudo verificar que las distribuciones de probabilidad no son normales, lo que conlleva a buscar métodos más apropiados para realizar su medición. Además, es muy común utilizar el coeficiente de correlación lineal de Pearson para encontrar en qué medida una variable mueve a la otra, pero esto asume normalidad de las distribuciones individuales, cuando en realidad la mayoría de las series financieras no presentan esta característica.

Las cópulas se pueden usar para resolver situaciones de riesgo multivariado, es una técnica más apropiada para analizar la dependencia de los datos extremos de pérdidas por riesgo operativo, sin tener que utilizar la correlación lineal que presenta algunas desventajas. La idea es determinar los datos extremos de los dos eventos, para la línea Banca personal y Minorista y calcular su dependencia, teniendo en cuenta que estos eventos extremos se encuentran en la cola superior.

Por las características de las cópulas, el teorema de (Sklar A., 1959) que las define, justifica el papel de las funciones de dependencia. En esta parte del trabajo, se pretende mostrar cómo las cópulas pueden ser usadas para describir la estructura de dependencia de las colas de los dos eventos en la línea de negocio, y además, determinar el capital necesario. Como se va a trabajar con datos extremos, de las familias de cópulas que existen sólo tendrían utilidad las que permitan distribuciones asimétricas como las cópulas Arquimedianas, por ejemplo Cópula Gumbel y la HRT.<sup>3</sup>

Teniendo en cuenta que en la teoría de valores extremos para determinar los datos de las colas se utilizan dos métodos: Bloques Máximos y Picos sobre el Umbral (POT), estos se aplicaron para obtener las distribuciones de las colas de los dos eventos de estudio para desarrollar el modelo de cópulas.

### Cópulas con datos Extremos de Bloques Máximos:

Como los datos obtenidos con Bloques Máximos se encuentran en periodicidad mensual, no es necesario hablar de la distribución de frecuencia, la metodología para la aplicación de las cópulas en el cálculo de las pérdidas no esperadas por riesgo operativo es la siguiente;

1. Datos de pérdidas extremas obtenidos por el método de Bloques Máximos con periodicidad mensual.

---

<sup>3</sup> Heavy Right Tail. Para esta investigación se utiliza la cópula de Gumbel.

2. Con los datos obtenidos extremos mensuales en el cruce de línea de negocio y tipo de evento se determinan las distribuciones de severidad de mejor ajuste, teniendo en cuenta que son distribuciones continuas.
3. Con la simulación de Montecarlo en Matlab se establece la distribución de pérdidas agregadas  $S_i$  para cada tipo de evento EAP y FE en la línea de negocio.
4. Con los resultados de la simulación de Montecarlo se calculan el OPVaR<sup>4</sup> al 95%, 99% y 99.9% de confianza y se calcula la pérdida esperada  $\mu$  y la no esperada ( $OPVaR - \mu$ ) por tipo de evento.
5. Se calculan las pérdidas totales para los dos eventos de la línea Banca Personal y Minorista asumiendo una dependencia perfecta.
6. Las pérdidas agregadas  $S_i$  para cada tipo de evento corresponden a las distribuciones marginales y con ellas se determina la función multivariante como insumo para la cópula.

Las distribuciones de ajuste de los datos extremos obtenidos por el método de bloques máximos para los dos eventos se encuentran en la Tabla 3.

Tabla 3: Parámetros Distribuciones Bloques Máximos

Tipo de evento/banca personal y minorista	Distribución de ajuste severidad/parámetros	
EAP	Gamma	Alfa: 0.2897 Beta 23 878 534.88
FE	Gamma	Alfa: 0.3494 Beta: 27 668 624.65

Fuente: Elaboración propia. Resultados Simulador de Riesgo

Con los parámetros de las distribuciones de ajuste para la severidad de los valores extremos se realiza una simulación de Montecarlo con el software de Matlab para obtener la distribución de pérdidas, que luego serán

<sup>4</sup> Value at Risk para Riesgo Operativo.

las distribuciones de entrada o marginales de la cópula. Los resultados de la simulación se muestran en las Tablas 4 y 5,

Tabla 4. Pérdidas agregadas evento EAP

<b>LÍNEA: BANCA PERSONAL Y MINORISTA</b>			
<b>Ejecución y Administración de Procesos</b>	<b>95%</b>	<b>99%</b>	<b>99.9%</b>
Pérdida Esperada ( $\mu$ )	\$ 6,810,567	\$ 6,810,567	\$ 6,810,567
VaR TVE BMAX	\$ 31,913,354	\$ 60,677,114	\$ 114,970,303
Pérdida No esperada (VaR - $\mu$ )	\$ 25,102,787	\$ 53,866,547	\$ 108,159,736

Fuente: Elaboración propia. Resultados Matlab

Tabla 5. Pérdidas agregadas evento FE

<b>LÍNEA: BANCA PERSONAL Y MINORISTA</b>			
<b>Fraude Externo</b>	<b>95%</b>	<b>99%</b>	<b>99.9%</b>
Pérdida Esperada ( $\mu$ )	\$ 9,591,973	\$ 9,591,973	\$ 9,591,973
VaR TVE BMAX	\$ 40,767,563	\$ 76,833,250	\$ 150,839,371
Pérdida No esperada (VaR - $\mu$ )	\$ 31,175,590	\$ 67,241,277	\$ 141,247,398

Fuente: Elaboración propia. Resultados Matlab

Ahora, asumiendo una dependencia perfecta lineal entre los dos eventos se suman los eventos en la Tabla 6, para compararlas con los resultados de la cópula que involucra la dependencia conjunta,

Tabla 6. Suma aritmética con dependencia lineal perfecta EAP y FE

<b>Pérdidas No Esperadas</b>	<b>95%</b>	<b>99%</b>	<b>99.9%</b>
Ejecución y Administración de Procesos	\$ 25,102,787	\$ 53,866,547	\$ 108,159,736
Fraude Externo	\$ 31,175,590	\$ 67,241,277	\$ 141,247,398
<b>Pérdida No Esperada TOTAL</b>	<b>\$ 56,278,377</b>	<b>\$ 121,107,824</b>	<b>\$ 249,407,134</b>

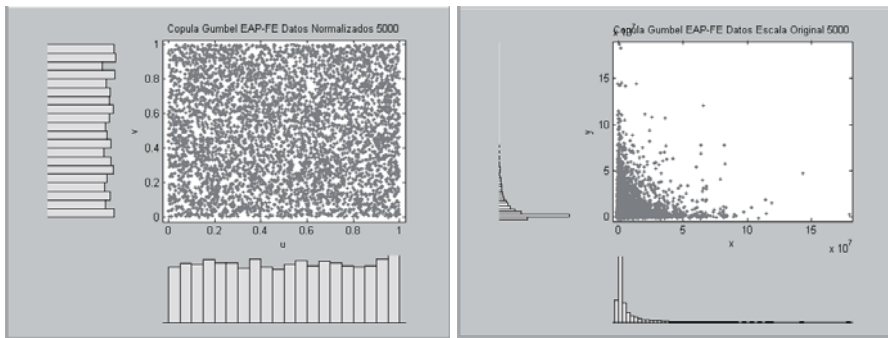
Fuente: Elaboración propia. Resultados Matlab.

Para la simulación con la cópula seleccionada se aplica la siguiente metodología:

1. Las pérdidas agregadas  $S_i$  corresponden a las distribuciones marginales de cada tipo de evento, y son los datos de entrada de la cópula.
2. Se normalizan las distribuciones marginales, es decir, se transforman los datos a la escala cópula con el estimador de Kernel.<sup>5</sup>
3. Se ajustan los datos transformados a la cópula seleccionada: Gumbel.
4. Se determinan los parámetros que identifican la cópula y el coeficiente de dependencia  $\tau$  de Kendall.
5. Se realiza la simulación con los parámetros de la cópula.
6. Y finalmente se transforman nuevamente los datos a su escala original, para este paso se utiliza el inverso del estimador de Kernel.

Los resultados para EAP y FE se muestran en la Figura 5, están los datos normalizados por la simulación de Montecarlo y los datos en escala original procesados en el software de Matlab,

Figura 5. Cópula Gumbel EAP y FE



Fuente: Elaboración Propia.

El parámetro Alpha y el coeficiente de dependencia no lineal la  $\tau$  de Kendall para cópula Gumbel se observa en la Tabla 7.

Además, en la Tabla 8, están los resultados de la simulación de Montecarlo con la cópula seleccionada.

Se observa que las pérdidas no esperadas representan entre el 71% y 89% de las pérdidas globales o VaR Cópula.

<sup>5</sup> Matlab trabaja con el estimador de Kernel gaussiano.

Tabla 7. Parámetros cópula Gumbel

Aspecto	Nombre	Valor
Parámetro Cópula Gumbel	Alpha	1.000001358
Dependencia no Lineal	$\tau$ de Kendall	1.36E-06

Fuente: Elaboración propia. Resultados Matlab

Tabla 8. Pérdidas conjuntas EAP y FE cópula Gumbel

Cópula EAP y FE	95%	99%	99.9%
VaR TVE Gumbel	\$ 59,273,862	\$ 97,169,457	\$ 159,576,470
Pérdida Esperada ( $\mu$ )	\$ 17,129,758	\$ 17,129,758	\$ 17,129,758
Pérdida No Esperada EAP-FE	\$ 42,144,105	\$ 80,039,699	\$ 142,446,712

Fuente: Elaboración propia. Resultados Matlab

A continuación se comparan las pérdidas totales de estos dos eventos de riesgo de la línea Banca Personal y Minorista con dependencia lineal perfecta y con la dependencia que generan cópulas, véase la Tabla 9.

Tabla 9. Comparación OPVaR dependencia lineal perfecta y Cópula

Pérdidas No Esperadas	95%	99%	99.9%
VaR (EAP + FE) Dependencia Perfecta Lineal	\$ 56,278,377	\$ 121,107,824	\$ 249,407,134
VaR (EAP + FE) Cópula Gumbel con Dependencia no Lineal	\$ 42,144,105	\$ 80,039,699	\$ 142,446,712

Fuente: Elaboración propia. Resultados Matlab

Se evidencia una disminución importante de las pérdidas con la cópula de Gumbel, lo que sería favorable para la entidad al momento de determinar el capital necesario para cubrir las pérdidas por riesgo operativo.

### Cópulas con datos extremos POT:

En esta parte el método de selección de datos extremos utilizado fue Picos sobre el Umbral (POT) aplicado en la Teoría de Valores Extremos (TVE), sin



embargo, no se trabaja con las diferencias entre el dato extremo y el umbral, sino con los valores de pérdidas que excedan el umbral. Se procesan los datos para obtener la distribución severidad y frecuencia que más se ajusta a cada tipo de evento.

La metodología utilizada para estos datos obtenidos de POT en el cálculo de las pérdidas no esperadas por riesgo operativo es la siguiente:

1. Datos de eventos extremos según la metodología POT en el cruce de línea de negocio y tipo de evento
2. Se determinan las distribuciones de frecuencia y severidad para los dos eventos.
3. Se buscan las distribuciones de mejor ajuste para severidad y frecuencia, teniendo en cuenta que la severidad es una distribución continua y la frecuencia es discreta.
4. Con la simulación de Montecarlo se determina la distribución de pérdidas agregadas para cada tipo de evento en la línea de negocio.
5. Con los resultados de la simulación de Montecarlo se calculan las pérdidas OPVaR al 95%, 99% y 99.9% de confianza y se calcula la pérdida esperada  $\mu$  y la no esperada por tipo de evento y la total como la suma.
6. Las pérdidas agregadas para cada tipo de evento son distribuciones que corresponden a las distribuciones marginales y con ellas se determina la función multivariante.
7. Se realiza la simulación de Montecarlo con la cópula seleccionada y se calculan las pérdidas.

Las distribuciones de frecuencia y severidad para los dos tipos de eventos, Tabla 10:

Con los parámetros de las distribuciones de ajuste para severidad y frecuencia se realiza una simulación de Montecarlo de tal forma que combine estas dos distribuciones en una sola llamada distribución agregada de pérdidas. Este procedimiento es posible a través de una convolución con los siguientes supuestos (Frachot, Roncalli, y Salomon, 2004).

- Los valores de frecuencia y severidad de las pérdidas por riesgo operativo son variables aleatorias independientes.

Tabla 10. Parámetros Frecuencia y Severidad POT

Tipo de evento/banca Personal y minorista	Distribución de ajuste severidad/parámetros		Distribución de ajuste frecuencia/parámetros	
EAP	Gamma	Alfa: 1.10514 Beta: 13835398.79	Poisson	$\lambda = 0.518283$
FE	Gamma	Alfa: 2.595570 Beta: 11302829.69	Poisson	$\lambda = 0.291077$

Fuente: Elaboración propia. Resultados Simulador de Riesgo

- Los valores de severidad de las pérdidas por riesgo operativo para un mismo evento de riesgo se distribuyen idénticamente.
- Los valores de severidad de las pérdidas por riesgo operativo para un mismo evento de riesgo son independientes.

Se realiza convolución con simulación de Montecarlo utilizando programación en Matlab, considerando la distribución de frecuencia y severidad para cada evento de riesgo. En esta simulación se utiliza el siguiente algoritmo:

- Se genera un valor  $n$  de la distribución de frecuencia.
- Con base en el valor de  $n$ , se generan  $n$  valores aleatorios de la distribución de severidad  $L_1, L_2, L_3, \dots, L_n$ .
- Se suman los valores de severidad y se forma la variable  $S_i$ , donde  $S_i = L_1, L_2, L_3, \dots, L_n$ .
- Los tres pasos anteriores se repiten un número de veces grande (5000, 10,000,...).
- Con los  $M$  resultados de  $S_i$ , se obtiene una distribución de pérdidas agregadas genera el OPVaR, y finalmente las pérdidas esperadas y no esperadas por riesgo operativo al 95%, 99% o 99.9% de confianza.
- Además, también se puede obtener con los  $M$  resultados de  $S_i$ , la distribución que mejor se ajusta a la distribución de pérdidas agregadas.

Los resultados de las pérdidas agregadas para los dos eventos se presentan en las Tablas 11 y 12. Los resultados son.

Tabla 11. Resultados OPVaR EAP

<b>LÍNEA: BANCA PERSONAL Y MINORISTA</b>			
<b>Ejecución y Administración de Procesos</b>	<b>95%</b>	<b>99%</b>	<b>99.9%</b>
Pérdida Esperada ( $\mu$ )	\$ 7,585,044	\$ 7,585,044	\$ 7,585,044
VaR TVE POT	\$ 38,160,969	\$ 65,940,287	\$ 95,285,483
Pérdida No esperada (VaR - $\mu$ )	\$ 30,575,925	\$ 58,355,242	\$ 87,700,439

Fuente: Elaboración propia. Resultados Matlab

Tabla 12. Resultados OPVaR FE

<b>LÍNEA: BANCA PERSONAL Y MINORISTA</b>			
<b>Fraude Externo</b>	<b>95%</b>	<b>99%</b>	<b>99.9%</b>
Pérdida Esperada ( $\mu$ )	\$ 8,446,322	\$ 8,446,322	\$ 8,446,322
VaR TVE POT	\$ 48,369,809	\$ 84,351,146	\$ 128,361,368
Pérdida No esperada (VaR - $\mu$ )	\$ 39,923,487	\$ 75,904,824	\$ 119,915,046

Fuente: Elaboración propia. Resultados Matlab

Ahora, asumiendo una dependencia perfecta lineal entre los dos eventos para compararlas con los resultados de la cópula que involucra la dependencia conjunta, las pérdidas se muestran en la Tabla 13.

Tabla 13. Pérdidas totales con dependencia lineal perfecta

<b>Pérdidas No Esperadas</b>	<b>95%</b>	<b>99%</b>	<b>99.9%</b>
Ejecución y Administración de Procesos	\$ 30,575,925	\$ 58,355,242	\$ 87,700,439
Fraude Externo	\$ 39,923,487	\$ 75,904,824	\$ 119,915,046
<b>Pérdida No Esperada TOTAL</b>	<b>\$ 70,499,411</b>	<b>\$ 134,260,066</b>	<b>\$ 207,615,485</b>

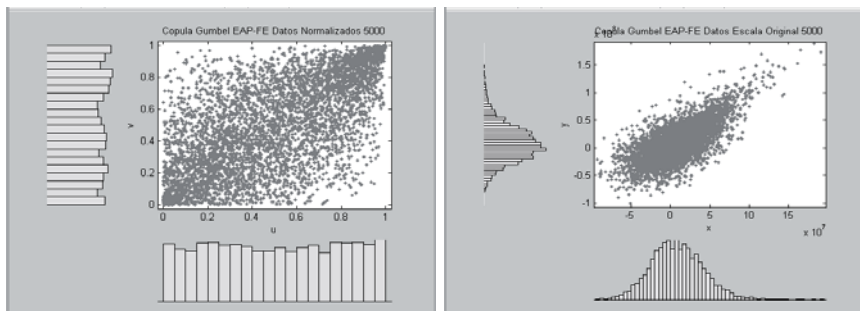
Fuente: Elaboración propia.

Para la simulación con la cópula seleccionada se aplica la siguiente metodología.

1. Las pérdidas agregadas  $S_i$  corresponden a las distribuciones marginales de cada tipo de evento, y son los datos de entrada de la cópula.
2. Se normalizan las distribuciones marginales con el estimador de Kernell.
3. Se ajustan los datos transformados a la cópula seleccionada: Gumbel.
4. Se determinan los parámetros que identifican la cópula y el coeficiente de dependencia  $\tau$  de Kendall.
5. Se realiza la simulación con los parámetros de la cópula.
6. Y finalmente se transforman nuevamente los datos a su escala original, para este paso se utiliza el inverso del estimador de Kernell.

Los resultados para Ejecución y Administración de Procesos y Fraude Externo, los datos normalizados por la simulación de Montecarlo y los datos en escala original con los eventos, se muestran en los siguientes gráficos que se realizaron en el software de Matlab, Figura 6.

Figura 6. Cópula Gumbel eventos EAP y FE



Fuente: Elaboración Propia.

En la Tabla 14 se observa parámetro Alpha y el coeficiente de dependencia no lineal  $\tau$  de Kendall,

Tabla 14. Parámetros cópula Gumbel

Aspecto	Nombre	Valor
Parámetro Cópula Gumbel	Alpha	1.898018
Dependencia no Lineal	$\tau$ de Kendall	0.473135

Fuente: Elaboración propia. Resultados Matlab

Los resultados de la simulación de Montecarlo con la cópula seleccionada aparecen en la Tabla 15,

Tabla 15. Resultados OPVaR cópula Gumbel EAP y FE

Cópula EAP y FE	95%	99%	99.9%
VaR TVE Gumbel	\$ 119,797,488	\$ 185,276,551	\$ 251,173,657
Pérdida Esperada ( $\mu$ )	\$ 16,874,277	\$ 16,874,277	\$ 16,874,277
Pérdida No Esperada EAP-FE	\$ 102,923,211	\$ 168,402,274	\$ 234,299,381

Fuente: Elaboración propia. Resultados Matlab.

Se observa que las pérdidas no esperadas representan entre el 86% y 93% de las pérdidas globales o VaR Cópula.

Ahora, se comparan las pérdidas totales de estos dos eventos de riesgo de la línea de negocio con dependencia lineal perfecta y con la dependencia que genera la cópula, Tabla 16.

Tabla 16. Comparación pérdidas dependencia lineal perfecta y cópula Gumbel

Pérdidas No Esperadas	95%	99%	99.9%
VaR (EAP + FE) Dependencia Perfecta Lineal	\$ 70,499,411	\$ 134,260,066	\$ 207,615,485
VaR (EAP + FE) Cópula Gumbel con Dependencia no Lineal	\$ 102,923,211	\$ 168,402,274	\$ 234,299,381

Fuente: Elaboración propia. Resultados Matlab

En estos resultados no se evidencia una disminución de las pérdidas con la cópula de Gumbel.

Teniendo en cuenta que los modelos aplicados tienen como propósito el cálculo de las pérdidas inesperadas para los dos eventos analizados, se realiza la comparación con las dos técnicas de selección de datos extremos y para los diferentes niveles de confianza, como se muestra en la Tabla 17 y la Figura 7.

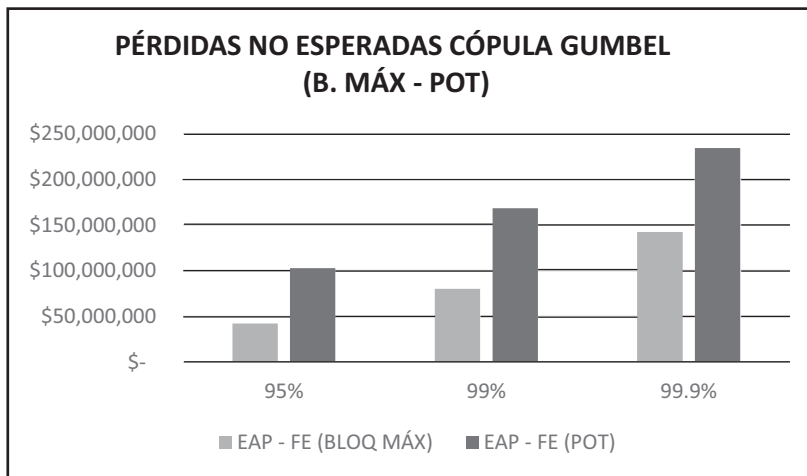
La teoría de cópulas es una metodología que se diferencia de las otras al incluir en el cálculo de las pérdidas la dependencia no lineal que existe entre los eventos de riesgo operativo. Se pudo determinar que al utilizar la cópula con datos de bloques máximos se reducen las pérdidas inesperadas hasta en un 33%.

Tabla 17. Comparación resultados pérdidas inesperadas con datos extremos Bloques Máximos y POT

	95%	99%	99.9%
EAP - FE (BLOQ MÁX)	\$ 42,144,105	\$ 80,039,699	\$ 142,446,712
EAP - FE (POT)	\$ 102,923,211	\$ 168,402,274	\$ 234,299,381

Fuente: Elaboración propia

Figura 7: Comparación resultados pérdidas inesperadas con datos extremos Bloq Max y POT



Fuente: Elaboración propia

## Conclusiones

La aplicación de la teoría de cópulas muestra que al calcular las pérdidas inesperadas por riesgo operativo con dependencia no lineal, en algunos eventos se está incumpliendo la propiedad de subaditividad, y esto se puede presentar debido a que la modelación se realiza con datos extremos, es decir, los que se encuentran en las colas de las distribuciones.

Por tanto, si para el cálculo de los requerimientos de capital se suman las pérdidas individuales de los eventos analizados, esto supone una dependencia perfecta positiva que conlleva a una posible subestimación de

las pérdidas para asignar el capital económico. Se infiere entonces, que al incluir la dependencia entre los eventos, no necesariamente se debe esperar que las pérdidas inesperadas por riesgo operativo sean inferiores a las obtenidas con el supuesto de dependencia lineal perfecta positiva.

Se evidencia en los modelos aplicados que el cambio en las pérdidas inesperadas se hace más grande a medida que se aumenta el nivel de confianza, es decir, un comportamiento exponencial por la característica de asimetría positiva y distribución asintótica que presentan los eventos por riesgo operativo, especialmente cuando se trabaja con datos extremos.

Los dos métodos utilizados para obtener los eventos de las colas, pueden ser válidos, lo que puede hacer uno mejor que el otro lo determinan las pruebas de Backtesting, que permite comparar lo estimado, versus las pérdidas reales.

## Referencias bibliográficas

- Arias, G. L. (2010). *Repository.eafit*. Obtenido de <https://repository.eafit.edu.co/handle/10784/136>.
- Bernal, S. (2013). *Repository.eia*. Obtenido de Modelo cuantitativo para la gestión de riesgos operativos: [repository.eia.edu.co/bitstream/11190/767/1/AD-MO0872.pdf](https://repository.eia.edu.co/bitstream/11190/767/1/AD-MO0872.pdf).
- Böcker, K. y Klüppelberg, C. (2005). *Cutting edge*. Obtenido de WWW.RISK.NET
- Castillo, E. (1988). *Estadística de Valores Extremos. Distribuciones Asintóticas*. España, Estadística Española.
- Chernoval, A., Svetlozar T. R. y Fabozzi, F. (2007). *Operational Risk. A Guide to Basel II Capital Requirements, Models, an Analysis*. New Jersey, Wiley.
- Comité de Basilea. (2004). *Convergencia internacional de medidas y normas de capital. Banco de Pagos Internacionales*. Suiza, Basilea.
- Embrechts, P., Kluppelberg, C. y Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Berlin, Springer.
- Fisher, R. y Tippett, L. (1928). Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample. *Proceedings of the Cambridge Philosophical Society*, pp. 180-190.
- Frachot, A., Georges, P. y Roncalli, T. (2001). *Social Science Research Network*. Obtenido de [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1032523](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1032523).

- Frachot, A., Roncalli, T. y Salomon, E. (2004). Obtenido de [http://www.maths-fi.com/malliavin/The\\_Correlation\\_Problem\\_in\\_Operational\\_Risk\\_01\\_23\\_2004.pdf](http://www.maths-fi.com/malliavin/The_Correlation_Problem_in_Operational_Risk_01_23_2004.pdf).
- Franco, L. C. y Murillo, J. G. (2008). Loss distribution approach (LDA): Metodología actuarial aplicada al riesgo operacional. *Revista Ingenierías*. Universidad de Medellín, pp. 143-156.
- Fréchet, M. (1927). "Sur la loi de probabilité de l'écart maximum". *Annales de la Société Polonaise de Mathématique*, Cracovie 6, pp. 93-116.
- Gallardo, B. (2009). *Teoría de cópulas y aplicaciones en simulación de riesgos financieros e ingeniería civil*. Granada, Universidad de Granada.
- Giacomini, E. (2005). *Risk management with copulae*. Berlin, Humboldt, Center for Applied Statistics and Econometrics, Universität zu Berlin, pp. 8-46.
- Jöhnemark, A. (2012). *Modeling Operational Risk*. Stockholm.
- Melo, L., y Becerra, O. (2008). "Medidas de riesgo financiero usando cópulas: teoría y aplicaciones". *Borradores de Economía. Banco de la República*. Núm. 489, p. 15.
- Mora, A. a. (2010). "Cuantificación del riesgo operativo en entidades financieras en Colombia". *Cuadernos de Administración*, pp. 185-212.
- Mora, A. b. (2010). "Un estudio comparativo de algunos estimadores del índice de cola  $\xi$ ". *Revista Innovar Journal*. Universidad Nacional de Colombia, vol. 21, núm, 40, p. 19.
- Plascencia Cuevas, T. N. (2010). *Valoración del riesgo utilizando cópulas como medida de la dependencia: Aplicación al sector financiero mexicano (2002-2008)*. México.
- Salinas, Sandra Milena. Maldonado, Diana A. y Díaz, Luis Guillermo. (2010). "Estimación del riesgo en un portafolio de activos". (s.f.). *Apuntes del CENES*. Núm. 50, vol. XXIX.
- Shevchenko, P. y Peters, G. (2013). *Loss Distribution Approach for Operational Risk Capital Modelling under Basel II: Combining Different Data Sources for Risk Estimation*.
- Sklar, A. (1959). "Fonctions de Repartition à n-dimensions et Leur Marges". *Inst. Statist.Univ. Paris*, pp. 229-231.
- Superfinanciera (2007). "Reglas relativas a la Administración del Riesgo Operativo". *Circular 041*, Bogotá.