

# Proyecto I: Métodos Monte Carlo Aplicados

## Cuantificación en Riesgo Operacional

Yael Fabián Olivares Cruz

December 3, 2024

En toda actividad de cualquier tipo existe una exposición inherente al riesgo, por lo que es necesario establecer estrategias para hacer frente a los posibles escenarios. Este proyecto busca abordar esta problemática mediante herramientas avanzadas de simulación para anticipar y gestionar eventos de alto impacto financiero.

El objetivo de este proyecto es estimar las pérdidas potenciales asociadas a eventos con impacto financiero, utilizando técnicas cuantitativas de modelado de riesgos. Se implementará el método de Monte Carlo, incorporando procesos de Poisson compuestos no homogéneos, cópulas para modelar dependencias y medidas de riesgo como la pérdida esperada, la volatilidad el VaR y el CVaR.

La simulación Monte Carlo es particularmente adecuada debido a:

- La aleatoriedad inherente a los eventos que generan pérdidas financieras.
- Su capacidad para modelar múltiples escenarios y condiciones complejas.
- La posibilidad de capturar dependencias entre diferentes fuentes de riesgo mediante cópulas.

Este enfoque asegura una evaluación cuantitativa de los eventos de riesgo, proporcionando bases sólidas para la toma de decisiones.

# Planteamiento del Problema

En la gestión de riesgo operacional existe un conjunto de estrategias conocidas como EMTA para hacer frente a las situaciones adversas, estas son

- 1 Evitar
- 2 Mitigar
- 3 Transferir
- 4 Aceptar

# Planteamiento del problema

Cuando se decide mitigar el riesgo, es necesaria la creación de un entorno de control, en el cual se definen herramientas para reducir el riesgo residual, algunas de estas son

- 1 Controles
- 2 **Gestión de eventos**
- 3 Indicadores
- 4 Action Drivers
- 5 Recomendaciones de auditoría

# Planteamiento del problema

Los riesgos son gestionados por diferentes áreas de, por ejemplo, riesgo tecnológico, riesgo cyber, riesgo de negocio, etc.

Dentro de esas áreas existen diferentes clasificaciones de riesgo, algunas de ellas son

- 1 Inadecuada Disponibilidad
- 2 Inadecuada gestión del ciclo de vida de activos
- 3 Inadecuada gestión del cambio

Para cada tipo de riesgo podemos hacer un análisis del comportamiento de eventos y posteriormente estudiar su dependencia.

Utilizando Python 3 se hicieron simulaciones de severidad por eventos con dos distribuciones utilizadas comunmente en este contexto, estas son

## 1 Distribución Lognormal

$$f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\ln(x) - \mu)^2}{2\sigma^2}\right\} \mathbb{I}_{(0,\infty)}(x)$$

$$\mu \in \mathbb{R}, \quad \sigma > 0$$

## 2 Distribución PERT

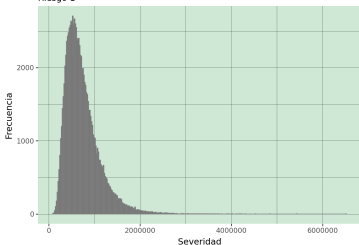
$$f_X(x) = \frac{(x-a)^{\alpha-1}(c-x)^{\beta-1}}{B(\alpha, \beta)(c-a)^{\alpha+\beta-1}},$$

$$\alpha = 1 + \lambda \frac{b-a}{c-a}$$

$$\beta = 1 + \lambda \frac{c-b}{c-a}$$



Distribución de severidad distribución Lognormal  
Riesgo 1

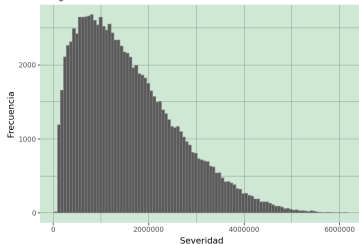


Distribución Lognormal

$$\mu = 650000$$

$$\sigma = 1 - 62$$

Distribución de severidad distribución PERT  
Riesgo 2



Distribución PERT

$$\alpha = 65000, \beta = 500000$$

$$c = 10000000, \lambda = 4$$

Para los casos donde no tenemos una función de densidad conocida para simular las pérdidas, podemos utilizar el método de la transformada inversa utilizando los datos empíricos para simular la severidad de algún riesgo.

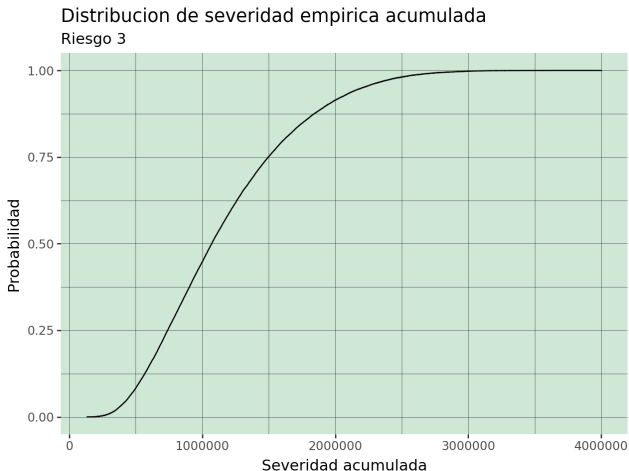
Para ello se utilizó el siguiente algoritmo

1. Obtenemos el número de eventos ( $n$ )
2. Ordenamos los eventos por severidad
3. La función de distribución acumulada (CDF) valdrá  $\frac{i}{n}$ , donde  $i$  es el  $i$ -ésimo evento
4. Generar un número aleatorio  $U \sim U(0, 1)$
5. Tomar  $X = F^{-1}U$

Para realizar las simulaciones Se utilizó un DataFrame llamado Tickets donde se tiene un registro de eventos, posteriormente se añadió una columna con la severidad simulada.

	Incident_ID	Impact	Category	Open_Time	Severidad
0	IM0000004	4	incident	2012-02-05 13:32:00	2.036414e+06
1	IM0000005	3	incident	2012-03-12 15:44:00	1.087896e+06
3	IM0000011	4	incident	2012-07-17 11:49:00	3.459975e+05
4	IM0000012	4	incident	2012-08-10 11:01:00	1.198250e+06
5	IM0000013	4	incident	2012-08-10 11:27:00	6.721282e+05

## Gráfico de la función de distribución empírica



Para simular la frecuencia de eventos se utilizó un proceso Poisson compuesto no homogéneo.

Las trayectorias se construyeron de la siguiente manera.

- 1 Construcción de  $\lambda(t)$
- 2 Calculo de las trayectorias del proceso
- 3 Iteraciones de los resultados

Para obtener la función de intensidad  $\lambda(t)$  se promediaron el número de eventos por mes, obteniendo la siguiente función.

$$\lambda(t) = \begin{cases} 3359.50 & \text{si } t \leq 1 \\ 2069.66 & \text{si } 1 < t \leq 2 \\ 1788.33 & \text{si } 2 < t \leq 3 \\ 16.00 & \text{si } 3 < t \leq 4 \\ 23.00 & \text{si } 4 < t \leq 5 \\ 34.00 & \text{si } 5 < t \leq 6 \\ 34.00 & \text{si } 6 < t \leq 7 \\ 42.50 & \text{si } 7 < t \leq 8 \\ 323.50 & \text{si } 8 < t \leq 9 \\ 3389.00 & \text{si } 9 < t \leq 10 \\ 3232.50 & \text{si } 10 < t \leq 11 \\ 2669.50 & \text{si } 11 < t \leq 12 \end{cases}$$

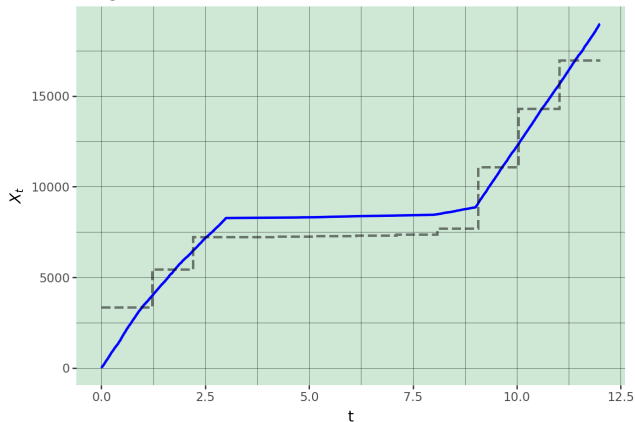
Las trayectorias del proceso se obtuvieron a través del siguiente algoritmo.

- 1 Genera un proceso poisson Homogéneo  $Xh_t$  con parámetro  $\lambda$  que acote las intensidades del proceso Poisson no Homogéneo
- 2 Para cada observación  $x$  del proceso  $Xh_t$  genera una variable aleatoria  $u \sim U(0, 1)$
- 3 Si  $u \leq \frac{\lambda(t)}{\lambda}$  acepta  $x$ , si no, rechaza

## Trayectoria del proceso Poisson no homogéneo

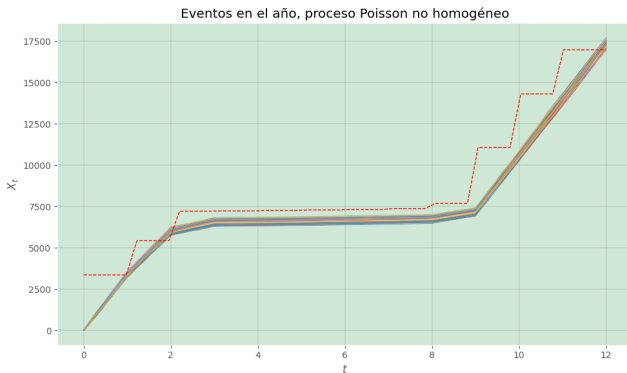
Simulación del proceso poisson no homogeneo  $\{X_t\}_{t \geq 0}$

Riesgo 3





Posteriormente se hicieron 1000 iteraciones del algoritmo para obtener diferentes escenarios del proceso.



## Histograma del total de eventos en el año



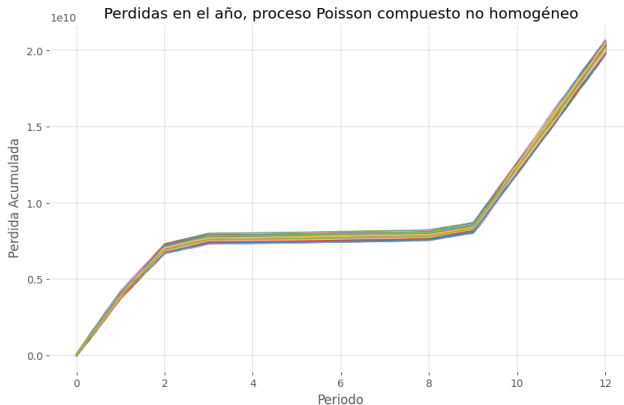
Una vez obtenidas las trayectorias del proceso poisson no homogéneo, se añadieron las severidades de eventos a la frecuencia, utilizando un proceso Poisson compuesto no homogéneo

$$X_t = \sum_{n=0}^{N_t} Y_n$$

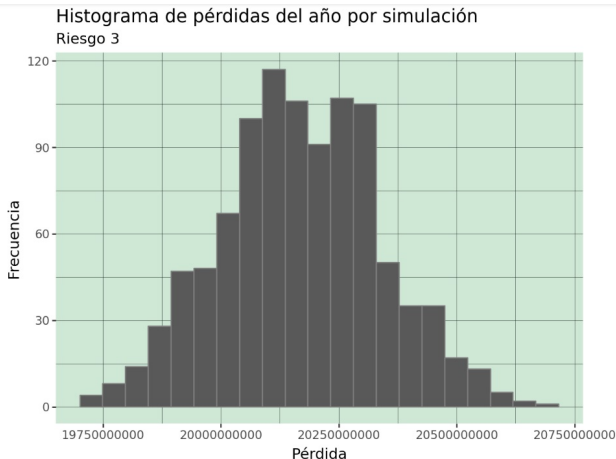
donde

- 1  $N_t$  es un proceso Poisson no homogéneo con función de intensidad  $\Lambda(t)$
- 2  $Y_n$  es una variable aleatoria independiente de  $N_t$

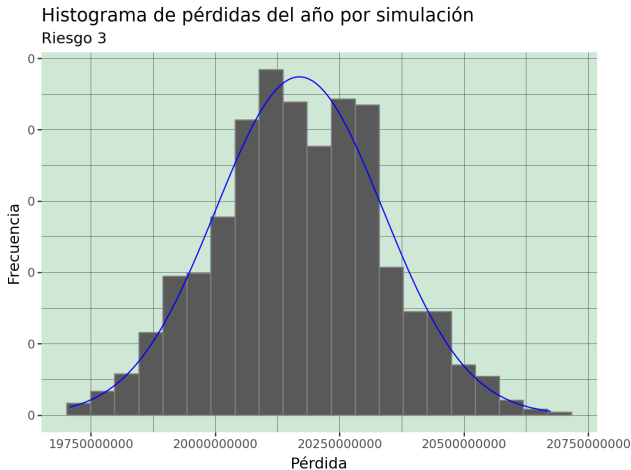
## Trayectorias del proceso poisson compuesto no homogéneo



## Histograma del total de pérdidas en el año

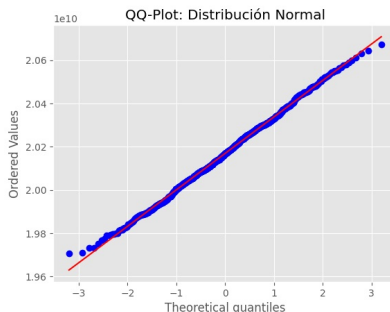


Al visualizar el histograma de pérdidas se pudo observar que aparentemente se sigue una distribución normal, por lo que se hicieron pruebas de normalidad.



Se hizo la prueba Shapiro Willks, así como un QQ-plot para evaluar la normalidad de la distribución de pérdidas totales a final del año.

```
stat, pvalue = stats.shapiro(perdidaFinal['perdidaFinal'])  
pvalue = 0.34962107166858747
```



Para el estudio de valores extremos se definió un umbral de 20500000000. Por el teorema de Fisher-Tippett, la distribución de las colas converge a una de tres distribuciones.

- Distribución Fréchet

$$F(x; \mu, \lambda, \alpha) = \exp \left( -\frac{x - \mu}{\theta} \right)^{-\alpha}, \quad x \geq \mu, \quad \alpha, \theta > 0$$

- Distribución Gumbel

$$F(t; \mu, \lambda) = \exp \left( -\exp \left( -\frac{t - \mu}{\lambda} \right) \right)$$

- Distribución Weibull

$$F(x; \mu, \lambda, \alpha) = \exp \left( -\left( -\frac{x - \mu}{\theta} \right)^{-\alpha} \right), \quad x \leq \mu, \quad \alpha, \theta > 0$$



Como la distribución Gumbel no está acotada, fué el primer candidato para hacer el ajuste de los valores extremos.

Tomamos la siguiente transformación

$$\begin{aligned} z &= \ln(-\ln(F(t))) \\ &= -\frac{t - \mu}{\lambda} \\ &= \frac{\mu}{\lambda} - \frac{1}{\lambda}t \end{aligned}$$

Tomamos el modelo  $y = \beta_0 + \beta_1 x$

Donde  $y = \ln(-\ln(F(t)))$

Entonces tenemos que

$$\beta_0 = \frac{\mu}{\lambda}$$

$$\beta_1 = -\frac{1}{\lambda}$$

$$x = t$$

Por lo que los parámetros serán

$$\lambda = -\frac{1}{\beta_1}$$

$$\mu = -\frac{\beta_0}{\beta_1}$$

A continuación se hizo una regresión lineal con los valores superiores al umbral establecido.

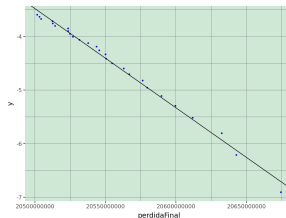
OLS Regression Results

Dep. Variable:	y	R-squared:	0.993
Model:	OLS	Adj. R-squared:	0.992
Method:	Least Squares	F-statistic:	3411.
Date:	Sun, 01 Dec 2024	Prob (F-statistic):	2.98e-28
Time:	03:43:29	Log-Likelihood:	33.209
No. Observations:	27	AIC:	-62.42
Df Residuals:	25	BIC:	-59.83
Df Model:	1		

Covariance Type: nonrobust

	coef	std err	t	P> t	[0.025	0.975]
const	375.9153	6.514	57.710	0.000	362.500	389.331
perdidaFinal	-1.851e-08	3.17e-10	-58.402	0.000	-1.92e-08	-1.79e-08

Omnibus: 3.967 Durbin-Watson: 0.929  
Prob(Omnibus): 0.138 Jarque-Bera (JB): 2.499  
Skew: -0.711 Prob(JB): 0.287  
Kurtosis: 3.449 Cond. No. 9.47e+12



Resultados de la regresión lineal

Gráfico del ajuste lineal

Se simuló otro proceso de eventos con diferentes parámetros y para simular la dependencia entre riesgos se utilizaron cópulas.

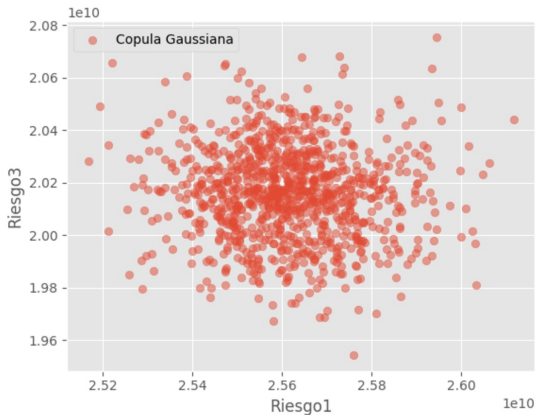
Una cópula  $C$  es una función de distribución conjunta de  $d$  variables aleatorias uniformes en el intervalo  $(0, 1)$ . Esto es

$$C : [0, 1]^d \longrightarrow [0, 1]$$
$$C(U_1, \dots, U_d) := \mathbf{P}[U_1 \leq u_1, \dots, U_n \leq u_n]$$

Inicialmente se utilizó una cópula gaussiana ya que ambos riesgos tienen una distribución de pérdidas totales normal

$$C(u_1, u_2) = \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right\} dy dx$$

## Muestras simuladas de la cópula Gaussiana

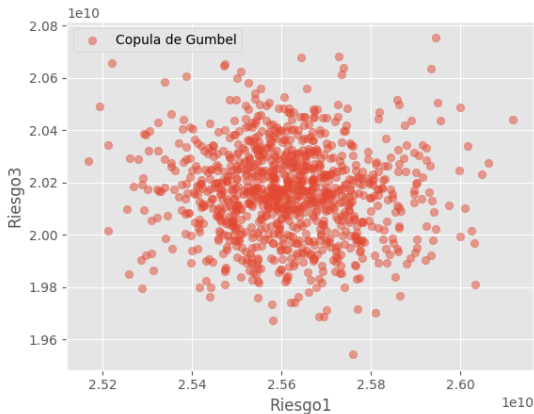


No se observa mucha dependencia entre ambos riesgos

Para estudiar los valores extremos se utilizó una cópula de Gumbel ya que en ambos riesgos se tuvo un buen ajuste a una distribución Gumbel para las colas.

$$C(u_1, \dots, u_d) = (u_1^{-\theta} + u_d^{-\theta} - d + 1)^{1/\theta}$$

## Muestras simuladas de la cópula de Gumbel



No se observa mucha dependencia entre ambos riesgos



# Resultados

Recordemos que utilizamos un proceso Poisson compuesto no homogéneo  $X_t$  para estudiar el comportamiento general de las pérdidas.

Las primeras medidas de riesgo que podemos obtener son la pérdida esperada  $\mathbf{E}(X_t)$  y la volatilidad  $Std(X_t)$

Entonces

1.  $\mathbf{E}(X_t) = \Lambda(t)\mathbf{E}(Y); \quad \Lambda(t) = \int_0^t \lambda(x)dx$

2.  $Var(X_t) = \Lambda(t)\mathbf{E}(Y^2)$

Utilizaremos como estimadores

$$\widehat{\mathbf{E}(Y)} = \bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i, \quad \widehat{Var(Y)} = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2$$

Notemos que  $\Lambda(12) = \int_0^{12} \lambda(x) dx = 16981.5$

Por lo que  $\widehat{\mathbf{E}}(Y_{12}) = 16981.5 * 1081694.923262261 = 18368802339.3780$   
Y además

$$\widehat{Var}(X_t) = 233962639949.6137 * \Lambda(t) + 1081694.9232^2 * \Lambda(t)$$

$$\begin{aligned}\widehat{Var}(X_{12}) &= 233962639949.6137 * 16981.5 + 1081694.9232^2 * 16981.5 \\ &= 1.986945860571556e + 16\end{aligned}$$

$$\widehat{Std}(X_t) = 140959067.1284$$

El **Valor en Riesgo (VaR)** de  $X$  al nivel de confianza  $100\%p$ , denotado  $VaR_p(X)$  ó  $x_p$  es el percentil  $100_p$  de la distribución de  $X$ .

Para variables continuas simplemente podemos buscar el valor  $x_p$  que satisface

$$\mathbf{P}(X > x_p) = p$$

```
var = np.quantile(perdidaFinal['perdidaFinal'], 0.95)
```

```
Out : 20450878141.960884
```

El **Valor de cola en Riesgo (TVaR)** de  $X$  al nivel de confianza  $100\%p$ , denotado por  $TVaR_p(X)$ , es la pérdida esperada dado que la pérdida excede el cuantil  $100_p$  de la distribución de  $X$ . Consideraremos únicamente distribuciones continuas.

$$TVaR_p(X) = \mathbf{E}(X|X > x_p)$$

Dado que la cola derecha se distribuye como una gumbel de parámetros

$$\lambda = 54033185.48415679 \text{ y } \mu = 20311902419.13288$$

entonces

$$TVaR_p(X) = \mathbf{E}(X|X > x_p) = \mu + \lambda\gamma = 20343091219.95402$$

Para medir la dependencia entre dos riesgos se utilizó la tau de Kendal, la cual se define como

$$\tau_k(X, Y) = \mathbf{P}[(X_1 - Y_1)(X_2 - Y_2) > 0] - \mathbf{P}[(X_1 - Y_1)(X_2 - Y_2) < 0]$$

Si la cópula es absolutamente continua puede reescribirse como

$$\tau_k(X, Y) = 4 \int_0^1 \int_0^1 C(u, v) \frac{\partial^2 C(u, v)}{\partial u \partial v} du dv - 1$$

La tau de Kendal de la cópula Gaussiana se obtuvo de la siguiente forma

$$\tau_k(X, Y) = \frac{2}{\pi} \arcsen(\rho)$$

Donde  $\rho$  es el coeficiente de correlación de Spearman.

De este modo

$$\begin{aligned}\tau_k(X, Y) &= \frac{2}{\pi} \arcsen(-0.034810178767449115) \\ &= -0.034810178767449115\end{aligned}$$

La tau de Kendal de la cópula de Gumbel se obtuvo de la siguiente forma

$$\tau_k(X, Y) = 1 - \frac{1}{\theta}$$

Donde  $\theta$  es el parámetro de la cópula de Gumbel.

De este modo

$$\tau_k(X, Y) = 0.4065934065934066$$

# Resultados

La pérdida esperada es 18,368,802,339.3780

La volatilidad es 140,959,067.1284

El VaR es 20,450,878,141.960884

El TVaR es 20,343,091,219.95402

La tau de Kendal para el comportamiento general es  
−0.034810178767449115




La tau de Kendal para las colas es 0.4065934065934066






Las simulaciones realizadas utilizando procesos de Poisson compuestos no homogéneos lograron modelar tanto la severidad como la frecuencia de eventos de riesgo. Esto permitió estimar métricas de riesgo como la pérdida esperada, la volatilidad el VaR, el TVaR y la tau de Kendal.

Las pruebas de normalidad y el análisis de valores extremos confirmaron que las distribuciones utilizadas son adecuadas para modelar las pérdidas. En particular, las distribuciones ajustadas para eventos extremos son esenciales para planificar escenarios de alto impacto financiero.

El uso de cópulas permitió estudiar la dependencia entre diferentes tipos de riesgos. Aunque algunas dependencias fueron bajas, los modelos lograron capturar las interacciones relevantes, ofreciendo un enfoque más completo para la gestión de riesgos.

-  Panjer, H. H. (2006). *Operational Risk: Modeling Analytics*. John Wiley & Sons, Inc.
-  Rodríguez Bermúdez, M. A. (2016). *Modelación Cuantitativa de Riesgo Operativo* (Tesis de maestría, Universidad Nacional de Colombia, Facultad de Ingeniería, Departamento de Ingeniería Industrial). Universidad Nacional de Colombia.
-  Baltazar-Larios, F., & López Ortega, S. I. (2024). *Simulación estocástica*. Facultad de Ciencias, Universidad Nacional Autónoma de México.

-  Erdely, A. (2009). Cópulas y dependencia de variables aleatorias: Una introducción. *Miscelánea Matemática*, 48, 7-28.
-  Lindskog, F., & McNeil, A. J. (2003). Common Poisson shock models: Applications to insurance and credit risk modelling. *ASTIN Bulletin*, Vol. 33 No. 2, pp. 209-238.
-  Rincón, L. (2012). Introducción a los procesos estocásticos (1a ed.). UNAM Facultad de Ciencias, Las Prensas de Ciencias.