

Análisis de Riesgo de una Compañía Aseguradora

Juan Pablo Echeagaray González, Verónica Victoria García De la Fuente,

Emily Rebeca Méndez Cruz, Eugenio Santiesteban Zolezzi,

Daniel De Zamacona Madero

Optimización Estocástica

MA2004B.101

Dr. Fernando Elizalde Ramírez

Dr. Jaime Eduardo Martínez Sánchez

Resumen—Una empresa aseguradora ha solicitado que alumnos del Tecnológico de Monterrey estimen la probabilidad de ruina de su empresa a horizontes de tiempo finito e infinito. Con la base de datos recibida se realizaron análisis estadísticos para aplicar el modelo de *Cramér-Lundberg* para la estimación de la probabilidad de ruina. Se calculó dicha probabilidad mediante simulación Monte Carlo y por la fórmula de *Pollaczek-Khinchine*.

Index Terms—Probabilidad de Ruina, *Pollaczek-Khinchine*, Monte Carlo

I. INTRODUCCIÓN

El riesgo se define como la posibilidad de perder algo o de tener un resultado no deseado, negativo o peligroso. Este se divide en dos componentes: la probabilidad de que un resultado negativo suceda, y el tamaño de dicho resultado; por lo que tenemos que mientras mayor sea la probabilidad y la pérdida potencial, mayor será el riesgo que se presente [1].

El análisis de riesgo financiero evalúa dicha probabilidad y mediante su estimación, se combina la ocurrencia del riesgo y las posibles pérdidas. Con esto buscamos investigar la probabilidad de ruina de una empresa de seguros realizando un modelo que simule distintas réplicas de la problemática.

Este escrito se organiza de la siguiente manera: en las secciones II y III se presenta la problemática a resolver y se plantean las preguntas de investigación que dirigirán nuestro análisis. En la sección IV se presentan algunos conceptos teóricos de Teoría de Riesgo necesarios para comprender la relevancia de nuestro proyecto.

En las secciones V y VI se presenta la metodología seguida así como los análisis estadísticos necesarios para el cálculo de la probabilidad de ruina; después presentamos los resultados de nuestras estimaciones en la sección VII para después discutir las áreas de oportunidad encontradas en la sección VIII. Finalmente presentamos las conclusiones de nuestro estudio en la sección X.

En las secciones IX y A se presentan datos técnicos del equipo utilizado y se proporciona una liga al código fuente implementado.

Juan Pablo Echeagaray González, Verónica Victoria García De la Fuente, Emily Rebeca Méndez Cruz, Eugenio Santiesteban Zolezzi, Daniel De Zamacona Madero pertenecen al Tec de Monterrey Monterrey, N.L. C.P. 64849, Mexico

II. DESCRIPCIÓN DE LA PROBLEMÁTICA

El reto se enfoca en el análisis del riesgo financiero concentrándose concretamente en las pérdidas económicas que surgen del pago que debe de hacer la compañía ante el caso de un siniestro, si el monto total de pagos exceden el capital que tiene disponible la compañía se producirá su ruina y provocará que salga del mercado.

II-A. Modelo clásico de Cramér-Lundberg

Para definir este modelo se toman en cuenta diversos aspectos como el capital inicial de la compañía, el monto recibido por unidad de tiempo, tamaño y número de reclamos en determinados momentos; todas estas variables proyectadas contra reclamaciones en tiempos aleatorios W_1, W_2, \dots de acorde a un proceso de Poisson homogéneo $\{N(t) : t > 0\}$ con una intensidad $\lambda > 0$. Este supuesto implica que los tiempos entre las llegadas de los eventos son variables aleatorias independientes e idénticamente distribuidas con distribución común exponencial [2]. Aunado a esto se hipotetiza que los tamaños de estas reclamaciones siguen una distribución común F de cola ligera, con las siguientes propiedades:

$$F(0) = 0$$

$$E(Z_k) = \mu > 0$$

La compañía recibe de sus asegurados una prima constante mayor a cero por unidad de tiempo, por lo que al denotar $X(t)$ como el capital de la compañía en tiempo t , se tiene que

$$X(t) = u + ct - \sum_{k=1}^{N(t)} Z_k \quad (1)$$

De forma intuitiva podemos decir que $\sum_{k=1}^0$ es igual a cero, esto hace alusión a que si no hubo ningún siniestro, no podemos suponer que la empresa tendrá que realizar un pago; y que el índice superior de la sumatoria, $N(t)$ es el número de reclamos en el intervalo $(0, t]$. Esta definición de $X(t)$ resulta como el modelo clásico de riesgo, el modelo de *Cramér-Lundberg*.

Para efectos prácticos de modelado y simulación, hemos optado por modificar la nomenclatura básica del modelo de *Cramér-Lundberg* a la siguiente fórmula iterativa:

$$X(0) = U_0$$

$$X(t+1) = X(t) + c - \sum_{k=1}^{N(t)} Z_k \quad (2)$$

Un desglose algebraico de la ecuación 2 es suficiente para ver que es equivalente a la ecuación 1 del modelo original.

III. PREGUNTAS DE INVESTIGACIÓN

Mediante el análisis de la base de datos que nos ha proporcionado el Socio Formador, nos planteamos 2 preguntas con el fin de guiar nuestro estudio:

1. ¿Cuál es la probabilidad de que la compañía aseguradora caiga en ruina en un tiempo finito T ?
2. ¿Cuál es la probabilidad de que la compañía aseguradora caiga en ruina para un horizonte de tiempo infinito?

Para ambas preguntas se adjuntan posibles alteraciones al modelo para mejorar la predicción y control de los episodios donde la compañía tenga que solventar gastos de siniestros y ampliar la predicción en un plazo de tiempo continuo con duración infinita.

Con esto introducimos las cargas de seguridad para las primas, las cuales se vincularían con las primas recolectadas y un factor de carga acompañado por la seriedad de los siniestros, mediante el factor de carga podemos indicar si las primas incitan un mayor riesgo de ruina o si la clientela preferiría buscar otras alternativa. Para la el estado continuo en tiempo infinito tiene como objetivo proporcionar un límite superior para la probabilidad de ruina, usando la desigualdad de Lundberg ajustamos mediante el coeficiente R una medida inversa de riesgo que mientras mayor sea menor es el límite superior. Para el proceso compuesto de Poisson con parámetro λ , obtenemos una R raíz única positiva de la ecuación:

$$\lambda + cr = \lambda M_x(r) \quad (3)$$

Con λ siendo el parámetro de poisson, c la prima de tasa de unidad de tiempo, y $M_x(r)$ es la función generadora de momentos para el indicador de reclamos en cierto punto, con esto es posible derivar el límite superior e inferior para R y así generalizar el modelo para incluir el impacto del reaseguro en caso de ruina.

IV. MARCO TEÓRICO

Cada que se tome una decisión se debe valorar la relación costos-beneficios, de lo contrario no estaríamos evaluando los riesgos que se corren al tomar dicha decisión y las ventajas o desventajas que esta nos traería [1].

Existen riesgos cuando los posibles escenarios con sus resultados son conocidos y existen antecedentes para la estimación de su distribución de frecuencia [3]. El riesgo financiero, también llamado riesgo de crédito o de insolvencia, es considerado como el riesgo de pérdidas en las posiciones dentro y fuera del balance proveniente de movimientos adversos en los precios de mercado. Este hace referencia a la incertidumbre asociada al rendimiento de la inversión debido a la posibilidad de que la empresa no pueda afrontar sus obligaciones financieras, como lo son el pago de los intereses y la amortización de las deudas. Las obligaciones financieras fijas en las que incurren es el único factor por la que ocurre el riesgo financiero [3].

En cuanto a la compañía aseguradora, el modelo de negocios se representa con la relación de ingresos por primas

devengadas e inversiones y de egresos por siniestros y gastos de emisión. Tenemos que la aseguradora aumenta sus ingresos mediante dos alternativas [4]:

- Aumento de la prima devengada, esta se logra a través del aumento de pólizas suscritas o mejora de cartera.
- Manejo óptimo del portafolio de clientes, esto se genera mediante decisiones de inversión que generen el mayor retorno posible.

Respecto a la disminución de los egresos de la aseguradora, esta se produce por medio de dos alternativas [4]:

- Optimización de los cálculos de reservas para cubrir los siniestros.
- Disminución de los gastos de emisión, reduciendo gastos como comisiones y exámenes médicos, así también las pólizas que se cancelan.

Durante el proyecto se hará uso del modelo de Cramér-Lundberg para la definición del modelo clásico de riesgo, como lo hemos descrito anteriormente II-A.

V. METODOLOGÍA

Se realizará una limpieza de la base de datos proporcionada por el Socio Formador; se comprobará la independencia de los montos de los siniestros y la cantidad de siniestros por día mediante una prueba de correlación, después se determinarán las distribuciones que modelan la frecuencia de ocurrencia de los siniestros así como los montos pagados por la aseguradora por cada uno de ellos mediante pruebas de bondad de ajuste; se procederá después a estimar los parámetros de dichas distribuciones con las pruebas de hipótesis correspondientes, se brindarán también intervalos de confianza para cada parámetro estimado.

Con los parámetros estimados con anterioridad se determinará primero una expresión para la prima c que garantice una probabilidad de ruina diferente a 1. Con estos parámetros se realizará una simulación de la probabilidad de ruina de la aseguradora a un horizonte finito T . Finalmente se calculará la probabilidad de ruina a un horizonte de tiempo infinito mediante la fórmula de *Pollaczek-Kinchine* [5].

VI. PROPUESTA METODOLÓGICA

Para la solución de esta problemática hemos usado el lenguaje de programación Python en su versión 3.10.4; para los procesos de lectura, limpieza, exploración y agregación de datos hemos usado la librería *pandas* [6] en conjunto con *matplotlib* [7] y *seaborn* [8], para las pruebas de hipótesis correspondientes así como para la generación de números aleatorios de acuerdo a las distribuciones encontradas se utilizó el módulo *stats* de la librería *scipy* [9] y la librería *statsmodels* [10]. Una liga al código fuente desarrollado se encuentra en el apéndice.

VI-A. Limpieza de la base de datos

Se recibió una base de datos de una empresa aseguradora de coches en formato *csv*, en esta se incluyen los diferentes siniestros ocurridos en determinado tiempo. Se tienen la fecha del mismo, el tipo de auto, el modelo, el monto del siniestro,

si se aplica la cobertura, la cantidad del deducible, si el cliente aplicó el seguro y si el siniestro representó una pérdida total.

Se aplicó un proceso de limpieza de datos en la que se removieron todos los siniestros en los que no se aplicó el seguro con el fin de tener solamente los datos que sí generaron una pérdida económica para la empresa aseguradora.

VI-B. Validación de los supuestos del modelo

Para este caso de estudio se necesitan demostrar 3 supuestos; el que los montos de los reclamos sean independientes de la frecuencia de los siniestros, que la frecuencia de siniestros pueda ser modelada como un proceso de Poisson Homógeno y que la distribución del monto de los reclamos F sea de cola ligera, estos supuestos vienen del artículo de la Dra. Ekaterina *Probabilidad de ruina en el modelo clásico de Cramér-Lundberg* publicado en el 2011 [2].

En cada prueba de hipótesis y cálculo de intervalos de confianza en el presente escrito se utilizará un nivel de significancia α de 0.05.

VI-B1. Prueba de independencia: El primer paso en nuestro estudio será verificar que exista independencia entre el monto de los reclamos y el número de reclamos por día; la modelación que seguimos originalmente propuesta por la Dra. Ekaterina en [2] asume que existe independencia entre estas 2 variables, en caso contrario, no podremos utilizar el modelo de *Cramér-Lundberg*. Como un primer heurístico, podemos visualizar estas 2 variables en el diagrama de dispersión que presentamos en la figura 1.

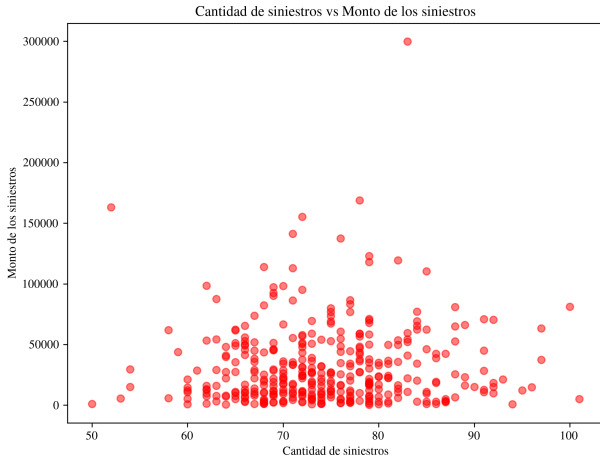


Figura 1: Diagrama de dispersión del monto de los siniestros contra la cantidad de reclamos por día

De forma intuitiva, podemos ver que no existe una relación aparente con los datos; se formalizará ahora este supuesto mediante una prueba de hipótesis para el coeficiente de correlación ρ de Spearman ya que puede medir dependencias más allá de una lineal, con un nivel de significancia $\alpha = 0.05$

$$H_0 : \rho = 0$$

$$H_1 : \rho \neq 0$$

Encontramos que el coeficiente de correlación para estas 2 variables es $\rho = 0.06501$. El p -value obtenido de esta prueba fue 0.215, por lo tanto, no podemos rechazar la hipótesis nula, supondremos que existe independencia entre estas 2 variables.

VI-B2. Distribución de la frecuencia de los reclamos: Realizaremos una prueba de bondad de ajuste *chi-square* para las cantidad de siniestros, $N(t)$ que hay un dado periodo. En la figura 2 presentamos el comportamiento de este proceso de conteo en la muestra recibida.

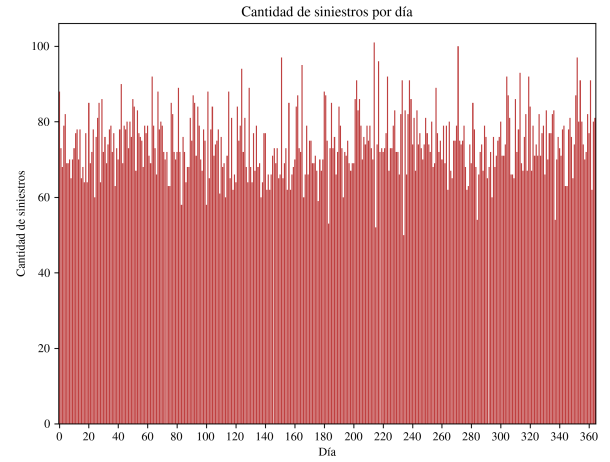


Figura 2: Distribución de la frecuencia de reclamos por día

Como esperamos modelar esta cantidad como proveniente de un proceso de Poisson homogéneo, se realiza una prueba de bondad de ajuste para la distribución de Poisson; realizamos esta prueba con un nivel de significancia $\alpha = 0.05$. Las hipótesis nula y la hipótesis alternativa que planteamos son entonces:

$$H_0 : N(t) \sim \text{Poisson}(\lambda t)$$

$$H_1 : N(t) \not\sim \text{Poisson}(\lambda t)$$

El estadístico de prueba obtenido fue de 16.1425 con un p -value de 0.373, el cual es mayor nuestro nivel de significancia, por lo que no se puede rechazar la hipótesis nula, podemos continuar nuestro estudio suponiendo que la cantidad de reclamos en cada día puede ser modelado como un proceso de Poisson homogéneo.

VI-B3. Distribución del monto de los reclamos: Para el caso de la distribución del monto de los reclamos haremos la prueba de *Anderson-Darling* [11] y aplicaremos la prueba de *Lilliefors* [12] para la distribución exponencial. Hemos decidido realizar 2 pruebas de bondad de ajuste dado que uno de los supuestos más fuertes de nuestro modelado es que la distribución del monto de los reclamos es de cola ligera; el que la cola sea pesada dificultará enormemente la determinación de la fórmula analítica de la probabilidad de ruina a un horizonte infinito. Se muestra una visualización de la distribución de los montos en la figura 3.

Para este estudio supondremos que esta variable aleatoria Z sigue una distribución exponencial con un parámetro μ . Para esta variable se hará la prueba de ajuste para la distribución

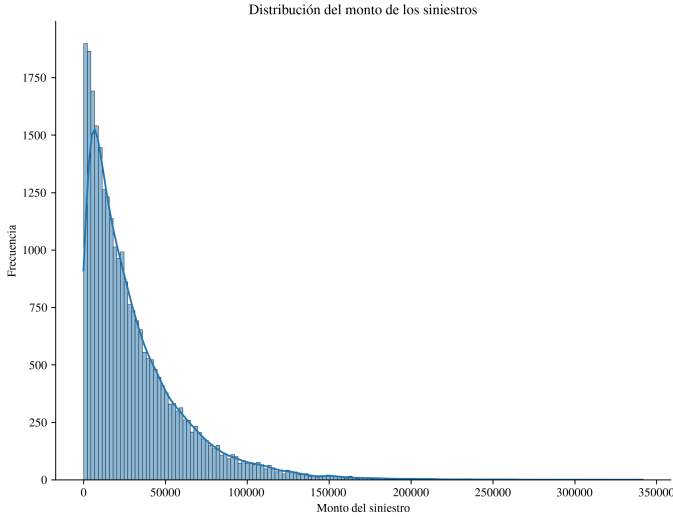


Figura 3: Distribución del monto de los reclamos

exponencial con el mismo nivel de significancia $\alpha = 0.05$. Las hipótesis son las siguientes:

$$\begin{aligned} H_0 : Z &\sim \text{Exp}(\mu) \\ H_1 : Z &\not\sim \text{Exp}(\mu) \end{aligned}$$

El p -value obtenido para la prueba de *Lilliefors* fue de 0.09746, y para la *Anderson-Darling* fue 0.109; ambas pruebas indican que no se puede rechazar la hipótesis nula, el modelado de esta variable aleatoria para nuestro caso de estudio será el de una distribución exponencial.

VI-B4. Discriminación de parámetros: Ya que conocemos las distribuciones que modelan los fenómenos de interés, nos queda solamente estimar los parámetros de dichas distribuciones; como ambos son en este caso las medias poblacionales de cada variable, realizaremos 2 pruebas de hipótesis sobre la media poblacional, con un nivel de significancia $\alpha = 0.05$ para ambos casos.

$$\begin{aligned} H_0 : \mu &= \bar{\mu} \\ H_1 : \mu &\neq \bar{\mu} \\ H_0 : \lambda &= \bar{\lambda} \\ H_1 : \lambda &\neq \bar{\lambda} \end{aligned}$$

Para la prueba de la media del monto de los reclamos estimamos que la media poblacional μ es \$30771.37 con p -value de 1, para la media poblacional de reclamos por día λ estimamos un valor de 74.30 con un p -value de 1. Como estos valores no son menores que el nivel de significancia establecido, fallamos el rechazo de la hipótesis nula, y de ahora en adelante supondremos que estos son los parámetros de la distribuciones encontradas.

El intervalo de confianza al 95 % de estos dos parámetros se puede calcular mediante la distribución normal (debido a que no se conoce la varianza poblacional de estos 2 parámetros). Los intervalos que hemos obtenido son:

$$\begin{aligned} \text{CI}_\lambda &= (73.439, 75.169) \\ \text{CI}_\mu &= (30, 399.680, 31, 143.071) \end{aligned}$$

Por lo que cualquier λ y μ dentro de estos intervalos darán un nivel de significancia del 5 % al modelo utilizar.

VI-B5. Cálculo de prima: Introducimos ahora una expresión para el cálculo de la prima que debe recibir la empresa aseguradora como una función del factor de recargo θ ; heurísticamente podríamos ver el factor de recargo como una medida de que tan elevada será la prima que la aseguradora le cobre a sus clientes. Se sabe que para que la probabilidad de ruina de la aseguradora no converja a 1, se tiene que cumplir la siguiente condición [13]:

$$c > \lambda\mu \quad (4)$$

De esta expresión podemos definir ahora la prima como una expresión en función de un parámetro θ que denominaremos *factor de recargo*:

$$c = (1 + \theta)\lambda\mu \quad (5)$$

Para que se cumpla la restricción de la expresión 4 θ tiene que ser estrictamente positivo. Para efectos de experimentación uno podría hacer este valor negativo pero no menor a -1 y observar como se comporta la probabilidad de ruina en las aproximaciones frecuentistas.

El parámetro que proponemos es similar a la *carga de seguridad* que la Dra. Ekaterina define en su artículo [2].

VI-C. Simulación de probabilidad de ruina por Monte Carlo Crudo

Con los parámetros estimados en las secciones anteriores se realizará una simulación de la probabilidad de ruina mediante Monte Carlo Crudo. Nos interesará aproximar dicha probabilidad para un horizonte finito de tiempo T mediante un número de simulaciones N . Nuestra estimación de la probabilidad de ruina para un capital inicial U_0 y un factor de recargo θ será la proporción de trayectorias en las que alguna vez el capital de la empresa U_t fue menor a 0.

Para dicha aproximación se calculará a su vez un intervalo de confianza al 95 % para después calcular su margen de error ME . El Socio Formador nos ha solicitado que dicho margen de error sea estrictamente menor a 0.05. En caso de que el margen de error calculado sea mayor al estipulado, se incrementará el número de simulaciones N hasta que satisfagamos las necesidades de la empresa aseguradora.

VI-D. Cálculo de la probabilidad de ruina con fórmula de Pollaczek-Khinchine

Finalizaremos nuestro caso de estudio al obtener una expresión analítica para la probabilidad de ruina a un horizonte de tiempo infinito empleando la fórmula de *Pollaczek-Kinchine* [5]; para su desarrollo supondremos que la frecuencia de los reclamos sigue un proceso de Poisson homogéneo de intensidad λ , y que la distribución que sigue el monto de los reclamos es una exponencial con parámetro μ .

La fórmula de *Pollaczek-Khinchine* es un resultado obtenido al analizar el monto de superávit que generaría la empresa aseguradora [14], es decir:

$$\begin{aligned} Y(t) &= u - U(t) = u - (u + ct - \sum_{k=1}^{N(t)} Z_k) \\ &= \sum_{k=1}^{N(t)} Z_k - ct \end{aligned}$$

Cuando $Z(t) > u$ se dice que la compañía ya quebró de igual manera cuando $U(t) < 0$. Esto es debido a que

$$\begin{aligned} U(t) &< 0 \\ u + ct - \sum_{k=1}^{N(t)} Z_k &< 0 \\ ct - \sum_{k=1}^{N(t)} Z_k &< -u \\ \sum_{k=1}^{N(t)} Z_k - ct &> u \\ Z(t) &> u \end{aligned}$$

Ahora definamos Y_k como sigue:

$$Y_k = \begin{cases} \max\{Z(t) : Z(t) \geq 0\} & \text{si } k = 0 \\ \max\{Z(t) : Z(t) \geq 0\} - Y_{k-1} & \text{si } k > 0 \end{cases} \quad (6)$$

Por lo que la condición de ruina se puede escribir como sigue:

$$M = \sum_{k=0}^K Y_k \geq u \quad (7)$$

La cual conlleva a la fórmula de *Pollaczek-Khinchine*:

$$\psi(u) = P(M > u) = \frac{\theta}{1 + \theta} \sum_{n=0}^{\infty} \left(\frac{1}{1 + \theta} \right)^n \bar{B}_0^{*n}(u) \quad (8)$$

En donde $\psi(u)$ es la probabilidad de ruina con horizonte infinito con utilidad inicial u . θ cumple que $c = (1 + \theta)\lambda\mu$. $\bar{B}_0^{*n}(u)$ es la n -ésima convolución de la distribución cola de los reclamos [15]. La 0-ésima de $\bar{B}_0(u)$ convolución se define en este caso es 0 si $u \geq 0$, y 1 si $u < 0$. La convolución, en el caso que $n = 1$, se define como sigue:

$$\bar{B}_0^{*(1)}(u) = \bar{B}_0(u) \quad (9)$$

Y en caso de que $n > 1$:

$$\begin{aligned} \bar{B}_0^{*(n)}(u) &= (\bar{B}_0^{*(n-1)}(u) * \bar{B}_0(u)) \\ &= \int_{-\infty}^{\infty} \bar{B}_0^{*(n-1)}(x) \bar{B}_0(u - x) dx \end{aligned}$$

Si se sustituye \bar{B}_0 como la respectiva distribución cola de una exponencial se lleva a la siguiente fórmula para la probabilidad de ruina en un horizonte infinito [5]:

$$\psi(u) = \frac{1}{1 + \theta} \cdot \exp\left(-\frac{u\theta}{(1 + \theta)\mu}\right) \quad (10)$$

VII. EXPERIMENTACIÓN Y RESULTADOS

VII-A. Simulación de probabilidad de ruina por Monte Carlo Crudo

Respondiendo a la primera de nuestras preguntas de investigación, se realizará una simulación de la probabilidad de ruina mediante Monte Carlo Crudo. Nos enfocaremos en estimar la probabilidad de ruina para un horizonte de tiempo T de 5,000 días, y se realizarán 10,000 simulaciones (N). Para esta simulación se usaron los parámetros λ y μ estimados en la sección VI-B4, se supuso que la empresa aseguradora comenzó sus operaciones con un capital inicial $U_0 = 1,000,000$, y asignamos un factor de recargo $\theta = 0.05$, que resultó en una prima $c = 2,400,761.65$. Presentamos una visualización de las trayectorias simuladas en la figura 4.

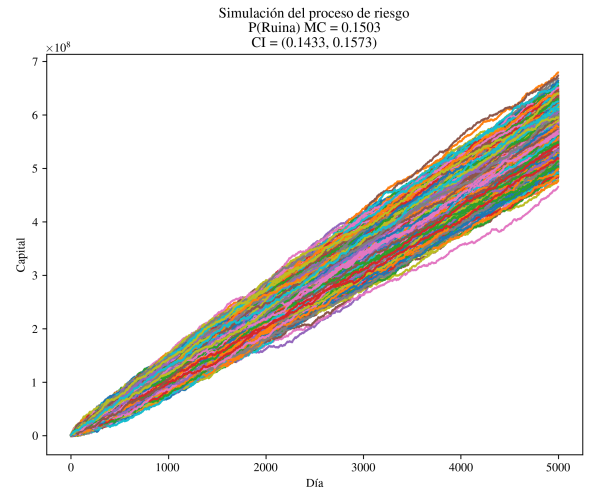


Figura 4: Trayectorias generadas de una simulación Monte Carlo Crudo, $\theta = 0.05$

Para esta simulación hemos estimado una probabilidad de ruina $\bar{\psi}_{5000}(U_0) = 0.1503$. Hemos calculado también un intervalo de confianza al 95 % para esta proporción aprovechando el Teorema Central del Límite, obtuvimos un intervalo de confianza $CI_{\bar{\psi}} = (0.1433, 0.1573)$; el margen de error del intervalo de confianza fue $ME = 0.007$, que es menor al margen de error solicitado por el Socio Formador.

VII-B. Cálculo de la probabilidad de ruina con la fórmula de Pollaczek-Khinchine

Para resolver nuestra segunda interrogante, calcularemos la probabilidad de ruina para la empresa aseguradora para un tiempo infinito utilizaremos la ecuación 10 que definimos en la sección VI-D. Utilizaremos como capital inicial u \$1,000,000, el factor de recargo θ se mantendrá en 0.05 y el parámetro de la distribución exponencial del monto de los reclamos μ será el estimado en la sección VI-B4.

Sustituyendo estos datos en la ecuación 10 encontramos que la probabilidad de ruina a un horizonte infinito es $\psi(u) = 0.202$. Notamos que este valor difiere del que estimamos en la sección VII-A.

VIII. DISCUSIÓN DE RESULTADOS

A pesar de que el presente escrito ha cumplido con todas las especificaciones solicitadas por el Socio Formador, hemos encontrado algunas áreas de oportunidad que de ser resueltas añadirían aún más valor a nuestra investigación.

VIII-A. Complejidad computacional

El método de Monte Carlo Crudo que hemos usado en este trabajo se vuelve computacionalmente ineficiente cuando se quieren estimar probabilidades a horizontes de tiempo cada vez más grandes, su complejidad computacional de tiempo en el peor de los casos posibles es de orden $O(NT)$, para las especificaciones de nuestra investigación, nuestro algoritmo requeriría como máximo un total de 50,000,000 de operaciones, tarea que se vuelve en extremo lenta cuando se usan lenguajes de programación dinámicos como lo es Python.

Creemos que el uso de un lenguaje de programación más eficiente en conjunto con la implementación de un método de reducción de varianza, ayudarían a reducir la complejidad computacional de nuestras simulaciones a la vez que nos facilitan la estimación de probabilidades a horizontes de tiempo cada vez más grandes.

VIII-B. Diferencia entre probabilidad frecuentista y analítica

Como hemos observado en la sección VII-B, la probabilidad de ruina frecuentista que hemos obtenido mediante el proceso de simulación difiere de la calculada con la fórmula de *Pollaczek-Khinchine* que hemos derivado para nuestro caso de estudio. Conjeturamos que este fenómeno ocurre ya que un horizonte de tiempo T de 5,000 días no es lo suficientemente grande como para acercarse a la probabilidad analítica de que haya ruina en un horizonte infinito; sin embargo, para resolver este problema, se requeriría aumentar el horizonte T y el número de simulaciones N para garantizar un margen de error pequeño en nuestra estimación, lo que nos lleva de nuevo al área de oportunidad anterior.

IX. RECURSOS UTILIZADOS

La implementación del código fuente fue corrida en una máquina de sistema operativo Ubuntu 22.04.1 LTS x86_64. El equipo disponía de un procesador Intel(R) Core(TM) i5-1035G1 CPU @ 3.6GHz, y tenía disponible 8 GB de RAM. Especificaciones sobre las versiones de los paquetes usados en este proyecto pueden ser encontrados en la documentación de nuestro repositorio A.

X. CONCLUSIONES

El análisis de la probabilidad de ruina seguirá siendo una de las principales tareas de la Teoría de Riesgo, el buen análisis de los registros históricos de una empresa en conjunción con la implementación de algoritmos de simulación eficientes es imperativo para cualquier departamento de gestión de riesgos.

Dada la dificultad de encontrar fórmulas matemáticas cerradas para la probabilidad de ruina a horizontes de tiempo finitos, el uso de simulaciones como la que proponemos

representa una buena alternativa para aquellas empresas que deseen entrar a este riesgoso mercado.

La metodología desarrollada solamente se vería beneficiada de la implementación de un mejor método de simulación que no tenga una complejidad computacional tan alta; esto le daría el poder a las aseguradoras de estimar la probabilidad de ruina de su empresa a largo plazo con un alto grado de fiabilidad.

APÉNDICE A CÓDIGO FUENTE

El código implementado puede ser consultado en el siguiente repositorio aquí.

REFERENCIAS

- [1] B. Echemendía Tocabens, "Definiciones acerca del riesgo y sus implicaciones," *Revista Cubana de Higiene y Epidemiología*, vol. 49, pp. 470 – 481, 12 2011. [Online]. Available: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1561-30032011000300014&nrm=iso
- [2] J. d. C. Jiménez Hernández, A. D. Maldonado Santiago *et al.*, "Probabilidad de ruina en el modelo clásico de cramer-lundberg," *REPOSITORIO NACIONAL CONACYT*, 2011.
- [3] C. L. Bazzani C. and E. A. Cruz Trejos, "Análisis de riesgo en proyectos de inversión un caso de estudio," *Scientia et Technica*, vol. 1, no. 38, jun. 2008. [Online]. Available: <https://revistas.utp.edu.co/index.php/revistaciencia/article/view/3771>
- [4] E. A. Chíncharo and M. M. Puente, "Aplicación de un modelo de evaluación crediticia para reducir el riesgo en la cartera de clientes de una compañía aseguradora," *Industrial data*, vol. 14, no. 2, pp. 59–66, 2011.
- [5] f. given i=David, "61/64 Probabilidad de ruina en tiempo continuo usando la fórmula de Pollaczek-Khinchine." 6 2020. [Online]. Available: <https://www.youtube.com/watch?v=hbfbaC5S4wY>
- [6] Wes McKinney, "Data Structures for Statistical Computing in Python," in *Proceedings of the 9th Python in Science Conference*, Stéfan van der Walt and Jarrod Millman, Eds., 2010, pp. 56 – 61.
- [7] J. D. Hunter, "Matplotlib: A 2d graphics environment," *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90–95, 2007.
- [8] M. L. Waskom, "seaborn: statistical data visualization," *Journal of Open Source Software*, vol. 6, no. 60, p. 3021, 2021. [Online]. Available: <https://doi.org/10.21105/joss.03021>
- [9] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, Í. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, pp. 261–272, 2020.
- [10] S. Seabold and J. Perktold, "statsmodels: Econometric and statistical modeling with python," in *9th Python in Science Conference*, 2010.
- [11] D. L. Evans, J. H. Drew, and L. M. Leemis, "The distribution of the kolmogorov-smirnov, cramer-von mises, and anderson-darling test statistics for exponential populations with estimated parameters," *Communications in Statistics—Simulation and Computation*, vol. 37, no. 7, pp. 1396–1421, 2008.
- [12] H. W. Lilliefors, "On the kolmogorov-smirnov test for the exponential distribution with mean unknown," *Journal of the American Statistical Association*, vol. 64, no. 325, pp. 387–389, 1969.
- [13] T. Rolski, H. Schmidli, V. Schmidt, and J. L. Teugels, *Stochastic processes for insurance and finance*. John Wiley & Sons, 2009.
- [14] K. Burnecki, P. Miśta, and A. Weron, "What is the best approximation of ruin probability in infinite time?" *Applicaciones Mathematicae*, vol. 2, no. 32, pp. 155–176, 2005.
- [15] f. given i=Marco, "Convolutions," 2021. [Online]. Available: <https://www.statlect.com/glossary/convolutions>
- [16] I. Minitab, "Minitab," 2020. [Online]. Available: <http://www.minitab.com/en-US/products/minitab/>