Proyecto de Inteligencia Artificial

David Sánchez Iglesias

2023

1. Introducción

En el presente proyecto se simula y estudia el comportamiento de un modelo de seguros de riesgo (Insurance Risk Model). El objetivo es estudiar la evolución del capital de la firma de seguros. Dicho problema se encuentra en el libro "Simulation, Fifth edition" de Sheldon M. Ross, en el capítulo 7, epígrafe 7.6, página 122.

2. Problema

Se tiene una compañía de seguros de riesgo con la cantidad de clientes que poseen un seguro de esta. Se tiene que se generan, contra la empresa, reclamaciones de acuerdo a un proceso de Poisson independiente, con una tasa común λ, por una cantidad de dinero cuya distribución F se conoce. Se tiene también que los nuevos cliente firman para la obtención de un seguro de acuerdo a otro proceso independiente de Poisson, con tasa v, y que cada uno de los asegurados se mantiene con la compañía por un tiempo cuya distribución es exponencial, con tasa μ. Se sabe que cada asegurado paga a la compañía una cantidad de dinero fija c por unidad de tiempo. Empezando con una cantidad fija de clientes y un capital inicial, la compañía está interesada en simular este sistema para saber si el capital de la firma siempre se mantiene no negativo todo el tiempo hasta un momento dado.

3. Modelación del problema

3.1. Variables

Para simular el sistema, se creó la clase *InsuranceRiskModel*, con todas las variables y procesos que se usarán durante la simulación. Entre sus variables se encuentran:

- *time*: la unidad de tiempo en la que se encuentra el sistema en un momento dado.
- *n_pols*: el número de asegurados (policyholders) en un tiempo específico.

- *capital*: recoge el capital que posee la empresa en el tiempo actual.
- *lost_pols*: cantidad de asegurados que dejaron la empresa. Este valor aumenta con cada ciclo y al final recoge la totalidad de asegurados que dejaron la empresa durante toda la simulación.
- *events_queue*: un heap, una cola con prioridad que almacena los eventos y los ordena según el tiempo en el que les toca ser ejecutado.
- *events*: diccionario que recoge los diferentes tipos de eventos junto a la función que los genera.
- *pol_list*: contiene los id de los asegurados que posee la empresa en una unidad de tiempo. Se actualiza con cada cambio que ocurre entre los asegurados, o sea, si aparece o se pierde un asegurado.

Para reiniciar las variables en caso de que se quiera ejecutar la simulación varias veces, se crearon otras variables de referencia:

- *limit_time*: tiempo límite, introducido por el usuario al crear la clase.
- *pol_pay*: lo que paga cada asegurado (c)
- *starting_capital*: el capital inicial
- *initial_n_pols*: la cantidad inicial de asegurados que posee la empresa.
- *current_pol_id*: para establecer un valor, un codigo de identificación a cada asegurado.

El resto de las variables tienen como objetivo almacenar las tasas de los diferentes procesos aleatorios así como las funciones para la generación de los valores aleatorios y al almacenamiento de estadísticas recogidas durante el proceso.

4. Flujo del programa

La clase cuenta con tres métodos principales:

- *run*: se encarga de sacar los eventos de la cola, producir los cambios que estos conllevan al sistema y actualizar el tiempo. Solo para si la cola está vacía (lo cual significa que ya no quedan eventos en lo que resta de tiempo) o si se llega al tiempo límite establecido por el usuario.
- *initialize*: restablece los valores de todas las variables de lo simulación a sus valores iniciales.
- *simulate*: ejecuta la cantidad de simulaciones que se le pasaran como argumento a la función.

5. Resultados y experimentos

Se escogieron los siguientes valores para las pruebas:

```
model = InsuranceRiskModel(
    new_pol_rate=20,              # 15 nuevos clientes/mes
    lost_pol_rate=0.05,           # 5% de clientes se pierden
    starting_capital=1_00,
    starting_policy_holders=500,
    pol_time_fun= policyholder_time_in_the_firm,
    next_claim= new_claim,
    claim_rate= 5,
    limit_time=120,                # 120 meses
    pol_pay=150                    # $150 por poliza/mes
)
```

Se realizaron varias pruebas y gráficos para estudiar el comportamiento del sistema y se observó un resultado interesante. Acontinuación se muestra un gráfico que ilustra cómo cambia la variable objetivo con el paso del tiempo:

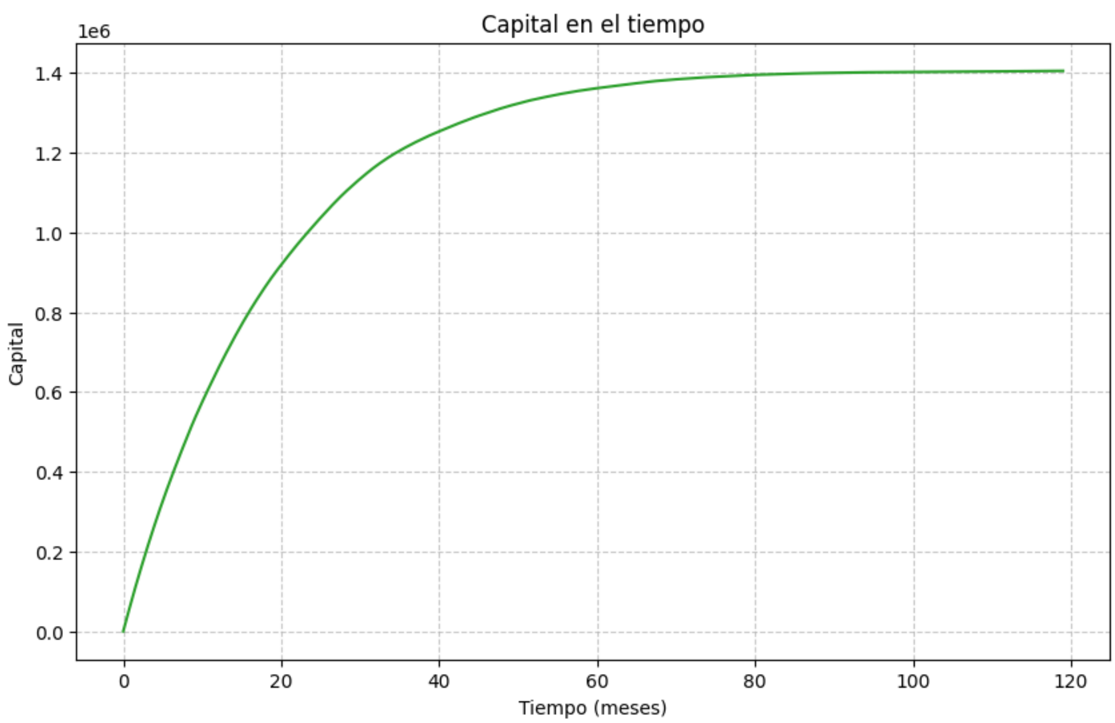

Figura 1: Capital por unidad de tiempo

En la figura claramente se ve que el capital de la empresa nunca es negativo, sin embargo tiende a estabilizarse hasta casi no variar. La razón detrás de este comportamiento se debe en parte al flujo de asegurados que entran y salen de la empresa.

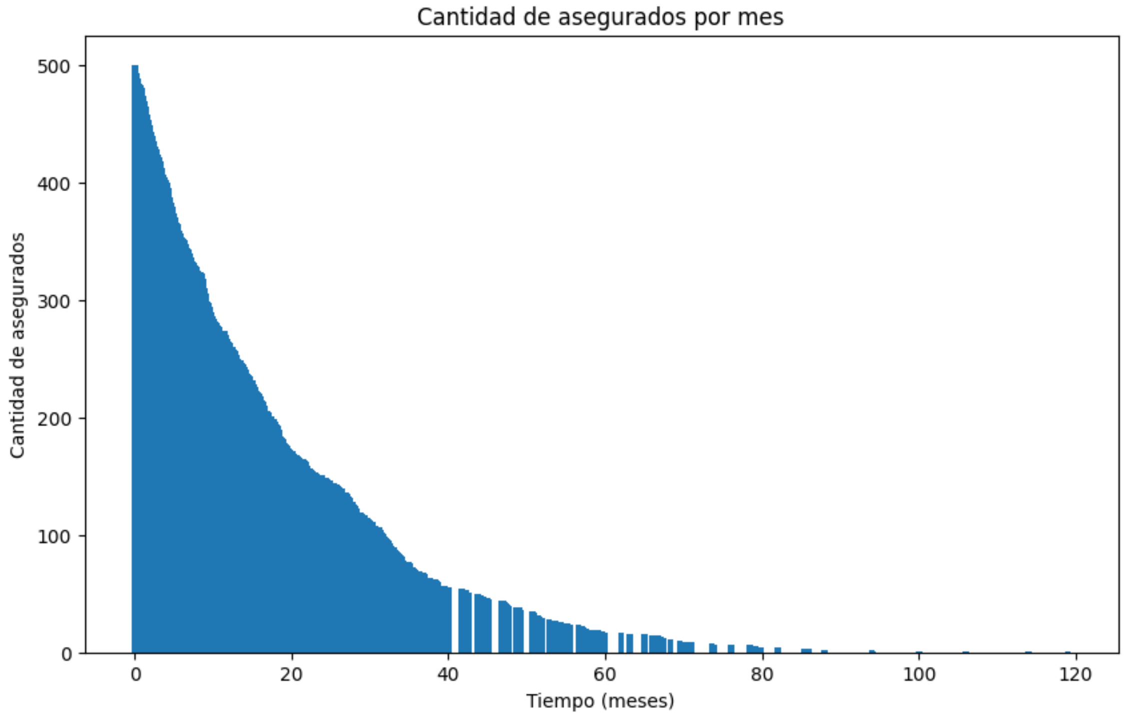

Figura 2: Cantidad de asegurados por unidad de tiempo

Esta gráfica muestra cómo hay una disminución clara de la cantidad de asegurados a medida que pasa el tiempo. Ya que, según los datos con los que se cuenta, el dinero que pagan los asegurados es la única fuente de ingresos de la empresa, no sorprende que el capital se comporte de esa manera. De hecho, al realizar más simulaciones queda en evidencia que esta situación siempre ocurre, y la siguiente figura así lo muestra:

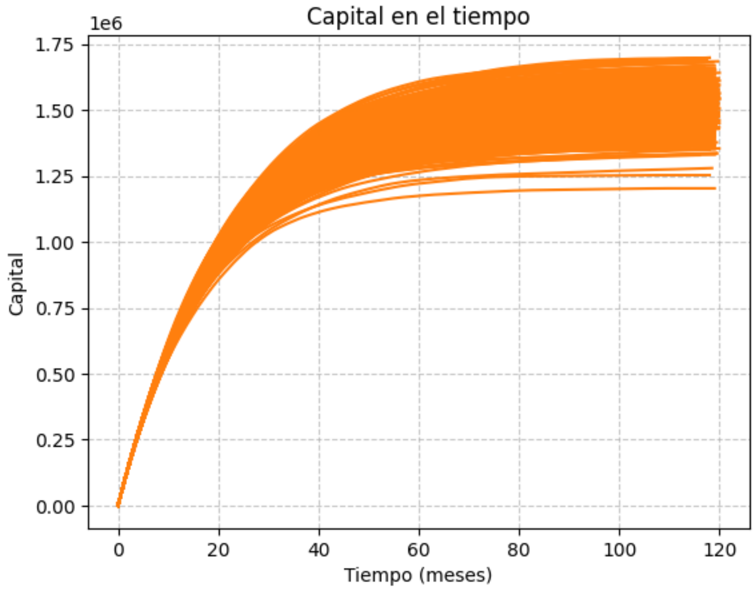

Figura 3: Capital por unidad de tiempo, en 1000 simulaciones

Esta figura muestra que el capital parece comportarse igual, como si distribuyera de una forma específica.

5.1. Hipótesis: El capital de la empresa distribuye de acuerdo a una función.

A vista, la gráfica se asemeja mucho a una exponencial negativa, una función logística o una logarítmica. Se realizó entonces un ajuste de modelos para verificar si alguna de estas describían el comportamiento del capital.
Fórmulas probadas:

- Logística: $C(t) = \frac{L}{1+e^{-k(t-t_0)}}$
- Exponencial negativa: $C(t) = L(1 - e^{-kt})$
- Logarítmica: $C(t) = a + b\ln(t)$

El resultado:

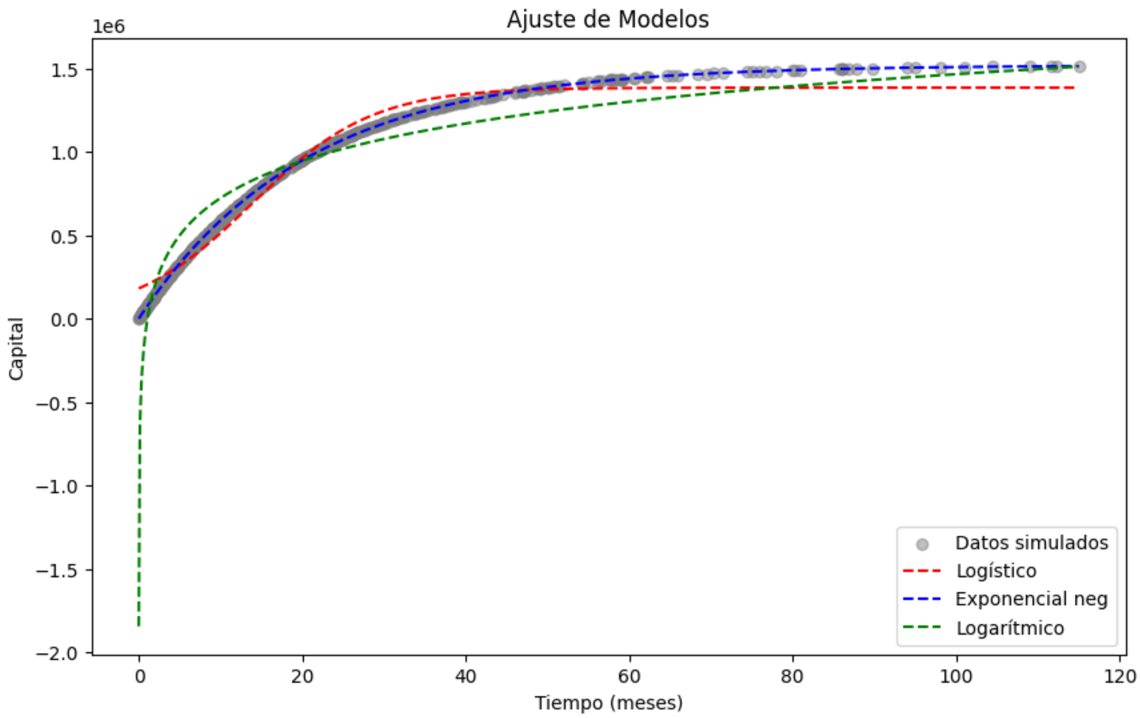

Figura 4: Ajustes de diferentes modelos a los datos

Evidentemente la exponencial negativa parece ser la que más se ajusta, seguida de la logística, mientras la logarítmica se desvía. Una última comprobación gráfica, esta vez con los residuales, nos muestra que, en efecto, la exponencial es la que parece ajustarse más al comportamiento de los datos extraídos al encontrarse todos sus residuos sumamente cercanos a cero.

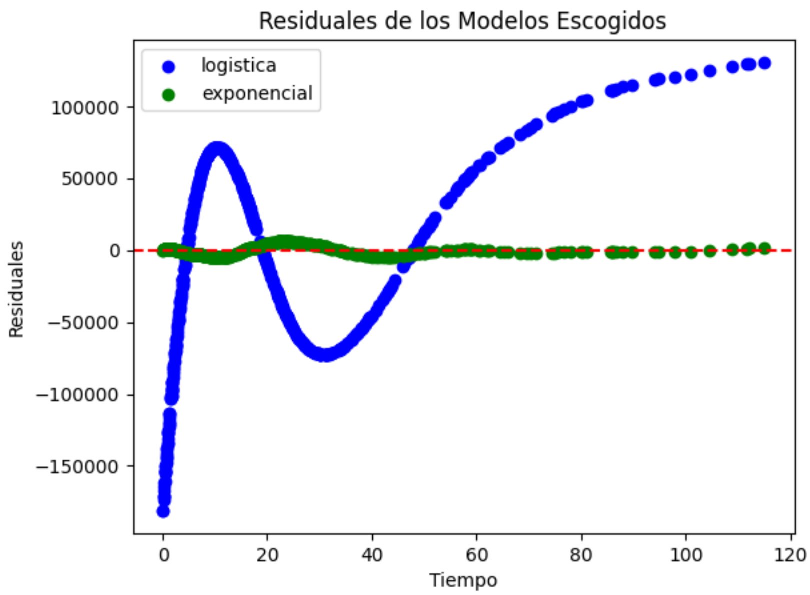

Figura 5: Residuales de la logística y la exponencial

Para comprobarlo se realizó el cálculo del error medio cuadrático.

$$MSE(F) \equiv E_F[(g(X_1, X_2, \ldots X_n) - \theta(F))^2] \tag{1}$$

Dicho resultado se normalizó con la varianza para una mejor visualización. Como resultado, el error da aproximadamente 0.00505, lo que se traduce en un 0.005 % de la varianza de los datos. En otras palabras, la función exponencial propuesta describe el comportamiento del capital durante la simulación. Se puede considerar que esta función sirve como función de distribución para la variación del capital de esta empresa.

6. Conclusión

Con los datos proporcionados y los valores establecidos para la simulación, el capital nunca será negativo pues su función de distribución (la exponencial recién estudiada) nunca es negativa.