



COLEGIO
NACIONAL
DE ACTUARIOS

ACTUARIOS TRABAJANDO

REVISTA MEXICANA DE INVESTIGACIÓN ACTUARIAL APLICADA

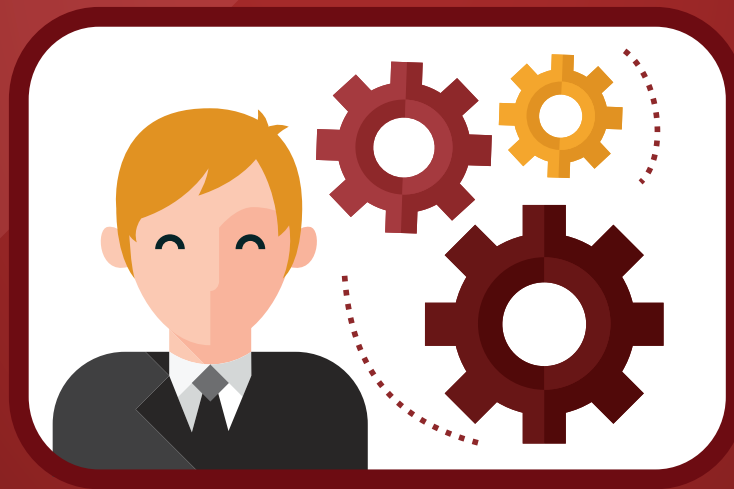
 q_x μ_x d_x ${}_tV_x$  A_x l_x $\ddot{a}_x$ p_x

Tabla de Contenidos

Carta editorial del presidente	4
---------------------------------------	----------

ACTUARIOS OPINANDO

Plan de Negocios: La brújula del Empresario	8
--	----------

Tomás Cortés Zulbarán y Francisco Ortiz Arango

ACTUARIOS TRABAJANDO

On a predictive model of the box office revenue	13
--	-----------

Gerardo Ortiz Arellano, Manuel Jiménez Lizárraga and Viacheslav Kalashnykov

Análisis del Riesgo país de México a partir de variables económicas	22
--	-----------

Ana María Hernández Piña y José Eduardo Becerra López

Retail Pricing Optimization with Monte Carlo Tree Search	41
---	-----------

Mauricio Hanono Gómez, Celeste Rodríguez Carreón and Nataliya Kalashnykova

Huracán Otis. ¿Cisne Negro o error en los modelos predictivos?	49
---	-----------

Marco A. De La Rosa Rodríguez

INVITADOS TRABAJANDO 15

Modelando Datos con Palitos y Bolitas: Algoritmo de Mapper	63
---	-----------

Salvador Mancilla, Eymard Hernández López y Giovanni Wences

PREMIO FUNDADORES **78**

Carta del Presidente del Consejo Directivo de la AMA

Act. Fernando Patricio Belaunzarán Barrera

3er lugar. Propuesta de aplicación de economía del comportamiento y aprendizaje máquina para incrementar el ahorro a largo plazo **82**

Antonio Cruz Torres

2do lugar. Una nueva aproximación a la constitución de una reserva actuarial de arriba p'abajo **91**

José Daniel López Barrientos y Humberto Plata Gallegos

CURSOS Y TALLERES **113**

Estimados Colegas,

Comenzamos un acelerado año 2024 con numerosos proyectos y el gran entusiasmo que proviene de iniciar un nuevo ciclo que otorga siempre la oportunidad de rectificar, mejorar y cambiar.

En esta ocasión invito a una reflexión que me viene a la mente cuando toca revisar con cada Año Nuevo los planes de trabajo y confirmar los avances en las metas esperadas. Y un objetivo que resulta inevitable en nuestras asociaciones y gremios, es la búsqueda incesante de incrementar la membresía y lograr un mayor voluntariado en las actividades que se realizan.

En ese sentido, uno de los factores fundamentales de atracción para el CONAC se centra en el intercambio de ideas, el impulso a la investigación y la dispersión de conocimiento. Fomentar estas actividades son y serán piedra angular de un Colegio de Profesionistas.

Bajo esos principios, con mucho gusto presentamos nuestra edición número 15 de la **Revista Actuarios Trabajando** que conjuga diversos temas que estamos seguros serán de su interés y muestran la flexibilidad y capacidad de nuestra profesión. Disfrutarán de la lectura de variados artículos que incluyen desde una revisión de las empresas para incursionar en nuevos proyectos, las variables económicas en el análisis del riesgo país, el sorpresivo Huracán Otis, los algoritmos de Mapper, la venta de artículos por menudeo a través de apps digitales o el desarrollo de modelos predictivos para la inversión de recursos en el área del entretenimiento (primeros artículos que publicamos en inglés).

Asimismo, este volumen publica con mucho orgullo los trabajos ganadores del *Tercer Premio Fundadores de la Asociación Mexicana de Actuarios*, donde podrán leer sobre la economía del comportamiento y relevancia de incrementar el ahorro a largo plazo y la revisión en la constitución de las reservas actuariales como fundamento de la seguridad social.

Agradecemos la participación a los autores por ayudarnos en nuestro cometido y abrimos la invitación para la publicación de más trabajos entre nuestros miembros, así como la adhesión de actuarios en todas las áreas para que a través de nuestros foros, seminarios, cursos y comités tengan una plataforma de difusión.

Mucha salud y éxito para el 2024.

Act. Elsa L. González Franyutti
Presidente del CONAC 2023-2025

EDITOR EN JEFE

José Daniel López Barrientos

CONSEJO EDITORIAL

Elsa Lorena González Franyutti

Mauricio Gabriel Arredondo Fernández Cano

Mónica Rocío Soler Alatorre

Francisco Javier Ruíz de la Peña Olea

Martha Guadalupe Pichardo Rojas

Patricio Fernando Belaunzarán Barrera

Anayelli Montagner Miranda

Mabel Ramírez Camacho

Miguel de la Garza Camacho

Beatriz Brambila Treviño

EDITORES ASOCIADOS

María de Lourdes Díez Gutiérrez Igartúa

Martha Reyes Villa

Jennifer Rangel Madariaga

Claudia López Soto

Juan Sadurni Gil

Manuel Jiménez Lizárraga

CORRECCIÓN DE ESTILO

Ana Pamela Flores Herrera

DISEÑO GRÁFICO

Priscilla Camargo Bacha

Prof. Miguel Serrano 21 PH.
Col. Del Valle Centro, 03100. CDMX
info@conacmexico.org.mx.

ISSN / DOI EN TRÁMITE

ENERO 2024 - NÚMERO 15



COLEGIO
NACIONAL
DE ACTUARIOS

¡Queremos oírte!

¡Tu aporte es necesario!

¡Queremos saber de ti!

Actuarios Trabajando
es una revista hecha por
Actuarios, con Actuarios y
para Actuarios, y necesita de
*tu aporte profesional con tu
toque personal.*

No dejes pasar la
oportunidad de que tu
voz profesional sea oída,
y tus trabajos sean
conocidos por los
demás miembros de
nuestro gremio.

Los Actuarios certificados
del CONAC obtienen horas
de educación continua por
sus contribuciones
aceptadas para la revista.



Envía tus contribuciones a alguno de los miembros del Consejo Editorial,
o escribe por correo-e a: info@conacmexico.org.mx

ACTUARIOS OPINANDO

Plan de Negocios: La brújula del Empresario

Tomás Cortés Zulbarán y Francisco Ortiz Arango



Plan de Negocios: La brújula del Empresario

Tomás Cortés Zulbarán^{*,1} y Francisco Ortiz Arango^{†,2}

^{*}InvVal Capital, [†]Facultad de Ciencias Económicas y Empresariales de la Universidad Panamericana

RESUMEN Este artículo hace hincapié en las condiciones que los empresarios deben tener en cuenta cuando desean que sus empresas incursionen en nuevos proyectos. Ejemplos de esto son: tener visión de alcance global, considerar las fuerzas dinámicas de su mercado, las nuevas tendencias, así como su entorno económico, entre otras.

Palabras clave

Planeación financiera, Vida de la empresa, Análisis Empresarial.

1. INTRODUCCIÓN

En el mundo actual se requiere que las empresas, cuando tienen la intención de llevar a cabo nuevos proyectos, además de tener una visión local y aislada de las características de su producto, servicio o manejo adecuado de su operación, también tengan una visión de alcance

global que tome en cuenta prioritariamente las necesidades de sus clientes, las fuerzas dinámicas de su mercado, las nuevas tendencias y su entorno económico, solo por mencionar algunos de los factores que rigen, modifican y evolucionan constantemente dentro de cualquier sector.

Según el Censo Económico publicado por el Instituto Nacional de Estadística y Geografía (INEGI), llevado a cabo en 2019, hoy día existen en México alrededor de cinco millones de empresas, de las cuales más del 99.5 % son pequeñas y medianas, las cuales concentran más de las tres cuartas partes del personal empleado, aportando un monto superior al 40 % del Producto Interno Bruto. Estas cifras determinan un factor crucial para la permanencia operativa de cada empresa, y que le permita alcanzar un mayor equilibrio y estabilidad económica; sin embargo, de acuerdo al INEGI, en sentido inverso se encuentra el hecho de que la esperanza de vida promedio de cualquier empresa en México es menor a ocho años (datos

Derechos reservados © 2024 por el Colegio Nacional de Actuarios

Última actualización del manuscrito: 5 de enero de 2024

Este artículo es publicado por el Colegio Nacional de Actuarios (CONAC) y contiene información de una variedad de fuentes. Es un trabajo cuyo fin es únicamente informativo y no debe interpretarse como asesoramiento profesional o financiero. La visión y comentarios contenidos en este documento corresponden exclusivamente a su autor y, por tanto, son ajenos a cualquier entidad pública o privada, incluyendo aquella para la que actualmente colabora. El CONAC no recomienda ni respalda el uso de la información proporcionada en este estudio. El CONAC no ofrece ninguna garantía, expresa o implícita, ni representación de ningún tipo y no asume ninguna responsabilidad en relación con el uso o mal uso de este trabajo.

¹ Correo electrónico: direccion@invvalcapital.com



previos al COVID-19), lo cual representa alrededor de una tercera parte de lo que se considera la duración de una generación humana.

El tener un proyecto o empresa con una sostenibilidad de largo plazo, permite dar sustento y solidez a las familias que participan en ella, por lo cual es fundamental que una empresa –independientemente del sector, su tamaño o en la etapa de maduración en la que se encuentre– considere como parte de sus prioridades el definir una estrategia con visión duradera, aun cuando maneje su operación de manera impecable o que conozca profundamente a sus clientes. Las herramientas que se establezcan, el monitoreo que se implemente y la proyección que se determine sobre el futuro que soporten a dicha visión duradera, deberá principalmente ser sustentada y confrontada mediante un plan de negocios y una planeación financiera de corto, mediano y largo plazos, pero la pregunta que surge en muchos de los casos es ¿por dónde comenzar?

2. UN PLAN DE NEGOCIO

Un plan de negocio debe ser soportado por múltiples herramientas teóricas que pueden ser adquiridas en las escuelas a través de diplomados, seminarios o programas formales de posgrado, o bien mediante lecturas que podamos encontrar al respecto. Sin embargo tener este conocimiento es solo el comienzo, ya que uno de los grandes retos de las empresas es aplicar ese conocimiento, enfocándolo al mundo práctico y real, observando y analizando al interior de la empresa, sin descuidar su entorno y las dinámicas que mueven al nicho específico en donde la empresa o el proyecto particular se desarrolla. Cada empresa debe entender profundamente su propia esencia y la de su mercado, adecuando los elementos del plan de negocios a sus propias condiciones, filosofía, misión, visión y naturalmente a las condiciones de su mercado. Esto hará que cada visión de negocio y por consiguiente de plan de negocios se convertirá en única, como un traje hecho a la medida de cada circunstancia.

Asimismo, el definir una estrategia adecuada permitirá a cada empresa crear su rumbo de manera realista acorde con su plan de negocios, soportada por proyecciones financieras de un plazo inmediato a varios años. Dichas proyecciones financieras deberán contener una serie de análisis sobre las premisas macroeconómicas del entorno global, local y propio del sector, así como las previsiones microeconómicas propias del mercado y de las variables materiales de la empresa o el proyecto. Lograr lo anterior implica considerar más allá de las circunstancias

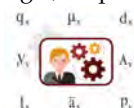
propias de la empresa y sus expectativas de crecimiento, se requiere llevar a cabo un análisis estratégico, de negocios, comercial y financiero a fondo; para lo cual es muy útil tomar en cuenta las siguientes consideraciones o Componentes Fundamentales de Análisis Empresarial, que se ilustran en la Figura 1.



Figura 1 Componentes fundamentales del análisis empresarial.

- Esencia de la empresa, visión, misión y valores
- Estructura organizacional y definición de un gobierno corporativo con contrapesos
- Valor agregado del producto o servicios que se ofrece, con respecto a los competidores del sector
- Dinámicas del mercado y competencia, así como la importancia de conocer dicho mercado y adecuar el producto/servicio
- Plan comercial enfocado en el diseño del producto y su posicionamiento
- Tamaño de la empresa y participación de mercado
- Conocimiento de los clientes actuales y prospectos, así como la comunicación del producto y un comparativo contra la competencia
- Aprovechamiento de las fortalezas y oportunidades, así como la importancia de mitigar o eliminar las debilidades y amenazas
- Los nichos de mercado existentes o nuevos, así como la búsqueda de una oferta disruptiva
- Cuestionar las barreras de entrada en el mercado y las fuerzas que mueven dicho mercado.
- Alguna otra consideración particular relevante de la empresa o proyecto en específico

Todos los elementos mencionados en la Figura 1, independientemente de que su naturaleza sea cuantitativa o cualitativa deberán confluir en una estrategia, un plan



de negocios y un modelo de proyecciones financieras. En particular los componentes cualitativos deberán cruzar, en la medida de lo posible, por el tamiz de los números, asignarle un valor y vislumbrar su viabilidad. Esta unión de los elementos cualitativo y cuantitativo de una manera objetiva, permitirá potenciar y apoyar la toma de mejores decisiones.

La capacidad de analizar, alinear, sintetizar y eventualmente presentar todas estas ideas con personas externas, como pueden ser prospectos de inversionistas, aliados, bancos, proveedores o alguna entidad con propósito específico, nos reta, no solo a ser precisos en nuestra comunicación para la ejecución del plan de negocios por parte del cuerpo directivo y sus equipos, sino, también a convencer a los externos con los que se busque algún tipo de acuerdo. La precisión y claridad en la elaboración y comunicación de un proyecto, soportada por un plan de negocios óptimamente fundamentado en términos financieros y estratégicos, será una pieza fundamental para el logro de los objetivos que se pretendan alcanzar.

Esperamos haber infundido en el lector la importancia de elaborar un plan de negocio adecuado y específico a las necesidades, capacidades, visión, valores y nicho de mercado de la empresa en cuestión. En futuras entregas analizaremos con mayor detalle y ejemplos prácticos cada uno de los componentes fundamentales del análisis empresarial mencionado en este artículo.



InvVal Capital y también se dedica a ser Consejero Empresarial.

Tomás Cortés Zulbarán se graduó como Ingeniero Industrial con especialidad en Finanzas por la Universidad Panamericana, y tiene una Maestría en Administración y Negocios por la Universidad de Northeastern of Boston. Es Socio Director de la División Financiera y de Inversiones en



Francisco Ortiz Arango se doctoró en Ciencias Financieras por el Instituto Tecnológico de Estudios Superiores de Monterrey. Se desempeña como Profesor-Investigador de la Facultad de Ciencias Económicas y Empresariales de la Universidad Panamericana. Es miembro del Sistema Nacional de Investigadores.



Maestría en Riesgo

Proceso de admisión 2024 abierto

Modalidad

Presencial

Duración

Ocho trimestres

Horario

Entre lunes y jueves
de 19:00 a 22:00 hrs.

ACTUARIOS TRABAJANDO

On a predictive model of the box office revenue

Gerardo Ortiz Arellano, Manuel Jiménez Lizárraga and
Viacheslav Kalashnykov



On a predictive model of the box office revenue

Gerardo Ortiz-Arellano^{*,1}, Manuel Jiménez-Lizárraga^{†,2}, and Viacheslav Kalashnykov^{†,3}

^{*}Nielsen, [†]Universidad Autónoma de Nuevo León

ABSTRACT The purpose of this project is to develop a model to predict the box office performance of a theatrical movie. Since it is very risky to invest in this area of entertainment, this tool can help determine how much to invest to make a profit. Also, due to the pandemic we experienced in 2019-2022, the means of distribution are expected to change, so our model determines the best direction of distribution. We develop a model that predicts first weekend revenues using a ranking model called “hard voting”. Our main tools are decision trees, Light Gradient-Boosting Machine (LightGBM), and k -nearest neighbors classification algorithms. Our model yields a balanced accuracy of 84.62%.

Key words

Decision trees, Light Gradient-Boosting Machine,
 k -nearest neighbors classification algorithms

1. INTRODUCTION

The multimedia entertainment business is one of the largest in the world today, and cinema is one of its major

contributors. Even with the size of the industry, investing in the production of motion pictures is very risky, this is since the financial success of a film is very uncertain. Indeed, only between 30 and 40 percent of the films manage to recover their investment, and only one in ten makes a profit (see Ghiassi *et al.* (2015)).

For these reasons, being able to predict box office performance has been a challenge, especially because of the uncertainty of the audiences’ behavior; and also because the biggest movies nowadays are part of existing franchises. An example of the latter is that, in 2021, the four highest-grossing films (in the United States of America and Canada) belong to the Marvel franchise, while the highest-ranked movie not belonging to an established

All rights reserved © 2023 by the Mexican Colegio Nacional de Actuarios
Last update of the manuscript: Friday 15th December, 2023
This article is published by the Mexican Colegio Nacional de Actuarios (CONAC) and contains information from a variety of sources. It is a work intended for informational purposes only and should not be construed as professional or financial advice. The vision and comments contained in this document correspond exclusively to its author and, therefore, are unrelated to any public or private entity, including the one for which he currently collaborates. CONAC does not recommend or endorse the use of the information provided in this study. CONAC makes no warranty, express or implied, or representation of any kind and assumes no liability in connection with the use or misuse of this work.

¹E-mail of the corresponding author: gera.o.arellano@gmail.com



franchise obtained only the eleventh place in the classification (see Table 1).

The cinematic industry was directly affected by the COVID-19 crisis, forcing studios to change the original release dates of their films or change the means of distribution. On the other hand, this exponentially boosted the growth of streaming platforms, which could be considered a direct competitor to movie theaters. This caused distribution strategies to change from day to day, shortening the theatrical exclusivity window of several films or even canceling their theatrical distribution altogether. An example of these new means of distribution was “Warner Brothers”, which decided to release all its films in theaters at the same time as on its streaming platform “HBO Max” in 2021.

All the codes are available in <https://github.com/DonDisparates/cinema.git> for replication and study purposes.

2. THEORETICAL FRAMEWORK AND METHODOLOGY

We carried out the practical part of the project in three main parts. First, we harvested our information from [The Numbers](#), [The Internet Movie Database](#) and [Rotten Tomatoes](#) websites. We focused on the financial information of the films (such as recollection and box office of the first weekend), as well as their genre, distributor, and source.

Once the database had been collected, we carried out an analysis of the data. According to this analysis, we added some variables that were not necessarily collected at the beginning but were implicit in the data. An example of this is the season premiere. This variable is important since, as can be seen in Figure 1, there is an increase in the recollection variable during the summer (from week 18) and also a drop in the fall (week 36).

We also adjusted all the monetary data for inflation. These include the recollection of the first weekend, the total recollection, and the budget. To this end, we used the average price of a movie ticket each year from 1999 to August 15, 2022, published by the Motion Picture Association (MPAA) in its annual report and shown in Table 2. Because the annual report is published in March of the following year, there is no average for 2022, so the same price is assumed as in 2021.

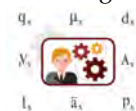
Finally, we proceeded to implement the model to obtain the prediction of the film’s performance at the box office during the first weekend. We also attempted to match the performance in the remainder of the season of



Figure 1 Seasonal revenue from movies at the box office (see [Hennig-Thurau y Houston \(2018\)](#)).

its projection, but unfortunately, the results we obtained were not as accurate as those for the opening weekend. However, it is widely known that the largest amount of money raised in a movie happens during its first weekend. Moreover, since it is then that the film receives reactions from both critics and the general public, it is very difficult to forecast the performance of a film.

The “opening weekend” phenomenon has been studied and described with an exponentially decreasing pattern, in which product revenue peak immediately after launch and then decline at a constant rate over time (so that the decline becomes less pronounced as time goes by). This is a characteristic feature of most entertainment products. In the case of cinema, this is measured with the first weekend of release because, most commonly, a film is released on Friday and, as can be seen in Figure



Position	Title	Franchise	Box Office	Market share
1	Spider-Man: No Way Home	Marvel	\$781,623,905	24.4%
2	Shang-Chi and the Legend...	Marvel	\$224,543,292	7.0%
3	Venom: Let There be Carnage	Marvel	\$213,550,366	6.7%
4	Black Widow	Marvel	\$183,651,655	5.7%
5	F9: The Fast Saga	F&F	\$173,005,945	5.4%
6	Eternals	Marvel	\$164,870,264	5.2%
7	No Time to Die	James Bond	\$160,891,007	5.0%
8	A Quiet Place: Part II	A Quiet Place	\$160,215,764	5.0%
9	Sing 2	Sing	\$151,904,445	4.7%
10	Ghostbusters: Afterlife	Ghostbusters	\$128,885,266	4.0%
11	Free Guy	Original	\$121,626,598	3.8%
TOTAL			\$2,464,768,507	77.0%

■ **Table 1** Domestic Box Office (2021). Data obtained from [The Numbers](#) as of March 3, 2022.

Year	Average Price	Year	Average Price
2021	\$9.57	2009	\$7.50
2020	\$9.37	2008	\$7.18
2019	\$9.16	2007	\$6.88
2018	\$9.11	2006	\$7.55
2017	\$8.97	2005	\$6.41
2016	\$8.65	2004	\$6.21
2015	\$8.43	2003	\$6.03
2014	\$8.17	2002	\$5.81
2013	\$8.13	2001	\$5.66
2012	\$7.96	2000	\$5.39
2011	\$7.93	1999	\$5.08
2010	\$7.89		
TOTAL		\$2,464,768,507	77.0%

■ **Table 2** Average price of a ticket at the box office. Data obtained from [The Numbers.com](#) and [Statista](#) and provided by the MPAA as of September 14 of 2022.



2, it is the weekend (Friday, Saturday and Sunday) of the highest recollections at the box office (see [Hennig-Thurau y Houston \(2018\)](#)).

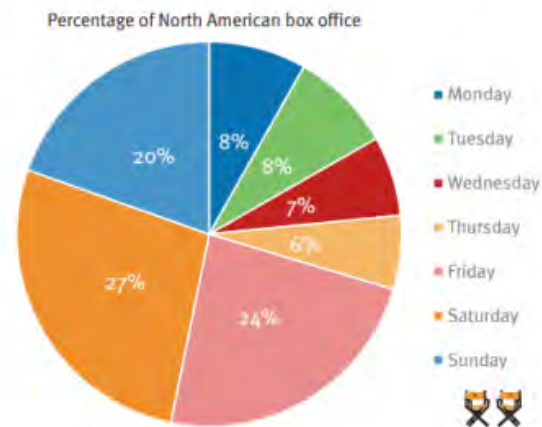


Figure 2 Average box office receipts by day of the week. See [Hennig-Thurau y Houston \(2018\)](#).

3. HARD VOTING

In this model, we used the “hard voting” assembly. This is an assembly method for multiclass predictive models where the outcome is the option with the highest number of votes. Figure 3 shows the model. Level (A) represents the tickets for a specific instance of a movie: budget, genre, etc. Level (B) represents a subset of optimized models. Level (C) represents the prediction given the input from each of those models. These predictions are added to find the mode. Level (D) is the final level, exemplifying the majority vote of the selected classifiers (see Phelps (2020)).

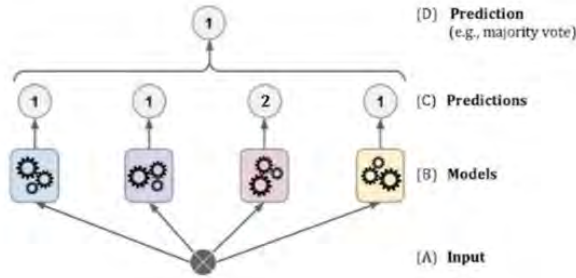


Figure 3 Hard voting classification tree (see Phelps (2020)).

For the assembly, three multiple classification algorithms will be used: classification trees, LightGBM, and k -nearest neighbors.

Classification Trees

Classification tree algorithms are recursive partitioning methods that use the statistical learning technique of decision trees. *Grosso modo*, the algorithm consists of first quantifying the optimization criterion for the data partition for each variable, and then selecting the best partition. The process is repeated until a Split that belongs to the same value as the target variable is reached (see Zelterman (2015)). Algorithm 1, which we borrowed from Software testing help, delves into the details of these ideas.

Algorithm 1 Classification tree

Algorithm 1: Classification tree

Data: Data partition D , $attributes_list$, $attributes_selection_method$ (i.e. $split_attribute$, and $split_point$ or $split_subset$)

Result: Decision tree

```

1 Create a node  $N$ ;
2 if all tuples in  $D$  are of class  $C$  then
3   return  $N$  as a leaf node labeled as class  $C$ 
4 end
5 if list_of_attributes =  $\emptyset$  then
6   return  $N$  as leaf node with majority class  $D$ : majority voting
7 end
8 Apply  $attributes\_selection\_method$  to  $D$  and  $attributes\_list$  to find the best split criteria;
9 Label node  $N$  with split criterion;
10 if  $split\_attribute$  is of discrete-type and Multidirectional splits are allowed > Non-binary trees are allowed
11 then
12    $attributes\_list \leftarrow attributes\_list \setminus split\_attribute$ : Eliminate  $split\_attribute$ 
13 end
14 for  $j \in split\_attribute$  do
15    $D_j \leftarrow$  subset of  $D$  that meets  $j$ ;
16   if  $D_j = \emptyset$  then
17     Attach labeled leaf with majority class in  $D$  to node  $N$ 
18   else
19     Attach decision tree for  $D_j$ ,  $attributes\_list$  to node  $N$ 
20   end
21 end
22 return  $N$ ;

```

Light Gradient Boosting Machine

The main advantage of LightGBM is the speed at which it is trained, which is where the “light” prefix comes from. It differs from other tree-based algorithms in that it grows its tree vertically, while other algorithms grow trees horizontally (see Phelps (2020)). Algorithm 2 presents the details (see Dunbray *et al.* (2021)).

Algorithm 2 Classification tree.

Algorithm 2: Classification tree

Data: Training data set $D = \{(x_1, y_1), \dots, (x_N, y_N)\} \subset \mathbb{R} \times \{-1, 1\}$, loss function $L(y, \theta(x))$, M , a -Number of sampled data of large gradient, b -Number of sampled data of small gradient.

Result: Minimal loss function θ_M

```

1 Combine mutually exclusive features of  $x_1, \dots, x_N$  by means of the exclusive features bundling technique;
2  $\theta_0(x) \leftarrow \arg \min_c \sum_{i=1}^N L(y_i, c)$ ;
3 for  $m = 1$  to  $M$  do
4    $r_i \leftarrow \left| \frac{\partial L(y_i, \theta(x_i))}{\partial \theta(x_i)} \right|$  for  $i = 1, \dots, N$ ;
5   Re-sample dataset with the gradient-based one side sampling method:
6    $topN \leftarrow a|D|$ ;
7    $randN \leftarrow b|D|$ ;
8    $\ast |\cdot|$  stands for the size of the set
9   sort indices in  $r$ ;
10   $A \leftarrow$  sort indices from 1 to  $N$ ;
11   $B \leftarrow$  random index from the sorted set  $topN, \dots, |D|$ ;
12   $D' \leftarrow A \cup B$ ;
13   $V_j(d) \leftarrow \frac{1}{2} \left( \sum_{x_i \in A, r_i} r_i + \frac{1}{d} \sum_{x_i \in B, r_i} r_i \right)^2 \left( \frac{1}{n_j(d)} + \frac{1}{n_j(d)} \right)$ ;
14   $\ast V$  represents an information gain
15  Build a new decision tree  $\theta_m(x)$  over the set  $D'$ ;
16   $\theta_m(x) \leftarrow \theta_{m-1}(x) + \theta_m(x)$ ;
17 end
18 return  $\theta'(x) \leftarrow \theta_M(x)$ ;

```



***k*-nearest neighbors**

The *k*-nearest neighbors' algorithm is a nonparametric supervised learning classifier that uses proximity to make classifications or predictions about the clustering of an individual data point. It is normally used as a classification algorithm, assuming that similar points can be found close to each other. In classification problems, the class label is assigned on the basis of a majority vote, that is, the label that occurs most frequently around a given data point is used (see [Dasarathy \(1991\)](#)). Algorithm 3 presents the details.

Algorithm 3 *k*-nearest neighbors

Algorithm 3: *k*-nearest neighbors

Data: D -Matrix of distances of dimension $n \times n$ and s -index of point of departure

Result: route list of vertices with the path of minimal length

```
1 for  $i = 1$  to  $n$  do
2    $visited(i) \leftarrow 0$ 
3 end
4 Initialize the route list with  $s$ ;
5  $visited(s) \leftarrow 1$ ;
6  $current \leftarrow s$ ;
7 for  $i = 2$  to  $n$  do
8   Find smallest index
9 end
10 return  $\theta'(x) \leftarrow \theta_M(x)$ 
```

4. WORK DEVELOPMENT

Database Acquisition

For data collection, we consulted the platform [The Numbers](#) platform, where the data of the 100 films with the highest domestic grossing for each year from 1999 to July 2022 were captured.

This database was used as a first approach and as a basic dataset for obtaining additional data for each film. These attributes were obtained through a program written in Python using the "selenium" library. This method was used on [Rotten Tomatoes](#) and [the Internet Movie Data Base](#).

Next, we show the attributes obtained in the three consulted platforms:

- Name: The title of the film, this was not used in the models.
- Release date: Data in a date format that indicates the first day the film was available to the public, from which the season variable emerges.
- Year of release: Year in which the film was released, used for inflation adjustment in monetary variables.



- Distributor: It oversees promoting and distributing the films, it can be the same production company.
- Genre: A way of categorizing films and a variable commonly used to predict the performance of a film. Within this are genres such as adventure, musical, action, drama, comedy, etc.
- Source: Material on which the film is based, be it a book, historical, or original idea.
- Production method: Indicates whether the film is animated or not.
- Creative Type: Another way of categorizing movies based more on *the universe* where the movie takes place as fantasy, science fiction, etc.
- Budget: All expenses used to produce the film.
- Cinemas on the first weekend: Number of establishments where the film was shown on the first weekend of release.
- Maximum number of cinemas: Maximum number of establishments where the film was shown on a certain date.
- First weekend income: Box office receipts for the film the first weekend it was shown in the United States and Canada.
- Domestic income: Box office receipts of the film in the United States and Canada.
- IMDB Rating: Rating from 1 to 10 given by the users of the [the Internet Movie Data Base](#).
- Metascore: Rating from 1 to 100 given by a weighted average of the rating given by critics published on [Metacritic](#).
- MPAA: Movie rating system in the United States published by the MPAA, these can be G, PG, PG-13, R and NC-17.
- Duration: The time that the film lasts in hours and minutes.
- Tomatometer: Aggregate rating of a movie given by certified critics on the [Rotten Tomatoes](#) page, it takes values from 0 to 100.
- Audience Score: Aggregate rating of a movie given by the user of the [Rotten Tomatoes](#) page; this takes values from 0 to 100.

Database settings

Adjustments are made to the same database by first removing films with incomplete data such as budget or number of theaters in which they are shown. In addition to adding variables obtained through an analysis of the database and adjusting the monetary data for inflation.

Binary variables refer to categorical variables, where a 1 is identified if this movie belongs to the indicated

category.

The variables indicated as *average* are obtained with the data from the same base obtained.

The variables that will be handled in the initial classification model with the recollection response variable on the first weekend are shown below:

- Post-COVID (binary): Identify if the film was released after March 31, 2021.
- During COVID (binary): Identify if the film was released between March 13, 2020, and March 31, 2021.
- Release season: Identifies in which season the film was released, this can be winter (January and February), spring (March and April), summer (from May to the end of August), autumn (from September to the 17th weekend before Thanksgiving) and Holidays (from Thanksgiving to New Years).
- Distributor Average (numerical): Average recollection of films distributed by the distribution company.
- Genre: Indicates the genre belonging to the film, this can be action, adventure, black comedy, comedy, documentary, drama, horror, musical, romantic comedy, thriller/suspense and "western".
- Source: Indicates the origin of the film's story, this can be comic strip/graphic novel, fact book/article, folk tale/legend/fairy tale, video game, play, events from real life, television, original script, "remake" and "spin-off".
- Production method: Indicates whether the film is animated, "live action" or a combination of both.
- Creative type: Categorizes the films more focused on its environment. They can be contemporary fiction, dramatization, factual events, historical fiction, children's fiction, science fiction, and superheroes.
- MPAA: Identifies the classification of films given by the "Motion Picture American Association", these include G, PG, PG-13, and R.
- Duration: Time that the film lasts in minutes.
- Franchise (binary variable): Identifies whether the film belongs to a franchise or not.
- Franchise average (numerical): Average of the recollection of the films within this franchise, in case of not belonging to a franchise it takes the value of 0.
- Average director (numerical): Average revenue from films directed by the same director.
- Average Actor 1 (numerical): Average of the recollection of the films where the main actor appears.
- Average Actor 2 (numerical): Average of the recollection of the films where the second most important actor appears.

- Competition (numerical): How many films had their release date 14 days or less from the film's release date.
- Budget (numerical): How much money it cost to make the film adjusted for inflation.
- Cinemas first weekend (numerical): Number of cinemas where the film was shown on its first weekend of release.
- Limited release (binary): Identifies if the film opened its first weekend in less than 1000 theaters.

For the regression model, where the response variable is domestic revenue, all the variables used in the previous model and the following are used:

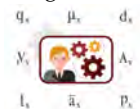
- Maximum number of theaters (numerical): Maximum number of theaters where the film was shown in a weekend.
- Income first weekend (numerical): Recollection adjusted for inflation during its first weekend.
- IMDB Rating (numeric): Rating from 0 to 10 of a movie by IMDB users.
- Metascore (numeric): Rating from 0 to 100 of a film by certified critics.
- Tomatometer (numeric): Rating of a movie from 0 to 100 by certified critics on Rotten Tomatoes.
- Audience Score (numeric): Rating from 0 to 100 of a film by users of Rotten Tomatoes

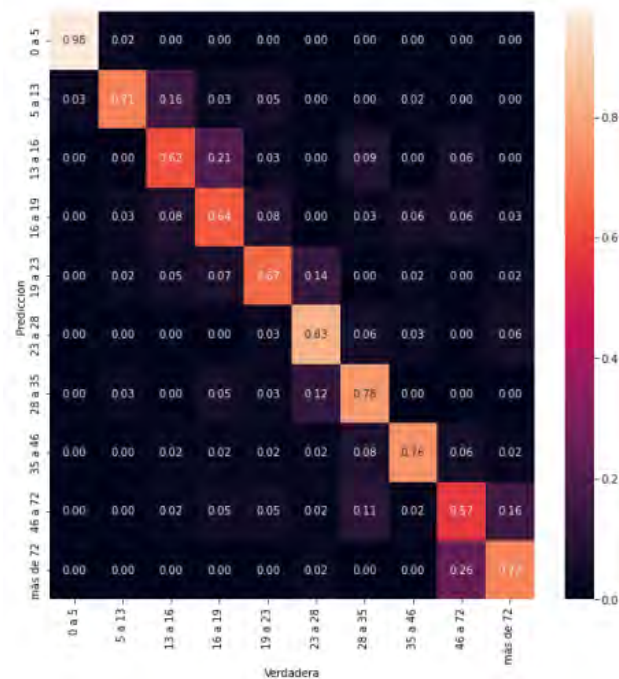
These variables are not the final ones because some may be eliminated due to a high correlation between them.

5. RESULTS

Table 3 shows a normalized confusion matrix for the model. The confusion matrix, also known as *error matrix*, is a table that allows visualizing the performance of a classifier. It shows the ways in which the classifier gets confused when making predictions. The columns of this confusion matrix represent the predicted classes, while the rows represent the actual classes. The intersection point of these cells represents the correct classification of the data. A perfect classifier has 1.0 in each row that matches the true label and the predicted label.

Table 4 shows the performance of each methodology used in the first model including the Hard-Voting assembly showing the balanced precision of each one. Balanced precision prevents inflated performance estimates on unbalanced data sets. It is defined as a metric that measures what fraction of the class in question has been correctly predicted. It is calculated as the sum of true positives divided by the sum of true positives and false negatives of





■ **Table 3** Confusion matrix.

a class. Algebraically, the mean sensitivity can be formulated as $\left(\frac{1}{M}\right) \sum_{m=1}^M \frac{r_m}{n_m}$ where M is the number of classes, n_m is the size of the 21 data in class m , r_m is the number of accurately predicted data items that belong to class m (see [Luque et al. \(2019\)](#)).

Figure 4 shows us which variables are more important in model 1, with the highest being the distributor's average followed by the number of movie theaters on its first weekend. It can be observed that the variables with more weight are the numerical ones, while the categorical ones are losing importance. Due to the number of variables (52) the graph only shows the first 30 variables.

6. CONCLUSIONS

Movie production requires a great deal of time and money. Most of these films generate income that is less than what it costs to produce them. This especially affects 24 small distributors and producers where today's competitive film market due to the production of franchised films and competition with streaming platforms in the world has made it more difficult for them to recover the



■ **Table 4** Performance of the model

Classifier	Balanced Precision
k -Nearest Neighbors ($k = 45$)	59.56%
Decision tree	62.92%
LightGBM	82.99%
Hard-Voting	84.62%

investment.

The models ended up giving satisfactory results while there is still room for improvement. In the first model, improvement can be sought by testing different methodologies in the assembly to increase its precision. In the same way, the model could improve if there were a more extensive database, on the one hand, with the number of observations to improve the precision of films with poor performance, as well as variables more focused on distribution and advertising methods.

REFERENCES

- Dasarathy, B. V., 1991 *Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques*. IEEE Computer Society Press.
- Dunbray, N., R. Rane, S. Nimje, J. Katade, y S. Mavale, 2021 A novel prediction model for diabetes detection using gridsearch and a voting classifier between lightgbm and knn. pp. 1–7.
- Ghiassi, M., D. Lio, y B. Moon, 2015 Pre-production forecasting of movie revenues with a dynamic artificial neural network. *Expert Systems with Applications* **42**: 3176–3193.
- Hennig-Thurau, T. y M. Houston, 2018 *Entertainment Science: Data Analytics and Practical Theory for Movies, Games, Books, and Music*.
- Luque, A., A. Carrasco, A. Martín, y A. de las Heras, 2019 The impact of class imbalance in classification performance metrics based on the binary confusion matrix. *Pattern Recognition* **91**: 216–231.
- Phelps, B., 2020 *Predicting Motion Picture Box-Office Performance*. Southern Connecticut State University, CT.
- Zelterman, D., 2015 *Applied Multivariate Statistics with R*. Springer Cham.

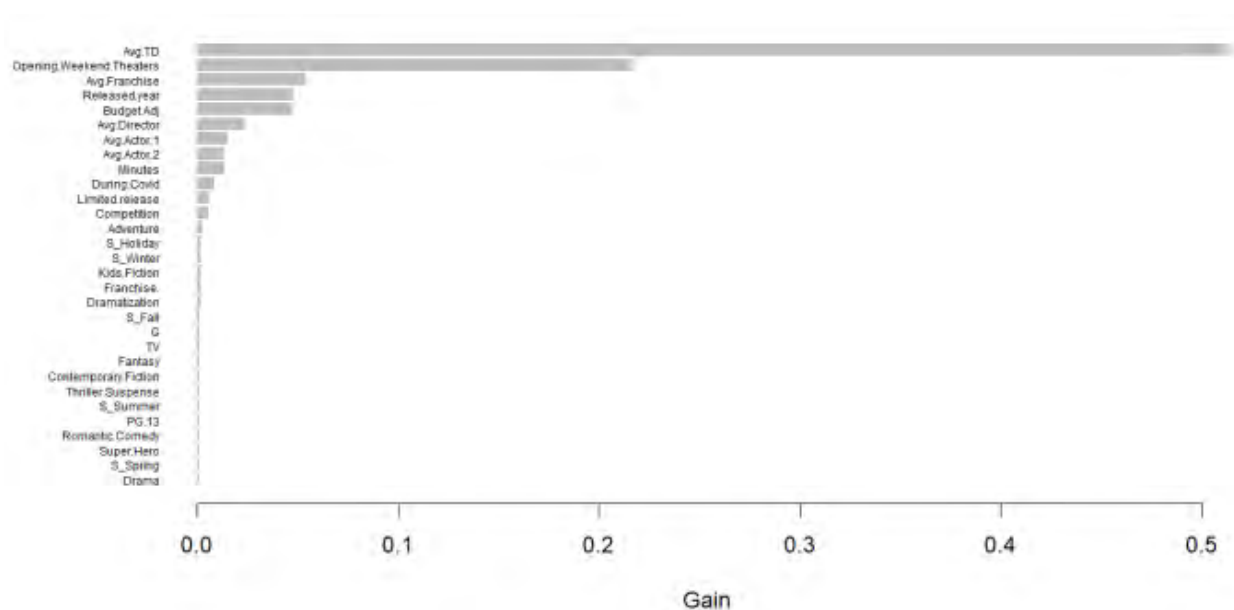


Figure 4 Importance of variables.



2022, he has been a data analyst at Nielsen for its Marketing Mix Model area in the United States.

Gerardo Ortiz Arellano has a degree in Actuarial Science from Universidad Anáhuac México, where he obtained the Academic Excellence Award. He was a speaker at the 2023 edition of the Canwest Research Conference in Madrid, Spain. Since



Ohio State University, USA from 2008-2009. He is currently with the Faculty of Physical and Mathematical Sciences of Autonomous University of Nuevo Leon, Mexico. His research interests include differential games, ro-

Manuel Jimenez-Lizarraga received the BSc degree in electrical engineering from Culiacan Institute of Technology, Sinaloa Mexico, and the MS and PhD degrees in Automatic Control from CINVESTAV-IPN Mexico in 1996, 2000 and 2006 respectively. He was postdoc fellow at the ECE Department of the

bust, optimal and sliding mode control and applications.



Viacheslav Kalashnykov has a Doctoral degree in Operations Research. He is the author of more than 700 published research papers and five monographs. Currently, he is a professor-researcher at the Universidad Autónoma de Nuevo León in Monterrey, Mexico.



ACTUARIOS TRABAJANDO

Análisis del Riesgo país de México a partir de variables económicas

Ana María Hernández Piña y José Eduardo Becerra
López



Análisis del Riesgo país de México a partir de variables económicas

Ana María Hernández Piña^{*,1} y José Eduardo Becerra López^{†,2}

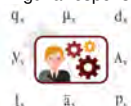
*PepsiCo, †Allianz México

RESUMEN Hoy en día, la incertidumbre en los mercados es muy común y la medición de los riesgos en cualquier ámbito se ha vuelto de suma importancia. El riesgo país se ha vuelto un indicador fundamental para monitorear la estabilidad de los países y su desempeño económico, social o política. Por ello, el análisis de este indicador es clave para encontrar áreas de mejora y evitar o minimizar, hasta cierto punto, posibles consecuencias por un mal desempeño. Mediante el uso de distintos tipos de indicadores se realizó un modelo para determinar el riesgo país a partir de los bonos gubernamentales y una predicción del comportamiento de este riesgo para México este 2023.

Palabras clave

Global equity, Método LASSO, Decision Tree, Random Forest

Derechos reservados © 2023 por el Colegio Nacional de Actuarios
Última actualización del manuscrito: 14 de diciembre de 2023
Versión preparada por Andrés Antonio Mantecón Araujo.
Este artículo es publicado por el Colegio Nacional de Actuarios (CONAC) y contiene información de una variedad de fuentes. Es un trabajo cuyo fin es únicamente informativo y no debe interpretarse como asesoramiento profesional o financiero. La visión y comentarios contenidos en este documento corresponden exclusivamente a su autor y, por tanto, son ajenos a cualquier entidad pública o privada, incluyendo aquella para la que actualmente colabora. El CONAC no recomienda ni respalda el uso de la información proporcionada en este estudio. El CONAC no ofrece ninguna garantía, expresa o implícita, ni representación de ningún tipo y no asume ninguna responsabilidad en relación con el uso o mal uso de este trabajo.



1. INTRODUCCIÓN

Este proyecto puede tener distintas vertientes y enfoques ya que es un tema bastante amplio. Algunas de ellas serían analizar el por qué países con alto riesgo, tienen ese comportamiento y si siguen alguna tendencia desde el punto de vista económico y político; identificar las posibles medidas que estos países podrían implementar para prevenir o mitigar las consecuencias de tener un alto riesgo; enfocarse en las consecuencias de que esto ocurra tanto a nivel nacional como internacional, y cómo perjudicaría tanto a la nación, como a los inversionistas.

¹Correo electrónico: ana.hernandezp0@gmail.com

²Correo electrónico: jose.becerra@anahuac.mx

Entendiendo los puntos de vista que se podrían tomar, el objetivo principal de este trabajo es poder calcular qué tan probable es que México eleve el riesgo país, a tal grado que la economía del país se vea gravemente afectada. Asimismo, se busca que, con base en información relevante sobre las condiciones y circunstancias de cada país, cálculos estadísticos y probabilísticos, podamos identificar las variables que afectan más al comportamiento de estos, así como comprender la situación actual y los factores causales que pudieran ocasionar un incremento en el riesgo. Finalmente, se analizarán diversos factores y variables económicas, políticas y sociales que tienen mayor relevancia en la estabilidad y bienestar de los países, para poder realizar un modelo estadístico mediante el uso de inteligencia artificial que nos permita detectar las variables que afectan más al riesgo de los países, para así poder predecir el riesgo país de México en el 2023.

Según [Gonzalez \(2020\)](#), “a pesar de atravesar una de las recesiones más grandes en décadas y haber sufrido una baja en su nota crediticia este año, México ha tomado ventaja frente a otros países emergentes en recuperar la confianza de los inversionistas, según el desempeño de su moneda y de instrumentos que miden riesgo soberano.” Es bien sabido que México ha enfrentado en su historia varias crisis económicas y problemas políticos, sanitarios y sociales que han afectado la estabilidad económica del país. Un claro ejemplo es la crisis económica de 1994 o Crisis del Tequila, la crisis hipotecaria del 2008, o la crisis sanitaria de COVID-19. Esto ha ocasionado problemas de financiamiento, declive en la inversión extranjera, y estancamiento del crecimiento económico. Esto lo limita en su actuar, provocando que no tenga acceso a mercados de deuda, generando desconfianza entre los inversionistas, aumentando la emisión monetaria, la cual impacta en la inflación y el estilo de vida de sus ciudadanos, devaluando la moneda y el poder adquisitivo del país, por mencionar algunos. Las crisis conllevan a la inestabilidad política, económica y social. La inestabilidad provoca problemas tanto internos como externos, y, por ende, una crisis mal manejada o imprevista implica que un país pueda incrementar de manera considerable su riesgo país. Por estas razones, es importante analizar el comportamiento de México ante el evento que sacudió la estabilidad de los países como fue la crisis de COVID-19.

Es posible tomar medidas para tener un control del riesgo país y disminuirlo. Existen distintas variables macroeconómicas utilizadas para conocer la posición en la que se encuentra un país. El objetivo fundamental de este trabajo es analizar las variables de mayor importancia para medir y predecir el riesgo país en México mediante

un modelo estadístico basado en el análisis y comparación de aportaciones, así como analizar las consecuencias que se podrían presentar tanto interna como externamente. El documento se estructura en hipótesis, desarrollo y conclusión. Entre nuestras hipótesis plantearemos de lo que se tratará el trabajo, así como la interrogante del riesgo de México. Continuando con el desarrollo, este se subdivide en cinco secciones, donde se analizará el modelo predictivo a utilizar, las variables predictoras y de respuesta, y finalmente se hará la predicción junto con su análisis.

2. HIPÓTESIS

Dentro de este trabajo se determinará si el riesgo de México como país, está en peligro o no, considerando las diez variables macroeconómicas más relevantes que afectan el desempeño de este.

3. DESARROLLO

El riesgo país puede tener diferentes interpretaciones. Para fines de este trabajo, usamos la definición obtenida de [Silvia Iranzo \(2008\)](#):

“El riesgo-país es la probabilidad de que se produzca una pérdida financiera por circunstancias macroeconómicas, políticas o sociales, o por desastres naturales, en un país determinado. El riesgo-país comprende el riesgo de impago de la deuda externa soberana (riesgo soberano), y de la deuda externa privada cuando el riesgo de crédito se debe a circunstancias ajenas a la situación de solvencia o liquidez del deudor privado.”

De esta manera, se considera el riesgo país como la probabilidad de ocurrencia de algún suceso que ocasione un desequilibrio en la economía del país y que está formado por distintos riesgos a los que puede estar expuesto un país. Ahora bien, para determinar el riesgo país de México es necesario construir un modelo de predicción que se irá presentando a lo largo de este documento.

Modelo Predictivo

Comenzaremos dando una breve explicación de las utilidades y la función del modelo que nos ayudará a medir el riesgo dentro del proyecto. A grandes rasgos, un modelo predictivo es un modelo basado en estadística, donde su principal función es predecir la respuesta a cierta interrogante mediante el reconocimiento de patrones y tendencias. Generalmente lo que se quiere predecir es en el futuro, sin embargo, se puede utilizar para modelar un evento desconocido, independientemente de la



fecha de ocurrencia. Este modelo consta de dos tipos de variables, la o las variables predictoras y la variable de respuesta, en donde la variable de respuesta es expresada en función de una o más variables predictoras. Las variables predictoras, también conocidas como variables independientes o explicativas, se utilizan para explicar la variable de respuesta. En cambio, la variable de respuesta, también conocida como variable dependiente u objetivo, es la cual se desea predecir. Este modelo es de gran utilidad ya que permite tomar decisiones sociales, políticas y empresariales. Algunos ejemplos de aplicación son el predecir la demanda de cierto producto, si un empleado/cliente tiene alta probabilidad de renunciar o terminar su contrato con la compañía, y en este caso, el predecir el riesgo de un país para poder así, analizar sus consecuencias e implicaciones.

Variable de respuesta

Al realizar un modelo de predicción, es necesario definir la variable de respuesta. La elección de esta es fundamental ya que afecta a la elección de las variables predictoras, así como al tipo de modelo que se utilizará y su eficacia. Para poder medir y explicar el riesgo país se pueden usar varios indicadores o instrumentos financieros que permiten cuantificarlo. Uno de estos indicadores o instrumentos son las calificaciones crediticias soberanas otorgadas por las agencias calificadoras como Standard & Poor's (S&P), Fitch o Moody's, "una calificación crediticia es una opinión educada sobre la probabilidad de que un emisor cumpla con sus obligaciones financieras en tiempo y forma" [Standard and Poors Global Ratings \(S.F\)](#), en este caso, el emisor es el país en cuestión. Otro instrumento considerado fue el Swap de Incumplimiento Crediticio (o CDS por sus siglas en ingles), que "es un instrumento financiero derivado que brinda al comprador protección contra el incumplimiento a la deuda emitida por un país a cambio de una prima" [Tim Vipond \(2022\)](#). Es utilizado como herramienta para medir el riesgo de un país, así como para especular sobre el incumplimiento en las obligaciones de este. Por último, se puede usar la diferencia entre tasas de interés de los bonos gubernamentales y un bono de referencia, como son los bonos del Tesoro de Estados Unidos (también conocido como el *Yield Spread*). Este indicador es comúnmente usado para medir la percepción del mercado sobre el riesgo crediticio de un país. Para el modelo de predicción de este trabajo, se usó este último indicador como variable de respuesta. Se intentó realizar el modelo de predicción con las calificaciones crediticias, sin embargo, estas presentan un cambio casi imperceptible (estadísticamente hablando) a

lo largo de tiempo, por lo tanto, se consideró que el modelo funcionaría mejor con el *Yield Spread*. En el caso de los CDS, la información de este instrumento para algunos países es escasa y si bien, para los países más grandes y desarrollados se contaba con la información, para la gran mayoría no. De igual manera se consideró que el modelo de predicción tendría un sesgo al no contemplar una muestra de países más amplia. De esta manera, el *Yield Spread* presenta ventajas considerables ya que presenta una volatilidad más notable a través del tiempo (comparada con las calificaciones crediticias), los bonos gubernamentales son emitidos por muchos más países, por lo cual la muestra de países es mucho más variada y es un indicador comúnmente usado para la medición del riesgo. Una vez definida la variable de respuesta, la base de datos utilizada consiste en 60 países, los cuales, si poseen información del *yield spread* para la ventana de tiempo seleccionada, siendo esta del periodo 2015 al 2020. Con el fin de realizar el modelo de predicción, se optó por transformar el *Yield Spread* a una variable binaria. Primero se calculó el cambio porcentual a través de los años mediante la siguiente formula:

$$C\% = \frac{Y_{st} - Y_{st-1}}{Y_{st-1}}, \quad (1)$$

donde Y_{st} es el valor del *Yield Spread* en el tiempo t y $C\%$ es el cambio porcentual en el precio. Para hacer la transformación de la variable, se optó por usar el percentil 95 como punto para determinar si la variable toma valor de 0 o 1. Se consideró que este percentil indica un cambio porcentual bastante elevado en el *Yield Spread*. Lo cual haría referencia a que, si algún país que estuviera por arriba de ese percentil sufriera algún evento desestabilizante, implicaría que el riesgo en ese país aumentaría y podría conllevar a caer en *default*. Para obtener el percentil 95 del total de datos del cambio porcentual se siguió la siguiente metodología:

1. Se ordenan los datos de menor a mayor.
2. Se busca la posición k del valor que corresponde al percentil q .
3. Se obtiene la posición correspondiente al percentil deseado, siendo P , en este caso es 0.95, utilizando la siguiente fórmula:

$$P = \frac{95n}{100}, \quad (2)$$

Donde n es el número de elementos de la muestra.



- Se utiliza interpolación lineal para estimar el valor del percentil dado que los datos pertenecen a una distribución continua, mediante:

$$P = a(f(k)) + [k - f(k)] * [a(f(k) + 1) - a(f(k))], \quad (3)$$

Donde $f(k)$ es la función que redondea hacia abajo la posición x y $a(i)$ es el i -ésimo valor en la lista ordenada de datos

- Se determina que

$$x = \begin{cases} 1 & \text{si } x > P, \\ 0 & \text{si } x \leq P, \end{cases} \quad (4)$$

donde x es el cambio porcentual del *Yield Spread* para cada país.

En la gráfica, podemos observar en el eje y el conteo de países, mientras que el eje x presenta el cambio porcentual. Se observa que la gran mayoría de los países tienen un cambio porcentual del *Yield Spread* alrededor del 0, lo cual indicaría cierta estabilidad en ellos. Para los países que se encuentran más hacia la derecha de la distribución implicarían que tienen un cambio porcentual muy grande, que se podría interpretar como un aumento en su riesgo. Por el contrario, los países que se encuentran a la izquierda de la distribución no presentan un problema de riesgo mayor.

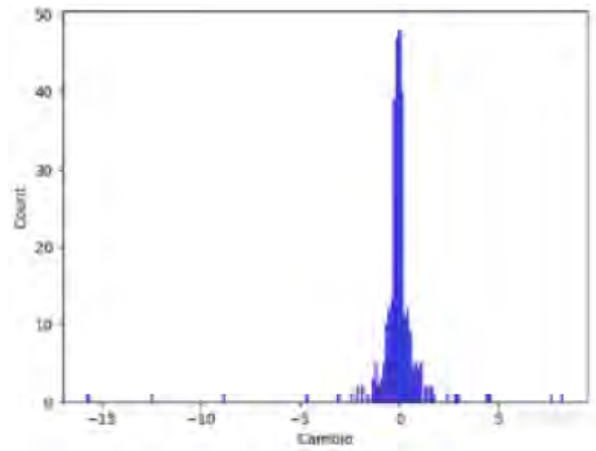


Figura 1 Distribución de la variable *Yield Spread*.

En el caso de México, el promedio del cambio porcentual del *Yield Spread* es de 0.0379. Por lo que la posición del país se mantuvo estable en el periodo de 2015 a 2020.

Variables Predictoras

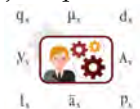
Una variable predictora desempeña un papel fundamental en la construcción de modelos predictivos, ya que, en este caso, nos permite predecir el riesgo que México tendrá en el actual año. Es por ello que identificar y seleccionar las variables que explican de mejor manera a la variable de respuesta es clave para obtener un buen modelo que se ajuste a los datos reales de los países seleccionados y obtener una predicción precisa. Para la creación de dicho modelo, se tomaron las variables macroeconómicas como las variables predictoras ya que el riesgo de un país se relaciona con las condiciones dentro de este. Asimismo, estas variables generan un impacto significativo en el riesgo de un país ya que reflejan la estabilidad económica, mercado laboral, inflación, capacidad de acceder a financiamiento en los mercados internacionales, la deuda pública, el comercio, entre otras. Entre mejor sea el comportamiento de las variables macroeconómicas, más saludable y estable será el país. Dentro de la base de datos del Banco Mundial se encuentran distintas categorías de las variables económicas, siendo: deuda, tecnología, economía, comercio, desigualdad social, desarrollo urbano, educación, eficacia, energía, infraestructura, finanzas, sector privado, y sector público. En total, se obtuvo una muestra de 956 variables macroeconómicas a considerar para la realización del modelo. Para poder realizar una selección de variables en forma arbitraria, se decidió utilizar tres modelos de selección:

- Método LASSO:** *Least absolute shrinkage and selection operator* por sus siglas en inglés. Este realiza la selección mediante un modelo de regresión lineal donde se penalizan las variables menos importantes, otorgando mayor peso a las variables que explican mejor el modelo. Este se expresa de la siguiente manera:

$$\underset{\beta_0}{\text{Min}} \left[\frac{1}{N} \sum_{i=1}^N (y_i - \beta_0 - x_i^T \beta)^2 \right] \text{ sujeto a } \sum_{j=1}^p |\beta_j| \leq t \quad (5)$$

Donde β_0 es el coeficiente constante, $\beta_i := (\beta_1, \beta_2, \dots, \beta_p)$ el vector de coeficientes, t es un parámetro preestablecido que determina la cantidad de regularización, N es el número de observaciones casos, cada una con p variable, y_i la variable independiente y $x_i := (x_1, x_2, \dots, x_p)^T$ es el vector con variables para el caso j

- Decision Tree:** Es un algoritmo que se utiliza para selección, clasificación y predicción de variables. Se presenta con una estructura de árbol jerárquica, la



cual consiste en un nodo raíz, ramas, nodos internos y nodos hoja.



Figura 2 Árbol de decisión IBM (S.Fc).

Como se muestra en el diagrama, un árbol de decisión empieza con un nodo raíz, donde a partir de este se toman decisiones y se generan ramas que conectan con los nodos internos llamados nodos de decisión, para finalmente obtener los nodos terminales, los cuales representan todos los resultados posibles. Para hacer más sencillo el modelo se utiliza la poda, este es un proceso que consiste en eliminar las ramas con poca importancia. La principal ventaja de utilizar este modelo es su facilidad de interpretación debido a su diseño.

3. *Random Forest*: Es un algoritmo que combina los posibles resultados de múltiples árboles de decisión para llegar a un solo resultado. Por lo que este modelo es más preciso.

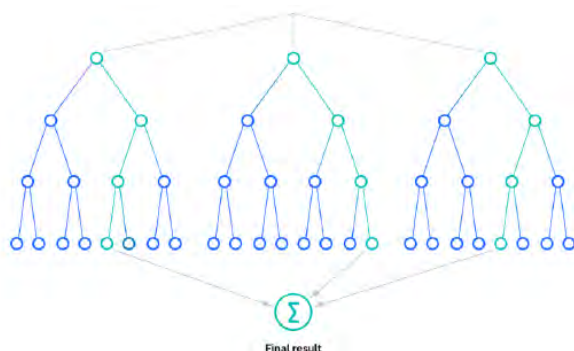


Figura 3 *Random Forest* IBM (S.Fb).



Como se muestra en el diagrama anterior, este modelo posee la misma estructura que un árbol de decisión, donde en cada nodo interno se generan subconjuntos de variables para poder observar su comportamiento y generar resultados diferentes. Asimismo, los nodos de color verde muestran los mejores resultados, para finalmente llegar a un solo resultado.

Cada modelo arrojó variables distintas, por lo que, utilizando nuestro criterio, escogimos las diez variables que creímos más relevantes, siendo:

Gasto en bienes y servicios (% del gasto)

“Los bienes y servicios incluyen todos los pagos del Gobierno realizados a cambio de bienes y servicios utilizados para la producción de bienes y servicios de mercado y ajenos al mercado.” *IndexMundi* (S.F). Este nos indica el dinero que se destina a la adquisición de producto, desde alimentos hasta transporte.

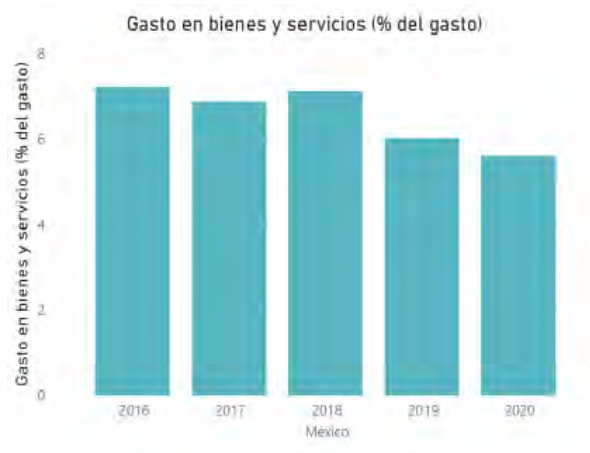


Figura 4 Gasto en Bienes y servicios (% del gasto) en México.

En esta gráfica podemos ver el gasto en bienes y servicios como porcentaje del gasto de México. Se nota estable pero decreciente en los últimos años. Una causa de esto puede ser la reciente pandemia de COVID-19, donde debido a las restricciones y cierre de muchos países, se generó incertidumbre económica y cambios en los patrones de consumo ya que cambiaron las necesidades de la población.

Capacidad de financiamiento neto (+), necesidad de financiamiento neto (-) (% del PIB)

Este se refiere a la capacidad que tiene dicho país de financiar sus operaciones entre la necesidad de financiamiento que se tiene. Esta variable es útil para analizar la capacidad financiera de los países.

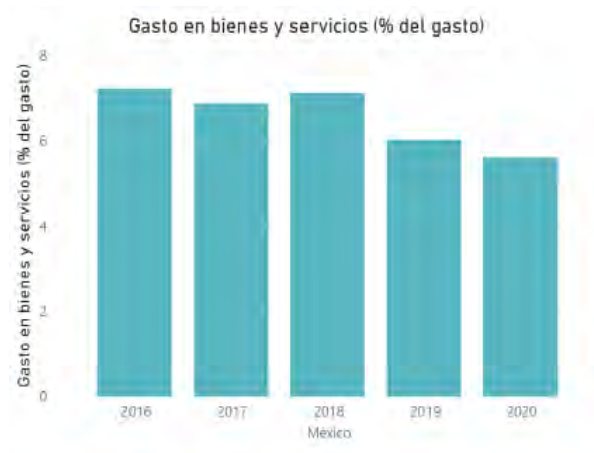


Figura 5 Capacidad de financiamiento neto (+), necesidad de financiamiento neto (-) (% del PIB)) en México.

Esta gráfica muestra la capacidad de financiamiento neto (+) y necesidad de financiamiento neto (-) (% del PIB) de México. Este ha presentado altas y bajas, pero la mayor fue en el 2020 con una caída de aproximadamente 125 % con respecto al año anterior. Esto significa que el país experimentó una reducción grave en su capacidad de financiamiento a través de ingresos internos y tuvo una mayor dependencia de fuentes externas de financiamiento. Nuevamente, esto se presentó debido a la pandemia de COVID-19 y la reducción en la calificación soberana por parte de S&P y Fitch ante la posible caída económica. Esto pudo haber llevado al país a buscar fuentes externas de financiamiento a un mayor costo, como préstamos internacionales o emisión de deuda para hacer frente a sus obligaciones.

Mercaderías importadas desde economías en desarrollo de Asia oriental y el Pacífico (% del total de mercaderías importadas)

Es bien sabido que uno de los principales países exportadores a nivel mundial es China, perteneciente a Asia, por lo que es un indicador importante.

En la gráfica de serie de tiempo de México sobre las mercaderías importadas desde economías en desarrollo de Asia oriental y el Pacífico se nota un claro incremento de estas. México tiene acuerdos comerciales con distintos países de Asia oriental y el Pacífico, un ejemplo de estos es Japón y los países miembros del Tratado Integral y Progresista de Asociación Transpacífico. Dicho acuerdo fue firmado en febrero de 2016 y consiste en un tratado de libre comercio entre once países, donde se reducen las barreras arancelarias, facilitando el comercio entre ellos. De igual manera, el crecimiento de marcas y productos importados de países dentro de estas regiones ha incrementado rápidamente, logrando un gran alcance en México.

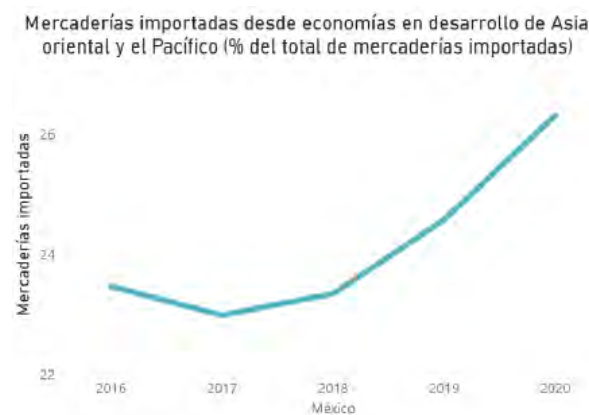
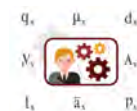


Figura 6 Mercaderías importadas desde economías en desarrollo de Asia oriental y el Pacífico en México.

Índice de términos netos de intercambio (2000=100)

“La relación porcentual de los índices de valor de la unidad de exportación con respecto a los índices de valor de la unidad de importación, medidos con relación al año 2000 tomado como base.” *IndexMundi* (S.F). Este indica la relación entre el precio de las exportaciones e importaciones. Cuando este número es grande significa que las exportaciones han aumentado, en precio o tamaño, en comparación con las importaciones y viceversa.

En la gráfica se aprecia el cambio en el Índice de términos netos de intercambio de México. Es evidente que ha estado estable pero decreciente en los últimos años, esto dice que las exportaciones han disminuido mientras las importaciones han aumentado. Si bien una de las causas puede ser la pandemia de COVID-19 con el cierre y



restricciones de distintos países, otra de las causas es la demanda interna dentro del país al tener una gran población que demanda de bienes y servicios, por lo cual se recurre a las importaciones para satisfacer las necesidades. Esta variable está relacionada con las mercaderías importadas desde economías en desarrollo de Asia oriental y el Pacífico, ya que, si estas aumentan, el índice de términos netos de intercambio disminuye. Es decir que tienen una correlación negativa, donde uno incrementa y el otro disminuye.

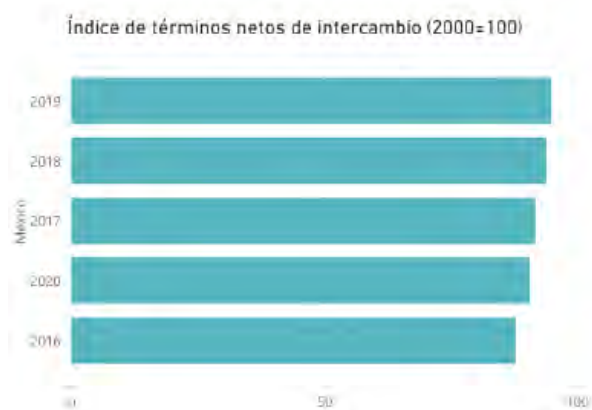


Figura 7 Índice de términos netos de intercambio de México.

Turismo internacional, recibos por artículos de viaje (US\$ a precios actuales)

Se refiere a todos los gastos relacionados con los visitantes internacionales dentro de dicho país, desde artesanías hasta gastos de hospedaje. La gráfica mostrada a continuación muestra el turismo internacional, medido mediante los recibos por artículos de viaje en México. Es incuestionable el cómo la crisis sanitaria de COVID-19 impactó negativamente al turismo internacional dentro del país debido al miedo e incertidumbre alrededor del mundo. Siendo el turismo una de las principales fuentes de ingreso del país, fue un golpe fuerte el haber disminuido casi por completo la entrada de extranjeros al país, provocando una desaceleración económica. Muchas empresas se vieron obligadas a cerrar o reducir su actividad y fuerza laboral, donde finalmente la economía de las familias padeció. A pesar de que el gobierno implementó medidas para mitigar el impacto económico, como programas de apoyo financiero a microempresas, estas



no satisficieron por completo los efectos negativos de la crisis sanitaria.



Figura 8 Índice de términos netos de intercambio de México.

Índices *Global Equity* de S&P (% anual de cambio)

La gráfica de Índices *Global Equity* de S&P de México no ha sido estable, independientemente de los efectos adversos de la pandemia de COVID-19, este ha fluctuado en el tiempo. Hablando del periodo de 2016, ocurrieron las elecciones de Estados Unidos, impactando directamente el cambio en el índice para México al ser el país vecino de Estados Unidos y la dependencia que se tiene con este. Para el periodo de 2018, este se pudo haber visto afectado por el cambio de presidencia, generando volatilidad en los mercados de valores.

Inversión extranjera directa, neta (US\$ a precios actuales)

Según El Gobierno de México, el propósito de este es "crear un vínculo duradero con fines económicos y empresariales de largo plazo, por parte de un inversionista extranjero en el país receptor." [Gobierno de México \(2016\)](#). Este tipo de inversión permite crecer a los países el desarrollo y mejorar su gobernanza. La siguiente gráfica muestra la inversión extranjera directa en México durante los últimos años. Ha habido un decremento de la inversión extranjera directa dentro del país. Esto puede ser interpretado de varias formas, puede significar que el país no recibió nuevas inversiones o una pérdida de las inversiones existentes, disminuyendo la inversión que



Figura 9 Índices Global Equity de S&P de México.

se tiene; asimismo, puede ser interpretado como desestabilización en la economía generada por un aumento de desconfianza de los inversionistas. México es un país altamente atractivo para los extranjeros debido a la riqueza natural que se posee, la flexibilidad respecto a las barreras de entrada al mercado y su fácil adaptabilidad. Se encuentran vigentes distintos tratados que permiten la facilidad de inversión extranjera como es el tratado T-MEC. Pero no todo es de color rosa, el gobierno, la incertidumbre e inconformidad dentro del país representa un gran peligro hacia los inversionistas, al no ser un mercado tan estable.



Figura 10 Inversión extranjera directa, neta (US\$ a precios actuales) en México.

Cambios en las reservas netas (US\$ a precios actuales)

Banxico lo define como “activos financieros que el banco central invierte en el exterior. Su característica principal es la liquidez; es decir, la facilidad de uso para saldar rápidamente obligaciones de pago fuera de nuestro país.” [Banxico \(2023\)](#). El objetivo de este es mantener la estabilidad de los precios ante eventos imprevistos o desequilibrios macroeconómicos y/o financieros. Esta gráfica muestra el cambio en las reservas netas (US\$ a precios actuales) de México. Sorprendentemente estas son crecientes. Este cambio es debido, según el Gobierno de México, “al cambio en la valuación de los activos internacionales del Instituto central” [Gobierno de México \(2022b\)](#).

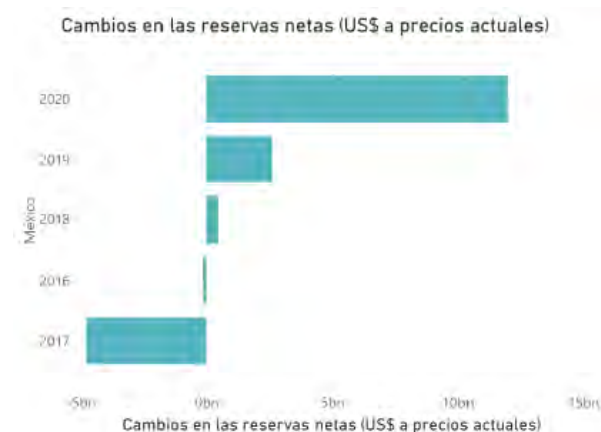


Figura 11 Reservas netas (US\$ a precios actuales) de México.

Desempleo, total (% de la población activa total)

Este se refiere al porcentaje de la población en edad de trabajar que no poseen un empleo. Esto se puede deber a diversas causas como la falta de empleos disponibles, ciclos económicos, educación dentro del país, entre otros. La gráfica que se presenta muestra el desempleo, total (% de la población activa total) de México. Esta había presentado decrementos en los últimos años, pero durante la pandemia de COVID-19 aumentó. Como se comentaba anteriormente, el decremento de la inversión extranjera y el turismo fueron afectados, y con ello, la empleabilidad también. Aunado a esto, el cambio en el mundo tecnológico ha evolucionado de una manera impresionante, experimentando cambios en la demanda de habilidades.



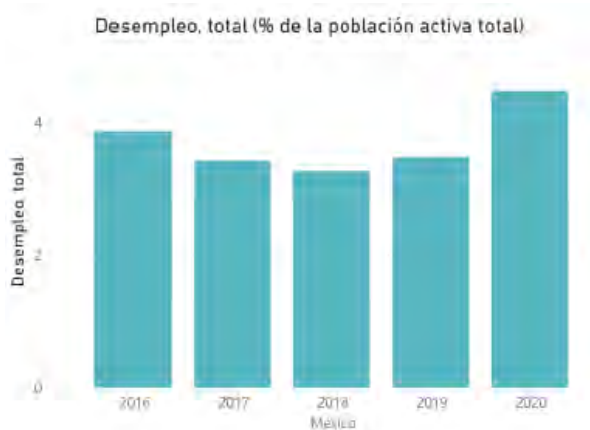


Figura 12 Desempleo, total (% de la población activa total) de México.

Hoy en día existen procesos automatizados que realizan procedimientos básicos, y están reemplazando a las personas que se encargaban de dicho trabajo, por lo que la demanda de empleos ha ido cambiando conforme la tecnología vaya evolucionando. Por este motivo, cada vez más se necesita de personas más calificadas o con estudios superiores, ya que los empleos que más rápido se están reemplazando son justo aquellos que necesitan menos preparación académica. Esto es un gran problema para países en desarrollo porque muchos de ellos presentan una alta tasa de abandono escolar, siendo así que en México se reportó una tasa de abandono escolar de grado media superior de 2020 al 2021 del 11.6% y de grado superior del 8.8% según información de INEGI.

Asimismo, el crecimiento de la población y la migración a las zonas urbanas debido al calentamiento global y falta de oportunidades incrementa la demanda laboral, lo cual dificulta la absorción de nuevos empleos.



Crecimiento de la población urbana (% anual)

De acuerdo con INEGI, “una población se considera rural cuando tiene menos de 2,500 habitantes, mientras que la urbana es aquella donde viven más de 2,500 personas.” INEGI (S.F). Este indicador es de suma importancia ya que evalúa el proceso de urbanización dentro de un país. Esta gráfica muestra el crecimiento de la población urbana (% anual) en México. Dentro del país se ha experimentado un proceso de urbanización, generando conflictos de abastecimiento de productos y lugares donde vivir debido a la gran demanda de productos, por lo que el crecimiento de la población urbana se desacelera. Como se mencionaba anteriormente, la gran demanda y poca oferta laboral, la reactivación del sector turístico, la búsqueda de mejores oportunidades de vida en el extranjero, son algunos de los factores que han y seguirán afectando el desarrollo y evolución demográfica del país.



Figura 13 Crecimiento de la población urbana (% anual) en México.

Por último, se realizó una matriz de correlación de las variables predictoras entre ellas y la variable de respuesta utilizando la herramienta de programación Python. Este es una medida de dependencia lineal entre dos o más variables cuantitativas, es decir, muestra que tan relacionadas están las variables predictoras con la variable de respuesta en un rango de -1 a 1. Entre mayor sea este número, mayor será la correlación positiva. Por el contrario, entre menor sea este número, mayor será la correlación negativa.

Con respecto a la gráfica anterior, se puede observar como las variables predictoras no muestran una fuerte correlación entre ellas, siendo la mayor de 0.5, corres-



Figura 14 Correlación variables de respuesta con variables predictoras.

pondiente a la relación de Gasto en bienes y servicios y crecimiento de la población urbana, ya que, si aumenta la población urbana, aumenta la demanda de productos, por ende, aumenta el gasto en bienes y servicios.

Asimismo, se presenta una correlación ligeramente fuerte entre Mercaderías importadas desde economías en desarrollo de Asia oriental y el Pacífico con desempleo, esto se debe a que, si aumentan los productos importados, se requiere menos mano de obra, ocasionando desempleo dentro del país. Consideramos que este es uno de los retos más grandes a los que se enfrenta México ya que engloba el nivel educativo, forma de vida, y posibilidades laborales que se tienen dentro del país.

Finalmente, resalta que la variable de respuesta no muestra una fuerte correlación con ninguna de las variables predictoras. Esto puede decir que, al seleccionar únicamente 10 variables de las 956 totales, es claro que estas no expliquen en su totalidad a la variable de respuesta. Igualmente, el tamaño de la muestra no es tan significativo, se esperarían mejores resultados con una muestra mayor.

Modelo de predicción mediante *Machine Learning*

De acuerdo con BBVA, “*Machine Learning* o aprendizaje automático es una rama de la inteligencia artificial que permite que las máquinas aprendan sin ser expresamente programadas para ello.” BBVA (2019). Esta herramienta es ampliamente usada en el área de predicción y su funcionamiento está sencillamente explicado por IBM:

“Un modelo de *machine learning* es la salida de información que se genera cuando entrena su algoritmo de *machine learning* con datos. Después del entrenamiento, al proporcionar un modelo con una entrada, se le dará una salida. Por ejemplo, un algoritmo predictivo creará un modelo predictivo. A continuación, cuando proporcione el modelo predictivo con datos, recibirá un pronóstico basado en los datos que entrenaron al modelo.” IBM (S.Fa)

Dentro de *machine learning* existen distintas categorías que se usan según el problema que se está analizando.

Las categorías se dividen en: aprendizaje supervisado, definido un poco más adelante; aprendizaje no supervisado, basado en la clasificación de datos con base en los patrones que encuentra el algoritmo; aprendizaje de refuerzo, que consta en el aprendizaje a través de la prueba y error; y, por último, *deep learning*, que sirve para aprender de forma iterativa mediante redes neuronales en capas. IBM (S.Fa)

En este caso, se usó aprendizaje supervisado, los cuales utilizan datos previamente etiquetados o clasificados, para entrenar al modelo y posteriormente, el modelo tendrá que ser capaz de predecir la etiqueta o la clasificación de un conjunto de datos que no haya visto antes mediante los patrones y las características.

Se realizó un modelo de predicción para determinar el riesgo país. Para ello se usó el software *Python*, específicamente la librería *Sklearn*, la cual es una librería para el análisis de datos mediante aprendizaje automático. Este modelo toma dos muestras conjuntos de datos, *train* y *test*, divididas aleatoriamente a partir de la muestra entera de datos, donde *train* es una muestra de entrenamiento utilizada para estimar los parámetros del modelo y ajustar los datos. Por otro lado, *test* es la muestra de prueba para evaluar el rendimiento del modelo.

Existen diferentes tipos de modelos de predicción de *Machine Learning* como los modelos de regresión lineal o logística, los árboles de decisión (también conocidos como *Decision tree*), modelos *Random Forest* (conjunto de árboles de decisión) o máquina de vectores de soporte (SVM por sus siglas en inglés). Cada uno de estos modelos busca predecir con diferentes herramientas la variable de respuesta a partir de las variables predictoras, y la precisión de cada uno varía dependiendo tanto del tipo de variable de respuesta (binaria, continua, discreta), como la capacidad de predicción de las variables predictoras.

Se decidió utilizar los modelos de *Decision tree*, *Random Forest* y *SVM* ya que, del total de modelos probados, estos presentaban mejores resultados. Los primeros dos modelos descritos en la sección *Variables Predictoras*, se basan en árboles de decisión, de los cuales se llega al modelo que explica de mejor manera la variable de respuesta.

En el caso de *SVM*, este se basa en relacionar datos a dos espacios o conjuntos de características y se dividen los datos dependiendo de su similitud con los conjuntos. Una vez que se tiene un separador entre ambos conjuntos mediante un hiperplano, se transforman los datos con una función llamada *kernel*, que puede ser lineal, polinómica, función de base radial (RBF) o sigmoide. Cada función *kernel* es recomendable con algún tipo de datos en específico, y dependerá del desarrollador



probarlas con distintas funciones para obtener el mejor modelo.

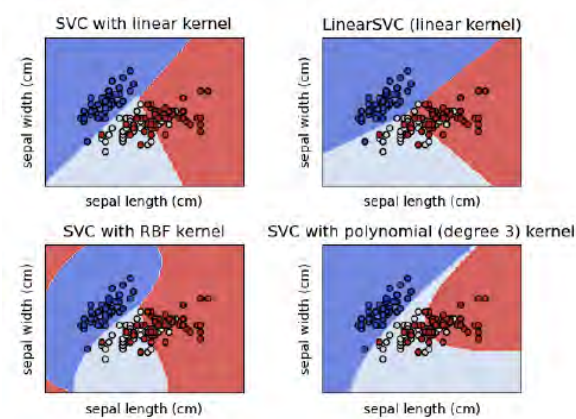


Figura 15 Ejemplo de clasificación mediante SVM con diferentes funciones *kernel*. [scikit learn](#) (S.F)

Se presentará una tabla con los resultados obtenidos de los modelos:

Modelo de predicción	Coefficiente de determinación	Error Medio Absoluto
Random Forest	85%	0.03
Decision Tree	83%	0.04
SVM	90%	0.11

Figura 16 Resultado de Modelos de predicción

De acuerdo con Minitab “El coeficiente de determinación es la proporción de la varianza de la variable de respuesta explicada por el modelo” [Minitab \(2019\)](#). En otras palabras, es que tan lejos o cerca están los datos al modelo que se le está ajustando. Si el coeficiente es cercano a 0 %, el modelo no se ajusta a los datos y no da



ninguna explicación. Y viceversa, si es cercano a 100 % se dice que tiene un ajuste bueno.

En cuanto al error medio absoluto, nos indica el promedio de las diferencias de las predicciones y los valores reales. Entre más bajo sea el error, el modelo está haciendo predicciones más precisas.

Se puede observar en la tabla que el modelo que obtuvo una mayor precisión para explicar los datos fue el SVM. Sin embargo, tiene el error medio más alto de los 3 modelos. Por estas razones, y porque no se quiere tener el riesgo de usar un modelo sobre ajustado, entendiéndose un modelo sobre ajustado como un modelo que proporciona predicciones muy acertadas para los datos de entrenamiento, pero no para los nuevos datos, se decidió usar el modelo *Random Forest* para la predicción del riesgo de México.

Predicción del riesgo país de México en 2023

Como fue mencionado, la ventana de tiempo utilizada fue de 2015 a 2020, por lo que para hacer la predicción del riesgo país de México en 2023, es necesario proyectar las variables predictoras hasta ese año.

Para las variables se hizo uso de la fórmula de Excel *pronóstico*, basada en la metodología *Suavizado Exponencial Simple*, que es un método de pronóstico que asigna un peso exponencial decreciente a los datos históricos que se usan. Por ende, otorga un mayor peso a los datos más recientes, y su peso irá decreciendo conforme los datos sean más antiguos.

La fórmula general para el cálculo del pronóstico siguiendo esta metodología está dada por:

$$Pr = \alpha * V_t + (1 - \alpha) * V_{t-1} \tag{6}$$

Donde α es un valor suavizado entre 0 y 1 que determina el peso que se le dará al valor, cercano a 1 para los valores más actuales y cercano a 0 para valores más antiguos.

V_i representa los valores para pronosticar, siendo V_i un valor más actual que V_{i-1} .

Asimismo, se calibraron los pronósticos dados con Excel utilizando pronósticos encontrados en distintas páginas como INEGI y BANXICO.

Gasto en bienes y servicios

BBVA Research indica que se reportó en el mes de enero de 2023 un incremento de 3.8 % con respecto al periodo anterior [Salazar \(2023\)](#). Este comportamiento se vio impulsado por la estabilización del mercado mexicano con respecto a la pandemia de COVID-19, la

mejora del salario real registrando 1 % de crecimiento vs el periodo anterior. Gracias a este comportamiento, se observa en la gráfica los valores pronosticados para 2021, 2022 y 2023.

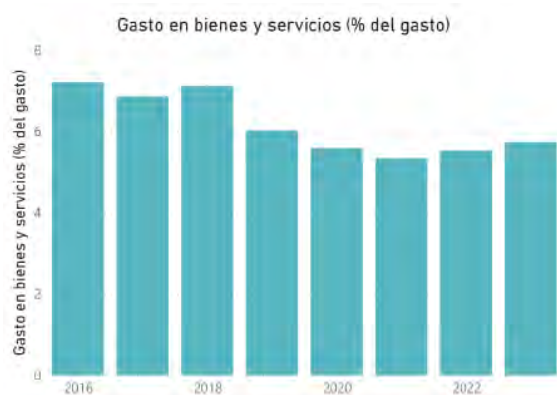


Figura 17 Gasto en bienes y servicios (% del gasto) proyección

Capacidad de financiamiento neto (+) /necesidad de financiamiento neto (-) (% del PIB)

Se asumió que la población seguirá creciendo, aumentando las necesidades y capacidad de financiamiento del país. Por ello, se pronosticó para 2021, 2022 y 2023:

Mercaderías importadas desde economías en desarrollo de Asia oriental y el Pacífico

En reunión con empresarios de México, Estados Unidos y Canadá, la secretaria Raquel Buenrostro comentó la importancia de fortalecer la región de América del Norte, sustituyendo importaciones de Asia con producción regional, para lograr un desarrollo justo, equitativo e incluyente” [Roberto Morales \(2023\)](#), dijo la Secretaría de Economía en un comunicado de prensa. A pesar de los distintos tratados de comercio, México sigue buscando sustituir las importaciones de economías de Asia oriental y el Pacífico, disminuyendo la dependencia que se tiene

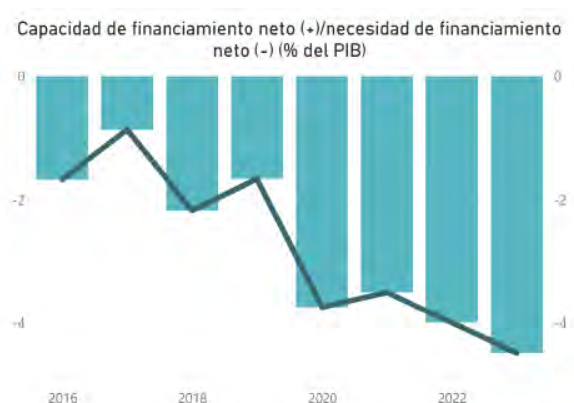


Figura 18 Capacidad de financiamiento neto (+), necesidad de financiamiento neto (-) (% del PIB) proyección

con ellos. Por lo que, si se decide reducir las importaciones y trabajar con producto nacional. Las predicciones para 2021, 2022 y 2023 se verían de la siguiente manera:

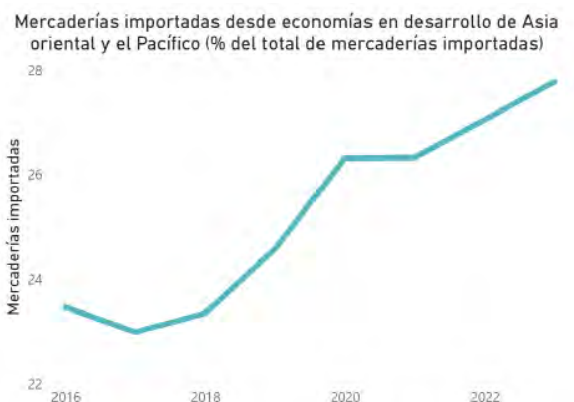


Figura 19 Mercaderías importadas desde economías en desarrollo de Asia oriental y el Pacífico proyección

Índice de términos netos de intercambio (2000=100)

Lo anterior va de la mano con el índice de términos netos de intercambio (2000=100), si las importaciones disminuyen, este índice aumentará.



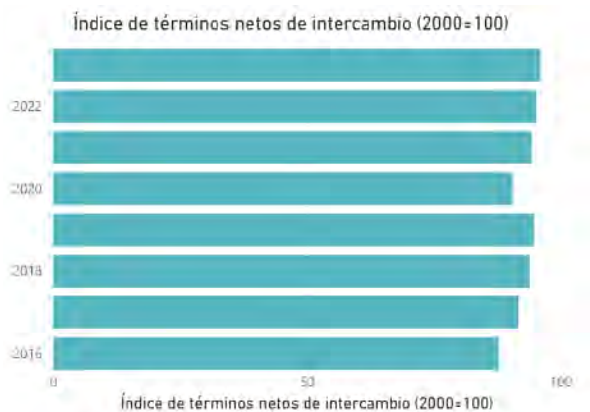


Figura 20 Índice de términos netos de intercambio (2000=100) proyección.

Turismo internacional, recibos por artículos de viaje (US\$ a precios actuales)

Sabiendo que el miedo y la incertidumbre causada por la pandemia de COVID-19 ha disminuido, México es uno de los destinos más baratos para países desarrollados y con mayor biodiversidad. Se dio a conocer por el Gobierno de México que “los principales países en llegadas de turistas internacionales vía aérea a México de enero a septiembre de 2022, fueron Estados Unidos, Canadá y Colombia, con 11 millones 440 mil turistas; esto es 51.1 % más respecto a 2021, y superando también un 14.3 % los 10 millones 10 mil turistas de 2019.” [Gobierno de México \(2022a\)](#).

Los destinos más turísticos con respecto a las llegadas aéreas durante periodo enero a septiembre de 2021 fueron:

1. Cancún, Quintana Roo con 7,075,000 turistas.
2. Ciudad de México con 3,050,000 llegadas.



Figura 21 Principales destinos turísticos en 2021 [Gobierno de México \(2022a\)](#).

3. Los Cabos, Baja California Sur con 1,602,000 llegadas de turistas.

Se espera que este siga aumentando, pero uno de los factores que podría perjudicar el incremento del turismo proveniente de Estados Unidos es el desempeño del dólar frente al peso mexicano, siendo más costoso el viajar a México. Suponiendo que incrementará cada año un 10 %, para el periodo 2022 y 2023 se consideraron los datos en la imagen anterior.

Índices *Global Equity* de S&P

Se pronosticó un incremento en este ya que no se espera la presencia de algún factor que desestabilice la economía de México. Se observó que en el año 2021 este fue de 16.62 %, mostrando un aumento de 567 % con respecto al año anterior. Esto es un buen indicio para los siguientes pasos de México, por lo que las predicciones para 2022 y 2023 pueden verse en la figura 23.

Inversión extranjera directa

Se consultó la Encuesta Sobre las Expectativas de los Especialistas en Economía del Sector Privado, donde se



Figura 22 Turismo internacional, recibos por artículos de viaje (US\$ a precios actuales) proyección.

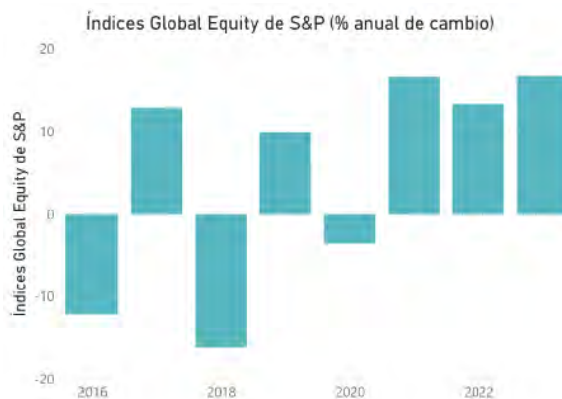


Figura 23 Índices Global Equity de S&P proyección.

trabajó con la media de las expectativas de los especialistas en este sector. Esto es de gran utilidad para comprender lo esperado basado en las tendencias económicas y posición del mercado. Utilizando las predicciones para 2022 y 2023 se observa la tendencia de la siguiente forma:

Cambio en las reservas netas

Con base en el Comunicado de Prensa de Banco de México del 9 de mayo del presente año, el saldo de la reserva internacional de 2018 hasta el 5 de mayo de 2023 se mostraba como en la imagen a la izquierda. Considerando que los datos que estamos midiendo es el cambio que presenta el saldo de la reserva, este es creciente, por



Figura 24 Inversión extranjera directa, neta proyección.

lo que se obtuvieron los cambios de cada año y el cambio con respecto al mes de mayo para el 2023, mostrados a continuación.

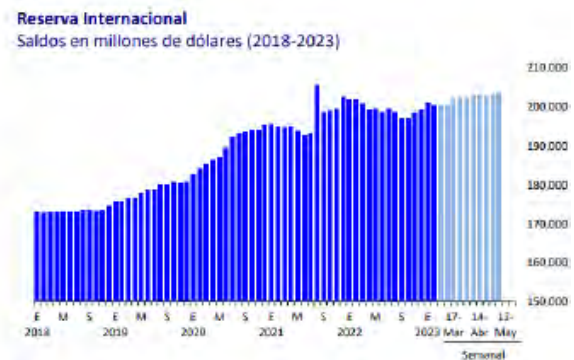
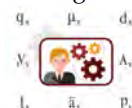


Figura 25 Reserva Internacional Banxico (S.F).

Desempleo, total (% de la población activa total)

De acuerdo con La Encuesta Nacional de Ocupación y Empleo (ENOE), "en enero de 2023, la Población No Económicamente Activa (PNEA) fue de 39.6 millones de personas, 1.6 millones menos que en el primer mes de 2022." INEGI (2023). Dentro de este documento se señala que se presentó una reducción en enero de 2023 de 0.7 puntos porcentuales con respecto al mes del año anterior. Se puede apreciar en la tabla anterior como la población desocupada disminuyó 0.3 puntos porcentuales, 0.2 puntos porcentuales correspondientes al género



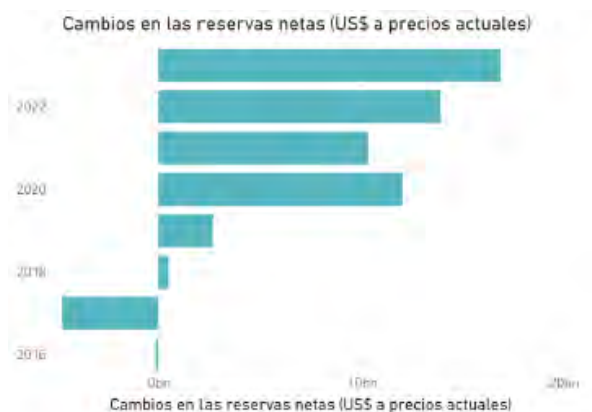


Figura 26 Cambio en las reservas netas (US\$ a precios actuales).

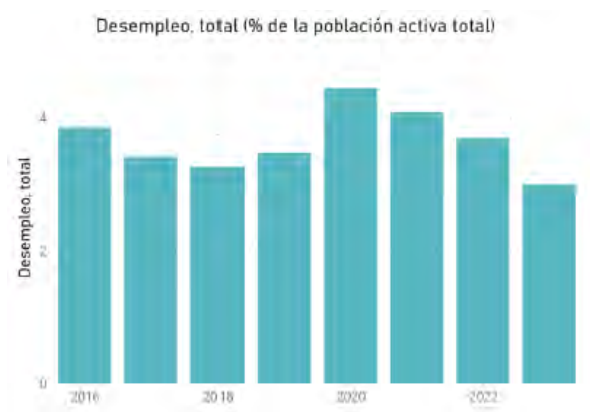


Figura 28 Desempleo, total (% de la población activa total).

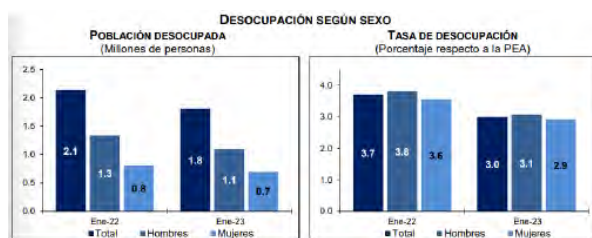


Figura 27 Desocupación según sexo en México durante 2022 y 2023 INEGI (2023).

masculino y 0.1 puntos porcentuales correspondientes al género femenino. Asimismo, la tasa de desocupación disminuyó. Considerando estos factores, se utilizó la tasa de desocupación dentro del modelo.

Crecimiento de la población urbana (% anual)

Hablando de la población en general, según las proyecciones de CONAPO de 2022 al 2050, se estima que el crecimiento de la población en México crezca a un ritmo menor al que se estaba presentando, con una cifra de 130.1 millones de personas en el 2022. Según la ONU "En las próximas décadas, buena parte del crecimiento demográfico en México será urbano. Esto significa que el país pasará de contar con 384 ciudades a 961 en 2030, en las que se concentrará 83.2 % de la población nacional y en donde muy probablemente, sea la población pobre la que predominará." ONU-Habitat (2017). Considerando que el cambio es estable y homogéneo cada año, las predicciones para el 2022 y 2023 son las siguientes:



A partir del pronóstico descrito anteriormente, se realizó la predicción usando el modelo seleccionado, *Random Forest*, previamente entrenado y el resultado fue el siguiente: Este resultado se interpreta como que México tiene un 99.2 % de probabilidad de que el cambio porcentual del *Yield Spread* no sea lo suficientemente grande para que se presente el riesgo país. En otras palabras, que el *Yield* de los bonos gubernamentales de México no aumentará de manera considerable con respecto al *Yield* de los bonos del tesoro de Estados Unidos y tendrá relativamente el mismo comportamiento que ha mantenido en los años anteriores.

Esto se traduce en que México mantendrá el nivel de riesgo, o por lo menos no aumentará de manera considerable como para causar desestabilidad en el país. Asimismo, se podría esperar que los inversionistas extranjeros volteen a ver a México, ya que mantener la estabilidad en tiempos en los que muchos países se encuentran tambaleando y teniendo problemas económicos y sociales es una clara muestra de la fortaleza del país. De esta manera, no es de sorprender la reciente noticia de la inversión millonaria que hará *Tesla Motors* en el norte del país para la construcción de una fábrica de autos eléctricos.

Por lo cual, se puede decir que México cerrará el año 2023 con estabilidad. El correcto comportamiento de la economía y de distintos factores como el turismo en el país, han llevado a México a que conserve, y puede ser que hasta fortalezca su posición en los mercados



Figura 29 Panorama demográfico de México, 2022. [Secretaría General del Consejo Nacional de Población \(2022\)](#)

globales.

Consideramos que el resultado del modelo es consistente con lo que el país ha registrado en 2021 y 2022, y lo que se espera de México para este 2023. Sin lugar a duda, el desempeño en los últimos 2 años ha sido bastante bueno y artículos como “México convence al mercado y reduce riesgos, pese a desplome de economía” de Reuters o “El riesgo país de México sin cambios durante la tercera semana del año” en la Gaceta económica del Gobierno Federal confirman la estabilidad que ha tenido el país.

Ahora bien, para 2023, la realidad se asemeja mucho al modelo. De acuerdo con Fitch Ratings “la perspectiva para fondos de inversión en instrumentos de deuda es ‘Neutra’.” [Fitch Ratings \(2023\)](#). De la misma forma, consideran que hay un panorama general de estabilidad en la gran mayoría de los sectores en México. De esta manera, se confirma que para este año se espera que México continúe con esa estabilidad que ha mantenido en los últimos años pese a los tiempos difíciles.

Por último, si consideramos también la fortaleza que ha mostrado la divisa mexicana frente a la gran mayoría de las monedas, es de reconocer la estabilidad que ha mostrado el país. En un artículo de Bloomberg publicado en El Financiero dice que

“Antes de que surgieran los problemas bancarios, la que había sido la moneda más castigada del mundo al inicio de la pandemia se vio impulsada por un banco central agresivo, inversiones en el sector manufacturero en medio del reequilibrio de la cadena de suministro mundial y una política fiscal sorprendentemente austera”



Figura 30 Crecimiento de la población urbana (% anual)

Yield Spread	Probabilidad
0	99.2%
1	0.8%

[Michael O’Boyle \(2023\)](#)

Con lo cual, podemos reafirmar que la estabilidad económica de México se ve reflejado en muchos aspectos y no solo es una cuestión de perspectiva. Por lo cual, podemos decir que realmente el modelo se asemeja a la realidad.

4. CONCLUSIÓN

El modelo de predicción de machine learning nos permitió determinar el nivel de riesgo país utilizando las variables más importantes que afectan su cálculo. Es importante mencionar que, a pesar de lograr un modelo capaz de explicarlo, se considera que la medición del riesgo país no es algo tan trivial, y depende de gran cantidad de factores, tanto internos como externos. Por lo que, la inclusión de otro tipo o cantidad de variables podría derivar en otra conclusión diferente a la presentada en este trabajo.

Por otro lado, se considera que los algoritmos de machine learning sin lugar a duda, pueden ser de gran utilidad en la toma de decisiones, ya que permiten analizar grandes cantidades de datos y obtener resultados relativamente precisos. En este sentido, son una herramienta de gran capacidad que permiten al usuario, concentrarse en el análisis de los resultados y en la interpretación del modelo según se necesite.



Dicho esto, a partir de lo presentado, se puede decir que en el caso específico de México existe una buena estabilidad económica. El riesgo país se considera adecuado para la atracción de inversionistas extranjeros, para el fortalecimiento del sistema financiero, entre muchos otros aspectos. Estos mismos puntos, son los que permiten un gran desarrollo y ayudan a México con la diversificación de su economía y reducir esa dependencia que se tiene de algunos sectores como el sector petrolero.

Siempre hay áreas de mejora que deben ser abordadas como la transparencia y redención de cuentas, que generan una mayor confianza entre los inversionistas extranjeros y el pueblo mexicano; y como se mencionó anteriormente, permite y fomenta el desarrollo del país. Así se puede mantener, y no solo eso, sino que se puede mejorar la posición del país en el mercado internacional.

Asimismo, la innovación tecnológica puede ser de gran utilidad para seguir avanzando con las tendencias globales y económicas que se presentan actualmente. Otro punto que se debería considerar para mejorar al país sería fortalecer su sistema financiero ya que genera estabilidad económica y protección a los consumidores.

México no solo posee de riqueza natural, sino también de riqueza intelectual. El fomento del emprendimiento hoy en día es una de las herramientas que les ha servido a distintos países para mejorar su economía y darse a conocer entre los demás países. El seguir brindando herramientas de emprendimiento sensato, generando ideas novedosas, puede ser un distintivo de México.

Por otro lado, hablamos del turismo en México y su importancia monetaria. Una manera de poder impulsar el sector turístico y atraer más visitantes podría ser la diversificación de productos turísticos y propaganda de estos. La biodiversidad es una de las principales herramientas con las que México podría trabajar para ofertar más destinos turísticos, resaltando su riqueza natural, cultural y gastronómica. Asimismo, la mejora de servicios de transporte facilitaría el desplazamiento de extranjeros como nacionales.

En conclusión, la medición del riesgo país es crucial para el éxito y mejora económica de un país y garantizar una mayor estabilidad a nivel mundial. Por lo tanto, es importante seguir explorando y analizando distintas herramientas y enfoques que permitan comprender mejor el riesgo país y con ello, tomar decisiones más acertadas de acuerdo con las necesidades que cada país tiene.

REFERENCIAS

Banxico, 2023 Reservas internacionales. Banxico **Mayo**,



Recuperado el 24 de abril de 2023, [aquí](#).

Banxico, S.F Reservas internacionales. Banxico **Mayo**, Recuperado el 01 de marzo de 2023, [aquí](#).

BBVA, 2019 'machine learning': ¿qué es y cómo funciona? BBVA Innovación **Noviembre**, Recuperado el 10 de marzo de 2023, [aquí](#).

Fitch Ratings, 2023 Fitch publica perspectiva 2023 para fondos de deuda en México. Fitch Ratings **Enero**, Recuperado el 05 de abril de 2023, [aquí](#).

Gobierno de México, 2016 ¿qué es la inversión extranjera directa? Secretaría de Economía **Mayo**, Recuperado el 24 de abril de 2023, [aquí](#).

Gobierno de México, 2022a Ingresan a México más de 15 millones de turistas internacionales vía aérea de enero a septiembre de 2022. Secretaría de Turismo **Octubre**, Recuperado el 01 de marzo de 2023, [aquí](#).

Gobierno de México, 2022b Las reservas internacionales aumentaron 3.4

Gonzalez, A., 2020 México convence al mercado y reduce riesgos, pese a desplome de economía. Reuters **10 de Diciembre**, Obtenido de [aquí](#).

IBM, S.Fa Machine learning. IBM Recuperado el 10 de febrero de 2023, [aquí](#).

IBM, S.Fb Random forest. IBM Recuperado el 15 de abril de 2023, [aquí](#).

IBM, S.Fc Árboles de decisión. IBM Recuperado el 15 de abril de 2023, [aquí](#).

IndexMundi, S.F Gasto en bienes y servicios (INEGI, 2023 Indicadores de ocupación y empleo. INEGI **Marzo**, Recuperado el 31 de marzo de 2023, [aquí](#).

INEGI, S.F Indicadores financieros. INEGI Recuperado el 07 de enero de 2023, [aquí](#).

Michael O'Boyle, M. A., 2023 El 'superpeso' es invencible: estos son los riesgos para la moneda mexicana. El Financiero **Marzo**, Recuperado el 05 de abril de 2023, [aquí](#).

Minitab, 2019 Análisis de regresión: cómo puedo interpretar el r-cuadrado y evaluar la bondad de ajuste. Minitab Blog **Abril**, Recuperado el 03 de febrero de 2023, [aquí](#).

ONU-Habitat, 2017 Tendencias del desarrollo urbano en México. ONU-Habitat México **Junio**, Recuperado el 01 de marzo de 2023, [aquí](#).

Roberto Morales, 2023 México insta al sector empresarial a sustituir importaciones desde Asia a Norteamérica. El Economista **Enero**, Recuperado el 27 de abril de 2023, [aquí](#).

Salazar, S. A., 2023 México: resiliencia del consumo en enero, mayor gasto en bienes y servicios. BBVA Research **Febrero**, Recuperado el 27 de abril de 2023, [aquí](#). scikit learn, S.F Support vector machines (svm). scikit-

learn Recuperado el 10 de febrero de 2023, [aquí](#).
 Secretaría General del Consejo Nacional de Población,
 2022 La situación demográfica de México. CONAPO
 Recuperado el 01 de marzo de 2023, [aquí](#).
 Silvia Iranzo, 2008 Introducción al riesgo-país. Banco de
 España: Documentos Ocasionales N.º 0802 Recuperado
 el 20 de febrero de 2023, [aquí](#).
 Standard and Poors Global Ratings, S.F Understanding
 ratings. Standard and Poors Global Ratings Recupe-
 rado el 20 de febrero de 2023, [aquí](#).
 Tim Vipond, 2022 Credit default swap (cbs). Corporate
 Finance Institute. Recuperado el 23 de enero de 2023,
[aquí](#).



Ana María Hernández Piña estudió Actuaría en la Universidad Anáhuac México. Tiene un diplomado en Liderazgo Empresarial, uno en tecnología en negocios, y uno más en informática actuarial. Fue becaria por ocho meses en Philip Morris en el área de Da-

ta, y actualmente es se analista del sector de informática y tecnología en PepsiCo. Sus intereses profesionales están en la ciencia de datos, la creación de modelos de aprendizaje máquina, y en el área de investigación, específicamente del sector salud. Junto con José Eduardo Becerra López, realizó un modelo de optimización de inversión basado en años de vida ajustados por discapacidad para el sector salud en México con el objetivo de encontrar el mayor beneficio al menor costo.



José Eduardo Becerra López estudió Actuaría de la Facultad de Ciencias Actariales de la Universidad Anáhuac México. En ese período cursó un diplomado en Informática Actuarial y obtuvo un reconocimiento a la Exce-

lencia Académica por su alto promedio. Sus principales intereses están en las áreas de riesgos, análisis de datos y el aprendizaje automático. Actualmente se encuentra laborando como Analista de Riesgos en Allianz México. Junto con Ana María Hernández Piña, realizó un modelo de optimización de inversión basado en años de vida ajustados por discapacidad para el sector salud en México con el objetivo de encontrar el mayor beneficio al menor costo.



ACTUARIOS TRABAJANDO

Retail Pricing Optimization with Monte Carlo Tree Search

Mauricio Hanono Gómez, Celeste Rodríguez Carreón and
Nataliya Kalashnykova



Retail Pricing Optimization with Monte Carlo Tree Search

Mauricio Hanono-Gómez^{*,1}, Celeste Rodríguez-Carreón^{†,2}, and Nataliya Kalashnykova^{†,3}

^{*}Red Bull, [†]Universidad Autónoma de Nuevo León

ABSTRACT Dynamic price changes in the growing retail industry can be determined by geography, inventory, and competition, among other relevant factors. A Monte Carlo Tree Search (MCTS) model is proposed to determine the optimal retail sales price for each day, to maximize profits with limited inventory. The expected units based on the defined price are estimated through a multiple linear regression model, in which random variables are added to make the simulations effective. The experiment is performed in real-time, on three products of a grocery shopping delivery startup, for five days. The results indicate an increase of 12% in net income and 33% increase in units sold within the test days, in comparison to last week.

Key words

Dynamic Pricing, Retail Pricing, Monte Carlo Tree Search, Reinforcement Learning, Markov Decision Process.

All rights reserved © 2024 by the Mexican Colegio Nacional de Actuarios
Last update of the manuscript: Thursday 25th January, 2024
This article is published by the Mexican Colegio Nacional de Actuarios (CONAC) and contains information from a variety of sources. It is a work intended for informational purposes only and should not be construed as professional or financial advice. The vision and comments contained in this document correspond exclusively to its author and, therefore, are unrelated to any public or private entity, including the one for which he currently collaborates. CONAC does not recommend or endorse the use of the information provided in this study. CONAC makes no warranty, express or implied, or representation of any kind and assumes no liability in connection with the use or misuse of this work.



1. INTRODUCTION

Dynamic pricing has become an important practice within the retail industry to respond to market behavior proactively. As the importance of retail grows in the e-commerce industry, pricing software has become more relevant to adjusting prices based on supervised strategies. Commercial pricing strategies can be based on competition, elasticity, cost, and seasonality, among other popular techniques.

The lack of information on any of these factors leads

¹E-mail of the corresponding author: mauriciohanono@gmail.com

to a non-optimal pricing strategy, leaving the cost of opportunity of being able to execute other scenarios with high potential. The same permanence in a single commercial strategy reduces the generation of information for the evaluation of other alternatives. It is possible to approximate the optimal dynamic pricing strategy through simulation and their interaction over time to evaluate alternatives without repercussions.

Machine learning and artificially intelligent algorithms are supporting important decision-making in a wide variety of applications. Progressively, decision-making algorithms are taking advantage of advances in Artificial Intelligence (AI) to solve problems of greater complexity at lower computational cost (Kochenderfer *et al.* 2022).

A dynamic pricing change model is proposed based on Monte Carlo Tree Search (MCTS), to maximize profits based on the location and demographic information of different groups of people, on their willingness to pay for the same product, and adjusting prices according to inventory control and demand behavior over time. The model also seeks to address the problem of minimizing the opportunity cost of abandoning potential pricing strategies.

The MCTS algorithm is a decision-making algorithm that combines the accuracy of a search tree with random sampling through Monte Carlo simulations. It was originally proposed in 2016 in the AI field called Computer Go, and since then, it has become a state-of-the-art technique for Combinatorial Game Theory (Couëtoux *et al.* 2013). The MCTS algorithm solves reinforced learning problems through weighting experienced returns. It is an iterative algorithm that relies on intelligent tree search to find promising states and builds statistical evidence through random sampling, and, therefore, make informed decisions.

As for the state-of-the-art, usage of the method under study, it is worth noting that MCTS has become a novel and unified framework for game artificial intelligence, after being tested in most combinatorial games to seek the ideal play in sequential games with perfect information. Regarding the application of the algorithm for purposes other than non-deterministic games, MCTS has been applied in different contexts, including cybersecurity (Wang *et al.* 2020), airport patrolling (Jain *et al.* 2010), poaching prevention (see (Segler *et al.* 2018)), chemical synthesis with deep neural networks (see Mañdziuk (2018)), vehicle routing with AI planning (see Neto *et al.* (2020)), among other relevant applications (Świechowski *et al.* 2021). Finally, López-Barrientos *et al.* (2022) used a Monte Carlo-

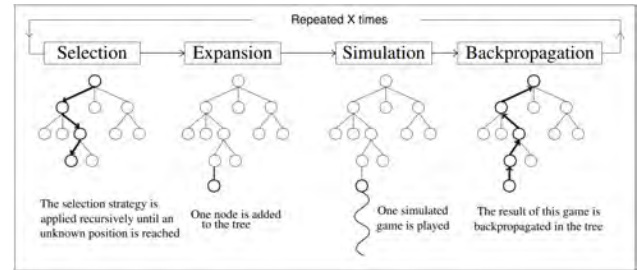


Figure 1 Outline of Monte Carlo tree search.

based approach to study and forecast football matches.

Regarding dynamic pricing, MCTS has been applied to maximizing revenue or utilization for electric vehicle charging station operators (see Mrkos y Basmadjian (2022)). The authors propose a Markov Decision Process (MDP)-based approach and implement it using MCTS to address the large-scale problem. The former approach proves near-optimal pricing decisions in milliseconds, significantly increasing their reward over the flat-rate baseline. The work by Puterman (1994) gives a concise introduction to MDPs.

Likewise, Monte Carlo Simulations have been applied to different fields of study within retail operations. Leepaitoon y Bunterngrachit (2019) estimated the optimal economic order quantity to save inventory cost in a retail store, under the uncertainty of lead time and demand. Taylor *et al.* (2019) developed a practical approach for applying Game Theory combined with Monte Carlo Simulation modeling to retail marketing strategies.

Unlike other contributions, the proposed dynamic pricing model is based on an MDP adapted to the retail industry and solved by the MCTS method. To the best of our knowledge, this is the first time this approach has been applied with dynamic pricing in the retail industry with MCTS. This constitutes our major contribution to the field.

In concrete, for MCTS, a search tree is built, node by node, according to the possible outcomes within the simulations. Naturally, as with the Monte Carlo method, the more iterations, the more certain the statistics are, and therefore, the probability that the best-recommended action is actually optimal increases. MCTS consists of repeated iterations (looped until a predefined stop condition is satisfied), on which each iteration consists of four phases, as shown in Figure 1, which we borrowed from Chaslot (2010).

The remainder of our work is organized as follows.



The next section delves into the description of the model for the retail industry, then in section ??, we present the experiments we performed and the corresponding results are presented in section 4. We give our conclusions in section 5.

2. MODEL DESCRIPTION

Formally, MCTS is applicable to an MDP, which allows modeling a simulated environment of sequential decisions to obtain rewards according to the agent's policy. An MDP is defined as a 4-tuple (S, A_s, P_a, R_a) where S is a finite set of possible states, A_s is a finite set of admissible actions when the process is in state s , $P_a(s, s')$ is the one-step transition probability of going from state s to s' when selecting action a ; and $R_a(s)$ is the immediate reward for reaching state s' by action a (see [Puterman \(1994\)](#)). Now, we describe how we use these elements to implement the MCTS.

- **Selection.** Recall that the state space S_t at time t is defined by the competitive set. Thus, the competitors within the industry are chosen to be considered as a point of comparison. Therefore, it is proposed that the possible states in S at time t , are initially a set of five retail sales price (RSP) options, which include the minimum and maximum, and the quartiles of the distribution of the competing RSP. Additionally, a "win" state is added, which fixes an RSP lower than the minimum competitor by 3%, and a "margin builder" state, which fixes an RSP higher than the maximum competitor by 3%.
To determine which nodes to add to the tree, the popular Upper-Confidence Bound method is used, since it maintains an appropriate balance between exploration and exploitation (see [Auer et al. \(2002\)](#)). This tree policy allows the agent to do both: improve its knowledge about each action in the long term and choose actions to try to obtain the highest reward in the short term.
- **Expansion.** Same as with S_t , the feasible action space A_s can be defined by the competitive set. Thus it is a set of the possible prices in state s' . The proposed limit on S and A is established so as not to iterate on prices that are not profitable, nor on prices that will have no demand at all. To forecast the expected demand with the simulated price changes, a multiple linear regression model is developed, considering elasticity and seasonality. Additionally, random variables that are also statistically significant are added to the model to create a random environment for

the simulations to be effective. These random variables can be related to discount strategies, marketing strategies, and traffic, among other relevant factors. The defined random variables are fitted to a theoretical distribution through a goodness-of-fit test. The Kolmogorov-Smirnov (K-S) test allows a comparison of a sample with a reference probability distribution (one-sample K-S test), and the parameters can be estimated through Maximum Likelihood Estimation (MLE). Afterward, a random simulation of the problem is performed until a terminal state is reached. Two conditions are proposed for a state to be considered terminal: whether the inventory in state $s_t \leq 0$, or $t > n$, where n is the parameter to define the number of days to be simulated.

Finally, a phase of backpropagation enters into place. Each node visited during the simulation is updated. There, the transition matrix P_a for each s is updated during each simulation. The reward function R_a is defined by the objective to be achieved. Since the model proposes to maximize profits, R_a will be a function of the Net Income. That is:

$$\text{Net Income} = \text{GMV} - \text{COGS} - \text{Taxes},$$

where GMV stands for the Gross Merchandise Volume, and COGS is the Cost of Goods Sold.

The pseudocode of the proposed MCTS model is shown in Algorithm 1. Regarding the input values, the root values are the last recorded values of units sold, RSP, GMV, COGS, Taxes, and the added random variables. In addition, the best-performing linear regression model must be provided, as well as the estimators of the fitted theoretical distribution. The output is the path with the highest expected reward, where its child nodes have also the best outcome within the path.

Algorithm 1 MCTS General Pseudocode

Input: Root values

Output: Best Route in the tree

```

while  $i \leq x$  do
   $t \leftarrow 0$ 
  while  $IsTerminalNode = FALSE$  do
    Selection
    Expansion
   $t \leftarrow t + 1$ 

```



3. EXPERIMENTS

The experiment considers a two-month database containing both daily sales and inventory of a grocery shopping delivery startup in Mexico City. We chose three SKUs from different categories so that their rotations behaved differently from each other. These SKUs are tested in only two of the +20 warehouses located in the city. In this way, the performance of the model is evaluated in real-time, compared to a control group of warehouses. The path of RSPs that maximize the expected Net Income over the next five days was calculated with 1,000 iterations each.

- Available States and Actions. Among the selected SKUs, one item belongs to the bread category. Comparing the prices of the same product in other retail stores, a range between \$1.75 and \$2.40 dollars is found. Therefore, both our S and A_s are a set of RSPs within $\{1.65, 1.75, 1.9, 2.05, 2.20, 2.35, 2.45\}$, rounded to the nearest multiple of 0.05 or 0.1 in US dollars.
- MLR and Fitting Random Variables. To develop the best non-supervised performing model regarding the expected units sold, a step-wise selection is made to lower the prediction error. MLR assumptions are validated.

Among the random variables added to the multiple linear regression model is the number of active users per day browsing the website.

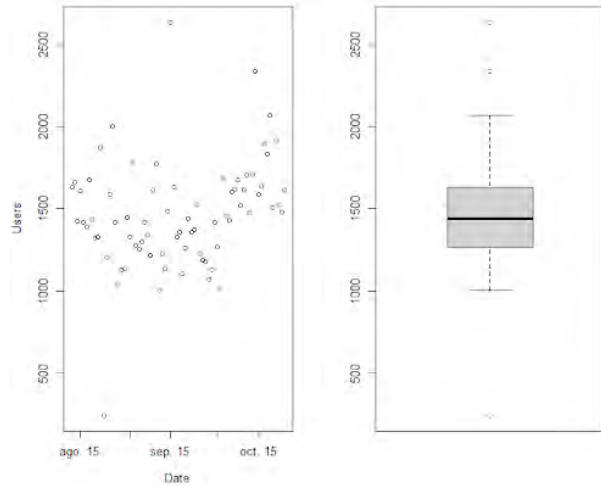


Figure 2 Active Users in two months

In Figure 2, the active users within the study period are plotted. Outliers are removed to continue with the fitting of the empirical distribution.

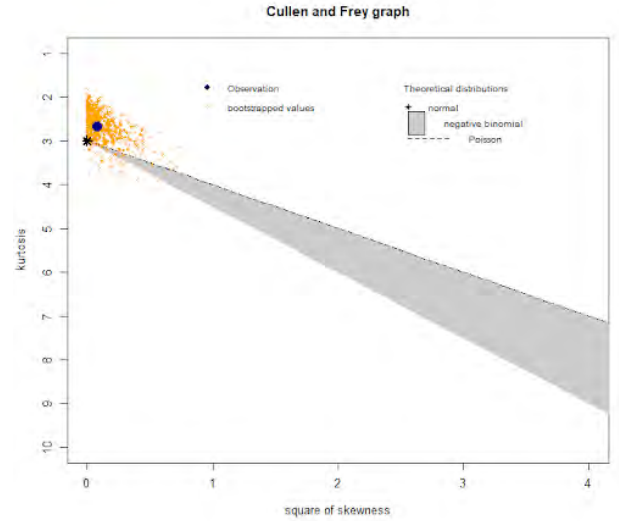


Figure 3 Cullen and Frey graph of Active Users

In Figure 3, performing a supervised analysis on the Cullen-Frey plot, the bootstrapped values suit the theoretical Normal and Poisson distributions.

A one-sample Kolmogorov-Smirnov test is performed to determine whether or not the null hypothesis that the sample comes from a normal distribution is rejected. Likewise, the parameters are estimated through MLEs.

- Before Testing. Once the root values are available, the proposed MCTS model is executed, and compared with a fixed price scenario, considering the initial RSP before trial, in terms of the expected units sold (again with 1,000 iterations) in the following five days. The iterations follow both the MLR model and the simulation of the random variables previously identified. See Figure 4.

Note in Table 1 that the proposed model identifies that the RSP can be slightly lowered on days with lower demand, in order to maximize the Net Income at the end of the five days of simulation. Hence there is an expected Net Income increase of 23%, compared to the fixed-price scenario.



Day	MCTS: RSP	MCTS: Units Sold	Fixed Price: RSP	Fixed Price: Units Sold
1	\$2.35	2	\$2.40	2
2	\$2.35	2	\$2.40	1
3	\$2.45	2	\$2.40	2
4	\$2.45	2	\$2.40	2
5	\$2.35	2	\$2.40	1
Net Income		\$7.45		\$6.05

■ **Table 1** Comparison of Expected Net Income.

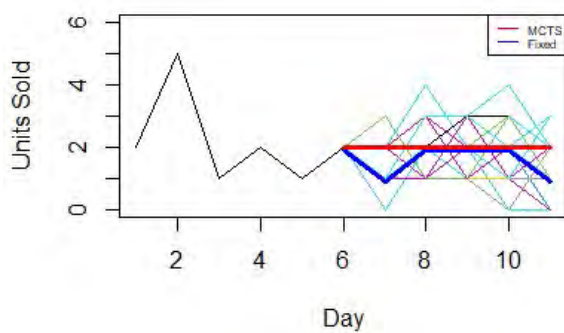


Figure 4 Simulation of Sold Units.

4. RESULTS

As mentioned, it has been proposed to conduct a 5-day test starting on a Monday. Overall, in comparison with the previous week, the model achieved an increase of 33% in units sold and a 12% increase in Net Income, whereas in comparison to the control group, the model achieved an increase of 66% in units sold and a 40% increase in Net Income. See Figures 5-6.

Considering the SKU that belongs to the bread category, the first tested hub, in comparison to the assigned control group, achieved a 51% net income increase and a 75% in units sold. Likewise, the second tested HUB, in comparison to the assigned control group, achieved a 35% net income increase and a 63% in units sold.

5. CONCLUSIONS

In essence, the underlying concept of the proposed dynamic pricing model is to sell the selected product at the

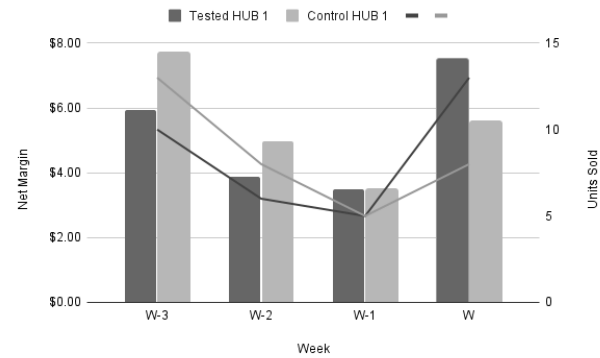
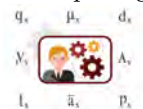


Figure 5 Five-day MCTS Test Results in hub 1.

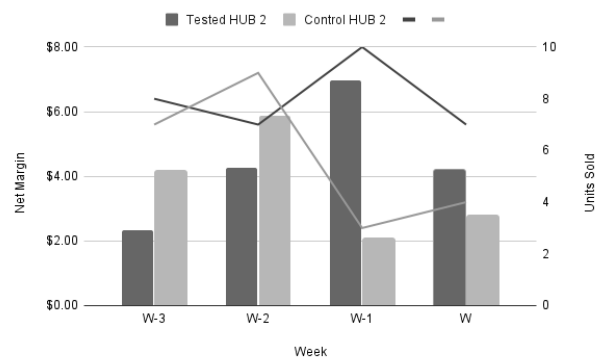


Figure 6 Five-day MCTS Test Results in hub 2.

right price, at the right time, to the right customer; with

the objective of maximizing profits. As the retail industry grows, dynamic pricing models are becoming more promising.

Through statistical and machine learning models, unsupervised and data-driven pricing strategies can be developed. Over time, the model itself will learn from the experience of its own outputs, based on real results. Therefore, the model is expected to iteratively improve over time.

Limitations

Since the proposed model is expected to be executed on a catalog of more than a thousand SKUs, at a warehouse level; the development of the LRM model must necessarily be unsupervised. Regardless of the variable selection method (Forward, Backward, Step-wise), there will be cases in which the RSP variable is not statistically significant for the model. Likewise, there will be cases where the R-squared is not statistically significant. For both cases, it is proposed to define criteria and thresholds in terms of p-values, R-squared, and other statistical values. It is suggested that the SKUs that are withdrawn from the model should have their RSP changed continuously to measure elasticity.

Next steps

The proposed MCTS model can be changed to address larger catalogs efficiently. The increase in simulation days or possible RSPs also increases the computational cost. There is an existing wide area of opportunity in developing dynamic pricing models to achieve commercial objectives within the retail industry.

REFERENCES

- Auer, P., N. Cesa-Bianchi, y P. Fischer, 2002 Finite-time analysis of the multiarmed bandit problem. *Machine learning* **47**: 235–256.
- Chaslot, G. M. J.-B. C., 2010 *Monte Carlo tree search*, volumen 24. Maastricht University.
- Couëtoux, A., M. Müller, y O. Teytaud, 2013 *Monte Carlo Tree Search in Go*. University of Alberta.
- Jain, M., J. Tsai, J. Pita, C. Kiekintveld, S. Rathi, *et al.*, 2010 Software assistants for randomized patrol planning for the lax airport police and the federal air marshal service. *Interfaces* **40**: 267–290.
- Kochenderfer, M. J., T. A. Wheeler, y K. H. Wray, 2022 *Algorithms for decision making*. MIT Press.
- Leepaitoon, S. y C. Bunterngrachit, 2019 The application of Monte Carlo simulation for inventory management: A case study of a retail store. *International Journal of the Computer, the Internet, and Management* **27**: 67–83.
- López-Barrientos, J. D., D. A. Zayat-Niño, E. X. Hernández-Prado, y Y. Estudillo-Bravo, 2022 On the Élö-Runyan-Poisson-Pearson method to forecast football matches. *Mathematics* **10**: 4587.
- Mańdziuk, J., 2018 New shades of the vehicle routing problem: emerging problem formulations and computational intelligence solution methods. *IEEE Transactions on Emerging Topics in Computational Intelligence* **3**: 230–244.
- Mrkos, J. y R. Basmadjian, 2022 Dynamic pricing for charging of EVs with Monte Carlo tree search. *Smart Cities* **5**: 223–240.
- Neto, T., M. Constantino, I. Martins, y J. Pedroso, 2020 A multi-objective Monte Carlo tree search for forest harvest scheduling. *European Journal of Operational Research* **282**: 1115–1126.
- Puterman, M. L., 1994 *Markov Decision Processes. Discrete Stochastic Dynamic Programming*. Wiley, New York.
- Segler, M. H., M. Preuss, y M. P. Waller, 2018 Planning chemical syntheses with deep neural networks and symbolic ai. *Nature* **555**: 604–610.
- Taylor, M., V. Kwasnica, D. Reilly, y S. Ravindran, 2019 Game theory modeling of retail marketing discount strategies. *Marketing Intelligence and Planning* **37**: 555–566.
- Wang, K., A. Perrault, A. Mate, y M. Tambe, 2020 Scalable game-focused learning of adversary models: Data-to-decisions in network security games. *International Conference on Autonomous Agents and Multi-agent Systems* pp. 1449–1457.
- Świechowski, M., K. Godlewski, B. Sawicki, y J. Mańdziuk, 2021 Monte Carlo tree search: A review of recent modifications and applications. *arXiv preprint arXiv:2103.04931*.





Mauricio Hanono-Gómez graduated from the actuarial sciences Bachelor's program of the Universidad Anáhuac México. Prior to JOKR's exit from the Mexican retail industry, he had his first experience there as an analyst. He presented the material contained in this research paper

at the CANWEST conference held in Madrid Spain, in September 2023. The video can be found [here](#).



Celeste Rodríguez-Carreón received her Bachelor's degree in mathematics in 2008, and holds a Doctoral degree in Industrial Physics Engineering, which she obtained in 2013 at the Faculty of Physics and Mathematics Science from the Universidad Autónoma de Nuevo Leon, Mexico. Currently, she is a professor-researcher at the same

Faculty. Her research interests include differential games, stochastic systems, and robust control problems.



Nataliya Kalashnykova has a Doctoral degree in Operations Research and a Master's Degree in Industrial Economics. She is the author of more than 500 published research papers and two monographs. Currently, she is a professor-researcher at the Universidad Autónoma de Nuevo

León in Monterrey, Mexico.



ACTUARIOS TRABAJANDO

Huracán Otis. ¿Cisne Negro o error en los modelos predictivos?

Marco A. De La Rosa Rodríguez



Huracán Otis. ¿Cisne Negro o error en los modelos predictivos?

Marco A. De La Rosa Rodríguez*,¹

*DRASolutions Consulting

RESUMEN El siguiente artículo aborda la vulnerabilidad de México ante eventos naturales catastróficos y el inesperado "Cisne Negro", Otis. De igual forma se examinan y analizan las estadísticas de seguros relacionadas con riesgos naturales, evaluando la efectividad de las pólizas existentes, así como las primeras estimaciones de daños y los esfuerzos de reconstrucción. El artículo concluye con lecciones aprendidas, recomendaciones y la importancia que tienen los seguros para enfrentar estos eventos climáticos.

Palabras clave

Seguros, Eventos Catastróficos, Riesgos hidrometeorológicos

Derechos reservados © 2023 por el Colegio Nacional de Actuarios
Última actualización del manuscrito: 20 de diciembre de 2023
Versión preparada por Maximiliano Tapia Díaz.
Este artículo es publicado por el Colegio Nacional de Actuarios (CONAC) y contiene información de una variedad de fuentes. Es un trabajo cuyo fin es únicamente informativo y no debe interpretarse como asesoramiento profesional o financiero. La visión y comentarios contenidos en este documento corresponden exclusivamente a su autor y, por tanto, son ajenos a cualquier entidad pública o privada, incluyendo aquella para la que actualmente colabora. El CONAC no recomienda ni respalda el uso de la información proporcionada en este estudio. El CONAC no ofrece ninguna garantía, expresa o implícita, ni representación de ningún tipo y no asume ninguna responsabilidad en relación con el uso o mal uso de este trabajo.
¹ Correo electrónico: marcodelarosa@drasolutions.com.mx



1. HURACANES Y RIESGOS DE LA NATURALEZA EN MÉXICO

México y los riesgos de la Naturaleza

De acuerdo con la Organización de las Naciones Unidas, México es uno de los diez países con mayor riesgo catastrófico por fenómenos naturales. Vea la figura 1.

Escala para clasificar huracanes

Es importante mencionar cómo se miden y se catalogan los huracanes. La escala Saffir-Simpson, fue diseñado en 1969 por el ingeniero civil Hebert Saffir y el director del Centro Nacional de Huracanes de Estados Unidos, Robert Simpson. Fue desarrollada para clasificar del 1 al 5, la intensidad de los ciclones tropicales de acuerdo con

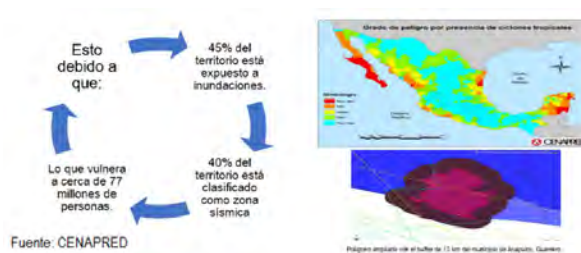


Figura 1 Fuente: Centro Nacional de Prevención de Desastres (2023).

la velocidad del viento, presión y la altura de la marejada. También mide los daños potenciales que estos factores pueden generar en las zonas afectadas. Vea la figura 2.



Figura 2 Fuente: Centro Nacional de Huracanes (2020).

Temporada de Huracanes 2023

Se esperaba una temporada activa, más de lo habitual en el océano Pacífico un 30 % más que el 2022; en el océano Atlántico, se esperaba normal y hasta 10 % por debajo del promedio. Vea la figura 3.



Figura 3 Fuente: Comisión Nacional del Agua (2023).

2. MÉXICO Y LOS EVENTOS DE LA NATURALEZA CATASTRÓFICOS EN CIFRAS

Eventos en México donde las aseguradoras han pagado más

Con base en la información reportada por la Asociación Mexicana de Instituciones de Seguros (AMIS), a lo largo de los últimos años, los siniestros pagados por las aseguradoras son en su mayoría por huracanes. Sin embargo, dentro de los eventos en las que las aseguradoras han pagado más, están los sismos del 19 de septiembre de 1985 y el del 7 de septiembre de 2017, así como la Pandemia derivada de la enfermedad COVID-19, que ocupa el primer lugar. Previamente, este lugar correspondió durante muchos años por los daños pagados ocasionados por el huracán Wilma en 2005.

Dentro de los eventos que las Aseguradoras han pagado más reclamaciones e indemnizaciones tenemos a (Vea la figura 4):



Figura 4 Fuente: Comisión Nacional del Agua (2023).

Riesgos hidrometeorológicos (incluye huracanes, inundaciones, lluvias, granizadas, entre otras...)

- 18 eventos
- 107,091 siniestros pagados
- Monto Siniestros Pagados USD \$7,219,000,000

Terremotos

- Cinco eventos
- 62,125 siniestros pagados
- Monto Siniestros pagados USD \$2,907,000,000

Pandemia (COVID-19)

- 1 evento
- 208,516 familias beneficiadas con un seguro
- Monto Siniestros Pagados USD \$3,472,000,000



Indemnizaciones por terremotos y riesgos hidrometeorológicos Sector Seguros

Del 2017 al 2022 el monto de siniestros retenidos pagados acumulado, de las coberturas de terremoto, erupción volcánica y riesgos hidrometeorológicos, ascendió a \$6,973,000,000. 40 % terremotos y 60 % riesgos hidrometeorológicos. Vea la figura 5 y la tabla 1.

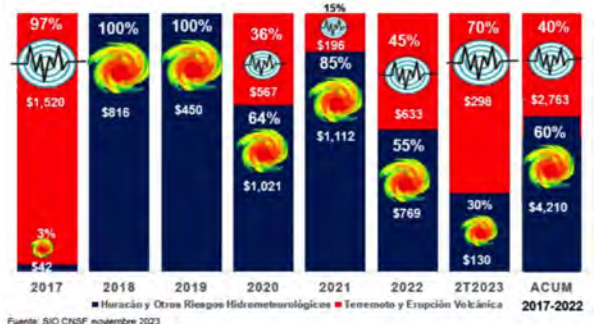


Figura 5 Fuente: **Comisión Nacional de Seguros y Fianzas (2023).**

SISMOS	07S17 y 19S17	2022
	Daños totales por USD \$4,607 mil. de los cuales el Gobierno pagó USD \$2,458 mil. (53.3%) y el Sector Asegurador USD \$2,157 mil. (46.7%)	34 sismos con magnitud mayor a 7.7. Entre 2022 y 2023 se han pagado \$931 mil. (USD \$52 mil.)
RIESGOS HIDRO	A diferencia de los terremotos, cada año hay daños por huracanes, inundaciones, lluvias, entre otros.	Si quitáramos lo pagado por terremoto en 2017, los riesgos hidro tendrían una participación total acumulada al 2022 del 77% y Terremoto

Tabla 1 Fuente: Elaboración Propia.

3. ESTADÍSTICAS DE SEGUROS DE LOS RIESGOS CATASTRÓFICOS DE LA NATURALEZA EN MÉXICO

Cúmulo de sumas aseguradas de riesgos hidrometeorológicos por Giro/Uso 2022

La suma asegurada que reportan las aseguradoras por la cobertura de riesgos hidrometeorológicos en el 2022 fue de \$59,138,000,000,000. Vea la figura 6.

En seis giros/sectores se acumula el 80 % de la suma asegurada:

- Fábricas
- Oficinas
- Casa Habitación
- Bodegas



- Tiendas Departamentales
- Vivienda Crédito Hipotecario



Figura 6 Fuente: **Comisión Nacional de Seguros y Fianzas (2022).**

Cúmulo de sumas aseguradas de riesgos hidrometeorológicos por Giro/Uso 2022–Guerrero

Participación por giro o uso de inmueble del cúmulo de la suma asegurada que representan más del 90 % de la cobertura de riesgos hidrometeorológicos, que contienen los seguros contra daños reportados en la información del 2022. Vea la figura 7.

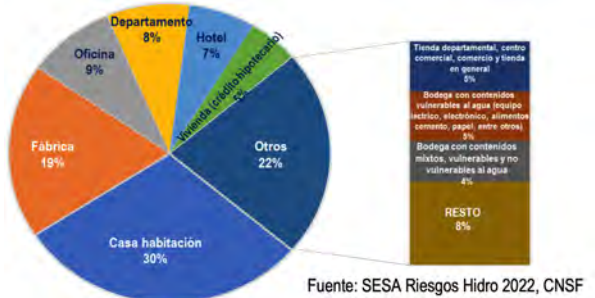


Figura 7 Fuente: **Comisión Nacional de Seguros y Fianzas (2022).**

A diferencia de la información total del mercado, la suma asegurada de la cobertura de riesgos hidrometeorológicos para la entidad de Guerrero, tiene una participación diferente en su composición, el giro/uso que ocupa

el primer lugar es casa habitación (30 %), si lo juntamos con los departamentos tendríamos una participación del 38 %, esto es importante ya que a pesar de la poca penetración del seguro, la suma asegurada se concentra en el giro de casa habitación, seguido de las fábricas (19 %), oficinas (9 %) y hoteles con el 7 %.

Suma asegurada de la cobertura de riesgos hidrometeorológicos por entidad 2022

La Ciudad de México es quien ocupa el primer lugar con el 25 %, en cambio Guerrero está en los últimos lugares con sólo el 0.63 % de participación. Esto es lamentable, ya que una de sus ciudades más importantes (que es Acapulco) sufrió daños severos por Otis. Vea la figura 8.



Figura 8 Fuente: Comisión Nacional de Seguros y Fianzas (2022).

Riesgos Asegurados y Prima Emitida de las Coberturas de Terremoto y Riesgos hidrometeorológicos

Uno de los temas importantes que surgen cuando ocurren este tipo de eventos de la naturaleza que impactan con daños importantes a la población es conocer cuántas casas, negocios, restaurantes, oficinas, hoteles, tiendas de conveniencia, tiendas departamentales, tiendas comerciales y demás giros de negocios que conforman la economía de un estado, municipio o ciudad.

En el argot de los seguros denominamos “riesgos o unidades aseguradas”, en cuanto a la forma de administrar y manejar esta información en las bases de datos de las aseguradoras puede variar, principalmente porque algunos negocios están asegurados de manera individual, es decir, casa por casa, negocio por negocio, hotel por hotel. Sin embargo, en otros casos se agrupan para hacer la administración y gestión de los riesgos de manera más fácil y sencilla, y precisamente aquí es donde surgen las diferencias, hay aseguradoras que en sus sistemas registran la información de forma desglosada y otras de



Figura 9 Fuente: Comisión Nacional de Seguros y Fianzas (2023).

forma agrupada, por lo que el número de póliza, riesgos o ubicaciones puede variar en su número.

El número de riesgos asegurados muestra decremento en ambas coberturas: -66.4 % en riesgos hidrometeorológicos y terremoto del 72 % del 2018 al 2022. Sin embargo, la prima emitida muestra un aumento: para riesgos hidrometeorológicos el incremento es del 41.8 % y para terremoto del 29.7 % del 2022 vs. 2018. Vea las figuras 9 y 10.

Al no contar con la explicación antes mencionada, se puede generar un análisis que puede ser no adecuado y suponer que estas coberturas no han tenido un crecimiento en los últimos años. Observamos en el cuadro que a pesar de que los riesgos o unidades aseguradas tanto de terremoto como de riesgos hidrometeorológicos han decrecido en un 42 % y 48 % respectivamente del 2018 al 2022, en cuanto a la prima emitida se ve un incremento del 42 % para riesgos hidrometeorológicos y del 30 % para terremoto del 2018 al 2022; además, observamos que la prima promedio es de \$4,978 para riesgos hidrometeorológicos y \$5,340 para terremoto en 2022, siendo que en el 2018 la prima promedio era de \$1,118 para terremoto y \$41,116 para riesgos hidrometeorológicos, esto indica que la prima emitida ha crecido a pesar que los riesgos han decrecido.



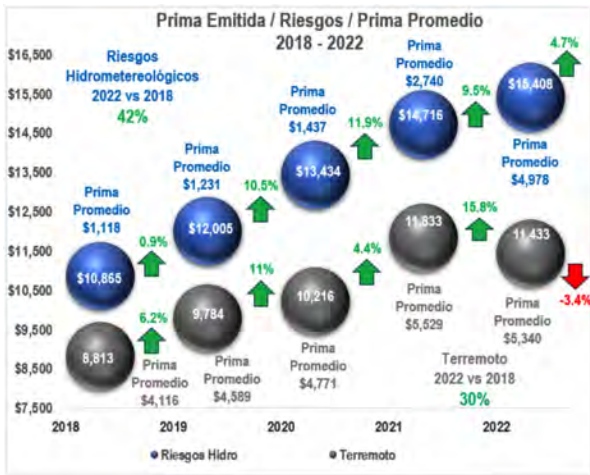


Figura 10 Fuente: Comisión Nacional de Seguros y Fianzas (2023).

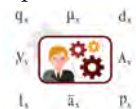
Riesgos y prima devengada retenida de la cobertura de riesgos hidrometeorológicos 2022 por Estado

Referente al número de riesgos y prima devengada retenida de la cobertura de riesgos hidrometeorológicos, las entidades costeras o que tienen ciudades importantes ocupan los primeros lugares, el total del mercado al 2022 fue de \$14,820,000,000 y el 72 % se acumuló en los siguientes estados: Ciudad de México (15.3 %), Quintana Roo (14.3 %), Querétaro (13.9 %), Baja California Sur (6 %), Jalisco (5.3 %), Nuevo León (5.1 %), Estado de México (4.4 %), Veracruz (3.7 %) y Tamaulipas (3.6 %).



Figura 11 Fuente: Comisión Nacional de Seguros y Fianzas (2022).

Guerrero está en el décimo octavo lugar con una participación del 1.4 % que representa una prima devengada



retenida de \$200,000,000 y un total de 55,000 riesgos asegurados, lo que da una prima promedio devengada de \$3,648, una prima relativamente baja considerando que en la información incluye hoteles, grandes negocios, restaurantes, tiendas comerciales, pero concuerda que el mayor número de riesgos asegurados está en casa habitación. Vea la figura 11.

Número y monto de siniestros pagados por riesgos hidrometeorológicos 2022 por estado

Guerrero tiene el lugar 18 en cuanto al pago de siniestros retenidos con un monto de \$17,000,000 distribuidos en 302 siniestros indemnizados, lo que indica que en promedio por cada siniestro pagaron \$56,000; en cuanto a siniestralidad retenido, la entidad cuenta con tan sólo un 8.6 %, lo que indica que por cada \$1,000 devengados y retenidos, las aseguradoras en promedio pagan \$86.

El 2022 fue un año relativamente tranquilo para las aseguradoras en cuanto al pago de daños por terremotos, huracanes, inundaciones, lluvias entre otros riesgos de la naturaleza, ya que en promedio pagaron un monto de siniestro por \$128 de cada \$10,000 de prima cobrada devengada retenida. Vea la figura 12.



Figura 12 Fuente: Comisión Nacional de Seguros y Fianzas (2022).

Siniestros pagados retenidos de la cobertura de riesgos hidrometeorológicos al 2022

De los 32 estados de México, 17 están situados junto a los océanos Atlántico o Pacífico, los golfos de México, Tehuantepec, de Cortés o el Mar Caribe. En esas 17 entidades se acumula el 56 % del total de reclamaciones pagadas durante 2022 de la cobertura de riesgos hidrometeorológicos.

Los estados no costeros que acumulan el 32 % del total de reclamaciones pagadas retenidas son: Ciudad de

México, Querétaro, Coahuila, Estado de México y Nuevo León. Vea la figura 13.



Figura 13 Fuente: Comisión Nacional de Seguros y Fianzas (2022).

Siniestros por causa de la cobertura de riesgos hidrometeorológicos al 2022

Con base en lo reportado por las aseguradoras a la CNSF en el 2022, en tres causas se acumula el 99 % del total de las reclamaciones pagadas durante el 2022. Vea la figura 14.

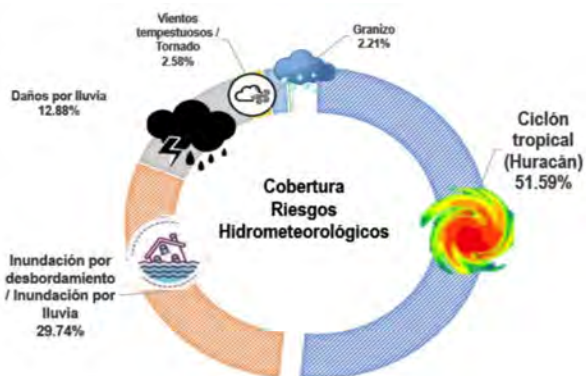


Figura 14 Fuente: Comisión Nacional de Seguros y Fianzas (2022).

- En 2022 se pagaron 16,627 reclamaciones de la cobertura de riesgos hidrometeorológicos.
- Se pagó en promedio un monto de siniestros retenidos de \$216,000.
- Entre huracanes, daños por lluvia e inundaciones se pagó el 85 % del total de las reclamaciones pagadas.

- Inundación por lluvia es la causa que más se pagó por cada siniestro, en promedio se pagaron \$441,000 por cada reclamación.
- El que menos pagó en promedio fueron las avalanchas con un monto de \$21,000 por cada siniestro reclamado.

4. PAULINE (1997) Y MANUEL (2013)

El huracán Pauline (1997)

Tocó tierra en el sur de México el 8 de octubre de 1997. Se trató de uno de los huracanes más mortíferos, destructivos y costosos durante la segunda mitad del siglo XX. Pauline fue el tercer huracán más intenso de la temporada de 1997.

- Estimación de muertes: más de 300 entre Guerrero y Oaxaca.
- Daño económico total: \$448,000,000 (cifra a 1997) - USD \$916,000,000 (actualizado sep. 2023).
- Daño económico con seguro: USD \$92,000,000 (actualizada sep. 2023) / Representa 10 % del daño económico total.
- FONDEN: USD \$824,000,000 (actualizados sep. 2023) / Representó el 90 % del daño económico total.

El huracán Manuel (2013)

Fue el decimosexto ciclón tropical formado durante la temporada de huracanes en el Pacífico de 2013, así como el séptimo de ellos en alcanzar la categoría 1 de huracán. Fue el más destructivo de dicha temporada tras afectar seriamente el estado de Guerrero.

- Estimación de muertes: 169 / 10,500 viviendas dañadas / 510 escuelas afectadas / 35 unidades de salud dañadas.
- Daño económico total: \$2,270,000,000 (cifra al 2013) - USD \$2,560,000,000 (actualizado sep. 2023).
- Daño económico con seguro: USD \$248,000,000 (cifra al 2013) - USD \$309,000,000 (cifra 2023) / Representa 12 % del daño económico total.
- FONDEN: USD \$2,251,000,000 (actualizados sep. 2023) / Representó el 88 % del daño económico total.

A diferencia de los huracanes Wilma 2005 (Cancún) y Odile 2014 (Los Cabos y La Paz), la baja penetración del seguro en el estado de Guerrero influyó para que las Aseguradoras tuvieran una participación promedio del 12 % en cuanto a los pagos por los daños ocasionados por Pauline y Manuel en sus respectivos años.



- En ambos huracanes la participación del sector asegurado fue mínima, Pauline 1997 – 11.9 % y Manuel 2013 – 14.2 %.
- De acuerdo a la información extra oficial los daños económicos para el huracán Pauline (1997) fueron de USD \$7,486,000,000 y el huracán Manuel (2013) USD \$4,941,000,000.
- Los daños ocasionados por el huracán Manuel casi cuatuplicaron las registradas durante el huracán Pauline, a pesar de que durante Manuel se registraron menos lluvias diarias.
- Esto explica que en 2013 haya existido una mayor exposición de personas e infraestructura que con respecto a 1997 y, por lo tanto, mayor vulnerabilidad.

Vea la tabla 2 y la figura 15

Evento	Año	Daños totales 2023 millones USD	Seguro Daños 2023 millones USD	FONDEN 2023 millones USD
Huracán Pauline %	1997	\$916 26.3%	\$92 10%	\$824 90%
Huracán Manuel %	2013	\$2,560 73.7%	\$309 14%	\$2,251 86%
TOTAL %		\$3,476 12%	\$401 12%	\$3,075 88%

■ **Tabla 2** Fuente: Elaboración Propia.



Figura 15 Fuente: Fondo de Desastres Naturales (2023).

5. ¿UN «CISNE NEGRO»?

Mucho sea cuestionado si es que el huracán Otis fue un evento esperado o uno denominado como “Cisne Negro”, la cual es una metáfora que, en el ámbito económico, describe aquellos sucesos que ocurren por sorpresa, que ningún analista había previsto ni tenido en cuenta porque, a priori, eran improbables y que, para bien



o, generalmente, para mal, terminan teniendo un gran impacto y repercusiones trascendentales. Para que un acontecimiento se considere «Cisne Negro», debe contemplar:

1. **Que sea inesperado.** Se trata de un hecho a priori improbable, para el que no hay ninguna evidencia de que vaya a suceder y que, por tanto, es una sorpresa para los analistas y para el mercado.
2. **Tienen un gran impacto.** Son acontecimientos que afectan de forma importante a la economía o a la política mundial.
3. **Se caracterizan por tener predictibilidad retrospectiva.** Es decir, una vez que han sucedido, y sólo entonces, se dan evidencias de que dicho hecho se podía haber evitado y se crean teorías que explican por qué se llegó a producir.

Para ejemplificar, así como Otis, los siguientes eventos han sido considerados «Cisnes Negros» por su impacto y característica de irreversibles:

- 11/9 en Nueva York
- Crisis de las Puntocom
- Crisis COVID-19

Estos eventos inesperados tuvieron efectos duraderos, planteando la pregunta de cómo las organizaciones y naciones pueden navegar un mundo lleno de sorpresas inesperadas.

Trayectoria esperada del huracán Otis 2023

De acuerdo a las estimaciones realizadas por Tropical Storm Risk, el huracán Otis impactaría las costas Mexicanas como huracán tipo 3 degradándose a huracán tipo 1 y tormenta tropical. Sin embargo, la realidad fue otra, pues la violencia del huracán Otis fue inesperada. Escaló de tormenta tropical a evento de categoría cinco en menos de doce horas. Vea la figura 16.

Trayectoria real del huracán Otis 2023

Nadie detectó a tiempo que Otis se convertiría en la peor tormenta de los últimos 30 años en el Pacífico. Los modelos no predijeron la intensificación del fenómeno con la antelación usual, lo que dejó a Acapulco, Guerrero, una región mexicana densamente poblada, con poquísim tiempo para prepararse ante la catástrofe, ¿qué falló?

Jorge Zavala Hidalgo, director e investigador del Instituto de Ciencias de la Atmósfera y Cambio Climático en la UNAM, señala del huracán Otis:

Intense Hurricane OTIS: Current Data									
Time		Position		Strength		Wind Probabilities		Wind Model	
GMT	Lead	Lat	Long	Peak Wind	Cat	TS	Cat 1 or Above	Wind Field	
25 Oct, 0:00	0 hrs	15.7 N	99.6 W	125 kts	4	Click Here	Click Here	Click Here	

FORECAST DATA									
Time		Position		Strength		Wind Probabilities		Wind Model	
GMT	Lead	Lat	Long	Peak Wind	Cat	TS	Cat 1 or Above	Wind Field	
25 Oct, 6:00	6 hrs	16.3 N	100.0 W	140 kts	5	To 6 hrs	To 6 hrs	Click Here	
25 Oct, 18:00	18 hrs	17.5 N	100.8 W	90 kts	2	To 18 hrs	To 18 hrs	Click Here	
26 Oct, 6:00	30 hrs	18.1 N	101.5 W	50 kts	1	To 30 hrs	To 30 hrs	N/A	
26 Oct, 18:00	42 hrs	18.4 N	102.1 W	30 kts	0	To 42 hrs	To 42 hrs	N/A	

Forecast lead times used by TSR are referenced to the current data time. These lead times differ from those used by the US National Hurricane Center which are referenced to the current data time minus 3 hours.

Figura 16 Fuente: Fondo de Desastres Naturales (2023).

- Fue categoría 5 y los fenómenos de este nivel son pocos y raros.
- El huracán tocó tierra con la más alta clasificación, cuando lo usual es que estos fenómenos alcancen la máxima categoría lejos de las costas y se debiliten al entrar al continente.
- Cambió tanto en tan poco tiempo.
- En 2023 vimos situaciones similares. Por ejemplo, en Florida vimos a Idalia convertirse rápidamente en un monstruo.

La realidad de Guerrero

Independientemente de la fuerza destructiva con la que impactó el huracán Otis la costa de Guerrero, hubo varios temas de prevención y previsión, así como de avisos oportunos por parte de las autoridades para que la gente se preparara y resguardara antes de la llegada del huracán. Existen factores económicos y sociales ya desde hace muchos años que conllevan a que la reconstrucción de Guerrero, por los daños ocasionados por Otis, llevarán varios meses, ya que, por lo regular, a nivel internacional, desde el punto de vista de seguros, el ajuste lleva de un año y medio a dos años para finalizar la atención de los daños. Vea la figura 17.

6. ¿ESTABAN BIEN ASEGURADOS?

Guerrero tiene una penetración y cultura del seguro de las más bajas en México, esto debido a que la gente no tiene el poder adquisitivo para contratar un seguro tradicional, que, a pesar de no tener un alto costo pa-



Figura 17 Fuente: Instituto Nacional de Estadística y Geografía (2023).

ra asegurar una casa, no es viable para la gente de la entidad.

Los datos duros del Seguro en Acapulco, Guerrero

Guerrero participa con el 2.8 % del total de población que vive en México y Acapulco participa con el 22 % del total de población que vive en Guerrero. Vea la figura 18.



Figura 18 Fuente: Instituto Nacional de Estadística y Geografía (2023).

- Sólo el 13 % de población que vive en Guerrero cuenta con un seguro de vida, 460,000 personas. A nivel nacional representan el 0.37 %.
- En cuanto al tema de salud, en Guerrero hay 1,630 establecimientos de salud, de los cuales 92 % es de consulta externa; 1.99,000,000 de personas se atienden en el seguro popular, 445,000 en el seguro social y 435,000 consultorios de farmacias.
- Siendo un estado con poca infraestructura de salud, se esperaría una mayor penetración en el seguro de gastos médicos mayores, pero sólo 12 % de los Guerrerenses tienen este seguro.
- El parque vehicular en circulación en México es de 55,000,000, de los cuales Guerrero está en el décimo



primer lugar, con 1.5,000,000 de vehículos. A pesar de que el seguro de responsabilidad civil es obligatorio en México, tan sólo el 32 % de los autos están asegurados, de los cuales Guerrero participa con el 7.2 %.

- Acapulco cuenta tan sólo con 20,000 vehículos asegurados lo que representa el 1.4 % del parque vehicular de Guerrero.

Siendo de los estados con más pobreza, a pesar de que cuenta con uno de los centros turísticos más famosos de México, el nivel de aseguramiento es muy bajo, tanto en el estado como en el puerto de Acapulco.

- En México, aprox. el 10 % de las casas tienen seguro privado, el 25 % tienen un seguro contra daños ligado a un crédito hipotecario, esto implica que unos 12,000,000 de casas tienen seguro.
- De acuerdo con la información del INEGI en Guerrero hay 942,000 casas, y con base en la información de la CNSF, 41,000 tienen seguro, esto representa el 4.3 % del total de viviendas de Guerrero.
- Acapulco tiene 16,000 viviendas aseguradas, de acuerdo a lo reportado por la AMIS, mismas que representan el 40 % de las viviendas aseguradas en Guerrero. Sin embargo, tan sólo representan el 7.2 % de las viviendas totales en Guerrero, lo que indica que la mayor parte de la población que sufrió daños en su vivienda, dependerán de la ayuda del Gobierno y la sociedad.

De acuerdo a la Cámara Mexicana de la Industria de la Construcción, en Acapulco el 100 % de los hoteles tuvieron algún daño; De acuerdo al DENUE del INEGI, en México hay 23,000 hoteles, de los cuales Guerrero tiene el 4.9 % (1,118), y Acapulco cuenta con 387 mismos que representan el 35 % del total en Guerrero.

En cuanto al tema del seguro:

- La CNSF reporta que a nivel nacional hay 15,000 hoteles con algún seguro, lo que representaría el 65 % del total de hoteles registrados en México.
- De acuerdo a la AMIS, los hoteles Gran Turismo entre el 90 % al 95 % cuenta con un seguro o alguna protección económica. Sin embargo, los medianos y pequeños hoteles pueden, o no, contar con un seguro, esto represente que entre el 50 % y 70 % de dichos hoteles pueden tener seguro, lo que implicaría que 233 hoteles podrían hacer la reconstrucción debido al seguro.



- Cerca del 80 % de las unidades de negocio quedaron con daños severos, de los cuales menos del 30 % contaban con un seguro o alguna garantía económica y financiera para hacerle frente a los daños ocasionados por Otis.
- Con base en la información que da el DENUE del INEGI, en México hay 2.4,000,000 de unidades de negocio, de las cuales 1.6 % (37,000) están en Acapulco.
- Con base en lo que se ha comentado en diversos de comunicación, del total de negocios de Acapulco se estima que unos diez mil (27 %) tienen seguro y 27,000 no cuentan con alguna protección económica, por lo que será importante la ayuda del Gobierno y la población para que reconstruyan su negocio y lo echen a andar nuevamente para que se reactive la economía de la zona.
- En cuanto a las tiendas, sin importar su tamaño, actualmente en Acapulco hay registradas aproximadamente unas seis mil, lo que representa tan sólo el 0.81 % del total de tiendas que hay en México, una penetración casi nula, lo cual va de acuerdo con la falta de poder adquisitivo del estado de Guerrero.
- Referente al aseguramiento, es muy similar al tema de las unidades de negocio, de las seis mil tiendas, se estima que unas 1,700 tengan seguro o alguna protección económica, mientras que el 73 % restante tendrán que hacer uso de recursos propios para la reconstrucción de su negocio.

7. LA RECONSTRUCCIÓN Y PRIMERAS ESTIMACIONES DE LOS DAÑOS

¿Cuánto tiempo llevará la reconstrucción de los daños ocasionados por Otis?

Aún es muy rápido para saber la cifra exacta de los daños económicos que ocasionó Otis. Sin embargo, tanto el Gobierno Federal, expertos en temas de modelación de pérdidas y reaseguradores, han publicado en diferentes medios de comunicación sus estimaciones, mismas que podrían ser similares a los daños ocasionados por otros huracanes como Wilma (2005), Odile (2014) o los sismos del 2017.

Según la iniciativa privada, Acapulco tardará dos años en ser reconstruido, este periodo concuerda con los tiempos estimados en el que el sector asegurador ha participado en estos eventos. El costo aproximado de del pago de daños oscila entre \$1,800,000,000 y \$33,000,000,000.

- Autos: 4 %

- Hogar: 7 %
- Hoteles: 60 %
- Negocios y tiendas: 27 %

Estimaciones de los Daños Totales y Asegurados: Gobierno, AMIS, ERN, Reaseguradores y Modeladores de Pérdidas de Riesgos

En este tipo de eventos, en este caso el impacto ocasionado por el huracán Otis, tanto el Gobierno a través de la Secretaría de Hacienda, CNSF, la AMIS y modeladores de riesgos comienzan a emitir las estimaciones de los daños y pérdidas materiales ocasionadas por el evento. Vea la tabla 3.

ENTIDAD	PERDIDA TOTAL		SEGURO	
	\$ (millones)	USD (millones)**	\$ (millones)	USD (millones)**
 Secretaría de Hacienda y Crédito Público	\$61,000	\$3,400	Programa FONDEN	\$17,000 / \$945
			Seguro CAT:	\$5,000 / \$278
			CAT Bond:	\$8,730 / \$485
 Asociación Mexicana de Instituciones de Seguros	NO APLICA		\$11,400	\$833
 Gallagher Re	* \$76,000 a \$117,000	\$4,223 a \$6,500	\$22,000 a \$33,000	\$1,200 a \$1,800
 Moody's	NO APLICA		** \$45,000 a \$81,000	\$2,500 a \$4,500
 Verisk	NO APLICA		** \$54,000 a \$108,000	\$3,000 a \$6,000

Notas: * NO incluye pérdidas por embarcaciones, autos, infraestructura del Gobierno de Guerrero
** NO incluye Pérdidas: por mercancías cíclicas costeras, en el transporte marítimo interior, en el transporte marítimo de carga y cascos, y en embarcaciones y yates de recreo, propiedades no aseguradas, a la infraestructura, por obligaciones extraccontractuales, por impunidad de desechos peligrosos, vandalismo o corrupción civil, ya sea directa o indirectamente causada por el evento y Gastos de ajuste de pérdidas
*** Tipo de cambio \$18 por dólar

■ **Tabla 3** Fuente: Elaboración propia.

Uno de los primeros modeladores que emiten sus estimaciones es ERN (Evaluación de Riesgos Naturales), empresa con muchos años de experiencia en modelos los daños y pérdidas ocasionadas por fenómenos naturales; De la misma forma el Gobierno, a través de Hacienda y CNSF emiten las estimaciones de los daños totales y con seguro, de igual forma la AMIS comunica el anticipo y finiquitos de pagos de las reclamaciones que llevan al momento de emitir su comunicado y los reaseguradores y modeladores de daños de pérdidas hace lo propio.

Derivado de esto, Hacienda, AMIS, ERN, Gallagher, Moody´s, Verisk emitieron ya sus primeras estimaciones, en el siguiente cuadro observamos cómo cada uno estima montos muy diferentes, destacando el del Gobierno ya que mencionan que los daños económicos podrían estar arrina de los \$61,000,000,000, lo que se traduce en USD \$3,400.

ERN informó en un Seminario (vía video conferencia) realizado el 9 de noviembre 2023, que las pérdidas económicas de los daños ocasionaos por Otis (Casas, Comercios, Industria y Hoteles), mismas que pagarán las aseguradoras, oscilan entre \$22,000,000,000 y \$33,000,000,000 (USD \$1,200,000,000 y USD \$1,800,000,000) no incluye, Autos, Barcos e Infraestructura.

Se estima que la pérdida total sea de unas cuatro o cinco veces la estimación de la pérdida asegurada. Por su parte, los modeladores de pérdidas de daños comentaron que:

Moody's RMS:

- Las pérdidas aseguradas en el mercado privado a causa del huracán Otis oscilan entre USD \$2,500,000,000 (\$45,000,000,000) y USD \$4,500,000,000 (\$81,000,000,000).
- Su estimación refleja daños a la propiedad y pérdidas por interrupción de negocios en líneas de negocios residenciales, comerciales, industriales y automotrices, y considera el potencial de amplificación de pérdidas post-evento, tendencias inflacionarias y fuentes de pérdidas no modeladas, incluyendo daños a la infraestructura.

Verisk:

- Estima las pérdidas aseguradas entre USD \$3,000,000,000 (\$58,000,000,000) y USD \$6,000,000,000 (\$108,000,000,000).
- En la estimación de pérdidas aseguradas de la industria se incluyen las sufridas por propiedades residenciales, comerciales e industriales en tierra y automóviles por su construcción, contenido y cobertura de elementos de tiempo por inundaciones inducidas por el viento y las precipitaciones.

Lo que es una constante entre las modeladoras es que es un evento sin precedentes y que los daños son cuantiosos, tanto las pérdidas con seguro y los daños totales.

Por otro lado, la Secretaría de Hacienda y Crédito Público informó que estiman la pérdida total de los daños en aproximadamente \$61,000,000,000 (USD \$3,400,000,000); Uno de los temas importantes son las herramientas financieras con las que cuenta el Gobierno Federal para afrontar estos daños, mismas que se hablarán más adelante.

AMIS, emitió un primer comunicado el 16 de noviembre de 2023 a través del cual informó que la primera estimación de los daños ocasionados por Otis que pagarán las Aseguradoras era de \$11,400,000,000/USD \$633,000,000. El 23 de noviembre publicó un segundo



comunicado a través del cual informaron la actualización de las estimaciones y anticipos de los pagos de los daños que cuentan con algún seguro contra daños o autos.

La actualización de los daños a pagar por las Aseguradoras aumentó un 77 %, pasando:

- Primer reporte: \$11,400,000,000 / USD \$633,000,000
- Segundo Reporte: \$20,141,000,000 / USD \$1,119,000,000

Derivado de las reclamaciones de daños a casas, autos, hoteles, negocios, restaurantes, tiendas departamentales, entre otros giros:

Las primeras valuaciones estiman \$950,000,000

Reclamaciones:

- Se aumentó el pago de siniestros: 12,035 reclamaciones
- 7,096 (44 %) - Relacionadas con algún Seguro contra Daños
- 9,032 (56 %) – Relacionado con el Seguro de Autos

Hoteles:

- 86 reclamaciones de hoteles
- El pago de anticipos aumentó a \$806,000,000
- Estimación total pérdida aumentó a: \$3,981,000,000

Negocios:

- Pagos de anticipos en 1,983, empresas, centros comerciales, restaurantes y actividades económicas aseguradas

Hogares:

- Sin Cambios: 3,891 reclamaciones
- De cada \$300 pagados \$100 corresponde a Seguro de Hogar

Embarcaciones:

- Registros Daños a 190 Embarcaciones Aseguradas

Con la actualización de a estimación de los daños ocasionados por el huracán Otis, AMIS posiciona este huracán como el evento #6 de los primeros #15 en los que el Sector Asegurador ha participado con pagos importantes; el evento en el cual la aportación de las aseguradoras ha sido importante:

1. Pandemia (COVID-19) 2019-2021 con USD \$3,472,000,000,
2. Wilma 2005 con USD \$2,675,000,000,



3. Sismo de 2017 con USD \$1,449,000,000,
4. Odile de 2014 con USD \$1,439,000,000,
5. Gilberto de 1998 con USD \$1,299 y 6. Otis de 2023 con una estimación de la pérdida por USD \$1,119,000,000.

Eventos donde las aseguradoras han pagado más en México

La gráfica muestra los primeros 15 eventos en los que las aseguradoras han participado con la pérdida de pagos millonarios. El primer lugar ocupa la Pandemia (COVID-19), seguido del huracán Wilma (2005) y el sismo del año 2017. Con base en el segundo comunicado publicado por la AMIS (23/11/2023), en el que informan el avance de pagos y atención de siniestros, el huracán Otis estaría ocupando el lugar seis, subiendo dos lugares después del informe anterior. Vea la figura 19.



Figura 19 Fuente: Asociación Mexicana de Instituciones de Seguros (2023).

Con esto se muestra una vez más que ante fenómenos como el huracán Otis, el seguro se convierte en una herramienta clave para que familias, empresas y el gobierno tengan la capacidad de salir adelante de manera más rápida y eficiente, para la reactivación de la actividad económica y reconstrucción de los centros de trabajo de Acapulco, centro turístico de gran importancia en México.

¿De dónde saldrá el dinero si no hay FONDEN?

Mucho se ha especulado que derivado a que ya no hay el fideicomiso del FONDEN, ¿de dónde saldrá el dinero para la reconstrucción? Hay que aclarar que el fideicomiso del FONDEN fue el que se extinguió. Sin embargo, es importante mencionar que se extinguió el fideicomiso y en su lugar se creo el programa "FONDEN", a través del cual es que se asignará y administrará dinero para la reconstrucción.

«FONDEN dejó de ser un fideicomiso y ahora es un programa que está en el Presupuesto de Egresos de la Federación»

- El presupuesto aprobado fue de \$17,156,826,205.00 pesos. [Fondo de Desastres Naturales \(2023\)](#).

Adicional al Programa FONDEN, el Gobierno Federal cuenta con protección:

- **Seguro Catastrófico:** En los Criterios Generales de Política Económica (CGPE) para el 2023, la Secretaría de Hacienda y Crédito Público (SHCP) expone que el seguro catastrófico tendrá una cobertura de \$5,000,000,000 de pesos con una vigencia del 5 de julio de este año al 5 de julio del 2023.
- **Bono catastrófico:** Lanzado en el 2020 en conjunto con el Banco Mundial, tiene una cobertura de hasta USD \$485,000,000 y provee respaldo contra pérdidas derivadas de sismos de diferentes magnitudes y ciclones tropicales.

8. LECCIONES APRENDIDAS, RECOMENDACIONES Y LOS SEGUROS ANTE ESTOS EVENTOS

Lecciones aprendidas

Otis pasará a la historia como uno de los eventos climáticos más impactantes que México haya enfrentado en décadas. Ha dejado valiosas lecciones, enfatizando la importancia de la preparación, la mitigación de desastres y la resiliencia comunitaria, además de contar con herramientas financieras y económicas para hacer frente a los daños, si, hablamos de los Seguros.

Recomendaciones

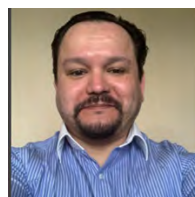
Lo primero e inmediato es ayudar a la población de Guerrero, en específico la de Acapulco, y en seguida o a la par se dará el tema de la reconstrucción, lo cual es necesario para reactivar la economía del puerto de Acapulco; posteriormente se deben de atender varios temas relacionados con construcción, previsión, prevención, modelaje de los huracanes, simulacros de evacuación y la penetración del Seguro en el estado y la zona de Acapulco.

¿Y los seguros ante estos eventos?

Otis ha demostrado la importancia de estar asegurados, ya que en tan sólo unos instantes puedes perder tu patrimonio que te ha llevado años obtenerlo, cuidarlo e invertir en él; recuerda que el seguro protege el patrimonio de las familias y las empresas.

REFERENCIAS

Asociación Mexicana de Instituciones de Seguros, 2023 AMIS, Disponible [aquí](#).
 Centro Nacional de Huracanes, 2020 Departamento de Comercio de los Estados Unidos, Disponible [aquí](#).
 Centro Nacional de Prevención de Desastres, 2023 Sistema Nacional de Protección Civil, Disponible [aquí](#).
 Comisión Nacional de Seguros y Fianzas, 2022 Gobierno de México, Disponible [aquí](#).
 Comisión Nacional de Seguros y Fianzas, 2023 Gobierno de México, Disponible [aquí](#).
 Comisión Nacional del Agua, 2023 Gobierno de México, Disponible [aquí](#).
 Fondo de Desastres Naturales, 2023 Disponible [aquí](#).
 Instituto Nacional de Estadística y Geografía, 2023 INEGI, Disponible [aquí](#).



Marco de la Rosa Rodríguez realizó sus estudios de Actuaría en la Facultad de Ciencias de la UNAM. Cuenta con 26 años de experiencia en el sector asegurador, de los cuales los ha dedicado a desarrollar proyectos de seguros para la expansión del mercado, para la inclusión financiera

en seguros de la sociedad mexicana, a través de la transformación de la información en datos útiles para la toma de decisiones tanto en las aseguradoras, agentes de seguros y la sociedad mexicana. Ha tenido múltiples experiencias laborales en la industria de los seguros. Actualmente presta sus servicios profesionales en la consultoría DRA-Solutions Consulting, firma de actuarios consultores especialistas en riesgos, soluciones financieras y de seguros. Ha impartido conferencias internacionales sobre temas relacionados con el seguro de casa habitación, coberturas catastróficas y transporte de mercancías, así como entrevistas relacionadas con los seguros de daños catastróficos para medios nacionales e internacionales.





**COLEGIO
NACIONAL DE
ACTUARIOS, A.C.**



¡Conviértete en miembro del
CONAC para que impules tu
carrera profesional!

La Actuaría es esencial para el país porque proporciona las herramientas y el conocimiento necesarios para gestionar riesgos financieros, garantizar la sostenibilidad de los sistemas de seguros y pensiones, y contribuir a la estabilidad económica. **El Colegio Nacional de Actuarios A.C. (CONAC)** representa desde hace 57 años al gremio actuarial mexicano ante autoridades, organismos empresariales y asociaciones profesionales internacionales.

Requisitos para ser miembro:

- Llena la solicitud de inscripción en nuestra página.
- Cubre la cuota de recuperación.
- Sube en PDF la siguiente documentación escaneada:
- Cédula Profesional
- Título
- CURP
- RFC (Constancia de Situación Fiscal actualizada)
- Estudiantes: con credencial vigente de la Universidad
- Comprobante de Pago

También te invitamos a formar parte de nuestros Comités y/o secciones profesionales:

- Investigación y desarrollo actuarial
- Certificación
- Certificación en Pasivos Laborales
- Educación continua
- Asuntos Académicos
- Relaciones internacionales
- Peritajes Actuariales Oficiales
- Profesionalismo
- Ciencia de Datos
- Capital humano

Síguenos en nuestras redes



conacoficial



conac_mx

WWW.CONACMEXICO.ORG.MX



conacmexico



conacmexico

INVITADOS TRABAJANDO

Modelando Datos con Palitos y Bolitas: Algoritmo de Mapper

Salvador Mancilla, Eymard Hernández López y Giovanni Wences



Modelando Datos con Palitos y Bolitas: Algoritmo de Mapper

Salvador Mancilla^{*,1}, Eymard Hernández-López^{†,2}, and Giovanni Wences^{‡,3}

^{*}Departamento de Matemáticas, Instituto Tecnológico de Estudios Superiores de Monterrey, [†]Posgrado en Ingeniería Industrial, TecNM-TESOEM, [‡]Escuela Superior de Matemáticas No.2, Universidad Autónoma de Guerrero

RESUMEN En el presente trabajo, se expone una herramienta de análisis topológico de datos, específicamente, el algoritmo de Mapper, el cual fue introducido por Singh *et al.* (2007). Se presenta una exposición detallada de los fundamentos algorítmicos, con énfasis en su implementación computacional. Adicionalmente, se ilustra el uso práctico del algoritmo mediante la presentación de ejemplos concretos desarrollados en el lenguaje de programación Python.

Palabras clave

Preimagen inversa, métrica, grafo, agrupamiento de datos.

1. INTRODUCCIÓN

El motivo del siguiente artículo es mostrar una de las tantas herramientas que hay en el análisis topológico de datos. Así que antes de comenzar a dar un breve repaso histórico y mencionar sus aplicaciones en diversos artículos de investigación del algoritmo que hablaremos en este trabajo, intentemos dar un poco de contexto a la situación. Imaginemos que tenemos un conjunto de datos que representa la estatura y el peso de un grupo de estudiantes. Así que podemos pensar que esta información de cada estudiante viene codificada en pares ordenados, es decir, en elementos de la forma (estatura, peso). Posteriormente,

Derechos reservados © 2023 por el Colegio Nacional de Actuarios
Última actualización del manuscrito: 14 de diciembre de 2023
Este artículo es publicado por el Colegio Nacional de Actuarios (CONAC) y contiene información de una variedad de fuentes. Es un trabajo cuyo fin es únicamente informativo y no debe interpretarse como asesoramiento profesional o financiero. La visión y comentarios contenidos en este documento corresponden exclusivamente a su autor y, por tanto, son ajenos a cualquier entidad pública o privada, incluyendo aquella para la que actualmente colabora. El CONAC no recomienda ni respalda el uso de la información proporcionada en este estudio. El CONAC no ofrece ninguna garantía, expresa o implícita, ni representación de ningún tipo y no asume ninguna responsabilidad en relación con el uso o mal uso de este trabajo.



¹Correo electrónico: salvador_mank@tec.mx

²Correo electrónico: eyuard.hl@otemexico.tecnm.mx

³Correo electrónico: giovanniwences@uagro.mx

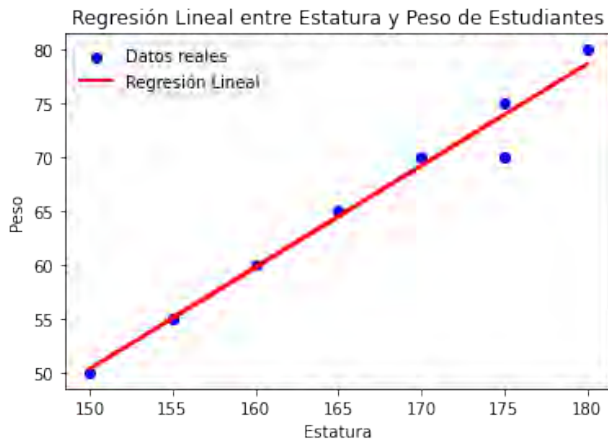


Figura 1 Diagrama de dispersión de la estatura y peso de cada estudiante y recta obtenida mediante regresión lineal.

se procede a representar gráficamente estos elementos en el plano con el objetivo de discernir la configuración inherente del conjunto de datos. La visualización resultante se presenta en la Figura 1, proporcionando una perspectiva visual que busca esclarecer la forma y estructura del conjunto de datos en cuestión.

La forma que presentan los datos en el plano nos sugiere que el conjunto de datos se podría modelar mediante una recta (¡aunque no sea exactamente una recta!). Es decir, que nuestros datos tienen la forma muy parecida a una recta, con lo cual sabemos desde nuestros cursos de bachillerato que una manera de representar esa recta es mediante la forma pendiente-intersección,

$$y = mx + b. \quad (1)$$

Por lo tanto, podemos emplear métodos previamente establecidos para identificar la recta que mejor se adapte a los datos y, a través de ella, realizar predicciones; pues si la variable dependiente y denota el peso y la variable x representa la estatura, entonces conociendo el peso de una persona podríamos conocer su estatura. Por ejemplo, si conocemos que una persona pesa 68 kilogramos, podemos utilizar nuestra ecuación de regresión lineal $y = m(68) + b$ para obtener una estimación de su estatura, aunque este valor sea aproximado. Un tema relevante es estudiar lo bien que se ajusta el modelo a los datos lo cual es un tema que también es muy importante pero que no se abordará en estas notas.

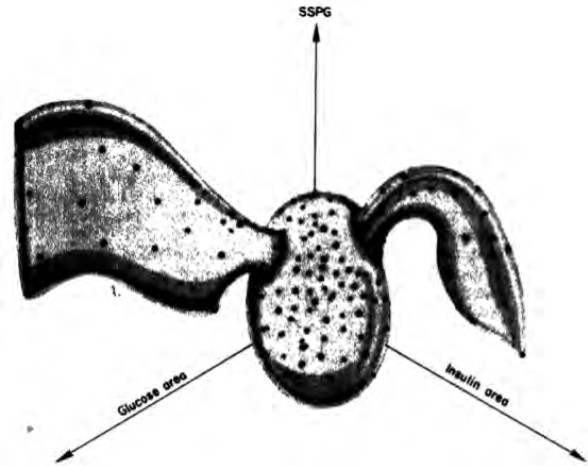


Figura 2 Cada paciente fue clasificado como normal, diabético químico, o diabético manifiesto. Se observó que los pacientes normales ocupaban el objeto redondeado central, mientras que los diabéticos químicos y los diabéticos manifiestos correspondían a las dos orejas en la imagen, ver [Reaven y Miller \(1979\)](#).

La importancia de la forma que presentan nuestros datos es clave para inferir información sobre ellos. En este sentido, la rama de la matemática que se dedica al estudio de los espacios abstractos y sus formas, se le conoce como topología. Aunque, no todo es tan sencillo en realidad. La mayoría de los problemas que nos enfrentamos presentan muchísimas variables en cuestión; así que que para obtener una representación en dos dimensiones o tres dimensiones es básicamente imposible. Aunque se han desarrollado herramientas de reducción de dimensiones para visualizarlos, como el algoritmo de análisis de componentes principales (PCA) o *projection pursuit* (vea [Kurita \(2019\)](#), y [Reaven y Miller \(1979\)](#)), donde es utilizado en el caso de estudio de Diabetes que clasifica tres tipos de pacientes como se ven en la Figura 3.

Por lo tanto, conocer la forma de los datos podría revelarnos patrones interesantes en ellos. El algoritmo de *Mapper* introducido por Carlsson, Mémoli y Singh en [Singh et al. \(2007\)](#), es un algoritmo no supervisado que ayuda en la representación y organización de un conjunto de datos multidimensional mediante un complejo simplicial como modelo. Mientras que, en su versión formal (topológica), no hay restricciones sobre el complejo simplicial asociado, este artículo se centra en casos donde se asigna un complejo simplicial de dimensión a lo sumo 1.



Estos objetos se les denomina grafos⁵

En estas notas descubriremos cómo esta herramienta nos ayuda a mapear las complejas estructuras de datos en representaciones visuales comprensibles. Desde la identificación de *clusters* (conglomerados de datos), hasta la revelación de conexiones no evidentes, además, el algoritmo Mapper desentraña capas de información que podrían pasar desapercibidas de otra manera. Finalmente, organizamos la estructura del artículo en cuatro secciones: la Sección 2 introduce los conceptos elementales que el lector necesita para entender el algoritmo. La Sección 3 presenta un ejemplo del algoritmo de Mapper implementado en Python mediante la librería *kmapper* van Veen *et al.* (2019), y van Veen *et al.* (2020). La Sección 4 presenta un caso de estudio detallado sobre los precios hedónicos mediante el algoritmo de Mapper. Por último, en Sección 5 hacemos un breve resumen de lo realizado en este documento.

2. FUNDAMENTOS DEL ALGORITMO DE MAPPER

En esta sección abordaremos los conceptos matemáticos que se necesitan para implementar el algoritmo de Mapper en su versión estadística: preimagen inversa de una función, métrica de un espacio (en nuestro caso, del conjunto de datos) y algoritmos de agrupación de datos.

Preimagen inversa de una función

Dada una función $f : X \rightarrow Y$ entre dos conjuntos y B cualquier subconjunto de Y . Se define la preimagen inversa de B bajo la función f como el conjunto:

$$f^{-1}[B] = \{x \in X : f(x) \in B\}. \quad (2)$$

Por lo tanto, la preimagen inversa de un conjunto B , denotada como $f^{-1}[B]$ en (2), está compuesta por todos los elementos en el dominio X que son mapeados a B por la función f . Veamos algunos ejemplos que ilustran este concepto.

Ejemplo 2.1. Considere el conjunto de calificaciones de un grupo de 30 estudiantes que consta de 3 exámenes parciales. Consideremos la función que nos entrega el promedio de cada estudiante: $f : X \rightarrow \mathbb{R}$ definida como

$$f((x_1, x_2, x_3)) = \frac{x_1 + x_2 + x_3}{3}.$$

Dicha función nos entrega el promedio de cada estudiante. Por lo tanto, si queremos representar en términos de la preimagen

inversa de la función promedio a todos aquellos estudiantes cuyo promedio es exactamente 70, entonces usando (2), escribiríamos

$$f^{-1}[\{70\}] = \{v \in X : f(v) = 70\}.$$

Nota: Si quisiéramos representar a todos aquellos estudiantes que aprueban el curso, simplemente se consideraría $f^{-1}[[70, 100]]$.

Es crucial comprender el concepto de preimagen inversa (2) para implementar con éxito el algoritmo de Mapper. En este contexto, el siguiente ejemplo tiene como propósito ilustrar y ampliar la explicación detallada.

Ejemplo 2.2. Considere el desplazamiento de una partícula el cual está descrita por la función cuadrática definida por $s(t) = t^2 + 1$, donde t representa el tiempo medido en segundos y $s(t)$ la cantidad de metros que se desplaza la partícula después de t segundos. Entonces,

$$s^{-1}[\{[20, 30]\}],$$

representa el conjunto de valores en el tiempo en el cual la partícula se ha desplazado entre 20 y 30 metros.

La noción de preimagen inversa es un concepto fundamental en matemáticas, especialmente útil al estudiar propiedades de las funciones, como la continuidad. En nuestro caso, este concepto es clave y forma una parte integral del algoritmo de Mapper.

Métrica sobre nuestro conjunto de datos

Otro ingrediente clave en el algoritmo de Mapper es el concepto de métrica. A continuación, se presenta la noción de métrica, la cual es un tipo especial de función definida sobre el conjunto $X \times X$, dado un conjunto X .

Definición 2.3. Dado un conjunto X . Decimos que una función $d : X \times X \rightarrow \mathbb{R}$ es una métrica si cumple las siguientes condiciones:

1. $d(x, y) = 0$ si y sólo si $x = y$.
2. $d(x, y) = d(y, x)$.
3. $d(x, y) \leq d(x, z) + d(z, y)$.

Aunque la Definición 2.3 pudiera parecer algo muy abstracto es algo con lo que estamos más familiarizados de lo que pensamos. Veamos, una manera de entender ese tipo de funciones que les llamamos métricas, es pensarlas como una generalización de la función que estudiamos en un primer curso de cálculo:

$$f(x, y) = |x - y|, \quad (3)$$

⁵ Un grafo es un objeto matemático que consta de vértices (puntos) y aristas (líneas).



que es la función distancia en la recta real entre los puntos x e y . Notemos que los valores que toma la función $f(x, y)$ en (3) son siempre mayor o igual a cero. Además, es el mismo valor $f(x, y) = f(y, x)$ para cada valor de x e y , y además cumple la famosa desigualdad triangular, que nos dice que $f(x, y) \leq f(x, z) + f(z, y)$ para cualesquier valor de x, y y z . Tal métrica se le conoce como la métrica Euclideana sobre los números reales.

Cabe notar que sobre un mismo conjunto X es posible definir métricas diferentes como lo muestran los siguientes ejemplos.

Ejemplo 2.4. Otra manera de medir en el espacio euclideo $\mathbb{R}^n$. La función

$$d_\infty(x, y) = \max\{|x_1 - y_1|, \dots, |x_n - y_n|\}, \quad (4)$$

donde $x = (x_1, x_2, \dots, x_n)$, $y = (y_1, \dots, y_n) \in \mathbb{R}^n$, la expresión (4) es una métrica para el espacio euclideo $\mathbb{R}^n$.

Ejemplo 2.5. Correlación de Pearson. Sean $x, y \in \mathbb{R}^n$. Definimos la distancia de correlación de Pearson entre x, y como

$$d(x, y) = 1 - \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (5)$$

donde $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ son las medias muestrales de x, y respectivamente.

La correlación de Pearson también se utiliza en el análisis exploratorio de datos para entender las relaciones entre variables. Esto puede proporcionar información valiosa sobre cómo las diferentes características del conjunto de datos están interrelacionadas.

Ejemplo 2.6. Métrica del Coseno. Considere el conjunto

$$S^n = \{x \in \mathbb{R}^{n+1} : \|x\| = 1\}. \quad (6)$$

Entonces, la función $d(x, y) = 1 - \frac{\cos^{-1}(x \bullet y)}{\pi}$ define una métrica sobre el conjunto S^n , donde $\bullet$ representa el producto punto entre los vectores y $\|x\|$ representa la norma o longitud del vector. Por lo tanto, S^n contiene a todos los vectores del espacio Euclideo $\mathbb{R}^{n+1}$ cuya norma es exactamente 1.

Después de comprender el concepto de métrica y revisar algunos ejemplos, es relevante destacar que la métrica del coseno se utiliza ampliamente en Machine Learning, especialmente en problemas relacionados con el procesamiento de texto y la representación vectorial. Este punto establece las bases para adentrarnos en el tema de agrupación de datos. La agrupación de datos es fundamental tanto en inteligencia artificial como en Machine Learning.

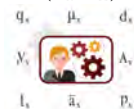
Agrupación de datos

La agrupación de los datos o *clustering* juega un papel muy importante en el mundo de la informática y la estadística. Supongamos que tenemos una habitación con juguetes de diferente tipo, algunos son coches, otros son muñecas y tal vez otros son pelotas. Ahora si se requiere poner en orden la habitación habrá que agrupar los juguetes que son parecidos en cajas separadas. Ahora, en el contexto de los datos es algo muy similar. Supongamos que se tiene una gran cantidad de datos sin organizar, y se requiere encontrar grupos o cajas de datos similares entre sí. Estos grupos se llaman *clusters*, o conglomerados. Considere el escenario en el que se trabaja con datos de clientes de una tienda en línea. Cada cliente tiene información como edad, historial de compras, y la frecuencia con la que visitan la tienda. Se busca comprender los tipos de clientes existentes, por lo que se utilizan algoritmos de agrupamiento para clasificarlos en distintas categorías.

El algoritmo examina los datos y encuentra automáticamente grupos basados en similitudes. Puede ser que encuentre un grupo de clientes jóvenes que compran con frecuencia, otro grupo de clientes mayores que compran ocasionalmente, y así sucesivamente. El objetivo es que los miembros de un mismo grupo sean más parecidos entre sí, que con los de otros grupos. Es como si cada grupo fuera una **caja** que contiene datos similares. Existen diversos algoritmos para organizar datos, como el agrupamiento espacial basado en densidad, el aglomerativo, y el de K-medias, entre otros. En este contexto, cobra vital importancia el concepto de métrica, ya que esta nos proporciona el sentido de cercanía según la función elegida en nuestro algoritmo. A continuación, nos sumergiremos brevemente en el algoritmo aglomerativo, utilizando el concepto de métrica que hemos introducido en la sección anterior.

Supongamos que tenemos un conjunto de datos X en el que se ha definido una métrica, es decir una función que cumple la Definición 2.3. Ahora introducimos el concepto de relación de equivalencia sobre el conjunto $X \times X$ como sigue. Fijemos un número real ϵ y $x, y \in X$. Entonces definimos la siguiente relación $R \subseteq X \times X$ como sigue: x está relacionado con y (denotado como xRy) si la distancia de x a y es menor que ϵ según nuestra métrica d , lo cual lo denotamos por $d(x, y) < \epsilon$. Por relación entendemos cualquier subconjunto del producto $X \times X$.

Ahora definimos la relación R' : $xR'y$ (x está relacionado con y según R') si existe una sucesión finita de elementos $x_0, x_1, \dots, x_n \in X$ de tal modo que $x_0 = x$, $x_n = y$ y además se cumple que $d(x_0, x_1) < \epsilon, d(x_1, x_2) < \epsilon, \dots, d(x_{n-1}, x_n) < \epsilon$.



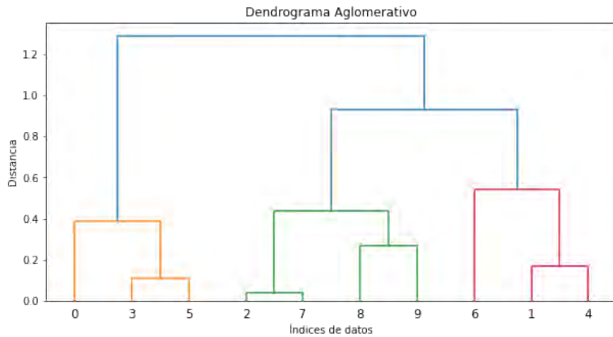


Figura 3 Cluster Aglomerativo o Jerárquico.

$\epsilon, \dots, d(x_{n-1}, x_n) < \epsilon$. No es difícil verificar que R' satisfice las siguientes propiedades:

1. $xR'x$ para todo $x \in X$.
2. Si $xR'y$ entonces $yR'x$.
3. Si $xR'y$ y $yR'z$, entonces $xR'z$.

Las propiedades 1), 2) y 3) convierten a R' en una relación de equivalencia. Finalmente se introduce el concepto de clase de equivalencia. Dado un elemento $x \in X$, la clase de equivalencia para el elemento x es el conjunto

$$[x] = \{y \in X : yR'x\}, \quad (7)$$

es decir, la clase de equivalencia del elemento x según R' , son todos los elementos y en nuestro conjunto de datos X que está en relación con x . El conjunto de todas las clases de equivalencia es lo que conocemos como los *clusters* obtenidos mediante el algoritmo aglomerativo.

Cabe notar la importancia de la métrica definida en nuestro conjunto de datos X y el valor fijo ϵ , ya que estos parámetros modifican sustancialmente los elementos que contiene cada clase de equivalencia.

Algoritmo de Mapper

A continuación se presenta el algoritmo de Mapper que sólo puede generar a lo más complejos simpliciales de dimensión 1, es decir, puntos y líneas. Por lo tanto, el modelo matemático que se asignará al conjunto de datos que se desea estudiar, será un grafo. A continuación, presentamos el algoritmo de Mapper en su versión estadística para la generación del grafo.

Definición 2.7 (Algoritmo de Mapper).



1. (Filtro) Escoger una función que proyecta el conjunto de datos X en algún $\mathbb{R}^d$, i.e., una función $f : X \rightarrow \mathbb{R}^d$, donde d es algún número natural mayor que cero.
2. (Cubiertas y *Overlap*) Dado un número natural $k > 0$, se propone fijar una cubierta finita de intervalos cerrados $\{U_i : i \leq n\}$ de $f(X)$, donde cada intervalo U_i satisface la condición de tener una longitud constante y uniforme a lo largo de la cobertura. Es decir, se busca una colección de intervalos tal que la intersección entre intervalos adyacentes, $U_i \cap U_{i+1}$, tenga la misma longitud para cada $i \leq n$.
3. Seleccionar un algoritmo para el agrupamiento de datos (clustering). Entonces, para cada $f^{-1}[U_i]$ hacer una clusterización $\{C_{ij} : j \leq m_i\}$.
4. Formar el grafo compuesto por los vértices $\{C_{ij} : i \leq n, j \leq m_i\}$ y si $C_{ij_0} \cap C_{ij_1} \neq \emptyset$, entonces se añade una arista (no orientada) entre estos dos vértices.

3. ALGORITMO DE MAPPER SOBRE UN CONJUNTO DE FLORES

El conjunto de datos Iris es un conjunto de datos muy conocido en el campo del aprendizaje automático y la estadística ver [Fisher \(1988\)](#). Es utilizado comúnmente para prácticas y ejemplos debido a su simplicidad y relevancia en la clasificación de especies de flores iris. Este conjunto de datos consta de 150 observaciones de iris, cada una de las cuales pertenece a una de las tres especies: setosa, versicolor o virginica. Cada observación tiene cuatro características: longitud del sépalo, ancho del sépalo, longitud del pétalo y ancho del pétalo, todas medidas en centímetros.

A continuación compartimos la implementación del algoritmo de Mapper en Python usando la librería Kmapper sobre el conjunto de datos Iris. Primero importamos los datos desde la web usando la librería `ucimlrepo`. Usando el siguiente comando instalamos la librería de datos disponibles en el repositorio **UC Irvine Machine Learning Repository**.

```
pip install ucimlrepo
```

La biblioteca Pandas se emplea para descargar la base de datos y asignarla al objeto `data`, que posteriormente será referenciado en el análisis subsiguiente. Para adquirir y procesar esta base de datos, se utiliza el siguiente código en Python.

```
from ucimlrepo import fetch_ucirepo
import pandas as pd

#dataset
data=pd.read_csv('
    https://archive.ics.uci.edu/static/public/53/data.csv')
print(data)
```

Vamos a emplear las variables de esta base de datos. Para hacerlo, es necesario referenciarlas de manera individual y separarlas en variables explicativas, y variables de respuesta. En el siguiente código, se proporciona un ejemplo de cómo realizar esta tarea.

```
sepalength = data['sepal length']
sepalwidth = data['sepal width']
petallength = data['petal length']
petalwidth = data['petal width']
species= data['class']
```

Hasta este punto, hemos obtenido la información de la base de datos iris y realizado algunas asignaciones de preprocesamiento cruciales. El siguiente paso consiste en importar la paquetería necesaria para implementar el algoritmo de Mapper usando el siguiente código en Python.

```
import numpy as np
import sklearn
import kmapper as km
from sklearn.preprocessing import StandardScaler
from sklearn.decomposition import PCA
import matplotlib.pyplot as plt
import scipy.cluster.hierarchy as shc
from sklearn.cluster import AgglomerativeClustering
from sklearn.cluster import DBSCAN
from ripser import Rips
```

Como ya hemos importado la base de datos en el objeto data, el siguiente paso implica normalizar cada variable numérica. En otras palabras, ajustaremos cada columna para que tenga norma unitaria.

```
np_vector1 = sepalength.to_numpy()
np_vector2 = sepalwidth.to_numpy()
np_vector3 = petallength.to_numpy()
np_vector4 = petalwidth.to_numpy()

normalized_vector1= np_vector1/np.linalg.norm(np_vector1)
normalized_vector2= np_vector2/np.linalg.norm(np_vector2)
normalized_vector3= np_vector3/np.linalg.norm(np_vector3)
normalized_vector4= np_vector4/np.linalg.norm(np_vector4)

df_vector1 = pd.DataFrame(normalized_vector1)
df_vector2 = pd.DataFrame(normalized_vector2)
df_vector3 = pd.DataFrame(normalized_vector3)
df_vector4 = pd.DataFrame(normalized_vector4)

df_vector1.insert(loc = 1 , column = '1', value = df_vector2)
df_vector1.insert(loc = 2 , column = '2', value = df_vector3)
df_vector1.insert(loc = 3 , column = '3', value = df_vector4)

df_vector1 = pd.DataFrame(df_vector1)
```

Construimos el grafo utilizando como algoritmo de *clustering* DBSCAN con la métrica Euclideana, establecemos con el valor de 5 el número de cubiertas (n_cubes) y como solapamiento o traslape entre las cubiertas ($perc_overlap$) 0.2.

```
mapper = km.KeplerMapper(verbose=1)

graph = mapper.map(
    df_vector1,
    df_vector1,
    clusterer=sklearn.cluster.DBSCAN(metric="euclidean"),
    cover=km.Cover(n_cubes=5, perc_overlap=0.2),
)
```

El código para visualizar el grafo se muestra en la Figura 4, y es como sigue:

```
mapper.visualize(
    graph,
    path_html="Mapper Graph on Iris Scaled Data Set.html",
    title="Mapper Graph on Iris Scaled Data Set",
    nbins=8,
    custom_tooltips=species,
    node_color_function=['mean','std','median','max']
)

km.draw_matplotlib(graph)

import matplotlib.pyplot as plt

plt.show()
```

En la Figura 4 se encuentra el grafo obtenido por Mapper, mientras que en la Figura 5, se encuentra la proyección de

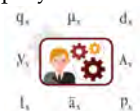




Figura 4 Resumen topológico obtenido por algoritmo de Mapper.

los datos en el plano utilizando el algoritmo de análisis de componentes principales, un algoritmo muy utilizado en *machine learning*.

En el grafo obtenido de Mapper, Figura 4, del conjunto de Iris podemos notar como el grafo muestra la división entre las tres clases de flores. Además de eso, muestra la distribución dentro de cada cluster o conglomerado en el grafo. Una peculiaridad a destacar del algoritmo es que revela las relaciones que se presentan entre los clusters (a través de las aristas) a diferencia de algoritmos de reducción de dimensiones como el análisis de componentes principales (PCA), ver Figura 5. Dichas relaciones podrían mostrar un patrón de comportamiento difícil de detectar con herramientas clásicas de estadística o *machine learning*. Una aplicación en el contexto de la política donde muestran las ventajas de Mapper sobre PCA puede consultarse a detalle en el artículo [Vejdemo-Johansson et al. \(2012\)](#).

4. EL MODELO DE LOS PRECIOS HÉDONICOS Y MAPPER

En este ejemplo deseamos mostrar cómo a través del algoritmo de Mapper es posible visualizar y complementar, en un ejemplo, la relación entre una variable respuesta y variables que son consideradas como amenidades o desamenidades en un modelo de precios hedónicos.

El modelo de los precios hedónicos

Existe una teoría que fundamenta esta modelación, la cual es llamada *teoría de los precios hedónicos*. Según [Rosen \(1974\)](#), esta teoría de los precios hedónicos se basa en el supuesto de que los bienes se valoran por la

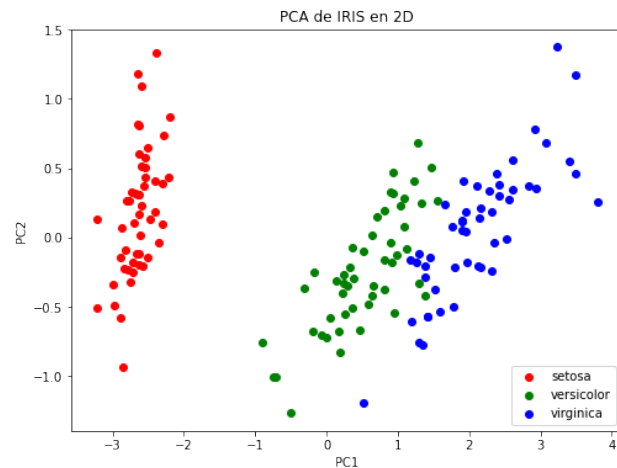


Figura 5 La proyección de los datos en el plano utilizando el algoritmo de análisis de componentes principales (PCA).

satisfacción o el placer que sus atributos o características producen en los consumidores y propone explicar el valor de un bien heterogéneo, es decir, un bien con n características o atributos, en función de cada una de estas características; así, con esta teoría podemos identificar el efecto de cada atributo sobre el valor de mercado asignado a un bien heterogéneo. Ejemplos de un bien heterogéneo podemos pensar en viviendas, coches, computadoras, etc. En el mercado las características de ese bien no son comercializables directamente, aunque tienen un precio implícito, así que una vez se encuentran los precios de las características del bien, se puede determinar el precio del bien en el mercado observado. En el caso de viviendas se consideran características físicas de la unidad (el número de recámaras, el número de baños, el área total, etc.), el entorno de la unidad o las características de la ubicación (por ejemplo, la proximidad a centro comercial, cercanía a universidades) y otras variables de interés que puedan producir amenidad o desamenidad, como por ejemplo, la proximidad a una zona de altos índices de crimen.

La metodología para encontrar los precios hedónicos de las características de un bien se lleva a cabo mediante una regresión, llamada *regresión hedónica*, la cual consiste en la estimación estadística de la relación entre las características de un bien y su precio de mercado, determinando así el valor del bien mismo. Para ello se tiene que tener una base de datos de precios del bien en cuestión con diferentes características, además, se tiene que proponer

una forma funcional con buen ajuste de los datos. La variable dependiente, llamada respuesta, será el precio del bien y las variables independientes, las cuales se llaman variables explicativas, son cantidades o calidades de las características, amenidades o desamenidades.

Un ejemplo de los precios hedónicos con TDA

En este ejemplo vamos a trabajar con una base de datos que contiene precios de viviendas en la ciudad de Acapulco, Gro., correspondientes al período 2015-2016. La variable a explicar es el precio de la vivienda. Esta base de datos consisten en una muestra de 339 viviendas ofertada en los sitios web de empresas inmobiliarias de la ciudad de Acapulco, ubicadas en 57 colonias. En la tabla 1 se muestran la variable explicativa y las variables explicativas. El objetivo de este ejemplo es mostrar visualmente, mediante resúmenes topológicos, las correlaciones de la variable respuesta con cada una de las variables explicativas, esto nos sirve para darnos una idea de como cambia la variable respuesta al cambiar la variable explicativa. El uso de TDA podemos considerarlo como un método complementario a la regresión hedónica. Iniciamos descargando la base de datos DATOSSCTC.csv (vea Wences y Hernández-López (2018)) disponible [aquí](#).

Ejemplo 4.1.

Cargamos los datos con la siguiente instrucción.

```
df=pd.read_csv('DATOSSCTC.csv',encoding='utf-8')
```

Ahora convertimos los datos en arreglos numpy, lo hacemos mediante la siguiente instrucción de Python:

```
Xt = np.array(df.fillna(0))
```

La librería kmapper nos permite visualizar el resumen topológico con la función `mapper.map(proyecciondatos, datos, cover)`. En los siguientes resúmenes topológicos usamos una función de filtrado con análisis de componentes principales, PCA, con una componente:

```
mapper = km.KeplerMapper(verbose=0)
lens = mapper.fit_transform(Xt, projection=PCA(1))
lens.shape
```

Antes de crear los resúmenes topológicos creamos un nuevo data frame para visualizar 2 variables:

```
def Concatenar(N1,N2):
    #N1 y N2 deben ser columnas de pandas dataframe
    #N1=df['H']
    #N2=df['IT']
    N=pd.DataFrame(np.zeros(np.size(N1)))
    for i in range(np.size(N1)):
        N.loc[i]=str(N1.loc[i])+ '/' +str(N2.loc[i])
    return N
```

Para visualizar las correlaciones de las características con la variable respuesta, procedemos a concatenar la variable respuesta con cada una de las variables características. A continuación mostramos sólo el código de dos concatenaciones:

```
precios=pd.DataFrame(df['P'])
precios["PRECIOS,SUPERFICIE"]=Concatenar(df['P'],df['S'])
precios["PRECIOS,HOMICIDIOS"]=Concatenar(df['P'],df['H'])
```

La Figura 6 muestra el resumen topológico de la variable respuesta, precio de la vivienda, concatenada con la variable superficie. Se puede observar que el cluster señalado con la línea en negro corresponde a los datos desplegados en la izquierda de la figura. Para este nodo señalado se observa que los datos corresponden a viviendas con superficies menores y viviendas con precios bajos. Los nodos siguientes de arriba hacia abajo sobre el mismo nervio contienen datos de precios y superficies de viviendas que van aumentando a medida que se avanza. En la Figura 7 se muestran los datos desglosados correspondientes al último nodo de arriba hacia abajo. Los datos en este conglomerado contiene precios altos de viviendas y superficies mayores de viviendas, es decir, las viviendas con mayor superficie tienden a tener un precio más alto, luego podemos decir que una vivienda con una mayor superficie produce una amenidad en el comprador.

En las Figuras 8, 9 se muestran los resúmenes topológicos de la variable respuesta concatenada con la variable homicidios. Es preciso mencionar que el nivel más alto de homicidios es de 66 (66 homicidios ocurridos en el período de estudio en la colonia donde se ubica la vivienda observada) y el nivel más bajo es 0. En la Figura 8 se señala el conglomerado con precios altos de viviendas que corresponden a bajos niveles de homicidios, mientras que en Figura 9 se señala el conglomerado con precios bajos de viviendas correspondientes a altos niveles de homicidios. A medida que nos desplazamos de arriba hacia abajo los precios aumentan mientras tanto, los niveles de homicidios decrecen, así podemos interpretar a los homicidios como una desamenidad de la vivienda.

En las Figuras 10, 11 se puede ver como viviendas con más número de baños corresponden a viviendas con



	Variable	Descripción
1.	P	precio de la vivienda observada
2.	NR	número de recámaras de la vivienda
3.	NB	número de baños de la vivienda
4.	NE	número de estacionamientos de la vivienda
5.	A	antigüedad de la vivienda
6.	S	superficie total de la vivienda
7.	TC	superficie construida en la vivienda construidas antes de 1940
8.	H	número de homicidios ocurridos en la colonia donde se ubica la vivienda observada en el período 2015-2016
9.	IT	índice de turismo

■ **Tabla 1** Variables respuesta y explicativas

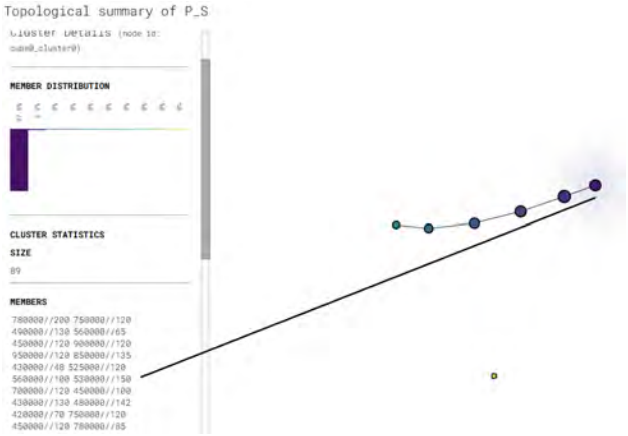


Figura 6 Resumen topológico con proyección PCA, de la variable respuesta precios de las viviendas concatenada con la variable superficie. Superficies menores de viviendas corresponden a bajos precios de viviendas



Figura 7 Resumen topológico con proyección PCA, de la variable respuesta precios de las viviendas concatenada con la variable superficie. Superficies mayores de viviendas corresponden a precios altos de viviendas



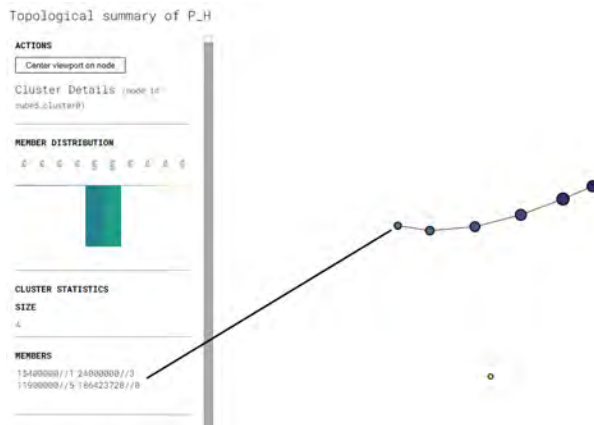


Figura 8 Resumen topológico con proyección PCA, de la variable respuesta precios de las viviendas concatenada con la variable homicidios. Bajos niveles de homicidios corresponden a precios altos de viviendas

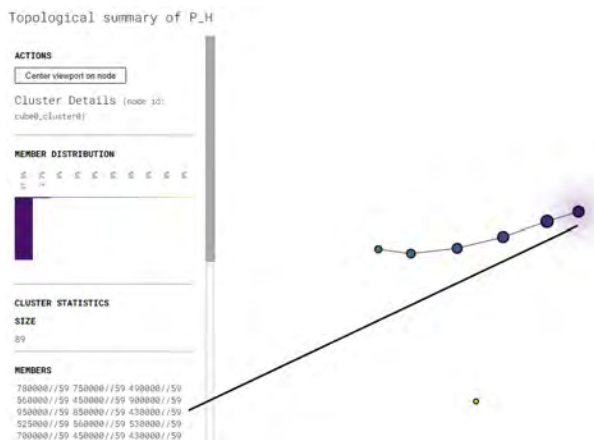


Figura 9 Resumen topológico con proyección PCA, de la variable respuesta precios de las viviendas concatenada con la variable homicidios. Altos niveles de homicidios corresponden a precios bajos de viviendas

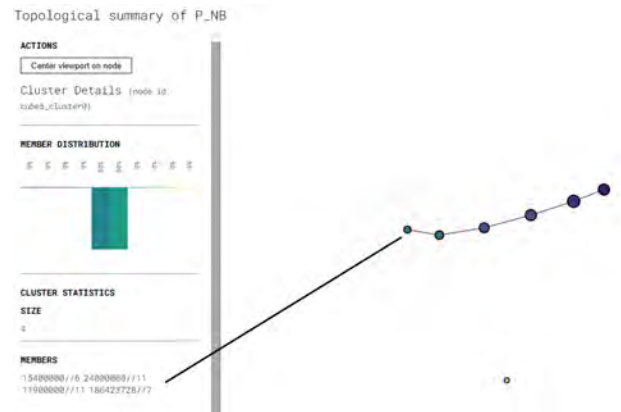


Figura 10 Resumen topológico con proyección PCA, de la variable respuesta precios de las viviendas concatenada con la variable número de baños. Entre más baños tenga la vivienda su precio aumenta

precios más altos, y viviendas con menos número de baños corresponden a viviendas con precios más bajos.

En las Figuras 12 Y 13 se puede observar que a mayor número de estacionamientos en la vivienda, el precio de ésta aumenta, y a menor número de estacionamientos en la vivienda, su precio decrece.

Es posible que para alguna de las variables explicativas no se vea con claridad en el resumen topológico el comportamiento de la variable respuesta en función de los valores de la variable explicativa, ello podría significar que dicha variable no es significativa para explicar la variable respuesta y se puede verificar con la regresión hedónica. Este tipo de situación se ejemplifica con la variable antigüedad de la vivienda. En la Figura 14 el cluster señalado corresponde a viviendas con precios altos, sin embargo, la variable explicativa A contiene tanto valores pequeños como valores grandes y esto también puede observarse en la Figura 15, en el cluster señalado correspondiente a viviendas con precios bajos, A contiene valores pequeños como altos.

Podemos examinar detenidamente cada una de las variables explicativas junto con la variable de respuesta en este contexto. En todos los resúmenes topológicos, se destaca un clúster o conglomerado aislado en color amarillo, separado del conjunto principal, el cual consiste en un único valor de P y un correspondiente valor de la variable explicativa. Este único valor representa una observación atípica en la variable de respuesta.



Topological summary of P_NB

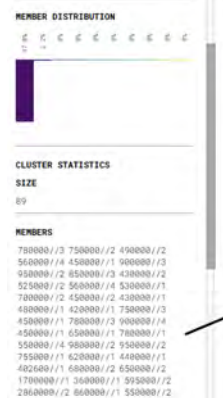


Figura 11 Resumen topológico con proyección PCA, de la variable respuesta precios de las viviendas concatenada con la variable número de baños. Entre menos baños tenga la vivienda su precio decrece

Topological summary of P_NE

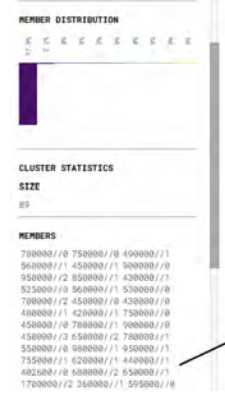


Figura 13 Resumen topológico con proyección PCA, de la variable respuesta precios de las viviendas concatenada con la variable número de estacionamientos. Entre menos estacionamientos tenga la vivienda su precio decrece

Topological summary of P_NE



Figura 12 Resumen topológico con proyección PCA, de la variable respuesta precios de las viviendas concatenada con la variable número de estacionamientos. Entre más estacionamientos tenga la vivienda su precio aumenta

Topological summary of P_A



Figura 14 Resumen topológico con proyección PCA, de la variable respuesta precios de las viviendas concatenada con la variable antigüedad. No se aprecia si hay una correlación entre estas variables



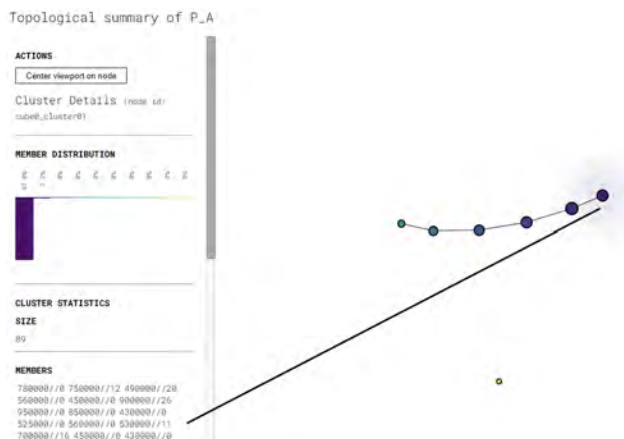


Figura 15 Resumen topológico con proyección PCA, de la variable respuesta precios de las viviendas concatenada con la variable antigüedad. En el cluster señalado se observan valores bajos y altos de la variable A

5. COMENTARIOS FINALES

En resumen, este escrito ha abordado el fascinante mundo del modelado de datos mediante el algoritmo de Mapper. Este enfoque no convencional nos permite representar la forma de los datos y descubrir relaciones entre variables en dimensiones significativamente grandes. A lo largo del texto, se han explorado conceptos y definiciones fundamentales en matemáticas, estableciendo conexiones entre estas ideas matemáticas y las innovadoras técnicas de análisis topológico de datos. La aplicación práctica de estos conceptos se ha presentado a través de ejemplos implementados en programación orientada a objetos mediante el lenguaje de Python. Este lenguaje, reconocido por su naturaleza de alto nivel, multiplataforma, interpretado y de baja complejidad, facilita la comprensión y aplicación de los principios topológicos en el análisis de datos. Para asegurar la comprensión del código, se han utilizado bibliotecas idóneas que permiten una implementación rápida y eficiente. La intención es que el lector pueda seguir cada ejemplo de manera fluida y autocontenida, proporcionando así una introducción práctica y aplicada a la topología de datos.

Aunque este documento ofrece una base sólida, se sugiere que los lectores interesados puedan explorar aún más los temas mediante las referencias presentadas en cada sección, y también al final del documento, se proporciona un listado completo de referencias para aquellos que deseen profundizar en áreas específicas. En conjunto,

este escrito busca no sólo informar, sino también inspirar a profesionales ajenos a matemáticas aplicadas, ciencias de datos o temas afines a explorar nuevas perspectivas y herramientas, contribuyendo así al crecimiento continuo de su disciplina particular, y agregando nuevos enfoques para analizar conjuntos de datos complejos y dinámicos.

REFERENCIAS

- Fisher, R. A., 1988 Iris. UCI Machine Learning Repository, DOI: <https://doi.org/10.24432/C56C76>.
- Kurita, T., 2019 Principal component analysis (pca). Computer Vision: A Reference Guide pp. 1–4.
- Reaven, G. M. y R. G. Miller, 1979 An attempt to define the nature of chemical diabetes using a multidimensional analysis. *Diabetologia* **16**: 17–24.
- Rosen, S., 1974 Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy* **82**: 34–55.
- Singh, G., F. Mémoli, y G. Carlsson, 2007 Topological methods for the analysis of high dimensional data sets and 3d object recognition pp. 91–100.
- van Veen, H. J., N. Saul, D. Eargle, y S. W. Mangham, 2019 Kepler mapper: A flexible python implementation of the mapper algorithm. *Journal of Open Source Software* **4**: 1315.
- van Veen, H. J., N. Saul, D. Eargle, y S. W. Mangham, 2020 Kepler Mapper: A flexible Python implementation of the Mapper algorithm.
- Vejdemo-Johansson, M., G. Carlsson, P. Y. Lum, A. Lehman, G. Singh, *et al.*, 2012 The topology of politics: voting connectivity in the us house of representatives. En *NIPS 2012 Workshop on Algebraic Topology and Machine Learning*.
- Wences, G. y E. Hernández-López, 2018 Precios hedónicos. DOI: <https://doi.org/10.24432/C56C76>.



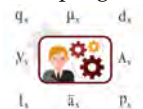


Salvador Mancilla Hernández obtuvo el Doctorado en la Universitat de Barcelona en el área de lógica pura y aplicada bajo la dirección de Joan Bagaria i Pigrau. Su investigación se centra en los grandes cardinales y su interacción directa con la matemática, especialmente con la topología y la teoría de grupos. También está interesado en el desarrollo de algoritmos mediante topología y sus aplicaciones a las grandes cantidades de datos. Realizó una estancia de investigación con el proyecto Propiedades de Reflexión en Espacios Topológicos en la Universidad Autónoma Metropolitana bajo la dirección de Rodrigo Hernández Gutiérrez. Actualmente es coordinador de la unidad de formación “Uso de geometría y topología para ciencia de datos” en el Instituto Tecnológico de Estudios Superiores de Monterrey (campus Monterrey). Miembro del Sistema Nacional de Investigadores. Ha sido parte del comité organizador de European Set Theory en la provincia Mòn Sant Benet de Catalunya en 2013 así como parte del comité organizador del seminario de Topología en la Universidad Autónoma Metropolitana 2021.

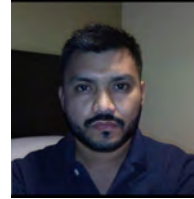


Eymard Hernández-López es un investigador en matemáticas aplicadas. Obtuvo su Licenciatura y su Maestría en Matemáticas en la Universidad Autónoma de Metropolitana (UAM) unidad Iztapalapa y completó su Doctorado en la unidad Cuajimalpa de la misma Universidad, y trabajó como investigador en el Instituto Mexicano del Petróleo. De 2011

al 2022, colaboró como profesor en la UAM-Iztapalapa, donde se destacó por sus investigaciones sobre matemáticas aplicadas. Sus intereses abarcan la teoría de bifurcaciones, problemas inversos, optimización, topología de datos y machine learning. Es miembro del American Institute of Mathematical Science y la Sociedad Matemática Mexicana, el Dr. Hernández-López ha aplicado machine learning para mejorar la interpretación de imágenes radiográficas y desarrollar métodos innovadores de detección del cáncer. Ha compartido su experiencia en conferencias internacionales y fue nombrado Investigador SNI nivel I en 2022 por sus destacadas contribuciones a la investigación en México. Actualmente es profesor en el posgrado en ingeniería industrial en el Tecnológico



Nacional de México, campus TESOEM.



Giovani Arquímedes Wences Nájera es profesor investigador de tiempo completo en la Escuela Superior de Matemáticas de la Universidad Autónoma de Guerrero, México. Obtuvo su Doctorado en Matemáticas en 2020 en el área de transporte óptimo y econometría bajo la supervisión del Profesor Joaquín Delgado. Sus líneas de investigación incluyen la teoría del transporte óptimo, el aprendizaje automático, los modelos de regresión lineal y la econometría. Además es candidato a ser mimebro del Sistema Nacional de Investigadores. Actualmente, realiza una estancia posdoctoral bajo la modalidad académica en la Universidad Autónoma Metropolitana-Iztapalapa.

La Facultad de Ciencias Actuariales de la Universidad Anáhuac México te invita a participar como ponente en sus sesiones mensuales de: **VISIÓN ACTUARIAL SOBRE LA INNOVACIÓN, LA INVESTIGACIÓN Y EL DESARROLLO**



Vale por horas de Educación Continua del CONAC

Facultad de
Ciencias Actuariales

Informes:

Dr. Daniel López Barrientos
daniel.lopez@anahuac.mx
anahuac.mx/mexico

VAIID

VISIÓN ACTUARIAL SOBRE LA
INNOVACIÓN, LA INVESTIGACIÓN Y
EL DESARROLLO

PREMIO FUNDADORES

Carta del Presidente del Consejo Directivo de la AMA

Act. Fernando Patricio Belaunzarán Barrera



Estimados colegas

Estamos por cerrar un 2023 que fue un año muy importante para la Asociación Mexicana de Actuarios, lleno de grandes eventos como nuestro **XXXI Congreso Nacional de Actuarios** y diversos cursos de interés para el gremio actuarial con el objetivo de tener una visión de un "Actuario Integral".

Este año realizamos la convocatoria de nuestro tercer premio fundadores, cuyos trabajos ganadores forman parte esta revista. Los ganadores del segundo lugar fueron los actuarios José Daniel López Barrientos y Humberto Plata Gallegos, con su trabajo: *Una nueva aproximación a la constitución de una reserva actuarial de arriba p'abajo*, donde se muestra una alternativa a la tendencia actual de modelación matemática de la constitución de reservas matemáticas de seguros y pensiones.

Entre las técnicas actuariales que se usaron están el modelo de Cramer-Lundberg con primas y derechos contingentes, y las recomendaciones de Bühlmann para calcular las primas en los niveles superior e inferior de una cartera. De la teoría del control estocástico, recurrieron a la optimización de un proceso Markoviano de decisión donde el conjunto de estados es un espacio métrico separable y localmente compacto bajo el criterio del pago promedio en horizonte infinito, y el uso de una versión de la ley fuerte de los grandes números para considerar las trayectorias del proceso de riqueza, en lugar de sus valores esperados.

En el tercer lugar, el Act. Antonio Cruz Torres participó con el trabajo titulado *Propuesta de aplicación de economía del comportamiento y aprendizaje máquina para incrementar el ahorro a largo plazo*.

En dicho trabajo se toman recursos de la Economía del Comportamiento y el aprendizaje máquina argumentando que la aportación de estas disciplinas puede ser particularmente efectiva para abordar problemas relacionados con el ahorro personal a largo plazo, aplicable en algunos contextos como el de los seguros con componente de ahorro a largo plazo, planes para el retiro o el ahorro voluntario en las AFORE. En la temática de inclusión financiera, además de brindar acceso a los productos financieros, resulta relevante entenderlos y usarlos de manera eficaz, lo cual es crucial para optimizar los resultados de ahorro.

Los invitamos a que nos hagan sus comentarios al correo electrónico presidencia@actuarios.org.mx, nos dará mucho gusto ver sus ideas y sugerencias.

Act. Fernando Patricio Belaunzarán Barrera
 Presidente del Consejo Directivo
 de la Asociación Mexicana de Actuarios
 2022-2024

La Facultad de Ciencias Actuariales y
el Colegio Nacional de Actuarios te
invitan a participar como expositor
en las:

JORNADAS DE ACTUARIZACIÓN

Vale por horas de
Educación Continua
del CONAC

Facultad de
Ciencias Actuariales

Informes:
Dr. Daniel López Barrientos
daniel.lopez@anahuac.mx
anahuac.mx/mexico



COLEGIO
NACIONAL
DE ACTUARIOS

PREMIO FUNDADORES

**3er lugar. Propuesta de
aplicación de economía
del comportamiento y
aprendizaje máquina para
incrementar el ahorro a largo
plazo**

Antonio Cruz Torres



Propuesta de aplicación de economía del comportamiento y aprendizaje máquina para incrementar el ahorro a largo plazo

Antonio Cruz Torres

RESUMEN Este artículo presenta un mapeo del proceso conductual de una persona frente al ahorro. Nuestra meta es aterrizar el problema de la falta de ahorro para identificar los puntos cruciales y aquellos que deberían modificarse a fin de incentivar un mejor resultado. Para alcanzar esta meta, proponemos una metodología que abarca las fases del análisis, el diagnóstico y la implementación de una herramienta amigable con el usuario ahorrador.

Palabras clave

Marco CREATE, k -means, aprendizaje máquina

1. INTRODUCCIÓN

El retiro consiste en una etapa crítica en la vida de las personas, resulta en la conclusión de la vida laboral

Derechos reservados © 2023 por el Colegio Nacional de Actuarios
Última actualización del manuscrito: 17 de noviembre de 2023
Este artículo es publicado por el Colegio Nacional de Actuarios (CONAC) y contiene información de una variedad de fuentes. Es un trabajo cuyo fin es únicamente informativo y no debe interpretarse como asesoramiento profesional o financiero. La visión y comentarios contenidos en este documento corresponden exclusivamente a su autor y, por tanto, son ajenos a cualquier entidad pública o privada, incluyendo aquella para la que actualmente colabora. El CONAC no recomienda ni respalda el uso de la información proporcionada en este estudio. El CONAC no ofrece ninguna garantía, expresa o implícita, ni representación de ningún tipo y no asume ninguna responsabilidad en relación con el uso o mal uso de este trabajo.
¹ Correo electrónico: antonio.cruz@hotmail.com



e idealmente de descanso y tranquilidad, tanto en el ámbito global personal como financiero. No obstante, en la actualidad el fenómeno general con respecto al retiro es preocupante debido a la baja proporción que representa el monto de pensión estimado en relación con el salario actual de los trabajadores, comúnmente llamada tasa de reemplazo, situación que plantea desafíos importantes para los actores involucrados en el sector del retiro. En el contexto de esquemas de cuentas individuales para el retiro, el instrumento complementario para mejorar esas tasas es el ahorro mismo del trabajador realizado de forma voluntaria.

El estímulo hacia el ahorro además de buscar mejorar las perspectivas de jubilación como en este caso, puede generalizarse y fomentar su familiarización e interacción

con otros instrumentos del sistema financiero, trasladando de manera directa algunas de las estrategias que se plantearán para impulsar el manejo y penetración de otros productos, como los relacionados con el sector asegurador, en especial seguros con componentes de ahorro a largo plazo.

Una forma de enfrentar esta temática es con apoyo interdisciplinario, en este caso con apoyo de la economía conductual y la ciencia de datos en la búsqueda de soluciones. La economía conductual, como rama de la economía busca entender las decisiones humanas incorporando conceptos psicológicos en los modelos económicos, en ese sentido incorpora la premisa de que los individuos no siempre actúan de manera racional o en su mejor interés financiero. Comprendiendo estas desviaciones, es posible diseñar intervenciones más efectivas que motiven cambios de comportamiento positivos en la sociedad, en específico se toman algunas estrategias de esa disciplina para localizar obstáculos que pueden estar impidiendo a las personas ahorrar de manera efectiva y ofrecer alternativas personalizadas para fomentar la decisión de realizar el ahorro. Por otro lado, la ciencia de datos, para la búsqueda de soluciones a través de la identificación de patrones y relaciones en conjuntos de datos y en específico el aprendizaje máquina facilitando el procesamiento y análisis de información. En el contexto del ahorro, mientras la economía conductual identifica barreras psicológicas que complican a las personas realizarlo, el aprendizaje máquina puede segmentar a la población en grupos homogéneos basados en sus características, permitiendo la creación de estrategias personalizadas para cada segmento.

2. SITUACIÓN ACTUAL

En 1997 se realizó en México una reforma relevante a la Ley del Seguro Social, que implicó la creación de cuentas individuales para los trabajadores a fin de que en ellas se depositaran sus ahorros para el retiro, cuentas que ahora serían administradas por las Administradoras de Fondos para el Retiro (AFORE). Esto contrastaba con el sistema anterior, en que las pensiones eran administradas por el Estado. Una de las premisas relacionadas fuertemente con el cambio tiene que ver con incentivar el ahorro por parte del trabajador, es decir, hacerlo partícipe de su ahorro para el retiro, misma que aún a la fecha está siendo complicada de cumplir tanto por las AFORE como por el regulador (CONSAR). Esta es la razón del surgimiento del presente trabajo que busca abordar la cuestión con apoyo de herramientas de reciente desarrollo, como

la economía conductual y la ciencia de datos.

El ahorro voluntario como herramienta complementaria a las aportaciones obligatorias ha tenido un crecimiento en los últimos años. Sin embargo, continua con una participación muy baja a pesar de los esfuerzos que se han realizado para facilitar procesos operativos e incremento de la oferta y difusión.

Una pieza importante para destacar en este episodio es la reciente implementación de calculadoras de proyecciones de pensión, derivadas a partir de la publicación de la metodología de CONSAR. Vea el manual elaborado por [Instituto Mexicano del Seguro Social \(2023\)](#).

Ahora varias de las AFORE cuentan con su propia herramienta de proyección variando en algunos supuestos, pero en general bajo la misma nota técnica del regulador, lo que permite estandarizar relativamente los cálculos de pensión. En este análisis el uso de esta herramienta es crucial para la generación de resultados que podrán ser visualizados por el trabajador.

Para dimensionar el impacto del ahorro voluntario en las pensiones, se revisó un grupo de 1.5 millones de personas registradas por voluntad propia en una AFORE, que tienen la característica de presentar aportaciones de manera activa en la actualidad. Se asumen supuestos que pueden variar como la tasa de rendimiento, la tasa asociada al cálculo de la unidad renta vitalicia o la densidad de cotización pero que consideran niveles posibles acorde a históricos. El 67 % alcanzaría una pensión mínima correspondiente a su rango salarial, el 31.5 % alcanzaría una pensión superior a la mínima y el 1.5 % corresponde a negativa de pensión.

El grupo de cotizantes que alcanzan una pensión superior a las mínimas garantizadas tienen una tasa de reemplazo alrededor del 39.5 %.

Ahora, suponiendo que todos realizaran un 5 % de ahorro voluntario de manera mensual sobre su sueldo actual, el grupo de personas con pensión superior a la mínima garantizada se incrementaría de 31.5 % a 46.2 % y su tasa de reemplazo alcanzaría niveles de 50 %. Es importante resaltar que estos resultados dependen de los supuestos asociados y pueden modificarse en escenarios temporales con situaciones económicas distintas, en lo que corresponde a tasas de rendimiento y tasa técnica para el cálculo de la unidad de renta vitalicia al momento de cálculo de la pensión, sin embargo, el ejercicio guarda consistencia pues considera los mismos supuestos en el ejercicio de realizar o no el ahorro voluntario. El impacto del ahorro voluntario en la mejora de pensiones es relevante, por ello resulta importante buscar alternativas de incentivar el hábito de manera adecuada.



3. ANÁLISIS SISTÉMICO DEL COMPORTAMIENTO

Una parte relevante del análisis consiste en el mapeo del proceso que vive la persona en relación al comportamiento de frente al ahorro, que ayudará para aterrizar el problema e identificar aquellos puntos cruciales y aquellos que se buscarán modificar a fin de incentivar un mejor resultado, claramente lo ideal sería poder ajustar todo lo identificado que es susceptible a mejora, pero en términos prácticos se priorizará lo que está al alcance en este análisis aprovechando las herramientas disponibles de reciente elaboración como la calculadora de proyecciones de pensión.

Para esa función se toma como base el enfoque de "triada de elección" que ha sido utilizada para el diseño de políticas de salud pública y consiste en colocar el proceso de la decisión a través de 3 dimensiones; la postura de elección, arquitectura e infraestructura [Schmidt et al. \(2022\)](#). En nuestro caso, se toma la información de respuestas sobre el tema específico de ahorro voluntario observado en algunas encuestas que se han realizado con los usuarios sobre hábitos de ahorro, percepción sobre las AFORE y conocimiento sobre el sistema de ahorro para el retiro.

En el proceso de decisión del ahorro voluntario en la AFORE, las tablas 1-3 describen algunas de las premisas que pudieron identificarse y algunas observaciones generales de lo que podría usarse para mitigar estas premisas.

Premisas en la postura de decisión	Observaciones
Falta de interés en mejorar la pensión mediante ahorro voluntario	Asumir la responsabilidad de ahorro para aquellas personas que pueden realizarlo.
Visión de ahorro para el retiro únicamente como responsabilidad del estado.	Sí hay personas quienes realizan ahorro adicional en su cuenta de retiro adicionales a las obligatorias buscando conscientemente incrementar su pensión. Realizar una aportación puede ser sencilla una vez que sea ha realizado anteriormente.

■ **Tabla 1** Posturas vs. observaciones.

A partir de las tablas 1-3, se identifica la falta de una herramienta que pueda apoyar en la dimensión de la arquitectura de decisión, pues de entrada se carece de un monto de pensión estimado personal, por lo que no se tiene una idea de que tan bajo puede ser y aun más que tanto sería posible mejorarlo en cada caso, o si en la vida práctica en el contexto social actual puede llevarse a cabo y en que nivel, es decir, cual sería el potencial de ahorro voluntario de cada persona.

Para aterrizar de mejor forma una posible solución



Premisas de arquitectura de decisión	Observaciones
Identificación perfiles de clientes, algunos por regulación no alcanzarán pensión, a algunos no ayuda realizar ahorro pues solo lograrán pensión garantizada, algunos son altamente potenciales a poder realizar el ahorro	Se pueden realizar segmentaciones de clientes incluyendo un criterio de su potencial ahorro adicional.
Los clientes desconocen cual es su estimado de pensión y cuanto podrían llegar a mejorarla, suelen sobreestimar ese valor.	Informar sobre su pensión calculada, aprovechando la promoción del ahorro voluntario de una manera funcional

■ **Tabla 2** Arquitectura vs. observaciones.

Premisas en la infraestructura de decisión	Observaciones
Gran parte de las personas desconocen los canales para realizar aportaciones	Las comunicaciones oficiales como estados de cuenta son mas vistos que cualquier otra promoción de publicidad, incluir en ellos ligas de acceso a esos canales e información sobre ellos.
Se enfrentan dificultades durante el proceso operativo de transferencia a pesar de que ya se tenía la decisión de ahorrar. Problemas con descargas de aplicaciones, validaciones de identidad.	Realizar análisis sobre puntos de abandono en los procesos operativos para simplificar o implementar incentivos para finalizar el proceso.

■ **Tabla 3** Infraestructura vs. observaciones.

se sigue una estrategia basada en el modelo de toma de decisiones CREATE, que sigue un proceso en las fases: señal, reacción, evaluación, habilidad y urgencia. Vea el libro de [Wendel \(2013\)](#). (El nombre CREATE es un acrónimo que corresponde a las fases del proceso en Inglés: Cue-Reaction-Evaluation-Ability-Timing.)

- **Señal.** Se ha identificado que un bajo porcentaje de clientes atiende las comunicaciones sobre el ahorro voluntario, en contraste, el índice de apertura de mensajes que abordan el estado de cuenta es considerablemente mayor. Adicional, las comunicaciones de ahorro voluntario no incluyen elementos monetarios que ayuden al cliente a visualizar su caso específico por lo que es difícil ubicarse en que situación de ahorro podrían estar.
Se propone enviar la comunicación que incluya montos de pensión personalizados y de posible ahorro voluntario en el mismo mensaje de estado de cuenta.
- **Reacción.** Existen prejuicios o preconcepciones que tienen los clientes sobre el ahorro. Se han identificado opiniones sobre la promoción del ahorro voluntario, entre las principales están; que el cliente manifiesta desinterés por la poca información que recibe o de baja calidad pues no siente identificado

su caso, tiene desconfianza en el ahorro formal, cree que los mecanismos son complicados y los que han llegado a ahorrar no son consistentes.

Se agregará el cálculo de pensión personalizado y un comparativo con un punto de referencia buscando que el cliente se sienta identificado y tenga una mejor respuesta a la comunicación.

- **Evaluación.** Una de las grandes dificultades en el ahorro a largo plazo es que resulta complicado tangibilizar los beneficios y costos asociados al ahorro, a pesar de que se tiene la idea positiva sobre el ahorro, al no visualizar una cantidad monetaria específica y que ésta sea alcanzable resultará complicado tomar una acción.

Es crucial que cada persona visualice el beneficio, costo y una meta alcanzable, por ello se van a agregar esos tres elementos a la propuesta de decisión, es decir, se mostrará el incremento en nivel de pensión, cuanto ahorrar adicional y sobre todo que ese monto sea alcanzable, por ello se tomará como referencia a personas con características demográficas y económicas equivalentes pero que sí realizan el ahorro voluntario actualmente.

- **Habilidad.** Se tiene una idea muy general sobre el ahorro en México, y regularmente se tiende a sobreestimar el conocimiento, el ejemplo está en que a pesar de que hoy ya se cuenta con las herramientas de cálculo de pensión en los portales de varias de las AFORE y CONSAR, y que con ellas se permite realizar diversos escenarios para darse una mejor idea de como se modifica la pensión con el ahorro, poca es la gente que las utiliza. Con eso en mente es que se decide acercar el cálculo a los clientes con una sugerencia de posible ahorro que puede realizar y de esa manera reducir la carga cognitiva y el tiempo en aprender a usar herramientas. No obstante, se deja la liga a la calculadora para quien desee realizar sensibilidades a su cálculo.

- **Urgencia.** Se tiene poco sentido de urgencia sobre el asunto pensional, debido a que la mayoría de las personas ubican el retiro en un punto lejano del tiempo. hasta ahora se han realizado esfuerzos cualitativos como enviar mensajes tales como “comienza tu ahorro ya” pero no se ha acompañado de un resultado numérico o “tangible”.

Mostrar el monto de pensión bajo y la brecha que se tiene en esa pensión en comparación con quienes en circunstancias relativamente equivalentes si realizan ahorro, acompañado de un mensaje directo sobre esa situación, se espera que despierte ese sentido de

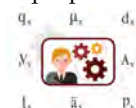
llamado a la acción.

4. METODOLOGÍA

Con los hallazgos y observaciones derivados de las metodologías revisadas, se arma una propuesta de respuesta que se centra en la creación de una herramienta que permita personalizar un resultado y ofrecerle a cada cliente una recomendación puntual sobre el ahorro voluntario, incorporando elementos de comunicación para incentivar el llamado a la acción.

Para incorporar apropiadamente elementos en la comunicación, y conforme a lo hallado en los modelos revisados, se consideran algunos de los sesgos cognitivos identificados por la economía del comportamiento que guardan relación con las barreras identificadas en el ahorro, como los son el “anclaje” (vea el libro de [Tversky y Kahneman \(1974\)](#)) o punto de referencia, que es la tendencia a depender de una información inicial o ancla al tomar decisiones, en este sentido, se incorporará en el diseño el monto de pensión de referencia que corresponderá a cliente que hace ahorro actualmente. En el mismo sentido del punto de referencia, pero con respecto a alguien influyente o autoridad, es decir el sesgo de “autoridad” (vea el trabajo de [Milgram \(1963\)](#)), en el que las personas tienden a seguir o creer en figuras de autoridad, se incorpora un punto de referencia con respecto a lo que regularmente se menciona como tasa de reemplazo comentado en algunos foros internacionales como OCDE. Otro de los sesgos relacionados a los hallazgos es el correspondiente al sesgo de “descuento hiperbólico” (vea el trabajo de [Ainslie \(1975\)](#)), referente a que las personas suelen preferir recompensas inmediatas, aunque sean menores en lugar de recompensas mas grandes a futuro, por tanto, en nuestro planteamiento se presentará directamente el costo o aportación que se requiere hacer para obtener un incremento en la pensión, monto que también se visualizará de tal manera que sea comparativo. Finalmente, el sesgo de “recarga cognitiva” (vea el trabajo de [Sweller \(1988\)](#)) que se refiere a cuando las personas tienen demasiada información o tareas que procesar pueden sentirse abrumadas y pueden tomar decisiones erróneas u optar por la procrastinación citada en el artículo de [Steel \(2007\)](#).

En lo correspondiente a la implementación de estos planteamientos, a partir de que se cuenta con la herramienta de proyección de pensiones, se tiene pendiente definir ese punto de referencia como meta alcanzable personalizada. Para la determinación de esa información se hará uso de un algoritmo de aprendizaje máquina, específicamente se determinan grupos de personas que puedan



ser comparables, esta construcción se realizará con el algoritmo conocido como *K-Means*. El planteamiento consiste en establecer grupos de clientes según sus características en común que hoy ya cuentan con ahorro voluntario, de esta manera se plantea la hipótesis de que clientes no ahorradores podrían hacerlo pues hay quienes con características similares ya lo hacen, es decir, existe un monto de ahorro potencial que no están realizando y que podría estar a su alcance. Ese monto se informa para cada cliente con respecto a su referente de grupo correspondiente.

El proceso operativo

- I. Estimación de pensiones de todos los clientes, se sigue la metodología definida para la calculadora desarrollada por CONSAR.
- II. Construcción de grupos o clusters de clientes ahorradores que tienen características similares. Se seleccionan clientes que hoy cuentan con saldo o aportaciones al ahorro voluntario y haciendo uso de su información económica demográfica (edad, género, sueldo, saldo en cuenta de retiro y semanas cotizadas) se utiliza un algoritmo de aprendizaje máquina (*K-Means*), dentro de cada grupo, se selecciona el caso que se encuentra en el percentil 70 %, de esta forma se define como referencia un caso superior a la mediana por grupo de clientes ahorradores pues el objetivo es una mejora continua.
- III. Se realizan las predicciones de clientes no ahorradores con el modelo definido o entrenado con el grupo de ahorradores, es decir, la predicción del grupo al que pertenecería cada cliente que no tiene saldo o aportaciones de ahorro voluntario. De esta manera se busca asociar a los clientes que no realizan ahorro con un caso de referencia superior a la mediana de los que ahorran.
- IV. Se compara la pensión de cada cliente no ahorrador con la pensión (superior) del cliente ahorrador, acción que se ejecuta para cada grupo. De esta manera se cuenta con el diferencial de pensión, sin embargo, debe estimarse el nivel de aportación a realizar de manera voluntaria.
- V. Con apoyo de un método numérico se obtiene el monto de ahorro requerido por cada persona para alcanzar a su pensión de referencia, acorde a las características asociadas.

En términos generales, la metodología adopta un enfoque mixto cuantitativo y cualitativo que busca incentivar

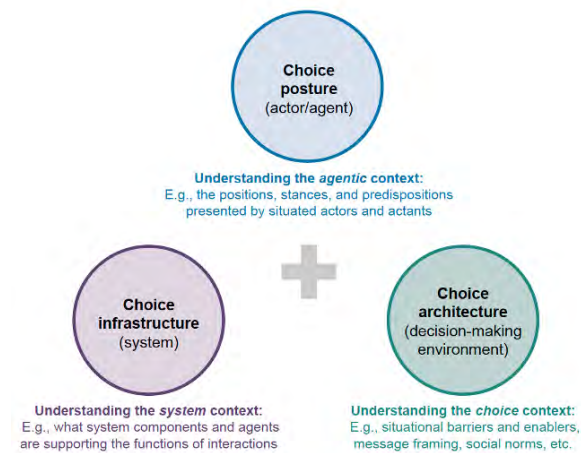


Figura 1 Fuente: Schmidt *et al.* (2022).

a clientes no ahorradores pero que tienen la posibilidad a hacerlo, presentándoles una referencia alcanzable y personalizada al mismo tiempo influenciada de por otros clientes del mismo perfil.

5. MARCO TEÓRICO

Economía del comportamiento

La tríada de elección: postura de elección, arquitectura de elección, y elección de infraestructura

El modelo de la "tríada de elección" propone tres dimensiones clave al diseñar la toma de decisiones y el comportamiento en contextos de sistemas complejos como la salud pública en lugar de centrar o diseñar un servicio producto o experiencia específica. La primera es la postura de elección, que representa la posicionalidad y predisposición de los agentes del sistema humanos y no humanos. El segundo es la arquitectura de elección, o el diseño de entornos de toma de decisiones para fomentar comportamientos deseables. Finalmente, el tercero es la infraestructura de elección, en forma de estructuras de sistemas subyacentes que apoyan y amplían comportamientos o intervenciones (vea el trabajo de Schmidt *et al.* (2022)).

Modelo CREATE. Es un marco de trabajo que facilita el desarrollo de intervenciones de cambios de comportamiento efectivas de manera práctica sin la necesidad de contar con un conocimiento detallado y profundo de las ciencias del comportamiento.

El modelo CREATE plantea una estrategia de cambio de conducta mediante la estructuración de una secuencia de pasos que guardan relación con el entorno y el individuo (vea el libro de [Wendel \(2013\)](#)).

- **Señal.** Es necesario que algo haga que la persona empiece a pensar en la acción. Puede ser algo que ya está en la vida diaria de la persona o algo que el producto o comunicación hace para llamar la atención.
- **Reacción.** La mente intuitiva reaccionará automática y rápidamente a la idea de realizar la acción. Esa reacción incluye una sensación básica sobre si la acción es interesante y placentera, también activará pensamientos sobre otras ideas y acciones posibles y relacionadas. El producto debe superar esta reacción sin ser rechazado ni distraer a la persona.
- **Evaluación.** La mente consciente evaluará los costos y beneficios de la acción, incluido el valor que proporciona el producto y las fricciones u otros desafíos al usarlo. Esa evaluación es relativa, no absoluta: la acción debe ser lo suficientemente valiosa y mejor, en términos netos, que las otras cosas que esa persona está pensando en hacer en ese momento.
- **Habilidad.** La persona realmente tiene que poder actuar en ese momento. Las posibles barreras incluyen no saber qué hacer, no tener lo necesario para actuar o sentir que fracasará, estos pueden superarse, pero lleva tiempo. Cuando se resuelve una barrera, se comienza de nuevo; todos los demás factores aún deben estar presentes.
- **Urgencia.** Tiene que haber una razón para actuar ahora, en lugar de hacer algo más urgente. Si el mensaje puede esperar y hay algo más que se debe hacer, lo pospondrá incluso si quiere y tiene la capacidad de hacerlo.

Aprendizaje máquina

K-means Clustering. La segmentación de datos en grupos homogéneos es un desafío fundamental en el campo del reconocimiento de patrones. Uno de los algoritmos más destacados para abordar este problema es el algoritmo *k*-means.

k-means es una técnica que busca dividir un conjunto de N observaciones en K clústeres distintos, de tal manera que cada observación se asocie con el clúster cuya media (o centroide) sea más cercana. La cercanía se mide comúnmente mediante la distancia Euclidiana, aunque otras métricas de distancia también pueden ser aplicadas. Vea el libro de [Bishop \(2006\)](#).

Denotemos a cada observación como un vector $x_n \in$

$\mathbb{R}^D$, donde D es la dimensionalidad del espacio de características.

El objetivo es encontrar K centroides $\{\mu_k\}$, con $k = 1, \dots, K$, que minimizan la función objetivo J :

$$J = \sum_{n=1}^N \sum_{k=1}^K r_{nk} \|x_n - \mu_k\|^2,$$

donde $r_{nk} \in \{0, 1\}$ es una variable binaria que indica si la observación n es asignada al clúster k . Específicamente $r_{nk} = 1$ si la observación n es asignada al cluster k y $r_{nk} = 0$ en caso contrario.

El proceso del algoritmo *k*-means puede describirse de la siguiente manera:

- **Inicialización:** Se escogen K puntos iniciales μ_k como centroides del conjunto de datos. Estos pueden ser seleccionados aleatoriamente, aunque existen diversas heurísticas que buscan mejorar la convergencia del algoritmo mediante una inicialización más informada.
- **Asignación:** Cada n -ésimo dato es asignado al centroide más cercano según:

$$r_{nk} = \begin{cases} 1 & \text{si } k = \arg \min_j \|x_n - \mu_k\|^2 \\ 0 & \text{en otro caso} \end{cases}$$

Esto resulta en una partición del conjunto de datos en K clústeres.

- **Actualización:** Los centroides μ_k se recalculan. Para cada clúster, esto se logra promediando todos los puntos asignados a ese clúster.

$$\mu_k = \frac{\sum_{n=1}^N r_{nk} x_n}{\sum_{n=1}^N r_{nk}}$$

- **Convergencia:** Los pasos de asignación y actualización se repiten hasta que los cambios en J sean insignificantes, es decir, que los centroides no cambien significativamente entre iteraciones o hasta que se alcance un criterio de parada predefinido.

El objetivo primordial de *k*-means es minimizar una función objetivo conocida como la suma de las distancias cuadradas entre los puntos y sus centroides asociados.

Sin embargo, *k*-means presenta algunos desafíos, es sensible a la inicialización, y existe el riesgo de que converja a mínimos locales. Esto ha llevado al desarrollo de múltiples variantes y técnicas de inicialización para mejorar su robustez y eficacia.

En el contexto del reconocimiento de patrones, el algoritmo *k*-means es frecuentemente utilizado como un



paso preliminar para la reducción de la dimensionalidad, la cuantización de vectores, entre otras aplicaciones. La simplicidad conceptual de k-means, junto con su capacidad de segmentar eficientemente grandes conjuntos de datos, lo hacen una herramienta valiosa en este campo.

6. INTERVENCIÓN

La hipótesis para probar consiste en: Se abren mas contratos para las aportaciones voluntarias asociados a los clientes que recibieron la comunicación con la pensión de referencia y sugerencia de monto de aportación que aquellos que solo reciben cálculo de pensión.

En la fase inicial de la experimentación, se presenta el diseño de comunicación en sesiones de “focus group” para conocer las percepciones y opiniones sobre la propuesta, se obtuvo que los participantes presentaron reacciones de interés y sorpresa en ambas comunicaciones pues el monto de la pensión resulta ser menor de lo que esperaban. Se destacó el interés en el mecanismo de ahorro voluntario en todos los clientes que vieron la propuesta de pensión de referencia, señalaron que además de que los alivia saber que es un fenómeno general el bajo nivel de pensión, los motiva saber que es posible mejorar en la medida de sus posibilidades, es decir, el monto sugerido a ahorrar voluntariamente está a su alcance. En clientes a los que solo se muestra el resultado de la pensión se tiene una reacción de preocupación, una actitud que se conserva al no haber una propuesta de solución puntual o específica, se pasa por alto el consejo que invita a realizar el ahorro voluntario en una leyenda al final del mensaje.

Con ese antecedente, se comienza con la prueba A/B seleccionado muestra de 2000 clientes, a fecha actual aún no se realiza el envío, pues se sugiere acompañarlo con el próximo envío de estado de cuenta, recordando la relevancia de la comunicación periódica asociada al estado de cuenta y no en una comunicación específica o aislada, pues la apertura de la información referente a los estados de cuenta ha resultado mayor.

Finalmente, se presenta un ejemplo del diseño para la comunicación que atiende a los planteamientos acorde a los análisis realizados, claramente cada cliente tiene una estimación específica con su correspondiente referencia personal del grupo al que corresponde. Vea la figura 2.

7. CONCLUSIÓN

En un contexto en que las tasas de reemplazo de las pensiones a menudo resultan insuficientes para mantener el nivel de vida de los trabajadores, el ahorro voluntario



Figura 2 Ejemplo de elemento de comunicación y pensión de referencia

se presenta como una alternativa que estructuralmente esta preparada para activarse, la promoción del ahorro voluntario resulta relevante para complementar y mejorar los resultados actuales. Este trabajo además de resaltar la relevancia en sí del ahorro destaca la importancia de realizar investigación de las causas, así como incorporar el análisis para el entendimiento de la conducta humana relacionado con las decisiones para desarrollar una respuesta efectiva que ayude a superar las barreras cognitivas y de comportamiento que pudieran estar inhibiendo el ahorro. La intervención propuesta además de buscar incrementar el saldo ahorrado para el retiro de las personas brinda una idea a utilizarse como alternativa para el incentivo de la cultura de previsión y responsabilidad financiera, la metodología y estrategias empleadas tienen directa aplicabilidad en industrias que enfrentan una problemática similar como lo es el caso de seguros con componente de ahorro.

REFERENCIAS

Ainslie, G., 1975 Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin* 82: 463.

Bishop, C. M., 2006 *Pattern Recognition and Machine Learning*.

Instituto Mexicano del Seguro Social, 2023 *Nota metodológica de calculadora de ahorro y retiro para trabajadores que cotizan al IMSS*. Comisión Nacional para el Retiro, Disponible [aquí](#).

Milgram, S., 1963 Behavioral study of obedience. *The Journal of Abnormal and Social Psychology* 67: 371.

Schmidt, R., Z. Chen, y S. V. Paz, 2022 Choice posture, architecture, and infrastructure: Systemic behavioral design for public health policy. *She Ji: The Journal of Design, Economics, and Innovation* 8.

Steel, P., 2007 The nature of procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychological Bulletin* **133**: 65.

Sweller, J., 1988 Cognitive load during problem solving: Effects on learning. *Cognitive Science* **12**: 257–285.

Tversky, A. y D. Kahneman, 1974 *Judgment under uncertainty: Heuristics and biases*, volumen 185. Science.

Wendel, S., 2013 *Designing for Behavior Change: Applying Psychology and Behavioral Economics*. O'Reilly.



Antonio Cruz Torres se graduó de la Licenciatura en Actuaría de la Facultad de Ciencias de la Universidad Nacional Autónoma de México en 2007; y de la Maestría en Ciencia de Datos de la Universidad Panamericana en 2021. Es un profesional de seguros, pensiones, AFORE y master en ciencia de datos con mas de 15 años de experiencia en riesgos financieros y técnica actuarial dentro de empresas líderes del sector financiero y asegurador. Ha diseñado e implementado estrategias para gestión de riesgos de las unidades de negocio de seguros de vida y pensiones, formando y consolidando el área responsable. Impulsó la reapertura de la cartera de rentas vitalicias. Desarrolló estrategias y modelos para potenciar la generación de valor en sectores asegurador y de retiro. Su formación y experiencia abarca riesgos financieros, ciencias actuariales, finanzas corporativas, inversiones y ciencia de datos.

Es un profesional de seguros, pensiones, AFORE y master en ciencia de datos con mas de 15 años de experiencia en riesgos financieros y técnica actuarial dentro de empresas líderes del sector financiero y asegurador. Ha diseñado e implementado estrategias para gestión de riesgos de las unidades de negocio de seguros de vida y pensiones, formando y consolidando el área responsable. Impulsó la reapertura de la cartera de rentas vitalicias. Desarrolló estrategias y modelos para potenciar la generación de valor en sectores asegurador y de retiro. Su formación y experiencia abarca riesgos financieros, ciencias actuariales, finanzas corporativas, inversiones y ciencia de datos.



PREMIO FUNDADORES

2do lugar. Una nueva aproximación a la constitución de una reserva actuarial de arriba p'abajo

José Daniel López Barrientos y Humberto Plata
Gallegos



Una nueva aproximación a la constitución de una reserva actuarial de arriba p'abajo

José Daniel López Barrientos^{*,1} y Humberto Plata Gallegos^{*,2}

*Facultad de Ciencias Actuariales de la Universidad Anáhuac México

RESUMEN Presentamos una alternativa a la tendencia actual de la modelación matemática de la constitución de reservas en seguros y pensiones. Utilizamos herramientas comunes en el análisis actuarial, así como de algunas técnicas de control estocástico con el fin de elaborar una metodología para constituir un fondo de pensiones que permita a sus afiliados cobrar las rentas vitalicias que les corresponden, manteniendo una probabilidad prescrita de ruina. Entre las técnicas actuariales usadas están el modelo de Cramér-Lundberg con primas y derechos contingentes, y las recomendaciones de Bühlmann para calcular las primas en los niveles superior e inferior de una cartera. De la teoría del control estocástico, recurrimos a la optimización de un proceso Markoviano de decisión donde el conjunto de estados es un espacio finito bajo el criterio del pago promedio y el uso de una versión de la ley fuerte de los grandes números para considerar las trayectorias del proceso de riqueza, en lugar de sus valores esperados. El hilo conductor del artículo es la implementación computacional del método de valuación de las pensiones en el caso Mexicano con datos recientes. Hemos creado una herramienta computacional que permite replicar los resultados de esta investigación, así como generar otros modificando nuestros supuestos.

Palabras clave

Modelo de Cramér-Lundberg con primas estocásticas, Simulación estocástica, Teoría de filas para procesos w -geométricamente ergódicos.



Derechos reservados © 2023 por el Colegio Nacional de Actuarios

Última actualización del manuscrito: 17 de noviembre de 2023

Este artículo es publicado por el Colegio Nacional de Actuarios (CONAC) y contiene información de una variedad de fuentes. Es un trabajo cuyo fin es únicamente informativo y no debe interpretarse como asesoramiento profesional o financiero. La visión y comentarios contenidos en este documento corresponden exclusivamente a su autor y, por tanto, son ajenos a cualquier entidad pública o privada, incluyendo aquella para la que actualmente colabora. El CONAC no recomienda ni respalda el uso de la información proporcionada en este estudio. El CONAC no ofrece ninguna garantía, expresa o implícita, ni representación de ningún tipo y no asume ninguna responsabilidad en relación con el uso o mal uso de este trabajo.

1. INTRODUCCIÓN

Este artículo propone una visión novedosa sobre uno de los problemas centrales en el *corpus* de las ciencias actuariales mismas: encontrar una política de contribuciones a un fondo de pensiones (o a la reserva de una aseguradora) tal que la probabilidad de la ruina no rebase un nivel fijado con anterioridad.

El Estado del Arte

Existen varios estudios acerca del estado y la sostenibilidad de los planes de pensiones en México. Los trabajos de Banco de México (2021); Cervantes González y Sedano Díaz (2023) y de Villarreal y Macías (2020) son ejemplos de esto. Las pensiones de nuestro interés pertenecen al sistema contributivo en México (vea las secciones 1.B y 2.A.1 del artículo de Villarreal y Macías (2020)). Partiremos del modelo clásico de Cramér-Lundberg como fue presentado en el décimo tercer capítulo del libro de Bowers *et al.* (1997), el cuarto capítulo de la obra de Kaas *et al.* (2008), el apéndice D del texto de Schmidli (2008), el cuarto capítulo del volumen de Mikosch (2009), el segundo capítulo del trabajo de Ramasubramanian (2009), el vigésimo primer capítulo del tomo de Promislow (2011), etcétera; y de la variante propuesta por Boikov (2002) para reflejar la posibilidad de que haya incertidumbre sobre el pago de las primas. Acto seguido, nos valdremos de los resultados presentados por López-Barrientos *et al.* (2023), el undécimo capítulo del libro de Hernández-Lerma y Lasserre (1999) y la heurística presentada en el artículo de Cano Ramos y López Barrientos (2021) para exhibir una política de contribuciones al fondo en su nivel superior, y un capital inicial tales que la probabilidad de que el proceso de la riqueza caiga por debajo de cero sea conocida con antelación.

Históricamente, el problema en consideración ha recibido muchos tratamientos, pero la conexión de los resultados analíticos con la práctica actuarial cotidiana se reduce al cálculo de una cota superior para la probabilidad de la ruina. Entre los desarrollos teóricos más importantes –y famosos–, podemos destacar los capítulos 1.3 y 4.2.2 en los libros de Schmidli (2008) y Mikosch (2009), respectivamente. El primero usa las mismas técnicas de programación dinámica que nosotros para estudiar el problema de encontrar una estrategia de reaseguro óptimo (con primas deterministas) para minimizar la probabilidad de la ruina de una compañía de seguros; mientras que el segundo presenta cálculos explícitos para obtener la probabilidad de la ruina de una reserva en un tiempo

finito. Por otro lado, las contribuciones de Aguirre-Farías (2012, 2001); Anderson (1992); Bühlmann (1985); Cano Ramos y López Barrientos (2021); Iyer (1999), (Daykin *et al.* 1993, Chapter 16) y Plata Gallegos y López Barrientos (2023) resultan referencias importantes en términos de la metodología realmente empleada por los actuarios a nivel profesional para valorar pasivos contingentes y/o carteras de seguros.

Nuestro plan de ataque

Nuestro cometido es determinar una combinación de los parámetros de un fondo de tal suerte que su probabilidad de insolvencia en horizonte finito no rebase un nivel fijado con antelación. Para lograrlo, partiremos del modelo de Cramér-Lundberg con primas estocásticas mostrado por Boikov (2002) para implementar el método de simulación de Monte Carlo propuesto en el artículo de Cano Ramos y López Barrientos (2021) usando la técnica que Belolipetskii y Lepskaya (2018, 2019) aplicaron a la Federación Rusa, pero con las proyecciones poblacionales de Consejo Nacional de Población *et al.* (2019) aplicadas a los datos estudiados por Saad *et al.* (2017), y bajo las hipótesis actuariales de Aguirre-Farías *et al.* (2019). Lo anterior nos dará una medida del desempeño del fondo a lo largo del tiempo. Posteriormente, garantizaremos que las condiciones del teorema 11.2.1 del libro de Hernández-Lerma y Lasserre (1999) se cumplen y así estimaremos el capital inicial necesario para sostener el fondo en un horizonte *lo suficientemente largo*. Finalmente, encontraremos la estructura del portafolios con un enfoque de arriba hacia abajo, como lo sugirió Bühlmann (1985). Nuestra contribución supone una novedad técnica con respecto a los métodos clásicos expuestos en el trabajo recepcional de Plata Gallegos y López Barrientos (2023) y los capítulos 15 y 16 del libro de Daykin *et al.* (1993).

El resto del artículo...

Dedicamos la sección 2 a establecer las definiciones básicas necesarias para construir nuestro trabajo. La sección 3 nos presenta la implementación numérica de nuestro modelo, y la correspondiente medición de su desempeño, bajo las hipótesis que aplican al caso Mexicano. Posteriormente, en la sección 4 enunciaremos una versión *ad hoc* del teorema central del límite para estimar la probabilidad de la ruina mediante los criterios de control estocástico de la esperanza y la varianza promedio; y calcularemos las primas en los niveles superior e inferior del portafolios. Planteamos nuestras conclusiones, y comentamos los alcances y limitaciones

¹ Correo electrónico: daniel.lopez@anahuac.mx

² Correo electrónico: humberto.plata@anahuac.mx



de nuestro estudio en la sección 5.

Hemos creado una herramienta computacional que permite replicar los resultados de esta investigación, así como generar otros modificando nuestros supuestos. Tal herramienta está disponible [aquí](#).

2. PRELIMINARES ACTUARIALES Y DE PROCESOS ESTOCÁSTICOS

Hipótesis demográficas

Comenzamos nuestro camino definiendo $\mathcal{L}(t+1, x+1)$ a la manera de la sección 3.3.1 del libro de Bowers *et al.* (1997): como el número de sobrevivientes a edad x de una población que tenía $\mathcal{L}(t, x)$ miembros al tiempo t . Así, $\mathcal{L}(t+1, x+1)$ es un representante de los llamados *sistemas aleatorios iterados* presentados en el artículo de Diaconis y Freedman (1999), el ejemplo 7.4.7 del libro de Hernández-Lerma y Lasserre (1999), y el capítulo 12 del libro de Lasota y Mackey (1994):

$$\mathcal{L}(t+1, x+1) := \sum_{j=1}^{\mathcal{L}(t,x)} \mathbb{1}_j(x), \quad (1)$$

donde

$$\mathbb{1}_j(x) = \begin{cases} 1 & \text{si la } j\text{-ésima vida subsiste a edad } x, \\ 0 & \text{si no.} \end{cases} \quad (2)$$

Desde luego, para todo j , tenemos que $\mathbb{1}_j(x)$ es una variable aleatoria Bernoulli cuyo parámetro es la función p_x del cálculo actuarial elemental. O sea, $\mathbb{P}(\mathbb{1}_j(x) = 1) = p_x$, y $\mathbb{P}(\mathbb{1}_j(x) = 0) = q_x$.

A continuación, seguimos a Belolipetskii y Leps kaya (2018, 2019) para definir el vector

$$\vec{\mathcal{L}}_t = (\mathcal{L}(t, 0), \mathcal{L}(t, 1), \dots, \mathcal{L}(t, \alpha), \dots, \mathcal{L}(t, r), \dots, \mathcal{L}(t, \omega))^T, \quad (3)$$

con $\mathcal{L}(t, x)$ independiente de $\mathcal{L}(t, y)$ para todos $t = 0, 1, \dots$, y $x \neq y$; y $\mathcal{L}(t, 0) = \ell_0(t)$, donde $\ell_0(t)$ es una variable aleatoria que representa el tamaño de la cohorte en el momento $t = 0, 1, \dots$. Se trata de la *radix* de la cohorte a la que damos seguimiento en el momento $t = 0, 1, \dots$

Notación. El símbolo A^T denota la transpuesta de la matriz A . Asimismo, usamos el símbolo $\vec{u}$ para denotar un vector cuyas componentes son $u_1, \dots, u_n$.

Ejemplo 2.1. Suponiendo que hacemos $t = 0$ al año 2020, el renglón 0 a 4 años de la tabla 1 nos muestra $\ell_0(0)$ para los hombres y las mujeres. A partir del renglón siguiente, la misma

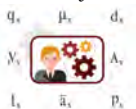


tabla presenta las radices $\ell_x(0)$ para $x = 5, 10, \dots, 100$ para hombres y mujeres.

La figura 1 es una representación gráfica de la información reflejada en la tabla 1 (o sea, del vector $\vec{\mathcal{L}}_0$ en México para hombres y mujeres). En el contexto demográfico, esa forma de presentar los datos se llama “pirámide de población” en 2020.

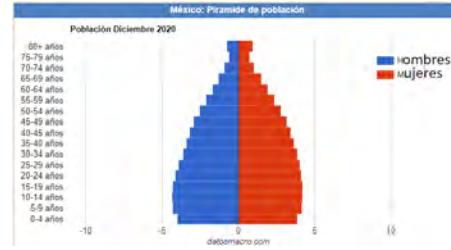


Figura 1 Pirámide poblacional para México en 2020. Fuente: Datos Macro (2021).

La relación (1) es importante porque, de acuerdo al resultado siguiente, demuestra que el proceso estocástico $\vec{\mathcal{L}} := (\vec{\mathcal{L}}_t : t = 0, 1, \dots)$ aludido en (3) es una cadena de Markov de dimensión $\omega + 1$ cuyo espacio de estados es $(\mathbb{N} \cup \{0\})^{\omega+1}$ (aquí, $\mathbb{N}$ denota el conjunto de los números Naturales). Incluimos la prueba para que nuestra presentación se mantenga completa.

Proposición 2.2. (Norris 1997, Exercise 1.1.6). Sean $X := (X_t : t = 0, 1, \dots)$ y $Y := (Y_t : t = 0, 1, \dots)$ dos procesos estocásticos definidos en el espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$; y con espacios de estados $(S_1, \mathcal{G}_1)$ y $(S_2, \mathcal{G}_2)$, respectivamente, donde $\mathcal{G}_1$ y $\mathcal{G}_2$ son las σ -álgebras de Borel de S_1 y S_2 . Suponga también que $Y_0, \dots, Y_t$ son mutuamente independientes. Si X_0 es independiente de Y y existe una función de Borel $G : S_1 \times S_2 \rightarrow S_1$ tal que $X_{t+1} = G(X_t, Y_{t+1})$ para toda $t = 0, 1, \dots$, entonces X_t es una cadena de Markov.

Demostración. Para verificar que X satisface la propiedad de Markov, note que para cada $t = 0, 1, 2, \dots$, las variables aleatorias $(X_0, \dots, X_t)$ y $(Y_{t+1}, \dots)$ son mutuamente independientes. Entonces, para todo conjunto de Borel $B \in \mathcal{G}_1$, lo siguiente rige.

$$\begin{aligned} & \mathbb{P}(X_{t+1} \in B | X_0 = x_0, \dots, X_t = x_t) \\ &= \mathbb{P}(G(x_t, Y_{t+1}) \in B | X_0 = x_0, \dots, X_t = x_t) \\ &= \mathbb{P}(G(x_t, Y_{t+1}) \in B) \text{ (por independencia)} \\ &= \mathbb{P}(G(X_t, Y_{t+1}) \in B | X_t = x_t) \\ &= \mathbb{P}(X_{t+1} \in B | X_t = x_t). \end{aligned}$$

Esto es lo que queríamos demostrar. $\square$

Grupo quinquenal de edad	2020		
	Total	Hombres	Mujeres
Total	126,014,024	61,473,390	64,540,634
0 a 4 años	10,047,365	5,077,482	4,969,883
5 a 9 años	10,764,379	5,453,091	5,311,288
10 a 14 años	10,943,540	5,554,260	5,389,280
15 a 19 años	10,806,690	5,462,150	5,344,540
20 a 24 años	10,422,095	5,165,884	5,256,211
25 a 29 años	9,993,001	4,861,404	5,131,597
30 a 34 años	9,420,827	4,527,726	4,893,101
35 a 39 años	9,020,276	4,331,530	4,688,746
40 a 44 años	8,503,586	4,062,304	4,441,282
45 a 49 años	7,942,413	3,812,344	4,130,069
50 a 54 años	7,037,532	3,332,163	3,705,369
55 a 59 años	5,695,958	2,692,976	3,002,982
60 a 64 años	4,821,062	2,257,862	2,563,200
65 a 69 años	3,645,077	1,706,850	1,938,227
70 a 74 años	2,647,340	1,233,492	1,413,848
75 a 79 años	1,814,582	847,898	966,684
80 a 84 años	1,175,364	523,812	651,552
85 a 89 años	659,245	283,351	375,894
90 a 94 años	266,806	107,358	159,448
95 a 99 años	95,205	36,615	58,590
100 años y más	18,295	6,644	11,651
No especificado	273,386	136,194	137,192

■ **Tabla 1** Población total por entidad federativa y grupo quinquenal de edad según sexo, serie de años censales de 1990 a 2020. Fuente: [Instituto Nacional de Estadística y Geografía \(2023a\)](#).

De hecho, la función G está dada por (1), y las variables aleatorias de Bernoulli ($\mathbb{1}_j(x) = 1$) desempeñan el papel del proceso estocástico Y . Naturalmente, de (1) y (2), se deduce que la distribución condicional de la variable aleatoria $[\mathcal{L}(t+1, x+1) | \mathcal{L}(t, x)]$ es Binomial con parámetros $\mathcal{L}(t, x)$ y p_x . Más aún, todos los sub-procesos de la forma $(\mathcal{L}(t+k, x+k) : k = 0, 1, \dots)$ son filas de la forma $M/M/2 : (SIRO/\infty/\infty)$ (como las estudiadas en el capítulo 1.4 del libro de [Llanos Reynoso \(2014\)](#) y en el artículo de [Lippman y Shaler Stidham \(1977\)](#)) para $t = 0, 1, \dots$ y $x = 0, 1, \dots, \omega - 1$.

De (1), está claro que $\mathcal{L}(t, x)$ es una variable aleatoria Binomial-compuesta-Bernoulli (para más detalles, vea por ejemplo el capítulo 17 del libro de [Promislow \(2011\)](#)). Este hecho, junto con el teorema de probabilidad total implica el resultado siguiente.

Lema 2.3. *Las funciones de esperanza y varianza del proceso*

estocástico $(\mathcal{L}(t+k, x+k) : k = 0, 1, \dots)$ están dadas por:

$$\mathbb{E}[\mathcal{L}(t, x)] = \begin{cases} x p_0 \mathbb{E}[\ell_0(t-x)] & \text{si } 0 < t-x, \\ t p_{x-t+1} \mathbb{E}[\mathcal{L}(0, x-t)] & \text{si } 0 \geq t-x, \end{cases} \quad (4)$$

$$\begin{aligned} \text{var}[\mathcal{L}(t, x)] = & p_x^{2x} \left(\text{var}[\ell_0(t-x)] - \frac{x p_0 \mathbb{E}[\ell_0(t-x)]}{1 - p_x^2} \right) \\ & + \frac{x p_0 \mathbb{E}[\ell_0(t-x)]}{1 - p_x^2}, \end{aligned} \quad (5)$$



cuando $0 < t - x$; y

$$p_x^{2(x-t)} \times \left(\text{var}[\mathcal{L}(0, x-t)] - \frac{tp_{x-t+1}\mathbb{E}[\mathcal{L}(0, x-t)]}{1-p_x^2} \right) + \frac{tp_{x-t+1}\mathbb{E}[\mathcal{L}(0, x-t)]}{1-p_x^2}, \quad (6)$$

cuando $0 \geq t - x$, respectivamente.

Demostración. El teorema de probabilidad total nos da que

$$\mathbb{E}[\mathcal{L}(t, x)] = \mathbb{E}[\mathcal{L}(t-1, x-1)]p_x, \quad (7)$$

$$\text{var}[\mathcal{L}(t, x)] = p_x^2 \text{var}[\mathcal{L}(t-1, x-1)] + \mathbb{E}[\mathcal{L}(t-1, x-1)]p_x^2 q_x \quad (8)$$

para $x = 1, \dots, \omega$ y $t = 1, \dots$. Resolver la ecuación en diferencias de primer orden (7) nos da (4). La sustitución de (4) en (8) hace que ésta última sea una ecuación no homogénea en diferencias de primer orden con coeficientes constantes. Entonces, por los resultados del capítulo 4.3 del libro de [Romeo \(2020\)](#), las relaciones (5) y (6) son ciertas. Esto es lo que queríamos demostrar. $\square$

Ejemplo 2.4. Considere las raíces $\ell_0^h(0) = 5077482$ y $\ell_0^m(0) = 4969883$ de la tabla 1, y las tablas de mortalidad para hombres y mujeres de las tablas 2 y 3, respectivamente.

Usamos el Algoritmo 1 –que tomamos prestado del capítulo 4.3 del libro de [Ross \(1997\)](#) o del tercer algoritmo del artículo de [López-Barrientos et al. \(2022b\)](#)– para simular variables aleatorias con distribución Binomial y así generar sendas trayectorias de las cadenas de muertes puras a partir de las raíces recién referidas. Puede ver el resultado de esto en la figura 2.

Algoritmo 1 Generación de una variable aleatoria Binomial con parámetros m y p_x .

Generar un número v a partir de la distribución uniforme $[0, 1]$
 $n \leftarrow 0$
 $c \leftarrow \frac{p_x}{1-p_x}$
 $p \leftarrow (1-p_x)^m$
 $F \leftarrow p$
mientras $v \geq F$ **hacer** $p \leftarrow c \frac{m-n}{n+1} p$
 $F \leftarrow F + p$
 $n \leftarrow n + 1$
retorna n



Edad/Age (x)	n	q(x,n)
0	1	0.019
1	4	0.005
5	5	0.002
10	5	0.002
15	5	0.004
20	5	0.006
25	5	0.008
30	5	0.010
35	5	0.012
40	5	0.015
45	5	0.020
50	5	0.028
55	5	0.040
60	5	0.059
65	5	0.089
70	5	0.134
75	5	0.199
80	5	0.289
85	5	0.409
90	5	0.551
95	5	0.709
100	ω	1.000

■ **Tabla 2** Tabla abreviada de mortalidad para hombres Mexicanos 2020. Fuente: [Saad et al. \(2017\)](#).

Las entradas $(\ell_0(t) : t = 0, 1, \dots)$ del proceso estocástico $\mathcal{L}$ constituyen un *proceso de nacimientos puros*, como los estudiados en el capítulo 1.4 del libro de [Llanos Reynoso \(2014\)](#). Las consideraremos como variables aleatorias de Poisson con parámetro igual a la tasa (bruta) de natalidad pronosticada por el [Consejo Nacional de Población et al. \(2019\)](#). La figura 3 muestra tales tasas (juntamente con las de mortalidad, que corresponden a las exhibidas en las tablas 2 y 3).

Las tasas brutas de natalidad (TBN) mostradas en la figura 3 para la sección sombreada son las que aparecen en la tabla 4.

El trazador cúbico interpolante (vea el tercer capítulo del libro de [Burden et al. \(2015\)](#)) para las tasas mostradas en la figura 3 a partir del año 2020 está dado por:

$$\lambda(t) = -0.0061 \left(\frac{t}{5} + 1 \right)^3 + 0.1411 \left(\frac{t}{5} + 1 \right)^2 - 1.7057 \left(\frac{t}{5} + 1 \right) + 18.399 \quad (9)$$

para $t = 0, 1, \dots$. En principio, usaremos (9) para nuestras proyecciones poblacionales hasta el año 2050. Asimismo,

Edad/Age (x)	n	q(x,n)
0	1	0.015
1	4	0.004
5	5	0.001
10	5	0.001
15	5	0.001
20	5	0.002
25	5	0.003
30	5	0.003
35	5	0.005
40	5	0.007
45	5	0.011
50	5	0.017
55	5	0.027
60	5	0.043
65	5	0.067
70	5	0.105
75	5	0.161
80	5	0.242
85	5	0.353
90	5	0.494
95	5	0.659
100	69	1.000

■ **Tabla 3** Tabla abreviada de mortalidad para mujeres Mexicanas 2020. Fuente: [Saad et al. \(2017\)](#).

propondremos diferentes maneras de ajustar las tasas brutas de natalidad para los años subsecuentes al 2050. Como es común en este tipo de trabajos, supondremos que los nacimientos en cada intervalo de años siguen un proceso de Poisson, y simularemos este número a través del algoritmo 2, que tomamos del capítulo 4.2 del libro de [Ross \(1997\)](#), o del segundo algoritmo del artículo de [López-Barrientos et al. \(2022b\)](#).

Ahora considere un plan de pensiones bajo las hipótesis siguientes, que tomamos del artículo de [Belolipetskii y Lepskaya \(2019\)](#).

Hipótesis 2.5. (i) Los participantes ingresan al plan de pensiones y comienzan a pagar contribuciones a la edad de α años.

(ii) Al alcanzar la edad de r , el participante se jubila y comienza a recibir pagos de pensión.

(iii) El número de participantes mayores de ω años no es significativo respecto al tamaño de la población, y está contabilizado en $\mathcal{L}(t, \omega)$.

(iv) El número de personas de edad α que se incorporan al plan

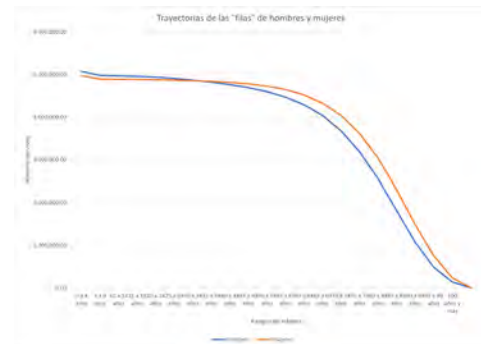


Figura 2 Posibles historias de las cohortes de damas y varones nacidos en 2020. Elaboración propia con información de [Saad et al. \(2017\)](#) e [Instituto Nacional de Estadística y Geografía \(2023a\)](#).

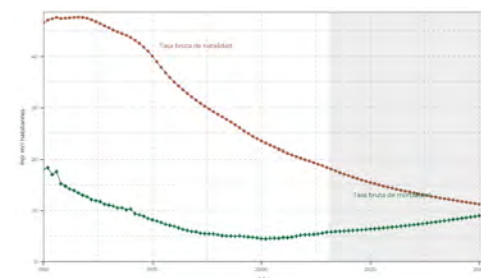


Figura 3 Tasas brutas de natalidad y mortalidad históricas y pronosticadas para el período 1950-2050. Fuente: [Consejo Nacional de Población et al. \(2019\)](#).

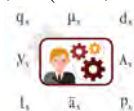
de pensiones en el año $t = 0, 1, \dots$ es una variable aleatoria cuyo soporte es el conjunto $\mathbb{N} \cup \{0\}$, y es proporcional a $\mathcal{L}(t, \alpha)$. Llamamos κ a la constante de proporcionalidad.

(v) No es posible entrar ni salir del plan de pensiones a una edad diferente a α .

(vi) Existe homogeneidad entre los integrantes del grupo en cuanto a género, evolución salarial y composición demográfica.

De acuerdo a las hipótesis 2.5(iii)-(iv), la estructura de las edades de los miembros del plan de pensiones está dada por

$$\vec{\mathcal{M}}_{t+1} := (\mathcal{M}(t+1, \alpha), \dots, \mathcal{M}(t+1, r), \dots, \mathcal{M}(t+1, \omega))^T$$



Año	2020	2025	2030	2035	2040	2045	2050
TBN	16.8283	15.5032	14.3871	13.4434	12.6355	11.9268	11.2807

■ **Tabla 4** Proyección de las tasas brutas de natalidad según el [Consejo Nacional de Población et al. \(2019\)](#).

Algoritmo 2 Generación de una variable aleatoria de Poisson con parámetro λ .

Genera un número aleatorio v
 $n \leftarrow 0$
 $p \leftarrow \exp(-\lambda)$
 $F \leftarrow p$
mientras $U \geq F$ **hacer**
 $p \leftarrow \lambda \frac{p}{n+1}$
 $F \leftarrow F + p$
 $n \leftarrow n + 1$
retorna n

$$:= \begin{pmatrix} \kappa \mathcal{L}(t+1, \alpha) \\ \vdots \\ \text{Bin}(\mathcal{M}(t, r-1), p_{r-1}) \\ \vdots \\ \text{Bin}(\mathcal{M}(t, \omega-1) \\ p_{\omega-1}) \end{pmatrix},$$

donde p_x son las probabilidades de supervivencia aludidas en las tablas 2 y 3. Esto supone una limitante a nuestro modelo, pues las tablas de mortalidad deben ser actualizadas cada cierto tiempo, y un cambio de tablas implicará la obtención de resultados distintos a los aquí presentados. Para efectos de nuestro estudio, nos referimos a los resultados presentados por la Encuesta Nacional de Ocupación y Empleo del segundo trimestre de 2023 conducida por el [Instituto Nacional de Estadística y Geografía \(2023b\)](#) para calcular la constante de proporcionalidad mencionada en la hipótesis 2.5(iv), y hacemos

$$\kappa = \frac{\text{Personas ocupadas}}{\text{Población de 15 y más años de edad}} \quad (10)$$

$$= 58.44 \, \%.$$

Los elementos recién presentados determinan por completo la estructura poblacional de nuestro interés.

Observación 2.6. *Nuestras metas son*



(a) *Medir el desempeño del fondo referido en la hipótesis 2.5 bajo diversos parámetros.*

(b) *Controlar la trayectoria que seguirá el proceso de la riqueza disponible en el portafolios a lo largo del tiempo.*

Hipótesis de ergodicidad

Grosso modo, decimos que un proceso estocástico es ergódico cuando admite una *medida de probabilidad límite* μ , también llamada *invariante*. Esto es, si X es una cadena de Markov, y P es su matriz de transición de un paso, entonces decimos que X es ergódica si

$$\lim_{n \rightarrow \infty} P^n = [\mu | \mu | \cdots | \mu]^\top \text{ y } \mu P = \mu. \quad (11)$$

Para efectos de la auto-contención de nuestro trabajo, a continuación presentamos algunos elementos técnicos que necesitaremos a lo largo de nuestros desarrollos. El lector interesado puede consultar la sección 7.D del libro de [Hernández-Lerma y Lasserre \(1999\)](#) para averiguar más detalles. Comenzamos con la definición 7.3.9 de esa obra adaptada a nuestro contexto.

Definición 2.7. Sean X una cadena de Markov con matriz de probabilidad de transición P , y $w \geq 1$ una función medible cuyo dominio coincide con el espacio de estados S_1 de X y tal que $\|P\|_w := \|P\|_\infty / \|w\| < \infty$, donde $\|\cdot\|_\infty$ es la norma del espacio de las matrices cuadradas, y $\|\cdot\|$ es una norma adecuada para la función w . Decimos que la cadena X es *w-geométricamente ergódica* si existe una medida de probabilidad μ (con $\|\mu\|_w < \infty$), y constantes no-negativas R y $\rho < 1$ que satisfacen

$$\|P^t - \mu\|_w \leq R\rho^t \text{ para todo } t = 0, 1, \dots \quad (12)$$

La definición 2.7 es muy poderosa. Es una versión fortalecida de (11), que se estudia en un curso básico de procesos estocásticos (como en la sección 1.10 del libro de [Norris \(1997\)](#)), y por lo tanto, extiende el concepto de *estado estable* de la teoría de filas presentada en el libro de [Llanos Reynoso \(2014\)](#). Como es de esperarse, cuando una cadena de Markov la cumple, podemos aseverar que si la medida de probabilidad límite μ existe, entonces es la única que satisface (12). El siguiente resultado es una consecuencia inmediata de nuestro lema 2.3; y

de la proposición 7.4.5 y el ejemplo 7.4.7 del libro de Hernández-Lerma y Lasserre (1999).

Proposición 2.8. Sea $w(\vec{\mathcal{L}}_t) := \|\vec{\mathcal{L}}_t\| + 1$. Si para toda $t = 0, 1, \dots$, se cumple que

$$\mathbb{E}[\mathcal{L}(t, x)] \leq p_0 \mathcal{L}(t, x) - q_0 \quad (13)$$

para toda $\mathcal{L}(t, x) \geq \max_t \lambda(t)$, donde $\lambda(\cdot)$ es la tasa bruta de natalidad a la que se refiere (9), entonces el sistema estocástico $(\mathcal{L}(t+k, x+k) : k = 0, 1, \dots)$ es w -geométricamente ergódico.

Boceto de la demostración. Note que

- Si $0 < t - x$, (4) nos da que

$$\mathbb{E}[\mathcal{L}(t, x)] = {}_x p_0 \mathbb{E}[\ell_0(t - x)]$$

- Si $0 \geq t - x$, (4) nos da que

$$\mathbb{E}[\mathcal{L}(t, x)] {}_t p_{x-t+1} \mathbb{E}[\mathcal{L}(0, x - t)],$$

que es una expresión recursiva.

La relación (13) representa las condiciones que impone la proposición 7.4.5 del libro de Hernández-Lerma y Lasserre (1999) para asegurar que

$$Pw(\vec{\mathcal{L}}_t) \leq p_0 w(\vec{\mathcal{L}}_t) + \left(q_0 + \sup_{\mathcal{L}(t, x) \leq \max_t \lambda(t)} \mathbb{E}[\mathcal{L}(t, x)] \right) \times \mathbb{1}_{\{\mathcal{L}(t, x) : \mathcal{L}(t, x) \leq \max_t \lambda(t)\}}(\vec{\mathcal{L}}_t).$$

Dado que (1) es un sistema aleatorio iterado, como el mostrado en el ejemplo 7.4.7 del libro de Hernández-Lerma y Lasserre (1999), el enunciado de nuestro interés es cierto (vea también la proposición 12.8.1 del libro de Lasota y Mackey (1994)). $\square$

El resultado siguiente corresponde al teorema 11.2.1 del libro de Hernández-Lerma y Lasserre (1999).

Teorema 2.9. Suponga que el proceso estocástico $(\vec{\mathcal{L}}_t : t = 0, 1, \dots)$ es w -geométricamente ergódico para alguna función w , y que μ es la medida a la que se refiere (12). Si $g : \mathbb{R}^{\omega+1} \rightarrow \mathbb{R}$ es una función tal que $\mu(g) := \mathbb{E}[g(\vec{\mathcal{L}}_t) | \mathcal{L}_0] < \infty$, entonces

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} \left[\sum_{t=0}^{n-1} g(\vec{\mathcal{L}}_t) | \mathcal{L}_0 \right] = \mu(g) \quad (14)$$

para μ -casi todo punto $\vec{\mathcal{L}}_0 \in \mathbb{R}^{\omega+1}$.

El teorema 2.9 es una ley fuerte de grandes números para procesos Markovianos, y como es el caso de la ley débil clásica en el contexto de la teoría moderna de probabilidades, da un sustento teórico al método de simulación de Monte Carlo que utilizaremos en la sección 3, y que de otro modo, no es sino un método heurístico. (De hecho, éste es el caso en la mayoría de las aplicaciones en la literatura existente sobre el tema.) Pueden ver más detalles sobre la técnica para algunas cadenas de Markov en la sección 5.5 del libro de Norris (1997).

Ejemplo 2.10. Considere que la función g es tal que nos da la proporción de personas por rango de edad en la población en el año $2020 + t$ para $t = 0, 1, \dots$. Las hipótesis demográficas mencionadas en la sección 2 –en particular, de los algoritmos⁴ 1 y 2 para simular el comportamiento de las cohortes y los nacimientos en cada año (con las tasas brutas de natalidad referidas en la tabla 4)- y la ley fuerte de los grandes números del teorema 2.9 nos da que, para el año 2050 la distribución de la población en México oscilará alrededor de los valores de la tabla 5.

Para efecto de comparación, la tabla 6 muestra la distribución exacta de la población Mexicana por segmento de edad en el año 2020.

Es interesante destacar lo siguiente.

- Bajo las hipótesis demográficas de nuestro modelo y el uso del teorema 2.9, la proporción de la población de hombres de menos de 15 años irá de un 26.2236 %, a aproximadamente un 20.6203 %; mientras que la de mujeres pasará de un 24.3317 %, a cerca de 20.0107 %.
- Bajo las hipótesis demográficas de nuestro modelo y el uso del teorema 2.9, la proporción de la población de hombres entre 15 y 64 años de edad irá de un 66.0388 %, a aproximadamente un 66.5062 %; mientras que la de mujeres pasará de un 67.0105 %, a cerca de 65.4087 %. Estos datos son relevantes porque, salvo por la constante κ referida en (10), se trata de la población susceptible de ser económicamente activa.

⁴ Vale la pena mencionar que, en la implementación computacional que ponemos a disposición del lector/usuario interesado, hemos empleado más bien el teorema central del límite. Efectivamente, para evitar sobrecargar de trabajo el procesador de una computadora convencional con cuatro núcleos, observamos que la variable aleatoria Poisson con media λ que modela los nacimientos puede escribirse como una suma de m variables aleatorias de Poisson independientes con parámetro común de intensidad λ/m , de manera que la aplicación del teorema central del límite es inmediata si m es lo suficientemente grande. El uso del teorema central del límite para simular las variables aleatorias Binomiales que representan la evolución de las cohortes está justificado a través de su versión más elemental: el teorema de DeMoivre-Laplace. De hecho, produjimos la figura 2 de esta manera. El artículo de López-Barrientos et al. (2022a) presenta una introducción lúdica –aunque un poco controversial- de estos resultados.



Rango de edades	Proporción de hombres	Proporción de mujeres
0-4	6.6698 %	6.5593 %
5-9	6.8300 %	6.6360 %
10-14	7.1205 %	6.8154 %
15-19	7.3546 %	6.9574 %
20-24	7.4775 %	7.0558 %
25-29	7.4972 %	7.1020 %
30-34	7.4132 %	7.0735 %
35-39	7.2272 %	7.0103 %
40-44	6.8933 %	6.8121 %
45-49	6.4896 %	6.5326 %
50-54	5.9935 %	6.1632 %
55-59	5.4006 %	5.6489 %
60-64	4.7595 %	5.0529 %
65-69	4.0410 %	4.3480 %
70-74	3.2427 %	3.5587 %
75-79	2.4158 %	2.7187 %
80-84	1.6131 %	1.8989 %
85-89	0.9338 %	1.1651 %
90-94	0.4405 %	0.5993 %
95-100	0.1531 %	0.2323 %
Mayores de 100	0.0338 %	0.0595 %

■ **Tabla 5** Distribución aproximada de la población Mexicana por segmento de edad en el año 2050.

Rango de edades	Proporción de hombres	Proporción de mujeres
0-4	8.2780 %	7.7168 %
5-9	8.8903 %	8.2469 %
10-14	9.0553 %	8.3680 %
15-19	8.9051 %	8.2985 %
20-24	8.4221 %	8.1614 %
25-29	7.9257 %	7.9679 %
30-34	7.3817 %	7.5976 %
35-39	7.0618 %	7.2803 %
40-44	6.6229 %	6.8960 %
45-49	6.2154 %	6.4128 %
50-54	5.4325 %	5.7534 %
55-59	4.3904 %	4.6628 %
60-64	3.6811 %	3.9799 %
65-69	2.7827 %	3.0095 %
70-74	2.0110 %	2.1953 %
75-79	1.3824 %	1.5010 %
80-84	0.8540 %	1.0117 %
85-89	0.4620 %	0.5837 %
90-94	0.1750 %	0.2476 %
95-99	0.0597 %	0.0910 %
Mayores de 100	0.0108 %	0.0181 %

■ **Tabla 6** Distribución exacta de la población Mexicana por segmento de edad en el año 2020.



- Bajo las hipótesis demográficas de nuestro modelo y el uso del teorema 2.9, la proporción de la población de hombres mayores de 64 años de edad irá de un 7.7376 %, a aproximadamente un 12.8785 %; mientras que la de mujeres pasará de un 8.6578 %, a cerca de 14.5806 %.
- La razón por la que enfatizamos que los resultados de la tabla 5 son sólo aproximaciones es que, si consideramos que la semilla del proceso de simulación es el año 2020, y que avanzamos de manera quinquenal, contamos con solamente siete iteraciones hasta ese año. De manera que la aproximación a la que se refiere (14) tiene $n = 7$. Para lograr un resultado más robusto, es imperativo simular en un horizonte más largo. Sin embargo, vale la pena destacar que los resultados mostrados en la tabla 5 no presentan gran variabilidad cuando se los compara con otras simulaciones bajo los supuestos que produjeron los aquí tabulados. Esto último puede servir como una comprobación intuitiva de la veracidad del teorema 2.9.
- Para simular la dinámica de la población después del año 2050, es posible considerar el trazador cúbico interpolante (9). Pero debe tenerse en cuenta que, para $t > 80.48$, la cualidad monótona de la estimación nos dará valores negativos, haciendo así imposible continuar las simulaciones. Como la proporción de la población que será económicamente activa se mantiene inmanente en $\kappa = 58.44$ %, las figuras 4-7 nos muestran diferentes momentos de la evolución de la población económicamente activa y de la población envejecida. Para hacer la figura 6, supusimos que la tasa bruta de natalidad $\lambda(t)$ se mantiene estable en su nivel mínimo a partir de 2050; mientras que para la figura 7, supusimos que la tasa bruta de natalidad descende prácticamente al nivel del país que menor tasa bruta de natalidad presentaba en 2021: Mónaco, con un índice de 6.63 nacidos por cada mil habitantes (vea el reporte de *World Population Review* (2022)).



Figura 4 Distribuciones de las poblaciones económicamente activa y envejecida en el año 2020.

Hipótesis financieras

Para lograr las metas establecidas en la observación 2.6, utilizamos la siguiente variación de la definición del

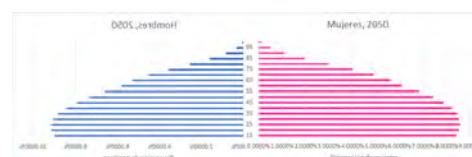


Figura 5 Distribuciones estables de las poblaciones económicamente activa y envejecida en el año 2050 suponiendo (9).

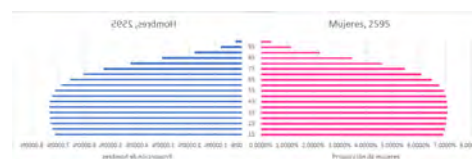


Figura 6 Distribuciones estables de las poblaciones económicamente activa y envejecida suponiendo (9) para $t < 35$, y $\lambda(t) = 11.0807$ para $t \geq 35$.

proceso de riqueza presentada en el artículo de [Boikov \(2002\)](#):

$$u(t) = u(0)(1+i)^t + \sum_{k=1}^t (1+i)^k (a(t-k) + b(t-k)), \quad (15)$$

con $u(0) =: u > 0$, para $t = 0, 1, \dots$; donde i es la tasa de interés a la que pueden invertirse los fondos, $a(t)$ es la aportación en el nivel superior del portafolios por parte de los contribuyentes y sus aliados⁵, y $b(t)$ representa los beneficios pagaderos a la población de jubilados en el tiempo t .

Una semejanza entre (15) y el modelo clásico de Cramér-Lundberg es que la magnitud del pago de los beneficios en el nivel superior del portafolios es contingente. Sin embargo, usamos los resultados presentados en el artículo de [Aguirre-Farías et al. \(2019\)](#) para suponer que el beneficio individual (o sea en el nivel inferior del portafolios) en cada tiempo es conocido. O sea que el beneficio para una persona de edad $x = r, \dots, \omega$ al tiempo

⁵ En México, el ideal es contar con los siguientes cuatro pilares para constituir una pensión de tipo contributivo:

- Una pensión obligatoria del sistema gubernamental.
- Planes de retiro corporativos.
- Planes individuales de retiro.
- Rentas, regalías, dividendos y otros ingresos pasivos.

Vea el artículo de [Cervantes González y Sedano Díaz \(2023\)](#).



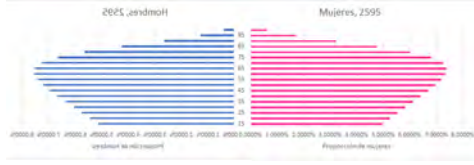


Figura 7 Distribuciones estables de las poblaciones económicamente activa y envejecida suponiendo (9) para $t < 65$, y $\lambda(t) = 6.6691$ para $t \geq 65$.

$t = 0, 1, \dots$ es

$$\begin{aligned} b(t, x) &:= f \cdot s(r) \cdot e^{\theta(t-(x-r))} \cdot e^{\pi(x-r)} \\ &= f \cdot s(r) \cdot e^{\theta t} \cdot e^{(\pi-\theta)(x-r)} =: e^{\theta t} F(x), \end{aligned} \quad (16)$$

donde $s(r)$ es el último salario recibido por un empleado en la edad de retiro r , $f \in [0, 1]$ es el nivel de reemplazo para quien se jubila a edad r , e^{θ} es el nivel de ajuste anual del salario, y e^{π} es el nivel de ajuste anual para la pensión.

Supongamos que un miembro del plan en edad x y sus aliados aportan un múltiplo $x \in [0, p]$ de su salario $s(x)$ al tiempo $t = 0, 1, \dots$ en el modelo de Belolipetskii y Lepskaya (2018, 2019), la regla es $p = 1$. Esto es

$$a(t, x) := z \cdot s(x) \cdot e^{\theta t} =: e^{\theta t} G(x). \quad (17)$$

Considere el vector de flujos de efectivo

$$\begin{aligned} V(\vec{a}) &:= (G(\alpha), \dots, G(r-1), -F(r), \dots, -F(\omega))^\top \\ &= \begin{pmatrix} z \cdot s(\alpha) \\ \vdots \\ z \cdot s(r-1) \\ -f \cdot s(r) \\ \vdots \\ -f \cdot s(r) \cdot e^{(\pi-\theta)(\omega-r)} \end{pmatrix}, \end{aligned} \quad (18)$$

defina las matrices cuadradas

$$E^+ := \text{diag}(\underbrace{1, 1, \dots, 1}_{r-\alpha}, \underbrace{0, 0, \dots, 0}_{\omega-r+1}), \quad (19)$$

$$\begin{aligned} &\text{entradas} \quad \text{entradas} \\ E^- &:= \text{diag}(\underbrace{0, 0, \dots, 0}_{r-\alpha}, \underbrace{1, 1, \dots, 1}_{\omega-r+1}), \quad (20) \\ &\text{entradas} \quad \text{entradas} \end{aligned}$$



y note que $E^+ + E^- = I \in \mathbb{R}^{(\omega-\alpha+1) \times (\omega-\alpha+1)}$. Por (16)-(17) y (19)-(20), las variables aleatorias escalares

$$a(t) := e^{\theta t} V(\vec{a}) E^+ \vec{\mathcal{M}}_t, \quad (21)$$

$$b(t) := e^{\theta t} V(\vec{a}) E^- \vec{\mathcal{M}}_t \quad (22)$$

corresponden a las aportaciones hechas por los participantes, y a los beneficios que el fondo otorga, respectivamente. La inserción de (21)-(22) en (15), nos da

$$\begin{aligned} u(t, \vec{\mathcal{M}}, \vec{a}) &:= u(0)(1+i)^t \\ &+ \sum_{k=1}^{t-1} e^{\theta(t-k)} (1+i)^k V(\vec{a}) \cdot \vec{\mathcal{M}}_{t-k}. \end{aligned} \quad (23)$$

Observación 2.11. (a) La diferencia crucial entre (23) y el modelo clásico de Cramér-Lundberg radica en que, como $\vec{\mathcal{M}}_t$ es un proceso estocástico, entonces el nivel de recaudación de primas (21) resulta incierto.

(b) Si relajamos los incisos (i) y (v) en la hipótesis 2.5, o sea, si permitimos que las personas entren y salgan del plan de pensiones a una edad que no sea α ; o si no usamos la hipótesis 2.5 (vi) para que el grupo poblacional sea heterogéneo en lo que se refiere a su nivel de ingreso, entonces el método que describimos aquí seguirá siendo factible, dada la naturaleza lineal de (15): bastará con sumar a (15) tantos sumandos como variantes homogéneas haya del sistema.

(c) Note que, a pesar de que el proceso (3), que describe cuántos participantes hay en cada edad es una cadena de Markov, el proceso de riqueza (23) no lo es. En efecto, aunque es posible escribir (15) de la forma que requiere la proposición 2.2:

$$\begin{aligned} u(t+1, \vec{\mathcal{M}}, \vec{a}) &= u(t, \vec{\mathcal{M}}, \vec{a}) (1+i) \\ &+ a(t+1) + b(t+1) \end{aligned}$$

para $t = 0, 1, \dots$, las variables aleatorias $a(t) + b(t)$ para $t = 0, 1, \dots$ no son mutuamente independientes, pues sus valores dependen de toda la trayectoria del proceso estocástico subyacente (3). Sin embargo, esto no limita nuestra acción.

(d) El término $b(t)$ definido en (22) incluye las pensiones de quienes ya estaban retirados al tiempo $t = 0$ y las pensiones de quienes se retiran en ese momento. Para financiar los pagos que se hacen a quienes estaban retirados cuando el fondo se establece, consideramos el capital inicial $u(0) = u > 0$.

- (e) Una diferencia fundamental entre nuestro modelo, y el presentado en el artículo clásico de Bühlmann (1985) y en el capítulo 5.2 del libro de Kaas et al. (2008), es que no existe un dividendo que considerar en (23).
- (f) De acuerdo a la observación 2.6, nuestra primera meta es medir el desempeño del fondo (23) en un horizonte establecido con hipótesis financieras realistas, mientras que la segunda es conocer una política $(a(t) : t = 0, 1, \dots)$ de aportaciones tal que podamos controlar la probabilidad de insolvencia. Lograremos lo primero considerando que:

- según la Secretaría de Hacienda y Crédito Público y Banco de México (2023), la tasa de retorno de los certificados de la tesorería de la Federación a 364 días al día 19 de septiembre de 2023 fue $i = 11.66\%$,
- de acuerdo al informe publicado por la Secretaría del Trabajo y Previsión Social y Banco de México (2023) (vea la figura 8), podemos tomar $\theta = 8.21527\%$, pues ésta es la tasa promedio de ajuste salarial contractual durante el período comprendido entre julio de 2022 y agosto de 2023,



Figura 8 Tasas de ajustes salariales de julio de 2022 a agosto de 2023. Fuente: Secretaría del Trabajo y Previsión Social y Banco de México (2023).

- como el porcentaje de incremento de las pensiones π se calcula con base en el Índice Nacional de Precios al Consumidor (vea el reporte de Comunicación Social (2022) y la figura 9), usamos la tasa de $\pi = 4.65\%$, que es la última disponible a la fecha de la redacción de este artículo.
- Usamos la Encuesta Nacional de Ocupación y Empleo conducida por el Instituto Nacional de Estadística y Geografía (2023b), y los datos de Secretaría de Economía (2023) para establecer la carrera salarial que consideraremos. La figura 10 y la tabla 7 muestran estos datos.



Figura 9 Índice Nacional de Precios al Consumidor. Fuente: Instituto Nacional de Geografía y Estadística (2023).

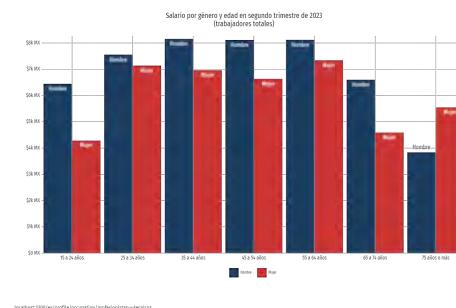


Figura 10 Salarios promedio mensuales clasificados por género y por grupo de edad. Fuentes: Instituto Nacional de Estadística y Geografía (2023b) y Secretaría de Economía (2023).



Edad	Hombres	Mujeres
15-24	\$6,437.23	\$4,274.70
25-29	\$6,437.23	\$4,274.70
30-34	\$7,547.07	\$7,131.04
35-39	\$7,547.07	\$7,131.04
40-44	\$8,146.36	\$6,955.90
45-49	\$8,146.36	\$6,955.90
50-54	\$8,105.63	\$6,624.72
55-59	\$8,105.63	\$6,624.72
60-64	\$8,110.85	\$7,333.40

■ **Tabla 7** Salarios promedio mensuales clasificados por género y por grupo de edad. Fuentes: [Instituto Nacional de Estadística y Geografía \(2023b\)](#) y [Secretaría de Economía \(2023\)](#).

- Citamos la página 140 del reporte de la *Organisation for Economic Co-operation and Development (2021)* para afirmar que el nivel de reemplazo (bruto) f^h para los hombres después de toda la carrera profesional será del 61.2 %, mientras que el de las mujeres f^m será del 58.2 %. (Es importante mencionar que estos datos ya incluyen las consideraciones necesarias tras la pandemia por COVID-19.)

Insertaremos estos valores en (16)-(23) y encontraremos los mejores valores de los demás parámetros en (23) desde el punto de vista del criterio de optimización del pago promedio a largo plazo (vea el artículo de [López-Barrientos et al. \(2023\)](#)).

3. DESEMPEÑO DEL FONDO DE PENSIONES

En esta sección, medimos el desempeño del sistema pensionario para una población dada bajo las hipótesis presentadas en las subsecciones 2-2. Note que esto implica utilizar las tasas *brutas* de natalidad de la tabla 4 y tomar alguna decisión sobre el uso del trazador cúbico interpolante (9) para los años subsecuentes al 2050.

La carrera salarial y el pago de beneficios

Ahora construiremos el vector de flujos de efectivo referido en (18) tomando los datos de la observación 2.11(f). Para este efecto, hacemos $\alpha = 15$ en la hipótesis 2.5, usamos la edad de jubilación vigente en México para el año



2023, o sea $r = 65$; y la edad límite de las tablas de mortalidad que aparecen en las tablas 2 y 3, o sea $\omega = 100$. Así, interpretamos el primer bloque del vector de flujos como una función de la proporción $z > 0$ del sueldo a reservar y construimos la tabla 8. Observe que una consecuencia de que hayamos elegido $\theta = 8.21527\% > 4.65\% = \pi$ es que las pensiones de los adultos mayores disminuyen a medida que la edad aumenta.

Edad	Flujo de efectivo hombres	Flujo de efectivo mujeres
15	\$77,246.74z	\$51,296.43z
20	\$77,246.74z	\$51,296.43z
25	\$90,564.80z	\$85,572.48z
30	\$90,564.80z	\$85,572.48z
35	\$97,756.33z	\$83,470.85z
40	\$97,756.33z	\$83,470.85z
45	\$97,267.54z	\$79,496.67z
50	\$97,267.54z	\$79,496.67z
55	\$97,330.24z	\$88,000.82z
60	\$97,330.24z	\$88,000.82z
65	-\$59,566.11	-\$51,216.48
70	-\$49,840.27	-\$42,853.95
75	-\$41,702.44	-\$35,856.84
80	-\$34,893.35	-\$30,002.20
85	-\$29,196.03	-\$25,103.50
90	-\$24,428.96	-\$21,004.65
95	-\$20,440.24	-\$17,575.05
100	-\$17,102.80	-\$14,705.43

■ **Tabla 8** Las entradas de la segunda y tercera columnas en el intervalo de edades de 15 a 60 años corresponden a la proporción del salario de hombres y mujeres, respectivamente que es necesario se aporte por cada persona para la constitución del fondo de pensiones. Las entradas de la segunda y tercera columnas en el intervalo de edades de 65 a 100 años corresponden a los beneficios pagaderos a los pensionarios hombres y mujeres, respectivamente.

Trayectorias del proceso de riqueza versus un modelo determinista de control

La modelación del pago de pensiones es un problema central para las ciencias actuariales. También es un problema abierto porque, pese a que admite elegantes soluciones parciales que pueden implementarse con una cantidad razonable de recursos computacionales (como los métodos clásicos de la prima media general, la prima escalonada, el reparto puro –expuestos en los libros de Aguirre-Farías (2001) y Anderson (1992), el capítulo 16 de la obra de Daykin *et al.* (1993), la tesis de Plata Gallegos y López Barrientos (2023) y el libro de Iyer (1999)–, o el método de las primas óptimas de liquidez presentado por Aguirre-Farías (2001, 2012)), nuestra propuesta complementa aquellas en términos de la solvencia del fondo de pensiones, y de su fundamento matemático.

Considere el proceso estocástico $(u(t, \vec{\mathcal{M}}, \vec{a}) : t = 0, 1, \dots)$ de (23) y defina el índice de rendimiento

$$J_T(\vec{\mathcal{M}}, \vec{a}) := u(T, \vec{\mathcal{M}}, \vec{a}). \quad (24)$$

La figura 11 presenta sendas trayectorias del proceso (24) en un horizonte de cien años para hombres, mujeres, y la suma de ambos, cuando el capital inicial $u(0)$ es nulo, los niveles de aportación sobre el salario de hombres y mujeres son del 12%; y la tasa bruta de natalidad $\lambda(t)$ obedece (9) para $t < 65$, y $\lambda(t) = 6.6691$ para $t \geq 65$. Ahí, el fondo común se vuelve insuficiente a partir del año 2050.

La figura 12 presenta sendas trayectorias del proceso (24) en un horizonte de cien años para hombres, mujeres, y la suma de ambos, cuando los capitales iniciales de los fondos de cada género igualan el monto de lo que debe pagárseles en el año 2020, esto es $u^h(0) = \$228,542,369,398.64$ y $u^m(0) = \$265,233,730,777.115$, los niveles de aportación sobre el salario de hombres y mujeres son del 12%; y la tasa bruta de natalidad $\lambda(t)$ obedece (9) para $t < 65$, y $\lambda(t) = 6.6691$ para $t \geq 65$. Ahí, el fondo común se vuelve insuficiente a partir del año 2085.

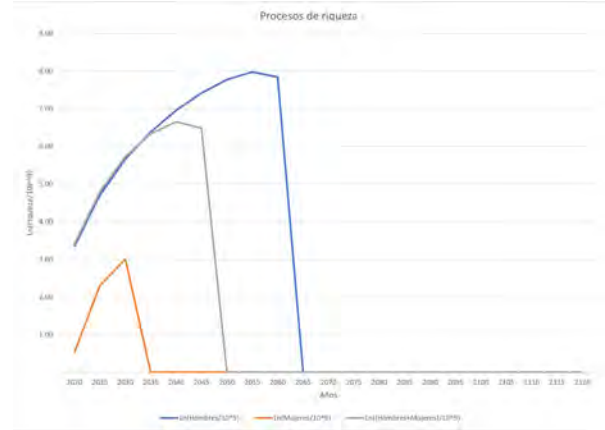


Figura 11 Trayectorias de (24) en un horizonte de cien años para hombres, mujeres, y su suma, con $u(0) = 0$, $z = 12\%$ y $\lambda(t)$ según (9) para $t < 65$, y $\lambda(t) = 6.991$ para $t \geq 65$.

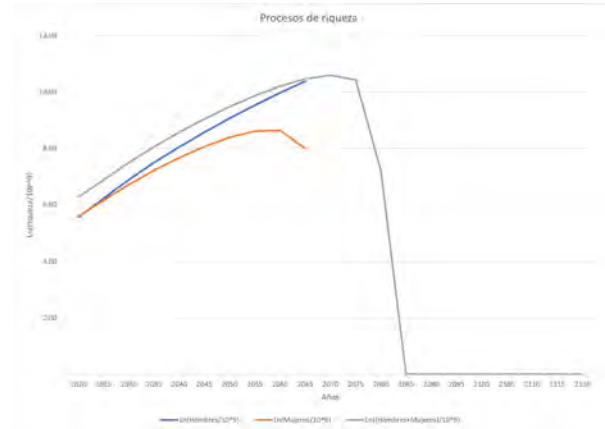


Figura 12 Trayectorias de (24) en un horizonte de cien años para hombres, mujeres, y su suma, con $u(0) > 0$, $z = 12\%$ y $\lambda(t)$ según (9) para $t < 65$, y $\lambda(t) = 6.991$ para $t \geq 65$.

La figura 13 presenta sendas trayectorias del proceso (24) en un horizonte de cien años para hombres, mujeres, y la suma de ambos, cuando el capital inicial $u(0)$ es nulo, los niveles de aportación sobre el salario de hombres y mujeres son del 12%; y la tasa bruta de natalidad $\lambda(t)$ obedece (9) para $t < 30$, y $\lambda(t) = 11.2807$ para $t \geq 30$. Ahí, el fondo común se vuelve insuficiente a partir del



año 2050.

La figura 14 presenta sendas trayectorias del proceso (24) en un horizonte de cien años para hombres, mujeres, y la suma de ambos, cuando los capitales iniciales de los fondos de cada género igualan el monto de lo que debe pagárseles en el año 2020, esto es $u^h(0) = \$228,542,369,398.64$ y $u^m(0) = \$265,233,730,777.115$, los niveles de aportación sobre el salario de hombres y mujeres son del 12 %; y la tasa bruta de natalidad $\lambda(t)$ obedece (9) para $t < 30$, y $\lambda(t) = 11.2807$ para $t \geq 30$. Ahí, el fondo común se vuelve insuficiente a partir del año 2085.

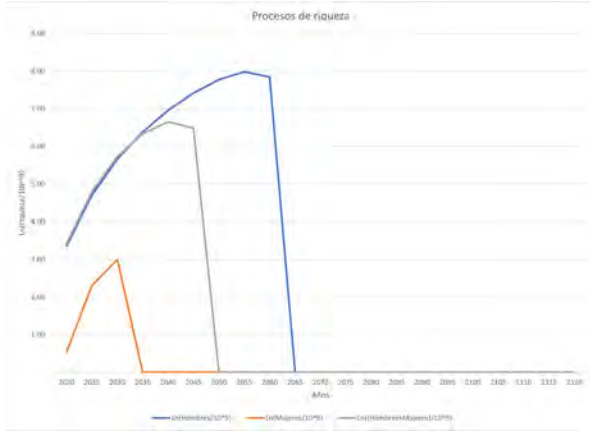


Figura 13 Trayectorias de (24) en un horizonte de cien años para hombres, mujeres, y su suma, con $u(0) = 0$, $z = 12\%$ y $\lambda(t)$ según (9) para $t < 30$, y $\lambda(t) = 11.2807$ para $t \geq 65$.

Considere ahora la dinámica (14) del teorema 2.9 y el proceso **determinista** $(u(t, \vec{m}, \vec{a}) : t = 0, 1, \dots)$, donde $\vec{m}$ es la sucesión de medidas de probabilidad aludidas en el ejemplo 2.10; y defina el índice de rendimiento

$$j_T(\vec{m}, \vec{a}) := u(T, \vec{m}, \vec{a}). \quad (25)$$

La figura 15 presenta la trayectoria del proceso (25) en un horizonte de cien años para hombres, mujeres, y la suma de ambos, cuando el capital inicial $u(0)$ es nulo, los niveles de aportación sobre el salario de hombres y mujeres son del 12 %; y la tasa bruta de natalidad $\lambda(t)$ obedece (9) para $t < 65$, y $\lambda(t) = 6.6691$ para $t \geq 65$. Ahí, el fondo común se vuelve insuficiente a partir del año 2085.

La figura 16 presenta la trayectoria del proceso (25) en un horizonte de cien años para hombres, mujeres, y

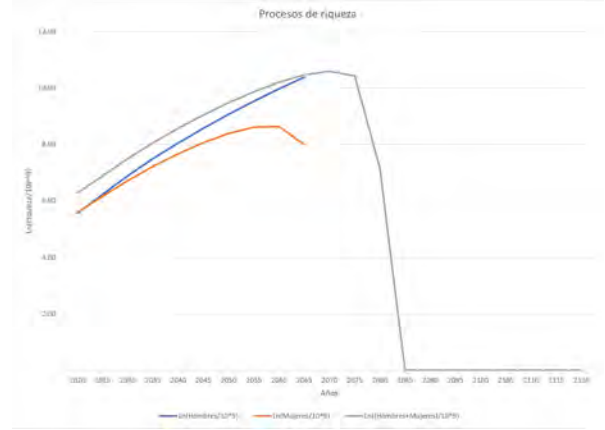


Figura 14 Trayectorias de (24) en un horizonte de cien años para hombres, mujeres, y su suma, con $u(0) > 0$, $z = 12\%$ y $\lambda(t)$ según (9) para $t < 30$, y $\lambda(t) = 11.2807$ para $t \geq 30$.

la suma de ambos, cuando los capitales iniciales de los fondos de cada género igualan el monto de lo que debe pagárseles en el año 2020, esto es $u^h(0) = \$8,042.19$ y $u^m(0) = \$8,612.35$, los niveles de aportación sobre el salario de hombres y mujeres son del 12 %; y la tasa bruta de natalidad $\lambda(t)$ obedece (9) para $t < 65$, y $\lambda(t) = 6.6691$ para $t \geq 65$. Ahí, el fondo común no se vuelve insuficiente durante los cien años que dura el horizonte.

La figura 17 presenta la trayectoria del proceso (25) en un horizonte de cien años para hombres, mujeres, y la suma de ambos, cuando el capital inicial $u(0)$ es nulo, los niveles de aportación sobre el salario de hombres y mujeres son del 12 %; y la tasa bruta de natalidad $\lambda(t)$ obedece (9) para $t < 30$, y $\lambda(t) = 11.2807$ para $t \geq 30$. Ahí, el fondo común se vuelve insuficiente a partir del año 2085.

La figura 18 presenta la trayectoria del proceso (25) en un horizonte de cien años para hombres, mujeres, y la suma de ambos, cuando los capitales iniciales de los fondos de cada género igualan el monto de lo que debe pagárseles en el año 2020, esto es $u^h(0) = \$8,042.19$ y $u^m(0) = \$8,612.35$, los niveles de aportación sobre el salario de hombres y mujeres son del 12 %; y la tasa bruta de natalidad $\lambda(t)$ obedece (9) para $t < 30$, y $\lambda(t) = 11.2807$ para $t \geq 30$. Ahí, el fondo común no se vuelve insuficiente durante los cien años que dura el horizonte.

Observación 3.1. Los resultados mostrados en esta subsección nos permiten obtener nuestras primeras conclusiones.



Figura 15 Trayectoria de (24) en un horizonte de cien años para hombres, mujeres, y su suma, con $u(0) = 0$, $z = 12\%$ y $\lambda(t)$ según (9) para $t < 65$, y $\lambda(t) = 6.991$ para $t \geq 65$.

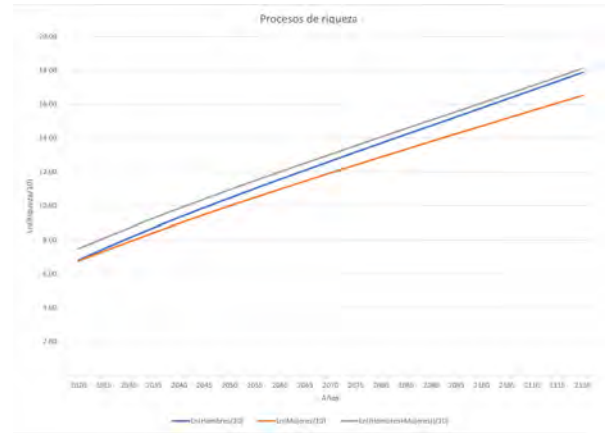


Figura 16 Trayectoria de (25) en un horizonte de cien años para hombres, mujeres, y su suma, con $u(0) > 0$, $z = 12\%$ y $\lambda(t)$ según (9) para $t < 65$, y $\lambda(t) = 6.991$ para $t \geq 65$.

- (a) Las figuras 11-18 son reveladoras. Ciertamente, la solvencia de los fondos (24) y (25) es mucho más sensible a las contribuciones de los participantes y al capital inicial, que a la estructura de la tasa bruta de natalidad.
- (b) Mientras que las gráficas de las figuras 11-14 corresponden a una de las trayectorias del proceso estocástico (24) para los hombres y para las mujeres, las gráficas de las figuras 15-18 son la representación geométrica de la trayectoria de la sucesión determinista (25). Esto último, en virtud de la ergodicidad w -geométrica del proceso $\mathcal{L}$ (definido en (1)) demostrada en la proposición 2.8.
- (c) El hecho de que hombres y mujeres tengan carreras salariales distintas entre sí no es un obstáculo para la implementación de nuestro modelo. Como argumentamos en la observación 2.11(b), basta con sumar los resultados que obtenemos para cada grupo.
- (d) A pesar de que las figuras 11-14 son sólo algunas trayectorias de los procesos que representan, todas ellas confirman la consecuencia de tener un sistema de salarios que perjudica a la mujer: los bajos salarios a la mujer son un lastre para la competitividad del fondo de pensiones.

4. UN ENFOQUE INGENUO AL CÁLCULO DE LA PROBABILIDAD DE LA RUINA

Somos actuarios. Independientemente de dónde hayamos estudiado, desde que empezamos la escuela, co-

menzaron a enseñarnos los fundamentos de nuestra disciplina. Uno de ellos es la recomendación de Bühlmann (1985) para obtener primas de riesgo al nivel superior de un portafolios dado, para acto seguido obtener las primas en el nivel inferior. Nuestra aproximación más cándida siempre implica tratar de utilizar el modelo de riesgos individuales a través del teorema central del límite, como en el ejemplo 2.5.1 del libro clásico de Bowers *et al.* (1997). En esta sección aplicaremos ese enfoque al problema que estamos estudiando.

Un teorema central del límite

Sea $S := \{\alpha, \dots, r, \dots, \omega\}$ y defina el conjunto de medidas de probabilidad con soporte en S como $\mathbb{P}(S)$. Recuerde (25) y llame a ése índice de rendimiento *reserva tras T años cuando usamos la política $\vec{a}$* , dado que partimos del vector inicial de proporciones $\vec{m}_0 = \frac{\vec{M}_0}{\sum_{x \in S} \vec{M}_0(0, x)}$. El teorema 4.7 del artículo de López-Barrientos *et al.* (2023) da condiciones suficientes para aseverar que

$$\frac{1}{T} \left| \mathbb{E} \left[J_T \left(\vec{\mathcal{M}}, \vec{a}_* \right) \mid \vec{\mathcal{M}}_0 \right] - j_T \left(\vec{m}, \vec{a}'_* \right) \right| \rightarrow 0 \quad (26)$$

cuando $T \rightarrow \infty$, donde $\vec{a}_*$ y $\vec{a}'_*$ son políticas que maximizan (24) y (25), respectivamente. En virtud de la proposición 2.8 y del teorema 2.9, podemos definir la **varianza límite promedio** del proceso estocástico (24) como

$$\sigma^2 \left(\vec{\mathcal{M}}, \vec{a} \right) := \liminf_{T \rightarrow \infty} \frac{1}{T} \text{var} \left[J_T \left(\vec{\mathcal{M}}, \vec{a}_* \right) \mid \vec{\mathcal{M}}_0 \right].$$



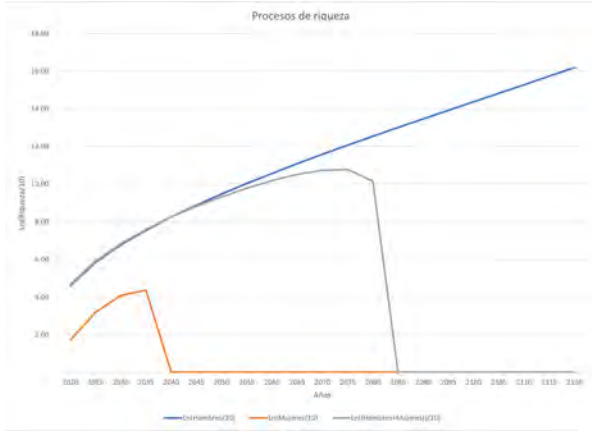


Figura 17 Trayectoria de (24) en un horizonte de cien años para hombres, mujeres, y su suma, con $u(0) = 0$, $z = 12\%$ y $\lambda(t)$ según (9) para $t < 30$, y $\lambda(t) = 11.2807$ para $t \geq 30$.

Una nueva aplicación del teorema 4.7 del trabajo de López-Barrientos *et al.* (2023), combinada con los resultados de la sección 11.B del libro de Hernández-Lerma y Lasserre (1999) nos da que la expresión

$$\frac{1}{T} \left| \text{var} \left[J_T \left(\vec{\mathcal{M}}, \vec{a}_* \right) \middle| \vec{\mathcal{M}}_0 \right] - \left(u(0)(1+i)^T \right)^2 - \left[\sum_{k=1}^{T-1} \left(e^{\theta(T-k)} (1+i)^k V \left(\vec{a}'_* \right) \vec{m}_{T-k} \right)^2 - \left(j_T \left(\vec{m}, \vec{a}'_* \right) \right)^2 \right] \right|$$

tiende a cero cuando $T \rightarrow \infty$, donde $\vec{a}_*$ y $\vec{a}'_*$ son como en (26).

Adaptamos a nuestro contexto la versión del teorema central del límite para procesos Markovianos de decisión de la p. 175 de la obra de Hernández-Lerma y Lasserre (1999) (vea también Meyn y Tweedie (1993); Doob (1994); Duflo (1990)):

Teorema 4.1. Suponga que la cadena de Markov $(\vec{\mathcal{M}}_t : t = 0, 1, \dots)$ es w -geométricamente ergódica y que $\sigma^2(\vec{\mathcal{M}}, \vec{a}_*)$ es la varianza límite promedio correspondiente. Entonces, para cada estado inicial $\vec{\mathcal{M}}_0$ de $\vec{\mathcal{M}}$, tenemos que

$$\mathbb{P} \left(\frac{J_T \left(\vec{\mathcal{M}}, \vec{a}_* \right) - T j_T \left(\vec{m}, \vec{a}'_* \right)}{\sqrt{T \sigma^2 \left(\vec{\mathcal{M}}, \vec{a}_* \right)}} \leq z \middle| \vec{\mathcal{M}}_0 \right) \rightarrow \Phi(z) \quad (27)$$

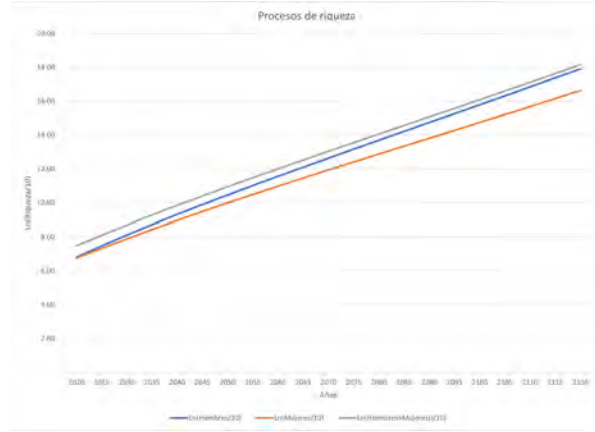


Figura 18 Trayectoria de (25) en un horizonte de cien años para hombres, mujeres, y su suma, con $u(0) > 0$, $z = 12\%$ y $\lambda(t)$ según (9) para $t < 30$, y $\lambda(t) = 11.2807$ para $t \geq 30$.

cuando $T \rightarrow \infty$, donde $\Phi(\cdot)$ representa la función de distribución acumulada de una variable aleatoria Normal estándar.

Es fácil ver que el símbolo $\Phi(0)$ representa la probabilidad de insolvencia del fondo (24) en el horizonte $[0, T]$.

Ejemplo 4.2. Considere el caso referido en la figura 11. El uso del teorema 4.1 para estimar la probabilidad de insolvencia del fondo (24) en un horizonte de cien años para hombres, mujeres, y su suma, con $u(0) = 0$, $z = 12\%$ y $\lambda(t)$ según (9) para $t < 65$, y $\lambda(t) = 6.991$ para $t \geq 65$ nos da $\Phi(0) = 98.1737\%$ para los hombres, $\Phi(0) = 98.4096\%$ para las mujeres, y $\Phi(0) = 98.3325\%$ para la combinación.

Constitución de una reserva actuarial de arriba p'abajo

Fijar la probabilidad referida en (27) a un nivel establecido con antelación es un modo ideal para encontrar las políticas que maximizan (24) y (25). Así, hacemos $\Phi(0) = 1\%$ e iteramos (27) cambiando el valor de $\vec{a}$ y $\vec{a}'$ hasta dar con un valor adecuado.

Suponiendo que $u(0) = V(\vec{a}) E^- \vec{\mathcal{M}}_0$ (esto es, que al principio de la proyección, los fondos para los hombres y las mujeres cuentan con el dinero que debe pagarse a la población jubilada durante ese año), obtenemos que durante la proyección, el monto a reservar por parte de la población masculina económicamente activa es del 13% de su salario; mientras que la población femenina económicamente activa debe aportar el 14% de su sueldo. Estas acciones redundan en sendas probabilidades de

Rango de edad	Hombres	Mujeres
15-24	$\$10,042.08e^{\theta t}$	$\$7,181.50e^{\theta t}$
25-34	$\$11,773.42e^{\theta t}$	$\$11,980.15e^{\theta t}$
35-44	$\$12,708.32e^{\theta t}$	$\$11,685.92e^{\theta t}$
45-54	$\$12,644.78e^{\theta t}$	$\$11,129.53e^{\theta t}$
55-64	$\$12,652.93e^{\theta t}$	$\$12,320.11e^{\theta t}$

■ **Tabla 9** Aportaciones que cada hombre y cada mujer deben hacer en el t -ésimo año del horizonte para garantizar una probabilidad de insolvencia del fondo de 0.2656 % para $t = 0, 1, \dots$

insolvencia de 0.5634 % para los hombres, y 0.0050 % para las mujeres, y una probabilidad combinada de ruina del 0.2656 %. Estas son las acciones que es preciso tomar en el nivel superior del portafolios.

La linealidad de nuestro modelo, juntamente con los resultados presentados en el artículo de [Aguirre-Farías et al. \(2019\)](#) nos permiten ver que las contribuciones que cada persona debe hacer en el nivel inferior del portafolios son como lo muestra la tabla 9.

5. ES QUE NO LE PUSIERON INTERÉS...

“Bueno, sí le pusieron. ¡Pero fue una sola tasa!”. Es verdad. Nuestro estudio considera que las tasas

- libre de riesgo i ,
- de ajuste salarial θ ,
- de reemplazo f ,
- de inflación anual π y
- de ocupación κ

son constantes. Esta es una limitante en nuestro estudio, porque de acuerdo a [Matthews \(2022\)](#), al moverse por ejemplo la tasa de descuento, también lo hará el nivel de los pasivos de las pensiones. Una extensión plausible de nuestro enfoque es la adopción de algún modelo estocástico de tasas de interés para analizar el comportamiento del fondo. Lo mismo puede decirse de las demás tasas que hemos mantenido constantes. Sin embargo, nuestra meta poco tiene que ver con la modelación de estas tasas. La contribución principal de este trabajo al Estado del Arte de la ciencia actuarial es que, usando una cantidad mínima de supuestos demográficos, financieros y estocásticos, hemos descrito un método sencillo para estimar la probabilidad de insolvencia de un fondo de pensiones

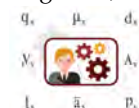
(vea la sección 4) y más, aún, calcular los montos que deben aportarse en los niveles superior e inferior para que esa probabilidad sea de, a lo sumo, un valor dado (vea la sección 4). Para lograrlo, hemos echado mano del modelo de Cramér-Lundberg con primas estocásticas presentado por [Boikov \(2002\)](#), la medición del rendimiento de un fondo de pensiones en la Federación Rusa estudiado por [Belolipetskii y Lepskaya \(2019, 2018\)](#), las recomendaciones clásicas de [Bühlmann \(1985\)](#) y algunos resultados de control estocástico a tiempo discreto (cf. los capítulos 7 y 11 del libro de [Hernández-Lerma y Lasserre \(1999\)](#), el libro de [Llanos Reynoso \(2014\)](#) y el artículo de [López-Barrientos et al. \(2023\)](#)).

Es importante destacar que, para simular la población económicamente activa del grupo de edad entre 15 a 64 años, utilizamos las mismas tablas de mortalidad de las tablas 2 y 3 que usamos para el resto de los grupos (vea la hipótesis 2.5 y la definición (10)). Dado que cada giro industrial tiene sus propias particularidades de riesgos, un análisis más fino requerirá contar con tablas adecuadas para cada giro. Por otro lado, los resultados que mostramos aquí son lo suficientemente generales como para aplicarlos en distintos contextos industriales. Para garantizar que el sistema demográfico es w -geométricamente ergódico (cf. el teorema 2.8), hemos usado tres supuestos sobre las tasas brutas de natalidad:

- el de [Consejo Nacional de Población et al. \(2019\)](#) en el período comprendido de 2020 a 2050,
- que la tasa bruta de natalidad se mantiene constante desde 2050,
- que el trazador cúbico interpolante (9) rige hasta alcanzar la tasa de natalidad mínima en el orbe al año de 2021.

Concluimos que el ejemplo 2.10, la tabla 8, la observación 3.1(d) y la tabla 9 ponen de manifiesto las diferencias señaladas en las obras de [Ayuso y Chuliá \(2018\)](#), y [Cebrián López y Moreno Raymundo \(2015\)](#) en términos de la brecha entre los beneficios a la jubilación para hombres y mujeres. Nuestro artículo las captura para dejar evidencia estadística y matemática sobre la importancia del rol de la mujer en la solidaridad intergeneracional de la dinámica demográfica del futuro. Verdaderamente, equiparar los sueldos de los géneros resultaría en una mitigación evidente de la contribución que este trabajo proyecta que debe hacerse para efectos de la solvencia de un plan de pensiones.

Finalmente, pero no menos importante, recalcamos que hemos creado una herramienta computacional que permite replicar los resultados de esta investigación, así



como generar otros modificando nuestros supuestos. Tal herramienta está disponible [aquí](#).

REFERENCIAS

- Aguirre-Farías, F., 2001 *Diseño conceptual del método de las primas óptimas de liquidez*. Instituto Nacional de Geografía y Estadística, Aguascalientes, México, Disponible [aquí](#).
- Aguirre-Farías, F., 2012 *Pensiones... ¿Y con qué?*. México: Fineo.
- Aguirre-Farías, F. M., F. M. Aguirre-Villarreal, y J. D. López-Barrientos, 2019 Comparison of the costs of the defined-benefit and the defined-contribution schemes under an actuarial methodology. *Revista Mexicana de Economía y Finanzas Nueva Época* 14: 671–691.
- Anderson, A. W., 1992 *Pension Mathematics for Actuaries*. Actex Publications, Connecticut.
- Ayuso, M. y E. Chuliá, 2018 ¿Hacia la progresiva reducción de la brecha de género en las pensiones contributivas? *BBVA: Documentos Mi jubilación* 22: 1–24.
- Banco de México, 2021 Reforma del Sistema de Pensiones de México: Posibles efectos sobre las jubilaciones, la dinámica del ahorro obligatorio y las finanzas públicas. Informe Trimestral Octubre-Diciembre 2020 3 marzo: 53–58, Recuperado de [aquí](#) el 24 de septiembre de 2023.
- Belolipetskii, A. y M. Lepskaya, 2018 A Mathematical Model of Pension Fund Operation and Methods of Fund Stability Analysis. *Computational Mathematics and Modeling* 29: 233–243.
- Belolipetskii, A. y M. Lepskaya, 2019 Mathematical modeling of the operation of pension funds in order to assess their sustainability. *Journal of Computer and Systems Sciences International* 58: 415–424.
- Boikov, A., 2002 The Cramer-Lundberg model with stochastic premium process. *Teoriya Veroyatnostey i eë Primeneniya* 47: 549–553.
- Bowers, N., H. Gerber, J. Hickman, D. Jones, y C. Nesbitt, 1997 *Actuarial mathematics*. The Society of Actuaries, Schaumburg, Illinois.
- Bühlmann, H., 1985 Premium calculation from top down. *ASTIN Bulletin* 15: 89–101.
- Burden, R. L., J. D. Faires, y A. M. Burden, 2015 *Numerical Analysis*. Cengage Learning.
- Cano Ramos, A. E. y J. D. López Barrientos, 2021 ¿“¡Te lo aseguro!” o “¡te lo apuesto!”? ¡He ahí el dilema!... Miscelánea Matemática de la Sociedad Matemática Mexicana 72: 1–20.
- Cebrián López, I. y G. Moreno Raymundo, 2015 Tiempo cotizado, ingresos salariales y sus consecuencias para

- las pensiones: diferencias por género al final de la vida laboral. *Cuadernos de Relaciones Laborales* 33: 311–328, Recuperado de [aquí](#) el 10 de agosto de 2023.
- Cervantes González, R. F. y L. F. Sedano Díaz, 2023 Actual sistema pensionario del Seguro Social en México y su comparación con otros sistemas pensionarios de América Latina. *Actuarios Trabajando* 14: 30–45, Consultado [aquí](#) el 30 de agosto de 2023.
- Comunicación Social, 2022 En 2022 aumenta en 7.36 por ciento el monto de pensiones del IMSS. Acercando el IMSS al Ciudadano 63, Recuperado de [aquí](#) el 24 de septiembre.
- Consejo Nacional de Población, Secretaría de Gobernación, y Secretaría General del Consejo Nacional de Población, 2019 *Colección. Proyecciones de la población de México y las entidades federativas 2016-2050*. República Mexicana. Gobierno de México, Ciudad de México, Recuperado el 9 de septiembre de 2023 de [aquí](#).
- Datos Macro, 2021 *Desciende la tasa de natalidad en México en 2021*. Expansión, México, Recuperado de [aquí](#) el 24 de septiembre de 2023.
- Daykin, C., T. Pentikäinen, y M. Pesonen, 1993 *Practical Risk Theory for Actuaries*. Chapman and Hall/CRC, Boca Raton.
- Diaconis, P. y D. Freedman, 1999 Iterated random functions. *Society for Industrial and Applied Mathematics Review* 41: 45–76.
- Doob, J. L., 1994 *Measure Theory*. Springer-Verlag, New York.
- Duflo, M., 1990 *Méthodes récursives aléatoires*. Masson, Paris.
- Hernández-Lerma, O. y J. B. Lasserre, 1999 *Further Topics on Discrete-Time Markov Control Processes*. Springer, New York.
- Instituto Nacional de Estadística y Geografía, 2023a *Población total por entidad federativa y grupo quinquenal de edad según sexo, serie de años censales de 1990 a 2020*. Instituto Nacional de Estadística y Geografía, Aguascalientes, Recuperado el 14 de septiembre de 2023 de [aquí](#).
- Instituto Nacional de Estadística y Geografía, 2023b Segundo trimestre de 2023. Encuesta Nacional de Ocupación y Empleo, población de 15 años de edad y más 2, Recuperado el 2 de septiembre de 2023 de [aquí](#).
- Instituto Nacional de Geografía y Estadística, 2023 Índice nacional de precios al consumidor. Ciclo de Actualización de la Información Económica **septiembre**, Recuperado de [aquí](#) el 24 de septiembre de 2023.
- Iyer, S., 1999 *Actuarial Mathematics of Social Security Pensions*. Geneva, Switzerland: International Labour Organization.



- Kaas, R., M. Goovaerts, J. Dhaene, y M. Denuit, 2008 *Modern Actuarial Risk Theory using R*. Springer-Verlag, Berlin.
- Lasota, A. y M. C. Mackey, 1994 *Chaos, fractals and noise: stochastic aspects of dynamics*. Springer-Verlag, New York.
- Lippman, S. A. y J. Shaler Stidham, 1977 Individual versus social optimization in exponential congestion systems. *Operations Research* **25**: 233–247.
- Llanos Reynoso, L. F., 2014 *Atiéndeme por favor. Técnicas para administrar y calcular el personal y el número de estaciones de servicio, con el objeto de proporcionar una estupenda experiencia en las filas de espera*. Editorial IMT, México.
- López-Barrientos, J. D., J. M. Mendoza-Madrid, y P. F. González-Vega, 2023 An Abelian theorem for a Markov decision process in a system of interacting objects with unknown random disturbance law. To appear in *Pure and Applied Functional Analysis* Available [here](#).
- López-Barrientos, J. D., E. Silva, y E. Lemus-Rodríguez, 2022a Un uso convulso de la fórmula de convolución (con una moraleja Normal). En *Investigación Estocástica y Estadística en la Educación*, editado por G. D. Salgado Suárez, F. S. Tajonar Sanabria, H. J. Reyes Cervantes, V. H. Vázquez Guevara, B. Juárez Hernández, H. A. Cruz Suárez, y F. Velasco Luna, pp. 37–53, Benemérita Universidad Autónoma de Puebla, Disponible [aquí](#).
- López-Barrientos, J. D., D. A. Zayat-Niño, E. X. Hernández-Prado, y Y. Estudillo-Bravo, 2022b On the Élö-Runyan-Poisson-Pearson method to forecast football matches. *Mathematics* **10**, Recuperado el 2 de septiembre de 2023 de [aquí](#).
- Matthews, K., 2022 Cómo afectan las tasas de interés y la inflación a los planes de pensiones. The PNC Financial Services Group, Inc. **22 de julio**, Recuperado de [aquí](#) el 24 de septiembre de 2023.
- Meyn, S. P. y R. L. Tweedie, 1993 *Markov Chains and Stochastic Stability*. Springer-Verlag, London.
- Mikosch, T., 2009 *Non-Life Insurance Mathematics: An Introduction with the Poisson Process*. Springer, Berlin.
- Norris, J. R., 1997 *Markov Chains*. Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.
- Organisation for Economic Co-operation and Development, 2021 *Pensions at a Glance 2021*. Organisation for Economic Co-operation and Development, Recuperado de [aquí](#) el 25 de septiembre de 2023.
- Plata Gallegos, H. y J. D. López Barrientos, 2023 *Comparación de metodologías de financiamiento en las valuaciones actuariales de planes de pensiones en la Seguridad Social*. Tesis de Maestría. Universidad Anáhuac México. Facultad de Ciencias Actuariales, México.
- Promislow, D., 2011 *Fundamentals of Actuarial Mathematics*. John Wiley and Sons Ltd., London.
- Ramasubramanian, S., 2009 *Lectures on Insurance Models*. Hindustan Book Agency.
- Romeo, G., 2020 *Elements of Numerical Mathematical Economics with Excel. Static and Dynamic Optimization*. Academic Press.
- Ross, S. M., 1997 *Simulation*. Academic Press Inc.
- Saad, P., G. Bay, M. C. Villarroel, y M. Acuña, 2017 *Tablas de Mortalidad*. Comisión Económica para América Latina y el Caribe, Santiago de Chile, Recuperado el 10 de septiembre de 2023 de [aquí](#).
- Schmidli, H., 2008 *Stochastic control in insurance*. Springer-Verlag, London.
- Secretaría de Economía, 2023 *Salario promedio y población ocupada según sexo y edad en el segundo trimestre de 2023*. Gobierno de México, Recuperado de [aquí](#) el 25 de septiembre de 2023.
- Secretaría de Hacienda y Crédito Público y Banco de México, 2023 Resultados de la subasta de valores. Subastas de valores **38**, Recuperado de [aquí](#) el 21 de septiembre de 2023.
- Secretaría del Trabajo y Previsión Social y Banco de México, 2023 Revisión de salarios contractuales. Sistema de Información Económica **8**, Recuperado de [aquí](#) el 8 de septiembre de 2023.
- Villarreal, H. y A. Macías, 2020 El sistema de pensiones en México. Institucionalidad, gasto público y sostenibilidad financiera. *Macroeconomía del Desarrollo* **210**: 7–70.
- World Population Review, 2022 *Top 10 Countries with the Lowest Birth Rate (per 1000 people, CIA World Factbook 2021 estimate)*. World Population Review, Recuperado de [aquí](#) el 7 de agosto de 2023.





José Daniel López Barrientos se graduó de la Licenciatura en Actuaría de la Universidad Anáhuac México en diciembre de 2003. Cuenta con una Maestría y un Doctorado en Ciencias, ambos en la especialidad en Matemáticas, por el Cinvestav. Hizo

dos postdoctorados; uno en la HEC de Montreal, y otro en la Facultad de Matemáticas Aplicadas de la Universidad Estatal de San Petersburgo. Ha impartido clases en varias universidades Mexicanas y extranjeras, en los niveles de Licenciatura, Maestría y Doctorado, y ha sido expositor en diversos foros nacionales e internacionales de investigación en Matemáticas. Actualmente se desempeña como Profesor-Investigador en la Facultad de Ciencias Actuariales de la Universidad Anáhuac México, es Investigador Nacional de nivel 1, preside el Comité permanente de Investigación y Desarrollo Actuarial del Colegio Nacional de Actuarios desde 2014, y sirve como Editor en Jefe de la revista de investigación actuarial aplicada del Colegio Nacional de Actuarios “Actuarios Trabajando” desde 2020. Uno de sus trabajos de investigación fue galardonado con el primer lugar de la edición de 2021 del Premio “Fundadores” de la Asociación Mexicana de Actuarios.



Humberto Plata Gallegos se graduó como Actuario por la Facultad de Estudios Superiores campus Acatlán de la Universidad Nacional Autónoma de México en 2010. En 2023 se graduó de la Maestría en Riesgos de la Facultad de Ciencias Actuariales de la Universidad Anáhuac México, y se encuentra cursando los estudios de Doctorado en Ciencias Actuariales en esa misma institución. Es un apasionado de la seguridad social y las pensiones. Se define por ser una persona con ganas de aprender, un gran líder de equipo, buena comunicación, analítico, comprometido, perseverante, práctico, sociable, activo, ambicioso, entusiasta. Trabajó en el grupo de consultoría Farell desde diciembre de 2010, y hoy día presta sus servicios profesionales en Lockton.

Es un apasionado de la seguridad social y las pensiones. Se define por ser una persona con ganas de aprender, un gran líder de equipo, buena comunicación, analítico, comprometido, perseverante, práctico, sociable, activo, ambicioso, entusiasta. Trabajó en el grupo de consultoría Farell desde diciembre de 2010, y hoy día presta sus servicios profesionales en Lockton.



CURSOS Y TALLERES



FEB 2024

Inscripciones abiertas

Evento Virtual por Zoom



Asociación Mexicana de
Actuarios Consultores, A.C.

Diplomado en Pensiones

Objetivo:

Proporcionar al estudiante de este diplomado un conocimiento ordenado y suficiente que constituya una base sólida para el diseño, valuación de planes públicos y privados de pensiones y su administración.

Duración: 143 hrs

Registro:

informes@amac.mx
55 5559 0514





**Curso
Live Stream**

REASEGURO

HECSE EN TRÁMITE

**INICIO: 6 DE FEBRERO
DURACIÓN 14 HRS**

CONTENIDO

- **ASPECTOS TÉCNICOS Y DEL MERCADO DEL REASEGURO**
- **REASEGURO PROPORCIONAL**
- **REASEGURO NO PROPORCIONAL**
- **REASEGURO EN LOS RIESGOS PERSONALES**
- **GESTIÓN ESTRATÉGICA ACTUARIAL**

@rhcecam



informes@rhcecam.com
WhatsApp: 5548663933

Auditoría de Reaseguro

**FEBRERO 9, 10, 16 Y 17
DE 17:00 A 21:00 HORAS**

PLATAFORMA ZOOM

MIEMBROS CONAC 8 HECSE

INFORMES E INSCRIPCIONES:



55 4809 7800



www.acrglobal.com.mx



capacitación@acrglobal.com.mx



COLEGIO
NACIONAL DE
ACTUARIOS, A.C.



ACTUARY
HUNTERS

CURSO

ONLINE

MEDICIÓN Y CONTROL DE RIESGOS

Fecha: 13 de Febrero de 2024

Total de Horas: 12

✉ informes@actuaryhunters.com

☎ 01 55 51 71 62 86

f @actuaryhunters

☎ 55 83 72 41 90

3 HECSE

Requerimientos de Capital y Prueba de Solvencia Dinámica

**FEBRERO 20, 21 Y 22
DE 17:00 A 21:00 HORAS**

PLATAFORMA ZOOM

**MIEMBROS CONAC
12 HECSE DE NORMATIVIDAD**

INFORMES E INSCRIPCIONES:



55 4809 7800



www.acrglobal.com.mx



capacitación@acrglobal.com.mx

Contabilidad de Instituciones de Seguros”

**FEBRERO 27, 28 Y 29
DE 17:00 A 21:00 HORAS**

PLATAFORMA ZOOM

**MIEMBROS CONAC
3 HECSE DE NORMATIVIDAD**

INFORMES E INSCRIPCIONES:



55 4809 7800



www.acrglobal.com.mx



capacitación@acrglobal.com.mx



COLEGIO
NACIONAL DE
ACTUARIOS, A.C.



ACTUARY
HUNTERS

CURSO

ONLINE

CONTRATOS DE REASEGURO

Fecha: 28 de Febrero de 2024

Total de Horas: 15

✉ informes@actuaryhunters.com

☎ 01 55 51 71 62 86

f @actuaryhunters

📞 55 83 72 41 90

15 HECSE
de las cuales
4 son de normatividad

Curso
Live Stream

PRICING NO VIDA CON MACHINE LEARNING EN R

INICIO: 2 DE MARZO / DURACIÓN: 30 HRS / 6 SÁBADOS

CURSO AVALADO
POR EL CONAC /
OTORGA 16.5 HECSE

- 1 **Exploración de datos**
USO DE GGLOT2 Y DPLYR
- 2 **Modelos GLM y Modelos GBM**
POISSON / BINOMIAL / GAMMA Y TWEEDIE
RANDOM FOREST Y BOOSTING
- 3 **Técnicas de penalización y selección de variables /
Técnicas de aprendizaje no supervisado**
K-MEAN, CLUSTERING, PCA
- 4 **Aplicaciones**
SCORE CREDITICIO
NÚMERO DE RECLAMOS
SEVERIDAD Y VARIABLE DE PÉRDIDA TOTAL
PRICING
- 5 **Análisis comparativo.**
INTERPRETACIÓN DE RESULTADOS.



ACTUARY
HUNTERS



COLEGIO
NACIONAL DE
ACTUARIOS, A.C.

CURSOS ONLINE

FUNDAMENTOS DE PROGRAMACIÓN CON PYTHON

Fecha: 04 de Marzo de 2024
Total de Horas: 30

20 HECSE

MODELACIÓN DE RIESGOS FINANCIEROS

Fecha: 12 de Marzo de 2024
Total de Horas: 12

3 HECSE

REPORTES REGULATORIOS

Fecha: 19 de Marzo de 2024
Total de Horas: 3

3 HECSE

✉ informes@actuaryhunters.com

☎ 01 55 51 71 62 86

fb @actuaryhunters

☎ 55 83 72 41 90

Inversiones de las Instituciones de Seguros

**MARZO 5 Y 7
DE 17:00 A 21:00 HORAS**

PLATAFORMA ZOOM

MIEMBROS CONAC 8 HECSE

INFORMES E INSCRIPCIONES:



55 4809 7800



www.acrglobal.com.mx



capacitación@acrglobal.com.mx



Curso
Live Stream

CONTABILIDAD DE SEGUROS

INICIO: 5 DE MARZO / 4 MARTES,
4 JUEVES / 6:00PM / DURACIÓN: 22 HORAS

CONAC OTORGA 16 HECSE / 6 HRS DE NORMATIVIDAD

PROGRAMA

CONTABILIDAD DE
SEGUROS I: CICLOS
DE LA OPERACIÓN.

CONTABILIDAD DE
SEGUROS II:
ANÁLISIS
FINANCIERO.

CONTABILIDAD DE SEGUROS
III: REPORTES FINANCIEROS
TRIMESTRALES Y ANUALES
OFICIALES.

@rhcecam



informes@rhcecam.com
WhatsApp: 5548663933

ANÁLISIS DE RIESGOS Y SIMULACIÓN DE PROCESOS ESTOCÁSTICOS

INICIO: 5 DE MARZO

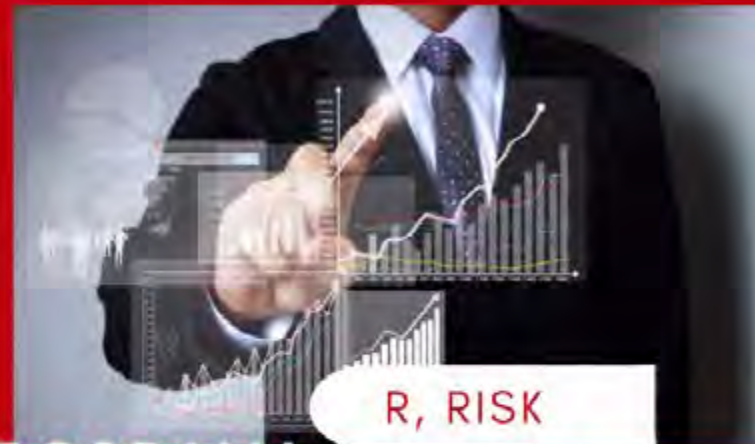
DURACIÓN 15 HRS
CONAC OTORGA 15 HECSE

1. INTRODUCCIÓN AL ANÁLISIS DE RIESGO

- 1.1 RIESGO VS INCERTIDUMBRE
- 1.2 ENFOQUE DETERMINÍSTICO VS MODELO MONTE CARLO (ESTOCÁSTICO)
- 1.3 EL PAPEL DE LAS DISTRIBUCIONES DE PROBABILIDAD EN LA SELECCIÓN DE FUNCIONES DE VARIABILIDAD: CLASIFICACIÓN Y CARACTERÍSTICAS.
- 1.4 EL SOFTWARE @RISK EN EL DISEÑO DE MODELOS DE SIMULACIÓN DE PROCESOS ESTOCÁSTICOS

2. MODELOS DE ANÁLISIS DE RIESGO EN AUSENCIA DE DATOS HISTÓRICOS

- 2.1 DISTRIBUCIONES DE PROBABILIDAD CONTINUAS: UNIFORME, TRIANGULAR, TRIGEN, PERT Y VARY
- 2.2 DETERMINACIÓN DE RIESGOS DE PÉRDIDA MEDIANTE EL USO DE PERCENTILES
- 2.3 ESTRATEGIAS DE MITIGACIÓN DE RIESGOS MEDIANTE HERRAMIENTAS AVANZADAS DE BÚSQUEDA DE OBJETIVOS
- 2.4 SUPERPOSICIÓN DE FUNCIONES DE RIESGO
- 2.5 SELECCIÓN DEL GENERADOR MONTE CARLO Y ASIGNACIÓN DE SEMILLAS
- 2.6 MUESTREO LATINO HIPERCÚBICO VS MONTECARLO SIMPLE : DIFERENCIAS METODOLÓGICAS
- 2.7 CÁLCULO DE PROBABILIDADES DE PÉRDIDA: MÉTODO GRÁFICO Y DE REGISTRO
- 2.8 ASIGNACIÓN DE SALIDAS MÚLTIPLES
- 2.9 FUNCIONES ESTADÍSTICAS DE RESULTADOS.
- 2.10 DISTRIBUCIONES DISCRETAS: BERNOULLI, BINOMIAL, DISCRETA Y POISSON
- 2.11 ANÁLISIS AVANZADO DE STRESS.



PROGRAMA

3. MODELOS DE ANÁLISIS DE RIESGOS CON INFORMACIÓN HISTÓRICA

- 3.1 DISTRIBUCIONES DE PROBABILIDAD CONTINUAS : NORMAL Y LOGNORMAL
- 3.2 AJUSTE PERSONALIZADO DE DISTRIBUCIONES
- 3.3 HERRAMIENTAS AVANZADAS DE JERARQUIZACIÓN DE RIESGOS : GRÁFICO DE TORNADO Y DE ARAÑA
- 3.4 ANÁLISIS AVANZADO DE SENSIBILIDAD
- 3.5 GENERACIÓN DE SIMULACIONES SIMULTÁNEAS PARA MODELOS CON PARÁMETROS MÚLTIPLES
- 3.6 AJUSTE DISTRIBUCIONES
- 3.6.1 PRUEBAS DE BONDAD DE AJUSTE : ESTADÍSTICOS DE PRUEBA VS. MÉTODO BOOTSTRAP PARA CÁLCULO DE VALORES DE SIGNIFICANCIA
- 3.7 AJUSTE DE DISTRIBUCIONES POR LOTES
- 3.8 AJUSTE DE SERIES DE TIEMPO
- 3.9 ANÁLISIS DE CORRELACIÓN AVANZADO MEDIANTE CÓPULAS : DIFERENCIAS ENTRE MODELOS GAUSSIANOS Y NO GAUSSIANOS
- 3.10 GRÁFICOS DE DISPERSIÓN PARA CRUCE DE VARIABLES DE RIESGOS EN ESCENARIOS
- 3.11 AJUSTE DE DISTRIBUCIONES MEDIANTE KERNELS
- 3.12 OPTIMIZACIÓN BAJO CONDICIONES DE RIESGO
- 3.12.1 OPTIMIZACIÓN INDIVIDUAL Y COLECTIVA (FRONTERA EFICIENTE)

4. APLICACIONES

- 4.1 MODELO DE FLUJOS DE EFECTIVO
- 4.2 MODELO DE SEGUROS FINANCIEROS
- 4.3 MODELO DE CARTERAS DE INVERSIÓN
- 4.4 MODELO DE RIESGO OPERATIVO

Gobierno Corporativo en Instituciones de Seguros / Roles y Responsabilidades

MARZO 12, 14, 19 Y 21
DE 17:00 A 21:00 HORAS

PLATAFORMA ZOOM

MIEMBROS CONAC
16 HECSE DE NORMATIVIDAD

INFORMES E INSCRIPCIONES:



55 4809 7800



www.acrglobal.com.mx



capacitación@acrglobal.com.mx



COLEGIO
NACIONAL DE
ACTUARIOS, A.C.



ACTUARY
HUNTERS

CURSO

ONLINE

MODELACIÓN
ACTUARIAL
Y TRANSFERENCIA
DE RIESGOS

Fecha: 03 de Abril de 2024

Total de Horas: 24

10.5 HECSE

✉ informes@actuaryhunters.com

☎ 01 55 51 71 62 86

f @actuaryhunters

IFRS 17

(21 HECSE, 3 HRS DE NORMATIVIDAD)

FSA CARLOS AROCHA

5 DE ABRIL / LÍMITE DE REGISTRO 17 DE MARZO

PROGRAMA

DURACIÓN: 30 HORAS

CONCEPTOS BÁSICOS Y
COMPARACIÓN CON IFRS 4

MODELO GENERAL DE
MEDICIÓN Y VARIANTES

POLÍTICAS CONTABLES,
PRESENTACIÓN DE ESTADOS
FINANCIEROS, NOTAS DE
REVELACIÓN Y
CONSIDERACIONES DE
APLICACIÓN E IMPLEMENTACIÓN

PROCESOS
ACTUARIALES DE
REPORTE FINANCIERO

@rhcecam



informes@rhcecam.com
WhatsApp: 5548663933

CIENCIA DE DATOS APLICADA A LA INGENIERÍA FINANCIERA CON R

1. INTRODUCCIÓN AL LENGUAJE R

- 1.1 CONCEPTOS BÁSICOS DE R EN R STUDIO
- 1.1 INTERFASE DE TRABAJO
- 1.2 COMANDOS Y OPERACIONES BÁSICAS DE EJECUCIÓN
- 1.3 INSTALACIÓN DE LIBRERÍAS DE APOYO

2. FUNDAMENTOS DE TEORÍA DE PORTAFOLIOS APLICADA EN R

- 2.1 CERTEZA, RIESGO E INCERTIDUMBRE
- 2.2 MEDICIÓN DE RENDIMIENTO Y RIESGO EN INVERSIONES: CASO PRÁCTICO EN R
 - 2.2.1 MODELO DE ESCENARIOS DISCRETOS
 - 2.2.2 CÁLCULO DE RENDIMIENTOS ESPERADOS Y CORRELACIÓN DE RIESGOS (COVARIANZAS).
- 2.3 EL MODELO DE PORTAFOLIOS DE INVERSIÓN PARA DOS ACTIVOS
 - 2.3.1 PONDERACIÓN DE ACTIVOS Y SU IMPACTO EN EL RIESGO CONJUNTO DEL PORTAFOLIO
 - 2.3.2 OPTIMIZACIÓN DEL PORTAFOLIO ÓPTIMO EN R
 - 2.3.3 EL CONCEPTO DE FRONTERA EFICIENTE: ENFOQUE DE TABULACIÓN
- 2.3 REPRESENTACIÓN MATRICIAL DEL MODELO GENERAL DE PORTAFOLIOS EN R
 - 2.3.1 EL MODELO GENERAL DE HARRY MARKOWITZ PARA N ACTIVOS
- 2.3 SIMULACIÓN MONTE CARLO PARA PROYECCIÓN DE GANANCIAS Y PÉRDIDAS
 - 2.3.1 ESQUEMA GENERAL DE MODELACIÓN
 - 2.3.2 GENERACIÓN DE ESCENARIOS ALEATORIZADOS EN R
 - 2.3.3 ALEATORIZACIÓN DE EVENTOS DISCRETOS
 - 2.3.3 ALEATORIZACIÓN A PARTIR DE FUNCIONES DE PROBABILIDAD



SOFTWARE R

PROGRAMA

3. AUTOMATIZACIÓN DE PORTAFOLIOS EN R EN TIEMPO REAL

- 3.1 EXTRACCIÓN DE PRECIOS HISTÓRICOS DEL MERCADO ACCIONARIO
 - 3.2.2 MANEJO DE VECTORES PARA CÓDIGOS BURSÁTILES
 - 3.2.3 SELECCIÓN DE FECHAS DE EXTRACCIÓN
- 3.3 CONSOLIDACIÓN DE PRECIOS DE CIERRE DE PORTAFOLIOS DE EMPRESAS
- 3.4 OBTENCIÓN DE RENDIMIENTOS LOGARÍTMICOS
- 3.5 OPTIMIZACIÓN Y GRAFICACIÓN DE LA FRONTERA EFICIENTE
 - 3.5.1 PORTAFOLIO DE VARIANZA MÍNIMA
 - 3.5.2 PORTAFOLIO DE TANGENCIA
 - 3.5.3 PORTAFOLIO MONTE CARLO
 - 3.5.4 PORTAFOLIO DE RAZÓN DE SHARPE
- 3.6 GENERACIÓN DE ARCHIVOS DE REPORTE

INICIO: 7 DE ABRIL

DURACIÓN 15 HRS
CONAC OTORGA 12 HECSE

Auditoría Interna del Proceso ARSI

**ABRIL 9, 11, 16 Y 18
DE 17:00 A 21:00 HORAS**

PLATAFORMA ZOOM

**MIEMBROS CONAC
16 HECSE DE NORMATIVIDAD**

INFORMES E INSCRIPCIONES:



55 4809 7800



www.acrglobal.com.mx



capacitación@acrglobal.com.mx

**Curso
Live Stream**

SOLVENCIA II CON R

(30 HECSE, 10 HRS DE NORMATIVIDAD)

INICIO: 23 DE ABRIL / 2 MARTES, 5 SÁBADOS
DURACIÓN: 30 HORAS

PROGRAMA

I: FUNDAMENTOS DE R
REGRESIÓN Y
SIMULACIÓN DE TABLAS
DE MORTALIDAD

II: MEJOR
ESTIMADOR (BEL)

III: MARGEN DE
RIESGO (MR)

IV: REQUERIMIENTO DE
CAPITAL DE SOLVENCIA (RCS)
Y OTROS TOPICOS

V: INTRODUCCIÓN A LA VALUACIÓN ACTUARIAL MARKET CONSISTENT
EMBEDDED VALUE Y AGREGACIÓN DE RIESGOS Y MÁS

Análisis e Interpretación de Estados Financieros de Instituciones de Seguros para no Financieros

**ABRIL 24, 25 Y 26
DE 17:00 A 21:00 HORAS**

PLATAFORMA ZOOM

**MIEMBROS CONAC
3 HECSE DE NORMATIVIDAD**

INFORMES E INSCRIPCIONES:



55 4809 7800



www.acrglobal.com.mx



capacitación@acrglobal.com.mx



COLEGIO
NACIONAL DE
ACTUARIOS, A.C.



ACTUARY
HUNTERS

CURSO

ONLINE

INFERENCIA DE DATOS PARAMÉTRICOS

Fecha: 30 de Abril de 2024

Total de Horas: 24

24 HECSE

✉ informes@actuaryhunters.com

☎ 01 55 51 71 62 86

f @actuaryhunters

📞 55 83 72 41 90

MACHINE LEARNING CON PYTHON

INICIO: 11 DE MAYO

DURACIÓN 30 HRS
CONAC OTORGA 20 HECSE

PROGRAMA

1

MODELOS SUPERVISADOS

- ALGORITMOS DE ML
- REGRESIÓN LINEAL, REGRESIÓN LOGÍSTICA Y STOCHASTIC GRADIENT DESCENT
- REGRESIÓN CON ÁRBOLES DE DECISIÓN Y BOSQUES ALEATORIOS
- CLASIFICACIÓN CON ÁRBOLES DE DECISIÓN Y BOSQUES ALEATORIOS
- K-NEAREST-NEIGHBOR (KNN)
- SUPPORT VECTOR MACHINE

2

MODELOS NO SUPERVISADOS

- HIERARCHICAL CLUSTERING
- K-MEANS
- PRINCIPAL COMPONENT ANALYSIS

3

OPTIMIZACIÓN DE PARÁMETROS

Cobertura de la Base de Inversión y Fondos Propios Admisibles

**MAYO 14, 15 Y 16
DE 17:00 A 21:00 HORAS**

PLATAFORMA ZOOM

**MIEMBROS CONAC
12 HECSE, 10 DE NORMATIVIDAD**

INFORMES E INSCRIPCIONES:



55 4809 7800



www.acrglobal.com.mx



capacitación@acrglobal.com.mx



COLEGIO
NACIONAL DE
ACTUARIOS, A.C.



ACTUARY
HUNTERS

CURSO

ONLINE

BEL Y MARGEN DE RIESGOS

Fecha: 15 de Mayo de 2024

Total de Horas: 24

✉ informes@actuaryhunters.com

☎ 01 55 51 71 62 86

f @actuaryhunters

☎ 55 83 72 41 90

24 HECSE
de las cuales
5 son de normatividad



COLEGIO
NACIONAL DE
ACTUARIOS, A.C.



ACTUARY
HUNTERS

CURSO

ONLINE

ACUERDO SERVICIOS DE ASISTENCIA

Fecha: 20 de Mayo de 2024

Total de Horas: 6

✉ informes@dactuaryhunters.com

☎ 01 55 51 71 62 86

f @dactuaryhunters

☎ 55 83 72 41 90

1 HECSE
de las cuales
1 es de normatividad

IFRS-17 Contratos de Seguros / Impacto Contable y Financiero de Implementación

**MAYO 29, 30 Y 31
DE 16:00 A 20:00 HORAS**

PLATAFORMA ZOOM

MIEMBROS CONAC 12 HECSE

INFORMES E INSCRIPCIONES:



55 4809 7800



www.acrglobal.com.mx



capacitación@acrglobal.com.mx



COLEGIO
NACIONAL DE
ACTUARIOS, A.C.



ACTUARY
HUNTERS

CURSO

ONLINE

GOBIERNO CORPORATIVO:
FUNCIÓN ACTUARIAL,
ADMINISTRACIÓN INTEGRAL
DE RIESGOS Y CONTRALORÍA
INTERNA

Fecha: 18 de Junio de 2024

Total de Horas: 3

✉ informes@actuaryhunters.com

☎ 01 55 51 71 62 86

📘 @actuaryhunters

📞 55 83 72 41 90

1 HECSE

de las cuales

1 es de normatividad



Global Solutions

NIF D-3

Beneficios a los empleados

**JUNIO 18, 19 Y 20
DE 16:00 A 20:00 HORAS**

PLATAFORMA ZOOM

MIEMBROS CONAC 3 HECSE

INFORMES E INSCRIPCIONES:



55 4809 7800



www.acrglobal.com.mx



capacitaci3n@acrglobal.com.mx



COLEGIO
NACIONAL DE
ACTUARIOS, A.C.



ACTUARY
HUNTERS

CURSO

ONLINE

ADMINISTRACIÓN DE RIESGOS FINANCIEROS

Fecha: 28 de Junio de 2024

Total de Horas: 24

✉ informes@actuaryhunters.com

☎ 01 55 51 71 62 86

f @actuaryhunters

☎ 55 83 72 41 90

24 HECSE



COLEGIO
NACIONAL
DE ACTUARIOS

MÉXICO