

Cadenas de Markov absorbentes para analizar el abandono escolar de estudiantes universitarios de matemáticas

Absorbing Markov chains for analyzing dropout of mathematics undergraduate students

Ana L. Olmedo-Alonso ^a, Axel I. García-Bernal ¹, Aarón V. Reyes-Rodríguez ^{1,*}, Margarita Tetlalmatzi-Montiel ¹

^aÁrea Académica de Matemáticas y Física, Universidad Autónoma del Estado de Hidalgo, 42184, Mineral de la Reforma, Hidalgo, México.

Resumen

El abandono de los estudios universitarios es un problema que tiene altos costos sociales y económicos para las naciones, los cuales se reflejan en el alcance de la cobertura en la educación superior. Los altos índices de deserción o abandono de los estudios universitarios disminuyen las oportunidades de la población para acceder este nivel educativo. En este contexto, se analizan las trayectorias académicas de estudiantes de la Licenciatura en Matemáticas Aplicadas, de la Universidad Autónoma del Estado de Hidalgo (LIMA-UAEH), mediante cadenas de Markov homogéneas a tiempo discreto; a partir de información publicada por la Dirección General de Planeación de la universidad. Específicamente, se calcularon las probabilidades de que los estudiantes de la LIMA-UAEH deserten o finalicen sus estudios universitarios; así como el tiempo promedio en que egresan o desertan, dado el semestre en el que se encuentran inscritos. Además, se realizó una comparativa entre cohortes que fueron afectadas, o no, por la pandemia derivada de la COVID-19.

Palabras Clave: Cadenas de Markov, abandono escolar, estudiantes de licenciatura, matemáticas, probabilidad.

Abstract

University students' dropout is a problem that has high social and economic costs, which are reflected in the higher education coverage. High rates of university dropout or abandonment decrease the opportunities for the population to have access to this educational level. In this context, the academic careers of students of the Undergraduate Program of Applied Mathematics at the Autonomous University of Hidalgo State are analyzed through discrete time homogeneous Markov Chains, based on information published by the Office of General Planning and Budget of this University. Specifically, probabilities of dropping out or finishing university studies were calculated; as well as the average time in which students graduate or drop out, given the semester they are enrolled in. In addition, a comparison between cohorts affected and not affected by the COVID-19 pandemic was carried out.

Keywords: Markov chains, university dropout, undergraduate students, mathematics, probability.

1. Introducción

El abandono y la deserción escolar, en los diversos niveles educativos, son problemas que aquejan a todos los países del mundo. De acuerdo con la *Organización para la Cooperación y Desarrollo Económico* (OECD, 2022), en promedio, solo el 39 % de las personas que ingresan a una licenciatura se gradúan en los tiempos marcados por los programas correspondientes. La tasa de graduación, tres años después del tiempo indicado, se incrementa a 68 %. Un dato interesante es que las

tasas de graduación para las mujeres son mayores que para los hombres en todos los lugares en los que se dispone de dicha información. En promedio, 12 % de quienes ingresan a licenciatura abandonan la universidad antes del inicio de su segundo año de estudios. El porcentaje de estudiantes que abandonan la universidad, sin graduarse, es en promedio de 21 % al final de la duración teórica del programa de licenciatura (OECD, 2022).

En el caso de México, las brechas de cobertura educativa, durante el ciclo escolar 2020-2021 aumentaron con respecto al ciclo previo, debido a una falta de políticas públicas para evitar

* Autor para correspondencia: aaronr@uaeh.edu.mx

Correo electrónico: ol333736@uaeh.edu.mx (Ana Luisa Olmedo-Alonso), ga419623@uaeh.edu.mx (Axel Iván García-Bernal), aaronr@uaeh.edu.mx (Aarón Víctor Reyes-Rodríguez), tmontiel@uaeh.edu.mx (Margarita Tetlalmatzi Montiel).

Historial del manuscrito: recibido el DD/MM/AAAA, última versión-revisada recibida el DD/MM/AAAA, aceptado el DD/MM/AAAA, en línea (postprint) desde el DD/MM/AAAA, publicado el DD/MM/AAAA. DOI: <https://doi.org/10.29057/icbi.vXXiYY.ZZZZ>



el abandono escolar durante la pandemia debida a la COVID-19, declarada oficialmente por la OMS el 11 de marzo de 2020 (WHO, 2020). Durante la pandemia, 5.2 millones de alumnos de entre 3 y 29 años no se inscribieron para cursar el ciclo escolar 2020-2021, siendo las principales razones el confinamiento en los hogares y la falta de recursos económicos (INEGI, 2022). En esta misma línea de ideas, la *Secretaría de Educación Pública* (SEP) reportó que la tasa de abandono escolar en el nivel superior en México, durante los años 2021, 2022 y 2023 fue de 8.8 %, 8.1 % y 6 %, respectivamente. En el caso del estado de Hidalgo, las tasas de abandono escolar, para el mismo nivel y los mismos años, fueron de 8.5 %, 9 % y 7.4 %, respectivamente (SEP, 2023).

El abandono escolar es un fenómeno de relevancia para las universidades, dado que los índices de deserción, retención y eficiencia terminal impactan en su posicionamiento en los rankings nacionales e internacionales; los cuales, a su vez, son un elemento considerado en las solicitudes de financiamiento público (Bahi *et al.*, 2015) y en las evaluaciones de organismos acreditadores como los *Comités Interinstitucionales para la Evaluación de la Educación Superior* (CIEES) o el *Consejo para Acreditación de la Educación Superior* (COPAES). En el caso de la Universidad Autónoma del Estado de Hidalgo (UAEH), los indicadores sobre eficiencia terminal, tasa de deserción y el tiempo medio de egreso de la licenciatura son componentes esenciales de las estrategias institucionales encaminadas a lograr la visibilidad y reconocimiento internacionales (UAEH, 2019). Al respecto, en la UAEH se han detectado algunas áreas de oportunidad, entre ellas que los indicadores globales de la institución, respecto de la eficiencia terminal y retención, se encuentran por debajo de la media nacional (UAEH, 2023b).

El abandono de los estudios universitarios es un problema ampliamente investigado (Ramón-Santiago *et al.*, 2023), pero aún existen aspectos que se desconocen o que no han sido clarificados. Las principales dificultades para abordar este tema radican en una carencia de precisión terminológica, así como la falta de información estadística confiable. Por ejemplo, los investigadores utilizan términos tales como deserción escolar o abandono escolar para referirse al mismo fenómeno de desafiliación de los estudios formales antes de culminarlos (Orozco-Rodríguez, 2022). Por otro lado, algunos estudios miden la deserción escolar con base en cohortes aparentes, pero lo más adecuado es realizar estimaciones basadas en cohortes reales (Vries *et al.*, 2011). Una *cohorte aparente* se refiere a un conjunto de estudiantes que, en cierto momento, se encuentran inscritos en el mismo semestre, pero que no ingresaron a la universidad en el mismo ciclo escolar; es decir, en una cohorte aparente puede haber tanto estudiantes rezagados, como estudiantes que hayan adelantado materias (Cabrera-González *et al.*, 2018).

La SEP utiliza el término *abandono escolar* para referirse al retiro, temporal o definitivo, del sistema educativo antes de terminar algún grado o nivel escolar (SEP, 2020). Otros autores, como Tinto (2006), utilizan el término *deserción escolar*, el cual definen como un proceso de abandono voluntario o forzoso, temporal o definitivo de la carrera en que se matricula un estudiante, por influencia de circunstancias internas o externas al alumno; sin embargo, en el presente artículo dichos términos se emplean como sinónimos.

Otra de las dificultades para el estudio del fenómeno de abandono escolar universitario consiste en el bajo uso de herramientas matemáticas para cuantificar el riesgo de abandono, lo cual trae como consecuencia una carencia de estrategias fundamentadas en información empírica de calidad que permitan minimizar tales riesgos (Rodríguez-Ríos, 2012). Por ejemplo, se ha reportado que diversos estudios utilizan únicamente estadísticas descriptivas para estimar el tiempo promedio en que estudiantes inscritos en licenciaturas de matemáticas abandonan sus estudios (Bahi *et al.*, 2015).

2. Antecedentes

La literatura de investigación sobre la deserción escolar es amplia, incluyendo investigaciones enfocadas en la deserción universitaria. La gran mayoría de estos trabajos identifican variables que influyen sobre el abandono escolar de licenciaturas en ingeniería industrial (Ruiz-Ramírez *et al.*, 2021; Sánchez-González y Ruiz, 2019) o administración (da Costa *et al.*, 2018).

Entre los principales resultados de los trabajos anteriores se destaca que los factores de la deserción o abandono incluyen la falta de recursos económicos, el sentimiento de decepción porque la carrera no es lo que el estudiante esperaba o las dificultades para compaginar el trabajo y el estudio (Ruiz-Ramírez *et al.*, 2021). Mención especial merecen las investigaciones que analizan la deserción o abandono escolar durante el confinamiento derivado de la COVID-19, las cuales reconocen como variables explicativas a características del perfil docente y las prácticas de enseñanza; así como aspectos coyunturales relacionados con la salud de los estudiantes o de sus familiares (Ramón-Santiago *et al.*, 2023).

En lo que respecta a las investigaciones que abordan la deserción en licenciaturas de matemáticas se identificaron preguntas, marcos y metodologías diversas. El interés por el tema se debe a que, generalmente, los índices de reprobación y deserción en estos programas educativos son mayores, en relación con los de otras licenciaturas (Martino y Gregorio, 2019; Orozco-Rodríguez, 2022). Se encontraron y revisaron trabajos que analizan la deserción escolar de estudiantes de matemáticas en diversas universidades de México, entre las que destacan: Universidad Autónoma de Yucatán (Aparicio, 2006; Batún *et al.*, 2023), Universidad Autónoma de Baja California (Jordan-Aramburo *et al.*, 2016), Universidad Autónoma Juárez de Tabasco (Bolívar-Cimé *et al.*, 2016), Benemérita Universidad Autónoma de Puebla (Muñoz, 2017) y Universidad de Guadalajara (Orozco-Rodríguez, 2022).

En las investigaciones enfocadas en los factores que influyen en la deserción escolar, los datos provienen principalmente de cuestionarios realizados a estudiantes (Orozco-Rodríguez, 2022). Entre las herramientas utilizadas para el análisis de la información empírica se encuentran estadísticas descriptivas (Bahi *et al.*, 2015), modelos de regresión logística (Bahi *et al.*, 2015; Orozco-Rodríguez, 2022), cadenas de Markov (Bolívar-Cimé *et al.*, 2016), modelos de supervivencia (Bahi *et al.*, 2015; Muñoz-Vargas, 2017; Ronco, 1994; Willett y Singer, 1991) o métodos cualitativos (Aparicio, 2006).

Entre las principales conclusiones de las investigaciones previas se resalta el impacto de la condición socioeconómica de

los estudiantes sobre la deserción (Orozco-Rodríguez, 2022). Se concluye también que las probabilidades de deserción disminuyen conforme se va avanzando en la licenciatura (Bolívar-Cimé et al., 2016). En el estudio elaborado por Bolívar-Cimé et al. (2016) la probabilidad de egreso para estudiantes de matemáticas fue de 0.63, el tiempo medio para la deserción de dos años y el tiempo promedio de egreso de cinco años, con un tiempo teórico de egreso de cuatro años (Bolívar-Cimé et al., 2016).

Con base en la revisión literatura se identificó que el tema del abandono escolar es relevante, así como el análisis de las trayectorias escolares de los estudiantes de matemáticas, dados los mayores índices de deserción de esta licenciatura, en relación con otras carreras. Así, en este artículo se utilizan cadenas de Markov discretas a tiempo discreto, con probabilidades de transición estacionarias, para analizar la deserción escolar de estudiantes de la LIMA-UAEH. Las cadenas de Markov discretas a tiempo discreto, son de utilidad para este propósito ya que permiten modelar un sistema y describir ciertas propiedades y, con base en dicha información, tomar decisiones fundamentadas que, a su vez, orienten las políticas del sistema; en este estudio son las políticas institucionales relativas al egreso, la deserción y retención de estudiantes de licenciatura. Además, hay evidencia del uso de modelos con base a cadenas de Markov para evaluar comportamiento a largo plazo de grupos de individuos (Zais y Zhang, 2016).

3. Cadenas de Markov

En esta sección se exponen los conceptos fundamentales relacionados con las cadenas de Markov con base en el texto *An Introduction to Stochastic Modeling* (Taylor y Karlin, 1998).

Un proceso de Markov $X = \{X_t\}_{t \in T}$ donde T es un conjunto de índices, es un proceso estocástico con la siguiente propiedad: dado el valor de X_t , los valores de X_s , para $s > t$ no están influenciados por los valores de X_u , para $u < t$. Lo anterior significa que la probabilidad de cualquier comportamiento futuro del proceso, dado que se conoce el estado actual o presente, no se modifica por el conocimiento que se tiene sobre el pasado de dicho proceso.

Una *cadena de Markov a tiempo discreto* y con espacio de estados discreto es un proceso de Markov, cuyo espacio de estados E es un conjunto numerable, y cuyo conjunto de índices es un elemento del conjunto $T = \{0, 1, 2, 3, \dots\}$. En otras palabras, es un proceso estocástico a tiempo discreto, con espacio de estados discreto.

En notación simbólica, la propiedad de Markov se expresa de la siguiente manera:

$$\begin{aligned} P[X_{n+1} = j | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i] = \\ P[X_{n+1} = j | X_n = i]. \end{aligned} \quad (1)$$

Para todos los valores de n , y para todos los estados $i_0, i_1, \dots, i_{n-1}, i, j$.

Generalmente, los elementos del espacio de estados se etiquetan con los enteros no negativos, es decir $E = \{1, 2, 3, \dots\}$, a menos que se especifique otra cosa. Se dice que X_n está en el estado i si $X_n = i$. Así, la probabilidad de que X_{n+1}

esté en el estado i , dado que X_n está en el estado j se denomina *probabilidad de transición en un paso* y se denota como $P_{ij}^{n,n+1} = P[X_{n+1} = j | X_n = i]$. La notación anterior expresa que las probabilidades de transición son funciones de los estados inicial y final, y también del tiempo en que se efectúa la transición. Cuando las probabilidades de transición en un paso son independientes del tiempo, se dice que la cadena de Markov tiene *probabilidades de transición estacionarias*.

En lo subsecuente, cuando nos refiramos a una cadena de Markov estaremos haciendo referencia a una cadena discreta a tiempo discreto y con probabilidades de transición estacionarias. Entonces $P_{ij}^{n,n+1} = P_{ij}$, es independiente de n , y P_{ij} es la probabilidad de que el proceso efectúe una transición del estado i al estado j en un paso.

3.1. Matriz de transición

Es frecuente presentar las probabilidades de transición en un arreglo matricial $\mathbf{P} = \|P_{ij}\|$, denominado matriz de probabilidades transición o, simplemente, matriz de transición de la cadena de Markov.

$$\mathbf{P} = \begin{pmatrix} P_{00} & P_{01} & P_{02} & P_{03} & \dots \\ P_{10} & P_{11} & P_{12} & P_{13} & \dots \\ P_{20} & P_{21} & P_{22} & P_{23} & \dots \\ \vdots & \vdots & \vdots & \vdots & \dots \\ P_{i0} & P_{i1} & P_{i2} & P_{i3} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (2)$$

El i -ésimo renglón de \mathbf{P} , para $i = 0, 1, 2, \dots$, es la distribución condicional de la variable aleatoria X_{n+1} dado que $X_n = i$. Si la cantidad de estados es finita, digamos k , entonces \mathbf{P} es una matriz cuadrada de dimensión $k \times k$. Las entradas P_{ij} satisfacen las siguientes condiciones:

$$P_{ij} = P[X_{n+1} = j | X_n = i] \geq 0 \quad (3)$$

y

$$\sum_{j=0}^{\infty} P_{ij} = 1 \quad (4)$$

para cualesquiera estados $i, j = 0, 1, 2, \dots$.

La relación expresada por la ecuación (4) indica que se efectúa una transición en cada ensayo. Por conveniencia, se dice que se ha efectuado una transición incluso si la cadena de Markov permanece en el mismo estado.

Una cadena de Markov se encuentra completamente definida una vez que se especifican las probabilidades de transición en un paso y la distribución de probabilidad del estado inicial X_0 .

El análisis de una cadena de Markov involucra principalmente el cálculo de las probabilidades y la matriz de transición en n pasos $\mathbf{P}^{(n)} = \|P_{ij}^{(n)}\|$, donde $P_{ij}^{(n)}$ denota la probabilidad de que el proceso pase del estado i al estado j en n pasos o transiciones. En símbolos,

$$P_{ij}^{(n)} = P[X_{n+m} = j | X_m = i]. \quad (5)$$

Como los procesos estocásticos considerados son homogéneos en el tiempo, entonces tienen probabilidades de

transición estacionarias, ya que, en caso contrario, el lado izquierdo de la ecuación (5) dependería también de m , que representa el tiempo en el que se lleva a cabo la transición. La propiedad de Markov permite expresar a las probabilidades de transición en n pasos en términos de la matriz de transición en un paso, como se expresa a continuación.

$$P_{ij}^n = \sum_{k=0}^{\infty} P_{ik} P_{kj}^{(n-1)}, \quad (6)$$

donde,

$$P_{ij}^{(0)} := \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j. \end{cases} \quad (7)$$

La justificación de la relación expresada en la ecuación (6) se bosqueja a continuación. En esta justificación resulta fundamental la propiedad de Markov y algunas propiedades de la probabilidad condicional.

$$P_{ij}^{(n)} = P[X_n = j | X_0 = i] \quad (8)$$

$$= \sum_{k=0}^{\infty} P[X_n = j, X_1 = k | X_0 = i] \quad (9)$$

$$= \sum_{k=0}^{\infty} P[X_1 = k | X_0 = i] P[X_n = j | X_0 = i, X_1 = k] \quad (10)$$

$$= \sum_{k=0}^{\infty} P_{ik} P_{kj}^{(n-1)}. \quad (11)$$

3.2. Análisis del primer paso

Se dice que un estado i de la cadena de Markov es *absorbente* si y sólo si $P_{ii} = 1$. El análisis del primer paso es útil cuando se tiene una cadena de Markov con espacio de estados finito, donde dos o más estados son absorbentes, si el estado inicial de una cadena es transitorio (no absorbente), esta técnica permite calcular el tiempo en el que el proceso se absorbe y el tiempo promedio de absorción en ese estado.

Así, para una cadena de Markov con espacio de estados finitos $E = \{0, 1, 2, \dots, N\}$, donde los estados $0, 1, 2, \dots, r-1$ son transitorios y los estados $r, r+1, \dots, N$ son absorbentes, la matriz de transición tiene la forma

$$\mathbf{P} = \begin{bmatrix} \mathbf{Q} & \mathbf{R} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \quad (12)$$

Donde $\mathbf{0}$ es una matriz de dimensión $(N-r+1) \times r$ cuyas entradas son ceros, \mathbf{I} es la matriz identidad de tamaño $N-r+1$, \mathbf{R} una matriz de dimensión $r \times (N-r+1)$ y $Q_{ij} = P_{ij}$ para $0 \leq i, j < r$.

Iniciando en uno de los estados transitorios $X_0 = i$ donde $0 \leq i < r$, el proceso continuará moviéndose hasta finalmente quedar atrapado en uno de los estados absorbentes $i = r, \dots, N$.

Algunos funcionales de una cadena de Markov que resultan relevantes son el tiempo promedio de transición hasta caer en un estado absorbente y la distribución de probabilidad entre los estados en los que tiene lugar la absorción.

Para conocer el tiempo promedio que le toma al proceso caer en alguno de los estados absorbentes, primero se fija un estado k absorbente ($r \leq k \leq N$). Así, la probabilidad de absorción en el estado k , a diferencia de algún otro estado absorbente,

depende del estado inicial $X_0 = i$. Dicha probabilidad se denotará como $U_{ik} = u_i$.

El análisis del primer paso inicia enumerando las posibilidades en la primera transición. Observe que si se parte del estado i , el proceso pasa inmediatamente al estado k con probabilidad P_{ik} para luego permanecer ahí, siendo así la primera posibilidad considerada. Alternativamente, el proceso podría pasar en su primer paso a un estado absorbente $j \neq k$, para $r \leq j \leq N$, en dicho caso, se excluye la absorción final en el estado k . Finalmente, el proceso podría pasar a un estado transitorio $j < r$. Debido a la propiedad de Markov, una vez en el estado j , la probabilidad de absorción final en el estado k es $u_j = U_{jk}$.

Así, tomando en cuenta las posibilidades enumeradas por sus respectivas probabilidades, mediante la ley de probabilidad total, se obtiene la relación

$$u_i = P[\text{Absorción en } k | X_0 = i] \quad (13)$$

$$= \sum_{j=0}^N P[\text{Absorción en } k | X_0 = i, X_1 = j] P_{ij} \quad (14)$$

$$= P_{ik} + \sum_{j=r}^N P_{ij} \times 0 + \sum_{j=0}^{r-1} P_{ij} u_j, \quad (15)$$

para $j \neq k$.

En resumen, para un estado absorbente fijo k , las cantidades

$$u_i = U_{ik} P[\text{Absorción en } k | X_0 = i], \quad (16)$$

para $0 \leq i < r$ satisfacen el sistema de ecuaciones lineales no homogéneo

$$U_{ik} = P_{ik} + \sum_{j=0}^{r-1} P_{ij} U_{jk}, \quad i = 1, 2, \dots, r-1. \quad (17)$$

Por otra parte, para $\{X_n\}$ una cadena de Markov con matriz de transición de la forma indicada en la ecuación (12), se dispone de una forma más general para calcular el tiempo promedio de absorción considerando el tiempo de absorción aleatorio T .

Formalmente, se define $T = \min\{n \geq 0 : X_n \geq r\}$. Suponiendo que a cada estado transitorio i se le asocia una tasa $g(i)$, el objetivo es determinar la tasa total media que se acumula hasta la absorción. Sea w_i esta cantidad media total, donde i denota la posición inicial $X_0 = i$, entonces

$$w_i := E \left[\sum_{n=0}^{T-1} g(X_n) | X_0 = i \right]. \quad (18)$$

Note que la elección de $g(i) = 1$ para toda i obtenemos que $w_i = E[T | X_0 = i]$, es decir, el tiempo medio de absorción. Si para un estado transitorio k elegimos que g sea de la siguiente forma

$$g(i) = \delta_{ik} := \begin{cases} 1 & \text{si } i = k, \\ 0 & \text{si } i \neq k. \end{cases} \quad (19)$$

esto nos da $w_i = W_{ik}$ el número promedio de visitas al estado k transitorio antes de entrar en algún estado absorbente.

La suma $\sum_{n=0}^{T-1} g(X_n)$ siempre incluye $g(X_0) = g(i)$, donde i es el estado inicial. Además, si una transición ocurre de i a otro

estado j entonces la suma incluye los términos futuros asociados a la trayectoria que empieza en j , usando la propiedad de Markov se deduce que el valor esperado de la suma empezando en j es w_j . Ponderando con la probabilidad de estas transiciones P_{ij} y usando el teorema de probabilidad total obtenemos

$$w_i = g(i) + \sum_{j=0}^{r-1} P_{ij} w_j \quad \text{para } i = 0, 1, \dots, r-1. \quad (20)$$

De lo anterior, expresado en términos matriciales

$$W_{ik} = \delta_{ik} + \sum_{j=0}^{r-1} P_{ij} W_{jk} \quad \text{para } i = 0, 1, \dots, r-1, \quad (21)$$

donde W_{ik} representa el número promedio de visitas al estado k antes de caer en un estado absorbente.

3.3. Aproximación matricial para el análisis del primer paso

En la sección anterior se indicó que la matriz de transición \mathbf{P} , para una cadena de Markov con espacio de estados finitos $E = \{0, 1, 2, \dots, N\}$, donde los estados $0, 1, 2, \dots, r-1$ son transitorios y los estados $r, r+1, \dots, N$ son absorbentes, se puede escribir como se muestra en la ecuación (12). Realizando una multiplicación de matrices por bloques se verifica que:

$$\mathbf{P}^2 = \begin{bmatrix} \mathbf{Q}^2 & \mathbf{I}\mathbf{R} + \mathbf{Q}\mathbf{R} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}. \quad (22)$$

Siguiendo el mismo procedimiento, es posible justificar que la matriz de transición en n pasos, es decir, \mathbf{P}^n tiene la forma:

$$\mathbf{P}^n = \begin{bmatrix} \mathbf{Q}^n & (\mathbf{I} + \mathbf{Q} + \mathbf{Q}^2 + \dots + \mathbf{Q}^{n-1})\mathbf{R} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}. \quad (23)$$

A continuación, se interpretará la ecuación (23). Sea $W_{ij}^{(n)}$ el número promedio de visitas al estado j hasta la n -ésima transición de la cadena de Markov que inició en el estado i . Formalmente,

$$W_{ij}^{(n)} = E \left[\sum_{l=0}^n \mathbf{1}\{X_l = j\} \mid X_0 = i \right], \quad (24)$$

donde

$$\mathbf{1}\{X_l = j\} = \begin{cases} 1 & \text{si } X_l = j, \\ 0 & \text{si } X_l \neq j. \end{cases} \quad (25)$$

es la función indicadora de que la cadena de Markov X , en el instante l , se encuentra en el estado j .

Ahora, $E[\mathbf{1}\{X_l = j\} \mid X_0 = i] = P[X_l = j \mid X_0 = i] = \mathbf{P}_{ij}^{(l)}$, y dado que la esperanza de una suma es la suma de los valores esperados, con base en la ecuación (24) se tiene que:

$$W_{ij}^{(n)} = \sum_{l=0}^n E[\mathbf{1}\{X_l = j\} \mid X_0 = i] = \sum_{l=0}^n \mathbf{P}_{ij}^{(l)}. \quad (26)$$

La ecuación (26) se satisface para todos los estados, pero tiene un significado relevante cuando los estados i, j son transitorios. La ecuación (23) asegura que $\mathbf{P}_{ij}^{(l)} = \mathbf{Q}_{ij}^{(l)}$ cuando

$0 \leq i, j < r$, entonces $W_{ij}^{(n)} = \mathbf{Q}_{ij}^{(0)} + \mathbf{Q}_{ij}^{(1)} + \dots + \mathbf{Q}_{ij}^{(n)}$, para $0 \leq i, j < r$, donde

$$\mathbf{Q}_{ij}^{(0)} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j. \end{cases} \quad (27)$$

En notación matricial $\mathbf{Q}^{(0)} = \mathbf{I}$, y dado que $\mathbf{Q}^{(n)} = \mathbf{Q}^n$, la n -ésima potencia de \mathbf{Q} , entonces

$$\mathbf{W}^{(n)} = \mathbf{I} + \mathbf{Q} + \mathbf{Q}^2 + \dots + \mathbf{Q}^n \quad (28)$$

$$= \mathbf{I} + \mathbf{Q}(\mathbf{I} + \mathbf{Q} + \mathbf{Q}^2 + \dots + \mathbf{Q}^{n-1}) \quad (29)$$

$$= \mathbf{I} + \mathbf{Q}\mathbf{W}^{(n-1)}. \quad (30)$$

Al escribir la ecuación matricial (30) en términos de los componentes de las matrices, es posible reconocer los resultados del análisis del primer paso. Se tiene entonces que

$$W_{ij}^{(n)} = \delta_{ij} + \sum_{k=0}^{r-1} Q_{ik} W_{kj}^{(n-1)} \quad (31)$$

$$= \delta_{ij} + \sum_{k=0}^{r-1} P_{ik} W_{kj}^{(n-1)}. \quad (32)$$

La ecuación anterior asegura que el número promedio de visitas al estado j durante las primeras n transiciones del proceso que empezó en el estado inicial i incluye la visita inicial si $i = j$ más las futuras visitas durante las restantes $n-1$ transiciones, ponderadas con las correspondientes probabilidades de transición.

Se toma el límite en la ecuación (30) para obtener que $W_{ij} = \lim_{n \rightarrow \infty} W_{ij}^{(n)} = E[\text{Total de visitas a } j \mid X_0 = i]$, $0 \leq i, j < r$. Lo cual corresponde a la ecuación matricial

$$\mathbf{W} = \mathbf{I} + \mathbf{Q} + \mathbf{Q}^2 + \dots, \quad (33)$$

$$\mathbf{W} = \mathbf{I} + \mathbf{Q}\mathbf{W}. \quad (34)$$

Considerando las entradas de las matrices en la ecuación (34) se tiene que

$$W_{ij} = \delta_{ij} + \sum_{l=0}^{r-1} P_{il} W_{lj} \quad \text{para } i, j = 0, 1, \dots, r-1. \quad (35)$$

Al reescribir la ecuación (34) se tiene que

$$\mathbf{W} - \mathbf{Q}\mathbf{W} = (\mathbf{I} - \mathbf{Q})\mathbf{W} = \mathbf{I}. \quad (36)$$

Entonces $\mathbf{W} = (\mathbf{I} - \mathbf{Q})^{-1}$; es decir, la matriz \mathbf{W} es la inversa de la matriz $\mathbf{I} - \mathbf{Q}$. A la matriz \mathbf{W} se le denomina la *matriz fundamental* asociada con \mathbf{Q} .

Sea T el tiempo de absorción, ya que los estados $r, r+1, \dots, N$ son absorbentes, entonces $T = \min\{n \geq 0 : N \geq X_n \geq r\}$. Así, la entrada W_{ij} de la matriz fundamental \mathbf{W} es

$$W_{ij} = E \left[\sum_{n=0}^{T-1} \mathbf{1}\{X_n = j\} \mid X_0 = i \right], \quad \text{para } 0 \leq i, j < r. \quad (37)$$

Sea $v_i = E[T | X_0 = i]$ el tiempo promedio de absorción dado que el proceso inició en el estado i . El tiempo de absorción se compone de estancias en estados transitorios; por lo que

$$\sum_{j=0}^{r-1} \sum_{n=0}^{T-1} \mathbf{1}\{X_n = j\} = \sum_{n=0}^{T-1} \sum_{j=0}^{r-1} \mathbf{1}\{X_n = j\} \quad (38)$$

$$= \sum_{n=0}^{T-1} 1 = T. \quad (39)$$

Combinando el resultado anterior con la ecuación (37), se tiene que

$$\sum_{j=0}^{r-1} W_{ij} = \sum_{j=0}^{r-1} E \left[\sum_{n=0}^{T-1} \mathbf{1}\{X_n = j\} | X_0 = i \right] \quad (40)$$

$$= E \left[\sum_{j=0}^{r-1} \sum_{n=0}^{T-1} \mathbf{1}\{X_n = j\} | X_0 = i \right] \quad (41)$$

$$= E[T | X_0 = i], \text{ para } 0 \leq i, j < r. \quad (42)$$

Súmese la ecuación (35) sobre los estados transitorios j de la siguiente manera

$$\sum_{j=0}^{r-1} W_{ij} = \sum_{j=0}^{r-1} \delta_{ij} + \sum_{j=0}^{r-1} \sum_{k=0}^{r-1} P_{ik} W_{kj} \quad \text{para } i, j = 0, 1, \dots, r-1. \quad (43)$$

Usando el hecho de que $v_i = \sum_{j=0}^{r-1} W_{ij}$, tenemos

$$v_i = 1 + \sum_{k=0}^{r-1} P_{ik} v_k, \quad \text{para } i = 0, 1, \dots, r-1. \quad (44)$$

Consideremos ahora las probabilidades de absorción. Recordéese que los estados $k = r, r+1, \dots, N$ son estados absorbentes. Dado que no se pueden abandonar estos estados, una vez que el proceso ha llegado ahí, la probabilidad de absorción al tiempo n en un estado absorbente particular k , dado que se inició en el estado i es

$$\begin{aligned} P_{ij}^{(n)} &= P[X_n = k | X_0 = i] \\ &= P[T \leq n, X_T = k | X_0 = i], \\ &\text{para } i = 0, 1, \dots, r-1; k = r, \dots, N. \end{aligned} \quad (45)$$

Donde $T = \min\{n \geq 0 : r \leq X_n \leq N\}$ es el tiempo de absorción. Sea

$$\begin{aligned} U_{ik}^{(n)} &= P[T \leq n, X_T = k | X_0 = i] \\ &\text{para } i = 0, 1, \dots, r-1; k = r, \dots, N. \end{aligned} \quad (46)$$

Considerando las ecuaciones (30) y (45) se tiene que la matriz $U^{(n)}$ se puede expresar como

$$\begin{aligned} U^{(n)} &= (\mathbf{I} + \mathbf{Q} + \mathbf{Q}^2 + \dots + \mathbf{Q}^{n-1}) \mathbf{R} \\ &= \mathbf{W}^{(n-1)} \mathbf{R}. \end{aligned} \quad (47)$$

Si se toma el límite cuando n tiende a infinito, las probabilidades de absorción son

$$\begin{aligned} U_{ik} &= \lim_{n \rightarrow \infty} U_{ik}^{(n)} = P[X_T = k | X_0 = i] \\ &\text{para } i = 0, 1, \dots, r-1; k = r, \dots, N. \end{aligned} \quad (48)$$

La ecuación (47) conduce entonces a una expresión para la matriz de probabilidades de absorción \mathbf{U} , en términos de la matriz fundamental \mathbf{W} , como $\mathbf{U} = \mathbf{WR}$.

4. Información estadística sobre la LIMA-UAEH

La Licenciatura en Matemáticas Aplicadas (LIMA-UAEH) es un programa de licenciatura ofertado por la Universidad Autónoma del Estado de Hidalgo desde el semestre julio-diciembre 2001, en su campus de Ciudad del Conocimiento, ubicado en el municipio de Mineral de la Reforma, Hidalgo. Las primeras clases de la LIMA tuvieron lugar el 30 de julio de 2001. El primer rediseño de este programa educativo fue aprobado por el Honorable Consejo Universitario de la UAEH en el mes de mayo de 2010, pero el programa rediseñado inició operaciones hasta al semestre julio-diciembre de 2014.

El programa de la LIMA-UAEH, versión 2010, está integrado por 51 asignaturas distribuidas en 10 semestres. La flexibilidad curricular del programa permite a los alumnos adelantar asignaturas no seriadas, a partir del segundo semestre, para concluir sus estudios antes del tiempo regular o bien hacer uso del tiempo máximo, en función de sus necesidades. Por lo anterior, es posible concluir la licenciatura en un mínimo de ocho semestres y en un máximo de 15 semestres (Avila-Pozos, 2010).

Es importante mencionar que el artículo 23 del Reglamento Escolar de la UAEH indica que el alumnado de nuevo ingreso que adeude como máximo el 50 % de las asignaturas, al concluir el primer periodo escolar, pasará a segundo semestre y las materias adeudadas se deben cursar nuevamente, en el periodo inmediato. En caso de adeudar un porcentaje mayor, el estudiante causará baja definitiva, por lo que no es posible recurrar el primer semestre. En el caso de los alumnos de la LIMA, durante el primer semestre cursan cinco asignaturas, por lo cual, si reprueban tres o más, serán dados de baja definitiva. Por otro lado, en el artículo 103 del mismo reglamento se establece que un alumno que acumule 10 resultados no aprobatorios en su historial académico causará baja definitiva por bajo rendimiento escolar (UAEH, 2019).

A continuación se muestra información básica de la LIMA-UAEH referida a la cantidad de aspirantes a ingresar al programa, por año, así como porcentaje de aceptación. Este programa de licenciatura se oferta anualmente, a diferencia de otros programas de la universidad, que emiten convocatorias para ingreso semestralmente. Lo anterior representa una desventaja para la LIMA, ya que los alumnos que reprueban alguna asignatura deben de esperar un año para volver a cursarla, y puede ser un factor que incrementa las tasas de abandono escolar y de rezago en este programa educativo.

De acuerdo con la Dirección General de Planeación de la UAEH la *retención* se define como “el conjunto de estudiantes de una misma cohorte o generación que permanecen inscritos a través de periodos escolares subsecuentes en un programa educativo”. Por otra parte, la *tasa de retención* es la “proporción

de alumnos que continúan en la escuela en cada grado consecutivo, a partir de la generación original o cohorte, que inició simultáneamente”. A su vez, la *eficiencia terminal* “es la relación cuantitativa entre los alumnos que egresan e ingresan de una cohorte, E.T. = Alumnos que egresaron $\times 100$ / alumnos que ingresaron en su generación”. Al respecto, una *cohorte* es el “conjunto de alumnos que se inscribieron al inicio de un programa educativo determinado” (<https://bit.ly/3QRFzF6>).

Tabla 1: Aspirantes y estudiantes aceptados en la LIMA

Año	Aspirantes	Aceptados	Porcentaje de aceptación
2014	62	57	92
2015	64	59	92
2016	95	86	91
2017	85	78	92
2018	105	75	71
2019	122	60	49
2020	107	60	56
2021	83	70	84
2022	96	86	90
2023	69	62	90

Tabla 2: Indicadores generales de retención del primer al segundo ciclo y eficiencia terminal en licenciatura, UAEH

Año	Retención (%)	Eficiencia terminal (%)
2014	74	40
2015	76	43
2016	73	43
2017	76	40
2018	74	41
2019	74	35
2020	76	38
2021	77	40
2022	73	41
2023	77	39

Tabla 3: Indicadores de retención del primer al segundo ciclo y eficiencia terminal de la LIMA

Año	Retención (%)	Eficiencia terminal (%)
2014	60.0	8.7
2015	46.0	12.5
2016	33.3	12.5
2017	60.5	5.0
2018	68.0	n.d.
2019	61.0	12.2
2020	60.0	14.6
2021	41.0	13.6
2022	55.0	17.0
2023	36.0	10.0

Las tasas de retención del primero al segundo ciclo, así como los porcentajes de eficiencia terminal reportadas en los anuarios estadísticos 2022 y 2023 de la universidad (UAEH, 2022, 2023a) se muestran en las tablas 2 y 3. Los indicadores de retención de primero al segundo ciclo de la LIMA, en general, son menores que los del resto de programas de licenciatura de

la UAEH, aunque no siempre son los más bajos. Por ejemplo, el indicador de retención en el año 2022 para las licenciaturas en antropología social, licenciatura en alimentación sustentable, ingeniería económica y financiera e ingeniería en tecnología del frío, fueron menores que los de la LIMA (UAEH, 2022). En el anuario estadístico 2018 no se reporta el indicador de eficiencia terminal de la LIMA (UAEH, 2018).

5. Estimadores de las probabilidades de transición

Para estimar las probabilidades de transición estacionarias de las cadenas de Markov, se tomó como base el procedimiento indicado en el artículo *Statistical Inference about Markov Chains* (Anderson y Goodman, 1957), en el cual se proponen estimadores útiles cuando se poseen observaciones del fenómeno a modelar en tiempos consecutivos. En este caso, dado que se tiene información sobre las cohortes reales, es posible utilizar los estimadores propuestos.

Denótese con $n_{ij}(t)$ el número de individuos que en el tiempo $t-1$ se encontraban en el estado i y que en el tiempo t pasaron al estado j , donde $t = 0, 1, \dots, T$, entonces los estimadores de las probabilidades de transición para cada tiempo t están dados por

$$\hat{p}_{ij}(t) = \frac{n_{ij}(t)}{\sum_{k=1}^m n_{ik}(t)} \quad (49)$$

Si se está considerando estimar probabilidades de una cadena estacionaria se proponen los siguientes estimadores

$$\hat{p}_{ij} = \frac{\sum_{t=1}^T n_{ij}(t)}{\sum_{k=1}^m \sum_{t=0}^{T-1} n_{ik}(t)} \quad (50)$$

En el caso de este trabajo $T = 15$, ya que es el máximo de semestres en que un estudiante puede concluir el plan de estudios de la LIMA-UAEH (Avila-Pozos, 2010).

5.1. Espacio de estados para el abandono escolar en la LIMA-UAEH

Los estados de un proceso estocástico indican los valores que pueden tomar las variables aleatorias que constituyen la cadena de Markov en cualquier instante del tiempo. En el caso de este artículo, el espacio de estados es el conjunto $E = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, D, F\}$, donde cada uno de los números indica el semestre en el que está inscrito el estudiante, la letra D representa la deserción escolar y la F la finalización de la licenciatura. Estos dos últimos estados tienen la característica de ser estados absorbentes del proceso.

5.2. Disponibilidad de datos

Los datos empíricos que se utilizaron para estimar las probabilidades estacionarias de transición de la cadena de Markov, se obtuvieron a partir de las estadísticas institucionales publicadas en la página oficial de la Dirección General de Planeación de la UAEH, en particular en el apartado sobre trayectorias escolares (<http://sgc.uaeh.edu.mx/planeacion/index.php/estainst?id=54>). En el archivo correspondiente el Instituto de Ciencias Básicas e Ingeniería, se encuentra información sobre las cohortes generacionales, por semestre, que incluyen el número de alumnos inscritos en cada semestre por grupos,

el número total de egresados, la tasa de retención, etc. de todas las licenciaturas que se imparte en el instituto, en particular de la Licenciatura en Matemáticas Aplicadas. La información se procesó utilizando el lenguaje de programación Python, para la estimación de las probabilidades y la implementación del análisis del primer paso.

6. Resultados

Para la estimación de probabilidades de transición en un paso, se consideraron los siguientes supuestos respecto de un alumno:

1. No puede retroceder del semestre en el que se encuentra inscrito.
2. No puede avanzar más de un semestre por periodo escolar.
3. Si abandona el programa, nunca terminará sus estudios.

Con base en el proceso descrito en la sección 5, se obtuvieron dos matrices de transición con el fin de analizar el comportamiento de las generaciones que ingresaron a la licenciatura para el plan educativo 2010 y observar la dinámica de los alumnos que estuvieron inscritos antes y durante de la pandemia debida a la COVID-19, con la finalidad de llevar a cabo un análisis comparativo. Por lo cual, las cohortes generacionales que se tomaron en cuenta son los correspondientes a los años 2014, 2015, 2016, 2017, 2018, 2019.

La primera matriz de transición P_1 (véase figura 1) corresponde a las generaciones que ingresaron en los años 2014, 2015 y 2016, las cuales concluirían la licenciatura, como alumnos regulares, en los años 2019, 2020 y 2021 respectivamente. Se consideró que la pandemia debida a la COVID-19 no afectó de manera significativa a los estudiantes de dichas generaciones, pues algunos ya habían egresado y otros estaban por egresar.

La segunda matriz de transición P_2 (véase figura 2) corresponde a las generaciones que ingresaron en los años 2017, 2018 y 2019, las cuales tienen como fecha teórica de término de sus estudios, los años 2022, 2023 y 2024, respectivamente. Se consideró que estos estudiantes sí fueron afectados por la pandemia.

Ambas matrices de transición (véase figuras 1 y 2), permiten visualizar la dificultad general que tienen los estudiantes de la LIMA-UAEH para avanzar de semestre, así como la probabilidad con la que un estudiante repite un semestre o abandona el programa de estudios. Como primer punto del análisis, es necesario aclarar que la columna del estado 1 solo contiene ceros, pues no es posible repetir el primer semestre dadas las normativas del programa educativo. Si un estudiante de nuevo ingreso de la LIMA reprueba tres o mas de las materias de su carga académica será dado de baja definitiva, a diferencia de los semestres posteriores en los que sí ocurre algo similar, el alumno deberá recurrir el semestre.

Por otra parte, el patrón descendente en las probabilidades en la columna que corresponden a la de deserción, nos indica que durante los primeros seis semestres conforme el alumno avanza, es menos probable que abandone la carrera, sin embargo, en la matriz P_1 de las generaciones que ingresaron en los años 2014, 2015 y 2016 respectivamente, al avanzar a séptimo semestre, la probabilidad de desertar vuelve a aumentar, lo que

podría indicar mayor dificultad de la licenciatura durante ese semestre. De igual manera, para la matriz P_2 de las generaciones que ingresaron en 2017, 2018 y 2019, se puede observar algo similar, pues al pasar a octavo semestre, la probabilidad de desertar aumenta, sin embargo, esto no solo podría deberse al aumento de dificultad de la licenciatura en ese periodo educativo, sino también, a que algunos estudiantes estuvieron aproximadamente dos años en clases en línea por pandemia y cuando regresaron a clases presenciales, algunos volvieron en los semestres más altos, teniendo que volver a acoplarse a dicha modalidad.

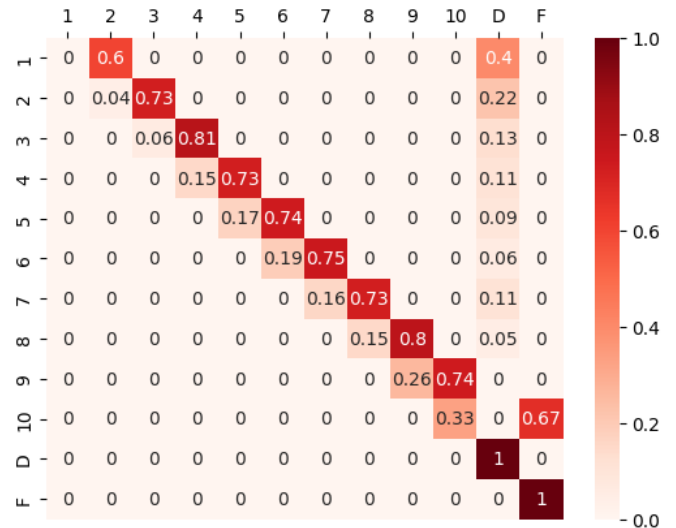


Figura 1: Mapa de calor representando P_1 , las probabilidades de transición se redondearon a dos dígitos decimales con fines de visualización.

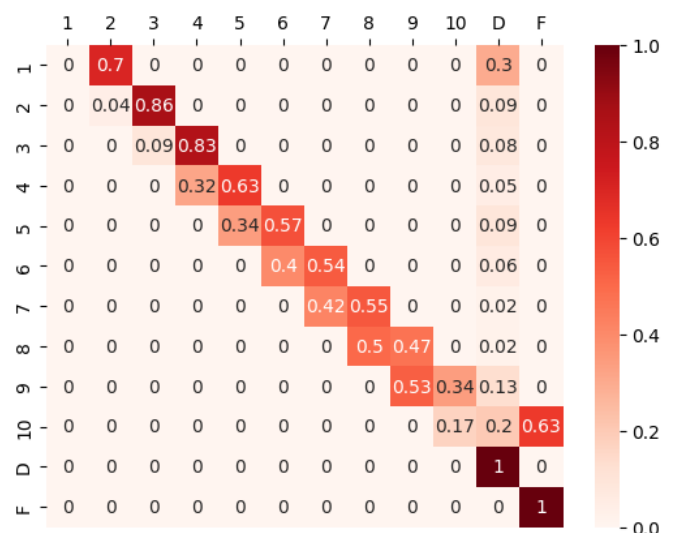


Figura 2: Mapa de calor representando P_2 , las probabilidades de transición se redondearon a dos dígitos decimales con fines de visualización..

Además, es importante notar que en la matriz P_1 dichas probabilidades disminuyen más lento a comparación de las probabilidades encontradas en la matriz de transición P_2 , lo cual

podría deberse a que las clases en línea ayudaron a los estudiantes a no desertar tan rápido.

Se puede observar también que a medida que los estudiantes avanzan de semestre las probabilidades de permanecer en ese semestre y no lograr avanzar al siguiente no disminuyen, especialmente para la matriz P_2 , en donde se observan probabilidades muy altas de permanecer por más tiempo, en los semestres de tercero a noveno. Posteriormente, al implementar el análisis del primer paso, con base en las matrices P_1 y P_2 , se obtuvieron los siguientes resultados respecto los tiempos promedios de absorción (véase tabla 4) y para las probabilidades de deserción (véase tabla 5):

Semestre	Sin pandemia	Con pandemia
1	4.6	7.5
2	6.1	9.3
3	6.5	9.2
4	6.4	8.8
5	6.0	8.0
6	5.4	7.4
7	4.5	6.4
8	3.9	4.8
9	2.9	3.0
10	1.5	1.2

Tabla 4: Tiempos promedio de absorción, en semestres (redondeados a una cifra decimal significativa)

Semestre	El estudiante deserta (sin pandemia)	El estudiante deserta (con pandemia)
1	0.88	0.79
2	0.62	0.70
3	0.50	0.67
4	0.42	0.64
5	0.33	0.61
6	0.25	0.55
7	0.18	0.50
8	0.06	0.47
9	0.00	0.45
10	0.00	0.24

Tabla 5: Probabilidades de absorción (redondeados a dos cifras decimales significativas)

En la tabla 5 se observa que los alumnos de nuevo ingreso tienen mayor probabilidad de desertar, pero a medida que se avanza de semestre, ésta probabilidad de abandono va disminuyendo. Sin embargo, para los estudiantes que ingresaron en los años 2014, 2015 y 2016, respectivamente, se puede notar que en los primeros 5 semestres de la licenciatura la deserción sigue siendo muy alta. En contraste a la probabilidad de finalizar los estudios, en donde ésta aumenta a partir de que los alumnos se encuentran en quinto semestre, más aún, a partir del octavo los estudiantes inscritos en este semestre, tenían una probabilidad muy alta de finalizar la licenciatura y al avanzar a noveno y décimo era seguro que egresarán.

Sin embargo, para los estudiantes que ingresaron en los años 2017, 2018 y 2019 respectivamente, esto cambió drásticamente, pues si se centra la atención en los estudiantes inscritos

a partir de quinto semestre, se puede notar que la probabilidad de finalizar sus estudios no es tan alta como se esperaría, incluso para los que se encuentran en décimo semestre existe una probabilidad positiva de desertar y no concluir la licenciatura.

Además, para los estudiantes que estuvieron inscritos antes de la pandemia derivada de la COVID-19, el tiempo promedio para que estando en primer semestre se pase a uno de los estados absorbentes **F** o **D** es de cuatro semestres y medio (ver tabla 4); esto es, aproximadamente, un alumno de nuevo ingreso tarda en promedio, casi cinco semestres en abandonar el programa educativo, ya que es imposible concluir la licenciatura en ese tiempo. De igual manera, se puede deducir lo mismo para los estudiantes de segundo y tercer semestre, ya que si un estudiante se encuentra en segundo semestre, es imposible terminar la licenciatura en seis semestres más y también esto es poco probable si se encuentra en tercer semestre (ver tabla 5), ya que se deben cubrir requisitos como el servicio social y las prácticas profesionales.

Por otra parte, en la misma tabla 4 podemos notar que para los estudiantes que fueron afectados por la pandemia derivada de la COVID-19 se obtuvieron resultados diferentes, pues el tiempo promedio en que los alumnos de nuevo ingreso tardan en desertar aumentó significativamente, ya que en comparación con los resultados obtenidos sin pandemia, este tiempo aumentó, en promedio, tres semestres más, permaneciendo en la licenciatura durante siete semestres y medio, antes de desertar.

Para complementar el análisis anterior se consideró el ángulo entre las matrices de transición en un paso, consideradas como vectores en \mathbb{R}^{n^2} , como indicador global de similitud de las probabilidades de transición entre las cohortes que cursaron la licenciatura antes de la pandemia, y aquellas que la cursaron durante la pandemia. Un ángulo cercano a cero indica paralelismo entre los vectores y, por tanto, similitud global entre las probabilidades de transición de ambas cohortes. Por otra parte, un ángulo cercano a $\frac{\pi}{2}$ significa ortogonalidad, entre los vectores; lo cual es indicador de una diferencia global máxima entre las probabilidades de transición entre ambas cohortes.

Los vectores de \mathbb{R}^{12^2} considerados para este análisis se obtuvieron concatenando los valores de las matrices P_1 y P_2 , por renglones, formando vectores x_1 y x_2 respectivamente. El ángulo entre x_1 y x_2 está dado por

$$\arccos \left(\frac{\langle x_1, x_2 \rangle}{|x_1| \cdot |x_2|} \right), \quad (51)$$

donde $\langle x_1, x_2 \rangle$ es el producto escalar usual de \mathbb{R}^n , mientras que $|\cdot|$ indica la norma euclidiana. Nótese que el ángulo obtenido por la expresión (51) es invariante respecto al ordenamiento de los valores de la matriz en el vector, mientras que el ordenamiento se mantenga consistente para ambos vectores. Este ángulo entre x_1 y x_2 es de 0.342 radianes (19.6°), indicaría diferencia moderada entre los valores de las probabilidades de transición entre ambas cohortes, y por tanto una afectación de la pandemia sobre la trayectoria escolar de los estudiantes de la LIMA.

7. Conclusiones

Las cadenas de Markov son una herramienta que es útil para modelar la deserción escolar, sin embargo, la falta de fuentes de

información adecuadas, o las dificultades administrativas para acceder a ellas limitan el conocimiento que podemos obtener respecto de este fenómeno. Es importante promover una cultura de información que permita obtener información precisa, la cual pueda ser la base para la toma de decisiones informadas.

Gracias a este modelo, fue posible observar cosas que no es posible notar mediante las estadísticas descriptivas proporcionadas por la universidad, como la permanencia de un estudiante en la licenciatura antes de desertar o egresar, las probabilidades de finalizar el plan educativo a partir de que se encuentra en un cierto semestre, etc. Es importante mencionar que, a diferencia de otros trabajos donde se analiza el abandono o deserción escolar mediante cadenas de Markov absorbentes, se utilizaron cohortes reales y no cohortes aparentes para la estimación de las probabilidades de transición, lo cual da una mayor precisión a los resultados.

Es interesante notar que los resultados de este trabajo coinciden parcialmente con los obtenidos por Bolívar-Cimé et al. (2016), en lo que se refiere a que las probabilidades de deserción van disminuyendo conforme se avanza en la carrera, pero esto solo ocurre para las cohortes pre-pandemia (véase columna D de la matriz P_1), ya que para las cohortes que cursaron la licenciatura durante la pandemia las probabilidades de deserción aumentan considerablemente en los semestres 9 y 10 (véase columna D de la matriz P_2), lo cual corresponde aproximadamente con el periodo de regreso a clases presenciales, que para el caso de la LIMA ocurrió a mediados del mes de marzo de 2022 (<https://www.uaeh.edu.mx/noticias/6895/>).

Un aspecto importante que se debiera de considerar, como política institucional, consiste en desagregar la información por sexo del alumno, ya que esto permitiría alinear y contrastar los resultados sobre aprobación, retención y abandono escolar, con la información que reportan organismos internacionales (OECD, 2022), y así llevar a cabo acciones que permitan vigilar y asegurar la presencia equitativa de las mujeres en las áreas STEM.

Referencias

- Anderson, T. W. y Goodman, L. A. (1957). Statistical inference about markov chains. *The Annals of Mathematical Statistics*, 28:89–110.
- Aparicio, E. (2006). Un estudio sobre factores que obstaculizan la permanencia, logro educativo y eficiencia terminal en las áreas de matemáticas del nivel superior: el caso de la facultad de matemáticas de la universidad autónoma de yucatán. En Martínez, G., editor, *Acta Latinoamericana de Matemática Educativa*, volumen I, capítulo 0, pp. 450–455. Comité Latinoamericano de Matemática Educativa, México D. F.
- Avila-Pozos, O. (2010). Programa educativo de la licenciatura en matemáticas aplicadas.
- Bahi, S., Higgins, D., y Staley, P. (2015). A time hazard analysis of students persistence: A us university undergraduate mathematics major experience. *International Journal of Science and Mathematics Education*, 13:1139–1160.
- Batún, J., Cool, R., y Pantí, H. (2023). Analysis of academic trajectories of higher education students by means of an absorbing markov chain. *Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales*, pp. 1–16.
- Bolívar-Cimé, A., Notario, C., y Pérez, A. (2016). Modelos de markov para la trayectoria académica de estudiantes de la ujat. *Miscelánea Matemática*, 62:29–43.
- Cabrera-González, G., Morales-García, M., Celso-Arellano, P. L., y Coronado, S. (2018). Análisis exploratorio de la trayectoria egreso-ingreso en instituciones de educación superior: un estudio de caso. En Villavicencio, C., editor, *Estudios latinoamericanos de enseñanza y aprendizaje: Compilación de casos*, volumen I, capítulo 1, pp. 167–178. Global Knowledge Academics, Madrid.
- da Costa, F. J., Bispo, M. S., y Pereira, R. C. F. (2018). Dropout and retention of undergraduate students in management: a study at a brazilian federal university. *RAUSP Management Journal*, 53:74–85.
- INEGI (2022). Demografía y sociedad: Características educativas de la población.
- Jordan-Aramburo, A., Rubí-Vázquez, G. E., y Rosa-Navarro, B. E. (2016). Comparación de los índices de deserción, retención, reprobación y aprobación: Licenciatura en matemáticas aplicadas, uabc. *Revista Iberoamericana de Producción Académica y Gestión Educativa*, 3(5):1–15.
- Martino, P. D. y Gregorio, F. (2019). The mathematical crisis in secondary–tertiary transition. *International Journal of Science and Mathematics Education*, 17:825–843.
- Muñoz-Vargas, B. X. (2017). Análisis de la deserción en las licenciaturas de la fcfm-buap mediante modelos de supervivencia.
- OECD (2022). *Education at a Glance 2022: OECD Indicators*. OECD Publishing, Paris.
- Orozco-Rodríguez, C. (2022). Factores que influyen en el abandono escolar de la licenciatura en matemáticas de la universidad de guadalajara. *Revista Mexicana de Investigación Educativa*, 27:259–287.
- Ramón-Santiago, P., García-Martínez, V., Aquino-Zúñiga, S. P., y Silva-Payró, M. P. (2023). *La deserción escolar de estudiantes universitarios*. Iconos Editorial, Villahermosa.
- Rodríguez-Ríos, C. Y. (2012). La deserción escolar de estudiantes universitarios. tesis de maestría no publicada.
- Ronco, S. L. (1994). *Meandering ways: Studying student stopout with survival analysis*. Annual forum of the association for institutional research, New Orleans.
- Ruiz-Ramírez, L. R., García-Vargas, M. L. E., Molina-Ruiz, H. D., y Reyes-Ruiz, G. (2021). Factores que inciden en la deserción escolar. *TEPEXI Boletín Científico de la Escuela Superior Tepeji del Río*, 8:1–5.
- SEP (2020). *Principales cifras del Sistema Educativo Nacional 2019-2020*. Secretaría de Educación Pública, Ciudad de México.
- SEP (2023). La deserción escolar de estudiantes universitarios.
- Sánchez-González, A. y Ruiz, H. D. M. (2019). Pronóstico de la deserción para un programa educativo de ingeniería industrial mediante el uso de mco, en un campus universitario del centro del país. *TEPEXI Boletín Científico de la Escuela Superior Tepeji del Río*, 6:43–47.
- Taylor, H. M. y Karlin, S. (1998). *An introduction to stochastic modeling*. Academic Press, San Diego.
- Tinto, V. (2006). Research and practice of student retention: What next? *Journal of college student retention*, 8:1–19.
- UAEH (2018). *Anuario estadístico 2018*. Universidad Autónoma del Estado de Hidalgo, Pachuca.
- UAEH (2019). *Reglamento escolar*. Universidad Autónoma del Estado de Hidalgo, Pachuca.
- UAEH (2022). *Anuario estadístico 2022*. Universidad Autónoma del Estado de Hidalgo, Pachuca.
- UAEH (2023a). *Anuario estadístico 2023*. Universidad Autónoma del Estado de Hidalgo, Pachuca.
- UAEH (2023b). *Plan de Desarrollo Institucional 2024-2029*. Universidad Autónoma del Estado de Hidalgo, Pachuca.
- Vries, W., León-Arenas, P., Romero-Muñoz, J. F., y Hernández-Saldaña, I. (2011). ¿desertores o decepcionados? distintas causas para abandonar los estudios universitarios. *Revista de la Educación Superior*, 40:19–50.
- WHO (2020). Timeline: Who's covid-19 response.
- Willett, J. B. y Singer, J. D. (1991). From whether to when: new methods for studying student dropout and teacher attrition. *Review of Educational Research*, 61:407–450.
- Zais, M. y Zhang, D. (2016). A markov chain model of military personnel dynamics. *International Journal of Production Research*, 54:1863–1885.