Miscelánea Matemática **62** (2016) 29-43 SMM

Modelos de Markov para la trayectoria académica de estudiantes de la UJAT

Addy Bolívar-Cimé, Carmen Notario, Aroldo Pérez
División Académica de Ciencias Básicas
Universidad Juárez Autónoma de Tabasco
addy.bolivar@gmail.com, carmita1090@hotmail.com, aroldopz2@gmail.com

Resumen

En este artículo, mediante el empleo de la teoría de cadenas de Markov homogéneas a tiempo discreto, específicamente, la teoría respecto a los tiempos y probabilidades de absorción, pronosticamos el tiempo promedio de egreso y el tiempo promedio de retiro para los estudiantes de la licenciatura en matemáticas de la Universidad Juárez Autónoma de Tabasco (UJAT), así como la probabilidad para cada semestre de, eventualmente, desertar o egresar.

1. Introducción

La dinámica asociada a la obtención de un título profesional es un proceso sujeto a múltiples factores, de manera que al iniciar sus estudios, ningún aspirante tiene completa seguridad sobre su futuro desempeño. De esta manera, el rendimiento y evolución académica no están exentos a la variabilidad que presentan los procesos estocásticos, en particular, las cadenas de Markov.

La teoría de cadenas de Markov a tiempo discreto es acreedora de una merecida importancia en la modelación de diversos fenómenos en muy variadas áreas del conocimiento (véase, por ejemplo, [2, 3, 4, 5, 6]). En particular, Ibarra [2] utilizó esta teoría para modelar el progreso en el tiempo de los alumnos de la carrera de ingeniería industrial de la Universidad Nacional Experimental del Táchira (Venezuela). En dicho artículo se formuló un modelo de cadena de Markov homogénea a tiempo discreto que permitió, mediante el uso de la teoría de cadenas de Markov, específicamente, la teoría respecto a los tiempos y probabilidades de absorción (véase por ejemplo [3, cap. III, §4] o [4, cap. I, §1.3]), pronosticar el tiempo promedio de graduación, así como la probabilidad para cada semestre de, eventualmente, desertar o graduarse. Debido a que en el caso estudiado en [2] los porcentajes de deserción son pequeños, puede considerarse, en ese caso, al tiempo medio de absorción como una estimación aceptable para el tiempo esperado de graduación, sin embargo, en la licenciatura en matemáticas de la UJAT, uno de los principales problemas, es precisamente el alto índice de retiro (deserción), por lo que en nuestro caso, la estimación del tiempo medio de absorción no es una estimación adecuada para el tiempo medio de egreso. De esta manera, en nuestro caso, usando los registros de los alumnos entre el 2003 y el 2009 de la licenciatura en matemáticas de la UJAT (como no todos los alumnos que ingresaron del 2010 al 2014 han egresado o han desertado, consideramos que no es adecuado usar esta información incompleta en nuestro estudio) obtenemos un primer modelo (sección 6) que nos permite encontrar las probabilidades de egreso y las de retiro desde cada semestre (intervalo de porcentaje de avance, sección 4), así como el tiempo promedio de absorción. Dividiendo los registros en dos partes (sección 7): una considerando solo los alumnos que lograron egresar y otra considerando solamente a los que desertaron (o fueron dados de baja), se obtienen, respectivamemte, dos modelos de Markov más, mediante el primero de estos dos se obtiene el tiempo medio de egreso y mediante el segundo de ellos, el tiempo promedio de retiro.

Los resultados en este estudio contribuyen considerablemente al conocimiento sobre la trayectoria académica de los estudiantes de la licenciatura en matemáticas de la UJAT y pueden ser fácilmente adaptados a otras carreras profesionales.

2. Conceptos básicos sobre cadenas de Markov homogéneas

Definición 2.1. Sea $X = \{X_n, n \geq 0\}$ una sucesión de variables aleatorias con valores en un conjunto finito o numerable S, llamado el *espacio de estados*.

a) Si para todo $n \geq 0$ y cualesquiera valores $x, y, x_0, \ldots, x_{n-1}$ se cumple

$$Pr[X_{n+1} = y | X_0 = x_0 \ldots, X_{n-1} = x_{n-1}, X_n = x] = Pr[X_{n+1} = y | X_n = x], \quad (1)$$

entonces X se dice ser una *cadena de Markov* o tener la propiedad de Markov.

b) Si además se cumple

$$Pr[X_{n+1} = y|X_n = x] = Pr[X_1 = y|X_0 = x],$$

entonces X se dice ser una *cadena de Markov homogénea.*

De aquí en adelante supondremos siempre que X es una cadena de Markov homogénea.

A la función $p_{xy}^{(n)}$ dada por

$$p_{xy}^{(n)} = Pr[X_n = y|X_0 = x], \quad x, y \in S$$

para $n \geq 1$ y por

$$p_{xy}^{(0)} = \begin{cases} 1, & \text{si } \; x = y, \\ 0, & \text{si } \; x \neq y \end{cases}$$

se le llama *función de transición* (en n pasos) de la cadena. Denotaremos a $p_{xy}^{(1)}$ simplemente por p_{xy}.

A la matriz $P_n = \left[p_{xy}^{(n)}\right]_{x,y\in S}$ se le conoce como la *matriz de transición en* n *pasos* de la cadena de Markov X. A P_1 la denotaremos simplemente por P y le llamaremos la *matriz de transición* de X.

Definición 2.2. Sea $A = [a_{xy}]$ una matriz. La matriz A se dice ser una *matriz estocástica* si

a) $a_{xy} \geq 0$ para cualesquiera x, y,
b) $\sum_y a_{xy} = 1$ para cualquier x.

Si en lugar de b) se tiene

b') $\sum_y a_{xy} \leq 1$ para cualquier x,

se dice que A es una *matriz subestocástica.*

Para la demostración de la proposición siguiente véase, por ejemplo, [5, pp. 40, 41].

Proposición 2.3. *Para todo $n \geq 1$, P_n es una matriz estocástica y $P_n = P^n$.*

De la proposición 2.3 se sigue directamente la ecuación de Chapman-Kolmogorov:

$$P_{n+m} = P_n P_m \quad \text{para } n, m \geq 1.$$

Definición 2.4. Para una cadena de Markov homogénea X con $X_0 = x$, definimos al tiempo medio de pasar del estado x al estado y por

$$m_{xy} = E(T_{xy}),$$

donde $T_{xy} = \text{mín}\{n > 0 : X_n = y|X_0 = x\}$.

Definición 2.5. Sea X una cadena de Markov homogénea con espacio de estados S.

a) Si $Pr[T_{xx} < \infty] = 1$, se dice que x es un estado *recurrente*.
b) Si $Pr[T_{xx} < \infty] < 1$, se dice que x es un estado *transitorio*.

Definición 2.6. Sea $x \in S$ un estado recurrente.

a) Si $m_{xx} < \infty$, se dice que x es un estado *recurrente positivo*.
b) Si $m_{xx} = \infty$, se dice que x es un estado *recurrente nulo*.

Definición 2.7. Sean $x, y \in S$. Si existen $n, m \geq 0$ tales que $p_{xy}^{(n)} > 0$ y $p_{yx}^{(m)} > 0$, se dice que x y y son estados *intercomunicantes*.

Definición 2.8. Sea C un subconjunto de S.

a) Se dice que C es *cerrado* si
$$Pr[X_k \in S \backslash C \text{ para algún } k \geq n | X_n \in C] = 0.$$
b) Se dice que C es *irreducible* si cualesquiera dos estados $x, y \in C$ son intercomunicantes.

Si para un estado x, $C = \{x\}$ es cerrado, se dice que el estado x es *absorbente*; y si el único subconjunto no vacío de S que es cerrado, es S, se dice que la cadena es *irreducible*, es decir, X es irreducible si para cualesquiera $x, y \in S$, existe $n \geq 0$ tal que $p_{xy}^{(n)} > 0$.

Para la demostración de la proposición siguiente véase, por ejemplo, [1, p. 107].

Proposición 2.9. *Sea X una cadena de Markov homogénea con espacio de estados S. Entonces*
$$S = T \cup \left(\bigcup_{j=1}^{N} C_j \right),$$
donde T es el conjunto de estados transitorios en S y cada C_j es un conjunto cerrado e irreducible de estados recurrentes con $C_i \cap C_j = \emptyset$ si $i \neq j$.

3. Tiempos de absorción y probabilidades de absorción

Supondremos aquí que X es una cadena de Markov homogénea con espacio de estados finito; en cuyo caso, es conocido (véase [5, p. 68]) que todo estado recurrente es recurrente positivo. Además, por la proposición 2.9, $S = T \cup \left(\bigcup_{j=1}^{N} C_j \right)$, donde T es el conjunto de estados transitorios y cada C_j es un conjunto cerrado e irreducible de estados recurrentes (positivos) con $C_i \cap C_j = \emptyset$ si $i \neq j$. Debido a que el espacio de estados es finito, existe (véase [5, p. 55]) al menos un conjunto cerrado e irreducible de estados recurrentes positivos; y supondremos que $T \neq \emptyset$. Así, reordenando los estados (si fuera necesario) podemos escribir a la matriz de transición de la cadena de Markov X en la forma bloque

$$P = \begin{bmatrix} P_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & P_2 & 0 & \cdots & 0 & 0 \\ . & . & . & \cdots & . & . \\ . & . & . & \cdots & . & . \\ 0 & 0 & 0 & \cdots & P_N & 0 \\ A_1 & A_2 & A_3 & \cdots & A_N & Q \end{bmatrix},$$

donde P_i denota a la matriz estocástica que corresponde a las transiciones en el subconjunto cerrado e irreducible C_i, $i = 1, \ldots, N$ y la matriz Q corresponde a las transiciones entre los estados transitorios T. La matriz Q es una matriz subestocástica ya que $\sum_{y \in T} q_{xy} \leq 1$ para cada $x \in T$. Finalmente, las matrices A_i, $i = 1, \ldots, N$ corresponden a las transiciones de los estados transitorios a los estados recurrentes. En general, las matrices A_i $(i = 1, \ldots, N)$ no son cuadradas, por lo que no tienen una interpretación como una matriz de transición de un espacio de estados en sí mismo. Es conocido además (véase [3, p. 84]), que en este caso, la cadena no puede permanecer por siempre entre los estados transitorios; así, una cuestión de interés aquí radica en determinar la probabilidad de que la cadena sea absorbida en cada conjunto cerrado e irreducible de estados recurrentes, habiendo iniciado en un estado transitorio. Sabiendo que la cadena será absorbida de los estados transitorios en un conjunto cerrado e irreducible de estados recurrentes, la segunda cuestión de interés radica en determinar el tiempo medio para la absorción. Notemos que esta segunda cuestión es de interés tanto para el caso en que exista un único conjunto cerrado e irreducible de estados recurrentes, como para el caso en que exista más de uno de tales conjuntos; mientras que la primera cuestión es solamente de interés en el caso de dos o más conjuntos cerrados e irreducibles. La razón de esto es que si existe solo un conjunto cerrado e irreducible de estados recurrentes, la cadena irá allí con probabilidad uno.

Analizaremos primero la cuestión que consiste en determinar el tiempo medio para la absorción habiendo iniciado en un estado $x \in T$.

Notemos que

$$Pr[X_n \in T | X_0 = x] = \sum_{y \in T} q_{xy}^{(n)}, \tag{2}$$

donde $q_{xy}^{(n)} = Pr[X_n = y | X_0 = x]$. Así, si μ_x denota al tiempo medio de absorción habiendo iniciado en $x \in T$, tenemos

$$\begin{aligned} \mu_x &= \sum_{n=1}^{\infty} n Pr\left[X_{n-1} \in T, X_n \in \bigcup_{i=1}^{N} C_i \middle| X_0 = x\right] \\ &= \sum_{n=1}^{\infty} \sum_{k=0}^{n-1} Pr\left[X_{n-1} \in T, X_n \in \bigcup_{i=1}^{N} C_i \middle| X_0 = x\right]. \end{aligned}$$

Debido a que todos los términos en esta serie doble son no negativos, se sigue del teorema de Fubini y (2) que

$$\begin{aligned} \mu_x &= \sum_{k=0}^{\infty} \sum_{n=k+1}^{\infty} Pr\left[X_{n-1} \in T, X_n \in \bigcup_{i=1}^{N} C_i \middle| X_0 = x\right] \\ &= \sum_{k=0}^{\infty} Pr\left[\bigcup_{n=k+1}^{\infty} \left\{X_{n-1} \in T, X_n \in \bigcup_{i=1}^{N} C_i\right\} \middle| X_0 = x\right] \\ &= \sum_{k=0}^{\infty} Pr[X_k \in T | X_0 = x] \\ &= \sum_{k=0}^{\infty} \sum_{y \in T} q_{xy}^{(k)}. \end{aligned}$$

Usando de nuevo el teorema de Fubini obtenemos

$$\mu_x = \sum_{y \in T} \sum_{k=0}^{\infty} q_{xy}^{(k)}.$$

(Es conocido, véase, por ejemplo, [3, pp. 94, 95], que la serie $\sum_{k=0}^{\infty} q_{xy}^{(k)}$ es convergente para cualesquiera $x, y \in T$.) Definamos la matriz

$$F = \sum_{k=0}^{\infty} Q^k,$$

donde $Q^0 = I$, la matriz identidad. Entonces la entrada xy-ésima de F es $\sum_{k=0}^{\infty} q_{xy}^{(k)}$ y así, si $1'$ representa a un vector columna con todas sus entradas igual a 1, tenemos que

$$\mu' = F1' \tag{3}$$

es un vector cuya x-ésima entrada es μ_x, el tiempo medio de absorción habiendo iniciado en $x \in T$.

Así, para encontrar μ_x para cada $x \in T$, basta conocer la matriz F, a la cual se le suele llamar *matriz fundamental* y al método para encontrar los tiempos medios de absorción determinada por (3), *método de la matriz fundamental*.

La proposición siguiente, cuya demostración puede consultarse en [3, pp. 93-95], nos da una forma para encontrar F sin necesidad de conocer $q_{xy}^{(k)}$ para todo $x, y \in T$ y todo k $(= 1, 2, \ldots)$.

Proposición 3.1. *La matriz fundamental F es igual a la matriz inversa de $I - Q$, esto es*

$$F = (I - Q)^{-1}.$$

Regresamos ahora a la cuestión de las probabilidades de absorción en los conjuntos cerrados C_i, $i = 1, \ldots, N$. Dados un estado transitorio x y un estado recurrente y, sea

$$r_{xy} = Pr[X_1 = y | X_0 = x],$$

la probabilidad de moverse de x a y en un paso.

La probabilidad de ser absorbido en el estado recurrente y a partir del estado transitorio x es

$$\begin{aligned} \alpha_{xy} &= Pr\left[\bigcup_{n=0}^{\infty} \{X_n \in T, X_{n+1} = y\} \,\middle|\, X_0 = x\right] \\ &= \sum_{n=0}^{\infty} Pr[X_n \in T, X_{n+1} = y | X_0 = x] \\ &= \sum_{n=0}^{\infty} \sum_{z \in T} q_{xz}^{(n)} r_{zy}. \end{aligned}$$

Como todos los términos en esta serie doble son no negativos, se sigue del teorema de Fubini que

$$\alpha_{xy} = \sum_{z \in T} r_{zy} \sum_{n=0}^{\infty} q_{xz}^{(n)}.$$

Notemos que α_{xy} es la entrada xy-ésima de la matriz FA, donde $A = [A_1\ A_2 \ldots, A_N]$ es la matriz que consiste de los elementos en A_1, A_2, …, A_N.

Obtenemos así, que los valores de la matriz fundamental son también suficientes para encontrar las probabilidades de absorción. Por ejemplo, para calcular $Pr\left[\bigcup_{n=0}^{\infty} \{X_n \in T, X_{n+1} \in C_1\} \middle| X_0 = x\right]$, la probabilidad de absorción en C_1, habiendo iniciado en el estado transitorio x, se deben sumar los valores α_{xy} del x-ésimo renglón de FA para los que $y \in C_1$, es decir

$$Pr\left[\bigcup_{n=0}^{\infty} \{X_n \in T, X_{n+1} \in C_1\} \,\middle|\, X_0 = x\right] = \sum_{y \in C_1} \alpha_{xy}.$$

% Avance	8.88	19.43	31.67	43.89	57.50	71.11	84.72	100	Retiro	Egreso
Semestre	1	2	3	4	5	6	7	8	R	E

Cuadro 1. Estados de la cadena de Markov según el porcentaje de avance acumulado

4. Identificación de los estados del proceso

Los estados de un proceso representan los aspectos que describen por completo al proceso en cualquier instante del tiempo. El modelo de Markov en el presente estudio, considera un número finito de estados que representan cada uno de los semestres, correspondientes a la licenciatura en matemáticas de la UJAT, en que puede encontrarse un alumno al inicio de cualquier ciclo académico largo, además de dos estados que simbolizan la condición de retiro (R) o egreso (E).

Cuando un alumno ingresa a la carrera de licenciado en matemáticas, inicia con una carga académica cuantificada en unidades de crédito, esta carga académica inicial es la misma para todos los alumnos de nuevo ingreso (y determina el primer intervalo dado en el cuadro 1). El alumno para egresar debe aprobar una determinada cantidad de unidades de crédito; pero después de cursar el primer ciclo largo, el alumno elige libremente su carga académica cuidando únicamente de cumplir con el mínimo y máximo de créditos permitidos por ciclo. El alumno tiene además, después de finalizar su primer ciclo académico largo en la UJAT, la posibilidad de cursar asignaturas en los llamados ciclos cortos (periodo entre la finalización de un ciclo largo y el inicio del siguiente ciclo largo), debiendo cumplir también aquí con un mínimo y máximo permitidos. Finalmente, el alumno después de contar con el 70 % o más de la cantidad de unidades de crédito que debe obtener para egresar, puede realizar su servicio social durante un ciclo largo. El servicio social tiene también, al igual que todas las demás asignaturas, una cuantificación en unidades de crédito. En este estudio, los créditos acumulados por el alumno en un ciclo corto se suman a los del ciclo largo recientemente cursado.

Con base en las anteriores particularidades de la licenciatura en matemáticas de la UJAT y al análisis de los datos bajo estudio que muestran una preferencia por parte de los alumnos de la licenciatura en matemáticas en cursar asignaturas en ciclos cortos a medida de que avanza su estadía en la carrera, así como el momento en que presentan su servicio social, decidimos dividir el avance porcentual de la manera siguiente:

En el cuadro 1 los intervalos se consideran cerrados por la izquierda y abiertos por la derecha, esto es, un alumno con un avance del 19.43 % se considera en el tercer semestre.

En este trabajo, X_n, $n = 1, 2, \ldots$ denota al semestre que cursa el alumno al inicio del tiempo (ciclo largo) n. De esta manera, el espacio de estados es entonces $S = \{1, 2, 3, 4, 5, 6, 7, 8, R, E\}$.

5. Matriz de probabilidades de transición estimadas

A lo largo de la creación de la licenciatura en matemáticas de la UJAT, el plan de estudios de esta carrera ha sido reestructurado en varias ocasiones. La última reestructuración fue en el 2010 y a la fecha, no todos los alumnos que ingresaron en este año han egresado o se han retirado. Por tal motivo, se decidió considerar solo los registros de los alumnos entre los años 2003 al 2009, los cuales pertenecen al mismo plan (después del 2003, la siguiente reestructuración fue en el 2010). Además, a partir del 2003 es que conseguimos los registros históricos con la información suficiente para poder realizar nuestros cálculos.

La propiedad de Markov (1) indica que la transición de un estado a otro depende solo del estado actual y no de su historia pasada. Esta propiedad se cumple en la dinámica de los estudiantes entre los semestres, ya que la transición de un semestre a otro de un estudiante solo depende del semestre actual en el que se encuentra. Además, debido a que los alumnos en nuestro estudio pertenecen a un mismo plan de estudios (estuvieron bajo las mismas reglas) es razonable suponer que la cadena de Markov que modela esta dinámica es homogénea.

Las probabilidades de transición p_{xy} se estimaron a partir de un conjunto de 327 registros generados por los estudiantes de la licenciatura en matemáticas de la UJAT durante los años 2003 al 2009. Debido a la flexibilidad del plan de estudios, que permite a los estudiantes elegir su carga académica en cada ciclo largo y poder cursar asignaturas en los ciclos cortos, es posible para un alumno pasar de un semestre x a un semestre $x+2$ (no solo al semestre consecutivo $x+1$), pero debido a los límites máximos permitidos, no se presentaron en este estudio saltos de más de dos semestres.

La estimación de las probabilidades de transición p_{xy} se realizó de la manera siguiente: sean $x, y \in \{1, 2, 3, 4, 5, 6, 7, 8\}$, si n_{xy} es el número total de transiciones ocurridas desde el estado x al estado y durante el periodo de observación y n_x es el número total de alumnos ubicados en el semestre x durante el periodo de observación, entonces un estimador con buenas propiedades estadísticas para p_{xy} está dado por la ecuación

$$\widehat{p}_{xy} = \frac{n_{xy}}{n_x}.$$

$$
\widehat{P} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ R \\ E \end{array}
\begin{array}{c}
\begin{array}{cccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & R & E \end{array} \\
\left[\begin{array}{cccccccccc}
0.390 & 0.431 & 0 & 0 & 0 & 0 & 0 & 0 & 0.179 & 0 \\
0 & 0.371 & 0.482 & 0.008 & 0 & 0 & 0 & 0 & 0.139 & 0 \\
0 & 0 & 0.377 & 0.532 & 0.028 & 0 & 0 & 0 & 0.063 & 0 \\
0 & 0 & 0 & 0.335 & 0.618 & 0.004 & 0 & 0 & 0.043 & 0 \\
0 & 0 & 0 & 0 & 0.355 & 0.551 & 0.030 & 0 & 0.064 & 0 \\
0 & 0 & 0 & 0 & 0 & 0.303 & 0.632 & 0.027 & 0.038 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.214 & 0.648 & 0.038 & 0.101 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.285 & 0.020 & 0.695 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}\right]
\end{array}
$$

Cuadro 2. Matriz de probabilidades de transición estimadas.

Así, $\widehat{p}_{xy}$ es la estimación de la probabilidad condicional de que un alumno quien al inicio de un ciclo largo se encuentra cursando el semestre x, en el siguiente ciclo largo se encuentre en el semestre y (aquí $y = x$ representa la situación en la que el alumno no acumula los créditos suficientes para pasar a un semestre superior y es clasificado entonces en el mismo semestre x en el siguiente ciclo largo). De manera similar se obtuvieron $\widehat{p}_{xR}$ y $\widehat{p}_{xE}$ (obviamente $p_{Ry} = 0$ para $y \neq R$ y $p_{RR} = 1$; también $p_{Ey} = 0$ para $y \neq E$ y $p_{EE} = 1$). En el cuadro 2 se presenta la matriz de transición estimada, que es la que contiene a todas las probabilidades de transición estimadas. En la matriz $\widehat{P}$ el valor 0.482 en la segunda fila y tercera columna es la probabilidad de transición estimada $\widehat{p}_{23}$, que indica la probabilidad de que un alumno con un porcentaje de avance que lo ubica en el segundo semestre, transcurrido el lapso del ciclo largo, al inicio del siguiente ciclo largo posee un porcentaje de avance que lo ubica en el semestre 3. Es decir, en la licenciatura en matemáticas de la UJAT, el 48.2 % de los alumnos que inician con un porcentaje de avance en el intervalo $[8.88, 19.43)$ logra en un periodo obtener un porcentaje de avance que lo ubica en el intervalo $[19.43, 31.67)$. El valor 0.139 es la probabilidad de transición estimada $\widehat{p}_{2R}$, que indica que el 13.9 % de los alumnos que tenían un porcentaje de avance en el intervalo $[8.88, 19.43)$ ya no se inscriben (o fueron dados de baja por cuestiones de reglamento) en el siguiente ciclo escolar largo.

De la matriz $\widehat{P}$ se observa que los estados $1, 2, \ldots, 8$ son transitorios y los estados R y E son absorbentes.

6. Probabilidades condicionales de absorción

Sea Q la matriz formada con los valores que corresponden a las probabilidades de transición solo entre los estados transitorios $1, 2, \ldots, 8$. De la proposición 3.1 se tiene que la matriz fundamental F está dada por

$$F = (I - Q)^{-1},$$

$$\widehat{F} = \begin{array}{c} \\ 1\\2\\3\\4\\5\\6\\7\\8 \end{array} \begin{array}{c} \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\ \left[\begin{array}{cccccccc}
1.639 & 1.122 & 0.869 & 0.708 & 0.716 & 0.570 & 0.486 & 0.462 \\
0 & 1.589 & 1.229 & 1.002 & 1.013 & 0.807 & 0.688 & 0.654 \\
0 & 0 & 1.605 & 1.282 & 1.298 & 1.034 & 0.881 & 0.837 \\
0 & 0 & 0 & 1.503 & 1.440 & 1.147 & 0.978 & 0.929 \\
0 & 0 & 0 & 0 & 1.550 & 1.225 & 1.045 & 0.992 \\
0 & 0 & 0 & 0 & 0 & 1.434 & 1.154 & 1.099 \\
0 & 0 & 0 & 0 & 0 & 0 & 1.272 & 1.152 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.398
\end{array}\right] \end{array} \begin{array}{c} \\ 6.572\\6.983\\6.938\\5.997\\4.812\\3.687\\2.424\\1.398 \end{array}$$

Cuadro 3. Matriz del número esperado de ciclos antes de la absorción en $\{R\}$ o $\{E\}$.

donde I es la matriz identidad de idénticas dimensiones que la matriz Q, en este caso 8×8. En el cuadro 3 se presenta la estimación de la matriz F correspondiente a los datos analizados. Los números que aparecen a la derecha de la matriz $\widehat{F}$ (suma de los valores del renglón) son los tiempos esperados de absorción en los estados absorbentes R o E a partir del semestre correspondiente.

A diferencia del artículo [2] donde los porcentajes de Retiro son pequeños, en nuestro caso, son bastante grandes, por lo que estos tiempos medios no representan el tiempo esperado de egreso, sino el tiempo esperado para la absorción, ya sea por $\{R\}$ o por $\{E\}$. Como en [2] las probabilidades de retiro para cada semestre son pequeños, ese tiempo de absorción por $\{R\} \cup \{E\}$ es una estimación aceptable para el tiempo esperado de absorción en $\{E\}$, pero no en nuestro modelo. Por este motivo, en la sección 7 analizaremos por separado los registros de los alumnos que lograron egresar de los que se retiraron, de esta manera, los espacios de estados correspondientes serán, respectivamente $S_R = \{1,2,3,4,5,6,7,8,R\}$ y $S_E = \{1,2,3,4,5,6,7,8,E\}$ y sus respectivas matrices fundamentales obtenidas de sus respectivas matrices de transición representarán los tiempos promedio de retiro y los tiempos promedio de egreso. Sin embargo (véase la exposición entre la proposición 3.1 y el final de la sección 3), los valores en la matriz

$$U = FA,$$

donde A es la matriz de 8×2 de las probabilidades de transición de los estados transitorios a los estados absorbentes R y E, nos dan las probabilidades de retiro para cada semestre (primer columna) y las probabilidades de egreso para cada semestre (segunda columna).

En el cuadro 4 se presenta la estimación de la matriz U de probabilidades de absorción desde cada semestre.

Observando la matriz $\widehat{U}$, se tiene que un alumno que inicia en el primer semestre de la licenciatura en matemáticas de la UJAT tiene una probabilidad de 0.630 de abandonar la carrera a lo largo del tiempo, y de egresar tiene una probabilidad de 0.370. Lo que es lo mismo que

$$\widehat{U} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{array} \begin{array}{c} \begin{array}{cc} R & E \end{array} \\ \begin{bmatrix} 0.630 & 0.370 \\ 0.476 & 0.524 \\ 0.329 & 0.671 \\ 0.256 & 0.744 \\ 0.205 & 0.795 \\ 0.120 & 0.880 \\ 0.071 & 0.929 \\ 0.028 & 0.972 \end{bmatrix} \end{array}$$

Cuadro 4. Matriz de probabilidades de absorción estimadas.

$$\widehat{P}_E = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ E \end{array} \begin{array}{c} \begin{array}{ccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & E \end{array} \\ \begin{bmatrix} 0.224 & 0.776 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.301 & 0.682 & 0.017 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.314 & 0.640 & 0.046 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.284 & 0.710 & 0.006 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.273 & 0.685 & 0.042 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.279 & 0.690 & 0.031 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.172 & 0.714 & 0.114 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.281 & 0.719 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

Cuadro 5. Matriz de probabilidades de transición estimadas de los que egresaron.

sostener que de cada 100 alumnos que inician la carrera se espera que logren egresar solo 37. Sin embargo, notemos que una vez obtenido un avance porcentual ubicado en el intervalo $[57.50, 71.11)$ (sexto semestre) la probabilidad de egreso es de 0.880 y la de retiro es solo de 0.120.

7. Tiempo promedio de egreso y tiempo promedio de retiro

Para este análisis separamos los registros de los alumnos que lograron egresar de los que no lo hicieron: de los 327 alumnos en el análisis 1) 121 lograron egresar y 2) 206 no egresaron.

Para el caso 1), el espacio de estados es $S_E = \{1, 2, 3, 4, 5, 6, 7, 8, E\}$. Utilizando los registros de los 121 alumnos que lograron egresar se obtuvo la estimación mostrada en el cuadro 5 de la matriz de transición P_E entre los estados $1, 2, \ldots, 8, E$.

De la matriz $\widehat{P}_E$ se observa que el único estado recurrente es E y los estados $1, 2, \ldots, 8$ son transitorios.

Ahora, si Q_E es la matriz cuyos valores corresponden a las probabilidades de transición solo entre los estados transitorios $1, 2, \ldots, 8$, se

$$\widehat{F}_E = \begin{array}{cc} & \begin{array}{cccccccc} \;\;1\;\; & \;\;2\;\; & \;\;3\;\; & \;\;4\;\; & \;\;5\;\; & \;\;6\;\; & \;\;7\;\; & \;\;8\;\; \end{array} \\ \begin{array}{c} 1\\2\\3\\4\\5\\6\\7\\8 \end{array} & \left[\begin{array}{cccccccc} 1.289 & 1.431 & 1.422 & 1.305 & 1.365 & 1.307 & 1.159 & 1.207 \\ 0 & 1.431 & 1.422 & 1.305 & 1.365 & 1.307 & 1.159 & 1.207 \\ 0 & 0 & 1.458 & 1.303 & 1.365 & 1.307 & 1.159 & 1.207 \\ 0 & 0 & 0 & 1.397 & 1.364 & 1.308 & 1.159 & 1.207 \\ 0 & 0 & 0 & 0 & 1.376 & 1.307 & 1.159 & 1.207 \\ 0 & 0 & 0 & 0 & 0 & 1.387 & 1.156 & 1.208 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1.208 & 1.199 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.391 \end{array}\right] \begin{array}{c} 10.485\\9.196\\7.799\\6.435\\5.049\\3.751\\2.407\\1.391 \end{array} \end{array}$$

Cuadro 6. Matriz del número esperado de ciclos antes del egreso.

$$\widehat{P}_R = \begin{array}{cc} & \begin{array}{ccccccccc} \;\;1\;\; & \;\;2\;\; & \;\;3\;\; & \;\;4\;\; & \;\;5\;\; & \;\;6\;\; & \;\;7\;\; & \;\;8\;\; & \;\;R\;\; \end{array} \\ \begin{array}{c} 1\\2\\3\\4\\5\\6\\7\\8\\R \end{array} & \left[\begin{array}{ccccccccc} 0.458 & 0.289 & 0 & 0 & 0 & 0 & 0 & 0 & 0.253 \\ 0 & 0.433 & 0.304 & 0 & 0 & 0 & 0 & 0 & 0.263 \\ 0 & 0 & 0.473 & 0.366 & 0 & 0 & 0 & 0 & 0.161 \\ 0 & 0 & 0 & 0.440 & 0.427 & 0 & 0 & 0 & 0.217 \\ 0 & 0 & 0 & 0 & 0.551 & 0.232 & 0 & 0 & 0.217 \\ 0 & 0 & 0 & 0 & 0 & 0.444 & 0.296 & 0 & 0.259 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.526 & 0.158 & 0.316 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.400 & 0.600 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \end{array}$$

Cuadro 7. Matriz de probabilidades de transición estimadas de los que se retiraron.

tiene de la proposición 3.1 que la matriz fundamental F_E está dada por

$$F_E = (I - Q_E)^{-1},$$

donde I es la matriz identidad de idénticas dimensiones que la matriz Q_E, es decir, 8×8. En el cuadro 6 se presenta la estimación de la matriz F_E correspondiente. Los números que aparecen a la derecha de la matriz $\widehat{F}_E$ (suma de los valores del renglón) son los tiempos esperados de egreso a partir del correspondiente semestre.

Obtenemos así, que en la licenciatura en matemáticas de la UJAT, el número esperado de ciclos largos para que un alumno egrese es de 10.485, es decir, aproximadamente 5 años. La suma de los valores en la quinta fila, 5.049, indica que una vez que los alumnos alcanzan un porcentaje de avance ubicado en el intervalo $[43.89, 57.50)$ (quinto semestre), estos egresan, en promedio, en cinco ciclos largos más (dos años y medio).

Para el caso 2), el espacio de estados es $S_R = \{1, 2, 3, 4, 5, 6, 7, 8, R\}$. Utilizando los registros de los 206 alumnos que no lograron egresar se obtuvo la estimación mostrada en el cuadro 7 de la matriz de transición P_R entre los estados $1, 2, \ldots, 8, R$.

De la matriz $\widehat{P}_R$ se observa que el único estado recurrente es R y los estados $1, 2, \ldots, 8$ son transitorios.

$$\widehat{F}_R = \begin{array}{c} \\ 1\\2\\3\\4\\5\\6\\7\\8 \end{array}\!\!\begin{array}{c} \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\ \left[\begin{array}{cccccccc} 1.845 & 0.942 & 0.544 & 0.355 & 0.338 & 0.141 & 0.088 & 0.023 \\ 0 & 1.764 & 1.018 & 0.666 & 0.632 & 0.264 & 0.165 & 0.043 \\ 0 & 0 & 1.898 & 1.241 & 1.178 & 0.492 & 0.308 & 0.081 \\ 0 & 0 & 0 & 1.786 & 1.696 & 0.708 & 0.443 & 0.117 \\ 0 & 0 & 0 & 0 & 2.226 & 0.929 & 0.581 & 0.153 \\ 0 & 0 & 0 & 0 & 0 & 1.800 & 1.126 & 0.296 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2.111 & 0.556 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.667 \end{array}\right] \end{array} \begin{array}{c} \\ 4.275\\4.552\\5.198\\4.749\\3.889\\3.222\\2.667\\1.667 \end{array}$$

Cuadro 8. Matriz del número esperado de ciclos antes del retiro.

Notemos también de la matriz $\widehat{P}_R$ que, de los alumnos que no lograron egresar, el 25.3 % se retira en el primer ciclo escolar largo y de los de este mismo grupo que no egresaron pero que lograron un porcentaje de avance en el intervalo $[8.88, 19.43)$ que los ubica en el segundo semestre, el 26.3 % se retiró durante o al final del ciclo escolar en que alcanzó dicho porcentaje de avance; y así, el penúltimo número de la columna R, 0.600, indica que de los que no lograron egresar pero que lograron un porcentaje de avance en el intervalo $[84.72, 100)$ que los ubica en el octavo semestre, el 60 % se retira en ese ciclo escolar largo y el 40 % se vuelve a inscribir a un ciclo escolar más.

Ahora, si Q_R es la matriz cuyos valores corresponden a las probabilidades de transición solo entre los estados transitorios $1, 2, \ldots, 8$, se tiene de la proposición 3.1 que la matriz fundamental F_R está dada por

$$F_R = (I - Q_R)^{-1},$$

donde I es la matriz identidad de 8×8. En el cuadro 8 se presenta la estimación de la matriz F_R. Los números que aparecen a la derecha de la matriz $\widehat{F}_R$ (suma de los valores del renglón) son los tiempos esperados de retiro a partir del correspondiente semestre.

Obtenemos así, que en la licenciatura en matemáticas de la UJAT, el número esperado de ciclos largos para que un alumno se retire es de 4.275, es decir, aproximadamente 2 años.

8. Conclusiones

1) De la matriz de transición estimada $\widehat{P}$, dada en el cuadro 2, se observa una probabilidad de reprobación alta en todos los semestres de la licenciatura en matemáticas de la UJAT, ligeramente mayor en los tres primeros semestres, es decir, antes de alcanzar el 31.67 % de avance.
2) De la matriz de probabilidades de absorción estimadas $\widehat{U}$, dada en el cuadro 4, se observa que el 63 % de los alumnos que inician

la carrera de licenciado en matemáticas en la UJAT no logran egresar. Notemos también que conforme el semestre aumenta, la probabilidad de retiro disminuye y la de egreso aumenta. Por ejemplo, una vez obtenido un avance porcentual ubicado en el intervalo $[57.50, 71.11)$ (sexto semestre) la probabilidad de egreso es de 0.880 y la de retiro es solo de 0.120.

3) De la matriz $\widehat{F}_E$, dada en el cuadro 6, se observa que de los alumnos que egresan, estos lo hacen en un tiempo promedio de aproximadamente 5 años, aunque el plan está diseñado para que un alumno regular termine su carrera en 4 años e incluso en un tiempo menor debido a la flexibilidad del plan.
4) De la matriz $\widehat{F}_R$, dada en el cuadro 8, se observa que el tiempo medio de retiro de los alumnos es de aproximadamente 2 años. Sin embargo, de los que logran, antes de retirarse, un porcentaje de avance en el intervalo $[19.43, 31.67)$ (tercer semestre) su tiempo medio de retiro a partir de allí es de aproximadamente 2.5 años, es decir, si un alumno no se retiró antes de alcanzar el 19.43 % de avance, este permanecerá, en promedio, a partir de que logre un avance en el intervalo $[19.43, 31.67)$, 2.5 años más de los que ya ha estado, antes de desertar o ser dado de baja.

   Estos resultados no intentan explicar la causalidad del proceso resultante, solo reflejan la forma en que los estudiantes transitan durante su estancia en la licenciatura en matemáticas de la UJAT, permitiendo describir el progreso mediante el uso de la teoría de cadenas de Markov.

Este trabajo se realizó con el apoyo del proyecto UJAT-2014-IB-06 del PFI.

Bibliografía

[1] Z. Brzeźniak y T. Zastawniak, *Basic stochastic processes: a course through exercises*, Springer-Verlag, Great Britain, 2002.

[2] L. F. Ibarra, «Predicciones de Markov aplicadas en el programa de ingeniería industrial de la Universidad Nacional Experimental del Táchira (unet)», *Ingeniería Industrial. Actualidad y Nuevas Tendencias*, vol. 1, núm. 2, 2009, 39–51, año 2.

[3] D. L. Isaacson y R. W. Madsen, *Markov chains theory and applications*, John Wiley&Sons, USA, 1976.

[4] J. R. Norris, *Markov chains*, Cambridge University Press, USA, 1997.

[5] L. Rincón, «Introducción a los procesos estocásticos», 2012, Disponible en `http://www.matematicas.unam.mx/lars`.

[6] D. W. Stroock, *An introduction to Markov processes*, 2.ª ed., Springer-Verlag, Germany, 2014.