Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales.
47(185):1008-1023, octubre-diciembre de 2023. doi: https://doi.org/10.18257/raccefyn.1973

Mathematics

Original article

Analysis of academic trajectories of higher education students by means of an absorbing Markov chain

Análisis de las trayectorias académicas de estudiantes de educación superior utilizando una cadena de Markov absorbente

José Luis Batún*, **Rubén Cool**, **Henry Pantí**

Facultad de Matemáticas, Universidad Autónoma de Yucatán. Anillo Periférico Norte, Colonia Chuburná Hidalgo Inn, Mérida Yucatán, México.

Abstract

The analysis of the academic trajectory is of the utmost importance for Scholar Program administrators, since it allows them to identify areas of opportunity for the Academic Program improvement. In this paper, we analyzed academic trajectory of a group students enrolled at a University Mathematics Program. To that aim, we utilize a stochastic process for modeling the academic trajectory. The model is defined in terms of a progressive Markov chain with two absorbing states. The inferential theory presented in this paper deals with the definition of a random sample for a Markov chain, the construction of the likelihood function and the estimation of the Markov chain parameters. Using these estimates and delta method, confidence intervals are calculated for the mean absorption time, the mean exit time of a state and the absorption probability into a state, these quantities correspond to expected time a student either concludes or drops out of the Program; the expected sojourn time in academic term and the probability a student either concludes or drops out of the Program, respectively.

Keywords: Markov chain model; maximum likelihood estimation; delta method; fundamental matrix; interval estimation; academic trajectory.

Resumen

El análisis de la trayectoria académica es de suma importancia para los administradores de Programas de estudio, ya que les permite identificar áreas de oportunidad para la mejora del Programa Académico. En este trabajo analizamos la trayectoria académica de un grupo de estudiantes matriculados en un Programa Universitario de Matemáticas. Para ello, proponemos un modelo estocástico, el cual se define en términos de una cadena de Markov progresiva con dos estados absorbentes. La teoría inferencial presentada en este artículo aborda la definición de una muestra aleatoria para una cadena de Markov, la construcción de la función de verosimilitud respectiva y la estimación de los parámetros del modelo. Mediante estos estimadores y el método delta, se derivan los intervalos de confianza para el tiempo medio de absorción, el tiempo medio de salida de un estado y la probabilidad de absorción en un estado, estas cantidades corresponden al tiempo esperado en que un estudiante concluye o abandona el Programa, tiempo de permanencia esperado en el semestre y la probabilidad de que un estudiante termine o abandone el Programa, respectivamente.

Palabras clave: Modelo de cadena de Markov; estimación por máxima verosimilitud; método delta; matriz fundamental; estimación por intervalo; trayectoria académica.

Citation: Batún JL, Cool R, Pantí H. 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. 247(185):1008-1023, octubre-diciembre de 2023. doi: https://doi.org/10.18257/raccefyn.1973

Editor: Luis Carlos Arboleda

***Corresponding author:**
Henry Pantí; henry.panti@correo.uady.mx

Received: August 17, 2023
Accepted: October 3, 2023
Published on line: November 1, 2023

Introduction

Stochastic processes usually model random phenomena evolving over time. A stochastic process is a collection of random variables with a certain dependency structure between them. Among the most important stochastic processes are Markov chains, whose dependency structure can be characterized by the Markov property. Roughly speaking, the Markov property says that the future given the present and the past depends only on the present and not on the past at all. This property, also known as the memoryless property, allows to establishing simple results for the calculation of probabilities and other quantities of interest, as well as for their interpretation in the mathematical modeling of random phenomena where Markov chains are used. Communications networks, Alzheimer's disease, DNA sequence analysis, and Education are some examples where Markov chain models have been used in recent decades (see (**Perera *et al*.**, 2019), (**Yu *et al*.**, 2013), (**Zakarczemny & Zajecka,** 2022), (**Muhammad *et al.,*** 2019), (**Wang et al.,** 2021) and references therein).

Absorbing Markov chains are the most known and most interesting examples of Markov chain theory. A Markov chain is absorbing if there is at least a state such that when the Markov chain reaches it, it stays there forever. When this occurs, the Markov chain is said to be absorbed. For this type of Markov chain, we are interested in calculating the absorption probabilities and the mean absorption times. In other words, the absorbing Markov chain theory allows us to answer the questions: What is the probability that the Markov chain will be absorbed in one state? What is the mean time it takes for the Markov chain to be absorbed? The answers to the questions are achieved by calculating the fundamental matrix, which is obtained from the transition probability matrix (see (**Kemeny** & **Snell**, 1976)). Therefore, statistical procedures are necessary to estimate the transition probability matrix for absorbing Markov chain models and thus determine the aforementioned quantities.

In their paper, Anderson and Goodman (**Anderson** & **Goodman**, 1957) establish a treatise on statistical inference for Markov chains. This work includes point estimation for a transition probability matrix, hypothesis tests for Markov homogeneity, asymptotic theory for point estimators, among others. However, explicit results related to absorption probabilities and mean absorption times for absorbing Markov chains have not been established in this paper. On the other hand, although the point estimation of the fundamental matrix was not studied in that paper, with the help of invariance property of maximum likelihood estimators, we can obtain a point estimator for this matrix. Similarly, the theory developed there does not allow the construction of confidence intervals for the absorption probabilities and the mean absorption times, however, with the help of delta method, in this paper we calculate asymptotic confidence intervals for these and other quantities, which are transformations of the fundamental matrix. The asymptotic confidence intervals are used to analyze the academic progress of students in a University Mathematics Program.

The purpose of this paper is to analyze academic trajectories using an absorbing Markov chain. It is important to mention this is not the first time that an analysis of this type has been done. There are works in which school trajectories analysis are carried out considering an absorbing Markov chain ((**Yahaha & Hasan**, 2021) (**Muhammad *et al.,*** 2019), (**Wang *et al.,*** 2021)). In all these works, the estimation of transition probabilities lies in counting the number of transitions from a state to another in a single unit of time. In our study, we consider an observation period and although we obtain an equivalent way to estimate transition probabilities, our initial approach is different and allows us to introduce the concept of a random sample for Markov chains. Furthermore, since point estimation is sometimes not sufficient to obtain conclusions on the true value of the population parameter, confidence interval estimation is preferable, since it provides a range of values that is likely to contain it. For this reason, in our work, as was mentioned before; we include the calculation of confidence intervals, which are used to analyze academic trajectories. Thus, our contribution lies in the application of absorbing Markov chains with confidence intervals to the case study. Finally, with the help of joint distribution of the absorbing time and the state where Markov chain is absorbed (Theorem 1 below), we determine the academic term in which the student is most probably either to dropping out or completing the Program.

The paper is organized as follows. The section Methodology introduces the main elements of Markov chain theory that are used throughout the paper. In this section, we calculate the joint distribution of the absorbing time and the state where the Markov chain is absorbed, statistical tools for absorbing Markov chain are presented and asymptotic confidence intervals are obtained. In the section Results and discussion, the statistical analysis of the academic progress of undergraduate students from a University mathematics Program is given using the theory developed in the previous sections. Finally, in section Conclusions we establish the final conclusions obtained from the statistical analysis of the academic trajectories.

Methodology

In this section, we present the main definitions and notations used in this paper. In the first part, we give a brief introduction to discrete time Markov chains and we present the quantities used in the application section. For a more in-depth study on the topic see (**Durrett,** 2016), (**Norris,** 1998), (**Pinsky & Karlin,** 2011). In the second part we introduce the progressive Markov chains.

Markov chains

Let $S \subset \mathbb{Z}$. We write $X = \{X_n, n \geq 0\}$ to denote the sequence of random variables $X_0, X_1, \ldots$, which are defined on the same probability space $(\Omega, \mathscr{F}, P)$. This sequence is called a stochastic process and the ones that are studied in this paper are the so-called Markov chains, which are defined below.

Definition 1 *A stochastic process $X = \{X_n, n \geq 0\}$ with state space S is called a Markov chain if satisfies the following: for any $n \geq 1$, $i, j, i_0, \ldots, i_{n-1} \in S$,*

$$P(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = P(X_{n+1} = j \mid X_n = i). \tag{1}$$

The property given in (1) is called the Markov property, which is generally interpreted as a conditional memoryless property. If the right-hand side in (1) does not depends of n, the Markov chain is said to be time-homogeneous or simply homogeneous. In this paper we only consider homogeneous Markov chains. Furthermore, although most of the results and definitions presented in this section are valid for Markov chains with a general state space, we will suppose that the state space S is finite and we will write $S = \{1, 2, \ldots, s\}$.

For $m \geq 0$, we define the m-step probability transition as

$$p_{ij}^{(m)} = P(X_{n+m} = j \mid X_n = i) = P(X_m = j \mid X_0 = i),$$

where the second equality follows from homogeneity of the Markov chain. The m-step probability transition is the probability of moving from state i to state j in m steps or m units of time. When $m = 0$, we write $p_{ij}^{(0)} = \delta_{ij}$, the Kronecker delta, which is defined as 1 whenever $i = j$ and 0 otherwise. We also note $p_{ij}^{(1)} = p_{ij}$.

The transition probability matrix $\mathbf{P}$, also called stochastic matrix, is a matrix whose elements are the one-step transition probabilities of the Markov chain $X = \{X_n, n \geq 0\}$. To be precise

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1s} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2s} \\ p_{31} & p_{32} & p_{33} & \cdots & p_{3s} \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ p_{s1} & p_{s2} & p_{s3} & \cdots & p_{ss} \end{pmatrix},$$

where

1. $0 \leq p_{ij} \leq 1$, for all $i, j \in S$.
2. $\sum_{j=1}^{s} p_{ij} = 1$, for all $i \in S$.

Another important element of the Markov chain theory is the initial distribution, which is the probability distribution of the random variable X_0. An important fact is the probabilistic behaviour of a Markov chain is completely determined by its transition probability matrix and its initial distribution. To be precise, for any $n \geq 1$, $i_0, i_1, \ldots, i_n \in S$:

$$P(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = p_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n},$$

where $p_{i_0} = P(X_0 = i)$.

The m-step transition probabilities satisfy the Chapman–Kolmogorov equations, i.e., for any $n, m \geq 0$, $i, j \in S$, it holds

$$p_{ij}^{(n+m)} = \sum_{k=1}^{s} p_{ik}^{(n)} p_{kj}^{(m)}.$$

The Chapman-Kolmogorov equations imply that transition probabilities in n steps can be obtained from the n-th power of the transition probability matrix $\mathbf{P}$, that is, $\mathbf{P}^n = (p_{ij}^{(n)})_{ij \in S}$.

In general, the states of a Markov chain can be classified in different ways. Transient and absorbing states are of special interest in this paper. We will say that the state i is transient if

$$P(X_n = i \text{ for some } n \geq 1 \mid X_0 = i) < 1.$$

In other words, if i is a transient state, starting in i, with probability positive the Markov chain does not return to state i. We will say that state i is absorbing if $p_{ii} = 1$.

Let us introduce an additional notation. Sometimes, we will write P_i to refer quantities calculated using the probability measure $P(\cdot \mid X0 = i)$, for instance, we will write $E_i(X)$ to denote the mathematical expectation of the random variable X calculated with the probability measure $P(\cdot \mid X0 = i)$.

Another amount of interest in this work is the following. Let T_i be given by

$$T_i = \inf\{n \geq 1 : X_n \neq i\},$$

where $\inf\{\emptyset\} = \infty$. If the Markov chain is in state i, then T_i is the first time that Markov chain leaves state i, for this reason we refer it as *exit time from state i*. By the Markov property, under P_i, T_i is a geometric random variable with parameter $1 - p_{ii}$, that is,

$$P_i(T_i = n) = p_{ii}^{n-1}(1 - p_{ii}), \quad n = 1, 2, 3 \ldots$$

Thus, starting in state i, the expected time that the Markov chain leaves state i is given by

$$E_i(T_i) = \frac{1}{1 - p_{ii}}.$$

Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales.
47(185):1008-1023, octubre-diciembre de 2023. doi: https://doi.org/10.18257/raccefyn.1973

Progressive Markov chains

Before introducing progressive Markov chains, we review some facts and results about Markov chains with absorbing states. From now on, we assume that the states $1,\ldots,r$, $r<s$ are transient states and $r+1,\ldots,s$ are absorbing states. Let T be the absorption time for X:

$$T=\inf\{n\geq 0 : X_n > r\}.$$

Since i is a transient state, for $i\leq r$, if the Markov chain starts in the state i, then the Markov chain is eventually absorbed. This means that the absorption time is finite, i.e., $P_i(T<\infty)=1$, for all $i=1,\ldots,r$. For Markov chains with absorbing states, we are particularly interested in the absorption probability in state k, starting in state i:

$$u_{ik}=P(X_T=k \mid X_0=i),$$

for $i=1,\ldots,r$, $k=r+1,\ldots,s$, and the mean absorption time:

$$v_i=E(T\mid X_0=i),$$

for $i=1,\ldots,r$.

The previously defined quantities can be organized in a matrix and a column vector as follows: $\mathbf{U}=(u_{ik})$ and $\mathbf{v}=(v_i)^t$.

The transition probability matrix $\mathbf{P}$ for this kind of Markov chains can be written as

$$\mathbf{P}=\begin{pmatrix}\mathbf{Q} & \mathbf{R}\\ \mathbf{O} & \mathbf{I}\end{pmatrix}, \tag{2}$$

where $\mathbf{O}$ is a zero matrix of dimension $(s-r)\times r$, $\mathbf{I}$ is the $(s-r)\times(s-r)$ identity matrix, $\mathbf{Q}$ and $\mathbf{R}$ are matrices of dimensions $r\times r$ and $r\times(s-r)$, respectively, given by

$$\mathbf{Q}=\begin{pmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1r}\\ p_{21} & p_{22} & p_{23} & \cdots & p_{2r}\\ p_{31} & p_{32} & p_{33} & \cdots & p_{3r}\\ \vdots & \vdots & \vdots & \ldots & \vdots\\ p_{r1} & p_{r2} & p_{r3} & \cdots & p_{rr}\end{pmatrix} \quad\text{and}\quad \mathbf{R}=\begin{pmatrix} p_{1r+1} & p_{1r+2} & \cdots & p_{1s}\\ p_{2r+1} & p_{2r+2} & \cdots & p_{2s}\\ p_{3r+1} & p_{3r+2} & \cdots & p_{3s}\\ \vdots & \vdots & \ldots & \vdots\\ p_{rr+1} & p_{rr+2} & \cdots & p_{ss}\end{pmatrix}.$$

The matrix $\mathbf{M}=(\mathbf{I}-\mathbf{Q})^{-1}=(m_{ij})$ is known as the fundamental matrix (**Kemeny** & **Snell**, 1976). The matrix $\mathbf{M}$ allows us to obtain the mean number of visit to a state before to be absorbed. To be precise, starting in the state i, the mean number of visits to state j before the Markov chain is absorbed is equal to m_{ij}. Two important identities can be established:

$$\mathbf{U}=\mathbf{MR}, \quad \mathbf{v}=\mathbf{M1},$$

where $\mathbf{1}$ is the column vector whose elements are equal to 1's, $\mathbf{U}=(u_{ik})$ and $\mathbf{v}=(v_i)^t$.

In (**Pinsky** & **Karlin**, 2011) is established the identity (3.90):

$$p_{ik}^{(n)}=P(X_T=k, T\leq n\mid X_0=i),$$

for $n\geq 1$, $i=1,\ldots,r$, $k=r+1,\ldots,s$. Hence

$$\begin{aligned}P(X_T=k,T=n\mid X_0=i)&=P(X_T=k,T\leq n\mid X_0=i)-P(X_T=k,T\leq n-1\mid X_0=i)\\ &=p_{ik}^{(n)}-p_{ik}^{(n-1)},\end{aligned}$$

for $n\geq 1$, $i=1,\ldots,r$, $k=r+1,\ldots,s$. Therefore, we have the following theorem on the joint distribution of (X_T,T) for an absorbing Markov chain.

Theorem 1 *Let X be a Markov chain with transition matrix given by (2). Then,*

$$P(X_T = k, T = n \mid X_0 = i) = p_{ik}^{(n)} - p_{ik}^{(n-1)},$$

for $n \geq 1$, $i = 1, \ldots, r$, $k = r+1, \ldots, s$.

From the latter we have that, for $i = 1, \ldots, r$, the absorbing probability in state k, is given by

$$u_{ik} = \lim_{n\to\infty} P(X_T = k, T \leq n \mid X_0 = i) = \lim_{n\to\infty} p_{ik}^{(n)},$$

for $k = r+1, \ldots, s$. In the same way, the distribution of T under P_i is

$$P(T = n \mid X_0 = i) = \sum_{k=r+1}^{s} \left(p_{ik}^{(n)} - p_{ik}^{(n-1)} \right),$$

for $n = 1, 2, 3, \ldots$.

Definition 2 *A Markov chain $X = \{X_n, n \geq 0\}$ with state space $S = \{1, \ldots, s\}$ is called progressive if its transition probability matrix* $\mathbf{P}$ *is an upper triangular matrix.*

For progressive Markov chains only forward transitions are possible. Furthermore, since the state space is finite, it follows that if X is a progressive Markov chain, then $p_{ss} = 1$, i.e., the state s is an absorbing state. In this paper we assume that we have more than one absorbing state, to be precise the progressive Markov chain has $s - r$ absorbing states, with $s - r > 1$ and we refer X as a *progressive Markov chain with $s - r$ absorbing states*. In this case, the transition probability matrix $\mathbf{P}$ is given by

$$\mathbf{P} = \begin{array}{cc} & \begin{array}{cccccccc} 1 & 2 & 3 & \cdots & r & r+1 & \cdots & s \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ \vdots \\ r \\ r+1 \\ \vdots \\ s \end{array} & \begin{pmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1r} & p_{1r+1} & \cdots & p_{2s} \\ 0 & p_{22} & p_{23} & \cdots & p_{2r} & p_{2r+1} & \cdots & p_{2s} \\ 0 & 0 & p_{33} & \cdots & p_{3r} & p_{3r+1} & \cdots & p_{3s} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & p_{rr} & p_{rr+1} & \cdots & p_{rs} \\ 0 & 0 & 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 1 \end{pmatrix} \end{array}.$$

The matrix $\mathbf{Q}$ given in the decomposition (2) is as well an upper triangular matrix. Taking this fact into account, it is straightforward to verify that $m_{ii} = (1 - p_{ii})^{-1}$, for $i = 1, \ldots, r$, which coincides with $E_i(T_i)$.

Point estimation

Let $X = \{X_n, n \geq 0\}$ an absorbing Markov chain with transition probability matrix $\mathbf{P}$. In order to estimate $\mathbf{P}$, it is necessary to observe realizations of the process X, that is, a set of times t and their respective states s_t.

Let h be a positive integer. If $\mathbf{x}$ is a finite set of points obtained observing the Markov chain X from time 0 to h, then $\mathbf{x}$ is called a trajectory or sample path (realization) of length h. In this case, $\mathbf{x} = \{(t, s_t), t = 0, 1, \ldots, h, s_t \in S\}$. A random sample of size m of X is a finite sequence of m independent trajectories of length h: $\mathbf{x}_1, \ldots, \mathbf{x}_m$, where

$$\mathbf{x}_a = \{(0, s_0^{(a)}), (1, s_1^{(a)}), \ldots, (h, s_h^{(a)})\}, \quad a = 1, \ldots, m.$$

Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales.
47(185):1008-1023, octubre-diciembre de 2023. doi: https://doi.org/10.18257/raccefyn.1973

Let $\vec{\mathbf{x}} = (\mathbf{x}_1, \ldots, \mathbf{x}_m)$. Since the trajectories are independent, the likelihood function $L(\mathbf{P};\vec{\mathbf{x}})$ is given by

$$L(\mathbf{P};\vec{\mathbf{x}}) = \prod_{a=1}^{m} P(X_0 = s_0^{(a)}, \ldots, X_h = s_h^{(a)}). \tag{3}$$

Let $n_{ij}^{(a)}$ be the number of transitions from state i to j in the a-th trajectory, since $r+1, \ldots, s$ are absorbing states, $n_{ij}^{(a)} = 0$, for $i = r+1, \ldots, s,\ j \neq i$. Furthermore, $p_{ij} = \delta_{ij}$, for $i = r+1, \ldots, s$. The above remarks and the Markov property imply

$$P(X_0 = s_0^{(a)}, X_1 = s_1^{(a)}, \ldots, X_h = s_h^{(a)}) = \prod_{i=1}^{r}\prod_{j=1}^{s} p_{s_0^{(a)}} (p_{ij})^{n_{ij}^{(a)}}.$$

Hence, equation (3) becomes

$$\begin{aligned} L(\mathbf{P};\vec{\mathbf{x}}) &= \prod_{a=1}^{m}\prod_{i=1}^{r}\prod_{j=1}^{s} p_{s_0^{(a)}} (p_{ij})^{n_{ij}^{(a)}} \\ &= \prod_{a=1}^{m} p_{s_0^{(a)}} \prod_{i=1}^{r}\prod_{j=1}^{s} (p_{ij})^{n_{ij}}, \end{aligned}$$

where $n_{ij} = \sum_{a=1}^{m} n_{ij}^{(a)}$ is the number of transitions from i to j considering all sample trajectories. Finally, if we assume that the initial point in every trajectory $\mathbf{x}_i$, $i = 1, \ldots, m$, is nonrandom, we obtain

$$L(\mathbf{P};\vec{\mathbf{x}}) = \prod_{i=1}^{r}\prod_{j=1}^{s} (p_{ij})^{n_{ij}}.$$

From (2) it follows that in this model, we have $r(s-1)$ parameters to be estimated. Finally, it is not difficult to verify that the MLEs for p_{ij}, denoted by $\widehat{p}_{ij}$ are given by

$$\widehat{p}_{ij} = \frac{n_{ij}}{n_i}, \tag{4}$$

where n_i is the number of transitions from i in all sample trajectories.

In the case of a progressive Markov chain with $s-r$ absorbing states and all trajectories start at state $i = 1$, the likelihood function $L(\mathbf{P};\vec{\mathbf{x}})$ is given by

$$L(\mathbf{P};\vec{\mathbf{x}}) = \prod_{i=1}^{r}\prod_{i \leq j \leq s} (p_{ij})^{n_{ij}}.$$

There are $(2rs - r^2 - r)/2$ parameters to estimate and the MLEs for the p'_{ij}s are the same as those given in equation (4).

In (**Anderson** & **Goodman**, 1957) the likelihood function is calculated for samples from a non-homogeneous Markov chain. They also calculate the MLEs in the homogeneous case and coincide with the estimators given in (4).

Interval estimation

The delta method is a general procedure to derive the variance of a function of asymptotically normal random variables with known variance. Here we present the version given in (**Casella** & **Berger**, 2002). Let $\mathbf{X} = (X_1, \ldots, X_{n'})$ be a vector of random variables such that

$$\sqrt{n}(\mathbf{X} - \Theta) \xrightarrow{d} N_{n'}(0, \Sigma), \quad n \to \infty,$$

where $\Theta = (\theta_1, \ldots, \theta_{n'})$ and Σ is the covariance matrix of $\mathbf{X}$. Then if $g : \mathbb{R}^{n'} \to \mathbb{R}$ is a function with continuous first partial derivatives respect to Θ, then it holds

$$\sqrt{n}(g(\mathbf{X}) - g(\Theta)) \xrightarrow{d} N_1(0, g'(\Theta)^t \Sigma g'(\Theta)),$$

where $g'(\Theta) = \left(\frac{\partial}{\partial\theta_1} g(\Theta), \ldots, \frac{\partial}{\partial\theta_{n'}} g(\Theta)\right)^t$. This result allows us to obtain asymptotic confidence intervals for $g(\Theta)$ in the following way. Suppose that Σ is unknown and let $\widehat{\Theta}$ and $\widehat{\Sigma}$ consistent estimators for Θ and Σ, respectively. Furthermore, suppose that $\widehat{\Theta}$ satisfies

$$\sqrt{n}(\widehat{\Theta} - \Theta) \xrightarrow{d} N_{n'}(0, \Sigma), \quad n \to \infty.$$

Then by continuity of the function g' and the consistency property of $\widehat{\Theta}$ and $\widehat{\Sigma}$, it follows that $\widehat{\sigma}_g^2 = g(\widehat{\Theta})^t \widehat{\Sigma} g'(\widehat{\Theta})$ is a consistent estimator for $\sigma_g^2 = g(\Theta)^t \Sigma g'(\Theta)$. Thus, the Slutsky's theorem implies

$$\frac{\sqrt{n}(g(\widehat{\Theta}) - g(\Theta))}{\widehat{\sigma}_g} = \frac{\sigma_g}{\widehat{\sigma}_g} \frac{\sqrt{n}(g(\widehat{\Theta}) - g(\Theta))}{\sigma_g} \xrightarrow{d} N_1(0, 1).$$

The latter equation establishes that an asymptotic $100(1-\alpha)\%$ confidence interval, for $g(\Theta)$, is given by

$$g(\widehat{\Theta}) \pm z_{\alpha/2} \frac{\widehat{\sigma}_g}{\sqrt{n}},$$

where

$$\widehat{\sigma}_g^2 = g(\widehat{\Theta})^t \widehat{\Sigma} g'(\widehat{\Theta}),$$

with $\widehat{\Theta}$ and $\widehat{\Sigma}$ being consistent estimators for Θ and Σ, respectively. For a further discussion of asymptotic normality under transformations see (**Mendoza**, 1994) and (**Serfling**, 1980).

By choosing suitable functions g we obtain asymptotic intervals for the following quantities:

i) Absorption probabilities,

ii) Mean absorption time,

iii) Mean exit time from a state.

The procedure for obtaining the confidence intervals of the aforementioned quantities is described below. Recall that $\mathbf{M} = (\mathbf{I} - \mathbf{Q})^{-1}$ and we introduce the following: $\mathbf{O}_k = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{R}_k = \mathbf{M}\mathbf{R}_k$, where $\mathbf{R}_k$ is the k-th column of $\mathbf{R}$, for $k = r+1, \ldots, s$. In the same way, for $i = 1, \ldots, r$, $\mathbf{M}_i$ and $\mathbf{q}_i^t$ denote the i-th column and row of $\mathbf{M}$ and $\mathbf{Q}$, respectively. Finally, $\mathbf{I}_{jl}$ is the matrix with a one in the (j, l)-th position and zeros elsewhere.

The procedure to construct the asymptotic intervals for every one of aforementioned quantities is to identify the function g and to calculate the consistent estimator for σ_g, denoted by $\widehat{\sigma}_g$. The estimator $\widehat{\sigma}_g$ is a function of $\widehat{\Theta}$ and $\widehat{\Sigma}$, consistent estimators for Θ and Σ, respectively.

In the examples, the consistency property of $\widehat{\Theta}$ and $\widehat{\Sigma}$ is satisfied since they are function of the elements of estimated transition probability matrix $\widehat{\mathbf{P}}$, which is a consistent estimator for $\mathbf{P}$ (see (**Anderson** & **Goodman**, 1957)).

Absorption probability

In (**Karson** & **Wrobleski**, 1976) an asymptotic confidence interval for the quantity $\mathbf{c}^t(\mathbf{I} - \mathbf{Q})^{-1}\mathbf{R}_k$ is presented, where $\mathbf{c}$ is any suitable vector. The delta method is used to obtain this asymptotic confidence interval, although it is not explicitly mentioned in this work. From

Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales.
47(185):1008-1023, octubre-diciembre de 2023. doi: https://doi.org/10.18257/raccefyn.1973

the result established there and using $\mathbf{c} = \mathbf{e}_i$, a unit vector with 1 in the component i, we can obtain the asymptotic confidence interval for the absorption probability in state k when starting in state i

$$u_{ik} = g(\Theta) = \mathbf{e}_i \mathbf{O}_k = o_{ik}, \tag{5}$$

for $i = 1, \ldots, r$, $k = r+1, \ldots, s$, where

$$\Theta = \left(p_{1k}, \mathbf{q}_1^t, p_{2k}, \mathbf{q}_2^t, \ldots, p_{rk}, \mathbf{q}_r^t\right).$$

The asymptotic $(1-\alpha)$ confidence interval is given by

$$\widehat{o}_{ik} \pm \frac{z_{\alpha/2}}{\sqrt{n}} \sqrt{\sum_{j=1}^{r} \widehat{\mathbf{a}}_j^t \widehat{\Sigma}_j \widehat{\mathbf{a}}_j}, \tag{6}$$

where $\widehat{\mathbf{a}}_j = (\widehat{m}_{ij}, \widehat{m}_{ij}\widehat{o}_{jk}, \ldots, \widehat{m}_{ij}\widehat{o}_{jk})^t$, $n = n_1 + \cdots + n_r$ and

$$\widehat{\Sigma}_j = \begin{pmatrix} \widehat{p}_{jk}(1-\widehat{p}_{jk}) & -\widehat{p}_{jk}\widehat{p}_{j1} & -\widehat{p}_{jk}\widehat{p}_{j2} & \cdots & -\widehat{p}_{jk}\widehat{p}_{jr} \\ -\widehat{p}_{jk}\widehat{p}_{j1} & \widehat{p}_{j1}(1-\widehat{p}_{j1}) & -\widehat{p}_{j1}\widehat{p}_{j2} & \cdots & -\widehat{p}_{j1}\widehat{p}_{jr} \\ -\widehat{p}_{jk}\widehat{p}_{j2} & -\widehat{p}_{j1}\widehat{p}_{j2} & \widehat{p}_{j2}(1-\widehat{p}_{j2}) & \cdots & -\widehat{p}_{j2}\widehat{p}_{jr} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ -\widehat{p}_{jk}\widehat{p}_{jr} & -\widehat{p}_{j1}\widehat{p}_{jr} & -\widehat{p}_{j2}\widehat{p}_{jr} & \cdots & \widehat{p}_{jr}(1-\widehat{p}_{jr}) \end{pmatrix}.$$

Remark 1 *When the actual absorption probability is very small or is almost one, the previous confidence interval could capture negative values or values greater than one respectively. In this case, using the logit transformation and the delta method again, an alternative asymptotic* $100(1-\alpha)\%$ *confidence interval is given by:*

$$\left(\frac{\widehat{o}_{ik}\exp\{-A\}}{1-\widehat{o}_{ik}+\widehat{o}_{ik}\exp\{-A\}}, \frac{\widehat{o}_{ik}\exp\{A\}}{1-\widehat{o}_{ik}+\widehat{o}_{ik}\exp\{A\}}\right),$$

where

$$A = \frac{z_{\alpha/2}}{\sqrt{n}} \sqrt{\frac{\widehat{V}(\widehat{o}_{ik})}{(\widehat{o}_{ik}(1-\widehat{o}_{ik}))^2}} \quad \text{and} \quad \widehat{V}(\widehat{o}_{ik}) = \sum_{j=1}^{r} \widehat{\mathbf{a}}_j^t \widehat{\Sigma}_j \widehat{\mathbf{a}}_j.$$

Remark 2 *Unlike absorption probabilities, to the best of our knowledge, no confidence intervals have been calculated for mean absorption times and mean exit times. This is carried out below.*

Mean absorption time

By following some ideas given in (**Karson** & **Wrobleski**, 1976), based on the delta method, we obtain an asymptotic confidence interval for the mean absorption time v_i:

$$v_i = g(\Theta) = \mathbf{e}_i(\mathbf{I}-\mathbf{Q})^{-1}\mathbf{1} = \mathbf{e}_i\mathbf{M}\mathbf{1} = \sum_{j=1}^{r} m_{ij}, \tag{7}$$

for $i = 1, \ldots, r$, where $\Theta = \left(\mathbf{q}_1^t, \mathbf{q}_2^t, \ldots, \mathbf{q}_r^t\right)$. In this case, we have

$$\Sigma = \begin{pmatrix} \Sigma_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \Sigma_2 & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \Sigma_r \end{pmatrix},$$

where

$$\Sigma_j = \begin{pmatrix} p_{j1}(1-p_{j1}) & -p_{j1}p_{j2} & \cdots & -p_{j1}p_{jr} \\ -p_{j1}p_{j2} & p_{j2}(1-p_{j2}) & \cdots & -p_{j2}p_{jr} \\ \vdots & \vdots & \cdots & \vdots \\ -p_{j1}p_{jr} & -p_{j2}p_{jr} & \cdots & p_{jr}(1-p_{jr}) \end{pmatrix}.$$

Since

$$\begin{aligned}\frac{\partial g(\Theta)}{\partial p_{jl}} &= \mathbf{e}_i(\mathbf{I}-\mathbf{Q})^{-1}\mathbf{I}_{jl}(\mathbf{I}-\mathbf{Q})^{-1}\mathbf{1}\\ &= \mathbf{e}_i\mathbf{M}\mathbf{e}_j\mathbf{e}_j^t\mathbf{M}\mathbf{1}\\ &= m_{ij}\mathbf{e}_j^t\mathbf{M}\mathbf{1}\\ &= m_{ij}\sum_{u=1}^{r} m_{lu},\end{aligned}$$

then, by making

$$\begin{aligned}\mathbf{b}_j &= (m_{ij}\sum_{u=1}^{r} m_{1u}, m_{ij}\sum_{u=1}^{r} m_{2u}, \ldots, m_{ij}\sum_{u=1}^{r} m_{ru})^t\\ &= m_{ij}\sum_{u=1}^{r}\mathbf{M}_u,\end{aligned}$$

it follows

$$\sigma_g^2 = \sum_{j=1}^{r}\mathbf{b}_j^t\Sigma_j\mathbf{b}_j. \tag{8}$$

Hence, an asymptotic $100(1-\alpha)\%$ confidence intervals for v_i is given by

$$\sum_{j=1}^{r}\widehat{m}_{ij} \pm \frac{z_{\alpha/2}}{\sqrt{n}}\sqrt{\sum_{j=1}^{r}\widehat{\mathbf{b}}_j^t\widehat{\Sigma}_j\widehat{\mathbf{b}}_j}. \tag{9}$$

Mean exit time

Finally, we have the mean exit time from state i is given by

$$E_i(T_i) = g(\Theta) = \frac{1}{1-p_{ii}},$$

where $\Theta = p_{ii}$, for $i = 1,\ldots,r$. Since $\sqrt{n_i}(\widehat{p}_{ii} - p_{ii})$ is asymptotically normally distributed with mean 0 and variance $p_{ii}(1-p_{ii})$ (see (**Anderson** & **Goodman**, 1957)), then $\sqrt{n_i}(g(\widehat{p}_{ii}) - g(p_{ii}))$ is asymptotically normally distributed with mean 0 and variance $\sigma_g^2 = g'(p_{ii})^2 p_{ii}(1-p_{ii})$. Hence, we can obtain an asymptotic $100(1-\alpha)\%$ confidence interval for $E_i(T_i)$, namely,

$$\frac{1}{1-\widehat{p}_{ii}} \pm \frac{z_{\alpha/2}}{\sqrt{n_i}}\sqrt{\frac{\widehat{p}_{ii}}{(1-\widehat{p}_{ii})^3}}. \tag{10}$$

Note that the asymptotic confidence intervals previously calculated were carried out considering the number of transitions in each state (in the first two considering all transitions from transient states) and not using the sample size. This implies that we could have a single path of length h, h large enough, and observe a large number of transitions n_i, to establish the asymptotic results. In summary, the important thing is to have a sufficient number of Markov chain transitions to obtain asymptotic confidence intervals through the delta method.

Results and discussion

Through a progressive Markov Chain with two absorbing states, we analyze the academic trajectories of a group of students from a Mathematics Academic Program in the Autonomous University of Yucatan. The Academic Program allows the students to choose

Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales.
47(185):1008-1023, octubre-diciembre de 2023. doi: https://doi.org/10.18257/raccefyn.1973

the subjects to take in each enrollment and there is no restriction for this, so there are no requirements between subjects that prevent the student's academic progress. The transition periods in the proposed discrete-time stochastic process model represent the academic terms in which students are enrolled. The monitoring of the student's academic trajectory is carried out by means of the number of accumulated credits and based on this number a semester is assigned to the student. Thus, in each academic term the student registers to a group of courses, at the end of the period, based on the courses they pass, the number of credits accumulated do date is calculated, based on this number a status is assigned (semester) to the period. Since the student's progress depends on the number of accumulated credits at the end of the academic term and the number of credits of a period are at most related to those of the previous period, we can assume that the stochastic process modeling student's academic trajectory satisfies the Markov property. In addition, the students are under the same characteristics of the Academic Program. Therefore, it is reasonable to assume that the Markov chain that describes the academic trajectory is homogeneous.

From the above, we propose ten states for the Markov chain model: states from 1 to 8 represent the semester of the Program in which students are enrolled. The state 9 represents the graduation status, namely, the student successfully has completed the Program, while state 10 represents the dropout status, this means, the student has withdrawn from the Program voluntarily or by University regulation. It is worth emphasizing the difference between academic term and semester; the former refers to a transition period and the latter to a state of the process.

In developing the model, the following assumptions are made: the student who is currently enrolled in a semester of the Program could, in the next academic term, either advance to a higher semester or repeat semester and stay at the same state, this means that the Markov chain is progressive. Furthermore, considering the Program, the student who has successfully completed it or has voluntarily withdrawn, cannot apply for the same Program. This fact ensures that states 9 and 10 are absorbing states. Therefore, the stochastic process describing student's academic trajectory is a progressive Markov chain with two absorbing states.

The data collected consists of the academic trajectories of 73 students in 16 academic terms (8 years). Considering the above assumptions and the formula (4), we obtain the estimated transition probability matrix $\widehat{\mathbf{P}}$ of the progressive Markov chain, which is

$$\widehat{\mathbf{P}} = \begin{array}{c} \\ 1\\ 2\\ 3\\ 4\\ 5\\ 6\\ 7\\ 8\\ 9\\ 10 \end{array} \begin{array}{c} \begin{array}{cccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \end{array} \\ \begin{pmatrix} 0.291 & 0.544 & 0.010 & 0 & 0 & 0 & 0 & 0 & 0 & 0.155 \\ 0 & 0.176 & 0.471 & 0.265 & 0 & 0 & 0 & 0 & 0 & 0.088 \\ 0 & 0 & 0.057 & 0.829 & 0.029 & 0 & 0 & 0 & 0 & 0.086 \\ 0 & 0 & 0 & 0.277 & 0.692 & 0.015 & 0 & 0 & 0 & 0.015 \\ 0 & 0 & 0 & 0 & 0.281 & 0.656 & 0.047 & 0 & 0 & 0.016 \\ 0 & 0 & 0 & 0 & 0 & 0.218 & 0.764 & 0 & 0 & 0.018 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.438 & 0.525 & 0 & 0.038 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.656 & 0.311 & 0.033 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.000 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.000 \end{pmatrix} \end{array}.$$

The off-diagonal non-zero elements of matrix $\widehat{\mathbf{P}}$ correspond to probability a student advances from a specific semester to another during an academic term, while elements on the main diagonal are the probabilities a student remains in the same semester, except for $\widehat{p}_{99}$ and $\widehat{p}_{1010}$ since states 9 and 10 are not semesters but absorbing states representing graduate and dropout status, respectively. For instance, if we consider a randomly selected student who is currently in the first state, namely, in the semester 1, the probability the student advances in the next academic term to semester 2 is given by $\widehat{p}_{12} = 0.544$. For the non-absorbing states probabilities in the main diagonal of $\widehat{\mathbf{P}}$, we have that the student does not accumulate enough credits to advance to the next semester, this means that student remains in the same semester. For example, $\widehat{p}_{44} = 0.277$ is the probability of remaining in the fourth semester.

Decomposing the matrix $\widehat{\mathbf{P}}$ as in (2), we obtain the matrices $\widehat{\mathbf{Q}}$ and $\widehat{\mathbf{R}}$, namely,

$$\widehat{\mathbf{Q}} = \begin{pmatrix} 0.291 & 0.544 & 0.010 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.176 & 0.471 & 0.265 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.057 & 0.829 & 0.029 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.277 & 0.692 & 0.015 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.281 & 0.656 & 0.047 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.218 & 0.764 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.438 & 0.525 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.656 \end{pmatrix}$$

and

$$\widehat{\mathbf{R}} = \begin{pmatrix} 0 & 0.155 \\ 0 & 0.088 \\ 0 & 0.086 \\ 0 & 0.015 \\ 0 & 0.016 \\ 0 & 0.018 \\ 0 & 0.038 \\ 0.311 & 0.033 \end{pmatrix}.$$

From the first column of $\widehat{\mathbf{R}}$, it follows that for the first seven semesters, the probability that an enrolled student successfully completes the Program is 0, while the probability that an enrolled student will drop out the Program decreases as the semester increases.

To obtain point and interval estimations for the quantities described in Section , we first estimate the fundamental matrix $\widehat{\mathbf{M}}$. Recall $\mathbf{M} = (I - \mathbf{Q})^{-1}$. Thus, we have

$$\widehat{\mathbf{M}} = \begin{pmatrix} 1.41 & 0.931 & 0.480 & 0.892 & 0.878 & 0.753 & 1.097 & 1.675 \\ 0 & 1.214 & 0.606 & 1.140 & 1.121 & 0.963 & 1.402 & 2.140 \\ 0 & 0 & 1.060 & 1.216 & 1.213 & 1.041 & 1.516 & 2.314 \\ 0 & 0 & 0 & 1.383 & 1.331 & 1.143 & 1.665 & 2.542 \\ 0 & 0 & 0 & 0 & 1.391 & 1.167 & 1.702 & 2.598 \\ 0 & 0 & 0 & 0 & 0 & 1.279 & 1.738 & 2.653 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1.779 & 2.716 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2.907 \end{pmatrix}.$$

For progressive Markov chain starting in a transient state, only transitions to the same state or forward are allowed. So that, the elements of fundamental matrix $\widehat{\mathbf{M}}$ have the following interpretation: being in semester i, the value $\widehat{m}_{ij}$ is the estimated expected number of academic terms that a student remains in semester j before completing or ropping out the Program.

Table 1 shows the estimated expected academic terms until to a student completes or drops out the Program. The same table shows the 95% confidence intervals for the quantity aforementioned, both are calculated using the formulas given in equations (7) and (9), with $n_1 = 103$, $n_2 = 68$, $n_3 = 35$, $n_4 = 65$, $n_5 = 64$, $n_6 = 55$, $n_7 = 80$, $n_8 = 122$ and $n = n_1 + \cdots + n_8 = 592$.

47(185):1008-1023, octubre-diciembre de 2023. doi: https://doi.org/10.18257/raccefyn.1973

Table 1. Estimated expected numbers of academic terms until a student completes or drops out the Program

Semester	Expected number of academic terms	Confidence Interval (95%)
1	8.11	$[7.65, 8.58]$
2	8.58	$[8.19, 8.98]$
3	8.36	$[7.98, 8.74]$
4	8.06	$[7.70, 8.42]$
5	6.85	$[6.51, 7.21]$
6	5.67	$[5.33, 6.01]$
7	4.49	$[4.16, 4.83]$
8	2.90	$[2.58, 3.23]$

The value 2.90 in row eight of Table 1 is the estimated expected academic term until a student completes or drops out the Program when is in the eighth semester. Further, with a confidence level of 95%, the real expected academic term value is between 1.99 and 3.83. We can observe that the width of confidence intervals decreases as the semester increases. The latter behavior is because of the estimated variance given in (8) decreases, that is, for any semester i, the vector $\widehat{\mathbf{b}}_j$ is null for all semester $j < i$, hence $\widehat{\mathbf{b}}_j^t \widehat{\Sigma}_j \widehat{\mathbf{b}}_j = 0$, for all $j < i$.

Using equation (5), we can estimate the probability that Markov chain is absorbed in a specific absorbing state when the student is in state i. The Tables 2 and 3 show the estimated absorption probabilities in graduate and dropout status, respectively.

Table 2. Estimated probabilities of graduate status when the student is currently in the i-th semester

Semester	Probability of graduate status	Confidence Interval (95%)
1	0.52	$[0.48, 0.56]$
2	0.66	$[0.62, 0.71]$
3	0.72	$[0.68, 0.76]$
4	0.79	$[0.75, 0.84]$
5	0.80	$[0.76, 0.85]$
6	0.82	$[0.78, 0.87]$
7	0.84	$[0.80, 0.89]$
8	0.90	$[0.86, 0.94]$

Table 3. Estimated probabilities of dropout status when the student is currently in the i-th semester

.

Semester	Probability of dropout status	Confidence Interval (95%)
1	0.48	$[0.44, 0.52]$
2	0.34	$[0.29, 0.38]$
3	0.28	$[0.24, 0.33]$
4	0.21	$[0.16, 0.26]$
5	0.20	$[0.15, 0.24]$
6	0.18	$[0.13, 0.22]$
7	0.16	$[0.11, 0.20]$
8	0.10	$[0.06, 0.14]$

From the Table 2, we infer that in semester eight a student has a high probability 0.9 to achieve graduate status. We can also observe the probability a student completes the Program increases as the semester increases. On the other hand, the Table 3 shows that in semester eight a student has a low probability of dropping out of the Program.

From the Scholar Program Administration point of view an important quantity to estimate is the expected time that a student sojourns in the j-th semester before moving to some upper semester, and also to obtain a 95% confidence interval for this quantity. These values corresponding to mean exit times. Therefore, we use (10) to estimate them. In this case also we have $n_1 = 103$, $n_2 = 68$, $n_3 = 35$, $n_4 = 65$, $n_5 = 64$, $n_6 = 55$, $n_7 = 80$, $n_8 = 122$.

Table 4. Estimated expected sojourn times in the semester i

Semester	Expected sojourn time	Confidence Interval (95%)
1	1.41	[1.24, 1.58]
2	1.21	[1.08, 1.35]
3	1.06	[0.97, 1.15]
4	1.38	[1.18, 1.59]
5	1.39	[1.18, 1.60]
6	1.28	[1.10, 1.46]
7	1.78	[1.44, 2.12]
8	2.91	[2.19, 3.62]

From Table 4, we observe a student in semester seven has an expected sojourn time in that semester of 1.78 academic terms, this means that a student takes almost two periods to move from semester seven and with a 95% confidence level, the real expected sojourn time value is between 1.44 and 2.12 academic terms.

To end this section, by Theorem 1, we estimate $P(X_T = 9, T = n|X_0 = 1)$ for $n = 1, 2, \ldots, 16$, namely, we estimate the probability of dropping out of the Program exactly in the n-th academic term. We organize these probabilities through a graph and determine the academic term in which the student has the highest probability to drop out of the Program. In Figure 1 we can observe that in the first academic term there is a higher probability of dropping out of Program, the probability is 0.15.

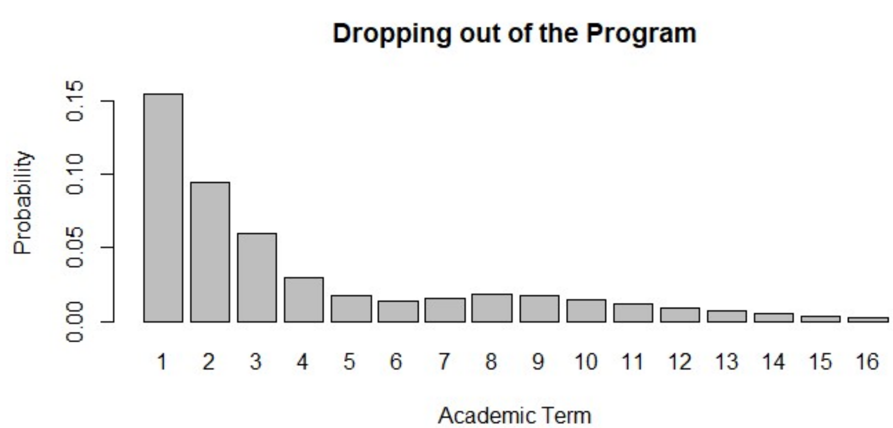

Figure 1. Probability of dropping out of the Program

On other hand, we also estimate probability of completing the Program exactly in n-th academic term: $P(X_T = 9, T = n|X_0 = 1)$ for $n = 1, 2, \ldots, 16$. We organize these probabilities through a graphic and we determine the academic term where it is most probable that the

Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales.
47(185):1008-1023, octubre-diciembre de 2023. doi: https://doi.org/10.18257/raccefyn.1973

student will complete the Program. In Figure 2 we can observe that in the tenth academic period there is higher probability of completing the Program, the probability is 0.077.

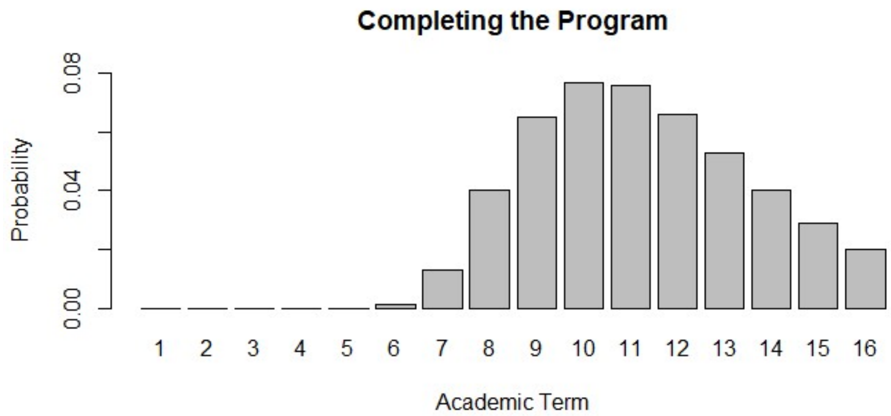

Figure 2. Probability of completing the Program

Conclusions

The statistical analysis of the academic trajectories provides useful and valuable information to identify weaknesses and improvement opportunities of the Program. For example, from our analysis and with the help of the administrators, we should be able to answer why the expected sojourn times per semester of a student for the seventh and eighth semesters are longer than previous ones.

The results obtained indicate that in the eighth semester, the probability of staying in a semester is higher compared to other semesters, this fact implies that in the eighth semester there is a greater expected sojourn time than in other semesters and its respective confidence interval is wider compared with the confidence interval of other semesters. The latter means that in the eighth semester students have difficulty accumulating credits to graduate. In the eighth semester there is an estimated expected sojourn time of 2.9 academic terms, and with a confidence level of 95%, the mean number of academic terms which a student has to accumulate enough credits to be able to graduate from the career is between 2.19 and 3.62 academic terms, which means, approximately between 1 and 1.5 years. Instead, the estimated probability of staying in third semester is lower compared to other semesters, this fact implies that in the third semester there is an expected sojourn time shorter than in other semesters and its respective confidence interval is shorter compared with the confidence interval of other semesters. The latter means, in the third semester, the average of academic terms that a student takes to accumulate the credits to advance is between 0.97 and 1.15 academic terms (with a 95% confidence level), that is, on average an academic term is maintained so that after that time the student advances to another semester.

From Figures 1 and 2, we conclude that the first two academic terms, which are related to the first semesters, there is a higher probability that a student will drop out of the Program, while the tenth and eleventh academic terms have a higher percentage that a student will complete the Program (between 5 and 5.5 years to graduate).

Finally, the Academic Program is designed in such a way that the credits assigned to subjects offered for the first semesters are greater than credits assigned to subjects offered for more advanced semesters. From this we deduce that in the first semesters there is the opportunity to accumulate enough credits that allow the student to advance two semesters (two states) in an academic period, as opposed in the last semesters, for example, a seventh-semester student cannot finish the program in that semester. This is observed in the estimated transition probabilities given in the transition matrix $\widehat{\mathbf{P}}$, for example, $\widehat{p}_{13} = 0.010$, $\widehat{p}_{24} = 0.265$, $\widehat{p}_{35} = 0.029$, $\widehat{p}_{46} = 0.015$, $\widehat{p}_{57} = 0.047$ and $\widehat{p}_{68} = 0$, $\widehat{p}_{79} = 0$.

Acknowledgements

This research has been supported by the CONACYT Research Project CB-2015/252996, “Modelos con estructuras de Dependencia y sus Aplicaciones II”.

Contribution of the authors in the development of the work

The authors participated according to the following: HP wrote the final conclusions, RC and HP worked on the preliminary ones, JL and RC wrote the introductory section, RC, JL and HP established the results in point and interval estimates, as well as carried out the statistical analysis of the academic progress of university students.

Conflict of interest

All authors declare that they have no conflict of interest.

References

Anderson, T.W., Goodman, L. A. (1957). Statistical Inference about Markov Chains. *The Annals of Mathematical Statistics, 28*(1), 89-110.

Casella, G., Berger, R. (2002). *Statistical inference* (2nd ed.). Thomson Learning.

Durrett, R. (2016). *Essentials of stochastic processes* (Third). Springer, Cham.

Karson, M. J., Wrobleski, W. J. (1976). Confidence intervals for absorbing markov chain probabilities applied to loan portfolios. *Decision Sciences, 7*(1), 10-17.

Kemeny, J. G., Snell, J. L. (1976). *Finite Markov chains*. Springer-Verlag, New York-Heidelberg.

Mendoza, M. (1994). Asymptotic normality under transformations. a result with bayesian applications. *TEST: An Official Journal of the Spanish Society of Statistics and Operations Research*, 3, 173-180.

Muhammad, M., Falgore, J. Y., Sani, U. (2019). Analysis of students’ academic performance and progression (using markov chain approach).

Norris, J. R. (1998). *Markov chains* (Vol. 2). Cambridge University Press, Cambridge.

Perera, S., Bell, M. G. H., Kurauchi, F., Kasthurirathna, D. (2019). Absorbing markov chain approach to modelling disruptions in supply chain networks. 2019 *Moratuwa Engineering Research Conference (MERCon)*, 515-520.

Pinsky, M. A., Karlin, S. (2011). *An introduction to stochastic modeling* (Fourth). Elsevier/ Academic Press, Amsterdam.

Serfling, R. J. (1980). Approximation theorems of mathematical statistics. John Wiley & Sons, Inc., New York.

Wang, L., Laird-Fick, H., Parker, C., Solomon, D. (2021). Using markov chain model to evaluate medical students¨aˆo trajectory on progress tests and predict usmle step 116 scores—a retrospective cohort study in one medical school. *BMC Medical Education, 21.*

Yahaya, K.H., Hasan, H. (2021). Application of markov chain in students´ assessment and performance: A case study of school of mathematical sciences, one of the public university in malaysia. *ITM Web Conf.*, *36,* 01004.

Yu, H.-m., Yang, S.-s., Gao, J.-w., Zhou, L.-y., Liang, R.-f., Qu, C.-y. (2013). Multistate markov model in outcome of mild cognitive impairments among community elderly residents in mainland china. International Psychogeriatrics, *25*(5), 797-804.

Zakarczemny, M., Zajecka, M. (2022) Note on DNA analysis and redesigning using markov chain. Genes, *13*(3).