

UNA RESPUESTA METODOLÓGICA PARA EL ESTUDIO DEL RENDIMIENTO ACADÉMICO EN LA FACULTAD DE CIENCIAS ECONÓMICAS Y ESTADÍSTICA, UNIVERSIDAD NACIONAL DE ROSARIO, ARGENTINA

Gabriela Boggio*
Nora Arnesi**
Guillermina Harvey***
Luciana Ruiz****
Julia Angelini*****

Resumen. Uno de los problemas que enfrenta la Facultad de Ciencias Económicas y Estadística de la Universidad Nacional de Rosario (Argentina) en relación al rendimiento académico de los alumnos es la demora en los tiempos previstos por los planes de estudio para poder elegir carrera. Esta primera meta académica es posible previa aprobación de cuatro asignaturas de un Ciclo Introductorio Común de las tres carreras de Ciencias Económicas. Con el objeto de estudiar qué factores de índole personal y familiar influyen en esta demora se analiza la cohorte de ingresantes del año 2008 a lo largo de cinco años. Dado que el tiempo hasta cumplir la mencionada condición se encuentra medido en años, con pocos valores diferentes, el análisis de datos de duración se puede realizar mediante modelos lineales generalizados previa reestructuración de los datos. La aplicación de esta metodología alternativa fue satisfactoria y entre los resultados hallados se puede mencionar que los estudiantes que presentan mayor demora en el cumplimiento de esta meta son los varones, los egresados de escuela pública, los estudiantes que no trabajan y los que no comenzaron sus estudios universitarios inmediatamente después de finalizada la escolaridad secundaria.

Palabras clave: Rendimiento Académico; Análisis de Datos de Supervivencia; Modelos Lineales Generalizados.

* Universidad Nacional de Rosario, Argentina

Contacto: gboggio@fcecon.unr.edu.ar

** Universidad Nacional de Rosario, Argentina

Contacto: narnesi@fcecon.unr.edu.ar

*** Universidad Nacional de Rosario, Argentina

Contacto: gharvey@fcecon.unr.edu.ar

**** Universidad Nacional de Rosario, Argentina

Contacto: l0ruiz@yahoo.com.ar

***** Universidad Nacional de Rosario, Argentina

Contacto: jangelini@fcecon.unr.edu.ar

A METHODOLOGICAL ANSWER FOR THE STUDY OF ACADEMIC PERFORMANCE AT THE FACULTAD DE CIENCIAS ECONÓMICAS Y ESTADÍSTICA, UNIVERSIDAD NACIONAL DE ROSARIO, ARGENTINA

Abstract. One of the problems regarding the academic performance of students at the Facultad de Ciencias Económicas y Estadística, Universidad Nacional de Rosario (Argentina), is the delay in the expected time limits stipulated by the curricula for choosing the course of study, which is possible only after passing four subjects of an Introductory Cycle Common for the three courses of study of this college. So as to study which family and personal factors affect this delay, the cohort of new students of 2008 is analyzed for a period of five years. Since the time until they meet this condition is measured annually and it assumes a few different values, the survival analysis could be accomplished by generalized linear models using suitably restructured dataset. The application of this methodology was successful and the results show that the students which present more delay in accomplishing this goal are male students, those who attended public schools, those who do not work and those who did not begin their university studies immediately after finishing high school.

Keywords: Academic Performance; Survival Data Analysis; Generalized Linear Models.

Original recibido el 21/12/2017

Aceptado para su publicación el 08/03/2018

1. Introducción

Una problemática que afrontan las autoridades de las instituciones de educación universitaria está relacionada con la demora de los estudiantes en completar sus estudios y el consiguiente riesgo de deserción. Numerosos trabajos han analizado desde diferentes perspectivas el problema del rendimiento, el rezago y el abandono en los estudios universitarios. Tinto (1993) identifica distintos enfoques teóricos para el estudio del abandono y propone un enfoque interaccional que integra factores individuales, como género, experiencia educativa previa y perfil familiar, con aquellos relacionados con el ambiente académico y social. El abandono de los estudios universitarios es, según este autor, el resultado de toda una trayectoria previa signada por factores personales y experiencias institucionales que devienen en la interrupción definitiva de los estudios.

Por otro lado, Garbanzo Vargas (2007) clasifica los factores que se asocian al rendimiento académico en determinantes personales, sociales e institucionales. Entre los determinantes personales menciona la motivación, el género y la formación académica previa. Como factores sociales, reconoce el entorno familiar, el nivel educativo de los padres, el contexto socioeconómico y otras variables demográficas que incluyen el lugar de procedencia y el de residencia del estudiante en época lectiva. Respecto del género, la autora señala que algunos trabajos le asignan a la mujer un rendimiento ligeramente superior que a los varones. En cuanto al nivel educativo de los padres, resalta las conclusiones de algunos autores que consideran que cuanto mayor es el nivel educativo de la madre, son mayores las exigencias académicas planteadas a sus hijos, pero a su vez, los hijos perciben un mayor apoyo hacia los estudios que se suele reflejar en el rendimiento académico alcanzado.

En un estudio realizado sobre las universidades nacionales de la Argentina, García de Fanelli (2014) postula que, si se analiza el abandono y el rendimiento académico en el proceso de admisión o a lo largo del primer año, los factores individuales seguramente mostrarán una alta significación en la explicación de estos fenómenos. Considera relevante tener en cuenta variables demográficas como sexo, edad y lugar de residencia, socioeconómicas como ingreso, nivel educativo y ocupacional de los padres y variables relacionadas con lo académico como gestión de la escuela secundaria (pública o privada), título alcanzado, aspiraciones y motivaciones al ingreso.

En particular en la Facultad de Ciencias Económicas y Estadística de la Universidad Nacional de Rosario (Argentina), se ha percibido un retraso en los tiempos previstos por los planes de estudio de las carreras de Ciencias Económicas para cumplir una primera meta relacionada con la posibilidad de elegir una de ellas: Contador Público, Licenciatura en Administración o Licenciatura en Economía. El plan de estudios de estas carreras comienza con un Ciclo Introductorio Común que tiene por finalidad brindar a los estudiantes los conceptos fundamentales de las disciplinas que conforman las Ciencias Económicas como así también las herramientas matemáticas necesarias para desarrollar asignaturas propias de cada una. La elección de la carrera es posible previa aprobación de cuatro de las seis asignaturas que conforman ese ciclo y ello permite al alumno iniciar el cursado de las asignaturas del denominado

Ciclo Básico correspondiente a la carrera elegida. Si bien el tiempo previsto por el plan de estudios para completar el primer ciclo es de un año académico, un alto porcentaje de estudiantes ha mostrado una demora importante en lograrlo, por lo que interesa no sólo medir este retraso sino también indagar acerca de qué factores tienen impacto directo sobre el mismo.

La Secretaría Académica de la Facultad lleva adelante un trabajo sistemático de monitoreo y seguimiento de los planes de estudio en diferentes cohortes de ingresantes, lo que permite analizar detalladamente las trayectorias académicas del alumnado (Cavallo *et al.*, 2017). Haciendo foco en el cumplimiento de la condición necesaria para poder elegir una de las tres carreras de Ciencias Económicas, resulta de interés estudiar qué factores de índole personal y familiar así como los relacionados con el desempeño académico del estudiante, influyen en la prolongación de los tiempos previstos por los planes de estudio para lograr estar en condiciones de elegir la carrera y así poder cursar las asignaturas propias de cada una de ellas. Las características de la información disponible hacen que sea un desafío en sí mismo dar una respuesta metodológica desde un enfoque estadístico que aproveche al máximo los datos con los que se cuenta, constituyéndose en el objetivo del presente trabajo.

2. Los datos

Para cumplir con este objetivo se utiliza la información relacionada con los alumnos ingresantes al Ciclo Introdutorio Común de las carreras de Contador Público, Licenciatura en Administración y Licenciatura en Economía de la Facultad de Ciencias Económicas y Estadística de la Universidad Nacional de Rosario en el año 2008. La población objeto de estudio, denominada “cohorte real”, está circunscripta a aquellos alumnos que han mostrado algún tipo de actividad académica durante su primer año en la Facultad, ya sea porque lograron la regularización o aprobación de alguna asignatura o al menos muestran intención de continuar los estudios universitarios mediante la reinscripción al año siguiente. Se utilizan los datos referidos a inscripción, reinscripción anual y trayectoria académica de los alumnos hasta marzo de 2013 cubriendo de esta forma la duración formal de cinco años de las carreras mencionadas.

Si bien los alumnos pueden elegir carrera finalizado cada turno de exámenes del año académico, en este trabajo se utiliza el registro de elección de carrera que figura en la reinscripción anual. De esta forma la variable “tiempo desde el ingreso hasta cumplir las condiciones para poder elegir carrera” toma los valores 1, 2, 3, 4 y 5 según la cantidad de años académicos que haya necesitado el alumno para cumplir los requisitos desde el momento de ingreso a la Facultad a comienzos del año 2008. Se registró también la pérdida de información provocada por el abandono o interrupción de los estudios durante el seguimiento así como la información sobre los alumnos que al momento de finalizar el estudio no habían alcanzado el objetivo. Cabe señalar que, dado que no es frecuente que el abandono se formalice a través de la comunicación explícita por parte del alumno de la decisión tomada, se resolvió considerar tres años consecutivos sin reinscripción como un abandono implícito. Por último, hubo situaciones especiales derivadas por ejemplo de períodos de pasividad

seguidos de reinscripción, las cuales fueron analizadas en forma particular para la asignación del correspondiente tiempo. Para ello se tuvieron en cuenta variables auxiliares correspondientes a cada uno de los años académicos considerados tales como condición del alumno (activo, pasivo, con cancelación de matrícula, sancionado, egresado) y cantidad de materias aprobadas.

Las variables incluidas en el análisis de los datos se presentan a continuación de acuerdo a la recategorización utilizada.

Entre las referidas a características personales se pueden mencionar: sexo, edad (medida en años), lugar de procedencia (Rosario, resto de Santa Fe, otras provincias), momento de egreso de la escuela secundaria (año 2007, antes de 2007), clase de escuela secundaria (privada, pública), título secundario otorgado (afín a las ciencias económicas, otro), orientación vocacional¹ (recibió, no recibió), estado ocupacional (trabaja, no trabaja).

En cuanto a las variables relacionadas con la familia del alumno se tuvo en cuenta: trabajo remunerado de la madre (sí, no), nivel de educación formal de la madre (sin estudios o escuela primaria incompleta, escuela primaria completa o secundaria incompleta, escuela secundaria completa o terciario incompleto, terciario completo o universitario incompleto o completo) y nivel de educación formal del padre con las mismas categorías que para el caso de la madre, todas ellas referidas al momento de inscripción en marzo de 2008².

Por último, en relación a la trayectoria universitaria del alumno, se considera el desempeño académico en categorías ordenadas de acuerdo al promedio de notas de exámenes finales acumulado a través de los años de estudio (promedio menor a 6, promedio entre 6 y 8, promedio mayor o igual a 8). Se asignan los *scores* no equiespaciados 1, 3 y 4 a las categorías de esta variable a fin de reflejar una distancia mayor entre un desempeño insuficiente (promedio menor a 6) y las otras categorías.

3. Metodología

En primer lugar, se estudia la trayectoria académica de la cohorte de ingresantes 2008 hasta marzo 2013, es decir se la observa en forma retrospectiva por cinco años consecutivos. Al finalizar cada año académico se consideran los logros alcanzados por cada alumno mediante la definición de los siguientes estados: estar en condición de elegir carrera (cuatro o más materias aprobadas del Ciclo Introductorio Común), abandonar los estudios y los datos por la cantidad de materias aprobadas al finalizar cada año académico (ninguna, una, dos o tres).

Haciendo uso del lenguaje de Cadenas de Markov, los dos primeros se corresponden con estados absorbentes debido a que cuando un alumno alcanza uno de esos estados permanece en él indefinidamente y desde ellos no se puede acceder a ningún otro estado. En cambio, los referidos a la cantidad de materias aprobadas son estados transitorios ya que desde ellos se puede pasar a otro estado diferente, ya sea

¹ Se refiere al acompañamiento de los jóvenes en la elección de una carrera, ya sea a través de *tests*, entrevistas, encuentros grupales con profesionales idóneos o cualquier otra herramienta orientadora.

² No se tuvo en cuenta el estado ocupacional del padre ya que prácticamente en todos los casos el padre del ingresante trabajaba.

otro transitorio o bien alcanzar uno absorbente. En base a los estados definidos se construyen las matrices de transición en un paso durante el período de seguimiento considerado, las cuales muestran los cambios de estado que se producen de un año académico al siguiente. Ello permite medir, no sólo la proporción de ingresantes que logran el objetivo en cada año considerado, sino también estudiar la dinámica de cambios de un estado a otro hasta alcanzar alguno de los estados absorbentes posibles: condición cumplida o bien abandono de los estudios (Privault, 2018).

En una segunda instancia, se hace foco en el estudio de la demora en la aprobación de cuatro asignaturas del Ciclo Introdutorio Común mediante el “análisis del tiempo desde el ingreso hasta cumplir las condiciones para poder elegir carrera”. Entre las características que distinguen a esta variable, se señala que si bien es de carácter continuo, se encuentra medida en años cumplidos asumiendo de esta forma pocos valores diferentes. Otra particularidad es que se pueden presentar casos con censura, es decir, contar para algunos individuos con información incompleta desconociendo el valor exacto que asume la variable y sólo saber, por ejemplo, que el tiempo es mayor que una determinada cantidad de años. Ello impone la utilización de métodos estadísticos para datos de duración agrupados, frecuentemente denominados como métodos de análisis de supervivencia para tiempos agrupados o discretos.

En principio los enfoques tradicionales de modelización para tiempos continuos (Collett, 2015; Hosmer, Lemeshow y May, 2008) se podrían aplicar también a este tipo de datos agrupados. Sin embargo es conveniente modelarlos en el marco de los modelos lineales generalizados (MLG) a través del “tratamiento del tiempo hasta el evento de interés” (cumplir las condiciones para elegir carrera) como un conjunto de variables dicotómicas que indican si el evento ocurrió o no en cada período de tiempo hasta el momento de ocurrencia del evento o censura.

Se presenta a continuación una breve descripción de estos modelos.

3.1. Modelos para tiempos agrupados o discretos

En un estudio de datos de duración agrupados se dispone de un resumen de lo que ha ocurrido durante un intervalo de tiempo. Formalmente el tiempo continuo se divide en intervalos:

$$[0, a_1), [a_1, a_2), \dots, [a_q, \infty).$$

Se pueden utilizar números enteros positivos para denotar estos intervalos. Así, si se define la variable T : “tiempo desde un origen prefijado hasta la ocurrencia del evento de interés”, $T=t$ ($t=1,2,\dots,q+1$) significa que el evento ha ocurrido en el intervalo $[a_{(t-1)}, a_t)$, también llamado período de tiempo t .

La herramienta que representa mejor el aspecto dinámico de esta variable respuesta es la denominada función *hazard* y en este contexto está definida como:

$$\lambda_t = P(T=t / T \geq t),$$

es decir, la probabilidad condicional de que el evento se produzca en el intervalo $[a_{(t-1)}, a_t)$ dado que no se produjo antes de él.

La función *hazard* así definida modela una respuesta binaria que distingue entre

los eventos que ocurren en la categoría t o no (siempre dado que $T \geq t$). En otras palabras, distingue entre la categoría t y las categorías $\{t+1, \dots, q+1\}$.

Para respuestas de este tipo, existe un amplio rango de modelos, en particular los correspondientes a la clase de MLG, los cuales toman la siguiente forma:

$$g(\lambda_{it}) = \gamma_t + \mathbf{x}'_i \boldsymbol{\beta}$$

donde:

\mathbf{x}'_i es el vector de valores de las covariables nominales, ordinales o continuas correspondiente al i -ésimo individuo,

$\boldsymbol{\beta}$ es el vector de coeficientes asociados y

$g(\cdot)$, la función de enlace que relaciona el predictor lineal con la función *hazard*³.

A diferencia del clásico MLG (Agresti, 2015), el parámetro γ_t es dependiente del tiempo, representando una constante fija más el efecto del tiempo asumido como una variable categórica a través de la consideración de q variables de diseño $\{D\}$. Debido a ello, el modelo puede expresarse del siguiente modo:

$$g(\lambda_{it}) = \gamma_0 + \gamma_1 D_{1i} + \dots + \gamma_q D_{qi} + \mathbf{x}'_i \boldsymbol{\beta}.$$

Esta representación del modelo permite obtener una medida del efecto de que el evento ocurra en un determinado momento en comparación con que ocurra en otro momento tomado como categoría de referencia.

La aplicación de este enfoque requiere de una reestructuración de los datos que consiste en tratar cada dato de supervivencia individual como un conjunto de observaciones dicotómicas de manera que la variable tiempo pasa a estar representada por un vector de variables indicadoras acerca de si cada individuo presenta o no el evento en cada unidad de tiempo. Ello es particularmente útil para incluir covariables dependientes del tiempo, es decir covariables cuyo valor para un individuo particular cambia durante el período de seguimiento (Allison, 2006; Hedecker, Ohidul y Frank, 2009).

Si bien el modelo de regresión logística, MLG con enlace “*logit*”, es el más conocido para respuestas dicotómicas, existen otras funciones de enlace que dan lugar a modelos alternativos. Una función de distribución utilizada es la distribución de Gompertz o distribución del valor extremo mínimo que conduce a la función de enlace conocida como “log-log del complemento”, dando lugar al siguiente modelo:

$$\log(-\log(1-\lambda_{it})) = \gamma_0 + \gamma_1 D_{1i} + \dots + \gamma_q D_{qi} + \mathbf{x}'_i \boldsymbol{\beta}.$$

Este modelo se suele llamar modelo de *hazards* proporcionales discretos porque puede considerarse como una versión discretizada del clásico modelo de *hazards* proporcionales de Cox⁴ (Allison, 2006; Hedecker *et al.*, 2009).

Esta formulación resulta atractiva ya que el efecto de las covariables sobre el *hazard*

3 Las funciones de enlace g se pueden definir de manera tal que g^{-1} es una función de distribución de probabilidad.

4 El modelo de Cox supone que dos grupos de individuos presentan *hazards* proporcionales cuando los cocientes entre los respectivos *hazards* son iguales a una constante para todo momento t .

del evento de interés se puede interpretar a partir de la relación que existe entre los coeficientes del modelo y la razón entre dos *hazards*. Así para el caso simple de un modelo con un sólo predictor binario (con valores 0 y 1), $\log(-\log(1-\lambda_{it}))=y_i+\beta x_i$, se puede demostrar que la razón de *hazards* (*RH*) toma la siguiente forma:

$$RH = \frac{\lambda_{it/x=1}}{\lambda_{it/x=0}} = \exp(\beta).$$

El ajuste de este modelo es satisfactorio en muchos casos. Sin embargo, es interesante flexibilizar la elección de la función de enlace recurriendo a una familia de funciones más amplia conocida como distribución log-Burr (Tutz y Schmid, 2016). La función de distribución acumulada correspondiente está dada por

$$F_{\xi}(u) = 1 - (1 + \xi \exp(u))^{-1/\xi}$$

con $\xi > 0$.

Esta distribución es asimétrica a la derecha para $\xi < 1$ y asimétrica a la izquierda para $\xi > 1$. Si $\xi = 1$, es simétrica obteniéndose la distribución logística y cuando $\xi \rightarrow 0$ se obtiene la distribución del valor extremo mínimo. Como puede apreciarse, esta familia comprende los dos modelos más ampliamente utilizados en la modelización de tiempos hasta el evento agrupados y permite captar otras asimetrías a través de los diferentes posibles valores que puede tomar el parámetro ξ . De este modo, el MLG más adecuado se puede buscar entre aquellos con función de enlace g igual a F_{ξ}^{-1} .

El procesamiento de los datos se realiza mediante el uso del programa SAS/STAT® (versión 9.3) y el Entorno “R” (R Core Team, 2017).

4. Resultados

La cohorte real de ingresantes en el año 2008⁵ está conformada por 1.723 alumnos, 1.078 (63%) mujeres y 645 (37%) varones. El 41% del alumnado ya residía en Rosario, el 28% proviene de otras localidades de la provincia de Santa Fe mientras que el 31% restante procede de otras provincias. El 51% manifiesta no trabajar al momento del ingreso en la Facultad y un alto porcentaje de alumnos (75%) finalizó los estudios secundarios el año anterior al ingreso en la Facultad con edades entre 18 y 19 años (98%)⁶.

En el afán de analizar año a año cómo esta cohorte de ingresantes a la Facultad va logrando la primera meta académica (la aprobación de cuatro materias del Ciclo Introductorio Común para poder entonces estar en condiciones de elegir una de las tres carreras de Ciencias Económicas disponibles) se presentan en la Tabla 1. En esta se puede observar las transiciones desde el estado alcanzado al finalizar cada

5 La cohorte de alumnos que han mostrado algún tipo de actividad académica durante su primer año en la Facultad, ya sea porque aprobaron o regularizaron alguna asignatura o, al menos, muestran intención de continuar los estudios universitarios mediante la reinscripción al año siguiente.

6 La variable edad del alumno no se tiene en cuenta en los análisis posteriores debido a su solapamiento con la variable momento de egreso de la escuela secundaria.

año académico hasta el correspondiente al siguiente año, sin tomar en consideración los alumnos que traían alguna materia aprobada por el sistema de equivalencias⁷.

Los diferentes estados representan las situaciones que van desde no haber logrado aprobar ninguna materia del Ciclo Introdutorio Común, pese a haber rendido exámenes parciales o bien finales pero con nota insuficiente, hasta haber cumplido la condición de aprobar cuatro materias de este ciclo o haber tomado la decisión de abandonar los estudios.

Tabla 1. Transiciones correspondientes al período 2008-2013. Facultad de Ciencias Económicas y Estadística.

Transición 2008-2009							
Situación al inicio del año académico 2008	Situación al inicio del año académico 2009						
	Ninguna materia aprobada	1 materia aprobada	2 materias aprobadas	3 materias aprobadas	Condición cumplida	Abandonó estudios	Total
Ninguna materia aprobada	298 (0,19)	170 (0,11)	114 (0,07)	112 (0,07)	907 (0,57)	907 (0,57)	1604 (1,00)

Transición 2009-2010							
Situación al inicio del año académico 2009	Situación al inicio del año académico 2010						
	Ninguna materia aprobada	1 materia aprobada	2 materias aprobadas	3 materias aprobadas	Condición cumplida	Abandonó estudios	Total
Ninguna materia aprobada	242 (0,81)	18 (0,06)	21 (0,07)	8 (0,03)	8 (0,03)	1 (0,00)	298 (1,00)
1 materia aprobada	-	100 (0,59)	30 (0,18)	17 (0,10)	22 (0,13)	1 (0,01)	17 (1,00)
2 materias aprobadas	-	-	74 (0,65)	12 (0,11)	27 (0,24)	1(0,01)	114 (1,00)
3 materias aprobadas	-	-	-	38 (0,34)	74 (0,66)	0 (0,00)	112 (1,00)
Condición cumplida	-	-	-	-	907(1,00)	-	907 (1,00)
Abandonó estudios	-	-	-	-	-	3 (1,00)	3 (1,00)
Total	242 (0,15)	242 (0,15)	125 (0,08)	75 (0,05)	1038 (0,65)	6 (0,00)	1.604 (1,00)

(Continúa)

7 El grupo de 1.723 alumnos queda reducido para este estudio de las transiciones a 1.604 por la eliminación de aquellos que inician los estudios con algunas materias aprobadas por sistema de equivalencias.

Tabla 1. (continuación)

Transición 2010-2011							
Situación al inicio del año académico 2010	Situación al inicio del año académico 2011						
	Ninguna materia aprobada	1 materia aprobada	2 materias aprobada	3 materias aprobadas	Condición cumplida	Abandonó estudios	Total
Ninguna materia aprobada	122 (0,50)	11 (0,05)	4 (0,02)	2 (0,01)	2 (0,01)	102 (0,42)	242 (1,00)
1 materia aprobada	-	63 (0,53)	8 (0,07)	0 (0,00)	1 (0,01)	46 (0,39)	118 (1,00)
2 materias aprobadas	-	-	77 (0,62)	5 (0,04)	18 (0,14)	25 (0,20)	125 (1,00)
3 materias aprobadas	-	-	-	35 (0,47)	28 (0,37)	12 (0,16)	75 (1,00)
Condición cumplida	-	-	-	-	1.038 (1,00)	-	1.038 (1,00)
Abandonó estudios	-	-	-	-	-	6 (1,00)	6 (1,00)
Total	122 (0,08)	74 (0,05)	89 (0,06)	42 (0,03)	1.086 (0,68)	191(0,12)	1.604 (1,00)

Transición 2011-2012							
Situación al inicio del año académico 2011	Situación al inicio del año académico 2012						
	Ninguna materia aprobada	1 materia aprobada	2 materias aprobada	3 materias aprobadas	Condición cumplida	Abandonó estudios	Total
Ninguna materia aprobada	52 (0,43)	3 (0,02)	1 (0,01)	1 (0,01)	0 (0,00)	65 (0,53)	122 (1,00)
1 materia aprobada	-	41 (0,55)	0 (0,00)	0 (0,00)	2 (0,03)	31 (0,42)	74 (1,00)
2 materias aprobadas	-	-	48 (0,54)	5 (0,06)	4 (0,04)	32 (0,36)	89 (1,00)
3 materias aprobadas	-	-	-	18 (0,43)	8 (0,19)	16 (0,38)	42 (1,00)
Condición cumplida	-	-	-	-	1.086 (1,00)	-	1.086 (1,00)
Abandonó estudios	-	-	-	-	-	191 (1,00)	191 (1,00)
Total	52 (0,03)	44 (0,03)	49 (0,03)	24 (0,02)	1.100 (0,69)	335 (0,21)	1.604 (1,00)

(Continúa)

Tabla 1. (continuación)

Transición 2012-2013							
Situación al inicio del año académico 2012	Situación al inicio del año académico 2013						
	Ninguna materia aprobada	1 materia aprobada	2 materias aprobadas	3 materias aprobadas	Condición cumplida	Abandonó estudios	Total
Ninguna materia aprobada	30 (0,58)	0 (0,00)	1 (0,02)	0 (0,00)	0 (0,00)	21 (0,40)	52 (1,00)
1 materia aprobada	-	23 (0,52)	2 (0,05)	1 (0,02)	2 (0,05)	16 (0,36)	44 (1,00)
2 materias aprobadas	-	-	26 (0,53)	4 (0,08)	0 (0,00)	19 (0,39)	49 (1,00)
3 materias aprobadas	-	-	-	13 (0,54)	4 (0,17)	7 (0,29)	24 (1,00)
Condición cumplida	-	-	-	-	1.100 (1,00)	0 (0,00)	1.100 (1,00)
Abandonó estudios	-	-	-	-	-	335 (1,00)	335 (1,00)
Total	30 (0,02)	23 (0,01)	29 (0,02)	18 (0,01)	1.106 (0,69)	398 (0,25)	1.604 (1,00)

Fuente: Elaboración propia.

A simple vista, se observan diferencias importantes en las transiciones del primer año de cursado respecto de las cuatro transiciones restantes, las cuales son bastante similares entre sí.

Se puede apreciar que la proporción de ingresantes que logran la condición para elegir carrera en los plazos previstos por los planes de estudio, es decir en el primer año académico, es igual a 0,57. Al finalizar el segundo año, esta proporción aumenta a 0,65. Es decir, sólo un 8% más (131 alumnos) de la cohorte inicial logra elegir carrera en el segundo año y se continúa desacelerando en los períodos siguientes: un 3% (48 alumnos) el tercer año, un 1% (14 alumnos) el cuarto año y menos del 1% (sólo 6 alumnos) en el quinto.

Observando al interior de cada matriz de transición, es notorio el estancamiento que se produce en algunos casos; por ejemplo, el 47% de los alumnos que habían aprobado tres materias al iniciar el año académico 2010 siguen en la misma situación después de transitar un año académico más, es decir no pudieron aprobar ninguna otra materia a pesar de estar tan cerca de alcanzar la condición.

Si se enfoca el análisis desde la perspectiva del tiempo transcurrido hasta lograr la condición para elegir carrera, resulta de interés calcular las razones de *hazards* correspondientes a cada uno de los factores que se piensa pueden estar asociados con una demora en la elección de la carrera. Dichas razones de *hazards*, presentadas en la Tabla 2, se interpretan en términos de razones de probabilidades condicionales en función de su definición para el caso de tiempos discretos, tal como se describió en la sección 3.1.

Las razones de *hazards* significativamente menores que uno ($RH < 1$) permiten concluir que la probabilidad estimada de lograr el objetivo en un determinado año académico

siendo que no pudo alcanzarse en los años anteriores, es menor para los varones, para los que egresaron de la escuela secundaria en años anteriores al 2007, para los que asistieron a una escuela pública, para los que no trabajan y para aquellos cuya madre no tiene un trabajo remunerado. En cambio, las razones de *hazards* significativamente mayores que uno muestran que dicha probabilidad estimada es mayor para los alumnos que provienen de otras provincias en comparación con los que ya residían en Rosario, para aquellos cuyo padre o madre tienen un nivel de educación más alto y para los que van mostrando un buen rendimiento académico a través del promedio de notas en las materias que aprueban año a año (Tabla 2).

Tabla 2. Razones de *hazards* estimadas. Análisis univariado

Covariable	Razón de <i>hazards</i>	IC95%
Sexo masculino vs. femenino	0,79	0,70 – 0,89
Lugar de procedencia resto de Santa Fe vs. Rosario otras provincias vs. Rosario	1,03 1,17	0,89 – 1,17 1,02 – 1,34
Momento de egreso de la escuela secundaria antes de 2007 vs. 2007	0,62	0,54 – 0,72
Clase de escuela secundaria pública vs. privada	0,72	0,64 – 0,81
Título secundario afín a las Cs. Económicas vs. otro	1,13	0,99 – 1,29
Orientación vocacional no recibió vs. recibió	0,94	0,82 – 1,10
Estado ocupacional del alumno no trabaja vs. trabaja	0,70	0,63 – 0,79
Trabajo remunerado de la madre no vs. sí	0,49	0,43 – 0,57
Nivel de educación formal de la madre	1,25	1,18 – 1,32
Nivel de educación formal del padre	1,24	1,18 – 1,31
Desempeño académico nota promedio >8 vs. entre 6 y 8 nota promedio entre 6 y 8 vs. <6	1,76 3,10	1,67 – 1,86 2,78 – 3,45

Fuente: Elaboración propia.

Para analizar la posible influencia de estos factores en forma simultánea se recurre a la modelización estadística descrita en la sección metodológica⁸. La selección del predictor se realiza en base a un MLG con enlace log-log del complemento por considerarse el más natural pensando en el tiempo continuo subyacente. Se tiene en cuenta la persistencia de la significación de los efectos de las covariables en

⁸ Para el ajuste de los MLG se reestructura la base de datos original de 1.723 individuos en una conformada por 2594 registros tiempo-individuo.

este contexto multivariado, la evaluación de la escala de medición más razonable en cada caso y la posibilidad de presencia de interacción entre covariables. Una vez elegido el predictor, se prueba el ajuste de modelos con diferentes enlaces, los convencionales en el marco de los MLG y algunos derivados de la familia log-Burr. Se calcula el valor del Criterio de Información de Akaike (AIC) para cada uno de ellos y se confirma en este caso la elección del enlace log-log del complemento por presentar en valor de AIC más chico (ver Tabla A1 del Anexo).

En relación a la evaluación del cumplimiento del supuesto de hazards proporcionales, los resultados hallados muestran que el nivel de educación del padre y el estado ocupacional del alumno no lo verifican. A pesar de ello, se decide incluirlas en el modelo interpretando las respectivas estimaciones como efectos promedios a lo largo de todo el período de seguimiento pudiendo ser mayor o menor en cada año académico en particular.

Se completa el estudio de la adecuación del modelo mediante la evaluación de la bondad de ajuste y de la capacidad predictiva. Si bien el *test* de bondad de ajuste utilizado -*test* de Hosmer y Lemeshow- supera levemente los niveles de significación habituales (estadística=14,61; probabilidad asociada=0,0671), la estadística “c” igual a 0,84 muestra un alto valor predictivo.

El modelo finalmente elegido incluye los efectos de las variables: sexo, clase de escuela secundaria a la que asistió, estado ocupacional del alumno, desempeño académico en la Facultad, nivel educativo del padre y la interacción entre momento de egreso de la escuela secundaria y trabajo remunerado de la madre. Es decir, la influencia del lugar de procedencia y del nivel de educación de la madre deja de ser significativa al considerar simultáneamente todos los factores (ver Tabla A2 del Anexo).

Con el objeto de dar una medida acerca de cuánto mayor o menor es la demora según las características de los alumnos, se estiman las razones de *hazards* a partir de los coeficientes del modelo ajustado (Tabla 3).

Se puede apreciar que la probabilidad estimada de poder elegir carrera para aquellos alumnos que aprueban las cuatro materias al cabo del segundo año es alrededor de la mitad (46%)⁹ que para los que lo logran en el primer año. En cambio si se comparan las probabilidades de cumplir los requisitos en los años siguientes en comparación con el año académico inmediato anterior, se observan disminuciones que no alcanzan a ser significativas (entre el 19% y 42%).

⁹ Este valor se obtiene como $(1-0,54) 100\%$ donde 0,54 es la razón de *hazards* estimada presentada en Tabla 3.

Tabla 3. Razones de *hazards* estimadas para el modelo con enlace log-log del complemento

Covariable	Razón de <i>hazards</i>	IC _{95%}
Tiempo		
2° año vs. 1° año	0,54	0,45 – 0,66
3° año vs. 2° año	0,81	0,58 – 1,14
4° año vs. 3° año	0,58	0,32 – 1,07
5° año vs. 4° año	0,62	0,24 – 1,61
Sexo		
masculino vs. femenino	0,82	0,71 – 0,94
Clase de escuela secundaria		
pública vs. privada	0,80	0,70 – 0,91
Estado ocupacional del alumno		
no trabaja vs. trabaja	0,74	0,65 – 0,85
Nivel de educación formal del padre	1,09	1,02 – 1,16
Momento de egreso de la escuela secundaria		
antes 2007 vs. 2007		
cuando la madre no tiene trabajo remunerado	0,74	0,55 – 0,99
cuando la madre tiene trabajo remunerado	0,46	0,38 – 0,55
Desempeño académico		
nota promedio >8 vs. entre 6 y 8	1,77	1,67 – 1,89
nota promedio entre 6 y 8 vs. <6	3,15	2,79 – 3,55

Fuente: Elaboración propia

Se estima que la probabilidad de que un alumno pueda elegir carrera en un año académico particular dado que no lo logró en años anteriores es un 18% menor para los varones, un 20% menor para los que egresan de escuela pública y un 26 % menor para los que no trabajan. Esta probabilidad también es más baja para aquellos que egresaron antes del año 2007, lo que se ve potenciado si la madre tiene trabajo remunerado (26% menor cuando la madre no tiene trabajo remunerado y un 54% menor cuando sí lo tiene). El desempeño académico que va teniendo el alumno año a año en la Facultad también, como era de esperar, influye sobre esta demora en alcanzar la meta. Así los alumnos con promedio de notas en los exámenes mayor a 8 tienen una probabilidad 77% mayor de lograrlo que los que tienen promedio entre 6 y 8. Si se comparan los alumnos con estas notas intermedias con los que no alcanzan un promedio de 6, esta probabilidad puede llegar a ser el triple, mostrando que no es fácil remediar los reveses derivados de exámenes insuficientes. Por último, la influencia de las características paternas se aprecia en un aumento de la probabilidad de elegir carrera en un año académico particular dado que no lo pudo lograr en años anteriores, a medida que el nivel de educación del padre es más alto. Se puede completar el análisis de los resultados obtenidos presentando las

probabilidades estimadas de lograr el objetivo planteado durante el primer año académico, tal como lo prevé el plan de estudios, para diferentes perfiles de alumnos desde la mejor situación a la más desfavorable.

El grupo de alumnos "ideal" -desde el punto de vista de su logro académico en esta primera etapa universitaria- está conformado por mujeres que finalizaron los estudios secundarios en el año anterior al ingreso a la Universidad (año 2007), en escuela privada, que trabajan y con promedio de notas en los exámenes de la Facultad alto, con padre con alto nivel de educación y madre con trabajo remunerado. La probabilidad estimada de lograr la meta en el primer año académico para este grupo ideal de ingresantes alcanza el valor uno. Si, por ejemplo, estas alumnas no trabajasen y el promedio de notas que lograrán en su trayectoria universitaria superara el 6, si bien no es sería mayor a 8, esa probabilidad se reduciría tan sólo a 0,94.

Por otro lado, el grupo de alumnos con menores expectativas de cumplir la meta objeto de estudio está formado por los varones que egresaron antes del 2007 de escuela pública, que no trabajan y con promedio de notas en los exámenes de la Facultad bajo, con padre de bajo nivel de educación y madre con trabajo remunerado. Para ellos, la probabilidad estimada de lograr el objetivo de acuerdo al plan de estudios es de 0,15. Esta probabilidad aumentaría a 0,20 si los alumnos trabajaran y a 0,24 si, además, egresaran de una escuela privada. En el plazo de 2 años las mencionadas probabilidades, como es de esperar, aumentan y casi se duplican si se considera un plazo de 3 años (0,28, 0,35 y 0,42 respectivamente). La identificación de estos perfiles de ingresantes que presentan grandes dificultades para alcanzar la meta resulta importante para direccionar hacia ellos la tarea de acompañamiento por parte de los tutores en los primeros años en la Facultad.

5. Discusión

En este trabajo se indaga acerca de los factores que pueden estar relacionados con la demora observada en los ingresantes a la Facultad de Ciencias Económicas y Estadística de la Universidad Nacional de Rosario en poder elegir carrera, condición que el plan de estudios prevé que se alcanza en el primer año de cursado pero que un grupo importante de alumnos no logra cumplirla.

Los resultados hallados muestran que los alumnos que presentan mayor demora en el cumplimiento de los requisitos para elegir carrera son los varones, los egresados de una escuela secundaria pública, los estudiantes que no trabajan, los que tienen (como era de esperar) un promedio bajo en las notas de las materias que van rindiendo en la Facultad y los que tienen padre con nivel de educación formal bajo. También, muestran mayor demora los que han terminado la educación secundaria años antes del ingreso a la Facultad, lo que se ve potenciado si además las madres trabajan. Este resultado amerita realizar una indagación adicional para lograr comprender el porqué de la significación estadística de esta interacción entre el momento de egreso de la escuela secundaria y el estado ocupacional de la madre.

Muchos de los hallazgos concuerdan con los obtenidos por otros grupos de investigación en donde se aborda el estudio del rendimiento académico universitario. Por ejemplo, Porto y Di Gresia (2004) encuentran que las mujeres más jóvenes

con padres de alto nivel de instrucción presentan mejor desempeño. No obstante, Garbanzo Vargas (2007) considera que el rendimiento académico es el resultado de la suma de diferentes y complejos factores donde entran en juego determinantes personales, sociales e institucionales en los que se desenvuelve el estudiante. En un estudio realizado en la Facultad de Ciencias Económicas de la Universidad Nacional de La Plata (Argentina), se identificaron factores que influyen en la probabilidad de completar el ciclo de formación inicial para las cohortes de ingresantes desde el año 2000 al 2003. Dicha probabilidad resultó mayor para las mujeres, para quienes poseen padres con niveles más altos de educación, para quienes no trabajan al inicio de la carrera y para los que tienen un nivel socio-económico relativamente alto (Carella, Ferreyra y Pron, 2007).

Asimismo, en otras Facultades del país se encontraron resultados semejantes al estudiar el rendimiento académico. A partir del análisis de la producción científica llevada a cabo entre el 2002 y el 2012 sobre esta problemática en las universidades nacionales de la Argentina, García de Fanelli (2014) concluye que los factores que resultaron significativos en un mayor número de publicaciones son, por orden de importancia, el género (mejor desempeño de las mujeres), el nivel educativo de los padres (mejor desempeño cuanto mayor es el nivel de educación), la actividad económica (la cantidad de horas trabajadas afecta negativamente el rendimiento), la formación académica previa (un buen promedio en la escuela media incide positivamente), la edad (los más jóvenes obtienen mejores resultados) y el tipo de escuela de la cual egresan (el haber estudiado en un colegio privado o en uno gestionado por una universidad incide positivamente). Respecto de las horas de trabajo señala que algunas investigaciones observan una relación no lineal; en un tramo de baja cantidad de horas el rendimiento del alumno es positivo y por encima de cierta cantidad incide negativamente. A su vez, si el trabajo está vinculado con la carrera, es posible obtener beneficios cuando el trabajo no es de tiempo completo. Estos últimos resultados están en sintonía con los hallados en el presente estudio donde se evidencia que el hecho de trabajar beneficia al rendimiento del alumno si bien en este estudio la carga horaria no pudo considerarse debido a los problemas de registro encontrados.

En cuanto a las posibles alternativas metodológicas disponibles para cumplir con el objetivo planteado, el análisis de datos de duración desde la perspectiva de los MLG resultó satisfactorio. Éste supone un tiempo continuo hasta la posibilidad de elegir carrera pero en el cual la ocurrencia de este evento se toma en cuenta una vez por año académico, lo que transforma la escala en puntos discretos en el tiempo. El enlace del MLG hallado como el más apropiado, log-log del complemento, ha permitido interpretar el efecto de los factores influyentes sobre la demora para estar en condiciones de elegir carrera en términos de razones de *hazards* debido a su relación con el modelo de *hazards* proporcionales desarrollado por Cox para tiempos en escala continua (Läärä y Matthews, 1985).

Cabe señalar que se podría haber considerado también un enfoque basado en el tratamiento del tiempo hasta el evento como una variable de escala ordinal mediante el ajuste de MLG para las probabilidades acumuladas de esa respuesta ordinal

(Agresti, 2013; Bennet, 1983). Este enfoque, si bien no requiere de ningún trabajo adicional con respecto a la disposición de los datos, impone una tarea de programación adicional para considerar la presencia de datos censurados en el ajuste de los modelos. Por el contrario, el enfoque elegido, una vez realizada la reestructuración de los datos, permite ajustar los modelos a partir de cualquier *software* convencional y es superior en el tratamiento de covariables que dependan del tiempo. En este caso la construcción de la nueva base de datos permitió incorporar sin dificultades la información sobre la variable dependiente del tiempo “desempeño académico” medido por el promedio de notas en los exámenes rendidos año a año.

A modo de cierre se puede concluir que la metodología utilizada da lugar al reconocimiento de perfiles con características particulares. Así, la identificación de los grupos de alumnos que presentan más dificultades para lograr este objetivo puede contribuir a la elaboración de estrategias por parte de los responsables del sistema de tutorías a fin de mejorar la inserción de los alumnos en su primer año académico.

Referencias bibliográficas

- Agresti, A. (2013). *Analysis of ordinal categorical data*. 2º ed. Nueva York, EEUU: John Wiley y Sons.
- Agresti, A. (2015). *Foundations of linear and generalized linear models*. Nueva York, EEUU: John Wiley & Sons.
- Allison, P. (2006). *Survival analysis using SAS. A practical guide*. Carolina del Norte, EEUU: SAS Institute Inc.
- Bennet, S. (1983). Analysis of survival data by the proportional odds model. *Statistics in Medicine*, 2, 273-277.
- Carella, L.; Ferreyra, G. y Pron, J. (2007). Desempeño en el ciclo de formación inicial: análisis de cohortes de la Facultad de Ciencia Económica de UNLP. En A. Porto (Ed), *Mecanismos de admisión y rendimiento académico de los estudiantes universitarios. Estudio comparativo para estudiantes de Ciencias Económicas* (117-140). La Plata, Argentina: Editorial de la Universidad de La Plata.
- Collett, D. (2015). *Modelling Survival Data in Medical Research*. Florida, EEUU: Chapman y Hall/CRC.
- Cavallo, M.; Fattore, N.; Medina, M.; Ruiz, L. Giustiniani, P. y Geli, M. (2017). *Seguimiento de implementación de Planes de Estudio 2003 de la Facultad de Ciencias Económicas y Estadística, Universidad Nacional de Rosario*.
- Garbanzo Vargas, G. M. (2007). Factores asociados al rendimiento académico en estudiantes universitarios, una reflexión desde la calidad de la educación superior pública. *Revista Educación*, 31(1), 43-63.

- García de Fanelli, A. M. (2014). Rendimiento académico y abandono universitario: Modelos, resultados y alcances de la producción académica en la Argentina. *Revista Argentina de Educación Superior*, 8, 9-38.
- Hedecker, D.; Ohidul, S. y Frank, B. (2009). Random effects regression analysis of correlated grouped-time survival data. *Statistical Methods in Medical Research*, 9, 161-179.
- Hosmer, D. W., Lemeshow, S. y May, S. (2008). *Applied Survival Analysis*. Nueva York, EEUU: John Wiley & Sons.
- Läärä, E. y Matthews, J. (1985). Equivalence of two models for ordinal data. *Biometrika*, 72(1), 206-207.
- Porto, A. y Di Gresia, L. (2004). Rendimiento de estudiantes universitarios y sus determinantes. *Revista de Economía y Estadística*, 42(1), 93-113.
- Privault, N. (2018). *Understanding Markov Chains. Examples and Applications*. 2º ed. Springer.
- R Core Team (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. Recuperado de <https://www.R-project.org/>
- Tinto, V. (1993). *Leaving College. Rethinking the Causes and Cures of Student Attrition*. Chicago, EEUU: The University of Chicago Press.
- Tutz G. y Schmid, M. (2016). *Modeling discrete time to event data*. Cham, Suiza: Springer Series in Statistics.

6. Anexos:

Anexo 1.

Tabla A1. Valores de AIC para modelos con diferentes enlaces

Enlace	AIC
log-log del complemento ($\xi \rightarrow 0$)	2342,6
F_{ξ}^{-1} con $\xi=0,5$	2351,3
logit ($\xi=1$)	2364,2
probit	2366,1
F_{ξ}^{-1} con $\xi=2$	2389,5
F_{ξ}^{-1} con $\xi=3$	2409,3
F_{ξ}^{-1} con $\xi=4$	2423,9
F_{ξ}^{-1} con $\xi=5$	2434,5

Fuente: Elaboración propia

Anexo 2.

Tabla A2. Estimaciones del modelo con enlace log-log del complemento

Parámetro	Estimación (error estándar)	Prob. asociada
Intercepto	-0,80 (0,13)	<0,0001
Tiempo		
2° año	-0,61 (0,10)	<0,0001
3° año	-0,82 (0,16)	<0,0001
4° año	-1,36 (0,27)	<0,0001
5° año	-1,84 (0,41)	<0,0001
Sexo masculino	-0,20 (0,07)	0,0055
Momento de egreso de la escuela secundaria antes 2007	-0,78 (0,09)	<0,0001
Clase de escuela secundaria pública	-0,23 (0,07)	0,0008
Estado ocupacional del alumno no trabaja	-0,30 (0,07)	<0,0001
Trabajo remunerado de la madre no	-0,66 (0,09)	<0,0001
Nivel de educación formal del padre	0,08 (0,03)	0,0096
Momento de egreso * Trabajo remun. madre antes 2007 y no tiene	0,49 (0,17)	0,0052
Desempeño académico	0,57 (0,03)	<0,0001

Fuente: Elaboración propia.