

ESTUDIO DEL ABANDONO EMPLEANDO UN MODELO DE RIESGOS PROPORCIONALES

Línea Temática: Factores asociados. Tipos y perfiles de abandono

Mariana Verónica Gonzalez

Instituto de Estadística y Demografía. Facultad de Ciencias Económicas. UNC

mvgonzalez@eco.unc.edu.ar

Resumen.

En este trabajo se aplica un modelo de riesgo semi-paramétrico para identificar algunas variables que contribuyen a explicar el fenómeno del abandono estudiantil en una unidad académica de la Universidad Nacional de Córdoba (UNC), Argentina. La inactividad académica prolongada se toma como *proxy* de la variable dependiente. Los resultados muestran que el promedio del secundario aparece como importante predictor de la deserción. Además, se detectó que los estudiantes que costean sus estudios con aporte familiar exclusivamente tienen mayor riesgo relativo de abandono respecto de aquellos que lo hacen con su trabajo personal únicamente. La relación cantidad de materias aprobadas respecto a cantidad de materias rendidas y el lugar de procedencia del estudiante también resultaron significativas.

Descriptor o Palabras Clave: abandono, modelo de Cox, riesgo

1. Introducción

El abandono de los estudios universitarios es un fenómeno extremadamente complejo (Cabrera *et al*, 1993), tanto por las dificultades que existen para identificar la situación del proceso educativo en la que debe encontrarse una persona para que pueda considerarse que hizo abandono de sus estudios (definición de deserción), como por la variedad de perspectivas con que puede abordarse la problemática (individual, institucional y estatal).

La mayoría de las Facultades de la Universidad Nacional de Córdoba (UNC) no tiene una definición operativa precisa que permita determinar el momento en que un estudiante abandona sus estudios¹. En el caso de la Facultad de Ciencias Económicas (UNC) la legislación sobre la

¹Una excepción es la Facultad de Ciencias Médicas (UNC) que establece un régimen según el cual un alumno es considerado activo si aprueba al menos dos materias durante el ciclo lectivo (para el primer año la exigencia se reduce a una materia) y si finaliza sus estudios en un lapso no mayor al doble de la duración prevista para la carrera

condición de alumno no prevé las circunstancias bajo las cuales se opera la caducidad, por lo que en el presente trabajo fue necesario adoptar un criterio práctico para intentar una aproximación al concepto de abandono. A tal fin, se consideró que el estado de inactividad académica² durante tres periodos lectivos consecutivos constituye una variable *proxy* del abandono (deserción) que un alumno hace de sus estudios. Para identificar los principales factores que se le asocian se aplicó un modelo semi-paramétrico de Cox.

2. Metodología

En el análisis de supervivencia se examina el tiempo o duración (t) de cada unidad de observación hasta que ocurre un fenómeno predefinido, en este caso, la deserción. De acuerdo a la definición adoptada, si transcurren seis semestres consecutivos (tres periodos lectivos) sin que un alumno exhiba actividad académica, se considera que ese estudiante abandonó los estudios, o sea, se presentó el evento objeto de análisis de riesgo.

El modelo de regresión con riesgos proporcionales de Cox (Cox 1972, Taucher 1999) es utilizado para evaluar la relación entre un conjunto de variables explicativas y el tiempo de supervivencia.

Se define mediante la siguiente función:

$$h(t) = h_0(t) \cdot \exp(b_1 \cdot x_1 + b_2 \cdot x_2 + \dots + b_k \cdot x_k)$$

donde:

$h_0(t)$ es la función de riesgo mínima para una duración t , también denominada riesgo base. Representa la tasa instantánea de riesgo para un sujeto con valor cero en todas las variables explicativas.

b_i parámetros estimados, que representan el cambio proporcional esperado en el riesgo, respecto al cambio marginal en la variable explicativa x_i .

Este modelo es semi-paramétrico porque mientras que el riesgo base puede tomar cualquier forma, las covariables entran en el modelo linealmente. Por lo tanto, con la aplicación de este modelo se tiene que la tasa de deserción es el producto de dos componentes, una $h_0(t)$ que depende del tiempo y representa el riesgo de supervivencia con independencia de los predictores y otra constituida por $\exp(b_1 \cdot x_1 + b_2 \cdot x_2 + \dots + b_k \cdot x_k)$ que depende de las variables de pronóstico o co-variables y no depende del tiempo. El riesgo de cada predictor es una proporción del riesgo base $h_0(t)$, de ahí el nombre de modelo de riesgos proporcionales.

El cociente entre dos riesgos se denomina *Hazard Ratio (HR)* y se obtiene mediante el exponencial del coeficiente de una regresión de Cox, siendo independiente del valor de t .

$$HR = \exp(b)$$

3. Fuente de datos

Los datos seleccionados para el análisis corresponden a la FCE de la que se extrajo una muestra aleatoria de una cohorte de ingresantes en el año 2000, para analizar el proceso de deserción durante un período de ocho años. Los datos sobre el desempeño de cada estudiante fueron obtenidos del sistema SIU-Guaraní y la información referida a la historia académica previa, en el nivel medio, de su legajo individual.

El modelo de Cox permite incorporar variables que varían con cada estudiante y que también lo hacen en función del tiempo. En este sentido fue posible incorporar como variable predictora que

²Se consideró que un estudiante está activo durante un semestre si ha realizado alguna actividad académica (cursar materia y/o rendir algún examen) durante el mismo.

varía con el tiempo un índice de eficiencia que surge de la relación entre la cantidad de materias aprobadas y la cantidad de materias rendidas al final de cada semestre.

Como sucede generalmente en los estudios de duración, dado que la cohorte considerada fue observada hasta el final del año 2011, las bases de datos con las que se trabajó contienen casos censurados por la derecha, que son aquellos estudiantes que al final del periodo de observación no concluyeron sus estudios ni hicieron abandono de la carrera que cursan (por lo que no se sabe si abandonarán o no en el futuro).

4. Resultados

Como se indicó anteriormente, se utilizó el modelo semi-paramétrico de riesgos proporcionales de Cox para intentar detectar factores que contribuyan a explicar el comportamiento del riesgo relativo de abandono. Las diferentes pruebas realizadas incluyeron como variables explicativas aquellas que se encuentran cargadas en el sistema Guaraní y que, además, presentan un razonable nivel de confiabilidad. A ellas se agregó la nota promedio en el secundario que debió capturada desde los legajos individuales de los alumnos.

Estudios anteriores realizados (regresión lineal, regresión logística) permitieron realizar una selección previa de las covariables a ser incluidas en el modelo de Cox, las que se detallan a continuación (Tabla 1):

Tabla 1. Definición de covariables

Nombre	Definición	Medición
Promedio	Promedio obtenido en el secundario	Escala 0 a 10
Procedencia	Lugar de donde procede el estudiante	Córdoba Capital=0, Resto del país =1
Aporte familiar	Costea sus estudios únicamente con aporte familiar	Si=1
Eficiencia	Cantidad de exámenes aprobados dividida por la suma de los exámenes aprobados más los aplazados (o sea, dividida por el total de exámenes rendidos).	Variable que varía con el tiempo
Trabajo	Trabaja	Si=1
Costea	Costea sus estudio únicamente con su trabajo personal	Si=1

Fuente: elaboración propia

El modelo de Cox aplicado supone proporcionalidad en el riesgo entre las diferentes categorías de cada una de las co-variables que se utilizan, hipótesis que debe ser previamente probada. Una de las formas de hacerlo es inspeccionando los gráficos Log-Log de cada una de las variables explicativas para comprobar que las curvas son aproximadamente paralelas entre sí. Esto es precisamente lo que ocurre en cada uno de los gráficos que se muestran en el Apéndice I, donde se observa el trazado muy similar que presentan las curvas, indicando que no hay motivos para dudar que exista proporcionalidad en el riesgo entre las categorías.

Al aplicar el modelo de riesgos proporcionales de Cox se obtuvieron las estimaciones de los parámetros que se presentan en la Tabla 2, mostrados en forma tasa de riesgo (Hazard Ratio). Como puede observarse, todas las covariables consideradas resultaron significativas y con efectos que están en línea con los obtenidos en investigaciones anteriores.

Se puede apreciar que la nota promedio del secundario tiene un efecto marginal muy fuerte sobre el riesgo de abandono en la universidad, ya que por cada punto adicional en dicha nota, el riesgo disminuye aproximadamente un 34%. Este hallazgo es concordante con las conclusiones de otros estudios que indican que el desempeño académico de un estudiante en el nivel medio es un poderoso predictor de su posterior performance en la universidad (Pascarella y Chapman, 1983;

Díaz, 2009). También surge de estos resultados que la variable eficiente, definida como la relación entre total de exámenes aprobados y el total de exámenes rendidos, única variable que cambia con el tiempo, constituye un factor que contribuye significativamente y en forma inversa sobre el riesgo de abandono. El efecto marginal de esta covariable se traduce en una disminución de aproximadamente 9% en el riesgo relativo de abandono. La explicación parece resultar simple si se tiene en cuenta que una mejor performance del alumno lo desalienta a abandonar la carrera y viceversa.

Tabla 2. Estimación del Modelo de Riesgos Proporcionales de Cox

Variable	HR	Error estándar	p- valor	Intervalo de confianza (95%)	
Promedio	0,66	0,06	0,000	0,56	0,79
Procedencia	0,79	0,09	0,048	0,62	0,99
Aporte familiar	2,53	0,60	0,000	1,59	4,04
Eficiencia	0,91	0,02	0,000	0,87	0,95
Trabajo	1,59	0,29	0,010	1,12	2,27
Costea	2,19	0,31	0,000	1,66	2,89

Fuente: elaboración propia

También se puede observar que los estudiantes que no provienen de la ciudad Capital de la provincia Córdoba, presentan un riesgo relativo de abandono que es alrededor de un 21% menor (HR=0,79) respecto de aquellos estudiantes de la ciudad Capital. Este resultado encuentra su explicación posiblemente en el mayor costo de oportunidad para abandonar los estudios que enfrentan los estudiantes que no son de la ciudad de Córdoba.

Un resultado que puede considerarse inesperado es que cuando el estudiante costea su carrera exclusivamente con aporte familiar tiene más del doble (HR=2,53) de riesgo relativo de abandonar respecto del caso que no lo costea de esa forma. Asimismo, si el alumno costea sus estudios exclusivamente con su trabajo personal afronta el doble (HR=2,19) de riesgo relativo de abandonar con relación a aquel estudiante que no costea sus estudios exclusivamente con su trabajo personal. La inferencia que surge de la comparación de estos dos efectos es que aquellos estudiantes que costean los estudios exclusivamente con su trabajo personal están menos expuestos al riesgo de abandono, respecto de aquellos que lo hacen exclusivamente con aporte de sus familiares³. Como se observa, además, los estudiantes que trabajan enfrentan aproximadamente un 50% más de riesgo relativo (HR=1,59) de abandonar la carrera universitaria que cursan, lo cual constituye un resultado esperado para esta variable.

5. Conclusiones

El modelo de riesgos proporcionales de Cox mostró ser apto e interesante para estudiar el comportamiento del abandono de los estudios universitarios, identificando factores que contribuyen a explicarlo. Si bien los datos empleados son anteriores al cambio de plan de estudios que experimentó la FCE en el año 2009, muestran la potencialidad del modelo, así como la posibilidad de efectuar comparaciones con datos de cohortes más recientes.

Uno de los factores con muy alto poder explicativo de la deserción es el desempeño del estudiante en la escuela secundaria reflejado en la nota promedio que obtuvo en todo ese ciclo. Dicha variable

³ Si bien este hallazgo puede aparecer como inesperado, en realidad está en línea con los resultados de otras investigaciones (ver Goldenhersch H *et al.* 2002)

reveló poseer un efecto marginal importante sobre el riesgo de abandono. Esta conclusión está en línea con los resultados encontrados por Giuliadori et al (2007), donde el desempeño académico en los estudios universitarios mostró estar fuertemente ligado a la nota promedio del secundario y a la nota promedio que los alumnos obtuvieron en el Curso de Nivelación que se imparte al momento de acceder a la Facultad.

Los estudiantes que trabajan tienen un mayor riesgo de abandonar sus estudios respecto de quienes no trabajan.

La cantidad de materias aprobadas con relación a la cantidad total de materias rendidas tiene un efecto marginal importante sobre el riesgo relativo de abandono, que se manifiesta en forma inversa: a mayor valor de esa relación, menor riesgo de abandono.

A partir de los resultados del trabajo es posible identificar perfiles de estudiantes con mayor riesgo de abandono. Estos alumnos son aquellos que pertenecen a una o más de las siguientes categorías: (i) costean sus estudios únicamente con aporte familiar, (ii) exhiben una débil performance en la carrera medida a través de la cantidad de materias aprobadas con relación al total de materias rendidas, (ii) arrastran pobres antecedentes académicos del nivel secundario, (iv) trabajan, (v) proceden de la ciudad Córdoba Capital.

6. Apéndice I

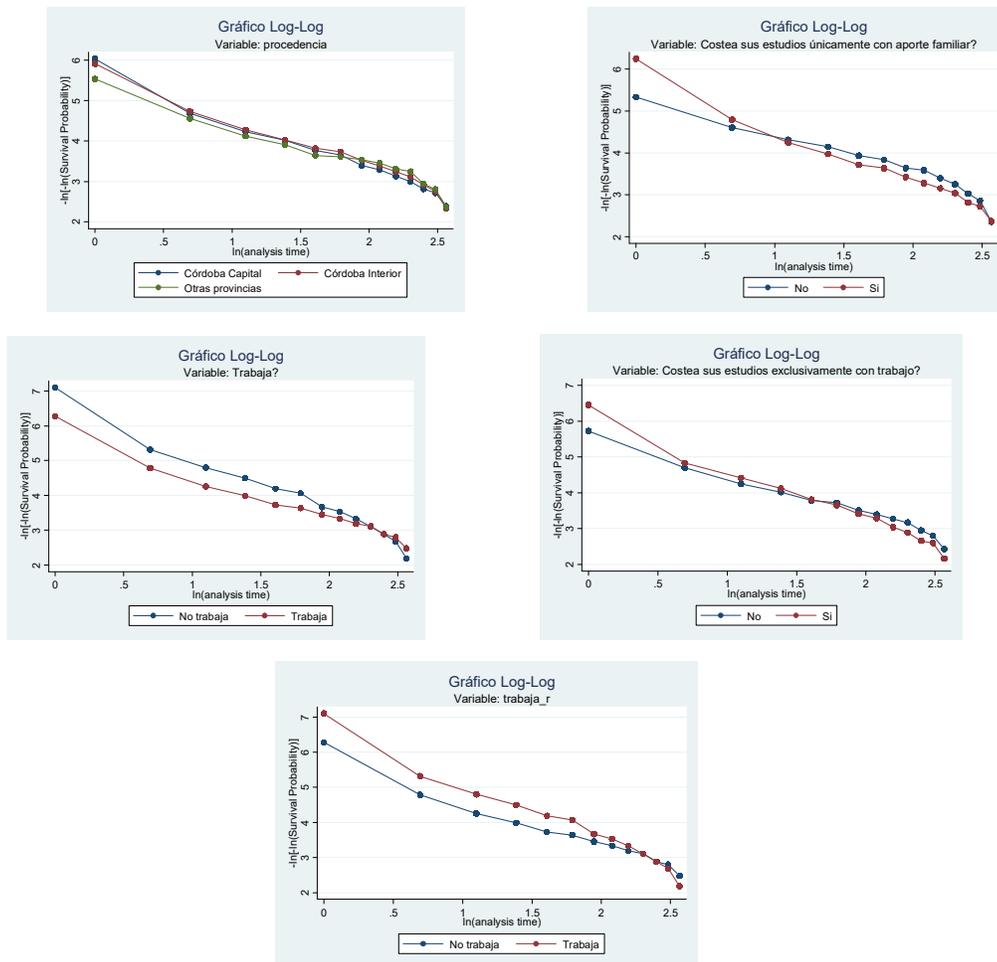


Figura 1. Análisis del supuesto de riesgo proporcional para las covariables

7. Referencias

- Cabrera, A., Nora, A. y Castañeda, M. (1993). "College Persistence: Structural Equations Modelling Test of Integrated Model of Student Retention". *Journal of Higher Education*: 64(2), 123-320.
- Cox, D. R. (1972). "Regression models and life tables". *Journal of Royal Statistical Society*, B 34, 187-220.
- Díaz, C. J (2009). "Factores de Deserción Estudiantil en Ingeniería: Una Aplicación de Modelos de Duración". Universidad Católica de la Santísima Concepción. Chile. *Información Tecnológica*. Vol 20(5), 129-145.
- Giuliodori, R., Gertel, H., Casini, R. y Gonzalez, M. (2007). "¿Qué factores influyen en el Rendimiento Académico de los Estudiantes Universitarios de Argentina?". Facultad de Ciencias Económicas. U.N.C. Proyecto SeCyT.
- Goldenhersch, H. et al. (2002). "Análisis de la deserción de estudiantes en la Facultad de Ciencias Económicas – U.N.C". Proyecto SeCyT.
- Grambsch, P. M. y Therneau, T. M. (1994). "Proportional hazards tests and diagnostics based on weighted residuals". *Biometrika* 81: 515–526.
- Lancaster, T. (1992). "The econometric analysis of transition data". Brown University. Cambridge University Press, p 364.
- Pascarella, E. y Chapman D. (1983). "A multi-institutional, path analytic validation of Tinto's model of college withdrawal". *American Educational Research Journal*, 20(1), 87-102.
- Taucher, E. (1999). "Análisis de supervivencia". *Bioestadística*. Santiago: Editorial Universitaria. 257 – 266.