

ANÁLISIS DEL RENDIMIENTO ACADEMICO MEDIANTE UN MODELO LOGIT*

ACADEMIC PERFORMANCE ANALYSIS THROUGH A LOGIT MODEL

María del Carmen Ibarra¹, Juan Carlos Michalus²

¹ Universidad Nacional de Misiones, Departamento de Matemática. Oberá, Misiones, Argentina.

² Universidad Nacional de Misiones, Departamento de Ingeniería Industrial. Oberá, Misiones, Argentina.

RESUMEN

Este trabajo analiza el rendimiento académico de los estudiantes de la Facultad de Ingeniería de la Universidad Nacional de Misiones; la población objetivo está conformada por los alumnos de las cohortes 1999 a 2003 (589 estudiantes). Se define al rendimiento académico como el promedio de materias aprobadas anualmente y mediante la técnica estadística multivariada de Regresión Logística, se determina la incidencia que tienen diferentes factores de índole personal, socioeconómica y académica.

Los resultados obtenidos permiten concluir que las variables significativas del rendimiento académico son: el promedio de calificaciones del nivel medio, el tipo de Institución donde cursó estos estudios y el número de asignaturas aprobadas en el primer año de carrera, siendo este último factor el más relevante, destacando la importancia de esta primera etapa de la carrera en los posteriores resultados académicos del estudiante.

Palabras clave: Análisis Multivariado, Estudiante universitario, Rendimiento Académico, Regresión Logística.

ABSTRACT

The purpose of this work is to analyse the determining factors which influence students' performance at university. The research has been carried out on five (5) engineering students' cohorts (1999-2003) from Universidad Nacional de Misiones (UNaM) and includes 589 students.

The academic performance is defined as average subjects approved annually. By means of the Logistic Regression technique, we determine the impact of different personal, socioeconomic and academic factors.

The main conclusion that can be drawn is that students' performance, is related to the grade point average (GPA) at high school, the kind of high school (public or private) students had attended, and the number of passing subjects in their first year at university. The latter being the most important factor, emphasizing the importance of this first stage at the University in the student's academic performance.

Keywords: Multivariate analysis - College student - students performance - logistic regression

* El presente artículo corresponde a avances del proyecto de investigación: "Análisis de factores incidentes en el desempeño académico y graduación de estudiantes universitarios", en el que actualmente trabajan los autores.

Autor para correspondencia: ibarra@fio.unam.edu.ar

Recibido: 08.06.2010 Aceptado: 15.10. 2010

INTRODUCCIÓN

En trabajos anteriores¹, los autores han realizado el análisis del Rendimiento Académico de los estudiantes de la FI – UNaM, mediante la utilización de la Técnica de Regresión Lineal Múltiple; tomando como población 594 alumnos, pertenecientes a las cohortes 1999 a 2003. Considerando como Rendimiento Académico el cociente entre el número de materias aprobadas y los años de permanencia en la Institución y como factores: características del entorno socioeconómico, condiciones académicas del nivel medio y número de materias aprobadas y promedio del primer año de carrera, concluyen que son factores determinantes del Rendimiento: el promedio general del nivel medio (a mayor promedio corresponde mejor rendimiento), el tipo de Institución donde cursó dichos estudios (mejor rendimiento para estudiantes provenientes de establecimientos privados) y el número de asignaturas aprobadas en el primer año de carrera (a mayor número de materias corresponde mejor rendimiento); siendo esta variable la más relevante del modelo.

Se sintetizan a continuación los resultados y conclusiones de algunos trabajos considerados relevantes para la problemática en cuestión, en los cuales se ha utilizado específicamente la Regresión Logística para analizar el Rendimiento Académico.

Rodríguez Fontes, Díaz Rodríguez, Moreno Lazo & Bacallao Gallestey (2000), tomando como muestra 114 ingresantes a la carrera de medicina (curso 1991 – 1992) en la Facultad de Ciencias Médicas Victoria de Girón (Cuba), analizaron la dicotomía éxito / fracaso académico, considerando el éxito como la obtención - durante el 1º año de carrera – de un promedio no inferior a 4 (sobre un máximo de 5). Concluyeron que el Índice Académico Preuniversitario (calificaciones de los últimos ciclos de este nivel) es el predictor más relevante y que el puntaje obtenido en el Examen de Ingreso no es relevante a la hora de predecir el desempeño en el 1º año de carrera.

García Jiménez, Alvarado Izquierdo & Jiménez Blanco (2000), realizaron un estudio sobre alumnos de primer año de Psicología de la Universidad Complutense de Madrid, España. Por un lado utilizan la técnica de Regresión Múltiple para analizar el rendimiento académico y por otro lado, la Regresión Logística para predecir el éxito / fracaso académico, entendido en este caso como la aprobación (o no) de una asignatura del 1º ciclo lectivo; en ambos casos concluyeron que son determinantes el promedio de calificaciones del nivel medio (bachillerato), la participación y asistencia a clases. También destacaron que la capacidad de predicción del Rendimiento Académico de la Regresión Logística es sustancialmente superior al de la Regresión Lineal.

Vélez van Meerbeke & Roa González (2005), realizaron un estudio sobre los ingresantes 2003 a la Facultad de Medicina de la Universidad del Rosario (privada) en Bogotá, Colombia. Definieron al Rendimiento Académico en términos de éxito / fracaso, este último entendido como la pérdida de materias o el abandono de los estudios; utilizaron la Técnica de Regresión Logística para predecir esta variable y concluyeron que el éxito (fracaso) está asociado principalmente al desempeño académico en el primer semestre de la carrera.

Di Gresia & Porto (2004), enfocaron su análisis en los logros académicos de los estudiantes de la cohorte 2000 de la Facultad de Ciencias Económicas de la Universidad Nacional de La Plata; mediante un Modelo Logit, analizaron la probabilidad que tiene un estudiante de no aprobar ninguna materia luego de dos años de permanencia en el sistema, y encontraron que dicha probabilidad es más elevada (alrededor del 76%) para un estudiante varón, casado, nacido y residente en La Plata y que trabaja 30 horas a la semana; mientras que dicho riesgo es menor

1 Ibarra & Michalus: “El Rendimiento Académico y sus factores determinantes – Aplicación a la Facultad de Ingeniería de la Universidad Nacional de Misiones” (2007); “Dos miradas al desempeño académico: Regresión Lineal / Regresión Logística” (2008); “Análisis del rendimiento académico mediante Modelos de Regresión” (2009).

(con probabilidad del 44%) para una mujer, soltera, no nacida en La Plata pero residente allí y que no trabaja.

MATERIAL Y METODOS

Caracterización de la población

La población estudiada consiste en cinco cohortes (1999 a 2003) de las cuatro carreras de grado que se dictan en la Institución²; totalizando 1183 alumnos, 594 de los cuales han abandonado los estudios³ (indicando un índice de deserción del orden del 50%). Queda así la población objetivo conformada por 589 estudiantes, 44 de los cuales son mujeres (7,47%) y 6 extranjeros (1%)⁴.

Definición de variables

El Modelo Logit es apropiado en aquellas situaciones donde se analiza una única variable dependiente categórica ó nominal y varias variables independientes.

La variable dependiente bajo estudio, en este caso, es el rendimiento académico definido desde un punto de vista cuantitativo, como el promedio de materias aprobadas anualmente, siguiendo a autores como Di Gresia, Porto & Ripani (2002) y Fazio (2004)⁵. El rendimiento académico definido de esta manera puede ser considerado un indicador del éxito / fracaso académico; así un alumno con un elevado número de asignaturas aprobadas da cuenta de un buen desempeño, mientras que un escaso número de materias estaría sugiriendo un bajo desempeño académico. Cuando la variable dependiente es dicótoma, se necesita utilizar una variable ficticia para definirla; dicha variable consiste en asignar dos valores (0 y 1) a la variable dependiente, según sea su categoría, de allí la denominación de variables categóricas. En este estudio se asigna a la variable dependiente – rendimiento académico (RA) – dos categorías, según el número promedio de asignaturas anuales que haya aprobado el estudiante; si ha aprobado 5 (cinco) ó más materias, le corresponde la categoría “1” y si ha aprobado menos de 5 (cinco) materias, se le asigna la categoría “0”; esto se indica en la siguiente expresión:

$$RA = \begin{cases} 1 & RA \geq 5 \\ 0 & RA < 5 \end{cases} \quad (1)$$

El valor de corte se ha establecido en 5 (cinco), ya que un estudiante que apruebe al menos ese número de materias al año, podría graduarse en un plazo aproximado de 7 (siete) años, lapso de tiempo razonable para una carrera de ingeniería.

Las variables independientes (factores) considerados en este estudio y cuyo grado de incidencia sobre el rendimiento se pretende determinar, pueden clasificarse en tres categorías:

- Condiciones personales y socioeconómicas (género, nivel instrucción de los padres)
- Condiciones académicas previas (titulación, promedio, tipo de establecimiento y localidad donde realizó estudios de nivel medio)
- Desempeño en el primer año de carrera (número de materias aprobadas y promedio).

2 Las carreras mencionadas son: Ingeniería Electromecánica, Electrónica, Civil e Industrial

3 Se considera que el alumno ha abandonado cuando no ha realizado actividad académica alguna en los dos últimos años, contados a partir de la fecha de relevamiento de los datos (fundamentado en el Art 50 de la Ley Nacional de Educación Superior 24.521/95).

4 Los datos están actualizados a febrero del año 2007.

5 Esta forma de definir el rendimiento se fundamenta en la Ley Nacional de Educación Superior 24.521/95, que en su Art. 50 establece como condición para que un alumno sea considerado regular, que haya aprobado al menos dos materias anuales.

A continuación se especifica cada una de ellas:

- Género: variable cualitativa (mujer / varón)
- Nivel de instrucción de los padres: variable cuantitativa que busca aproximar la condición socioeconómica del entorno familiar.
- Promedio de calificaciones del nivel medio (Promedio NM): variable cuantitativa
- Tipo de Institución donde cursó estudios de nivel medio (Institución NM): variable cualitativa que distingue si el establecimiento es de gestión pública ó privada.
- Localidad donde radica el establecimiento en que el estudiante cursó estudios de nivel medio (Localidad NM): variable cuantitativa que indica la densidad poblacional de la localidad⁶.
- Titulación obtenida en el nivel medio: variable cualitativa que hace referencia a la titulación con que egresa el estudiante, diferenciando entre técnicos (en sus diferentes especialidades) y otras titulaciones.
- Número de materias aprobadas en el primer año de carrera.
- Promedio de calificaciones del primer año de carrera.

METODOLOGÍA

El trabajo propone el análisis del rendimiento académico mediante la utilización de una de las técnicas del Análisis Multivariante, la Regresión Logística.

A continuación se expondrán los conceptos básicos acerca de la metodología empleada, para una descripción más acabada del tema, se puede recurrir, entre otros, a los siguientes textos, Hair *et al.* (1999), Pérez (2004), Uriel & Aldás (2005), Jonhson (2000).

En el caso de una variable dependiente categórica y variables independientes métricas (o factibles de ser transformadas mediante variables ficticias) son apropiadas las técnicas multivariantes de Análisis Discriminante ó Regresión Logística, se ha optado por esta última ya que el Análisis Discriminante requiere el cumplimiento de ciertos supuestos estrictos de normalidad y homocedasticidad de las variables, mientras que la Regresión Logística es más robusta en caso de que no se verifiquen tales supuestos y además sus coeficientes y estadísticos permiten un análisis e interpretación muy similares a los de la Regresión Lineal Múltiple (Hair *et al.*, 1999).

El modelo Logit permite calcular para cada entidad de la población, la probabilidad de pertenecer a una u otra de las categorías establecidas para la variable dependiente, si la probabilidad es alta se concluye que el evento tiene elevadas posibilidades de ocurrir, caso contrario, con valores reducido (cerca de cero) de probabilidad, las posibilidades de ocurrencia son prácticamente nulas; el valor de corte es 0,50.

Para un vector $\mathbf{x} = (x_1, x_2, \dots, x_n)$ de variables independientes, la probabilidad de ocurrencia del evento se calcula mediante una probabilidad condicional, mediante la expresión:

$$p(y=1|x) = \frac{e^{\beta_0 + \beta_i X}}{1 + e^{\beta_0 + \beta_i X}} \quad (2)$$

β_0 : término independiente

$\beta_i \mathbf{x} = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n$: combinación lineal de variables independientes

La ecuación (2) representa la probabilidad de ocurrencia del evento, dadas las características determinadas por las variables independientes (x_1, x_2, \dots, x_n) para cada sujeto de la población;

⁶ Fuente: Censo Nacional de Población, Hogares y Viviendas 2001, Instituto Nacional de Estadísticas y Censos (INDEC).

la cual se obtiene mediante una expresión que involucra funciones exponenciales de base “e” (2,7182...); se observa que el denominador de la ecuación es siempre mayor que el numerador, de forma que resulta en todos los casos un valor siempre positivo y menor que la unidad.

A partir de la ecuación (2) es posible plantear el cociente $\frac{p}{1-p}$:

$$\frac{p}{1-p} = \frac{e^{\beta_0 + \beta_i X}}{1 + e^{\beta_0 + \beta_i X}} = \frac{e^{\beta_0 + \beta_i X}}{1 + e^{\beta_0 + \beta_i X}} = \frac{e^{\beta_0 + \beta_i X}}{1 + e^{\beta_0 + \beta_i X}} \left(1 + e^{\beta_0 + \beta_i X}\right) = e^{\beta_0 + \beta_i X}$$

La Transformación Logit se define como el logaritmo natural (de base “e”) del cociente anterior, de manera que resulta:

$$\text{Ln}\left(\frac{p}{1-p}\right) = \text{Ln}\left(e^{\beta_0 + \beta_i X}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n \quad (3)$$

Como se aprecia de la ecuación (3) la Transformación Logit es lineal en los parámetros del modelo, de manera que permite realizar análisis muy similares a los de la Regresión Lineal; un coeficiente positivo aumenta la probabilidad de ocurrencia del evento, en tanto que uno con signo negativo la disminuye.

Para analizar la bondad de ajuste de un modelo Logit, se deben considerar: el *Odds Ratio*, el Estadístico de Wald, el Valor de Verosimilitud, el Coeficiente de Determinación y el Test de Hosmer y Lemeshow; los dos primeros hacen referencia al nivel de significación individual de cada variable independiente, mientras que los demás dan cuenta del ajuste global del modelo. A continuación se hace una breve referencia a cada uno de ellos siguiendo a Hair *et al.* (1999) y Johnson (2000).

- *Odds Ratio* (razón de posibilidades): siendo β_i el coeficiente correspondiente a la variable x_i , se define como *Odds Ratio* a la exponencial $\text{Exp}(\beta_i)$, un signo positivo en el exponente provoca un aumento en la probabilidad de ocurrencia del evento; un signo negativo reduce dicha probabilidad y un coeficiente cercano a cero da lugar a un valor próximo a la unidad, que prácticamente no incide en la probabilidad de ocurrencia (o no) del evento.
- Estadístico de Wald: hace referencia a la significación individual de cada variable y su interpretación es muy similar a la del Estadístico t para la Regresión Lineal Múltiple. Si el nivel de significación asociado es menor que el seleccionado (α), la variable en cuestión es relevante y debe ser tenida en cuenta en el modelo.
- Valor de Verosimilitud (-2LL): se obtiene como dos veces el logaritmo de la verosimilitud con signo contrario; un buen ajuste del modelo proporciona valores pequeños; un modelo ideal tendría una verosimilitud de uno, con lo cual su Valor de Verosimilitud sería nulo. Esta medida de ajuste global, compara las probabilidades predichas por el modelo con las observadas y mediante una prueba Chi - Cuadrado, indica si existe ó no diferencia significativa en la reducción del (-2LL) entre el modelo inicial y el final. Este contraste es equivalente al Estadístico F en un modelo de Regresión Lineal Múltiple.
- Coeficiente R^2_{logit} : este coeficiente puede asumir valores entre cero y uno, cuanto más cercano a la unidad se encuentre, tanto mejor será el ajuste del modelo. Esta medida está directamente relacionada con el Valor de Verosimilitud, ya que también mide el grado de reducción del mismo entre los modelos inicial y final.

- Test de Hosmer y Lemeshow: es la medida final para valorar el ajuste del modelo; analiza mediante una prueba Chi – Cuadrado, si existe ó no, diferencia significativa entre los valores observados y los predichos por el modelo para la variable dependiente.

RESULTADOS

Obtención del Modelo

Los resultados obtenidos para este modelo se resumen en las tablas 1 y 2. En la tabla 1 se presentan los coeficientes β correspondientes a la expresión (3), y los estadísticos asociados a cada uno de ellos, mientras que en la tabla 2 se presentan las medidas de resumen del modelo.

Tabla 1: Resultados del modelo de Regresión Logística

Variables Independientes	Coefficiente β	Estadístico de Wald	Nivel de Significación (p)	Exp (B) Odds Ratio
Género	-1,759	5,389	0,020	0,172
Instrucción padres	-0,012	0,354	0,552	0,988
Promedio NM	0,813	8,849	0,003	2,254
Institución NM	0,957	8,461	0,004	2,603
Localidad NM	-0,001	0,503	0,478	0,999
Titulación NM	-0,497	2,410	0,121	0,609
Materias 1° año	0,663	41,812	0,000	1,941
Promedio 1° año	0,278	2,277	0,131	1,320
Constante	-13,180	32,018	0,000	0,000

Tabla 2: Resultados del modelo de Regresión Logística

- 2LL Inicial	- 2LL Final	Test Global coeficientes	R ² Logit	Porcentaje Correcto	Test Hosmer Lemeshow
454,93	287,53	$\chi^2 = 167,40$ $p = 0,000$	0,368	85,3	$\chi^2 = 7,646$ $p = 0,469$

Los indicadores generales del modelo (tabla 2) permiten concluir que el conjunto de factores seleccionado contribuye significativamente a explicar la probabilidad de conseguir un buen rendimiento académico, lo cual también se ve reflejado en el importante porcentaje de predicción correcto alcanzado (del orden del 85%).

Para este modelo resultan estadísticamente significativas cuatro variables: el promedio general obtenido en el nivel medio, el tipo de Institución donde cursó estos estudios, el número de materias aprobadas en el primer año de carrera y el género. Resulta así que tienen mayor probabilidad de lograr un buen rendimiento académico aquellos estudiantes de género femenino, provenientes de establecimientos privados, con elevadas calificaciones en el nivel medio y que aprueben mayor cantidad de materias en el primer año de carrera.

ANÁLISIS DE RESULTADOS

En función de los estadísticos obtenidos para cada variable (Cuadro 1), caben las siguientes consideraciones:

- Para la variable que hace referencia al género del estudiante, el coeficiente (β) resulta negativo y su *Odds Ratio* menor a la unidad - $\text{Exp}(-1,759) = 0,172$ -, lo cual indica que a igualdad de todas las demás condiciones, los estudiantes varones tienen 5,81 veces menos probabilidad de conseguir un buen rendimiento académico (aprobar al menos 5 materias anuales) que

sus pares mujeres. Este resultado es coincidente con el encontrado mediante la metodología de Regresión Lineal, aportando ambos modelos evidencia a favor de un mejor desempeño académico para estudiantes de género femenino.

- La variable que hace referencia al promedio de calificaciones del nivel medio resulta con un coeficiente positivo y el *Odds Ratio* es mayor a la unidad – $\text{Exp}(0,813) = 2,254$ lo cual indica que por cada punto en que se incrementa el promedio, el estudiante tiene 2,25 veces más probabilidad de conseguir un buen rendimiento académico. Por ejemplo, si se compara la situación de dos alumnos, uno de ellos con un promedio de 6,00 y otro de 8,50, el último tiene 7,63 veces más probabilidad de conseguir un buen desempeño en su carrera; a igualdad de todas las demás condiciones.
- La variable que hace referencia al tipo de establecimiento donde cursó el alumno sus estudios de nivel medio resulta con coeficiente positivo y *Odds Ratio* es mayor a la unidad – $\text{Exp}(0,957) = 2,603$ – esto indica que los estudiantes provenientes de instituciones privadas tienen 2,60 veces más probabilidad de conseguir un buen rendimiento académico que los egresados de escuelas públicas, a igualdad de todas las demás condiciones.
- La variable que hace referencia al número de asignaturas aprobadas durante el primer año de carrera presenta coeficiente positivo y *Odds Ratio* mayor a la unidad – $\text{Exp}(0,663) = 1,941$ –, indicando así que por cada materia que consiga aprobar el alumno en esta etapa se incrementa en 1,94 veces la probabilidad de conseguir un buen rendimiento académico a lo largo de su carrera. Por ejemplo, si se compara la situación de dos alumnos, de manera que uno de ellos haya aprobado tres materias en su primer ciclo lectivo y el segundo tenga seis materias aprobadas, este último tendrá 7,31 veces más probabilidad de conseguir buen desempeño en su carrera; a igualdad de todas las demás condiciones.

Si bien las demás variables no han resultado significativas, su análisis permite extraer interesantes consideraciones:

- Las variables que hacen referencia al nivel de instrucción de los padres y a la densidad poblacional de la localidad en la cual cursó el alumno sus estudios secundarios, presentan *Odds Ratio* muy próximos a la unidad (0,988 y 0,999, respectivamente), de manera que prácticamente no ejercen influencia en la probabilidad de conseguir (ó no) un buen rendimiento académico.
- La variable que hace referencia a la titulación con que egresa el alumno del secundario, resulta con signo y negativo y *Odds Ratio* menor a la unidad – $\text{Exp}(-0,497) = 0,609$ – indicando que los estudiantes provenientes de escuelas técnicas tienen 1,64 veces menos probabilidad de conseguir un buen rendimiento académico que quienes provienen de instituciones con otro tipo de orientación; a igualdad de todas las demás condiciones. Este resultado también coincide con el encontrado para el modelo de Regresión Lineal.
- La variable que hace referencia al promedio de calificaciones en el primer año de carrera resulta con coeficiente positivo y *Odds Ratio* mayor a la unidad – $\text{Exp}(0,278) = 1,320$ – indicando que por cada punto que se incrementa el promedio obtenido en esta etapa aumenta 1,32 veces la probabilidad de conseguir buen rendimiento académico en la carrera. Por ejemplo si se compara la situación de dos estudiantes, el primero con un promedio de 6,00 y el segundo con 8,50, este último tiene 2 veces más probabilidad de lograr buen desempeño, a igualdad de todas las demás condiciones. También en el modelo de Regresión Lineal se había encontrado que mayores promedios en el primer año de carrera favorecía el rendimiento académico; de manera que ambos modelos aportan evidencia a favor de la relevancia del desempeño del alumno en la primera etapa de carrera. Los valores del Estadístico de Wald y del Nivel de Significación, indican que esta variable es la más relevante del modelo.

REFERENCIAS

- Bacallao, C., Parapar de la Riestra, M., Roque, M., & Bacallao, J. (2004).** Árboles de regresión y otras opciones metodológicas aplicadas a la predicción del rendimiento académico. *Revista de Educación Médica Superior*, Vol. 18, N° 3.
- Debera, L.; Machado, A., & Nalbarte, L. (2004).** Trayectoria y desempeño escolar de los estudiantes de la Facultad de Ciencias Económicas y Administración. Instituto de Estadística, Facultad de Ciencias Económicas y Administración, Universidad de la República, Uruguay.
- Di Gresia, L., & Porto, A. (2004).** Dinámica del desempeño académico. Departamento de Economía, Universidad Nacional de La Plata.
- Di Gresia, L., & Porto, A., & Ripani, L. (2002).** Rendimiento de los estudiantes de las Universidades Públicas Argentinas. Departamento de Economía, Universidad Nacional de La Plata. Argentina
- Di Gresia, L. (2007).** Rendimiento académico universitario. Trabajo de tesis doctoral, Doctorado en Economía, Universidad Nacional de La Plata. Argentina
- Díaz Martínez, L., & Toloza G, C. (2007).** Los indicadores de selección para el ingreso a la universidad y su valor para estimar el rendimiento académico en el primer semestre. *Ciencia e Investigación Médica Estudiantil Latinoamericana – CIMEL*. 12(2); 59 – 65.
- Fazio, M. (2004).** Incidencia de las horas trabajadas en el rendimiento académico de estudiantes universitarios argentinos. Trabajo de tesis de Maestría, Maestría en Economía, Universidad Nacional de La Plata. Argentina.
- Ferreyra, M. (2007).** Determinantes del desempeño universitario: Efectos heterogéneos en un modelo censurado. Trabajo de tesis de Maestría, Maestría en Economía, Universidad Nacional de La Plata. Argentina
- Foio, M. & Espínola, A. (2004).** Seguimiento académico de los ingresantes 2002. Universidad Nacional del Nordeste.
- García J, M., Alvarado I, J.; & Jiménez, A. (2000).** La predicción del rendimiento académico: Regresión Lineal versus Regresión Logística. *Psicothema*, 12(2); 248 - 252.
- Giovagnoli, P. (2002).** Determinantes de la deserción y graduación universitaria: Una aplicación utilizando modelos de duración. Trabajo de tesis de Maestría, Maestría en Economía, Universidad Nacional de La Plata. Argentina
- Guzmán G, C. & Serrano S, O. (2006).** Detrás del concurso de selección – Un análisis de los factores que inciden en el ingreso a la licenciatura de la UNAM. Universidad Nacional Autónoma de México.
- Hair, J.; Anderson, R.; Tatham, R., & Black, W. (1999).** Análisis Multivariante. Prentice Hall, Madrid. 5ª ed.
- Ibarra, M. & Michalus, J. (2007).** El Rendimiento Académico y sus factores determinantes – Aplicación a la Facultad de Ingeniería de la Universidad Nacional de Misiones. XX ENDIO, Mar del Plata, Argentina.
- Ibarra, M., & Michalus, J. (2008).** Dos miradas al desempeño académico: Regresión Lineal / Regresión Logística. EMCI XIV, Mendoza, Argentina.

Ibarra, M., & Michalus, J. (2009). Análisis del rendimiento académico mediante Modelos de Regresión. 3º Congreso Argentino de Ingeniería Industrial COINI 2009, Misiones, Argentina.

Johnson, D. (1998). Modelos Multivariados Aplicados al Análisis de datos. International Thomson Editores, México.

Maradona, G., & Calderón M. (2004). Una aplicación del enfoque de la función de producción en educación. Revista de Economía y Estadística, Universidad Nacional de Córdoba, XLII.

Montgomery, D.; Peck, E. & Vining, G. (2001). Introducción al Análisis de Regresión Lineal. Compañía Editorial Continental, México. 3ª ed.

Pérez, C. (2004). Técnicas de Análisis Multivariante de Datos. Aplicaciones con SPSS. Prentice Hall, Madrid. 646 pp.

Porto, A. & Di Gresia, L. (2000). Características y Rendimiento de estudiantes universitarios. El caso de la Facultad de Ciencias Económicas de la Universidad Nacional de La Plata. Departamento de Economía, Universidad Nacional de La Plata.

Porto, A.; Di Gresia, L., & López A, M. (2004). Mecanismos de admisión a la Universidad y rendimiento de los estudiantes. Departamento de Economía, Universidad Nacional de La Plata.

Rodríguez F, R.; Díaz R, P.; Moreno L, M., & Bacallao G, J. (2000). Capacidad predictiva de varios indicadores de selección para el ingreso a la carrera de medicina. Revista Cubana de Educación Médica Superior, 14(2).

Uriel, E., & Aldás, J. (2005). Análisis Multivariante aplicado. International Thomson Editores, Madrid.

Valdivieso, M.; Monar, K. & Granda, M. (2004). Análisis del rendimiento de los estudiantes de ESPOL- 2002. Revista Tecnológica, 17(1); 213 – 218.

Vélez van Meerbeke, A. & Roa G, C. (2005). Factores asociados al rendimiento académico en estudiantes de medicina. *Educación Médica*, 8(2).

