

Un Análisis de las generaciones 2000-2008 de la Facultad de Físico-Matemáticas de la Benemérita Universidad Autónoma de Puebla.

Ana Luisa Nieto Méndez¹, Hortensia Reyes Cervantes², Flaviano Godínez Jaimes³, Francisco Tajonar Sanabria⁴ y Víctor Hugo Vázquez Guevara⁵.

Resumen: En este trabajo se analizaron las generaciones de estudiantes de la Facultad de Físico-Matemáticas; de las carreras de Matemáticas y Matemáticas Aplicadas, de 2000-2008. Hay una deserción del 80%, sólo se revisaron algunas materias básicas de los primeros dos años ya que se consideran estas materias como representativas de las carreras debido a su grado de dificultad y significancia a lo largo de la carrera. Los alumnos se clasificaron en tres tipos: desertores al inicio de los cursos, egresados y titulados. Se observó que el promedio semestral y final son dos indicadores para saber si un alumno termina o no la carrera y más frecuentemente las mujeres terminan que los hombres.

Palabras clave: materias básicas, dificultades en el estudio, regresión logística.

Abstract: In this work the generations of students of the Faculty of Physics and Mathematics were analyzed; degree Mathematics and Applied Mathematics, 2000-2008. There is 80% dropout, just some basic materials of the first two years were reviewed and considered these matters as representative of degree due to its difficulty and significance throughout the degree. Students were classified into three types: Drop out to start the courses, graduates and graduates. It was observed that the semi-annual and final average are two indicators as to whether or not a student finishes the degree and more often women than men end up.

Keywords: basic materials, difficulty in the study, logistic regression.

Justificación

Se ha encontrado que la eficiencia terminal es un indicador importante en las metas

¹FCFM-BUAP, Puebla-México, ana_mate89@hotmail.com

²FCFM-BUAP, Puebla-México, hreyes@fcfm.buap.mx

³UAG, Guerrero-México, fgodinezj@gmail.com

⁴FCFM-BUAP, Puebla-México, ftajonar@fcfm.buap.mx

⁵FCFM-BUAP, Puebla-México, vvazquez@fcfm.buap.mx

y objetivos que tienen las universidades hoy en día, aunque se ha visto que con esto no se puede juzgar la calidad de las instituciones y el aprendizaje de los estudiantes. [3]. Evaluar una carrera de ciencias usando los mismos criterios usados para evaluar otras carreras de ciencias pasa por un proceso más complejo. Así, evaluar en términos de su eficiencia terminal (número de egresados y titulados) produce resultados insatisfactorios.

En ciencias los alumnos se enfrentan a un ambiente escolar diferente al que estaban acostumbrados. Ahora, los alumnos deben de ser responsables por sí mismos, de sus logros académicos pues se esmeran en sus tareas, estudios y exámenes de las materias que cursan. También los alumnos empiezan a conocer las matemáticas desde sus propiedades elementales, axiomas y se aprende a manejar la lógica para demostrar sus resultados

1 Introducción

Reconocer al alumno como el eje de la atención de los programas académicos de las Instituciones de Educación Superiores (IES) mexicanas, implica considerar su existencia y su transformación a lo largo de un espacio temporal que comprende tres momentos claramente identificables: la etapa anterior a su ingreso y hasta su integración a la institución educativa (primer momento), el transcurso de los estudios regulares durante su permanencia en la institución (segundo), y el lapso que se observa entre el momento de egreso y la obtención de un título que acredita sus aprendizajes y su calidad como profesionista en alguna o algunas áreas del conocimiento (tercero). En ese extenso recorrido, las instituciones educativas prevén o debieran prever un conjunto de actividades dirigidas a los alumnos, con la aspiración de atenderlos de manera integral, es decir, ocuparse de los estudiantes desde antes de su ingreso, con programas de información y orientación efectivos y mecanismos de selección justos; y hasta después de su egreso, con mecanismos de apoyo para la inserción laboral [2].

Los modelos de regresión logística son una herramienta que permite explicar el comportamiento de una variable respuesta binaria mediante una o varias variables independientes explicativas de naturaleza cuantitativa y/o cualitativa. Los modelos de respuesta discreta son un caso particular de los modelos lineales generalizados formulados por Nelder y Wedderburn en 1972, al igual que los modelos de regresión lineal o el análisis de varianza[4].

2 Marco Teórico

Los estudios de egresados constituyen una de las estrategias apropiadas para retroalimentar los programas de formación de profesionales e investigadores en las instituciones de educación superior. Estos también son considerados como mecanismos poderosos de diagnóstico de la realidad, con el potencial de inducir en las instituciones la reflexión a fondo sobre sus fines y sus valores. Los resultados de estos estudios pueden asimismo, aportar elementos para redefinir el proyecto de desarrollo de aquellas instituciones que se mantienen alerta ante las nuevas necesidades sociales, permitiéndoles reconocer y asumir las nuevas formas de práctica profesional que se requieren para sustentar un proceso social menos inequitativo y dependiente. Los estudios de egresados pueden ser una herramienta básica para la definición de políticas en el nivel regional, estatal e incluso nacional y para el diseño de estrategias tendientes a propiciar el desarrollo y el fortalecimiento de todas las instituciones educativas del país [13].

En la Facultad de Ciencias Físico Matemáticas (FCFM) de la Benemerita Universidad Autónoma de Puebla (BUAP) existen algunos estudios acerca de los factores que influyen en la aprobación de la materia de Matemáticas Básicas y/o Matemáticas Elementales, tales como los siguientes:

Una aplicación de regresión lineal en el aprovechamiento de los alumnos de nuevo ingreso en el área de matemáticas de la FCFM, en el que se encontró que los alumnos no tiene buenos hábitos de estudio, pues la mayoría estudia muy pocas horas diariamente y esto influye en sus calificaciones que los alumnos obtienen[1]. Algunas variables se encontraron importantes, las cuales son:

- a) Si la carrera fue su primera opción de ingreso y
 - b) La confianza que tienen con el profesor para preguntarle sus dudas.
-
- Un estudio de uso de la regresión logística para estudiar la aprobación de la materia de Matemáticas Básicas de la FCFM en las generaciones 2010 y 2011. Se encontró que los alumnos tienen malos hábitos de estudio. Las dos generaciones se comportan de manera diferente: la mejor en aprobación es la 2011 y se debe a que estos alumnos van estudiando diariamente los conceptos vistos durante la clase, mientras que los alumnos de la generación 2010 deciden estudiar hasta la fecha del examen[7].

- Un Análisis estadístico de algunos factores que afectan el proceso de enseñanza aprendizaje en la FCFM, usando técnicas estadísticas multivariadas, donde encontró que uno de los principales factores fue el hecho de que la Licenciatura en Matemáticas no fuera su primera opción, ya que esto afecta en el desempeño académico, otro factor que influye es la atención que tiene el maestro en sus cursos, así como los factores económico, social y cultural, entre otros[5].
- El estudio Proceso de enseñanza aprendizaje, en el que encontró que el lugar de Procedencia es un indicador, ya que los alumnos que provienen de otros estados son los que más acreditan la materia de Matemáticas Básicas[6].
- Un estudio de identificación de factores que intervienen en la reprobación del curso de Matemáticas Básicas de la FCFM de la BUAP. El análisis se realizó del período Primavera 2007 a Otoño 2010, y encontró que los principales factores son:
 1. El profesor;
 2. La falta de asistencia a asesorías que compete tanto a estudiantes como a profesores;
 3. La literatura empleada en el curso.Se hace mención de estos trabajos, ya que son importantes sus resultados y son antecedentes para investigar a las poblaciones que forman a los egresados y titulados de la FCFM [11].

3 Metodología

Este es un análisis de tipo inferencial, previamente se ha hecho un análisis descriptivo de la información y se observó que un 20% es egresado, así en este estudio nos enfocaremos a determinar los factores que determinan la eficiencia terminal por medio de la regresión logística.

Es conocido por los profesores que dan esos cursos (ME, CD, CI, CDV, CIV) que los primeros dos años, los alumnos desertan más que en cualquier otro periodo. Los alumnos que ya se titularon tienen mejores promedios que los alumnos que todavía están tomando cursos, y los alumnos que desertan sus promedios son menores en general a los otros dos (egresados y en cursos). El promedio que los alumnos tienen en la preparatoria no es un indicador si los alumnos serán capaces de terminar la carrera. El promedio que tienen en tercer año de la carrera está en relación al

promedio final del alumno[14].

3.1 Descripción de la información

El Universo consta de 1047 alumnos inscritos de dos de la licenciaturas Matemáticas y Matemáticas Aplicadas que se ofrecen en la Facultad de Ciencias Físico Matemáticas de la Benemérita Universidad Autónoma de Puebla, la información se obtuvo de la Dirección de Administración Escolar por medio de sus departamentos de Computo y de Titulación de la BUAP y de la Dirección General de Profesiones (SEP). El cohorte generacional fue hecho al periodo verano 2013; la información se analizó en SPSS[15].

De la información obtenida de todos los alumnos de las dos licenciaturas, estos se clasificaron como alumnos desertados (aquellos alumnos que no se inscribieron en dos semestres consecutivos [16]), alumnos en cursos (aquellos que estan inscritos actualmente), y alumnos "egresados" (alumnos que han alcanzado el 100 por ciento de créditos requeridos para la titulación). Así los datos con los que se trabajaron fueron con los alumnos "en cursos" y "egresados"; el total de estos alumnos son 308. **Nota:** observar que el tiempo mínimo que se toma para un alumno en la universidad son 5 años y el máximo 8 años.

Código para la variables:

La variable respuesta fue egreso que tuvo dos valores: egresado ($Y=1$) y no egresado ($Y=0$). Las variables independientes consideradas fueron Carrera con valores Licenciaturas de matemáticas (MAT) y matemáticas aplicadas (LMA). Las materias que se analizaran son: Matemáticas Elementales (ME), Cálculo Diferencial (CD), Cálculo Integral (CI), Cálculo Diferencial en Varias Variables (CDV), Cálculo Integral en Varias Variables (CIV), Años Cursados (AÑOS), Promedio Primer año (PROM1), Promedio Segundo año (PROM2), Promedio Tercer Año (PROM3), Promedio final (PROMAC), Promedio Preparatoria o Bachiller (PROMPREP), Puntaje exámen de admisión (PUNTAJE) y sexo (SEXO).

Al trabajar la información en SPSS, se asignaron etiquetas a varias de las variables que se toman en consideración para obtener el modelo tal es el caso de "carrera", se etiquetó como 1 a MAT y como 2 a LMA, además en la variable "sexo" se tomo como 0 para masculino y 1 para femenino.

3.2 Regresión Logística

El Análisis de Regresión Logística (teoría obtenida de Berenson y Levine, 1996; Hosmer y Lemeshow, 1989; Mather, 1976; Soares y Bartman, 1983), es una técnica estadística que permite, a través del planteamiento de un modelo matemático, establecer una relación funcional entre la variable dependiente y las variables independientes o regresoras para el establecimiento de los parámetros del modelo matemático.

El modelo de regresión logística se deriva de la función matemática:

$$\pi(x) = \frac{e^{g(x)}}{1+e^{g(x)}} \quad \text{con } -\infty < g(x) < \infty$$

La función $\pi(x)$ toma valores desde cero hasta uno de una forma creciente y tiene forma sigmoidea. La variable dependiente o de respuesta es una variable dicotómica que toma valores de 0 ó 1, con k variables regresoras evaluadas en un tiempo inicial[12].

4 Análisis de Datos

Realizando la comparación de medias para la variable respuesta Y (categórica) y las variables explicativas (cuantitativas), obtuvimos que para la variable CD, no se rechaza, la hipótesis nula de igualdad de medias y por tanto ésta variable ya no se considerara como variable explicativa.

Para el caso de la prueba de igualdad de medias, menores a $\alpha = 0.05$, en las variables restantes se observan valores de significancia .

Se realizó selección de variables considerando las siguientes: ME, CI, CDV, CIV, AÑOS, PROM1, PROM2, PROM3, PROMAC, PROMPREP, PUNTAJE y SEXO, y utilizando en método "introducir" en SPSS; se obtiene la Tabla 1, así tomando un nivel significancia de 0.05 observamos que las variables que estarán en nuestro modelo final propuesto son ME, CIV y PROMAC pues tienen alta significancia.

Tabla 1: Variables en la ecuación; salida de SPSS

Variables en la ecuación		B	ET	Wald	gl	Sig	Exp(B)
Paso 3º	ME	-.396	.179	4.872	1	.029	.673
	CIV	.344	.144	5.682	1	.017	1.410
	PROMAC	2.695	.587	21.060	1	.000	14.8
	Constante	-21.074	4.386	23.085	1	.000	.000

4.1 Interpretación de los coeficientes

La interpretación de los coeficientes en un modelo de regresión logística comúnmente se realiza mediante cambios en la escala logit, a través de razones de odds condicionales. Sin embargo dichos coeficientes no tiene una interpretación directa en términos de probabilidad, que podría considerarse como la mas frecuente para la mayoría de los usuarios[10].

Otra forma de presentar los resultados del modelo de regresión logística es a través de una estimación ajustada de las probabilidades asociadas a cada covariable. El calculo de probabilidades ajustadas puede ser mas conveniente para una ausencia no experta en las nociones de estimación de parámetros de regresión y probabilidad asociada[9].

El signo de los coeficientes estimados de la función logística en la Tabla 1 da una explicación de las variables predictoras usadas, luego los signos de CIV y PROMAC son positivos y por consiguiente producen un crecimiento en la probabilidad para egresar. En cambio el signo de la variable ME es negativo con lo cual produce un decrecimiento en la misma probabilidad.

La razón de momios (RM) se obtuvo exponenciando la estimación del parámetro y corresponde a la columna $\text{Exp}(B)$ en la Tabla 1 por lo tanto:

- La RM par ME es 0.673, lo que indica que un incremento de la calificación de ME disminuye la probabilidad de que un alumno egrese de alguna de las licenciaturas (MAT y LMA).
- La RM para CIV es 1.416, esto indica que un incremento de la calificación de CIV incrementa la probabilidad de que un alumno egrese de alguna de las lilenciaturas (LMA y MAT).
- La RM para PROMAC es 14.8, lo que indica que un incremento en el PROMAC incrementa la probabilidad de que un alumno egrese de alguna de las licenciaturas (LMA y MAT).

4.2 Bondad de Ajuste

Del modelo obtenido se debe verificar que efectivamente este modelo se ajusta correctamente a los datos, para esto en SPSS se tiene la opción que al momento de realizar el análisis de regresión logística se obtenga en la misma salida el resultado de la Prueba de Hosmer y Lemeshow. En esta prueba la hipótesis nula se define por: no hay diferencias entre los valores observados y los valores pronosticados. En la Tabla 2, el valor calculado del estadístico de Hosmer y Lemeshow de bondad de ajuste para el modelo fue $c = 8.22$, con un p valor de 0.394, y por tanto podemos

concluir que no se rechaza la hipótesis nula.

Tabla 2: Prueba de Hosmer y Lemeshow; salida de SPSS

Prueba de Hosmer y Lemeshow			
Paso	Chi cuadrado	gl	Significancia
3	8.420	8	.394

En la Tabla 3 observamos la tabla de Clasificación tenemos que: 40% de los alumnos no egresados se pronosticaron correctamente y 52 se clasificaron incorrectamente. El 96% de los alumnos egresados se pronosticaron correctamente, y por tanto esto nos da un 85 de porcentaje global pronosticado correctamente. Con lo cual el modelo se puede considerar bueno.

Tabla 3: Tabla de clasificación; salida de SPSS

Tabla de Clasificación				
Observado		Pronosticado		
		0	1	Porcentaje Correcto
0		20	30	40.0%
1		8	202	96.2 %
		Porcentaje global		85.4%

El valor de corte es .500

El área bajo la curva COR es frecuentemente usada para probar la hipótesis nula: área=0.5, vs. Ha: área \leq 0.5. Según [8], si COR es mayor a 0.7 se puede decir que hay una aceptable discriminación. De la tabla 4 obtenemos el área de la curva COR igual a 0.776, con lo cual podemos decir que nuestro modelo tiene una buena discriminación, es decir clasificará de manera correcta a los alumnos.

Tabla 4: Área bajo la curva COR; salida de SPSS

Área bajo la curva				
Variables resultado de contraste: Probabilidad pronosticada				
Área	Error típ ^a	Sig ^{asintoticab}	IC asintótica al 95% Límite inferior	Límite superior
.776	.032	0.000	.713	.839

a. Bajo el supuesto no paramétrico

b. Ho: área verdadera = 0.5

5 Conclusiones

En este análisis, se estudiaron los factores para que un alumno egrese o no de alguna de las dos licenciaturas Matemáticas o Matemáticas Aplicadas, usando el modelo de regresión logística.

Los factores usados fueron Matemáticas Elementales (ME), Cálculo Diferencial (CD), Cálculo Integral (CI), Cálculo Diferencial en Varias (CDV), Cálculo Integral en varias (CIV), Años cursados (AÑOS), Promedio primer año (PROM1), Promedio segundo año (PROM2), Promedio tercer año (PROM3), Promedio final (PROMAC), Promedio Preparatoria o Bachiller (PROMPREP), Puntaje examen de admisión (PUNTAJE), Y, CARRERA, Sexo (SEXO).

El análisis concluye que las materias ME y CIV influyen en la probabilidad de que un alumno egrese o no. Mientras la variable PROMAC aumenta de manera muy significativa esta probabilidad. El modelo de regresión logística es usado para estimar si un alumno egresa o no, las pruebas de bondad de ajuste nos indican que el modelo obtenido es un buen modelo para poder clasificar u obtener las probabilidades de los alumno para egresar o no de las licenciaturas analizadas.

Bibliografía

- [1] Arenas Y., Una Aplicación de Regresión Lineal en el Aprovechamiento de los alumnos de Nuevo Ingreso en el Área de Matemáticas de la FCFM. Tesis de licenciatura, 2011.
- [2] Díaz J., "Análisis Descriptivo de los Egresados y Titulados de las Licenciaturas de Matemáticas y Matemáticas Aplicadas de las Generaciones 2000 a 2004", tesis de matemáticas FCFM-BUAP, 2013
- [3] Romo A., Fresan M., Los factores curriculares y académicos relacionados con el abandono y el rezago.
- [4] De la Fuente S., Regresión Logística, Facultad de Ciencias Económicas y Empresariales, UAM, Madrid, 2011.
- [5] Hernández S., Reyes H., Linares G., Análisis Estadístico de algunos factores que afectan el proceso de enseñanza aprendizaje en la FCFM-BUAP, usando técnicas estadísticas multivariadas., VI Encuentro Participación de la Mujer en la Ciencia, CIO, 2009.
- [6] Hernández S., Reyes H., Ibarra M., Linares G., Proceso de enseñanza aprendizaje, VII encuentro Participación de la mujer en la ciencia, CIO, 2010.

- [7] Hernández, S. Uso del modelo de Regresión Logística para estudiar la aprobación de la materia de Matemáticas Básicas de la FCFM en las generaciones 2010 y 2011., tesis de licenciatura, México, 2013.
- [8] Hosmer D., Lemeshow S., Applied Logistic Regression, Wiley, 2013.
- [9] Iglesias Cabo T., " Métodos de Bondad de ajuste en regresión logística", Trabajo fin de Máster, 2013.
- [10] López J., García J., Eventos por variable en Regresión Logística y Redes Bayesianas para Predecir Actitudes Emprendedoras, Universidad de Almeria, 2011.
- [11] Maldonado A., Identificación de los Factores que intervienen en la reprobación del curso de Matemáticas Básicas de la FCFM de la BUAP, Tesis de Licenciatura, México, 2012.
- [12] Muller M., Generalized Linear Models, Gemany, 2004.
- [13] Ontiveros I., Seguimiento de Egresados de la Licenciatura en Artes Visuales de la Escuela de Pintura, Escultura y Artesanías de la UJED, Tesis de Maestría, 2006.
- [14] Reyes J., Canizo J., Meza E., Herrera A., Cruz H., Nieto A., Godínez J.; Descripción de las generaciones 2000-2008 para los alumnos que desertan en dos licenciaturas de la FCFM-BUAP, Encuentro Participación de la Mujer en la ciencia, CIO 2014.
- [15] SPSS(2010), IBM SPSS Statistics 19 para Windows.
- [16] Tinto V., Una consideración de las teorías de la deserción estudiantil en la trayectoria escolar en la educación superior, México: ANUIES 1987.