



# Análisis del rendimiento académico mediante regresión logística y múltiple

Analysis of academic performance through logistic and multiple regression

Jeylin Meybelin Pérez Obregón<sup>1</sup>  
Tonys Romero Díaz<sup>2</sup>

## Resumen

Este artículo presenta un análisis del rendimiento académico del Instituto Nacional Santa Teresita en la disciplina de matemática durante el período 2011-2016 mediante la regresión logística y múltiple para determinar el comportamiento y relación que hay entre las calificaciones de la asignatura de matemática con lengua extranjera - lengua y literatura. Se utilizó un enfoque mixto, el aspecto cuantitativo permitió examinar las calificaciones de años consecutivos para crear los modelos matemáticos que permiten predecir resultados; y la metodología cualitativa sirvió para explorar y comparar opiniones de personas relacionadas al contexto. El modelo de regresión múltiple aproximó la calificación de matemática considerando las de lengua y literatura, y lengua extranjera -. Se aplicó el modelo de regresión logística, el cual permitió calcular el intervalo promedio de los estudiantes, considerando grado, sexo y año; siendo este modelo un complemento del anterior, ya que fortalece los resultados para su respectivo análisis, ambos modelos resultaron significativos y superaron los supuestos estadísticos, por tanto, se recomienda usarlos para mejorar el rendimiento del centro educativo.

**Palabras claves:** Rendimiento académico, regresión logística, regresión múltiple, predicciones, análisis.

## Abstract

This article presents an analysis of the academic performance of the Santa Teresita National Institute in the subject of mathematics during the period 2011-2016 through logistic and multiple regression to determine the behavior and relationship between the qualifications of the subject of mathematics with a foreign language - language and literature. A mixed approach was used, the quantitative aspect allowed examining the ratings of consecutive years to create mathematical models that allow predicting results; and the qualitative methodology served to explore and compare opinions of people related to the context. The multiple regression model approximated the mathematics

- 1 Licenciada en Matemáticas. Profesora de Matemática del Departamento de Ciencias de la Educación y Humanidades de la Universidad Nacional Autónoma de Nicaragua-Managua-FAREM-Chontales. Correo: [pjeylinmeybelin@yahoo.com.mx](mailto:pjeylinmeybelin@yahoo.com.mx)  
 <https://orcid.org/0000-0002-6645-5150>
- 2 Doctor en Educación e Intervención Social. Profesor de Matemática y Estadística del Departamento de Ciencias de la Educación y Humanidades de la Universidad Nacional Autónoma de Nicaragua-Managua (FAREM-Chontales). Correo: [tonyromeong@hotmail.com](mailto:tonyromeong@hotmail.com) <https://orcid.org/0000-0003-3910-6980>

Recibido: 24/04/2018 Aprobado: 01/09/2018

Pérez-Obregón, J., & Romero-Díaz, T. (2018). Análisis del rendimiento académico mediante regresión logística y múltiple. *Revista Electrónica de Conocimientos, Saberes y Prácticas*, 1(2), 33-42. DOI: <https://doi.org/10.30698/recsp.v1i2.10>

qualification considering those of language and literature, and foreign language -. The logistic regression model was applied, which allowed to calculate the average interval of the students considering degree, sex and year; The logistic regression model was applied, which allowed to calculate the average interval of the students considering degree, sex and year; Being this model a complement to the previous one, since it strengthens the results for its respective analysis, both models were significant and exceeded the statistical assumptions, therefore, it is recommended to use them to improve the performance of the Educational center.

**Keywords:** Academic performance, logistic and multiple regression, predictions, analysis

## 1. Introducción

Estudiar el rendimiento académico en la asignatura de matemática, tiende a ser un tema muy amplio y a la vez complicado por los distintos factores que intervienen en los resultados de dichas evaluaciones dentro del proceso de enseñanza – aprendizaje, tales como calificaciones obtenidas en las asignaturas de español, matemática e inglés (Ibarra y Michalus, 2010). En este trabajo se analiza el comportamiento del rendimiento académico de los estudiantes del Instituto Nacional Santa Teresita (secundaria regular), del poblado de La Batea, municipio Muelle de los Bueyes, en la disciplina de matemática durante el período 2011-2016.

Es una investigación con enfoque mixto, donde se hace uso de técnicas estadísticas y entrevistas abiertas a maestros para tener una mejor percepción de lo que implica el análisis del rendimiento académico en este contexto. Con el fin que permita tanto a docente como a estudiantes realizar una autorreflexión sobre el trabajo que se está realizando en las aulas de clases, ya que un alto o bajo rendimiento académico no es producto únicamente de una evaluación, sino de un proceso.

Es relevante mencionar que, al estudiar el rendimiento académico, no siempre se obtienen resultados satisfactorios es por esto que se debe estar dispuesto a erradicar aquello que no está siendo fructífero en el proceso de enseñanza – aprendizaje.

## 2. Literatura

El rendimiento académico es un tema controvertido que transcurre en todas las instituciones escolares, ya que este termina determinando la calidad de la educación ante la sociedad y el número de deserción escolar. El rendimiento académico de forma empírica las personas lo definimos como un promedio de las notas obtenidas durante el período de clase y que estas fueron adquiridas a partir de evaluaciones a los estudiantes. En términos más científicos, León (2008) lo relaciona: “Como la productividad del sujeto,

el producto final de la aplicación de su esfuerzo matizado por sus actividades rasgos y la percepción más o menos concreta de los contenidos asignados” (p.2).

Este indica que es la expresión que da a conocer la voluntad y trabajo de los estudiantes en las clases, y que este de una u otra forma da una idea sobre cuanto conocimiento lograron alcanzar de acuerdo al nivel calificativo alcanzado, es decir que un bajo rendimiento podría expresar que no hay mucho aprendizaje adquirido por parte del alumno. El rendimiento académico también se mide por medio de las calificaciones obtenidas por los estudiantes y que este se da a conocer mediante porcentajes. Montero-Rojas, Villalobos-Palma y Valverde-Bermúdez (2007) expresa que:

Encontrar la medida válida de rendimiento académico es un reto, porque convergen distintas variables y formas de cálculo que dependen del objetivo de las materias o de cada profesor. Los indicadores más utilizados para el rendimiento académico han sido las calificaciones y las pruebas objetivas o test de rendimiento creados (p.3).

Puede entenderse que mide el conocimiento, y que este valor exacto del rendimiento académico puede variar de acuerdo a aspectos que condicionan estos resultados como lo es: el contenido, el maestro en cuanto a su forma de enseñar, evaluar y procesar los datos; incluso el contexto puede ir dentro de estos. La modelación matemática, sirve para pensar en representaciones de datos, estudios, ejercicios, de realización de operaciones matemáticas; a continuación, veamos como otros investigadores han tomado y expresado la definición de modelación matemática: Uloa y Rodríguez (2010) expresan que:

La modelación matemática se muestra como un eje medular en la predicción o toma de decisiones respecto de fenómenos sociales o naturales ya que una buena interpretación de un modelo matemático ayudara a tener buenos resultados futuros, de lo contrario las perdidas pueden ser grandes (p.3).

Ellos lo definen como una forma de realizar predicciones de fenómenos o situaciones reales, mediante un modelo matemático que permita obtener resultados, que al ser debidamente analizados se busquen alternativas de mejoras que ayuden en el progreso del objeto de estudio. Hasta el momento se ha visto un poco acerca de la modelación matemática y como funciona a partir de los diferentes modelos y en distintas ciencias. Pero ¿Qué papel juega como predictor de resultados?, al respecto Grass (2016) argumenta:

Los modelos matemáticos utilizan, en la derivación de los enunciados empíricos (teoremas), el cálculo matemático; en especial, la teoría de la probabilidad. En virtud del tratamiento matemático de los postulados del modelo y de la derivación de una serie de expresiones cuantitativas exactas, el modelo matemático permite la

predicción, con un pequeño margen de error, de las posibles relaciones que suelen presentar los datos experimentales. Así, mediante el modelo se puede obtener una estimación teórica de las características y ordenamientos que presentan los datos producidos secuencialmente. En el grado, pues, que el modelo permita una correcta predicción de un determinado estadístico y éste coincida con la estimación del mismo a partir de los datos, se concluye que el modelo es una buena descripción de la situación experimental (p.2).

De aquí se desglosa la importancia de los modelos matemáticos y el uso de esta ciencia a partir de diferentes métodos, unificando teorías y aplicaciones para la creación de estos. Para su creación se procesan datos, el cual permitirá la predicción de futuros resultados, los cuales no hay un 100% de probabilidad que ocurran exactamente ya que el modelo puede variar con los resultados, es decir son predicciones con aproximaciones a los resultados reales que permitan tener una generalidad.

### 3. Metodología de investigación

Dentro del proceso de investigación se revisó la literatura relacionada al tema para apoyar los aspectos o métodos relevantes y poder hacer contrastes. El enfoque considerado en esta investigación fue el mixto porque se utilizó el enfoque cuantitativo analizando el rendimiento académico de forma objetiva estudiando las calificaciones de años consecutivos y a partir de ello se relacionó variables para crear un modelo matemático que predice resultados. También se empleó el enfoque cualitativo para explorar y analizar opiniones de personas relacionadas al contexto.

Este estudio es de tipo transversal ya que se realizó en un tiempo determinado y único período (I semestre 2017), tomando en cuenta las calificaciones de matemática de los últimos 6 años. Para ambos enfoques el muestreo fue no probabilístico porque se tomó la muestra a conveniencia. Se seleccionaron 8 docentes para contrastar las opiniones y 938 estudiantes en la selección de las calificaciones escolares.

En esta investigación se incluyeron solo aquellos docentes que estuvieron durante todo el período seleccionado y las notas de estudiantes que tenían completas sus calificaciones de matemáticas. En la parte cuantitativa las calificaciones se obtuvieron del cuaderno de registro y se analizó el registro mediante el SPSS versión 21, tomando en cuenta el sexo, año lectivo, el grado y el promedio de la clase de matemática, lengua extranjera - lengua y literatura. En cuanto al proceso cualitativo se recurrió a aplicar una entrevista estructurada a los docentes.

Para el análisis de las calificaciones se utilizó el modelo de regresión lineal múltiple para establecer relaciones entre las calificaciones de la asignatura de matemática (variable dependiente), lengua extranjera - lengua y literatura (variables independientes).

También se empleó el modelo de regresión logística binaria como complemento a la regresión lineal.

Para el análisis de las entrevistas se clasificaron en cuadros comparativos, con sus respectivas preguntas y respuestas obtenidas, previamente depuradas. Se aprovechó la metodología cualitativa para darle salida a los objetivos mediante las respuestas de las entrevistas. Se extrajeron los aspectos negativos y positivos que los docentes expresaron al momento de responder, ya que al ser preguntas abiertas las expresiones llegan más allá de lo que se pretendía conocer. A continuación, se presenta el cuadro comparativo utilizado en el análisis de los resultados ya comprimidos de las entrevistas.

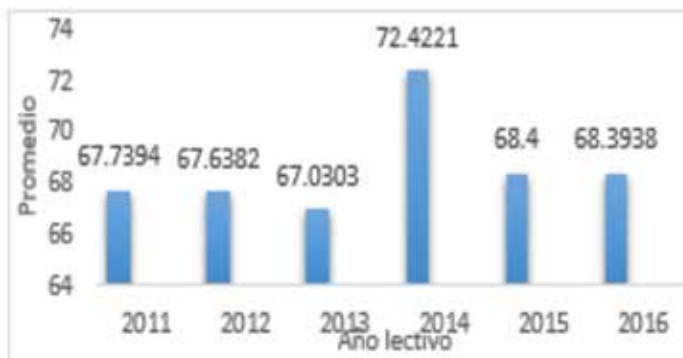
**Tabla 1:** Propuesta de análisis de entrevistas

Categorías respecto a los objetivos	Respuestas obtenidas en relación a los objetivos	Aspectos Positivos	Aspectos Negativos

Respecto a las calificaciones, se ajustaron en un modelo que visualizo el comportamiento del rendimiento académico de los estudiantes en la asignatura de matemática mediante el modelo de regresión múltiple (Escalante, Mayorga y Vergara, 2012). El modelo de regresión lineal múltiple busca una función de regresión poblacional a partir de una función de regresión muestral. La idea de regresión es mostrar un análisis cuantitativo. La regresión logística binaria se utilizó en la selección de dos intervalos en cuanto al promedio de la asignatura. Este modelo logístico permite estudiar el comportamiento de una variable no métrica, es decir de tipo cualitativa con otras variables métricas.

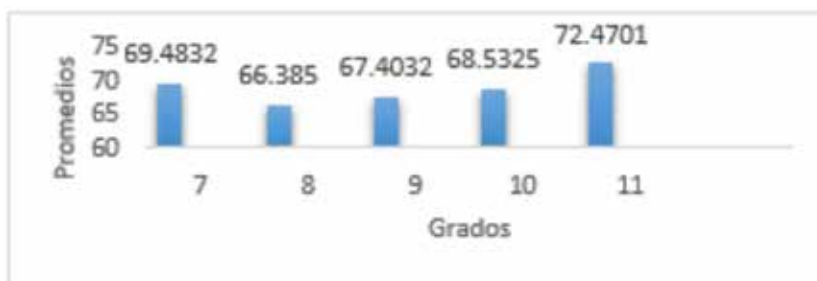
## 4. Resultados

En cuanto a los resultados del análisis cualitativo se relaciona mucho con los resultados obtenidos en la parte cuantitativa. El grado con mejores resultados el undécimo y el séptimo grado, en cuanto al año hubo inconsistencia, los docentes han notado que los mismos estudiantes que tienen problemas en matemática también los enfrentan en lengua extranjera, lengua y literatura. Algunos aspectos positivos son los reforzamientos y entre los negativos sistema de evaluación y falta de apoyo por los padres de familia.



**Figura 1:** Promedios de matemáticas respecto a cada año lectivo

Se puede observar que en el año 2014 se obtuvo un mayor rendimiento académico en términos numéricos respecto a los demás años.



**Figura 2:** Promedios de matemáticas respecto a cada grado

En el gráfico anterior podemos verificar que el grado con más alto rendimiento académico es el undécimo grado con 72.47 % y le sigue el séptimo grado con un 69.48%.

Se estableció un modelo de regresión múltiple para cada grado tomando en cuenta las calificaciones de lengua extranjera – lengua y literatura; y se ajustó el modelo de regresión logística para establecer en que intervalo de promedio se puede encontrar un estudiante de acuerdo a la nota de matemática, además el intervalo se realizó de 0 a 79 y de 80 a 100, debido a que un estudiante con un promedio de 80 se puede considerar como muy bueno, mientras que uno de 70 se podría decir que hay poco aprendizaje adquirido. Para ambos modelos se tomó un intervalo de confianza del 95%.

**Aplicación del modelo de regresión múltiple:** Para verificar la relación que existe entre las calificaciones de matemática con las de lengua extranjera - lengua y literatura. Se presentan datos del modelo por cada uno de los grados del Instituto Nacional Santa Teresita La Batea de secundaria regular, respecto a la ecuación:

Se tomó en cuenta que la variable dependiente Y son las calificaciones de matemática y las variables independientes son:

$$X_1 = \text{Lengua extranjera}, X_2 = \text{Lengua y literatura}.$$

**Tabla 2:** Resumen del modelo de regresión lineal múltiple

	R <sup>2</sup>	R <sup>2</sup> corregida	Error típico de la estimación
Séptimo	0.537	0.534	6.22468
Octavo	0.476	0.471	6.06716
Noveno	0.615	0.611	6.98532
Décimo	0.550	0.544	7.38262
Undécimo	0.588	0.582	6.88583

En el caso de los grados séptimo, noveno, décimo y undécimo, el modelo explico más del 50 % y una vez corregido por el efecto de la muestra y de las variables independientes resulta ser siempre más del 50%. En cambio, en octavo grado es menos, siendo de 47.6%.

**Tabla 3:** Coeficientes del modelo

		Coefficientes	Significancia
Séptimo	Constante	17.869	0.000
	Lengua extranjera	0.102	0.031
	Lengua y literatura	0.631	0.000
Octavo	Constante	19.457	0.000
	Lengua extranjera	0.113	0.013
	Lengua y literatura	0.555	0.000
Noveno	Constante	4.522	0.245
	Lengua extranjera	0.100	0.096
	Lengua y literatura	0.788	0.000
Décimo	Constante	10.070	0.022
	Lengua extranjera	0.180	0.008
	Lengua y literatura	0.638	0.000
Undécimo	Constante	9.519	0.043
	Lengua extranjera	0.245	0.000
	Lengua y literatura	0.606	0.000

Los dos coeficientes de regresión múltiple son mayormente significativamente distintos, por tanto, contribuyen significativamente a la ecuación de regresión. A partir de esta información denotamos que las ecuaciones de regresión múltiple son:

**Tabla 4:** Ecuaciones de regresión múltiple

Grado	Ecuación
Séptimo	$Y = 17.869 + 0.102X_1 + 0.631X_2$
Octavo	$Y = 19.457 + 0.113X_1 + 0.555X_2$
Noveno	$Y = 4.522 + 0.100X_1 + 0.788X_2$
Decimo	$Y = 10.070 + 0.180X_1 + 0.638X_2$
Undécimo	$Y = 9.519 + 0.245X_1 + 0.606X_2$

Estas ecuaciones luego fueron puestas en práctica con estudiantes activos del último año los resultados mostraron un alto porcentaje de acertar al contexto.

Aplicación del modelo de regresión logística binaria: La variable dependiente la sometimos a modelizar o predecir las calificaciones de la signatura de matemática, que será una categórica dicotómica, codificada con valores 0 y 1 para el intervalo donde se encontrara el promedio (0.5 es el valor de corte). Las variables independientes o modificadoras de efecto, y que fueron incluidas en el modelo, son el sexo, el grado y el año lectivo.

**Tabla 5:** Clasificaciones del modelo de regresión logística binaria

Observado		Pronosticado		
		Intervalo		Porcentaje correcto
		80 a 100	0 a 79	
intervalo	80 a 100	118	32	78.7
	0 a 79	23	765	97.1
Porcentaje global				94.1

Indica que el modelo es bueno para clasificar en el futuro ya que, si se aplica la clasificación a las observaciones ya conocidas, se obtiene un porcentaje de éxito del 94.1%, esto nos hizo suponer que las clasificaciones futuras mantendrán el mismo porcentaje de éxito.

**Tabla 6:** Coeficientes del modelo

	Sexo (1)	Grado	Año	Matemática	Constante
B	-1.537	-0.135	-0.110	-0.299	25.851
Significancia	0.000	0.200	0.253	0.000	0.000



Esta tabla 6, nos regala los coeficientes para determinar la ecuación del modelo de regresión logística binaria:

$$x_1 = \text{sexo } x_2 = \text{grado } x_3 = \text{año } x_4 = \text{nota de matemática}.$$

$$P = \frac{1}{1 + e^{-(25.851 - 1.537x_1 - 0.135x_2 - 0.110x_3 - 0.299x_4)}}$$

Luego se validaron ambos modelos usando Excel, se realizaron cálculos del modelo de regresión múltiple con notas del libro de registro y se comparó la aproximación de los datos obtenidos en el programa. En la siguiente tabla: A= Nota de matemática, B=Lengua extranjera, C= Lengua y literatura, D= Sexo, E= Grado, F= Año y G= Nota aproximada de matemática por la regresión múltiple.

**Tabla 7:** Validación de los Resultados de las ecuaciones de los modelos de regresión

Notas extraídas del libro de registro	Resultados de la regresión múltiple	Datos para la aproximación del intervalo	Aproximación de la regresión logística
A= 78, B= 88, C= 83	79.218	D= 1, E= 1, F= 6, G= 79.218	0.4583
A= 60, B= 46, C= 63	59.62	D= 1, E= 2, F= 6 G= 59.62	0.99616
A= 69, B= 83, C= 76	72.71	D= 1, E= 3, F= 6 G= 72.71	0.81887
A= 67, B= 60, C= 62	60.426	D= 1, E= 4, F= 6 G= 60.426	0.99361
A= 71, B= 79, C= 84	79.778	D= 1, E= 5, F= 6 G= 72.71	0.29429

A partir de estos resultados se pudo observar que, al realizar las operaciones para el modelo de regresión binaria logística, tomando en cuenta solo la nota de matemáticas obtenida del modelo de regresión múltiple, estas se aproximan mucho a los datos del libro de registro y coinciden con los intervalos a los que corresponde.

## 5. Discusión y conclusiones

De acuerdo a los resultados obtenidos del análisis general del rendimiento académico de la asignatura de matemática durante el período 2011 – 2016, se logró identificar que el año lectivo 2014 tiene mejores resultados, datos que fueron expresados al igual por docentes. Así también se deduce que el grado más sobresaliente es undécimo grado, siguiéndole el séptimo grado.

Se logró establecer un modelo matemático regresión múltiple que permite predecir resultados del rendimiento académico para el período lectivo actual del estudiante relacionando calificaciones de matemática respecto a lengua extranjera - lengua y literatura. Además de un modelo probabilístico regresión logística que complementa los resultados del modelo anterior determinando el intervalo de promedio en el cual se encuentra el estudiante respecto de la calificación de matemática. Se evidencian

las aproximaciones del modelo y que si hay relación entre las asignaturas de lengua extranjera - lengua y literatura con la asignatura de matemática, pudiendo verificar que se puede deducir esta calificación a partir de las ecuaciones obtenidas de los modelos analizados en el estudio.

## 6. Lista de referencia

- Escalante, R. D., Mayorga M., & Vergara, J. (2012). *Manual de aplicación del modelo de regresión lineal múltiple con correcciones de especificaciones*. E-Libro-Corporación.
- Grass, J., A. (2016). *Utilización de modelos matemáticos en psicología*. Barcelona: Departamento de Psicología Experimental.
- Ibarra, M., Michalus, J. "Análisis del rendimiento académico mediante Modelos de Regresión" (2010). *Ingeniería Industrial*, 9(2), 47-56.
- León, B. (2008). Atención plena y rendimiento académico en estudiantes de enseñanza secundaria. *European journal of education and psychology*, 1(3), 17-26.
- Montero-Rojas, E., Villalobos-Palma, J., & Valverde-Bermúdez, A. (2007). Factores institucionales, pedagógicos, psicosociales y sociodemográficos asociados al rendimiento académico en la Universidad de Costa Rica: Un análisis multinivel. *RELIEVE*, 13(2), 215-234.
- Sampiere, R., Collado, C., & Lucio, F. (2010). *Metodología de la investigación*. México: McGraw-Hill.
- Uloa y Rodríguez, U. (2010). *El modelo logístico: una alternativa para el estudio de crecimiento poblacional de organismo*. Madrid: Universidad Complutense.