



LILIANA VITOLA GARRIDO*

ARTÍCULO DE INVESTIGACIÓN
FECHA DE RECEPCIÓN: 30 DE MAYO DE 2014
FECHA DE EVALUACIÓN: 9 DE OCTUBRE DE 2014

REGRESIÓN LOGÍSTICA: UNA APLICACIÓN EN LA IDENTIFICACIÓN DE VARIABLES QUE INCIDEN EN EL RENDIMIENTO ACADÉMICO, EN EL AREA DE MATEMÁTICAS

Logistic regression: an application in the identification of variables affecting academic performance in the area of mathematics

Regressão logística: uma aplicação na identificação de variáveis que possam afetar o desempenho acadêmico, na área da matemática

* Licenciada en Matemáticas, de la Universidad de Sucre; Magister en Estadística Aplicada, de la Universidad del Norte, Docente Catedrática de la Universidad de Sucre. Correo electrónico: lilianavit_1@hotmail.com



Referencia: Vitola, L. (2015). Regresión Logística: una aplicación en la identificación de variables que inciden en el rendimiento académico, en el área de matemáticas. *Revista Educación y Desarrollo Social*. 9(1), 118-131.

RESUMEN

Este artículo es el producto de una investigación que buscaba identificar entre las variables socioeconómicas, demográficas, familiares y personales, aquellas que influyen en el rendimiento académico de los estudiantes en el área de matemáticas. A su vez, la investigación estuvo motivada por el bajo rendimiento que los estudiantes del nivel básico secundario registran en dicha área, tanto a nivel institucional como local. Aplicando un modelo de regresión logística, se analizaron las variables estructurales mencionadas. Los datos se obtuvieron con la realización de un cuestionario aplicado a una muestra de estudiantes del grado séptimo de la Institución Educativa Santa Rosa de Lima de la ciudad de Sincelejo, en el que se indagó sobre aspectos socioeconómicos, demográficos, familiares, personales y

físicos de la vivienda en la cual habitan los estudiantes y que pueden estar relacionados con su rendimiento académico. Se procesaron los datos usando el software Statgraphics y la información que éste arrojó se organizó en tablas. Los resultados mostraron que el rendimiento académico o desempeño de un estudiante está supeditado, no sólo a factores escolares, sino también a variables de tipo social, económico y familiar del contexto donde el estudiante se desarrolla. Entre esos factores están el número de hermanos que el estudiante tiene, sexo del estudiante, estado civil de los padres, raza del estudiante, ingresos económicos de los padres, con quien vive el estudiante e interés del alumno por estudiar.

Palabras clave: matemáticas, predicción, regresión logística, rendimiento académico, variables estructurales.

ABSTRACT

This article is the result of a research that sought to identify among socioeconomic, demographic, family and personal variables, those that affect the academic performance of students in the area of mathematics. The research was also motivated by the poor performance of secondary students in this area, both at an institutional and local levels. Using a logistic regression model, the structural variables mentioned were analyzed. The data were obtained from a questionnaire applied to a sample of seventh graders of Santa Rosa de Lima School at the city of Sincelejo. In the questionnaire the students were asked about socioeconomic, demographic, family, personal and physical aspects of housing in which they live and which may be related to their academic performance. Data were processed using the Statgraphics software and the information was organized in tables. The results showed that academic achievement and student performance is subject not only to school factors, but also to social, economic and family variables of the context where the student develops. Among these factors are number of siblings, sex, marital status of parents, race, parents' income, with whom the student lives and student's interest in studying.

Keywords: mathematics, prediction, logistic regression, academic achievement, structural variables.

RESUMO

O presente artigo é o produto de uma pesquisa que procura identificar entre as variáveis socioeconômicas, demográficas familiares e

personais, aquelas que influenciam no desempenho acadêmico dos alunos na área das Matemáticas; motivada por sua vez pelos baixos rendimentos que os alunos da primeira série do segundo grau têm registrado nesta área, a nível institucional tanto como a nível local. Usando um modelo de regressão logística, foram analisadas as variáveis estruturais mencionadas acima. Os dados foram obtidos com o preenchimento de um questionário aplicado a uma amostra de estudantes da primeira série do segundo grau de Instituição Educativa Santa Rosa de Lima da cidade de Sincelejo, no qual foram investigados aspectos socioeconômicos, demográficos, familiares, pessoais e físicos da habitação em que os alunos vivem e que pode estar relacionada com o desempenho acadêmico dos mesmos. Foram processados os dados usando o software Statgraphics e a informação fornecida por ele foi apresentada em tabelas. Os resultados mostraram que o desempenho acadêmico ou desenvolvimento de um aluno está sujeito não só a fatores escolares, mas também às variáveis de tipo social, econômico e familiar em que o aluno se desenvolve; entre esses fatores estão: o número de irmãos que tem o estudante; o sexo do estudante; o estado civil dos seus pais, a raça do estudante; a renda econômica dos seus pais; com quem vive o estudante; e se o aluno tem interesse por estudar.

Palavras-chave: matemáticas, predição, regressão logística, desempenho acadêmico, variáveis estruturais.

Actualmente los resultados de pruebas internas y externas aplicadas en las Instituciones Educativas (IE) de Básica Secundaria y Media de la ciudad de Sincelejo dan cuenta de las dificultades que, en general, presentan los estudiantes en diversas áreas, entre ellas la de Matemáticas (ICFES, 2010-2012). Lo anterior se evidencia posteriormente, con el bajo rendimiento que muchos estudiantes demuestran en las asignaturas que tienen que ver con dicha área, cuando ingresan a la universidad. Ante tal situación, y de acuerdo con los resultados que obtienen los estudiantes en las pruebas internas y externas que se realizan de manera anual en cada IE del nivel básico y medio de enseñanza, las IE diseñan, ejecutan y evalúan planes de mejoramiento para abordar las falencias o debilidades que se pueden estar presentando en cuanto a prácticas pedagógicas de los docentes y que posiblemente están afectando el rendimiento académico de los estudiantes. Dichos planes de mejoramiento, en su mayoría, van encaminados a reestructurar el quehacer diario de los docentes y que hacen parte del proceso enseñanza-aprendizaje.

Pero además de las prácticas pedagógicas, sobre las cuales tradicionalmente se hacen los planes de mejoramiento, el rendimiento académico de los estudiantes a nivel escolar en las IE de nivel básico secundario y medio se puede ver influenciado por diversas variables de tipo social, económico y familiar, propias del contexto donde se desenvuelven los alumnos. Entre estas variables están el estrato, los ingresos económicos en el hogar, la raza, el sexo, la edad, los problemas familiares, el abandono, el abuso sexual, el desplazamiento forzado, etc.

Adell (2002) menciona, que entre los factores por los que los niños pueden presentar bajo rendimiento escolar está el factor económico, pues los padres a lo mejor no tienen buenos ingresos, o porque no viven con alguno de sus padres, o están a cargo de sus abuelos u otro familiar.

Además Morales, Arcos, Ariza, Cabello, López, Pacheco, Palomino, Sánchez y Venzalá (2005) expresan que el rendimiento escolar es la resultante del complejo mundo que envuelve al estudiante: Cualidades individuales (actitudes, capacidades, personalidad), y su medio socio-familiar (familia, amistades, barrio), entre otros.

Pérez (citado por Adell, 2002) afirma que sobre el rendimiento académico influyen diversos factores; por eso se dice que es multicondicionado y multidimensional y uno de los factores determinantes es la familia.

De igual manera, Morales et al. (2005), después de realizar un estudio titulado “El Entorno Familiar y el Rendimiento Escolar” en estudiantes de educación primaria de los centros públicos de Lucena (Córdoba, España) menciona que el nivel cultural de los padres, el nivel económico, el número de hermanas y hermanos, los problemas familiares y el interés y expectativas de la familia, influyen significativamente en el rendimiento escolar de sus hijos.

Otra investigación realizada por Miguel Andrade, Christian Miranda e Irma Freixas en Chile en 2005 abordó el Rendimiento Académico y Variables Modificables en Alumnos de 2° de Liceos Municipales de la Comuna de Santiago. En este estudio se determinaron los niveles de relación de algunas variables intelectuales y del hogar, con el rendimiento académico de los estudiantes, en matemáticas

“Actualmente los resultados de pruebas internas y externas aplicadas en las Instituciones Educativas (IE) de Básica Secundaria y Media de la ciudad de Sincelejo dan cuenta de las dificultades que, en general, presentan los estudiantes en diversas áreas, entre ellas la de Matemáticas.”

y castellano, haciendo uso de un modelo de regresión lineal múltiple, además de identificar cuáles de las variables intelectuales y del hogar son las que tienen mayor capacidad predictiva sobre el rendimiento de los alumnos.

Por otra parte, la investigación Modelo Causal del Rendimiento en Matemáticas, realizada por Arrieta en 1998, propuso un modelo teórico que recogía jerárquicamente los factores que influyen significativamente en el rendimiento en el área de matemáticas y que permitió obtener un correcto diagnóstico de las causas del bajo rendimiento del alumno en esta área.

Por todo lo anterior, es necesario identificar de manera clara y concreta cuales, de las mencionadas variables, son las que inciden realmente en el rendimiento académico de los alumnos, de tal manera que una vez identificadas se diseñe un plan de mejoramiento que las trabaje para que los niveles en el rendimiento académico mejoren, especialmente en el área de matemáticas. Esto se evidenciara posteriormente con el desempeño que los estudiantes demuestren cuando ingresen a la universidad.

Se propone aplicar el Modelo de Regresión Logística para identificar las variables de tipo socioeconómico, familiar, personal, de equipamiento y demográfico, que ejercen influencia sobre el rendimiento académico de los estudiantes debido a que, según (Linás, 2010, pp. 1, 29), “los modelos de Regresión Logística son adecuados para aquellas situaciones en las que se desee encontrar la probabilidad “ p ” de ocurrencia de un evento, por medio de los valores ciertas variables explicativas”, es decir, la relación entre una variable dependiente categórica que representa el

evento de interés y variables independientes numéricas o categóricas como edad, sexo, raza, religión, etc. que representan las variables explicativas. Además, las variables independientes categóricas se pueden incluir en el modelo de manera codificada con 1 y 0 y esta codificación representará los niveles que tienen dichas variables; de igual manera se puede asociar a la variable dependiente, una variable dicotómica Y , que tomará los valores $Y=\{1, 0\}$.

Una vez identificadas las variables que influyen en el desempeño de los alumnos, con ayuda del modelo de Regresión Logística se determinará en qué medida éstas variables están relacionadas con el rendimiento académico de los estudiantes en el área de matemáticas y cuáles de éstas describen o explican, de mejor manera, el desempeño de los alumnos en el área. De esta manera se obtendrá un modelo predictivo de variables fundamentales, que ayudan a predecir el tipo de rendimiento que un estudiante va a tener, de acuerdo con las características del medio en donde se desarrolla como persona.

METODOLOGÍA

Población Objeto De Estudio

La población objeto de estudio son los alumnos pertenecientes a la IE Santa Rosa de Lima de la ciudad de Sincelejo.

Muestra

Para realizar el estudio se tomó una muestra de 76 estudiantes pertenecientes al grado séptimo de la IE Santa Rosa de Lima de la ciudad de Sincelejo.

Instrumento

Los datos se recogieron mediante una encuesta con 20 preguntas que recogían información acerca de las condiciones económicas y sociales en las cuales viven los estudiantes y que está conformada por las siguientes variables (Monreal, 1989, pp. 51); (Abarca, 2008, pp. 12-13); (Waldron, E., Salamanca, L. & González, C., 2008, pp. 5-6):

- 1) *Nombres y Apellidos:* _____
- 2) *Vives en zona:* Urbana Rural
- 3) *Edad:* _____
- 4) *Sexo:* Femenino Masculino
- 5) *Estado Civil (padre y/o madre):*
 Casado
 Soltero
 Separado
 Unión libre
 Vive con una pareja que no es tu Papá o Mamá
- 6) *Estrato:*
 1
 2
 3
 4
 5
- 7) *Raza:*
 Blanco
 Negro
 Indígena

8) Eres desplazado: Si No

9) *Máximo nivel académico alcanzado por los padres (Marca solo una opción):*

- Primaria
- Bachillerato
- Cursos Técnicos
- Profesional Universitario
- Magister
- Doctorado
- No estudiaron

10) *Ingresos económicos mensuales de los padres:*

- Menos de un salario mínimo (menos de \$616000)
- Un salario mínimo (\$616000)
- Entre uno y dos salarios mínimos (entre \$616000 y \$1232000)
- Más de dos salarios mínimos (más de \$1232000)

11) *Actividad económica de los padres:*

- Empleado
- Desempleado
- Trabajador Independiente

12) *Tipo de vivienda:*

- Propia
- Arrendada
- Familiar

13) *Servicios públicos básicos de la vivienda (Si tienes todos los servicios no los marques todos, marca la opción "Todas las anteriores"):*

- Luz, agua, gas e Internet
- Sólo luz, agua y gas natural
- Sólo dos de las anteriores

14) *Servicios de salud:*

- Régimen de salud contributiva
- Régimen de salud subsidiada

15) *Numero de hermanos:*

- 1
- 2
- 3
- 4
- 5
- 6 ó más

16) *En tu casa hay problemas de (Si en tu casa hay todos estos problemas no los marques todos, marca la opción "Todas las anteriores"):*

- Violencia Intrafamiliar
- Abuso Sexual
- Drogas
- Alcohol
- Abandono por parte de uno o ambos padres
- Todas las anteriores
- Ninguna de las anteriores

17) Vives con:

- Ambos padres y hermanos
- Solo mamá y hermanos
- Solo papá y hermanos
- Papá, madrastra y hermanos
- Mamá, padrastro y hermanos
- Otros (Abuelos, y/o tíos, y/o hermanos, y/o primos, solo)

18) Con respecto a las asignaturas que tienen que ver con matemáticas:

- Te gustan mucho
- Te gustan poco
- No te gustan

CONSIDERAS QUE TIENES

19) Interés por estudiar:

- Siempre
- Casi Siempre
- Casi Nunca
- Nunca

20) Responsabilidad y compromiso frente al estudio:

- Siempre
- Casi Siempre
- Casi Nunca
- Nunca

Procedimiento

Para recolectar la información necesaria se diseñó una encuesta con preguntas orientadas a los estudiantes, que buscaban indagar la situación del contexto en el que viven como, por ejemplo, estrato, si viven o no con los padres, si son desplazados o no, ingresos económicos de los padres, nivel de formación académica de los padres, entre otras.

En las variables que son de tipo categórico como sexo, raza o estado civil, por ejemplo, las respuestas se codificaron con 1 y 0 si sólo tenían dos niveles. Por ejemplo, para la variable sexo que tiene dos niveles (masculino y femenino), la codificación quedó de la siguiente manera:

SEXO	CÓDIGO
Femenino	1
Masculino	0

Para variables con más de dos niveles, por ejemplo actividad económica de los padres, que tiene tres niveles (empleado, independiente, desempleado), la codificación quedó de la siguiente forma:

ACTIVIDAD ECONÓMICA DE LOS PADRES	CÓDIGO
Empleado	1
Independiente	2
Desempleado	3

Para variables numéricas como número de hermanos, estrato y edad se dejaron las cantidades numéricas escritas por los estudiantes.

Seguidamente, y como variable dependiente, se tuvo en cuenta el rendimiento académico obtenido por los estudiantes en el área de matemáticas en lo corrido del primer periodo del año lectivo 2014, es decir, las notas que muestren si los estudiantes hasta el momento, aprueban o no el área de matemáticas durante el primer período del año en curso.

Debido a que en la institución los puntajes van de 1.0 a 5.0, la codificación para la variable dependiente rendimiento académico quedó de la siguiente manera:

Nota (1.0 a 2.9)(y=0)	Nota (3.0 a 5.0)(y=1)
No aprobado	Aprobado

Posteriormente se aplicó la Teoría de Regresión Logística para encontrar el modelo que muestre cuáles de las variables independientes (socioeconómicas, familiar, personal, demográfica y de equipamiento) se relacionan con la variable dependiente (rendimiento académico en el área de matemáticas). Dado que el modelo de Regresión Logística tiene la forma (Linás, 2010):

$$g(x) = \delta + \beta_1 x_1 + \dots + \sum_{m=1}^{k_j-1} \beta_{jm} D_{jm} + \beta_k x_k$$

Se procesaron y analizaron los datos con ayuda del Software Statgraphics, para identificar las variables que componen el modelo con el método hacia atrás. De acuerdo con el valor-p

asociado a las variables, se eliminaron las que mayor valor-p tenían, de tal manera que al final nos quedaron las variables más significativas. Los coeficientes de estas variables independientes se calcularon con el Método de Máxima Verosimilitud y se verificó el tipo de relación que tienen las variables independientes con la variable dependiente, de acuerdo al signo que tienen sus coeficientes. Si el signo es positivo, la relación entre las variables es directa; si es negativo la relación entre las variables es inversa.

Si la relación es directa y la probabilidad de aprobar el área de matemáticas es alta, entonces estas variables influirían de manera positiva sobre el desempeño académico de los alumnos. Por el contrario, si la relación es inversa y la probabilidad de aprobar el área de matemáticas es baja, se podrá concluir que estas inciden de manera negativa sobre el desempeño de los estudiantes.

Posteriormente se puso en marcha el modelo, asignándole valores a las variables que lo componen para calcular la probabilidad de que un estudiante apruebe el área de matemáticas y, de esta manera, poder predecir la ocurrencia de este evento, dadas las variables del contexto en las que se encuentra.

RESULTADOS

Análisis Descriptivo de la Muestra

La Tabla 1 muestra el análisis descriptivo de la muestra en donde aparecen las frecuencias y porcentajes de las variables socio demográficas y económicas, que conforman parte de la encuesta.

Estadísticas Descriptivas de las variables Socio demográficas y Económicas

VARIABLE	FRECUENCIA	PORCENTAJE
Edad		
< 12 a ños	5	7.81 %
12 – 14 años	57	89.1%
> 14 años	2	3.1%
Sexo		
Femenino	30	46.9%
Masculino	34	53.1%
Estrato socioeconómico		
Bajo (1-2)	54	84.4%
Medio (3-4)	9	14.1%
Alto (5-6)	1	1.5%
Zona en la que vive		
Rural	37	42.2%
Urbana	27	57.8%

Análisis Estadístico

La Tabla 2 muestra las variables independientes con su correspondiente valor p , el cual indica si dicha variable es significativa en el modelo, es decir, si influye o no sobre la variable dependiente *rendimiento académico en el área de matemáticas*. Para ello se debe ver que si el valor p es mayor que el nivel de significancia 0,05, la variable en cuestión no es significativa, mientras que si es menor que el nivel de significancia 0,05, es significativa y por lo tanto influye sobre la variable de respuesta.

FACTOR	CHI-CUADRADA	G.L.	VALOR-P
Hermanos	5,03009	1	0,0249
Sexo	10,8878	1	0,0010
Estado Civil	21,8463	4	0,0002
Raza	16,0724	2	0,0003
Ingresos	11,4978	3	0,0093
Vives con	21,1298	4	0,0003
Interés	12,195	2	0,0022

Se propone aplicar el Modelo de Regresión Logística para identificar las variables de tipo socioeconómico, familiar, personal, de equipamiento y demográfico, que ejercen influencia sobre el rendimiento académico de los estudiantes.

Como se puede observar, las variables número de hermanos, sexo, estado civil, raza, ingresos económicos de los padres, personas con quien vive e interés por el estudio tiene un valor p menor que 0,05, lo que lleva a concluir que dichas variables influyen sobre la variable dependiente rendimiento académico en el área de matemáticas con una confianza del 95%.

Por otro lado, la siguiente salida (Tabla 3) muestra los valores estimados de los coeficientes del modelo regresor con sus respectivos

niveles. Cada uno de estos coeficientes representa la variable independiente o explicativa que influye sobre la variable dependiente que para este estudio es rendimiento académico en el área de matemáticas, con sus correspondientes errores de estimación.

Con los valores de los coeficientes plasmados en esta tabla se puede armar el Modelo de Regresión Logística buscado. Los signos positivos y negativos en cada uno de los valores de los coeficientes hallados dan muestra del tipo de relación

Tabla 3: Modelo Estimado de Regresión (Máxima Verosimilitud)

		ERROR	RAZÓN DE MOMIOS
PARÁMETRO	ESTIMADO	ESTÁNDAR	ESTIMADA
Constante	-23,7072	316,257	
Hermanos	-1,0153	0,556	0,362293
Sexo=0	4,79576	2,05687	120,996
Estado Civil=1	-2,71768	3,43446	0,0660276
Estado Civil=2	-27,4652	189,026	1,18041E-12
Estado Civil=3	0,937055	2,99565	2,55245
Estado Civil=4	-6,541	4,29309	0,00144304
Raza=1	6,23411	3,23855	509,846
Raza=2	0,159115	1,94353	1,17247
Ingresos=1	-1,53612	2,86993	0,215216
Ingresos=2	4,06872	3,15547	58,4819
Ingresos=3	25,4408	146,434	1,11897E11
Vives con=1	-0,990465	2,81782	0,371404
Vives con=2	-5,40915	4,02092	0,00447546
Vives con=4	-9,90011	4,83369	0,0000501692
Vives con=5	-11,4845	5,94028	0,0000102887
Interés=1	27,6628	316,243	1,03234E12
Interés=2	31,1361	316,262	3,32839E13

que esas variables tienen con la variable independiente: Si el signo es positivo, la relación es directa pero si es negativo, la relación es inversa.

En consecuencia, las variables sexo (hombre), estado civil (separado), raza (blanca o negra), ingresos económicos (un salario mínimo, o entre uno y dos salarios mínimos), interés por el estudio (siempre o casi siempre), influyen de manera directa sobre el rendimiento académico de los estudiantes en el área de matemáticas, porque su coeficiente tiene signo positivo.

Mientras tanto, las variables número de hermanos, estado civil (casado, soltero, unión libre), ingresos económicos (menos del mínimo), con quién vives (ambos padres y hermanos o sólo mamá y hermanos o papá, madrastra y hermanos o mamá, padrastro y hermanos) influyen de manera inversa sobre el rendimiento académico de los estudiantes en el área de Matemáticas, porque su coeficiente tiene signo negativo.

A su vez, la Tabla 4 muestra la bondad de ajuste del modelo. En ella se debe ver que si el valor p del modelo es menor que 0,05, el modelo hallado es adecuado; de lo contrario, si es mayor que 0,05, el modelo hallado no es el adecuado y las variables implícitas en él, no

explican de manera completa la variable dependiente rendimiento académico en el área de matemáticas.

Como se puede observar, el valor p del modelo es menor que 0,05, lo que indica que el modelo hallado es el adecuado, con una confianza del 95%. Además, se puede observar el porcentaje de variabilidad explicada por el modelo el cual es de 68, 5035%, lo que indica que se explica gran parte de la variable dependiente con el modelo hallado.

En consecuencia y dado que el modelo de Regresión Logística tiene la forma

$$g(x) = \delta + \beta_1 x_1 + \dots + \sum_{m=1}^{k_j-1} \beta_{jm} D_{jm} + \beta_k x_k$$

el modelo queda así:

$$g(x) = -23,7072 - 1,0153 * X_1 + 4,79576 * X_2 - 2,71768 * X_3 - 27,4652 * X_4 + 0,937055 * X_5 - 6,541 * X_6 + 6,23411 * X_7 + 0,159115 * X_8 - 1,53612 * X_9 + 4,06872 * X_{10} + 25,4408 * X_{11} - 0,990465 * X_{12} - 5,40915 * X_{13} - 9,90011 * X_{14} - 11,4845 * X_{15} + 27,6628 * X_{16} + 31,1361 * X_{17}$$

dónde

X_1 = Numero de hermanos; X_2 = Sexo (hombre); X_3 = Estado civil casado; X_4 = Estado civil soltero; X_5 = Estado civil separado; X_6 = Estado civil unión libre; X_7 = Raza blanca; X_8 = Raza negra; X_9 = Ingresos menos de un mínimo; X_{10} = Ingreso un mínimo; X_{11} = Ingresos entre uno y dos mínimos; X_{12} = Vives con ambos padres; X_{13} = Vives con solo mamá; X_{14} = Vives con papá y madrastra; X_{15} = Mamá y Padrastro; X_{16} = Interés por estudio siempre; X_{17} = Interés por estudio casi siempre.

Tabla 4: Análisis de desviación

FUENTE	DESVIACIÓN	G.L.	VALOR P
Modelo	55,4905	17	0,0000
Residuo	25,5134	46	0,9938
Total (corr.)	81,0039	63	

Porcentaje de desviación explicado por el modelo = 68,5035

Debido a que con el modelo hallado se puede calcular la probabilidad de que un estudiante apruebe el área de matemáticas, dicha probabilidad está definida como:

$$P = \frac{e^{g(x)}}{1 + e^{g(x)}}$$

Entonces, la probabilidad que el estudiante apruebe el área de matemáticas dado que, por ejemplo: Es $X_2 =$ Hombre; los padres son $X_3 =$ Separados; es de raza $X_7 =$ Blanca, el ingreso económico de los padres es de un $X_{10} =$ Salario mínimo y tiene $X_{16} =$ Siempre interés por el estudio, tomando los valores de los coeficientes de dichas variables, es igual a:

$$P = \frac{e^{g(x)}}{1 + e^{g(x)}} = \frac{e^{-23,7072+4,79576*(0)+0,937055*(1)+6,23411*(1)+4,06872*(1)+27,6628*(1)}}{1 + e^{-23,7072+4,79576*(0)+0,937055*(1)+6,23411*(1)+4,06872*(1)+27,6628*(1)}}$$

$P = 95\%$, es decir que hay una alta probabilidad de aprobar el área de matemáticas, de acuerdo con estas variables.

Por otra parte, la probabilidad que el estudiante apruebe el área de matemáticas dado que, por ejemplo: El estudiante tiene $X_1 = 4$ hermanos; el estado civil de los padres es $X_3 =$ Casados; el ingreso económico de los padres es $X_9 =$ Menos de un mínimo; y vive con $X_{12} =$ Ambos padres, tomando los valores de los coeficientes de dichas variables, es igual a:

$$P = \frac{e^{g(x)}}{1 + e^{g(x)}} = \frac{e^{-23,7072-1,0153*(4)-2,71768*(1)-1,53612*(1)-0,990465*(1)}}{1 + e^{-23,7072-1,0153*(4)-2,71768*(1)-1,53612*(1)-0,990465*(1)}}$$

$P = 0,00046\%$, es decir que hay una baja probabilidad de aprobar el área de matemáticas, de acuerdo con estas variables.

CONCLUSIONES

Con los resultados expuestos en el análisis estadístico se puede concluir que:

El rendimiento académico de los estudiantes de la IE Santa Rosa de Lima, en el área de matemáticas, está influenciado por variables tales como el número de hermanos que el estudiante tiene, el sexo del estudiante, el estado civil de los padres, la raza del estudiante, los ingresos económicos de los padres, con quién vive estudiante y si el alumno tiene interés por estudiar.

Como las variables sexo (hombre), estado civil (separado); raza (blanca o negra); ingresos económicos (un salario mínimo o entre uno y dos salarios mínimos); interés por el estudio (siempre o casi siempre), influyen de manera directa sobre el rendimiento académico de los estudiantes en el área de matemáticas y la probabilidad de aprobar el área es alta, dada la presencia de estas variables, éstas inciden de manera positiva en el rendimiento académico de los alumnos en la mencionada área. Lo anterior tiene sentido en cuanto a

los ingresos económicos de los padres y el interés que el alumno le ponga al estudio. Pero por el contrario de lo que normalmente se piensa, cuando los padres de los estudiantes están separados, según este resultado, esta variable influye de manera positiva sobre el rendimiento académico de los escolares. Por otra parte, el hecho de ser hombre al parecer también tiene una incidencia positiva sobre el desempeño escolar de los alumnos, en el área de matemáticas. Cabe anotar que el coeficiente de la variable sexo (hombre) es igual al de la variable sexo (mujer), sólo que para esta última el signo es contrario, es decir negativo, lo que lleva a pensar que el hecho de ser mujer aparentemente afecta de manera negativa el rendimiento en el área de matemáticas.

Como las variables número de hermanos, estado civil (casado, soltero o unión libre), ingresos económicos (menos del mínimo), con quién vives (ambos padres y hermanos; solo mamá y hermanos; papá, madrastra y hermanos; mamá, padrastro y hermanos), influyen de manera inversa sobre el rendimiento académico de los estudiantes en el área de matemáticas y la probabilidad de aprobar el área es baja, dada la presencia de estas variables, éstas inciden de manera negativa en el rendimiento académico de los alumnos en la mencionada área. Lo anterior tiene sentido ya que si el ingreso económico de los padres es de menos de un salario mínimo y de con quien está viviendo el alumno, debido a que en muchos casos estos no viven con los padres sino con otro tipo de familiar o acudiente. De manera análoga, aparentemente el hecho de que los padres sean casados, solteros o vivan en unión libre afecta de manera negativa el desempeño de los escolares en el área.

REFERENCIAS

- » Abarca, Valeria. (2008). *Estudio cuantitativo sobre el Efecto de variables estructurales en el incremento entre el SINCE y la PSU*. Tesis de Maestría. Santiago: Universidad de Chile.
- » Adell, M. (2002). *Estrategias para mejorar el rendimiento académico de los adolescentes*. Madrid: Ediciones Pirámide.
- » Andrade, M. Miranda, C. Freixas, I. (2005). *Rendimiento académico y variables modificables en alumnos de 2° de liceos municipales de la comuna de Santiago*. Proyecto de Investigación. Santiago de Chile.
- » Arrieta, M. (1998). Modelo causal del rendimiento en Matemáticas. *Enseñanza de las Ciencias*, vol. 16 No 1, 63-71.
- » Llinás, H. (2010). *Guía Resumida de Regresión Logística*. Barranquilla: Universidad del Norte.
- » Monreal, J. (1989). Las variables socioeconómicas como Contexto Metodológico en la evaluación de Programas. *Anales de Psicología*, vol. 5, 43-51.
- » Morales, A. Arcos, P. Ariza, E. Cabello, M. López, M. Pacheco, J. Palomino, A. Sánchez, J. Venzalá, M. (2005). *El Entorno Familiar y el Rendimiento Escolar*. Proyecto de Investigación Educativa subvencionado por la Consejería de Educación y ciencia de la Junta de Andalucía. Recuperado el 16 de noviembre de 2005 de www.juntadeandalucia.es/averroes/publicaciones/investigacion/entorno_familia.pdf
- » Waldron, E. Salamanca, L. González, C. (2005). *Análisis de la estructura y composición de las principales variables demográficas y socioeconómicas del Censo 2005*. Censo de población y vivienda de Colombia DANE.
- » ICFES. (2010-2012). Resultados pruebas históricas. Recuperado el 1 de octubre de 2013 de <http://www.icfesinteractivo.gov.co>