

RENDIMIENTO EN MATEMATICAS I EN LA UNIVERSIDAD TECNOLOGICA DE PEREIRA. FACTORES DE PREDICCIÓN

Performance in Mathematics I at Pereira Technological University. Prediction Factors

RESUMEN

En este trabajo se presentan los resultados sobre el estudio de los factores que influyen de manera significativa sobre el rendimiento académico de los estudiantes de ingenierías y tecnologías de la Universidad Tecnológica de Pereira en la asignatura Matemáticas I. Se propone un modelo lineal de regresión múltiple que emplea las variables más relevantes halladas durante la investigación y que afectan el rendimiento de los estudiantes en la asignatura Matemáticas I, con el fin de predecir el rendimiento de un alumno dado.

PALABRAS CLAVES: Rendimiento académico, modelo regresión múltiple.

ABSTRACT

On this paper, the study of the factors that affect significantly the academic performance in the subject Mathematics I, for students of engineering and technological programs in Pereira Technological University is presented.

A linear multiple regression model that uses most relevant variables found due to the investigation and that affect student achievement in the subject Mathematics I is proposed in order to predict the performance of a given student.

KEYWORDS: Academic performance, multiple regression model.

PATRICIA CARVAJAL OLAYA

Estadística
Magíster en Investigación Operativa y Estadística
Profesora Asistente
Universidad Tecnológica de Pereira
pacarva@utp.edu.co

JULIO CÉSAR MOSQUERA M.

Ph. D en Física.
Profesor Asistente
Universidad del Quindío
jucemos@gmail.com
jcmosquera@uniquindio.edu.co

IRINA ARTAMÓNOVA

Estadística
Magíster en enseñanza de las matemáticas
Docente
Universidad de Quindío
artiri@gmail.com

INTRODUCCIÓN

En 2004 se demostró que existe poca la relación solo entre el puntaje ICFES tanto por áreas, como total, con el rendimiento en matemáticas I y II de los estudiantes de facultades de ingeniería de la Universidad Tecnológica de Pereira [1]. Otro estudio mostró que la función discriminante calculada [2] no logra diferenciar la media de los dos grupos para los estudiantes que perdieron o ganaron álgebra lineal. En ambos estudios se concluyó que además de los puntajes de Icfes deben existir otros factores que influyen en el rendimiento académico en las áreas de las matemáticas de los estudiantes de UTP. Igualmente, el alto grado de deserción y repitencia en algunos cursos básicos de las carreras de ingeniería y tecnologías de la UTP ha sido un campo de especial investigación en los últimos años [3]. Los estilos del aprendizaje utilizados por los estudiantes se analizan en [4]. La administración universitaria ha liderado procesos de investigación que conduzcan a generar herramientas que permitan medir el riesgo de bajo rendimiento académico en la asignatura de matemáticas, en el que incurren los estudiantes de ingenierías de primer semestre de la Universidad Tecnológica de Pereira. A partir de una base de datos suministrada por la oficina de deserción de la Universidad Tecnológica de Pereira, se ha realizado un análisis de regresión [5] que precise la manera como

ciertas variables pueden influir en la nota definitiva de la materia ya mencionada.

La base de datos en cuestión involucra información de los estudiantes de ingenierías de primer semestre del año 2007 segundo periodo académico. A estos estudiantes se les han medido por medio de test psicológicos y pruebas de aptitudes académicas ciertas características que se consideran podrían ser influyentes en la nota definitiva semestral del área de matemáticas. Sin embargo, sólo un análisis de regresión multivariado podría evidenciar el grado de influencia que cada uno de estos factores predichos por el grupo multidisciplinario de investigadores, tiene sobre el rendimiento académico del estudiante en el área de matemáticas.

2.1. RENDIMIENTO EN MATEMATICAS I.

Los resultados del rendimiento académico de los estudiantes en la asignatura de Matemáticas I durante el segundo periodo del año 2007, muestra que de 414 estudiantes de ingenierías y tecnologías en los programas diurnos y nocturnos, el 30% no pasaron la matemática I. Este porcentaje, sin embargo, para las tecnologías eléctrica y mecánica llega hasta 67% mientras que para el programa de química industrial alcanza 59%. Estos porcentajes son aun mayores si se tienen en cuenta los

estudiantes que cancelaron la asignatura matemática I. Los resultados por programas se pueden ver en la tabla 1.

Cod	Matemáticas I PROGRAMA	Pasó la materia		Total	Pasó la materia	
		No pasó	Pasó		No pasó	Pasó
24	tecnología mecánica	29	14	43	67,4%	32,6%
22	tecnología eléctrica	18	9	27	66,7%	33,3%
16	química industrial	13	9	22	59,1%	40,9%
37	Ing. de sistemas y computación (nocturna)	14	21	35	40,0%	60,0%
36	Ing. electrónica	13	20	33	39,4%	60,6%
38	Ing. industrial (nocturna)	11	30	41	26,8%	73,2%
23	tecnología industrial	7	21	28	25,0%	75,0%
28	Ing. de sistemas y computación	7	21	28	25,0%	75,0%
7	licenciatura en matemáticas y física	2	7	9	22,2%	77,8%
14	ingeniería mecánica	5	28	33	15,2%	84,8%
12	ingeniería eléctrica	2	42	48	12,5%	87,5%
13	ingeniería industrial	2	30	32	6,3%	93,8%
86	técnico profesional en mecatrónica	1	34	35	2,9%	97,1%
	TOTAL	124	286	414	30%	70%

Tabla 1. Rendimiento en Matemáticas I

2.2. MODELO DE REGRESION COMPLETO [5].

Para la realización del modelo de tendencia se definió como variable dependiente Y del modelo, la nota en matemáticas de los estudiantes de ingenierías de primer semestre de la Universidad Tecnológica de Pereira. Las variables predictoras o independientes del modelo, aparecen en la tabla 2.

Riesgo cobertura salud
Riesgo relaciones familiares
Riesgo salud física
Riesgo nutricional
Riesgo alteración mental
Riesgo sustancias psicoactivas
Riesgo uso de tiempo libre
Riesgo otras responsabilidades
lcfes
Nivel en lectura literal
Nivel de lectura inferencial
nivel de lectura intertextual

Tabla 2. Variables predictivas del rendimiento académico en matemáticas

Las variables predictoras en su mayoría son variables categóricas y deben ser convertidas a variables numéricas

nominales a través del grado de presencia del riesgo estudiado en cada registro de la base de datos.

Posteriormente se diseñó un modelo general que contempló todas las variables independientes dadas en la tabla 2, para posteriormente separar aquellas que aportan información más significativa en la explicación de la variable dependiente. A cada variable de la tabla 2 se le asignó un nombre X_j en el mismo orden que se encuentra en la tabla 2. A continuación en las tablas 3 y 4 se presenta el resumen para el modelo completo de regresión.

Modelo 1	B_j	Error típico	Estadístico t	Probabilidad
Intercepción	-2,316	1,594	-1,453	0,1489
lcfes (x_1)	0,079	0,020	3,994	0,0001
Rcobertura salud	-0,048	0,106	-0,449	0,6540
R relaciones familiares (x_3)	-0,135	0,134	-1,007	0,3162
R salud física	-0,144	0,121	-1,189	0,2369
R nutricional (x_5)	0,137	0,133	1,026	0,3072
R alteración mental (x_6)	-0,404	0,286	-1,413	0,1604
R sustancias psicoactivas (x_7)	0,646	0,246	2,626	0,0098
R uso del tiempo libre (x_8)	-0,130	0,165	-0,787	0,4329
Lectura nivel literal (x_9)	0,365	0,131	2,793	0,0061
Lectura nivel inferencial (x_{10})	0,162	0,129	1,253	0,2127
Lectura nivel intertextual (x_{11})	0,046	0,135	0,340	0,7346

Tabla 3. Coeficientes de regresión para el modelo completo

Anova	Grados de libertad	Suma de cuadrados	Promedio de los cuadrados	F	Valor crítico de F
Regresión	11	54,84	4,99	5,16	0,0000015
Residuos	116	112,17	0,97		
Total	127	167,01			

Tabla 4. Anova para el modelo completo

Las tablas anteriores aportan información sobre los coeficientes del modelo de regresión lineal múltiple. El modelo como tal es el siguiente:

$$Y = \sum_{j=1}^{12} B_j X_j + B_0 \quad [1]$$

Al analizar los estadísticos de prueba t asociados a cada coeficiente de las variables, se puede concluir que para

un nivel de confianza del 95%, solo de los parámetros B_1 , B_7 y B_9 se puede tener evidencia suficiente de que serán diferentes de cero y que por tanto aportan algún tipo de información al modelo.

El modelo visto de manera general no presenta problemas de auto correlación. Por otro lado el modelo general explica 32,8% de la variabilidad de las notas en matemáticas. Al realizar la prueba F se concluye que hay evidencia suficiente que garantice la aplicabilidad del modelo, aún cuando este no ofrezca un alto coeficiente de determinación.

La idea de esta aplicación no es ofrecer un modelo óptimo para el problema del pronóstico de la nota en matemáticas, si no más bien tener un modelo del cual se tengan garantías de que es parcialmente predictor de la variable ya mencionada.

2.3. MODELO DE REGRESION SIMPLIFICADO

Del modelo anterior para predecir la nota en matemáticas solo aportan la información significativa las variables: el porcentaje total del Icfes (y no sólo en matemáticas), riesgo de sustancias psicoactivas y nivel de lectura literal, dado que son las únicas que pasan las pruebas t correspondientes, por lo cuál deben ser tomadas como centrales, no sólo en la formación del modelo predictivo, sino también durante la formulación de programas de prevención del riesgo de repitencia.

En las tablas 5 y 6 se muestra el modelo simplificado donde solamente se utilizan estas tres variables

Modelo 2	Coefficientes	Error típico	Estadístico t	Probabilidad
Intercepción	-3,814	1,126	-3,387	0,001
Icfes	0,085	0,019	4,407	0,000
Rsubstancias psicoactivas	0,520	0,231	2,251	0,026
Lectura nivel literal	0,398	0,126	3,159	0,002

Tabla 5. Coeficientes de regresión para el modelo simplificado

Anova	Grados de libertad	Suma de cuadrados	Promedio de los cuadrados	F	Valor crítico de F
Regresión	3	47,290	15,763	16,33	0,00000
Residuos	124	119,716	0,965		
Total	127	167,007			

Tabla 6. Anova para el modelo simplificado

La ecuación de este modelo sería:

$$Y = -3,814 + 0,085X_1 + 0,520X_2 + 0,398X_3 \quad [2]$$

Para un estudiante en particular que registre por ejemplo valores en las variables del siguiente tipo, se tiene la siguiente predicción de la nota:

Puntaje Icfes..... 56,5
 Riesgo sustancias psicoactivas..... 3
 (1-alto, 2-medio, 3-bajo)
 Lectura a nivel literal..... 3
 (1- Deficiente, 2 - Insuficiente, 3 - Aceptable, 4 -Bueno, 5 - Excelente)

$$Y = -3,814 + 0,085 * (56,5) + 0,520 * (3) + 0,398 * (3)$$

$$Y = 3,7$$

En tales condiciones se espera que la nota del estudiante en la asignatura de matemáticas sea aproximadamente 3,7.

La relación entre el rendimiento académico en las matemáticas con en nivel de riesgo de consumo de sustancias psicoactivas se puede evidenciar en la figura 1 donde se gráfica la nota promedio en esta materia y nivel de riesgo de consumo de sustancias psicoactivas. Con mayor riesgo de consumo de sustancias psicoactivas menor es la nota en matemáticas.

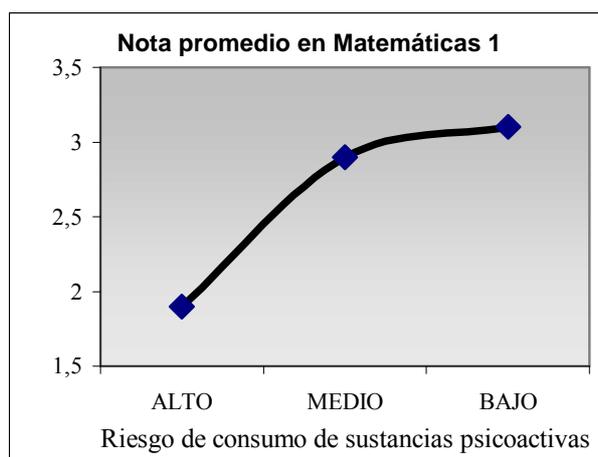


Figura 1. Relación entre riesgo de consumo de sustancias psicoactivas y la nota promedio en Matemáticas 1.

Nivel de lectura literal y la nota promedio en matemáticas I tienen una relación mostrada en la figura 2. Los estudiantes que presentan mejores niveles de lectura literal obtuvieron en promedio, las notas más altas en matemática I.

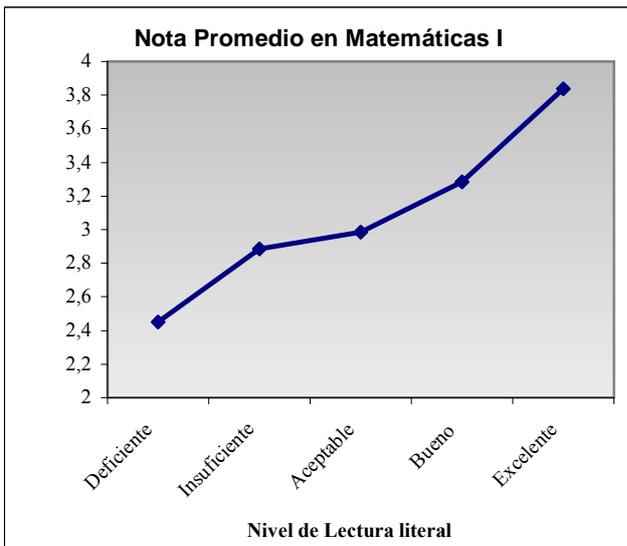


Figura 2 Relación entre el Nivel de lectura Literal y la nota promedio en Matemáticas I

2.4. APLICACIÓN DEL MODELO SIMPLIFICADO

El modelo simplificado en forma de ecuación 3 especificando las variables que aportan la información significativa fue aplicado para calcular la nota en la asignatura Matemática I para todos los estudiantes de ingenierías y tecnologías del segundo periodo del año 2007.

$$Y = -3,814 + 0,085 * Icfes + 0,520 * Riesgo_sustancias_psicoactivas + 0,398 * Nivel_de_lectura_literal \quad [3]$$

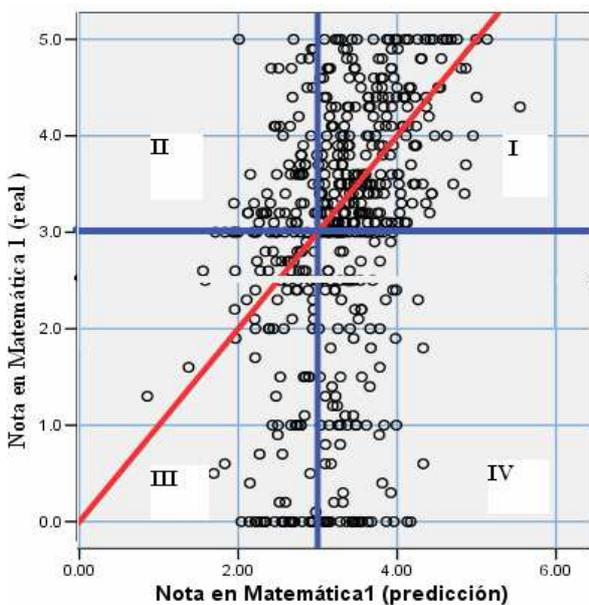


Figura 3. Gráfica nota matemáticas I (predicción) contra la nota real.

La figura 3 muestra la nota en matemáticas calculada con la fórmula de predicción (ecuación 3) contra la nota real que obtuvo el estudiante. En el cuadrante I están ubicados los estudiantes que pasaron la asignatura y la predicción fue que la deben ganar. Es el cuadrante de la predicción correcta. En el cuadrante III están ubicados los estudiantes que perdieron la materia y la predicción para ellos fue que la van a perder. De nuevo es el cuadrante donde el pronóstico se cumple. Los cuadrantes II y IV son donde la predicción fue incorrecta, es decir la zona de falsos positivos y falsos negativos.

Pasó Matemática 1 (real)	Pasó Matemática 1 (predicción)			% predicciones correctas
	No pasó	Pasó	Total	
No pasó	294	192	486	65,5%
Pasó	155	267	422	58,2%
Total	449	459	908	61,8%

Tabla 7. Resultados de las predicciones

Del total de 908 estudiantes que deberían ver la Matemática 1 se acertaron correctamente 65,5% de los estudiantes que no pasarán esta asignatura y 58,2% de los que la pueden pasar. En total 61,8% de las predicciones fueron correctas. En este resultado están incluidos los estudiantes que se salieron sin terminar la materia (nota real en la gráfica es igual a cero), por lo cual las predicciones del modelo pueden ser aun superiores al estudiar el factor de riesgo de que la asignatura sea cancelada.

3. CONCLUSIONES Y RECOMENDACIONES

Los análisis de regresión lineal múltiple arrojan información valiosa que permite hacer conjeturas sobre las variables de mayor significancia en el momento de la predicción de una variable de interés.

En el caso puntual de los modelos aquí aplicados, la exploración de la situación problema por medio de los métodos de regresión lineal múltiple permitió identificar la importancia del puntaje total de Icfes en el momento de explicar las notas de matemáticas de los estudiantes de ingenierías y tecnologías del segundo periodo del año 2007 de la Universidad Tecnológica de Pereira. La variable ICFES guarda relación directa con la nota definitiva en matemáticas.

Se encontraron en total tres variables que aportan información a los modelos planteados y pasan las pruebas que se les aplicó. Estas variables son: puntaje total de Icfes, Nivel de Lectura Literal y Riesgo de consumo de sustancias psicoactivas.

El riesgo de consumo de sustancias psicoactivas aumenta la probabilidad de la pérdida de matemáticas I igualmente sucede con un niveles deficientes o insuficientes en lectura literal.

Los programas de control de deserción estudiantil deben prestar especial atención a estas variables en el momento de desarrollar sus programas de mejoramiento.

Con el modelo de regresión simplificado se puede predecir al principio del semestre cuales de los estudiantes tienen alto riesgo de perder la asignatura de matemáticas para conformar grupos de monitoreo o tutorías, y programas individuales de lectura y análisis de textos.

4. BIBLIOGRAFÍA

[1] P. Carvajal, A. A. Trejos, " Búsqueda de la relación entre áreas Icfes en matemáticas, física, lenguaje y rendimiento en matemáticas I y matemáticas II a través del análisis de componentes principales", *Scientia et Technica* Año X, No 26, Diciembre 2004. UTP. ISSN 0122-1701.pp.133-138. [Online]. Available: <http://www.utp.edu.co/php/revistas/ScientiaEtTechnica/docsFTP/11363133-138.pdf>

[2] P. Carvajal, A. A. Trejos, Soto J. "Aplicación del análisis discriminante para explorar la relación entre el examen de ICFES y el rendimiento en algebra lineal de los estudiantes de ingeniería de la UTP en el período 2001-2003, " *Scientia et Technica* Año X, No 25, Agosto 2004. UTP. ISSN 0122-1701, pp.191-196. [Online]. Available: <http://www.utp.edu.co/php/revistas/ScientiaEtTechnica/docsFTP/14951191-196.pdf>

[3] P. Carvajal, A. A. Trejos, and C. A. Caro "Estudio de causas de deserción de los estudiantes de la Universidad Tecnológica de Pereira entre enero/2000 – diciembre/2004 utilizando la técnica de análisis de correspondencias simples," *Scientia et Technica* Año XII, No 30, mayo 2006. UTP. ISSN 0122-1701, pp.261-266. [Online]. Available: <http://www.utp.edu.co/php/revistas/ScientiaEtTechnica/docsFTP/101731261-266.pdf>

[4] P. Carvajal, A. A. Trejos, A. M. Barros "Clusters de estudiantes nuevos matriculados a la universidad tecnológica de pereira, según estrategias de aprendizaje", *Scientia et Technica* Año XIII, No 35, Agosto de 2007. Universidad Tecnológica de Pereira. ISSN 0122-1701, pp. 315-320. [Online]. Available: <http://www.utp.edu.co/php/revistas/ScientiaEtTechnica/docsFTP/233620315-320.pdf>

[5] Walpole, Ronald E *Probabilidad y Estadística para ingenieros*, vol. 6. Mexico : Wiley, 1999, p. 405-456