

**COMPONENTES PRINCIPALES Y REGRESIÓN
LOGÍSTICA: ANALIZANDO EL RENDIMIENTO
ACADÉMICO DE LOS ESTUDIANTES EN
MATEMÁTICA PREBÁSICA**

Por

Claudia Liliana Rodríguez Morantes

Tesis sometida en cumplimiento parcial de los requisitos para el grado de

**MAESTRÍA EN CIENCIAS
en
MATEMÁTICAS (ESTADÍSTICA)**

**UNIVERSIDAD DE PUERTO RICO
RECINTO UNIVERSITARIO DE MAYAGUEZ
2011**

Aprobado por:

Dámaris Santana Morant, Ph.D.
Miembro, Comité Graduado

Fecha

Tokuji Saito, Ph.D.
Miembro, Comité Graduado

Fecha

Edgardo Lorenzo González, Ph.D.
Presidente, Comité Graduado

Fecha

Henrick Mario Ierkic, Ph.D.
Representante de Estudios Graduados

Fecha

Silvestre Colón, M.S.
Director de Departamento

Fecha

ABSTRACT

In this research we analyze the factors associated with the academic performance of students in the course of Pre-Basic Mathematics (I semester 2010-2011), from the information obtained through three questionnaires, applied to a random sample of 404 students of the course. First, we did an exploratory analysis of the data by descriptive statistics, which allows the construction of the student's profile. Second, we apply the multivariate technique of principal component analysis, which allowed us to define three new uncorrelated variables, from the quantitative variables of the study. Third, these three variables and some highlighted categorical variables in the profile were used to construct a prediction model for approval of the students in the course, from the multiple logistic regression technique. It was found that the fitted model is adequate and has a classification rate of 75.1%. The research results indicate that the "previous academic preparation, "the "inertia towards mathematics", the "perceived exertion" and "anxious trend" are the most influential factors in the success of students in this course.

RESUMEN

En esta investigación se analizan los factores asociados con el rendimiento académico de los estudiantes, en el curso de Matemática Prebásica (I semestre 2010-2011), a partir de la información obtenida a través de tres cuestionarios, aplicados a una muestra aleatoria de 404 estudiantes del curso. Primero, se hizo un análisis exploratorio de los datos, mediante estadísticas descriptivas, que permitió la construcción del Perfil de los estudiantes. Segundo, se aplica la técnica multivariada de análisis de componentes principales, la cual permitió definir tres nuevas variables no correlacionadas, a partir de las variables cuantitativas del estudio. Tercero, estas tres variables y algunas variables categóricas que resaltaron en el perfil se utilizaron para construir un modelo de predicción de aprobación de los estudiantes en el curso, a partir de la técnica de Regresión Logística Múltiple. Se encontró que el modelo ajustado es adecuado y tiene una tasa de clasificación de 75.1%. Los resultados de la investigación indican que la “Preparación académica previa”, la “Inercia hacia las matemáticas”, la “Percepción del esfuerzo” y la “Tendencia ansiosa”, son factores que tienen mayor influencia en el éxito de los estudiantes de este curso.

Copyright © 2011

Por

Claudia Liliana Rodríguez Morantes

A Oscar Yesid, por su amor y apoyo incondicional.

A mis padres y hermanos, por animarme durante todo el proceso.

AGRADECIMIENTOS

A Dios, por sus bendiciones y por ser mi soporte a lo largo de esta experiencia.

Al Dr. Edgardo Lorenzo, mi asesor de tesis, por su orientación durante todo el proceso.

A Oscar Yesid por su constante apoyo y sugerencias en el análisis de algunos resultados.

Al Director del Departamento de Ciencias Matemáticas, profesor Silvestre Colón y al coordinador de los cursos de Matemática Prebásica, Dr. Pedro Vásquez, por permitir la aplicación de los cuestionarios y por la información suministrada.

A los estudiantes matriculados en el curso de MATE 0066, durante el I semestre de 2010, por la información proporcionada y que fue analizada en esta tesis.

Al personal administrativo del Departamento de Ciencias Matemáticas por la colaboración recibida durante la maestría.

A mis compañeros y amigos, estudiantes graduados, por hacer de esta experiencia un continuo aprendizaje.

TABLA DE CONTENIDO

ABSTRACT.....	ii
RESUMEN.....	iii
TABLA DE CONTENIDO.....	vii
LISTA DE TABLAS.....	viii
LISTA DE FIGURAS.....	ix
1. INTRODUCCIÓN.....	1
1.1. Justificación.....	3
1.2. Antecedentes.....	5
1.3. Objetivos.....	9
1.3.1. Objetivo General.....	9
1.3.2. Objetivos Específicos.....	9
1.4. Resumen de los siguientes capítulos.....	10
2. REVISIÓN DE LITERATURA.....	11
2.1. Componentes Principales.....	11
2.1.1. Descripción de la Técnica de Análisis de Componentes Principales.....	12
2.1.2. Propiedades de los Componentes Principales.....	16
2.1.3. Selección del número de Componentes Principales.....	17
2.2. Regresión Logística.....	18
2.2.1. Modelo de Regresión Logística Simple.....	18
2.2.2. Estimación de los parámetros del modelo.....	20
2.2.3. Modelo de Regresión Logística Múltiple.....	21
2.2.4. Estimación de los parámetros del modelo.....	22
2.2.5. Medidas de confiabilidad del modelo.....	23
3. METODOLOGÍA.....	25
3.1. Instrumentos.....	25
3.2. Población y Muestra.....	28
3.3. Procedimiento.....	30
4. RESULTADOS.....	32
4.1. Descripción del perfil de los estudiantes.....	32
4.1.1. Características sociodemográficas.....	32
4.1.2. Trasfondo Académico.....	36
4.1.3. Aspiraciones académicas.....	41
4.1.4. Gusto hacia las matemáticas.....	43
4.1.5. Ansiedad hacia las matemáticas.....	45
4.1.6. Actitud hacia las matemáticas.....	49
4.1.7. Motivación y planes al graduarse.....	52
4.2. Construcción de categorías a partir del Análisis de Componentes Principales.....	54
4.3. Construcción de un índice para predecir éxito o fracaso de los estudiantes a partir de la Regresión Logística.....	65
5. CONCLUSIONES Y TRABAJOS FUTUROS.....	71
REFERENCIAS	
ANEXOS	

LISTA DE TABLAS

Tabla 1.1 Relación de No Aprobados y Bajas - estudiantes de MATE 0066.....	6
Tabla 4.1 Preparación académica de los padres.....	36
Tabla 4.2 Preparación de los estudiantes en los cursos de matemáticas.....	37
Tabla 4.3 Variables académicas de Matemática Prebásica.....	39
Tabla 4.4 Resultados Prueba de Levence	40
Tabla 4.5 Promedio de Escuela Superior, Aptitud y Aprovechamiento según género.....	41
Tabla 4.6 Promedio de Escuela Superior, Aptitud y Aprovechamiento según escuela.....	41
Tabla 4.7 Factores Escala de Ansiedad.....	45
Tabla 4.8 Estadísticas descriptivas ansiedad según género de los estudiantes.....	47
Tabla 4.9 Razones más importantes para ingresar a la Universidad.....	52
Tabla 4.10 Matriz de Correlaciones.....	55
Tabla 4.11 Valores propios y varianza explicada por los componentes principales.....	56
Tabla 4.12 Vectores propios.....	57
Tabla 4.13 Variables que entraron al modelo.....	66
Tabla 4.14 Codificación de la variable.....	66
Tabla 4.15 Coeficientes del mejor modelo.....	67
Tabla 4.16 Clasificación de los estudiantes.....	68

LISTA DE FIGURAS

Figura 1.1 Gráfico del Modelo Logístico.....	19
Figura 4.1 Distribución de estudiantes en MATE 0066 según el género.....	33
Figura 4.2 Porcentaje de estudiantes de nuevo ingreso.....	34
Figura 4.3 Porcentaje de estudiantes de MATE 0066 según la escuela de procedencia.....	34
Figura 4.4 Distribución de la ayuda económica que reciben los estudiantes de MATE 0066..	35
Figura 4.5 Distribución preparación académica de los padres.....	36
Figura 4.6 Distribución horas de estudio/semana de matemática en la Escuela.....	38
Figura 4.7 Distribución horas de estudio /semana en el curso de MATE 0066.....	38
Figura 4.8 Distribución aspiraciones académicas de los estudiantes según el género.....	42
Figura 4.9 Distribución aspiraciones académicas de los estudiantes según preparación académica de los padres.....	43
Figura 4.10 Distribución “me preparo antes de presentar los exámenes de MATE 0066”... según horas dedicadas a estudiar.....	44
Figura 4.11 Distribución de ansiedad ante los exámenes.....	46
Figura 4.12 Distribución de ansiedad ante la eventualidad del examen.....	46
Figura 4.13 Distribución ansiedad ante situaciones matemáticas de la vida real.....	47
Figura 4.14 Ansiedad General por Facultad	48
Figura 4.15 Distribución tipo de mentalidad de los estudiantes, según el género.....	50
Figura 4.16 Distribución tipo de mentalidad según la facultad.....	50
Figura 4.17 Distribución tipo de mentalidad según horas de estudio de Matemáticas Prebásicas.....	51
Figura 4.18 Distribución ansiedad según tipo de mentalidad.....	51
Figura 4.19 Distribución razones para ingresar a la Universidad.....	52
Figura 4.20 Distribución planes de los estudiantes al graduarse.....	53
Figura 4.21 Gráfica del número de componentes según su valor propio.....	57
Figura 4.22 Plano definido por los componentes 1 y 2.....	60
Figura 4.23 Plano definido por los componentes 1 y 3.....	61
Figura 4.24 Plano definido por los componentes 2 y 3.....	62
Figura 4.25 Representación de las variables en el plano de los componentes 1 y 2.....	63
Figura 4.26 Representación de las variables en el plano de los componentes 1 y 3.....	64

CAPÍTULO 1

INTRODUCCIÓN

Uno de los problemas en el campo de la educación y áreas afines, que ha llamado la atención a investigadores (Carvajal, Mosquera & Artamonova, 2009; Petriz, Barona, López & Quiroz, 2010), es el bajo rendimiento académico de los estudiantes en los cursos de matemáticas, en todos los niveles. Problemática que parece incrementar con el paso del tiempo y que se ha buscado explicar estudiando diversas variables, como por ejemplo las relacionadas con la fobia o ansiedad hacia las matemáticas (Roty, 2008), la falta de interés o motivación (Petriz, Barona, López, & Quiroz, 2010), el nivel académico, la carencia de hábitos de estudio, entre otras. El análisis de dicha problemática ha conllevado la realización de cambios en términos de currículo, técnicas de enseñanza e implementación de estrategias buscando mejorar los resultados.

El Recinto Universitario de Mayagüez (RUM) no es ajeno a esta problemática, según un estudio de Calificaciones y Fracagos de la Oficina de Investigación Institucional y Planificación [OIIP], compilado con datos desde 1994 a 2005, durante ese periodo el 46% de las notas obtenidas en los cursos que ofrece el Departamento de Ciencias Matemáticas había resultado en D, F, W (bajas). A finales del 2005, el mismo departamento coordinó el *Primer Simposio sobre la retención en los cursos de precálculo y métodos cuantitativos en la Universidad de Puerto Rico (UPR)* ante la preocupación por el aumento del fracaso de los

estudiantes en los cursos de matemáticas (Santori, 2005). En ese sentido, se han realizado esfuerzos que apuntan a atender esta problemática, aún así, el panorama no es muy alentador.

En esta tesis se identificaron y analizaron una serie de variables, tanto cualitativas (escala ordinal y nominal) como cuantitativas, que influyen en el bajo desempeño académico de los estudiantes del curso de Matemática Prebásica que ofrece el Departamento de Ciencias Matemáticas del RUM. Algunas de ellas fueron: características sociodemográficas, trasfondo educativo, hábitos de estudio, ansiedad a las matemáticas, aptitud académica, actitud hacia las matemáticas, entre otras.

Primero se generó el perfil de los estudiantes de Matemática Prebásica, a partir de las estadísticas descriptivas sobre cada una de las variables del estudio. Segundo, se analizaron las variables cuantitativas: promedio de escuela superior, puntuación en aptitud matemática del College Board, puntuación en aprovechamiento del College Board, puntuación en escala de ansiedad y promedio de quices electrónicos, a partir de la técnica multivariada de Componentes Principales, lo que permitió la construcción de tres nuevas variables (componentes). Finalmente, se aplicó la técnica de Regresión Logística para predecir la aprobación o no del curso de Matemática Prebásica, para ello se definió la variable dependiente dicotómica (1 aprobó, 0 no aprobó) y se utilizaron como variables explicativas los tres componentes principales y las variables cualitativas mencionadas. Para la selección del mejor subconjunto de variables se utilizó el método de selección hacia atrás, determinando el modelo predictivo.

1.1 Justificación

Existe una realidad asociada con el bajo rendimiento académico de los estudiantes del RUM en los cursos de matemáticas que ofrece el Departamento de Ciencias Matemáticas, especialmente para los estudiantes de nuevo ingreso, en particular el curso de Matemática Prebásica.

Una evidencia de ello es la alta tasa de estudiantes que deben tomar este curso, aproximadamente el 55% de los estudiantes de nuevo ingreso, que han obtenido menos de 650 puntos en la prueba de aprovechamiento en matemáticas del Examen del College Board, ó, no han aprobado el examen diagnóstico administrado por el Departamento de Ciencias Matemáticas.

Son muchos los esfuerzos del departamento (taller de verano, implementación de quices electrónicos, implementación de un libro de texto, Centro de Apoyo, etc.) para lograr que los estudiantes adquieran las destrezas matemáticas mínimas, que les permita tener éxito en los cursos posteriores de matemáticas. Sin embargo, en la mayoría de los casos, estos esfuerzos se enfocan en el aspecto académico y hasta el momento no se muestra un cambio significativo en los resultados; lo que conlleva a la consideración de otras variables que también pueden estar afectando dichos resultados y a las que no se les ha prestado suficiente atención.

En ese sentido, estudios previos (Roty, 2008; Spybrook, 2008) han identificado que la preparación académica previa, la aptitud hacia la matemática, la ansiedad hacia las matemáticas, la actitud hacia la matemática, son factores determinantes en el rendimiento académico de los estudiantes y que influyen en los resultados que obtienen en sus cursos de matemáticas.

Se reconoce que el problema de rendimiento académico en matemáticas es complejo y que son muchas las variables que inciden en él, muchas de ellas posiblemente correlacionadas. Para analizar la problemática se siguieron las siguientes etapas: primero, elaboración del perfil de los estudiantes de Matemática Prebásica, identificando con ello algunos de los posibles factores que inciden en su rendimiento académico; segundo, definiendo unos componentes principales que revelaran la estructura de estas variables, reduciendo la dimensionalidad de los datos recopilados y tercero, analizando la importancia de éstos componentes y las demás variables del estudio, en un modelo de predicción del rendimiento académico de los estudiantes de Matemática Prebásica.

1.2 Antecedentes

En noviembre de 2005 se llevó a cabo en el Recinto Universitario de Mayagüez (RUM), el Primer Simposio sobre la retención en los cursos de precálculo y métodos cuantitativos en la Universidad de Puerto Rico (UPR). Una de las preguntas centrales era: ¿por qué la mayoría de los jóvenes que ingresa a la Universidad fracasan o se dan de baja, de los cursos de matemáticas, en primer año? Una de las hipótesis que se manejaba, es que muchos jóvenes de nuevo ingreso no cuentan con las destrezas necesarias en matemáticas ya que sólo toman cursos elementales de álgebra o geometría en las escuelas y no cursos de precálculo.

De acuerdo a la Certificación 99-15 del Senado Académico todo estudiante de nuevo ingreso que obtenga una puntuación menor de 650 en la prueba de Aprovechamiento Matemático del examen del College Board, deberá tomar un examen diagnóstico que evaluará sus conocimientos matemáticos básicos y de aprobarlo, podrá ingresar al primer curso de matemáticas según su concentración, de no aprobarlo ingresará al curso de Matemática Prebásica (MATE 0066). En la Tabla 1.1, se presenta el porcentaje de fracasos y bajas de los estudiantes del curso de Matemática Prebásica, en los últimos 5 años, según estadísticas del Departamento de Matemáticas.

Tabla 1.1 Relación de No Aprobados y Bajas – estudiantes MATE 0066

Semestre	Porcentaje No Aprobados	Porcentaje Bajas
I - 2006	40.0%	12.3%
I - 2007	35.1%	5%
I - 2008	40.5%	4%
I - 2009	60.9%	10%
I - 2010	56.77	3%

El interés de este estudio, como se mencionó antes, recae sobre los estudiantes que ingresan al curso de Matemática Prebásica, del Departamento de Ciencias Matemáticas, que corresponde a los estudiantes que obtuvieron menos de 650 puntos en el examen de aprovechamiento en matemáticas, del College Board y que no aprobaron el examen diagnóstico. Pero la problemática es más compleja porque a pesar de ofrecer ese curso de destrezas básicas a los estudiantes y proporcionar otras ayudas: la práctica a través de pruebas electrónicas, apoyo por parte de los instructores (estudiantes graduados), la implementación de un libro diseñado específicamente como apoyo al curso, entre otras, los resultados no son los esperados.

En la búsqueda de variables que expliquen los resultados anteriores se encontraron algunos estudios (Spybrook, 2008; Roty, 2008) que han identificado variables que afectan el rendimiento académico de los estudiantes. En el primer estudio (Spybrook, 2008), se investigaron principalmente dos factores: la memoria de trabajo y su relación con el rendimiento en matemáticas, y la ansiedad matemática y su relación con la memoria de trabajo y el rendimiento académico y una de las conclusiones de este estudio es que se deben implementar estrategias para disminuir la carga cognitiva de los estudiantes, en los cursos de

matemáticas, para acceder más eficazmente a la memoria de trabajo y aumentar el logro en dichos cursos. En el segundo estudio (Roty, 2008), se examinó la relación entre la ansiedad matemática y la inteligencia emocional, encontrando que los estudiantes que obtenían puntajes altos en la escala de ansiedad, tenían puntajes bajos en la inteligencia emocional, es decir, que la correlación entre las variables era negativa. Además resaltan que un alto nivel de ansiedad hacia las matemáticas, interfiere en la manipulación de números y en la solución de problemas, teniendo más dificultades estos estudiantes en los cursos de matemáticas e incluso, evitando carreras universitarias en las áreas de ciencias e ingeniería donde el mayor énfasis es la matemática.

Otra de las variables analizadas fue la actitud hacia las matemáticas, que se abordó a partir de los estudios sobre la Teoría de la Mentalidad de Dweck (2006). La autora, a partir de años de investigación, establece que las personas pueden tener una mentalidad fija o de crecimiento, la cual estaría asociada con la forma cómo se concibe la inteligencia; por ejemplo, una persona con mentalidad fija cree que su inteligencia es un rasgo fijo, por el contrario, las personas con mentalidad de crecimiento tienden a desarrollarse y trabajar mucho más, porque creen que sus habilidades básicas se pueden desarrollar.

En otra investigación (Díaz, González, Torres & Garnica, 1994) aplicaron el Análisis de Componentes Principales para analizar la influencia de la metodología de enseñanza y de la preparación previa del estudiante, sobre el aprendizaje. En el estudio se describen algunas de las aplicaciones más frecuentes de la técnica: análisis exploratorio, construcción de indicadores, utilización de los indicadores (componentes) en estudios posteriores, como por

ejemplo, el análisis de regresión, entre otros. En el análisis exploratorio obtuvieron dos grupos de variables: uno relacionado con la inteligencia y el otro con el rendimiento estudiantil. La estimación de los constructos (aprendizaje e inteligencia) se realizó mediante indicadores sintéticos, producto del análisis. Además, efectuaron un análisis de regresión múltiple con el que comprobaron la influencia significativa de la inteligencia y metodología de la enseñanza, sobre el aprendizaje.

Finalmente es importante resaltar que una de las técnicas que ha sido utilizada por algunos autores (García, 2000; Porcel, 2010) para el análisis de los factores asociados con el rendimiento académico es la regresión logística, recomendada por su alto porcentaje de clasificación. En el primer estudio, se confirmó que el rendimiento académico previo de los estudiantes o promedio de bachillerato, es un buen predictor del rendimiento académico futuro. También encontraron que variables como la asistencia y participación son significativas en el modelo de predicción del rendimiento académico, lo que consideraron relevante pues son variables sobre las que se puede intervenir. En el segundo estudio, utilizaron la técnica de regresión logística, para predecir el rendimiento académico de los estudiantes, en función de sus características socioeducativas y encontraron, por ejemplo, que el nivel académico de los padres resultó significativo y que la variable más significativa es el título secundario del alumno, es decir, el énfasis en su programa académico y su preparación académica pre universitaria.

1.3 Objetivos

1.3.1 Objetivo General

Aplicar la técnica multivariada de componentes principales para analizar los factores asociados con el bajo rendimiento académico de los estudiantes de Matemática Prebásica (MATE 0066), de modo que se entienda la estructura que gobierna esta problemática, desde una perspectiva cuantitativa, e identificar los componentes que tienen mayor incidencia para cuantificar la probabilidad de éxito de estos estudiantes a través de un modelo de regresión logística.

1.3.2 Objetivos Específicos

- Elaborar y aplicar un cuestionario a los estudiantes del curso de Matemática Prebásica que permita la construcción de su perfil.
- Aplicar la técnica de componentes principales para entender la realidad de los factores que inciden sobre su rendimiento académico.
- Predecir la probabilidad de éxito de los estudiantes de Matemática Prebásica, utilizando el modelo de regresión logística, a partir de las variables cuantitativas y cualitativas que se hayan identificado en el estudio.
- Contrastar los factores que ha atendido el Departamento de Ciencias Matemáticas en sus iniciativas, con los factores identificados en el estudio sobre el rendimiento académico de los estudiantes de Matemática Prebásica.

1.4 Resumen de los siguientes capítulos

- En el capítulo 2, se exponen las técnicas estadísticas utilizadas para el análisis. Se hace una descripción de la técnica de análisis multivariado: Componentes Principales y de la Regresión Logística Múltiple.
- En el capítulo 3, se presenta la metodología incluyendo los instrumentos utilizados en este estudio, se define la población y muestra, así como el procedimiento para su selección y posterior recolección de la información.
- En el capítulo 4, se presentan los resultados obtenidos: primero se hace un análisis descriptivo que permitió la elaboración del perfil de los estudiantes; segundo, se presentan los resultados del análisis de componentes principales, con la descripción de las nuevas variables y finalmente, los resultados del análisis de la regresión logística para predecir el éxito o fracaso de los estudiantes en el curso de Matemática Prebásica.
- En el capítulo 5 se presentan las conclusiones y recomendaciones de la tesis.

CAPÍTULO 2

REVISIÓN DE LITERATURA

A continuación se presenta una descripción de las técnicas estadísticas que se emplearon en este estudio:

2.1 Componentes Principales

El análisis de componentes principales es una de las técnicas de análisis multivariado más difundidas, que permite la reestructuración de un conjunto de datos multivariado, obtenido de una población cuya distribución de probabilidad no necesita ser conocida (Pla, 1986), en un conjunto de variables que son combinaciones lineales de las originales, llamadas componentes principales y que contienen la mayor parte de la variabilidad de los datos originales.

La técnica fue propuesta por Hotelling (1933), aunque sus orígenes se encuentran en los ajustes ortogonales por mínimos cuadrados introducidos por K. Pearson (1901) citado en Peña (2002). Entre los objetivos y usos más importantes del análisis por componentes principales se destacan:

- Estudiar interrelaciones entre las variables y de acuerdo con los resultados proponer los análisis estadísticos más apropiados, esto como técnica de análisis exploratorio.

- Representar en un espacio de menor dimensión observaciones de un espacio general p -dimensional.
- Transformar las variables originales, generalmente correlacionadas, en variables no correlacionadas, eliminando el problema de multicolinealidad y permitiendo una mejor interpretación de las variables en análisis posteriores, como el de la regresión.

2.1.1 Descripción de la técnica de análisis de componentes principales

Sea X una matriz de dimensiones $n \times p$, asociada a n observaciones (filas) y p variables (columnas), donde cada observación se ha centralizado, por lo que la matriz X tiene media cero y matriz de covarianzas $S = \frac{X'X}{n}$.

El problema que se desea resolver, desde un enfoque estadístico, es encontrar un espacio de dimensión más reducida que represente adecuadamente los datos. Es decir, representar puntos p -dimensionales con la mínima pérdida de información, en un espacio de dimensión $r < p$. Según Peña (2002) es equivalente a sustituir las p variables originales por nuevas variables z_1, z_2, \dots, z_r , las cuales tienen máxima correlación con las variables originales, pero no están correlacionadas entre ellas mismas, propiedad deseable, en análisis posteriores como el de regresión múltiple.

Para encontrar estas nuevas variables se deben obtener los vectores propios de la matriz de covarianzas S , denotados por α_i , $i = 1, 2, \dots, r$, ya que las proyecciones de las observaciones sobre estos vectores propios definen la dirección de los nuevos ejes o componentes

principales. La existencia de los r valores propios (reales y diferentes) está garantizada porque S es simétrica y definida positiva. A continuación se describe el proceso:

El primer componente principal será la combinación lineal de las variables originales, que tiene varianza máxima. Los valores de este primer componente para n observaciones se representarán por el vector z_1 , dado por $z_1 = X\alpha_1$, donde α_1 es el vector propio asociado al mayor valor propio de S . Para ver esto note que las variables originales tienen media cero, por lo que z_1 también tiene media cero y varianza dada por;

$$var(z_1) = \frac{z_1'z_1}{n} = \frac{\alpha_1'X'X\alpha_1}{n} = \alpha_1'S\alpha_1 \quad (1)$$

Esta varianza se puede maximizar sin límite, ya que depende del módulo del vector α_1 , así que sin pérdida de generalidad se introduce la restricción $\alpha_1'\alpha_1 = 1$, usando multiplicadores de Lagrange. Entonces

$$M = \alpha_1'S\alpha_1 - \lambda_1(\alpha_1'S\alpha_1 - 1) \quad (2)$$

y maximizando esta expresión, con respecto a α_1 se obtiene $\frac{\partial M}{\partial \alpha_1} = 2S\alpha_1 - 2\lambda_1\alpha_1 = 0$, cuya solución es:

$$S\alpha_1 = \lambda_1\alpha_1 \quad (3)$$

Lo que implica que α_1 es un vector propio de la matriz S , y λ_1 su correspondiente valor propio. Para determinar qué valor propio de S es la solución de (3) se multiplica por α_1' esta ecuación, $\alpha_1'S\alpha_1 = \lambda_1\alpha_1'\alpha_1 = \lambda_1$ y se concluye por (1) que λ_1 es la varianza de z_1 . Por lo que λ_1 es el mayor valor propio de S y su vector asociado α_1 define los coeficientes de cada variable en el primer componente principal.

Un razonamiento similar se usa para obtener el segundo componente principal. Para obtener el mejor plano de proyección de las variables en la matriz X, se establece como objetivo que la suma de las varianzas de $z_1 = X\alpha_1$ y $z_2 = X\alpha_2$, sea máxima, donde α_1 y α_2 definen el plano. Entonces, como se definió antes $var(z_1) = \alpha_1' S \alpha_1$ y similarmente $var(z_2) = \alpha_2' S \alpha_2$. Por lo que la función objetivo será:

$$\phi = \alpha_1' S \alpha_1 + \alpha_2' S \alpha_2 - \lambda_1(\alpha_1' S \alpha_1 - 1) - \lambda_2(\alpha_2' S \alpha_2 - 1) \quad (4)$$

que también impone restricciones sobre los módulos de los vectores α_1 y α_2 , $\alpha_i' \alpha_i = 1$, para $i = 1, 2$. Derivando e igualando a cero obtenemos

$$\frac{\partial \phi}{\partial \alpha_1} = 2S\alpha_1 - 2\lambda_1\alpha_1 = 0 \quad (5)$$

$$\frac{\partial \phi}{\partial \alpha_2} = 2S\alpha_2 - 2\lambda_2\alpha_2 = 0 \quad (6)$$

Las soluciones del sistema de ecuaciones son

$$S\alpha_1 = \lambda_1\alpha_1 \quad (7)$$

$$S\alpha_2 = \lambda_2\alpha_2 \quad (8)$$

las cuales indican que α_1 y α_2 son vectores propios de S. Tomando los vectores propios de norma unidad y sustituyéndolos en (4), se obtiene que en el máximo la función objetivo es

$$\phi = \lambda_1 + \lambda_2 \quad (9)$$

que corresponden a los dos valores propios mayores de la matriz S. Note que la covarianza entre z_1 y z_2 , dada por $\alpha_1' S \alpha_2 = 0$, ya que $\alpha_1' \alpha_2 = 0$, por tanto las variables z_1 y z_2 no están correlacionadas.

La idea anterior se puede generalizar para obtener los demás componentes principales. Así el espacio de dimensión r que mejor representa a los puntos viene definido por los vectores propios asociados a los r mayores valores propios de S . Estas direcciones se denominan *direcciones principales* de los datos y a las nuevas variables por ellas definidas *componentes principales*. En general la matriz X (y por tanto la matriz S), tiene rango p , existiendo entonces tantas componentes principales como variables que se obtendrán calculando los valores propios $\lambda_1, \lambda_2, \dots, \lambda_p$ de la matriz de covarianzas S mediante

$$|S - \lambda I| = 0, \quad (10)$$

con vectores asociados

$$(S - \lambda_i I)a_i = 0 \quad (11)$$

Llamando Z a la matriz cuyas columnas son los valores de los p componentes en los n individuos, estas nuevas variables están relacionadas con las originales mediante

$$Z = XA, \quad (12)$$

donde $AA' = I$. Por lo que calcular los componentes principales equivale a aplicar una transformación ortogonal A , a las variables X (ejes originales) para obtener unas nuevas variables Z no correlacionadas.

Cuando las variables originales X , están en distintas unidades, es conveniente utilizar la matriz de correlaciones en lugar de la matriz de covarianzas, debido a que la matriz de covarianzas es susceptible a los cambios de escala. El procedimiento es similar al descrito anteriormente, con la ventaja de que la matriz de correlación le da la misma importancia a todas las variables.

2.1.2 Propiedades de los componentes principales

Algunas de las propiedades de los componentes principales, que hacen que la técnica sea utilizada con frecuencia en el análisis de datos, se describen a continuación (Peña, 2002):

- La suma de las varianzas de los componentes es igual a la suma de las varianzas de las variables originales y la varianza generalizada de los componentes es igual a la original. Como $var(z_i) = \lambda_i, i = 1, 2, \dots, p$ y la suma de los valores propios es la traza de la matriz, entonces

$$tr(S) = var(x_1) + \dots + var(x_p) = \lambda_1 + \dots + \lambda_p .$$

- La proporción de la variabilidad explicada por un componente, es el cociente entre su varianza y la suma de los valores propios de la matriz S . Esto es, para un componente z_k , sabemos que $var(z_k) = \lambda_k$, entonces la proporción de varianza total explicada por z_k es $\frac{\lambda_k}{\sum_{i=1}^p \lambda_i}$.
- Las covarianzas entre cada componente principal y las variables originales están dadas por $Cov(z_i; x_1, x_2, x_3 \dots x_p) = \lambda_i \alpha_i = (\lambda_i \alpha_{i1}, \lambda_i \alpha_{i2}, \dots, \lambda_i \alpha_{ip})$, es decir, el producto de las coordenadas del vector propio α_i , con su respectivo valor propio.
- La correlación entre el i -ésimo componente y la j -ésima variable, es:

$$\rho(z_i, x_j) = \frac{Cov(z_i; x_j)}{\sqrt{var(z_i)var(x_j)}} = \frac{\lambda_i \alpha_{ij}}{\sqrt{\lambda_i s_j^2}} = \alpha_{ij} \frac{\sqrt{\lambda_i}}{s_j} \quad (13)$$

2.1.3 Selección del número de componentes principales

Algunos de los métodos sugeridos para seleccionar el número de componentes que mejor representan el espacio de las variables son:

- **Criterio de Catell** (Peña, 2002). Consiste en seleccionar componentes asociados a valores propios mayores que una cota fija, generalmente establecida como la varianza promedio $\frac{\sum \lambda_i}{p}$. En particular, cuando se trabaja con la matriz de correlación esta cota es de 1, así que se escogen los componentes principales cuyos valores propios están por encima de 1. Este fue el criterio utilizado en el análisis de componentes en esta investigación.
- **Criterio Gráfico**. Se realiza un gráfico de los valores propios λ_i , según $i = 1, 2, \dots, p$, del mayor al menor valor propio, llamado gráfico de sedimentación. Comenzar seleccionando componentes hasta que los restantes tengan aproximadamente el mismo valor de λ_i . Gráficamente se busca un “codo”, es decir, un punto a partir del cual los demás valores propios serán aproximadamente iguales. Este criterio se debe utilizar con precaución porque tiene a escoger pocos componentes (Pla, 1986).
- Seleccionar componentes hasta completar la proporción de varianza acumulada que satisfaga un requerimiento a priori, como por ejemplo el 80% o el 90% de la variabilidad total. Esa decisión le corresponde al investigador y dependerá de la importancia que tengan en la interpretación de los datos.

2.2 Regresión Logística

La Regresión Logística es una técnica multivariada que se utiliza para describir la relación existente entre una variable de respuesta binaria Y , y un conjunto de p variables explicativas x_1, x_2, \dots, x_p , que pueden ser cuantitativas o cualitativas. La variable respuesta, generalmente se codifica con los valores 0 y 1, para dos posibles categorías.

Se distinguen principalmente dos objetivos al utilizar la técnica de Regresión Logística:

- Construir un modelo que permita predecir la probabilidad de éxito de un suceso, para valores particulares del conjunto de variables explicativas.
- Determinar el modelo más parsimónico y ajustado, que describa la relación entre la variable respuesta y el conjunto de variables explicativas.

2.2.1 Modelo de Regresión Logística Simple

La variable respuesta binaria (codificada con 0, 1), tiene propiedades que no satisfacen los supuestos de la regresión múltiple, normalidad de los errores y varianza constante (Hosmer & Lemeshow, 1989), como se describe a continuación:

Supongamos que para una variable respuesta y , $P(y = 1|x) = \pi(x)$ y $P(y = 0|x) = 1 - \pi(x)$, para una variable independiente x . De modo que $E(y|x) = \pi(x) + 0(1 - \pi(x)) = \pi(x)$ por lo que $0 \leq E(y|x) = \pi(x) \leq 1$. Esto significa que $E(y|x)$ está acotado entre $[0,1]$, por lo cual una línea de regresión no sería un buen ajuste para este tipo de variable.

Por otro lado, si el modelo fuera lineal $y = E(y|x) + \varepsilon$, donde ε es el error aleatorio, con $\varepsilon \sim N(0, \sigma^2)$. Si $y = 1$, entonces $\varepsilon = 1 - E(y|x) = 1 - \pi(x)$, con probabilidad $\pi(x)$ y si $y = 0$ entonces $\varepsilon = -\pi(x)$, con probabilidad $1 - \pi(x)$, por lo que los errores aleatorios para la variable respuesta binaria se distribuyen con media cero y varianza $\pi(x)(1 - \pi(x))$. Es decir, la distribución condicional de la variable respuesta sigue una distribución binomial con probabilidad dada por la media condicional, $\pi(x)$.

Uno de los modelos que mejor se ajusta a las características descritas antes, es el modelo logístico, cuya distribución de probabilidad está definida por:

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} \quad (14)$$

Su gráfica en forma sigmoidea cumple con las propiedades requeridas para $\pi(x)$.

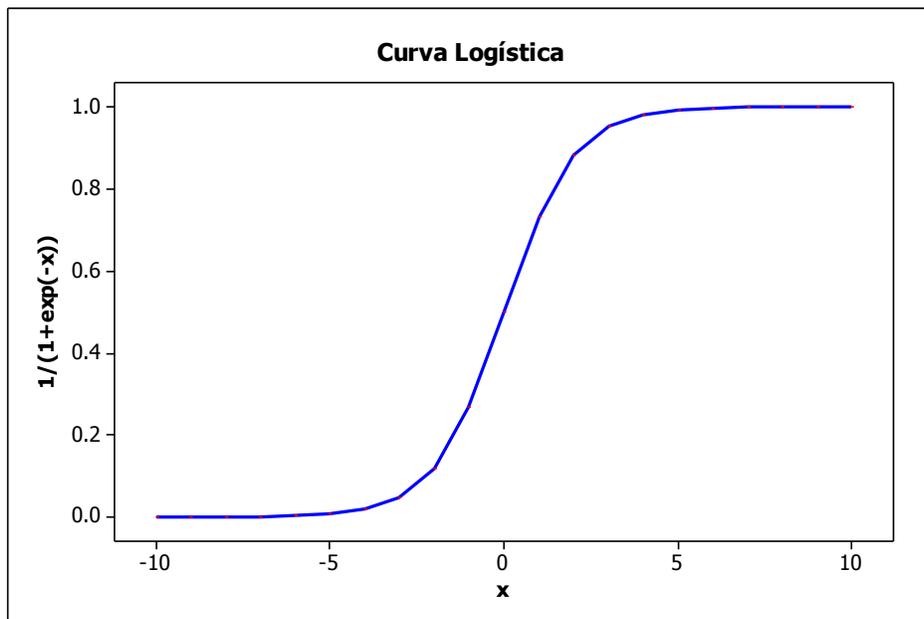


Figura 1.1 Gráfica del Modelo Logístico

Una transformación de $\pi(x)$, de utilidad en la regresión logística, es la transformación “logit”, definida por

$$f(x) = \ln \left[\frac{\pi(x)}{1-\pi(x)} \right] = \beta_0 + \beta_1 x \quad (15)$$

Esta transformación tiene las propiedades de un modelo de regresión lineal, lo cual facilita la interpretación de sus parámetros. La función $f(x)$ es continua, lineal en sus parámetros y con rango de valores $(-\infty, \infty)$.

2.2.2 Estimación de los parámetros del modelo

Se tiene una muestra aleatoria de n observaciones independientes (x_i, y_i) , con $i = 1, 2, \dots, n$.

La función de probabilidades para una respuesta y_i cualquiera estará dada por

$$P(y_i) = \pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i} \quad y_i = 0, 1$$

por lo que la distribución conjunta de las muestras es

$$P(y_1, y_2, \dots, y_n) = \prod_{i=1}^n P(y_i) = \prod_{i=1}^n \pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i} \quad (16)$$

Para estimar los parámetros se hallan los estimadores de máxima verosimilitud maximizando

$$\ln P(y_1, y_2, \dots, y_n) = \sum_{i=1}^n y_i \ln \left[\frac{\pi(x_i)}{1-\pi(x_i)} \right] + \sum_{i=1}^n \ln [1 - \pi(x_i)] \quad (17)$$

que es el logaritmo de la distribución conjunta.

Despejando $\pi(x_i)$ de la ecuación (15) se obtiene que $\pi(x_i) = (1 + e^{-(\beta_0 + \beta_1 x_i)})^{-1}$. Por lo que

$$1 - \pi(x_i) = (1 + e^{\beta_0 + \beta_1 x_i})^{-1} \quad (18)$$

Entonces la ecuación (17), en términos de los parámetros que se desean estimar es

$$L(\beta_0, \beta_1) = \sum_1^n y_i (e^{\beta_0 + \beta_1 x_i}) - \sum_1^n \ln(1 + e^{\beta_0 + \beta_1 x_i}) \quad (19)$$

Para encontrar los valores de β_0, β_1 , que maximiza $L(\beta_0, \beta_1)$ se hallan las derivadas parciales de $L(\beta_0, \beta_1)$ con respecto a los parámetros y se igualan a cero

$$\frac{\partial(L(\beta_0, \beta_1))}{\partial \beta_0} = \sum_1^n \left(y_i - \frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}} \right) = \sum_1^n (y_i - \pi(x_i)) = 0 \quad (20)$$

$$\frac{\partial(L(\beta_0, \beta_1))}{\partial \beta_1} = \sum_1^n x_i \left(y_i - \frac{X_i e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}} \right) = \sum_1^n x_i (y_i - \pi(x_i)) = 0 \quad (21)$$

Obteniendo una expresión no lineal en los parámetros y cuya solución requiere de métodos iterativos, como el caso de mínimos cuadrados ponderados o métodos numéricos como el caso de Newton Raphson (Hosmer & Lemeshow, 1989).

2.2.3 Modelo de Regresión Logística Múltiple

La Regresión Logística Múltiple es una generalización del caso univariado expuesto antes.

Consideremos una colección de p variables independientes, denotada por el vector

$\mathbf{X}' = (x_1, x_2, \dots, x_p)$ con probabilidad condicional $\pi(\mathbf{X}) = P(Y = 1|\mathbf{X})$, entonces el modelo

de regresión logística múltiple se define por

$$\pi(\mathbf{X}) = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p}} = \frac{e^{\beta_0 + \beta' \mathbf{X}}}{1 + e^{\beta_0 + \beta' \mathbf{X}}} \quad (22)$$

en cuyo caso:

$$\ln \left[\frac{\pi(\mathbf{X})}{1-\pi(\mathbf{X})} \right] = \beta_o + \beta' \mathbf{X} \quad (23)$$

donde β_o representa el intercepto del modelo, β' es un vector de dimensión $1 \times p$ que contiene los parámetros del modelo y \mathbf{X} es un vector de dimensión $p \times 1$ que contiene las variables independientes.

En algunas ocasiones se utilizan también como parámetros e^{β_o} y e^{β_i} , que se denominan los *odds ratios* (razones de probabilidades), e indican cuanto se modifican las probabilidades por unidad de cambio en las variables \mathbf{X} . En efecto, de la ecuación (23) se deduce que:

$$O_i = \frac{\pi(\mathbf{X}_i)}{1-\pi(\mathbf{X}_i)} = e^{\beta_o} \prod_{j=1}^p e^{\beta_j x_j} \quad (24)$$

Supongamos dos observaciones i, k , con todos los valores de las variables iguales, excepto la variable h y $x_{ih} = x_{kh} + 1$. El cociente de las razones de probabilidades para estas dos observaciones es $\frac{O_i}{O_k} = e^{\beta_h}$ e indica cuanto se modifica la razón de probabilidad cuando la variable x_h aumenta una unidad. Sustituyendo $\pi(\mathbf{X}_i) = 0.5$ en el modelo logístico, entonces,

$$\ln \left[\frac{\pi(\mathbf{X}_i)}{1-\pi(\mathbf{X}_i)} \right] = \beta_o + \beta_1 x_{i_1} + \dots + \beta_p x_{i_p} = 0 \text{ es decir, } x_{i_1} = -\frac{\beta_o}{\beta_1} - \sum_{j=2}^p \frac{\beta_j x_j}{\beta_1}$$

Note que para el caso univariado el valor de $x_1 = -\frac{\beta_o}{\beta_1}$.

2.2.4 Estimación de los parámetros del modelo

Similar al caso univariado, se estimarán los parámetros del modelo a partir de la función de máxima verosimilitud. Sea $n = n_1 + n_2$ una muestra de observaciones independientes, con n_1 observaciones que pertenecen a la clase 1 y n_2 observaciones que pertenecen a la clase 2;

y sea $\beta' = (\beta_1, \beta_2, \dots, \beta_p)$ el vector de p parámetros que se estimarán. Eligiendo un parámetro binomial, $\pi(\mathbf{X}) = \frac{e^{\beta_0 + \beta' \mathbf{X}}}{1 + e^{\beta_0 + \beta' \mathbf{X}}}$ la función de verosimilitud está dada por:

$$L(\beta_0, \beta) = \prod_{i=1}^{n_1} \frac{e^{\beta_0 + \beta' \mathbf{X}_i}}{1 + e^{\beta_0 + \beta' \mathbf{X}_i}} \prod_{j=n_1+1}^n \frac{1}{1 + e^{\beta_0 + \beta' \mathbf{X}_j}} \quad (25)$$

Los estimados $\hat{\beta}_0$ y $\hat{\beta}$ se obtienen maximizando el logaritmo de $L(\beta_0, \beta)$ y son encontrados aplicando métodos iterativos tales como Newton-Raphson o mínimos cuadrados ponderados iterativos.

2.2.5 Medidas de confiabilidad del modelo

Algunas de las medidas que cuantifican el nivel de ajuste del modelo logístico al conjunto de observaciones son:

- **La Devianza Residual.** Es una medida del grado de diferencia, entre las frecuencias observadas y las estimadas por el modelo, similar a la suma de cuadrados del error, de la regresión lineal. Se define como el negativo de dos veces la función de verosimilitud maximizada, obteniendo de este modo una cantidad cuya distribución es conocida y así tiene validez para usarla en la prueba de hipótesis (Hosmer & Lemeshow, 1989). Para los casos en que la variable Y no está agrupada, se tiene que:

$$D = -2 \left[\sum_{y_i=1}^n \ln(\hat{\pi}_i) + \sum_{y_i=0}^n \ln(1 - \hat{\pi}_i) \right] \quad (26)$$

donde $\hat{\pi}_i$ es el valor estimado en la ecuación (22). El estadístico D se distribuye como una Ji cuadrado con $n-p-1$ grados de libertad, donde p es el número de

variables predictoras. Si $D > \chi_{\alpha, n-p-1}^2$ para un nivel de significancia α dado, entonces el modelo logístico no es confiable.

- **El Pseudo- R^2 .** Es una versión similar al R^2 de la regresión lineal. El propuesto por McFadden (citado por Acuña, 2007) es:

$$\text{Pseudo} - R^2 = 1 - \frac{\text{Devianza Residual}}{\text{Devianza Nula}} \quad (27)$$

donde la devianza nula es la devianza considerando solamente el intercepto y que se distribuye como χ_{n-1}^2 . Un Pseudo- R^2 mayor que 0.3 es considerado como aceptable.

- **La Prueba de Hosmer-Lemeshow.** Evalúa la bondad de ajuste del modelo construyendo una tabla de contingencia, para lo cual divide la muestra en g grupos iguales, a partir de las probabilidades estimadas y compara las frecuencias esperadas con las observadas en cada uno de los grupos. La prueba es una del tipo χ^2 y se define por (Hosmer & Lemeshow, 1989) :

H_0 : El modelo ajustado es adecuado

H_1 : El modelo ajustado no es adecuado

$$\hat{C} = \sum_{k=1}^g \frac{(o_k - n_k \bar{\pi}_k)^2}{n_k \bar{\pi}_k (1 - \bar{\pi}_k)} \quad (28)$$

donde n_k es el número de observaciones en el k -ésimo grupo, $o_k = \sum_{j=1}^{n_k} y_j$ es la suma de las y 's en el k -ésimo grupo y $\bar{\pi}_k = \sum_{j=1}^{n_k} \frac{m_j \hat{\pi}_j}{n_k}$ es el promedio de las probabilidades estimadas $\hat{\pi}_j$. Si $\hat{C} > \chi_{\alpha}^2$ con $g-2$ grados de libertad, entonces se concluye que el modelo logístico no es adecuado.

CAPÍTULO 3

METODOLOGÍA

En este capítulo se presentará, de manera detallada, la forma en que se realizó el estudio; se discutirán los instrumentos utilizados, así como la forma de su aplicación, la descripción de los sujetos de estudio, los procedimientos y técnicas que formaron parte del análisis. Los instrumentos que se describen a continuación permitieron obtener información de los estudiantes de Matemática Prebásica para generar su perfil, analizar las variables cuantitativas asociadas a su rendimiento académico y construir el modelo de predicción de aprobación de los estudiantes de este curso.

3.1 Instrumentos

Se utilizaron tres cuestionarios para medir las variables de interés en el estudio. El primer cuestionario (Anexo 1), diseñado por la investigadora, “*Estudio de algunos factores que afectan el rendimiento académico de los estudiantes de MATE 0066*” indagó sobre características sociodemográficas, trasfondo educativo, hábitos de estudio, gusto hacia las matemáticas, metas y aspiraciones, entre otros.

El cuerpo de este cuestionario se dividió en tres partes: En la primera parte, se recopiló información sobre los aspectos sociodemográficos y académicos de los estudiantes, en la segunda parte, se analizaron los aspectos relacionados con su interés hacia las matemáticas y

el curso de prebásicas y en la tercera, las motivaciones para ingresar a la universidad. La mayoría de los reactivos se presentaron en escala tipo Likert, debido a que las preguntas asociadas eran categóricas.

La validez del contenido para este cuestionario se hizo a través de la revisión de un grupo de expertos (5 profesores del Departamento de Matemáticas, del Recinto Universitario de Mayagüez y 1 estudiante doctoral del área de Currículo y Enseñanza en matemáticas, del Recinto de Río Piedras), quienes hicieron algunas recomendaciones sobre estilo y redacción de preguntas, sobre las escalas de medición, inclusión y/o exclusión de algunos reactivos, adecuación a la población, entre otros. Una vez se hicieron las modificaciones recomendadas por el Comité, el cuestionario fue aplicado a un grupo piloto de 10 estudiantes de Matemática Prebásica, para evaluar si los reactivos tenían un lenguaje claro para ellos, medir el tiempo promedio en que lo podrían responder y tomar nota de las dudas que pudieran surgir durante su aplicación.

También se midió la confiabilidad del instrumento utilizando el coeficiente de consistencia interna, alfa de Cronbach. Como el cuestionario presentaba variedad en sus reactivos y en las escalas de medición de la respuesta, no se podía establecer un único coeficiente. Al respecto, se determinaron coeficientes para dos subescalas: la del interés hacia las matemáticas y el curso de prebásicas, con un coeficiente de 0.99 y la otra sobre las motivaciones para ingresar a la universidad, obteniéndose un coeficiente de 0.71, los cuales se consideran confiables.

El segundo cuestionario: “*Estudio sobre la Ansiedad hacia las matemáticas*” fue diseñado por los profesores Muñoz y Mato Vásquez (2007), de la Universidad de La Coruña, y adaptado y modificado por los profesores Wayland y Alvarado (2008). En su versión original, este cuestionario tenía 24 reactivos, repartidos en cinco factores: “ansiedad ante la evaluación de matemáticas”, “ansiedad ante la temporalidad”, “ansiedad ante la comprensión de problemas”, “ansiedad frente a los números y operaciones matemáticas” y “ansiedad ante situaciones matemáticas de la vida real” y su confiabilidad, medida a partir del coeficiente de consistencia interna alfa de Cronbach, era de 0.95.

Luego de la adaptación por Alvarado (2008), se añadieron al cuestionario 10 reactivos más y la confiabilidad del mismo, de acuerdo a los resultados obtenidos en su investigación, fue de 0.97. Esta versión del cuestionario, que había sido implementada con estudiantes de undécimo grado, en una escuela de Mayagüez, fue la utilizada para medir el nivel de ansiedad de los estudiantes de Matemática Prebásica.

El tercer cuestionario: “*Estudio sobre la Actitud hacia las matemáticas*” fue adaptado por Wayland y Alvarado (Alvarado, 2008) del cuestionario original MINDSET, de la Dra. Carol Dweck sobre tipos de mentalidad. Al igual que el cuestionario anterior, éste también había sido implementado con estudiantes de undécimo grado, de una escuela en Mayagüez. La versión utilizada consta de 6 reactivos de análisis cuantitativo y 1 reactivo de análisis cualitativo.

Previo a la aplicación de estos cuestionarios se obtuvo permiso de los autores, así como la autorización del Comité para la Protección de Seres Humanos en la Investigación [CPSHI] (Anexo 2). También fueron entregados los consentimientos informados a los estudiantes de Matemática Prebásica participantes en la investigación (Anexo 3).

3.2 Población y Muestra

De acuerdo a la matrícula de estudiantes en el curso de Matemática Prebásica para el I semestre 2010-2011 hubo 35 secciones, divididas en 5 megasecciones y 30 secciones regulares. La capacidad de las megasecciones es de 70 estudiantes y las secciones regulares 29 estudiantes, por sección. Sin embargo, de acuerdo a la matrícula, la población objetivo fue de 1094 estudiantes. Durante la administración de los cuestionarios se encontró que las secciones eran más pequeñas, algunas porque los estudiantes no asistían a clase y otras, porque los estudiantes se presentan solamente a tomar los exámenes.

Para seleccionar la muestra se utilizó muestreo por conglomerados combinado con muestreo estratificado, debido a que las secciones de Matemática Prebásica, como se indicaba antes, estaban divididas en megasecciones y secciones regulares. En general, los métodos de muestreo pueden combinarse dependiendo de las necesidades del estudio. Para ello se dividió la población en dos estratos, en el primero iban a estar las 5 megasecciones y en el segundo, las 30 secciones regulares. Posteriormente se seleccionaron de cada uno de los estratos los conglomerados (secciones), en forma aleatoria y de acuerdo al tamaño de cada estrato; del

primer estrato se seleccionaron 2 conglomerados y del segundo, 15 conglomerados. Éstos representan el 40% y 50% del tamaño total en cada estrato.

La utilización del muestreo por conglomerados se justifica en que los conglomerados (secciones) están establecidos de forma natural, al inicio del semestre, y por esta razón se esperaría una alta variabilidad dentro de cada sección y homogeneidad entre secciones. Además, al escoger secciones de Matemática Prebásica y dentro de ellas encuestar a todos los estudiantes, se estaría afectando lo menos posible el tiempo de los estudiantes y obteniendo la mayor información posible.

Los cuestionarios fueron aplicados a la muestra seleccionada, en la semana del 25 al 29 de octubre de 2010, tomando una hora de clase de los estudiantes durante esa semana. A los estudiantes se les explicó el objetivo del estudio, así como la importancia de que la información fuera verdadera, con el fin de aumentar la validez y confiabilidad de los resultados. Se les entregó la carta de consentimiento informado, la cual firmaron los estudiantes que accedieron participar de la investigación. En general, todos los estudiantes que estaban en clase, el día de la administración del cuestionario, participaron. Los tres cuestionarios fueron aplicados en el mismo momento, para evitar que los estudiantes respondieran sólo uno o dos de ellos y se perdiera información. A cada estudiante se le solicitó escribir en los cuestionarios los últimos 4 dígitos del número de identificación, para garantizar que al transcribir la información no se presentaran errores, como por ejemplo cruzar información de dos estudiantes.

3.3 Procedimiento

El análisis de la información se realizó en varias etapas: La primera consistió en la caracterización del perfil de los estudiantes a partir de la utilización de estadísticas descriptivas (de tendencia central y de variabilidad), análisis de tabulación cruzada, representación en gráficas (distribuciones de frecuencias, diagramas circulares, entre otros). En dicha caracterización se analizaron los siguientes aspectos: características sociodemográficas, trasfondo académico, aspiraciones académicas, gusto hacia las matemáticas, ansiedad y actitud hacia las matemáticas, así como motivaciones y planes al graduarse. También se hicieron algunas pruebas no paramétricas para analizar si había o no diferencias significativas entre los grupos de estudiantes, de acuerdo a ciertas variables. Es importante indicar que la variabilidad de las principales variables cuantitativas era similar entre Estrato 1 y Estrato 2, por lo cual se combinaron en el análisis.

En la segunda etapa se analizaron las variables cuantitativas: promedio de escuela superior, puntuación en aptitud matemática del College Board, puntuación en aprovechamiento del College Board, puntuación en escala de ansiedad y promedio de quices electrónicos, a partir de la técnica multivariada de Componentes Principales, lo que permitió la construcción de tres nuevas variables, las cuales son combinaciones lineales de las mencionadas antes y tienen la propiedad de no estar correlacionadas, lo cual es deseable para el análisis posterior.

En la tercera etapa se aplicó la técnica de Regresión Logística para predecir la aprobación o no del curso de Matemática Prebásica, para ello se definió la variable dependiente dicotómica

(1 aprobó, 0 no aprobó) y se utilizaron como variables explicativas los tres componentes principales y algunas variables cualitativas, resultado del estudio. Para la selección del mejor subconjunto de variables se utilizó el método de selección hacia atrás, determinando el modelo predictivo. Una vez determinado el modelo final de Regresión Logística, se analizaron características como: significancia del modelo, bondad del ajuste del modelo, tasa de buena clasificación y análisis de los correspondientes diagnósticos de regresión.

Finalmente, se hizo un análisis retrospectivo, sobre los resultados, obtenidos para contrastarlos con las posibles recomendaciones, a las autoridades del Departamento de Ciencias Matemáticas, en función de posibles cambios a las iniciativas tradicionales que ha tomado el Departamento para atender la problemática aquí analizada.

CAPÍTULO 4

RESULTADOS

4.1 Descripción del perfil de los estudiantes

4.1.1 Características Sociodemográficas

En total, 404 estudiantes respondieron los cuestionarios, lo que representa al 37% de la población de estudiantes en las secciones de Matemática Prebásica. De estos, el 51.7% son mujeres y el 48.3% son hombres. El 76.7% del total tienen edades entre 18 y 19 años. De acuerdo al programa de estudios, el 49.9% pertenece a la Facultad de Artes y Ciencias (A y C), el 21.5% pertenece a la Facultad de Ciencias Agrícolas (C. Agr.), el 15.2% pertenece a la Facultad de Administración de Empresas (Adm.) y el 13.4% pertenece a la Facultad de Ingeniería (Ing). Los departamentos de Biología, Ciencias Sociales y Matemáticas, de la Facultad de Artes y Ciencias tienen el mayor número de estudiantes matriculados en Matemática Prebásica, para un 33.2%, del total de estudiantes.

En la Figura 4.1 se muestra la distribución de estudiantes según género, en cada Facultad. Se observa que las mujeres predominaron en la Facultad de Artes y Ciencias, mientras que los hombres predominaron en la Facultad de Ingeniería.

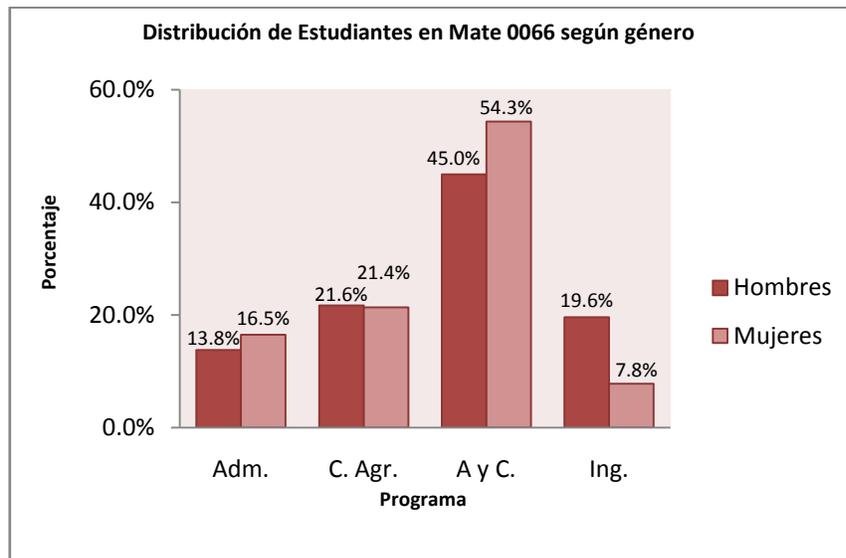


Figura 4.1

El 78.4% de los estudiantes son de nuevo ingreso, lo que implica que están tomando el curso de Matemática Prebásica por primera vez. Este alto porcentaje de estudiantes de nuevo ingreso le da pertinencia a los instrumentos utilizados para medir ansiedad y actitud hacia las matemáticas, puesto que habían sido validados con estudiantes de los últimos grados de escuela superior.

La Figura 4.2 muestra que el porcentaje de estudiantes de nuevo ingreso que están tomando el curso, es mayor al porcentaje de los estudiantes que llevan más de un semestre en los programas del RUM, en las cuatro facultades; predominando la Facultad de Artes y Ciencias, con un 39.1%. El porcentaje más bajo de estudiantes que están viendo el curso por segunda o tercera vez, se encontró en la Facultad de Ingeniería, lo que es razonable porque a sus programas ingresan estudiantes con mayor rendimiento académico en los cursos de matemáticas.

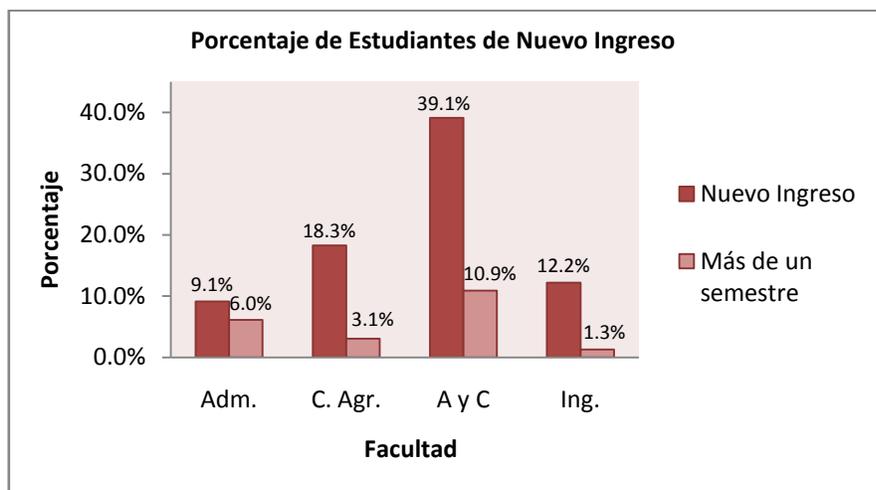


Figura 4.2

El 62% de los estudiantes provienen de escuela pública y el 38% de escuela privada. En la Figura 4.3 se observa que el porcentaje de estudiantes que provienen de escuela pública es mayor que el porcentaje de estudiantes que provienen de la escuela privada, en todas las facultades.

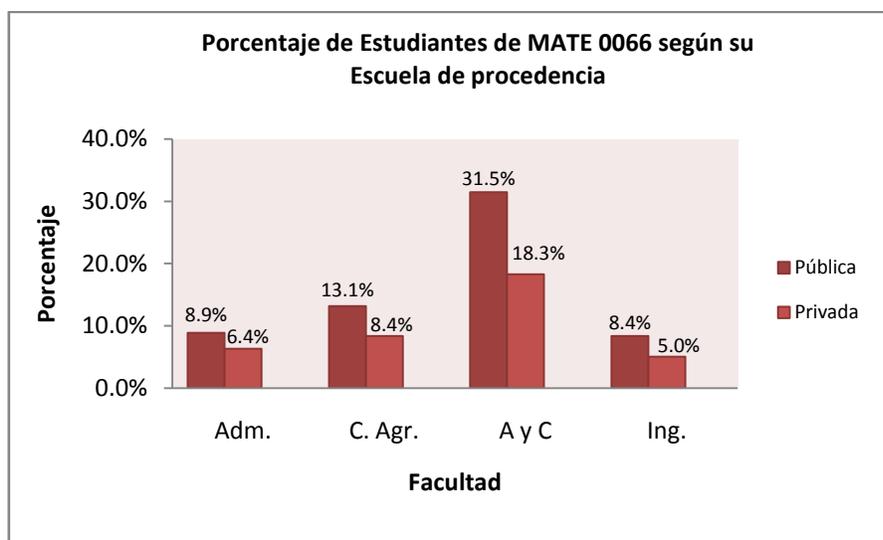


Figura 4.3

El 83% de los estudiantes de Matemática Prebásica recibe algún tipo de ayuda económica o beca, siendo predominante la beca PELL, con el 81%, como se observa en la Figura 4.4. Se encontró que del total de estudiantes en la muestra el 13.9% tiene un trabajo parcial. El 10.1% de los estudiantes recibe beca PELL y tiene un trabajo parcial. Por otra parte, el 17% de los estudiantes no recibe ningún tipo de ayuda económica en el Colegio, ni tampoco trabaja, lo que implicaría que corresponde a una minoría que depende económicamente de su familia.

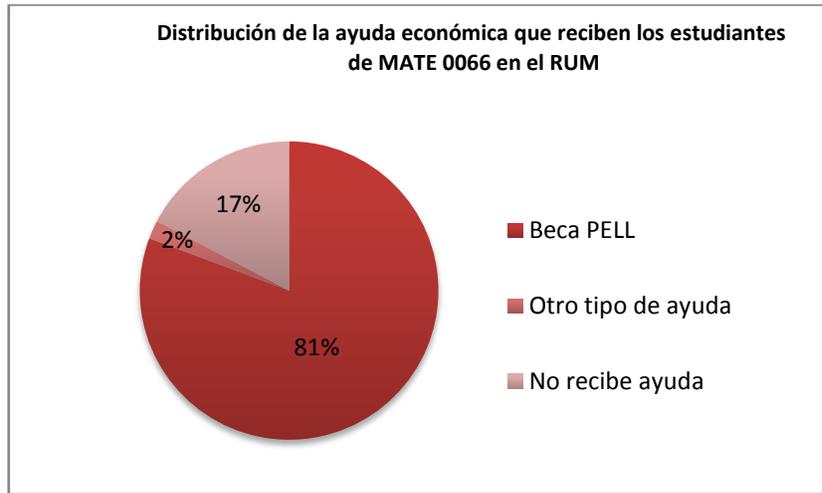


Figura 4.4

4.1.2 Trasfondo Académico

En cuanto al trasfondo académico de los estudiantes se analizaron aspectos relacionados con la preparación académica de los padres, horas de estudio dedicadas a matemáticas en la escuela superior, promedio de graduación, resultados del College Board en las pruebas de aprovechamiento y aptitud matemática, su percepción sobre que tan preparados estaban en matemáticas, si utilizaban o no textos de matemáticas en la escuela superior, entre otros. La distribución de la preparación académica de los padres se presenta en la Tabla 4.1:

Tabla 4.1 Preparación Académica de los Padres

Preparación Académica	Padre	Madre
Escuela Superior	33.9%	20.9%
Grado Asociado	21.6%	20.4%
Bachillerato	34.5%	45.5%
Maestría	7.00%	11.5%
Doctorado	3.00%	1.70%

En la Figura 4.5 se observa que el 58.6% de las madres de los estudiantes tienen como mínimo estudios de bachillerato, en contraste con el 44.4% de los padres; además el porcentaje de padres con sólo escuela superior es mayor que el de las madres.

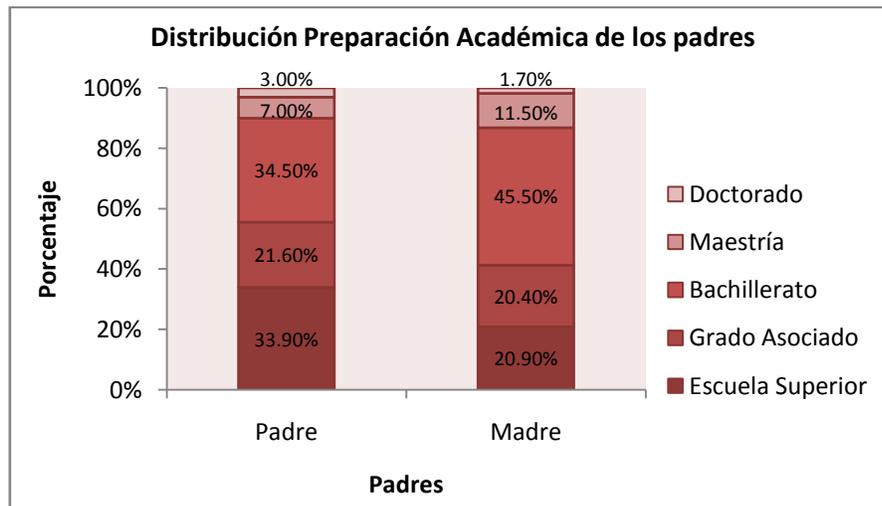


Figura 4.5

El 30.8% de los estudiantes opinan que su preparación para los cursos de matemáticas que tomarán en la universidad, es deficiente o regular, mientras que el 69.2% de los estudiantes afirman tener un nivel aceptable, bueno o excelente, como se muestra en la Tabla 4.2. Este resultado contrasta con el hecho de que los estudiantes ingresen al curso de Matemática Prebásica, ya que los estudiantes tienen una percepción de que están bien preparados, sin embargo, los resultados en MATE 0066 indican lo contrario.

Tabla 4.2 Preparación en los cursos de matemáticas

Preparación en los cursos de Matemáticas E.S	Por ciento
Deficiente	4.20%
Regular	26.6%
Aceptable	30.2%
Bueno	34.3%
Excelente	4.70%

Los resultados anteriores se pueden contrastar con el tiempo dedicado por los estudiantes (horas de estudio a la semana) para prepararse en sus cursos de matemáticas en la escuela superior; al respecto, el 50.7% dedicaba menos de una hora semanal y el 91.3% dedicaba 3 horas semanales o menos a su preparación, como se muestra en la Figura 4.6, lo cual es un indicador de los pobres hábitos de estudio que tienen los estudiantes participantes.

La misma pregunta, para las horas de estudio dedicadas al curso de Matemática Prebásica, proporciona resultados parecidos: el 50.3% de los estudiantes dedican menos de una hora a la semana para estudiar el material del curso y el 95.4% dedicaban 3 horas semanales o menos, como se muestra en la Figura 4.7, reflejando de este modo que no hay mejoría en los hábitos de estudio para los estudiantes de Matemática Prebásica.

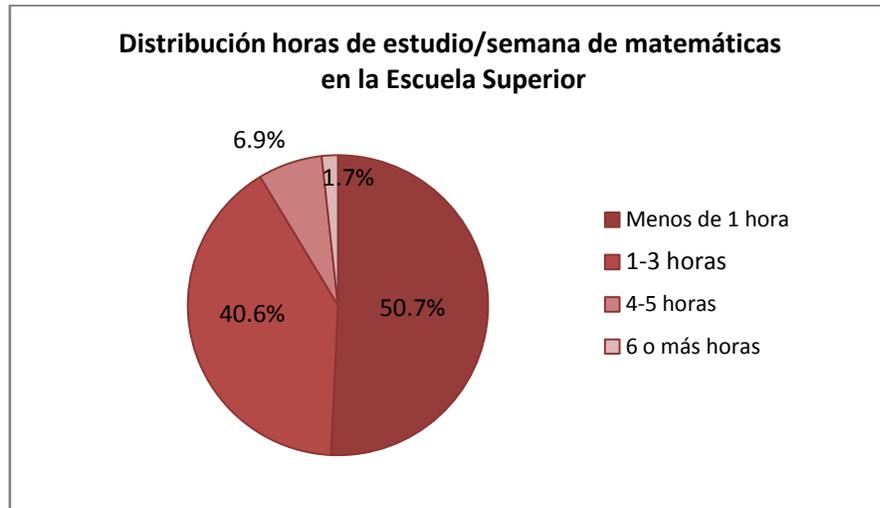


Figura 4.6

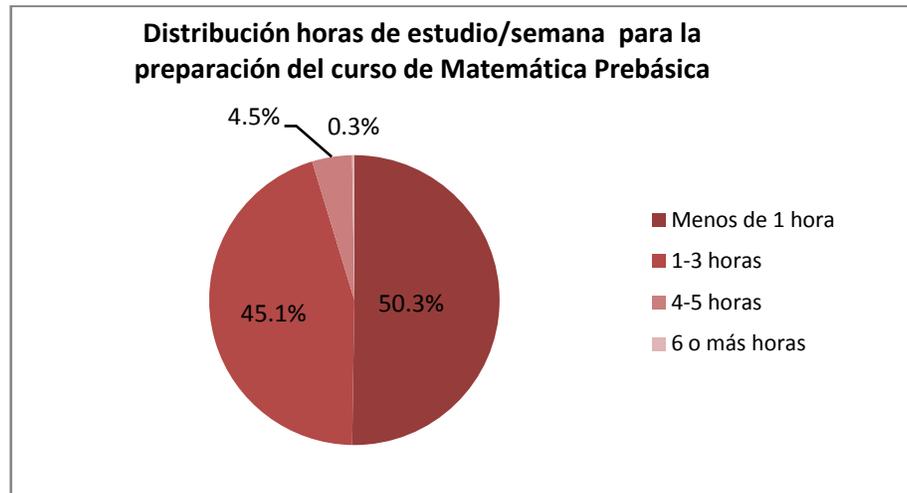


Figura 4.7

El 61% de los estudiantes manifestó tener a su disposición libro de matemáticas cuando estaba en la escuela superior; sin embargo, el 17.5% de ellos nunca utilizaba el libro y el 57.7% lo utilizaba una vez por semana.

Los resultados anteriores evidencian que los pobres hábitos de estudio son un factor predominante entre los estudiantes del curso de Matemática Prebásica y podrían derivarse, del tipo de mentalidad de los estudiantes; así por ejemplo, según las investigaciones de

Dweck (2006) sobre tipos de mentalidad, si un estudiante tiene mentalidad fija, él tendrá la creencia de que sus habilidades matemáticas no se pueden modificar, se es “bueno” en matemáticas, o se es “malo” en matemáticas, por lo que tienden a trabajar menos para tener a quién o qué (externo a él) culpabilizar de su fracaso. En cuanto a la opinión que tienen sobre la preparación que recibieron de sus maestros en la escuela superior, el 52.9% opina que sus maestros los prepararon adecuadamente.

En la Tabla 4.3 se presentan otras de las variables que ayudan a definir el perfil de los estudiantes, desde lo académico, como son el promedio de escuela superior y los resultados en la prueba de aptitud y aprovechamiento matemático, en el College Board. La tasa de respuesta para estas dos últimas variables fue 65.6%, a diferencia del promedio de escuela superior que fue respondido por el 96% de los estudiantes. Estos resultados fueron obtenidos estimando el promedio por estudiante, a partir del estimador propuesto por Scheaffer, Mendenhall y Ott (1987), en la forma de un estimador de razón combinado.

Tabla 4.3 Variables académicas de los estudiantes de Matemática Prebásica

Variable	Puntuación (Promedio)	Varianza	Coefficiente de Variación
Promedio Escuela Superior	3.67	0.12	9.00%
Aptitud Matemática	601.21	4065.78	10.6%
Aprovechamiento en Mat.	589.55	3919.12	10.6%

En general, un promedio de 3.67 en la escala de 0 a 4, corresponde a una valoración de A, lo que es cuestionable con la realidad de los estudiantes que ingresan a los cursos de Matemática Prebásica y con la percepción que tienen los estudiantes de cuán preparados se sienten en los cursos de matemáticas, así como los deficientes hábitos de estudio que se exponían antes.

Se analizó la homogeneidad de la varianza por estratos, utilizando la Prueba de Levene, ya que las variables en cuestión no se distribuyen en forma normal (Anexo 4). Se encontró con un 95% de confiabilidad que no existen diferencias significativas entre las varianzas de las variables promedio de escuela superior, la aptitud del College Board y el aprovechamiento en el College Board, en las megasecciones y las secciones regulares; por esta razón, no había necesidad de la estratificación. Los resultados de la Prueba de Levene para las variables se presentan en la Tabla 4. 4:

Tabla 4.4 Resultados Prueba de Levene

Variable	Prueba de Levene	
Promedio	Test Estadístico	0.220
Esc. Superior	P_valor	0.137
Aptitud	Test Estadístico	0.229
College Board	P_valor	0.132
Aprovechamiento	Test Estadístico	0.480
College Board	P_valor	0.487

También se analizó el promedio de las variables anteriores, según género y escuela de procedencia, como se indica en las Tablas 4.5 y 4.6, respectivamente. Para establecer si existía alguna diferencia significativa entre ellas, se utilizó la prueba no paramétrica de Mann-Withney, que se usa cuando se quieren comparar dos poblaciones, usando muestras independientes. La hipótesis nula es que las medianas de las dos poblaciones son iguales y la hipótesis alterna puede ser que la mediana de la población 1 sea mayor o menor que la mediana de la población 2. La prueba de hipótesis con un nivel de confianza de 95% indicó que existen diferencias significativas entre las medianas del puntaje de escuela superior,

según género y según escuela de procedencia, con valores p de 0.029 y 0.001, respectivamente.

Tabla 4.5 Promedio Escuela Superior, aptitud y aprovechamiento según género

Variable	Género	Puntuación (Promedio)	Mediana	Coefficiente Variación
Promedio Escuela Superior	Hombres	3.65	3.78	8.70%
	Mujeres	3.76	3.84	6.90%
Aptitud Matemática	Hombres	568.5	584	16.9%
	Mujeres	569.1	573	16.1%
Aprovechamiento - Matemática	Hombres	556.2	583	15.7%
	Mujeres	560.9	580	14.1%

Tabla 4.6 Promedio Escuela Superior, aptitud y aprovechamiento según Escuela de procedencia

Variable	Tipo de Escuela	Puntuación (Promedio)	Mediana	Coefficiente Variación
Promedio Escuela Superior	Pública	3.74	3.83	7.30%
	Privada	3.65	3.78	8.80%
Aptitud Matemática	Pública	567.0	567	15.3%
	Privada	570.8	589	18.3%
Aprovechamiento - Matemática	Pública	562.6	580	13.6%
	Privada	551.2	585	16.9%

De acuerdo a la escala del College Board la puntuación media es de 500, por lo que los estudiantes están por encima de la media. Además una puntuación de 550 a 640 es alta, por lo que el promedio de la puntuación de los estudiantes, en las dos pruebas, se encuentra en esa categoría.

4.1.3 Aspiraciones académicas:

En la Figura 4.8 se muestra el porcentaje de los estudiantes, en las secciones de Matemática Prebásica, según sus aspiraciones académicas, por género. Se observa que el 82.1% de las mujeres aspira llegar a estudios de maestría y doctorado, porcentaje que disminuye en el caso de los hombres a 75.8%. Es decir, que hay una diferencia de 6.3% entre las aspiraciones académicas de las mujeres y los hombres para los niveles de maestría y doctorado. La prueba

de independencia para estas dos variables categóricas, a un nivel de confianza del 95% indica que hay asociación entre las aspiraciones académicas de los estudiantes y el género.

Pearson Chi-Square = 8.383, DF = 2, P-Value = 0.015
Likelihood Ratio Chi-Square = 8.427, DF = 2, P-Value = 0.015

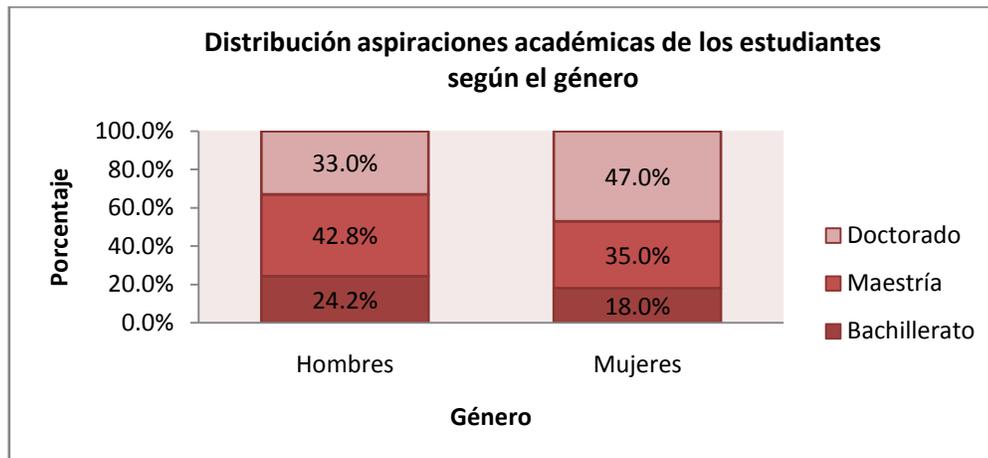


Figura 4.8

Lo anterior refleja que, en general, los estudiantes tienen una percepción de que están bien preparados, que sus maestros los prepararon bien y tienen altas aspiraciones académicas (maestría o doctorado), pero no conciben el esfuerzo y la dedicación que ello implica, haciendo referencia nuevamente, a los resultados expuestos antes sobre hábitos de estudio con relación al curso de Matemática Prebásica.

En la Figura 4.9, se muestra la distribución de las aspiraciones académicas de los estudiantes según la preparación académica de sus padres; para ello se escogió el nivel máximo de escolaridad entre padre y madre. De los estudiantes que tienen aspiraciones de maestría, el 48.7% de sus padres tienen estudios de bachillerato, el 12.3% estudios de maestría y el 5.2% estudios de doctorado y de los estudiantes que tienen aspiraciones de doctorado, el 53.4% de sus padres tienen estudios de bachillerato, el 15.5% de maestría y el 5% estudios de doctorado. En general, se observa que a mayor nivel académico de los padres, mayores

aspiraciones académicas de los estudiantes. Sin embargo estas diferencias no son estadísticamente significativas, como lo indica la prueba de independencia correspondiente

Pearson Chi-Square = 8.530, DF = 8, P-Value = 0.383
Likelihood Ratio Chi-Square = 8.671, DF = 8, P-Value = 0.371

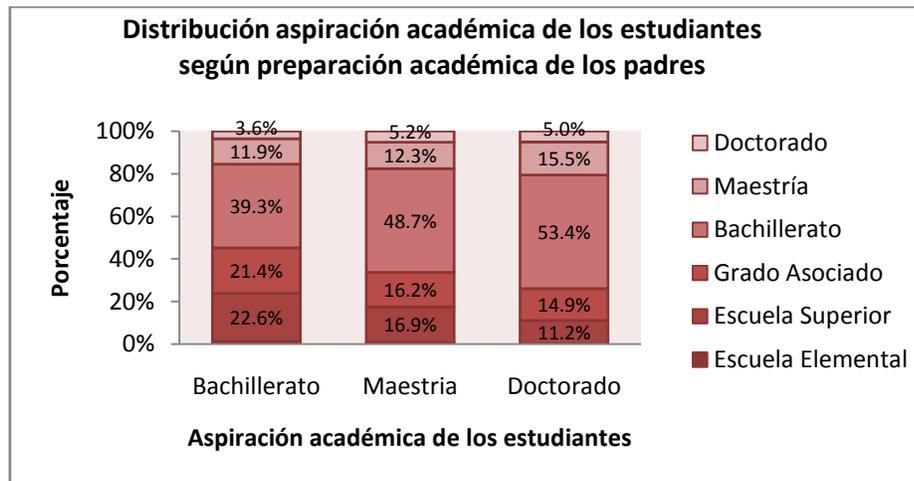


Figura 4.9

4.1.4 Gusto hacia las matemáticas

Para caracterizar esta variable se tuvieron en cuenta aspectos como la importancia dada al curso de Matemática Prebásica, su percepción sobre la dificultad del curso, el esfuerzo para aprobar el curso, el tiempo dedicado en la preparación para los exámenes, su percepción sobre los quices electrónicos y su motivación en el curso, entre otros.

El 51.3% de los estudiantes afirma estar en desacuerdo o es indiferente ante la pregunta de “*si le gustan las matemáticas*”. Similarmente, el 40% de los estudiantes afirma *no interesarse por el curso de Matemática Prebásica*, lo que preocupa porque ese bajo nivel de interés por parte de algunos estudiantes se ve reflejado en su actitud y en los resultados académicos.

Por otra parte, el 95% considera que *aprobar el curso es importante y le será útil para sus cursos posteriores de matemáticas* y el 75% de los estudiantes consideran que *deben esforzarse para aprobar el curso*; sin embargo, cuando se comparan éstos resultados con las *horas dedicadas a estudiar el material de clase* se encuentran serias contradicciones, como se observa en la Figura 4.10, ya que de los estudiantes que afirman prepararse al momento de presentar los exámenes, el 47.6% afirmó dedicar menos de una hora a la semana para su estudio y otro 47%, entre una y tres horas semanales. Nuevamente se está evidenciando, a partir del análisis descriptivo que el tiempo dedicado a estudiar el curso es una variable que afecta el rendimiento académico de los estudiantes de Matemática Prebásica.



Figura 4.10

En resumen, para el 95% de los estudiantes de Matemática Prebásica aprobar el curso es importante y el 75% reconoce que debe esforzarse, pero por otro lado, al 40% de los estudiantes no le “interesa” el curso, lo cual se ve reflejado en que cerca del 95% estudia 3 o menos horas, a la semana, como se muestra en la Figura 4.10.

4.1.5 Ansiedad hacia las matemáticas

En la escala de ansiedad hacia las matemáticas, se encontró que el promedio de ansiedad general de los estudiantes de Matemática Prebásica es de 37.6, con un coeficiente de variación de 57.8%. Esto significa que en promedio, de acuerdo a la medición de esta escala, los estudiantes de Matemática Prebásica tienen un nivel medio de ansiedad. En la Tabla 4.7 se presenta la puntuación promedio y el coeficiente de variación de cada uno de los factores asociados con esta escala, cada factor corresponde a una dimensión de la ansiedad, propuesta por Muñoz y Mato (2007):

Tabla 4.7 Factores Escala de Ansiedad

Variable	Puntuación Promedio	Coefic. Variación
Ansiedad ante los exámenes	44.9	52.1%
Ansiedad ante la eventualidad	37.8	61.5%
Ansiedad ante la comprensión de problemas	34.5	74.2%
Ansiedad frente a los números y operaciones	35.7	78.5%
Ansiedad ante situaciones matemáticas de la vida real	19.1	93.1%
Ansiedad General	37.6	57.8%

En las Figuras 4.11 y 4.12, respectivamente, se presenta la distribución de los estudiantes de Matemática Prebásica, según el nivel de ansiedad, frente a los exámenes y frente a la eventualidad de los mismos. El 57.2% de los estudiantes tienen un nivel “medio” o “alto” de ansiedad hacia los exámenes; y el 47% tienen un nivel “medio” o “alto” hacia la eventualidad. Esto significa, de acuerdo a la interpretación de los factores, que los estudiantes de Matemática Prebásica experimentan mayor nivel de ansiedad cuando van a ser evaluados ó, cuando la fecha del examen es próxima.

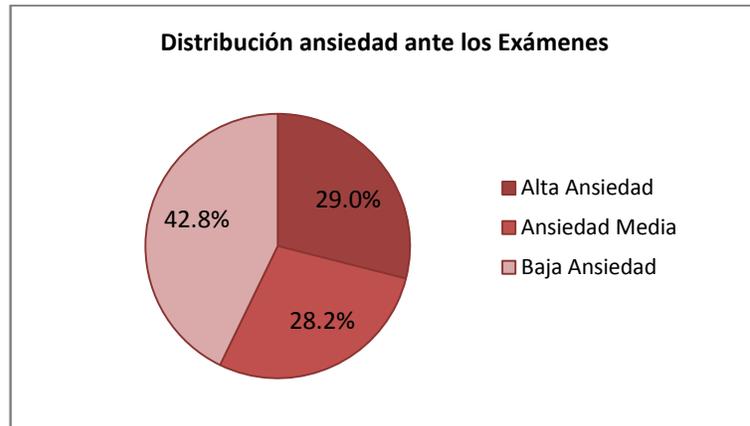


Figura 4.11

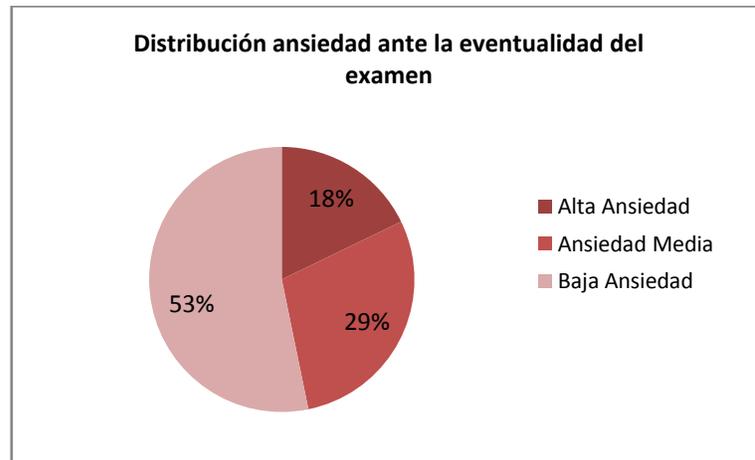


Figura 4.12

El factor donde los estudiantes manifestaron menos ansiedad, fue el de ansiedad ante situaciones matemáticas de la vida real, con un 88%, como se muestra en la Figura 4.13. Sin embargo, el coeficiente de variación de este factor es el más alto, con un 93.1%. Uno de los reactivos que constituye este factor es “*me pone nervioso resolver un problema de álgebra*”, y se encontró que el 45.4% de los estudiantes manifestaron ansiedad con respecto a él.

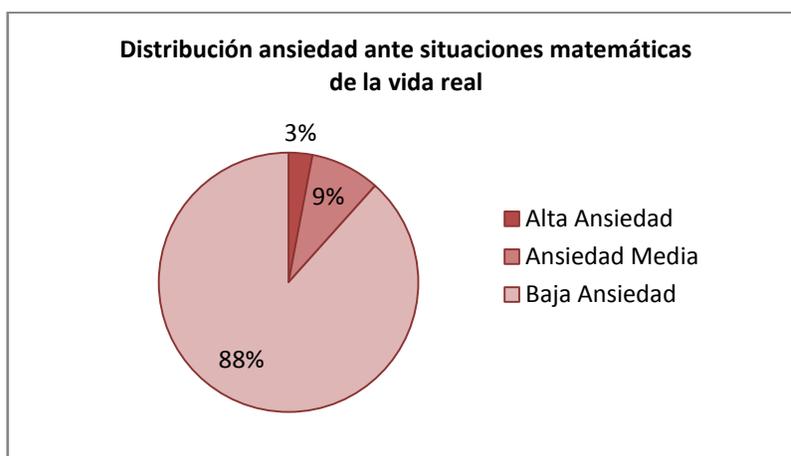


Figura 4.13

También se analizó la ansiedad matemática de los estudiantes, según el género y la escuela de procedencia. Este contraste de hipótesis, al igual que los anteriores, se hizo aplicando la prueba de Mann-Withney. Los resultados de las pruebas indicaron que con un nivel de confianza del 95%, existen diferencias de género para la mediana del nivel de ansiedad de los estudiantes (valor $p < 0.05$) y con respecto a la escuela de procedencia no se encontraron diferencias significativas (valor p de 0.102). En la Tabla 4.8 se presentan las estadísticas descriptivas de la ansiedad según el género de los estudiantes. Se observa que hay una diferencia de 11.8 puntos entre la mediana de la ansiedad de los hombres y las mujeres, siendo éstas la de mayor ansiedad.

Tabla 4.8 Estadísticas Descriptivas ansiedad según género de los estudiantes

Género	Puntuación (Promedio)	Puntuación (Mediana)	Coefficiente Variación	Puntuación Mínima	Puntuación Máxima
Hombres	33.4	30.1	63.3	0.0	80.2
Mujeres	41.4	41.9	52.2	1.5	94.9

Se concluye que existen diferencias significativas en el nivel de ansiedad general, de hombres y mujeres de Matemática Prebásica. Además, al separarlos por grupos de acuerdo a la Facultad a la que pertenece se observa que el nivel de ansiedad de las mujeres es más alto que el de los hombres, excepto en la Facultad de Artes y Ciencias, como se muestra en la Figura 4.14. La mayor diferencia se encuentra en la Facultad de Ingeniería, donde las mujeres tienen un porcentaje de ansiedad mayor que el de los hombres, con una diferencia de más del 30%.

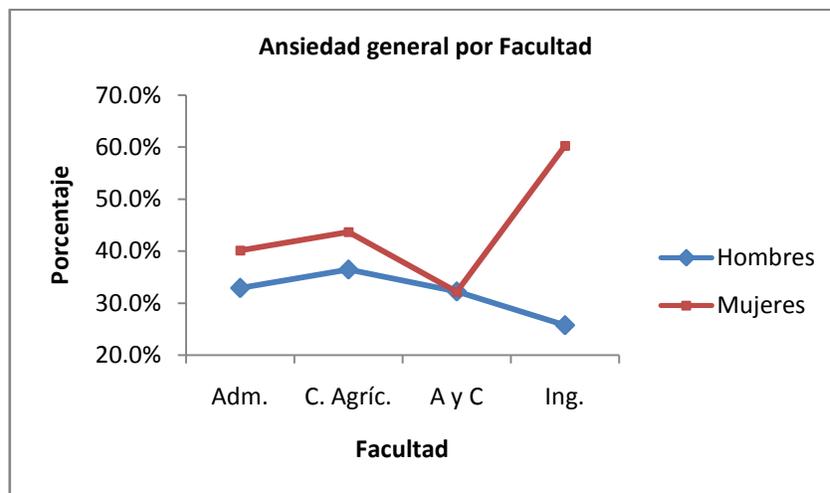


Figura 4.14

Como se mencionó antes, los estudiantes de Matemática Prebásica experimentan un nivel de ansiedad alto cuando van a ser evaluados o cuando la fecha del examen se aproxima, esto puede estar relacionado con los pobres hábitos de estudio y también con las deficiencias de conocimiento básico, ya que al no tener una buena organización del tiempo de estudio, terminan mecanizando algoritmos o procedimientos, sin desarrollar las destrezas matemáticas requeridas en el curso.

4.1.6 Actitud hacia las matemáticas

A partir de la Escala de Actitud hacia las matemáticas se estableció la distribución de los estudiantes, según el tipo de mentalidad fija o de crecimiento que poseen, de acuerdo a la teoría de la mentalidad de Dweck. Al respecto, se encontró que el 29.9% de los estudiantes tienen una mentalidad fija, el 47.7% tienen una mentalidad de crecimiento y el 22.4% restante, son estudiantes que muestran en sus respuestas tendencia hacia los dos tipos de mentalidad.

En sus investigaciones, Dweck (2006) ha encontrado que en general, hay un 40% de personas que tienen mentalidad fija, 40% de personas con mentalidad de crecimiento y el 20% restante los define como personas que tienen una combinación de ambos tipos de mentalidad. En la Figura 4.15 y Figura 4.16, se presentan la distribución del tipo de mentalidad de los estudiantes, según el género y según la facultad a la que pertenecen, respectivamente. Se observa que el 45.2% de los hombres y el 50% de las mujeres tienen mentalidad de crecimiento.

Al hacer una prueba de homogeneidad, a un nivel de confianza del 95%, se encontró que no hay asociación entre el tipo de mentalidad y el género. En el caso de la facultad, se encontró que el 43.9% de administración, el 40% de ciencias agrícolas, el 48% de artes y ciencias y el 50.9% de ingeniería, tienen mentalidad de crecimiento. Tampoco hay asociación entre el tipo de mentalidad y la facultad a la que ingresaron los estudiantes de Matemática Prebásica.

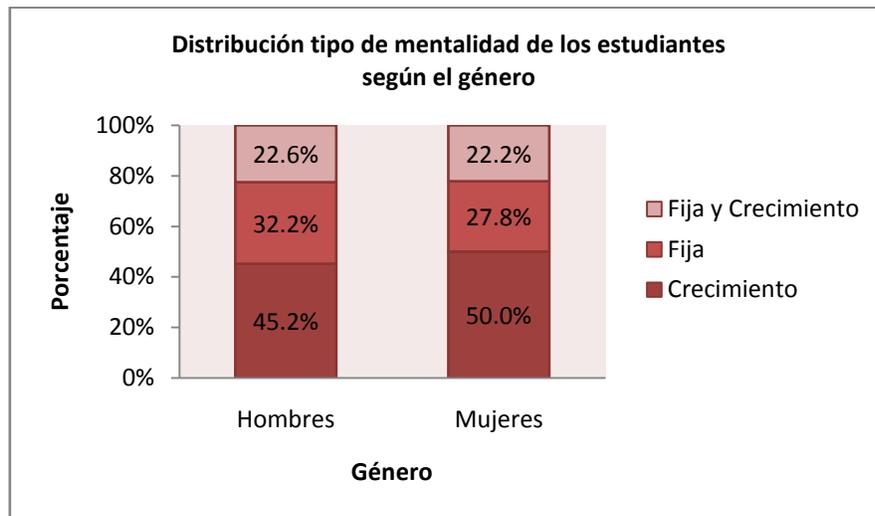


Figura 4.15

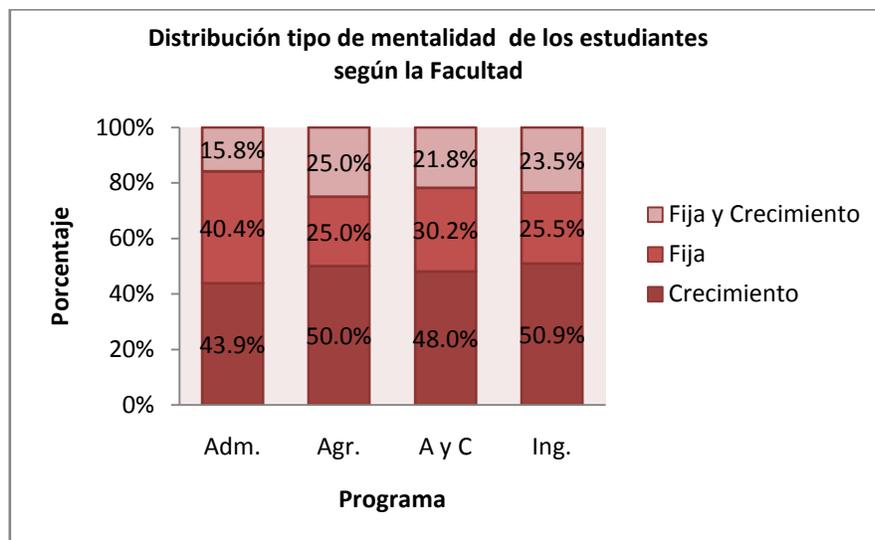


Figura 4.16

En la Figura 4.17 se presenta la distribución de las horas de estudio dedicadas a Matemática Prebásica según el tipo de mentalidad y se observa que el porcentaje más alto de mentalidad de crecimiento coincide con los estudiantes que estudian 4 o más horas. En general, el porcentaje de estudiantes con mentalidad de crecimiento va aumentando con respecto a las horas de estudio.

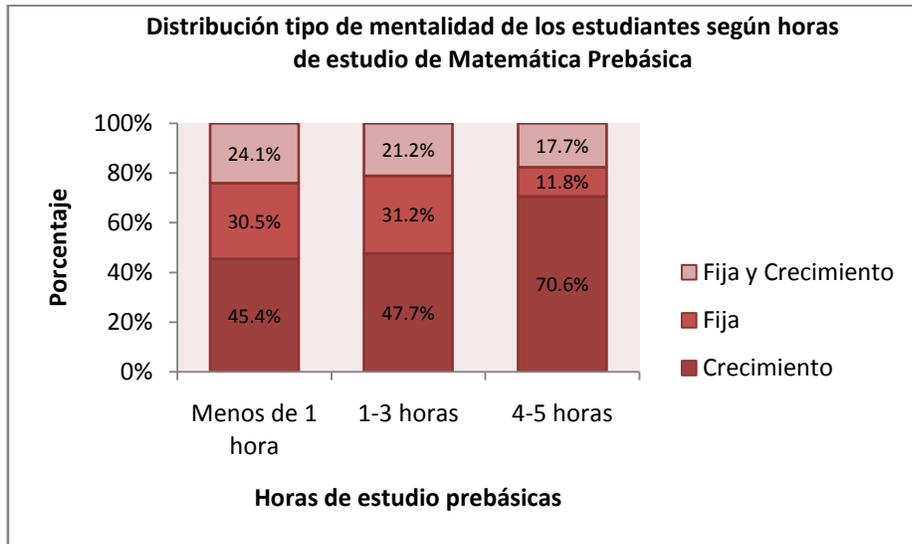


Figura 4.17

En la Figura 4.18 se observa la distribución del nivel de ansiedad de los estudiantes según el tipo de mentalidad. Una prueba de independencia para estas dos variables categóricas al 95% de confianza, concluye con un *valor p* de 0.356 que no hay asociación entre las variables.

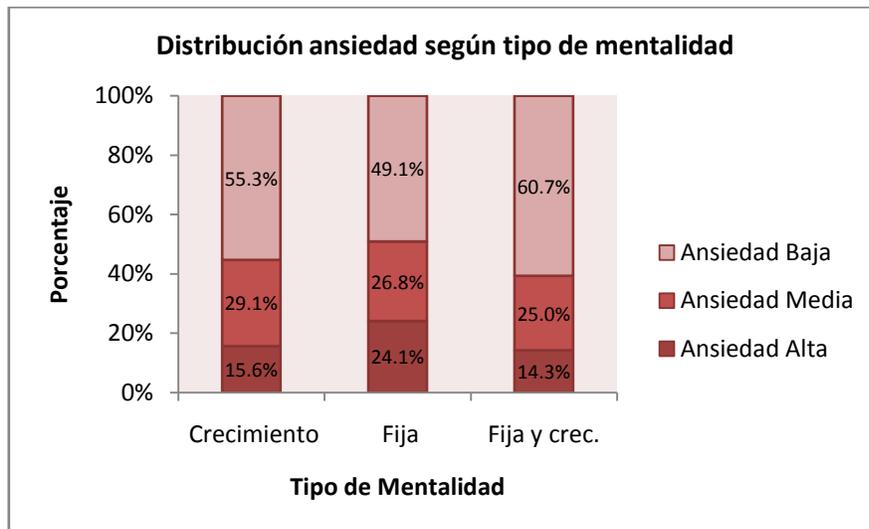


Figura 4.18

4.1.7 Motivación y planes al graduarse

En la Tabla 4.9 se muestran los reactivos 40 - 49 del cuestionario “*Estudio de algunos factores que afectan el rendimiento académico de los estudiantes de Mate 0066*”; con estos reactivos se examinaron las principales motivaciones de los estudiantes para ingresar a la universidad, destacándose, como se muestra en la Figura 4.19, que “alcanzar una profesión que me satisfaga” (P45), y “tener mejores oportunidades de empleo al graduarme” (P47) son las de mayor porcentaje, con 97.2% y 95%, respectivamente. También se encontró que lo menos importante a la hora de ingresar a la universidad fue “seguir a los amigos” (P42) y “no conseguir trabajo al salir de la escuela superior” (P43), como lo indican los estudiantes de Matemática Prebásica.

Tabla 4.9 Razones más importantes para ingresar a la Universidad

P40	Independizarme de mi familia	P45	Alcanzar una profesión que me satisfaga
P41	Presión de mis padres	P46	Expectativa de lograr un mejor ingreso económico
P42	Por seguir a mis amigos	P47	Tener mejores oportunidades de empleo al graduarme
P43	No conseguir trabajo al salir de la escuela superior	P48	Alcanzar un status y prestigio social
P44	Desarrollar mis talentos y habilidades	P49	Tener un mejor ambiente social y divertirme

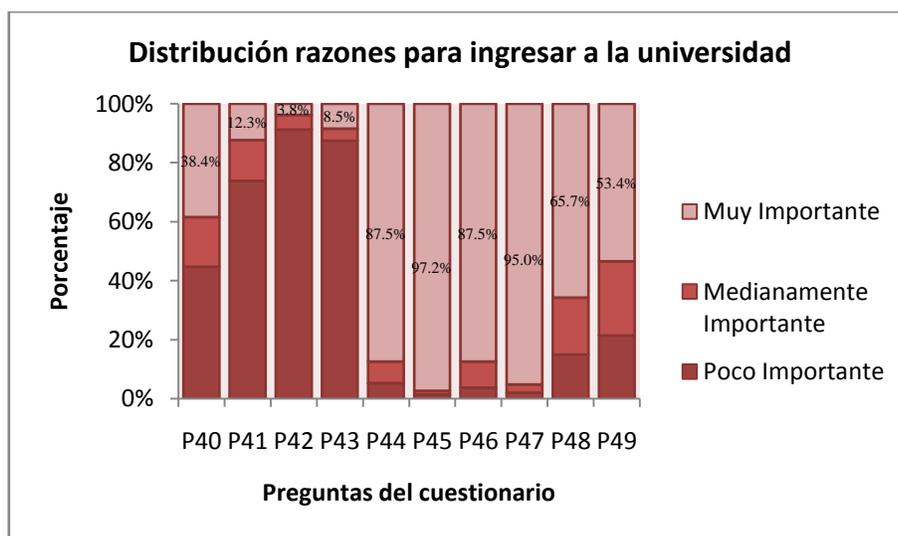


Figura 4.19

Con respecto a los planes de los estudiantes al graduarse, para el 76% es muy importante continuar con estudios graduados (P27). El 80.5% planea trabajar en una empresa o compañía de gran prestigio (P28) y el 60% afirma que es muy importante, establecer su propio negocio (P29). Estos resultados se presentan en la Figura 4.20:

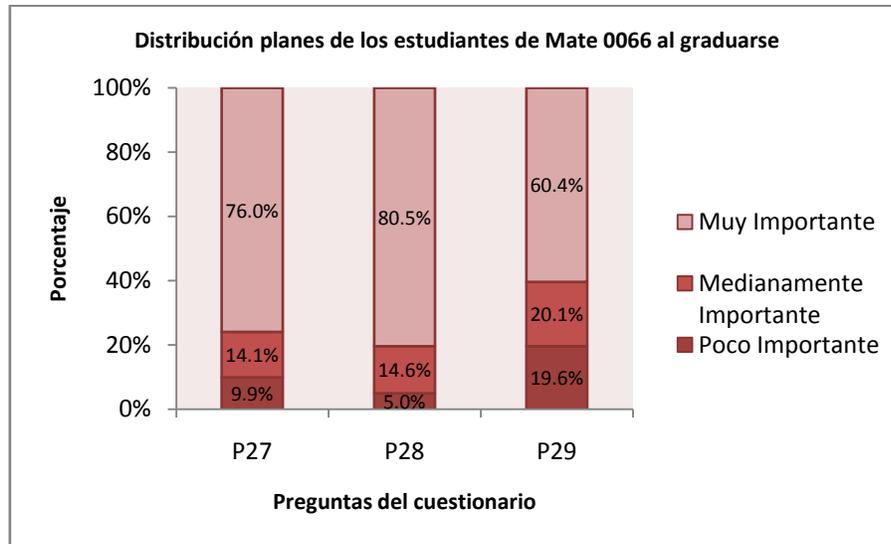


Figura 4.20

Cuando se analizaron los resultados anteriores según género, se encontró que el 72.8% de los hombres y el 87.6% de las mujeres, consideran muy importante continuar con estudios graduados. Resultados similares se habían encontrado cuando se discutieron las aspiraciones académicas de los estudiantes, según género. La prueba de hipótesis de independencia, establece con p valor de 0.354, que no hay asociación entre las variables. En cuanto a trabajar en una empresa o compañía de gran prestigio, el 72.8% de los hombres y el 80% de las mujeres, opinaron que es muy importante. Finalmente, el 69% de los hombres y el 61.7% de las mujeres consideran muy importante crear su propio negocio.

4.2 Construcción de categorías a partir del Análisis de Componentes

Principales (ACP)

Se realizó un análisis de componentes principales a partir de las variables cuantitativas del estudio, las cuales se definen a continuación:

- Promedio de escuela superior. Es el promedio que obtienen los estudiantes al terminar sus estudios de Escuela Superior, de acuerdo a su rendimiento académico durante todos los años. Esta variable se mide en un rango de 0 a 4.
- Promedio de quices. Es la nota promedio de los siete quices electrónicos que los estudiantes presentaron durante el curso de Matemática Prebásica. Esta puntuación se mide en una rango de 0 a 100 puntos.
- Puntuación en la prueba de aptitud del College Board. Esta prueba evalúa la habilidad para resolver problemas de razonamiento y problemas básicos de aritmética, álgebra y geometría, usando principios matemáticos básicos (College Board, 2002). Esta puntuación se mide en un rango de 200 a 800 puntos.
- Puntuación en la prueba de aprovechamiento del College Board. Esta prueba evalúa los conocimientos básicos de aritmética, álgebra y geometría, así como distintos niveles de conocimiento de las matemáticas e incluye ejercicios que requieren el uso de: memoria, comprensión, aplicación y análisis (College Board, 2002). Esta puntuación es la que se considera para ubicar al estudiante en su primer curso de matemáticas en la UPRM. Las puntuaciones para esta prueba también se miden en un rango de 200 a 800.

- Nivel de Ansiedad. Este se midió utilizando la Escala de Ansiedad (Muñoz y Mato, 2007), a partir de cinco factores distintos de la ansiedad. La puntuación corresponde al promedio de ansiedad general de los estudiantes, según las puntuaciones obtenidas en cada uno de los factores. Esta se mide en una escala de 0 a 136, la cual se llevó a la escala de 0 a 100, para facilitar la interpretación de los resultados.

A partir de la matriz de correlaciones de estas variables, se encontró que existe correlación lineal fuerte entre las puntuaciones de la prueba de aptitud y la prueba de aprovechamiento, del College Board (valor $p < 0.05$). Estas dos variables a su vez, están correlacionadas (correlación débil), con promedio de escuela superior y promedio de quices (valor $p < 0.05$). La variable ansiedad está correlacionada con aprovechamiento del College Board (valor $p < 0.05$). En la Tabla 4.10 se muestran los coeficientes de correlación de Pearson para las variables anteriores, los valores en “negrita” indican los coeficientes que resultaron significativos.

Tabla 4.10 Matriz de correlaciones

Variables	Promedio Escuela	Aptitud	Aprovechamiento	Ansiedad General	Promedio quices
Promedio Escuela	1.00				
Aptitud	0.18	1.00			
Aprovechamiento	0.32	0.78	1.00		
Ansiedad General	0.05	-0.12	-0.18	1.00	
Promedio quices	0.26	0.18	0.13	-0.09	1.00

Coeficiente de correlación de Pearson

El análisis por componentes principales es pertinente, para este estudio, por las correlaciones existentes entre las variables. De este modo, con la técnica se pueden establecer nuevas variables no correlacionadas o componentes, y también reducir la dimensionalidad de los datos, tomando en cuenta la máxima variabilidad. Es importante recordar que, en este estudio, las nuevas variables serán utilizadas en la regresión logística, junto con otras variables (cualitativas), para generar un modelo predictivo de la variable dependiente: aprueba o no aprueba el curso de Matemática Prebásica.

En la Tabla 4.11 se presenta la salida en MINITAB, de los valores propios y la proporción de varianza explicada por cada componente principal, así como la varianza explicada acumulada. Los resultados indican que los dos primeros componentes explican el 63% de la variabilidad total y los tres primeros componentes explican el 83% de la variabilidad total.

Tabla 4.11 Valores propios y varianzas explicadas por los componentes principales

	PC1	PC2	PC3	PC4	PC5
Valor propio	2.01	1.13	1.01	0.65	0.21
Proporción	0.40	0.23	0.20	0.13	0.04
Acumulado	0.40	0.63	0.83	0.96	1.00

La determinación del número de componentes, que se incluyeron en el análisis, se hizo a partir del criterio de Kaiser (Pla, 1986), que establece la selección de los componentes asociados a valores propios superiores a una cota, que suele fijarse como la varianza media $\frac{\sum \lambda_i}{p}$. En particular, cuando se trabaja con la matriz de correlación, el valor medio es 1. De acuerdo a ello cumplen con este criterio el primero, segundo y tercer componente, con

valores propios de 2.01, 1.13 y 1.01, respectivamente, como se observa en la Figura 4.21. Además, la inclusión del tercer componente, se debe a que explica el 20% de la variabilidad total de los datos, por lo cual se consideró relevante en el análisis.

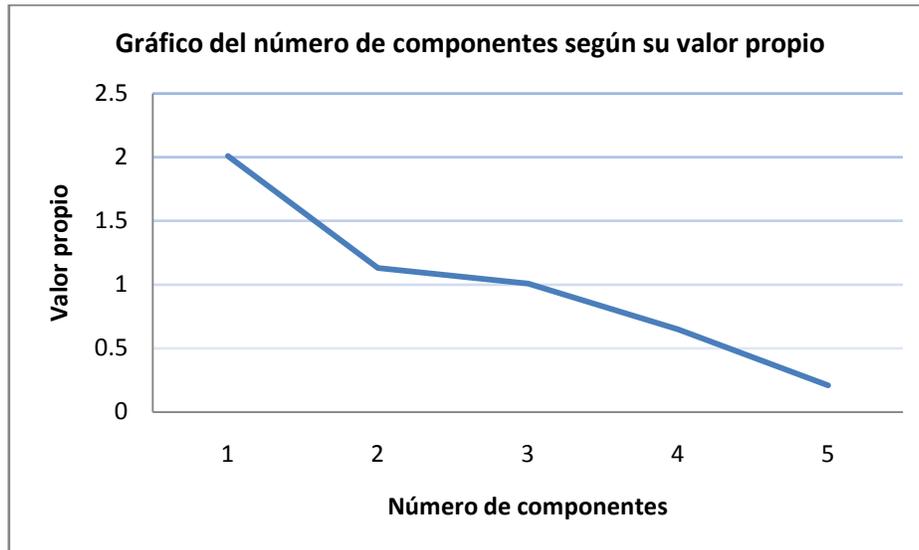


Figura 4.21

En la Tabla 4.12 se presentan los vectores propios para cada uno de los componentes principales, asociados a los valores propios, los cuales indican la correlación de cada una de las variables originales con los componentes y a partir de éstos, se crean las nuevas variables (componentes).

Tabla 4.12 Vectores Propios

	PC1	PC2	PC3	PC4	PC5
Prom. Quices (X_1)	0.29	-0.44	-0.64	0.56	-0.08
Prom. Esc. Sup (X_2)	0.35	-0.64	-0.01	-0.67	0.16
Aptitud CB (X_3)	0.61	0.22	0.27	0.26	0.67
Aprov. CB (X_4)	0.63	0.20	0.22	-0.03	-0.72
Ansiedad (X_5)	-0.16	-0.56	0.69	0.42	-0.10

Las tres variables nuevas quedan definidas por las siguientes combinaciones lineales:

$$PC1 = 0.29X_1 + 0.35X_2 + 0.61X_3 + 0.63X_4 - 0.16X_5$$

$$PC2 = -0.44X_1 - 0.64X_2 + 0.22X_3 + 0.20X_4 - 0.56X_5$$

$$PC3 = -0.089X_1 + 0.16X_2 + 0.67X_3 - 0.72X_4 - 0.10X_5$$

A partir del análisis por componentes principales se establecen las siguientes variables nuevas, de acuerdo a la interpretación de los primeros tres componentes:

1. **Preparación académica previa del estudiante (PC1).** El primer componente tiene correlación positiva con las cuatro variables académicas y correlación negativa con la variable ansiedad; predominando, la correlación positiva alta con las variables aptitud y aprovechamiento del College Board. Este componente podría considerarse como un indicador del rendimiento académico general del estudiante al ingreso del curso de Matemática Prebásica; el cual explica el 40% de la variabilidad total de los datos.
2. **Inercia hacia las matemáticas (PC2).** Este componente tiene la mayor correlación negativa con el promedio de escuela superior, seguido de una correlación igualmente negativa con la ansiedad y el promedio de quices, por lo que podría estar explicando el nivel de “inactividad” de los estudiantes hacia el curso de Matemática Prebásica, el cual se manifiesta en aquellos estudiantes que aún considerando el curso importante, su trabajo no es constante, los niveles de tensión necesarios para mantener su atención en la tarea son bajos (Víctor y Romper, 2002) y como resultado su promedio es bajo. Este componente explica el 23% de la variación total.

3. **Tendencia ansiosa hacia el aprendizaje de las matemáticas (PC3).** En el componente tres, se destaca la correlación positiva alta con el nivel de ansiedad y correlación negativa alta con el promedio de quices, seguido de correlación negativa con el promedio de escuela superior. Por esta razón, este componente podría explicar los resultados de los estudiantes que padecen de ansiedad (Roty, 2008), específicamente ansiedad ante la evaluación, en donde a mayor nivel de ansiedad menor es el rendimiento en los exámenes o pruebas. Este componente explica el 20% de la variación total.

Con el fin de interpretar los resultados obtenidos en el análisis, se realizaron dos tipos de gráficas: la gráfica que representa a los estudiantes en el plano de los componentes y la gráfica de las variables originales, también en el plano de los componentes. Desde el punto de vista geométrico, lo que se observa en las Figuras 4.22, 4.23 y 4.24, es que los puntos se sitúan siguiendo una elipse y se pueden describir por su proyección en la dirección del eje mayor de la elipse. Este eje corresponde a la recta que minimiza las distancias ortogonales. Así que la mejor aproximación en dos dimensiones, es la proyección sobre el plano de los dos ejes mayores del elipsoide, los cuáles vienen siendo las nuevas variables (componentes).

En la Figura 4.22 se presenta el diagrama de dispersión de los estudiantes, según los nuevos ejes (componentes 1 y 2). Las proximidades entre individuos se interpretan como similitud de comportamiento de estos, respecto a las variables. Se observa mayor homogeneidad en los puntos que se sitúan en valores positivos del primer componente. Esto significa que el grupo

de estudiantes con mayor “*preparación académica previa*” es más homogéneo, que el grupo de estudiantes con menor “*preparación académica previa*”, donde las variables aptitud y aprovechamiento del College Board son las que mayor contribución hacen al mismo.

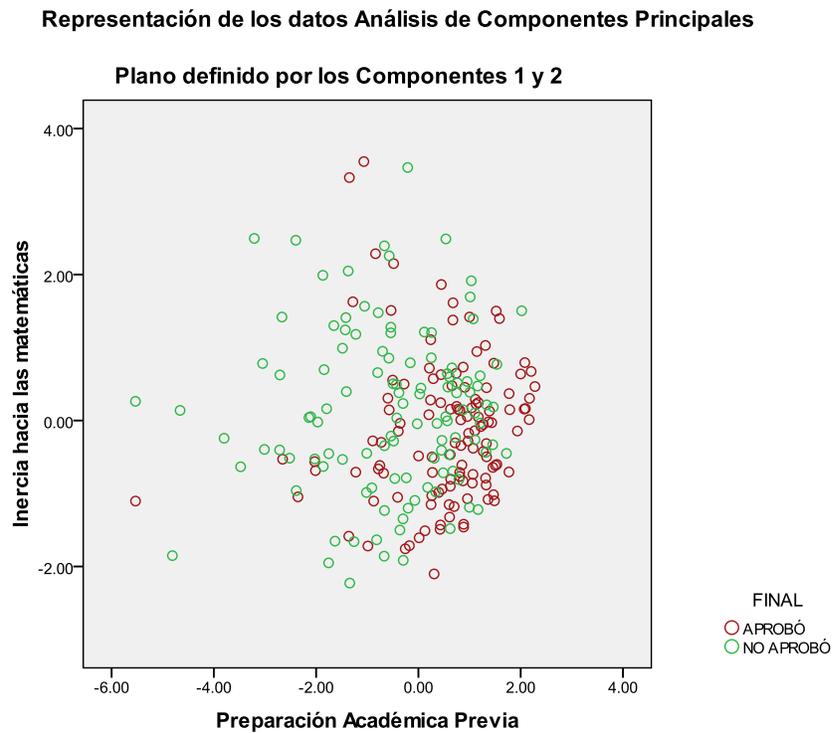


Figura 4.22

Los estudiantes que tienen puntuaciones positivas con respecto al segundo componente, representan los de mayor “*Inercia hacia las matemáticas*”, es decir, aquellos estudiantes que aún teniendo puntuaciones relativamente buenas en aptitud y aprovechamiento en matemáticas, tienen puntuaciones muy bajas en su promedio de graduación y puntuaciones bajas en su nivel de ansiedad. A partir de la gráfica se puede observar que la mayor concentración de estudiantes aprobados está en puntuaciones positivas de la preparación académica previa y puntuaciones negativas en la inercia hacia las matemáticas.

En la Figura 4.23 se muestra la dispersión de los estudiantes con respecto al primer y tercer componente, “Preparación académica previa” y “Tendencia ansiosa hacia el aprendizaje de las matemáticas”, respectivamente. En este gráfico se observa mayor dispersión e incluso la presencia de posibles outliers (valores alejados del resto de observaciones), en valores negativos del primer componente, que corresponden a los estudiantes con menor preparación académica, comportamiento similar al gráfico anterior. Además, los datos con mayor dispersión, son los que toman valores negativos para ambos componentes. Estos corresponden a estudiantes con menor preparación académica y también menor nivel de ansiedad, que podría corresponder al grupo de estudiantes que requiere atención porque su nivel de desinterés, sumado a su bajo nivel académico, podría rezagarlo en la Universidad.

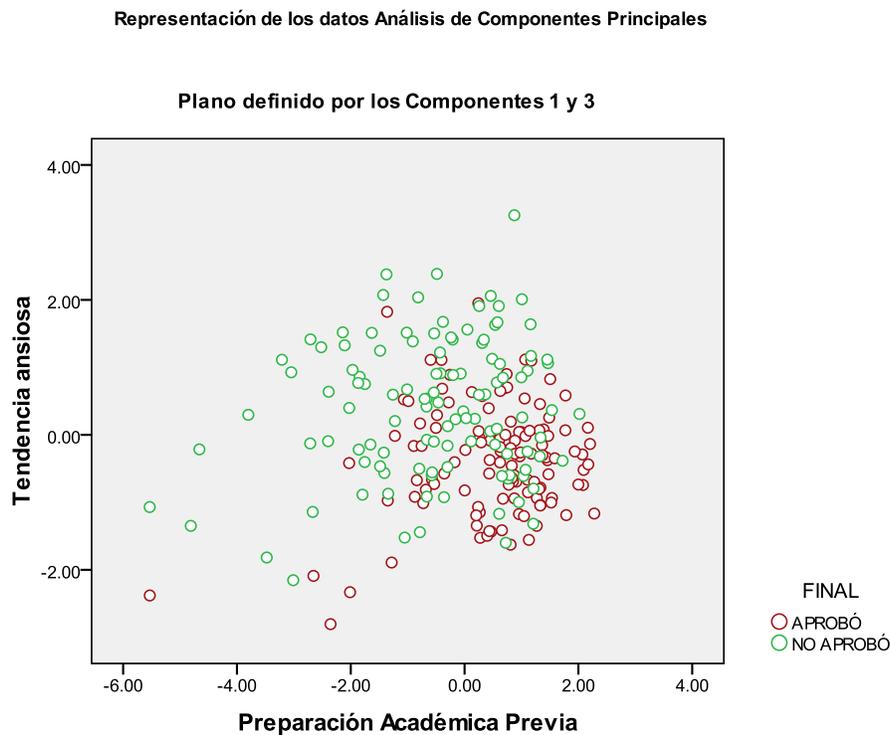


Figura 4.23

En la Figura 4.24 se muestra la dispersión de los estudiantes, en el plano del segundo y tercer componente. Se observa que la mayor proporción de estudiantes aprobados aparece en los valores negativos de inercia hacia las matemáticas y valores negativos de tendencia ansiosa. Esto es, aquellos estudiantes que estudian el material y se preparan para el examen, que serían los de menor inercia hacia las matemáticas, y además controlan la ansiedad al momento del examen.

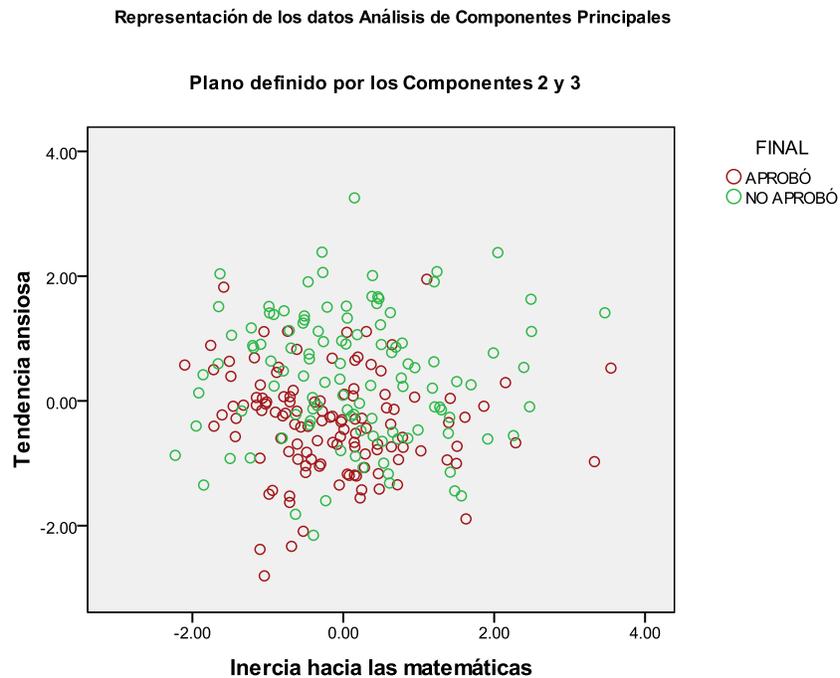


Figura 4.24

Para complementar la descripción anterior, se exploraron las variables originales en el plano generado por el primer y segundo componente. En esta gráfica las proximidades entre variables indican el grado de correlación que existe entre ellas. La Figura 4.25 muestra como las variables aptitud y aprovechamiento del College Board, están correlacionadas y el mayor

peso se lo están proporcionando al primer componente “*Preparación académica previa*”. Por otro lado, las variables promedio de graduación y promedio de quices también están relativamente cercanas y aportan, aunque de forma negativa al segundo componente: “*Inercia hacia las matemáticas*”. Además la variable ansiedad ejerce su mayor influencia sobre el segundo componente.

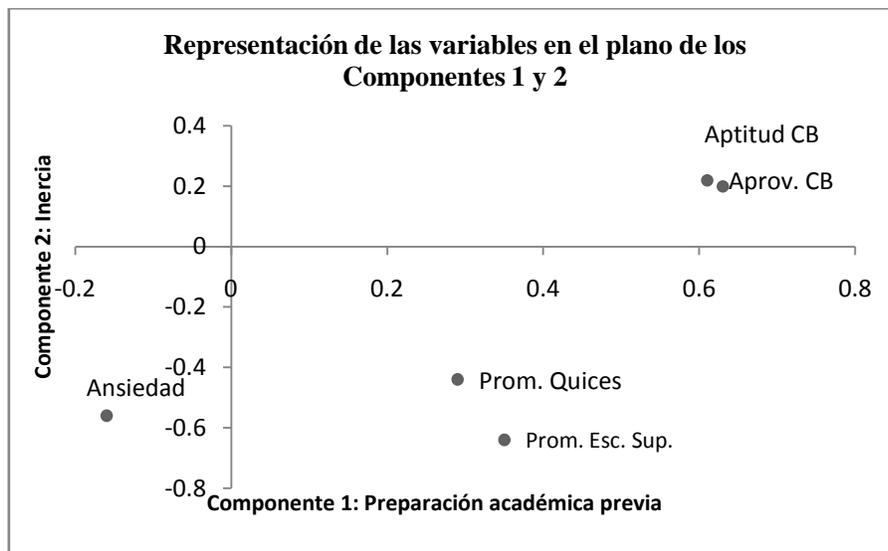


Figura 4.25

Del mismo modo, al representar las variables en el plano formado por el primer y tercer componente, como se muestra en la Figura 4.26, las variables mantienen una estructura similar al plano generado por el primer y el segundo componente, pero cambiando la dirección, según la correlación que tienen con las nuevas variables.

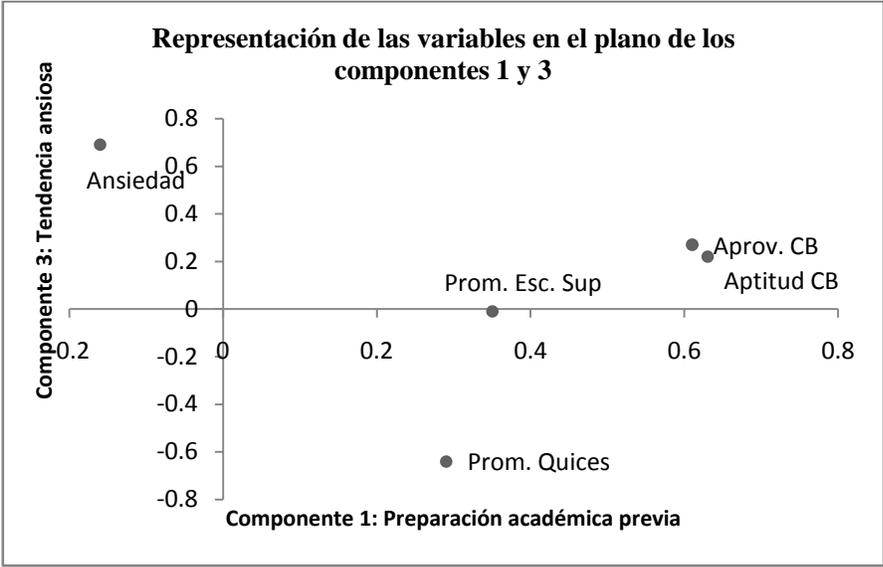


Figura 4.26

4.3 Construcción de un índice para predecir éxito o fracaso de los estudiantes a partir de la Regresión Logística

Para la predicción del éxito o fracaso de los estudiantes, en el curso de Matemática Prebásica, se utilizó el modelo de Regresión Logística Múltiple, debido a que es una de las técnicas recomendadas por su alto porcentaje de clasificación cuando la variable respuesta es de carácter dicotómica. En este caso la variable dependiente se categorizó en aprobados (121 estudiantes) y no aprobados (112 estudiantes).

Las variables que se utilizaron en un análisis exploratorio, para encontrar el mejor modelo de predicción del éxito o fracaso de los estudiantes fueron:

- **Variables Categóricas** (13). Género, Escuela de procedencia, Facultad a la que pertenece, Estudiante de Nuevo Ingreso, Gusto hacia las matemáticas, Facilidad hacia las matemáticas, Preparación de exámenes, Motivación, Importancia de los quices, Ser profesional, Tener buen empleo, Tipo de mentalidad y Percepción del esfuerzo.
- **Variables Ordinales** (3). Aspiraciones académicas, Nivel académico de los padres, Horas dedicadas a estudiar Matemática Prebásica.
- **Variables Continuas** (3). Preparación académica previa, Inercia hacia las matemáticas y Tendencia ansiosa, las cuales se habían obtenido a partir de la técnica de componentes principales.

El método utilizado para la selección del mejor modelo posible fue el “Bacward Stepwise (Wald)” o “*Eliminación progresiva*”, el cual comienza con el modelo completo y en cada paso va eliminando una variable, la menos significativa, es decir aquella que tiene el estadístico de t , en valor absoluto, más pequeño entre las variables incluidas en el modelo.

Una vez aplicada la técnica se obtuvo el siguiente modelo general:

$$\hat{y} = \frac{e^{-0.152+0.727x_1-0.395x_2+1.407x_{31}+0.380x_{32}-1.088x_4}}{1 + e^{-0.152+0.727x_1-0.395x_2+1.407x_{31}+0.380x_{32}-1.088x_4}}$$

o su equivalente con la función logit,

$$f(x) = -0.152 + 0.727x_1 - 0.395x_2 + 1.407x_{31} + 0.380x_{32} - 1.088x_4$$

En la Tabla 4.13 se presenta la descripción de las variables que entraron al modelo. La variable categórica, Percepción del Esfuerzo, fue codificada usando variables indicadoras en los niveles de su respuesta, como se muestra en la Tabla 4.14.

Tabla 4.13 Variables que entraron al modelo

Variable	Descripción	Niveles
x_1	Preparación académica previa	Continua
x_2	Inercia a las matemáticas	Continua
x_3	Percepción del Esfuerzo	Nominal (3)
x_4	Tendencia ansiosa	Continua

Tabla 4.14 Codificación de la variable categórica

Esfuerzo	Parámetro Codificado	
	(1)	(2)
Desacuerdo	1	0
Indiferente	0	1
De acuerdo	0	0

En la Tabla 4.15 se presentan los coeficientes de los parámetros estimados, con su correspondiente error estándar, valor p y odds ratio.

Tabla 4.15 Coeficientes del mejor modelo

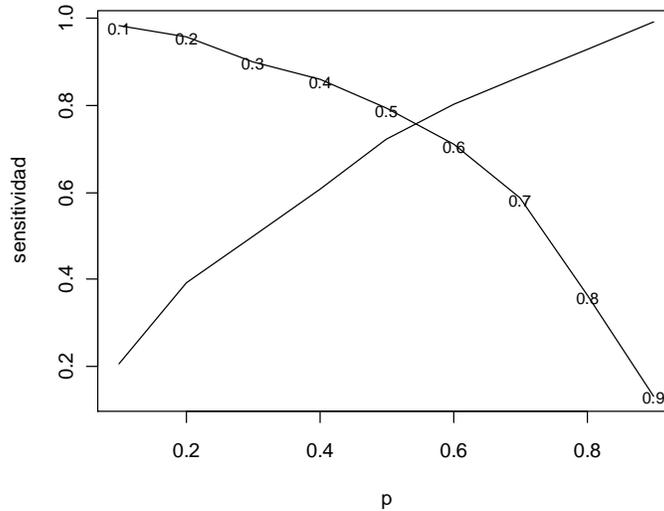
Variable	Coefficiente	Error Estándar	<i>p_valor</i>	Odds ratio
Preparación Académica	0.727	0.135	0.000	2.069
Inercia	-0.395	0.160	0.014	0.674
Percepción del Esfuerzo			0.077	
Percepción del Esfuerzo (1)	1.407	0.645	0.029	4.083
Percepción del Esfuerzo (2)	0.380	0.449	0.398	1.463
Tendencia Ansiosa	-1.088	0.185	0.000	0.337
Constante	-0.152	0.186	0.415	0.859

La bondad del ajuste del modelo se verificó a partir de la Prueba de Hosmer y Lemeshow. Como $\hat{C} = 3.080$ y $\hat{C} < \chi^2_{0.05,8} = 15.50$, se concluye que el modelo ajustado es adecuado.

Como el modelo de regresión logística se utilizaría para clasificar a los estudiantes de prebásicas, según aprobaran o no el curso, se buscó el valor del p-óptimo que consiguiera una buena clasificación. Para ello se elaboró una gráfica de la sensibilidad (porcentaje de estudiantes que habían aprobado el curso y fueron clasificados correctamente por el modelo) versus especificidad (porcentaje de estudiantes que no aprobaron el curso y que también fueron clasificados correctamente).

Figura 4.27

Sensitividad y Especificidad para varios valores de p



Utilizando el p – *óptimo* de la prueba de sensibilidad y especificidad ($p = 0.560$), indicado en la Figura 4.27, se tiene que el 75.1% de los estudiantes son clasificados correctamente, lo cual es satisfactorio. Notar que en nuestro caso las clases estaban bastante balanceadas, 121 estudiantes aprobados y 112 estudiantes no aprobados, por lo que el p – *óptimo* fue cercano a 0.50. En la Tabla 4.16 se muestra el porcentaje de clasificación:

Tabla 4.16 Clasificación de los estudiantes

	Valores Estimados		
Valores Observados	No aprobó	Aprobó	Porcentaje Correcto
No aprobó	84	28	75.0
Aprobó	30	91	75.2
	Porcentaje Promedio		75.1

Interpretación de los coeficientes del modelo

De acuerdo a lo anterior, tenemos un modelo con un 75.1% como tasa de buena clasificación. Para analizar el impacto de los coeficientes de las variables explicativas en el modelo, se consideraron las razones de probabilidad (odds ratio). La variable con mayor “odds ratio” fue la variable categórica Percepción del Esfuerzo, con 4.083; como la categoría de referencia para esta variable fue estar de acuerdo con la percepción del esfuerzo, se interpretaría como: la razón entre la probabilidad de éxito (aprobar el curso) vs. fracaso (no aprobar), en un estudiante que está en desacuerdo con la percepción del esfuerzo es cuatro veces mayor que la razón entre la probabilidad de éxito vs. fracaso en un estudiante que está de acuerdo con respecto al esfuerzo, manteniendo las demás variables constantes.

La variable continua con mayor “odds ratio” fue la preparación académica previa, con 2.069; esto significa que un cambio en una unidad de la variable “preparación académica previa”, manteniendo todas las demás constantes, aumentaría en un factor de 2, la probabilidad de aprobar el curso. Recordemos que el componente preparación académica previa es una combinación lineal de las variables originales. Un ejemplo con las variables del modelo indica que una unidad de diferencia en la variable preparación académica previa implicaría una diferencia de 5 décimas en promedio de escuela superior, 30 puntos en aptitud matemática, 30 puntos en aprovechamiento matemático, 20 puntos en ansiedad y 15 puntos en la nota de quices.

En particular, un estudiante con puntuaciones en el punto medio de todas las variables, esto es, 50 en promedio de quices, 2.0 en promedio de escuela superior, 500 en aptitud del College Board, 500 en aprovechamiento del College Board, 50 en el nivel ansiedad y con un nivel de indiferencia hacia el esfuerzo en matemáticas tendría tan sólo un 7% de probabilidad de aprobar el curso, de acuerdo al modelo.

Estadísticas influenciales en el modelo de regresión logística

Al analizar los diagnósticos de regresión: residuales de la devianza, los residuales de Pearson y puntos Leverage alto, se identificaron algunas observaciones atípicas. El diagnóstico de los residuales estudentizados indicó que las observaciones 50, 53, 74, 94, 108 y 157 son outliers. Las observaciones 44, 77, 94 y 158, son identificadas como puntos de Leverage alto. Finalmente, a partir del gráfico de la distancia Cook, según cada observación no se encontró ningún punto influyente en el modelo (Anexo 5).

5. CONCLUSIONES Y TRABAJOS FUTUROS

Este estudio se originó en la problemática de los cursos de Matemática Prebásica del Departamento de Ciencias Matemáticas del RUM, debido a que el porcentaje de estudiantes que fracasan cada semestre es muy alto, alrededor del 50%. La preocupación es latente y las iniciativas que el Departamento de Ciencias Matemáticas ha implementado se enfocan en el aspecto académico, sin embargo, se ha encontrado en este estudio, que además de la preparación académica previa, otras variables como la ansiedad frente a los exámenes, el desinterés hacia el curso y hacia las matemáticas, la “apatía” hacia las matemáticas, y sus hábitos deficientes de estudio influyen en los resultados académicos de los estudiantes en Matemática Prebásica. Estos se describen a continuación:

Los aspectos más destacados en el perfil de los estudiantes de Matemática Prebásica, fueron:

Percepción de que están bien preparados. Este aspecto se vio reflejado en que el 69.2% de los estudiantes afirmó tener una aceptable, buena o excelente preparación en matemáticas. El 53% opinó que sus maestros los habían preparado muy bien en los cursos de matemáticas de la escuela superior, y la puntuación obtenida en el promedio de escuela superior fue de 3.71, que corresponde a un promedio de A.

Altas expectativas académicas y profesionales. El 76% de los estudiantes tiene aspiraciones de continuar con estudios graduados y el 95% desea alcanzar una profesión que le satisfaga.

Pobres hábitos de estudio. El 50.3% de los estudiantes dedicaba menos de una hora a la semana para preparar su curso, el 17.5% nunca utiliza el libro de texto, el 48% dedicaba menos de una hora a la semana en la preparación del examen y tampoco presentaban los quices electrónicos de práctica.

Falta de motivación. El 40% de los estudiantes afirmó no interesarse en el curso y el 51% afirmó que no le gustan las matemáticas. Sin embargo, para el 95% de los estudiantes el curso de Matemática Prebásica es importante y el 75% reconoce que debe esforzarse.

Ansiedad hacia las matemáticas. El promedio de ansiedad de los estudiantes hacia las matemáticas fue de 37.5, lo que representa un nivel medio de ansiedad. Su mayor manifestación se presenta durante los exámenes o en la fecha próxima a los mismos.

A partir del análisis de componentes principales sobre las variables cuantitativas: promedio de escuela superior, puntuación en el Examen del College Board, Puntuación en el Examen de Aprovechamiento del College Board y Ansiedad, se identificaron tres nuevas variables, no correlacionadas: Preparación académica previa, Inercia hacia las matemáticas y Tendencia ansiosa.

La mayor correlación positiva del primer componente se obtuvo con las variables de aptitud y aprovechamiento matemático del College Board; quedando en un orden de menor importancia el promedio de escuela superior. Esto podría sugerir un replanteamiento en el cálculo del IGS, para ingreso a la Universidad, pues a la luz de los resultados no tiene el mismo peso en el “rendimiento académico real” de los estudiantes, que las otras variables.

El segundo componente, inercia hacia las matemáticas, estaba correlacionado negativamente con el promedio de quices, promedio de escuela superior y la ansiedad, básicamente describe el nivel de apatía de algunos estudiantes en el curso y el tercer componente, tendencia ansiosa, describe como los niveles de ansiedad durante los exámenes pueden estar afectando sus resultados académicos en el curso.

En los gráficos de dispersión generados por los planos de los componentes, se observa la agrupación de estudiantes según sus similitudes en las variables. Al respecto, se encontró que la mayor concentración de estudiantes que aprobaron el curso está en puntuaciones positivas de la preparación académica previa y puntuaciones negativas en la inercia hacia las matemáticas, cuando se analiza en el plano de los componentes 1 y 2. En los componentes 1 y 3, se tiene la misma relación, pero cambiando inercia por la variable tendencia ansiosa. Estos gráficos resaltan la preparación académica previa como una variable de interés, asociada al rendimiento académico de los estudiantes.

Finalmente se encontró un modelo de predicción de la aprobación de los estudiantes en el curso de Matemática Prebásica, utilizando la Regresión Logística Múltiple. Las variables significativas en este modelo fueron los 3 componentes principales: preparación académica previa, inercia hacia las matemáticas y tendencia ansiosa, y una variable categórica, la percepción del esfuerzo. La tasa de clasificación general, de este modelo fue de 75.1%.

Los resultados en el modelo también señalan la importancia de la preparación académica previa en matemáticas, como indicador del éxito de los estudiantes en el curso de Matemática Prebásica. Así por ejemplo, en los componentes el 40% de la variabilidad de los datos fue explicada por la preparación académica previa y en el modelo de regresión logística este coeficiente tiene un odds ratio de 2.069. Bajo esa perspectiva, recomendaríamos lo siguiente:

- Como las mayores deficiencias se observan a nivel de hábitos de estudio sería conveniente que la Universidad proporcione orientación a los estudiantes, en destrezas de estudio, manejo efectivo del tiempo, manejo de la ansiedad ante los exámenes, técnicas de motivación, entre otros temas, que fortalezca especialmente al estudiante de nuevo ingreso.
- Para disminuir el impacto de la inercia de los estudiantes hacia las matemáticas, es importante que los estudiantes reconozcan la necesidad del esfuerzo para obtener los resultados deseados. Continuar con estrategias de evaluación, como asignaciones, pruebas cortas, etc., sin embargo estas estrategias deberían formar parte de las

experiencias previas de los estudiantes, ya que mientras más arraigado se encuentre su nivel de inercia, más tiempo le tomará al estudiante cambiar ese patrón de conducta. Para ello, se sugiere iniciar estrategias conjuntas con el Departamento de Educación, a través de Programas de Preparación de maestros, para elaborar políticas sistemáticas a tales fines.

Como trabajo futuro se sugiere analizar otras variables que también pueden estar influyendo en el rendimiento académico de los estudiantes, como por ejemplo, el síndrome Burnout (Síndrome de “estar quemado”), que podría explicar o estar correlacionada con la “apatía” de los estudiantes hacia el curso. El síndrome de Burnout se define como un estado de agotamiento físico, emocional y cognitivo producido por el contacto permanente con situaciones generadoras de estrés (Barraza, 2008). El término originalmente fue utilizado por Freudenberger (1974) citado en Barraza, 2008 y estaba relacionado con trabajadores, sin embargo se ha generalizado a otros grupos, en este caso particular los estudiantes, quienes también se someten a situaciones de estrés como los exámenes, quices y exposiciones, entre otros.

REFERENCIAS

- [1]. Acuña, E. (2007). *Análisis de Regresión*. Departamento de Ciencias Matemáticas. Universidad de Puerto Rico – Recinto Universitario de Mayagüez.
- [2]. Alvarado M. J. (2008). *Evaluación de la Ansiedad ante un curso de matemáticas*. Universidad de Puerto Rico, Recinto de Mayagüez.
- [3]. Barraza, A. (2008). *Burnout estudiantil: un enfoque unidimensional*. Revista Electrónica Psicología Científica. Consultado en www.psicologiacientifica.com.
- [4]. Carvajal, P., Mosquera, J. y Artamonova, I. (2009). *Modelos de predicción del rendimiento académico en Matemáticas I en la Universidad Tecnológica de Pereira*. Scientia at Technica XV, (43), 258-263.
- [5]. García, M., Alvarado, J. y Jiménez, A. (2000). *La predicción del Rendimiento Académico: regresión lineal versus regresión logística*. Psicothema, 12, (2), 248-252. Consultado en <http://www.psicothema.com/pdf/558.pdf>.
- [6]. Dweck, C. (2006). *MINDSET The new psychology of success*. New York: Random House.
- [7]. Díaz A., Garnica, E., González M. P., & Torres, E. (1994). *Una aplicación del análisis de componentes principales en el área educativa*. Economía XIX (9), 55-72. Consultado en http://ies.faces.ula.ve/Revista/Articulos/Revista_09/Pdf/Rev09Gonzalez_Diaz.pdf
- [8]. Hosmer, D. y Lemeshow, S. (1989). *Applied Logistic Regression*. Estados Unidos de América: WILEY.
- [9]. Ospina, D. (2001). *Introducción al muestreo*. Bogotá: Editorial UNIBIBLOS, Universidad Nacional de Colombia.
- [10]. Peña, Daniel. (2002). *Análisis de datos multivariantes*. Madrid: Mc Graw Hill/ Interamericana de España.
- [11]. Petriz, M., Barona, C., López, R. y Quiroz, Jackeline. (2010) *Niveles de desempeño y actitudes hacia las matemáticas en estudiantes de la Licenciatura en administración en una universidad estatal mexicana*. Revista Mexicana de Investigación Educativa, 15 (47), 1223-1249. Consultado en <http://redalyc.uaemex.mx/redalyc/pdf/140/14015564012.pdf>.

- [12]. Pla, L. (1986). *Análisis Multivariado: Método de componentes principales*. Venezuela: Secretaría General de Educación.
- [13]. Porcel, E., Dapozo, G. y López, M. (2010). *Predicción del Rendimiento Académico de alumnos de primer año de la FACENA (UNNE) en función de su caracterización socioeducativa*. Revista Electrónica de Investigación Educativa, 12 (2). Consultado en <http://redie.uabc.mx/vol12no2/contenido-contenido.html>
- [14]. Rojo, J. (2007). *Regresión con variable dependiente cualitativa*. Consultado http://www.cchs.csic.es/web_UAE/tutoriales/PDF/Regresion_variable_dependiente_dicotomica_3.pdf
- [15]. Roty, M. (2008) *The relationship between mathematics anxiety and emotional intelligence*. <http://library.uprm.edu:2058/pqdweb?index=1&did=1537004321&SrchMode=1&sid=2&Fmt=2&VInst=PROD&VType=PQD&RQT=309&VName=PQD&TS=1272454569&clientId=45091>. Recuperado de la Base de Datos PROQUEST.
- [16]. Santori, M. (2005). Artículo Prensa del RUM. Recuperado el 25 de febrero de 2010. <http://www.uprm.edu/news/articles/as2005150.html>
- [17]. Spybrook, Janet. (2008) *The relationship among working memory, mathematics anxiety, and mathematics achievement in developmental mathematics courses in college community*. <http://library.uprm.edu:2058/pqdweb?index=0&did=1678560401&SrchMode=2&sid=3&Fmt=2&VInst=PROD&VType=PQD&RQT=309&VName=PQD&TS=1272455063&clientId=45091>. Recuperado de la Base de Datos de PROQUEST.
- [18]. Scheaffer, R., Mendenhall, W. y Ott, L. (1986). *Elementos de Muestreo*. México: Grupo Editorial Iberoamérica.
- [19]. UPRM (2005). Estudio de Calificaciones y Fracasos. Oficina de Investigación Institucional y Planificación del Recinto Universitario de Mayagüez (RUM).
- [20]. (2002) Usos e interpretación de la las pruebas de aprovechamiento académico del College Board. Vol. 16 (2). Consultado en <http://www.collegeboard.com/ptorico/noviembre02/aprovechamiento.html>

Anexo 1. Cuestionario aplicado a los Estudiantes



ESTUDIO DE ALGUNOS FACTORES QUE AFECTAN EL RENDIMIENTO ACADÉMICO DE LOS ESTUDIANTES DE MATE 0066

Protocolo

El propósito de este estudio es determinar algunos de los factores que inciden en el rendimiento académico de los estudiantes del curso MATE 0066. El cuestionario se divide en tres partes. La parte I consta de 27 reactivos, relacionados con aspectos sociodemográficos, preparación en la Escuela Superior, expectativas de ingreso al RUM, entre otros. La parte II consta de 10 reactivos para indagar sobre percepción de los estudiantes hacia las matemáticas y hacia el curso de prebásicas y la parte III, consta de 10 reactivos, donde se indaga sobre las razones por las cuáles ingresó a la Universidad. Estas dos últimas escalas se presentan en escala Tipo Likert. **No hay contestaciones correctas ni incorrectas. Sin embargo, le solicitamos sinceridad en sus respuestas, para obtener resultados confiables en la investigación.**

**ESTUDIO DE ALGUNOS FACTORES QUE AFECTAN EL RENDIMIENTO
ACADÉMICO DE LOS ESTUDIANTES DE MATE 0066**

A continuación encontrará una serie de preguntas relacionadas a características que pueden influir en su rendimiento académico. **Responda las preguntas con mucha confianza y honestidad.** Recuerde que no representa una nota para su curso de MATE 0066 y que sus contestaciones serán confidenciales.

PARTE I. Marque con X o responda, según sea el caso:

1. Género: Masculino Femenino
2. Edad: 16-17 años 18-19 años 20-21 años 22 o más años
3. Escuela de Procedencia: Pública Privada
4. Programa al que pertenece en el RUM: _____
5. Grado más alto que desea alcanzar:
Bachillerato Maestría Doctorado Otro ¿Cuál? _____
6. Preparación Académica del Padre:
Escuela Elemental Escuela Superior Grado Asociado
Bachillerato Maestría Doctorado
7. Preparación Académica de la Madre:
Escuela Elemental Escuela Superior Grado Asociado
Bachillerato Maestría Doctorado
8. Años de estudio universitario en el RUM:
Nuevo Ingreso 1 año 2 años 3 o más años
9. ¿Recibe Beca o algún tipo de ayuda económica? Sí No Si su respuesta es sí;
Indique el tipo de beca/ayuda y número de horas a la semana que dedica a la actividad.

<i>Tipo de Beca / Ayuda</i>	<i>No. de horas/ semana</i>	<i>Tipo de Beca/Ayuda</i>
<input type="checkbox"/> Actividad atlética	_____	<input type="checkbox"/> Ayuda suplementaria
<input type="checkbox"/> Banda/orquesta	_____	<input type="checkbox"/> PELL
<input type="checkbox"/> Estudio y Trabajo	_____	<input type="checkbox"/> AMS
<input type="checkbox"/> Otros Indique	_____	

10. ¿Trabaja mientras estudia? Sí No Si su respuesta es Sí,
A Tiempo Completo A Tiempo parcial

11. ¿Qué tan preparado está para los cursos de matemáticas que tomará en la Universidad?
 Excelente Bueno Aceptable Regular Deficiente
 (A) (B) (C) (D) (F)
12. ¿Cuál fue el último curso de matemáticas que llevó en la Escuela Superior, el grado en que lo tomó y la nota del curso? _____.
13. ¿Cuántas horas a la semana dedicaba a estudiar matemáticas en la Escuela Superior?
 Menos de 1 hora 1-3 horas 4-5 horas 6 o más horas
14. ¿Tenía a su disposición textos de matemáticas para sus cursos en la Escuela Superior?
 Sí No . Si respondió Sí, ¿con cuánta frecuencia usaba los libros de texto?
 Nunca 1 vez/semana 2 veces/semana 3 o más veces/semana
15. ¿Considera que sus maestros de la Escuela Superior lo prepararon adecuadamente en las matemáticas que requiere en la Universidad? S Nd
16. ¿Con cuánta frecuencia sus padres o encargados estaban al tanto de su desempeño académico en la Escuela Superior?
 Nunca Casi nunca Algunas veces Casi siempre Siempre
17. ¿En qué momento hubiera preferido tomar el curso de prebásicas?
 Último año Escuela Superior Sesión de Verano Curso Regular
18. Preferiría que su curso de prebásicas lo enseñara:
 Maestro/Escuela Superior Profesor/Universidad Instructor
19. Indique el promedio de graduación que obtuvo en la Escuela Superior: _____
20. Indique la puntuación obtenida en la prueba de Aptitud Matemática del College Board:
 21. Indique la puntuación obtenida en la prueba de Aprovechamiento en Matemáticas del College Board: _____
22. Indique el número de créditos matriculados durante este semestre: _____
23. ¿Está matriculado en MATE 0066 por primera vez? Sí No
 Si su respuesta es No, está tomando el curso por: Segunda vez Tercera vez o más
24. ¿Cuánto tiempo (horas) dedica para estudiar MATE 0066 durante la semana?
 Menos de 1 hora 1-3 horas 4-5 horas 6 o más horas
25. Estudia usualmente en forma: Individual En Grupo (2 o más personas)
26. ¿Qué recursos disponibles utiliza como apoyo a su curso de prebásicas?
 Pruebas Electrónicas Biblioteca Horario de Oficina/Instructores
 Tutorías Externas Otros, ¿cuáles? _____

27. Marque con X, la que mejor identifique su situación sobre ¿Cuáles son sus planes al graduarse? Utilice la siguiente escala:

1 No es importante	2 Poco importante	3 Medianamente importante	4 Importante	5 Muy importante
-----------------------	----------------------	------------------------------	-----------------	---------------------

		1	2	3	4	5
1	Trabajar en una empresa o compañía de gran prestigio					
2	Continuar estudios graduados					
3	Crear mi propio negocio					

PARTE II.

A continuación encontrará unas afirmaciones relacionadas con la matemática y el curso de prebásicas. Marque con X su opinión utilizando la siguiente escala:

1 Totalmente en desacuerdo	2 En desacuerdo	3 Indiferente	4 De acuerdo	5 Totalmente de acuerdo
-------------------------------	--------------------	------------------	-----------------	----------------------------

	AFIRMACIONES	1	2	3	4	5
1	Me gustan las matemáticas.					
2	Entiendo fácilmente las matemáticas.					
3	El curso de prebásicas es fácil.					
4	Me esfuerzo para aprobar el curso de prebásicas.					
5	Me intereso por estudiar el curso de prebásicas.					
6	Es importante para mí aprobar este curso.					
7	Me preparo antes de presentar los exámenes del curso.					
8	Me encuentro muy motivado para aprender matemáticas.					
9	Los "quizzes" electrónicos son útiles en el curso de prebásicas.					
10	El curso de prebásicas me ayudará para los siguientes cursos de matemáticas que debo tomar.					

PARTE III.

Marque con X, la que mejor identifique su situación sobre ¿cuáles fueron las razones más importantes para ingresar a la Universidad? Utilice la siguiente escala:

1 No fue importante	2 Poco importante	3 Medianamente Importante	4 Importante	5 Muy importante
------------------------	----------------------	------------------------------	-----------------	---------------------

	AFIRMACIONES	1	2	3	4	5
1	Independizarme de mi familia.					
2	Presión de mis padres.					
3	Por seguir a mis amigos(as).					
4	No conseguí trabajo al salir de la Escuela Superior.					
5	Desarrollar mis talentos y habilidades.					
6	Alcanzar una profesión que me satisfaga.					
7	Expectativa de lograr un mejor ingreso económico.					
8	Tener mejores oportunidades de empleo al graduarme.					
9	Alcanzar un status y prestigio social.					
10	Tener un mejor ambiente social y divertirme.					

ANEXO 2. AUTORIZACIÓN DEL COMITÉ CPSHI/IRB

UNIVERSIDAD DE PUERTO RICO EN MAYAGÜEZ
DECANATO DE ASUNTOS ACADÉMICOS
COMITÉ PARA LA PROTECCIÓN DE LOS SERES HUMANOS EN LA
INVESTIGACIÓN
(CPSHI/IRB-- 00002053)

11-10-CR-02

13 de octubre de 2010

Claudia L. Rodríguez
Estudiante Graduada
Estadística
UPRM

Estimada estudiante:

El comité revisó su proyecto: *"Aplicación de la técnica de Componentes Principales para identificar los factores que afectan el rendimiento académico de los estudiantes de matemáticas pre básicas del Recinto Universitario de Mayagüez"* y luego de evaluar la documentación sometida, aprueba de forma expedita su investigación al examinar que no existen riesgos significativos para los participantes.

La aprobación de su propuesta de investigación se extiende desde el 13 de octubre de 2010 hasta el 13 de octubre de 2011. Le recuerdo que cualquier modificación de su proyecto necesitaría pasar por una nueva revisión por parte de este Comité.

Le deseo mucho éxito en su trabajo de investigación y quedo a sus órdenes para cualquier pregunta o clarificación ulterior que estimase necesaria.

Cordialmente,



Dr. Brian Muñoz
Presidente CPSHI
UPRM

ANEXO 3. CONSENTIMIENTO INFORMADO

**UNIVERSIDAD DE PUERTO RICO
RECINTO UNIVERSITARIO DE MAYAGÜEZ
DEPARTAMENTO DE CIENCIAS MATEMÁTICAS**

HOJA DE CONSENTIMIENTO DE PARTICIPACIÓN (ESTUDIANTE)

Saludos. Soy Claudia Liliana Rodríguez Morantes, estudiante Graduada del Departamento de Matemáticas, del Recinto Universitario de Mayagüez. Estoy llevando a cabo un estudio, como parte de mi proyecto de Tesis cuyo título es "Aplicación de la técnica de componentes principales para identificar los factores que afectan el rendimiento académico de los estudiantes de matemáticas prebásicas del Recinto Universitario de Mayagüez", dirigida por el Dr. Edgardo Lorenzo González, Profesor del Departamento de Matemáticas.

En este estudio se analizarán una serie de variables que pueden estar influyendo en el desempeño académico de los estudiantes y que no han sido atendidas, motivo por el cual se dan un número alto de bajas y pérdida del curso. Estoy interesada en determinar dichos factores y sugerir algunas acciones que permitan atender la problemática.

Por esta razón te estoy invitando a participar de este proyecto de investigación, respondiendo tres cuestionarios, donde proporcionarás información sobre los factores que afectan el rendimiento académico, otro sobre la actitud hacia las matemáticas y el tercero, sobre la ansiedad hacia las matemáticas.

Tu participación es libre y voluntaria y aunque hayas respondido alguno de los cuestionarios puedes informar que ya no deseas participar en el estudio, en ese caso te comunicarías con la persona a cargo del estudio. La información que se obtenga a partir de los cuestionarios es libre y confidencial, únicamente será utilizada para propósitos de este estudio. No recibirá beneficios económicos por participar de este estudio. Los beneficios serán proporcionar y atender la problemática asociada con el bajo rendimiento académico. No se anticipa que el responder los cuestionarios te puedan generar algún daño físico o psicológico, ni tampoco te afectará la nota de tu curso MATE 0066.

Si tienes alguna duda acerca del estudio puedes aclararla ahora o en cualquier momento que lo desees comunicándote con Claudia Liliana Rodríguez Morantes al (787) 506 3417 o por email claudia.rodriguez3@upr.edu

Si tienes alguna pregunta o sientes que no has sido tratado como se indicó anteriormente puedes comunicarte con el Presidente del Comité de Participación de Sujetos Humanos en la Investigación, Dr. Brian Muñoz al email brian.munoz@upr.edu

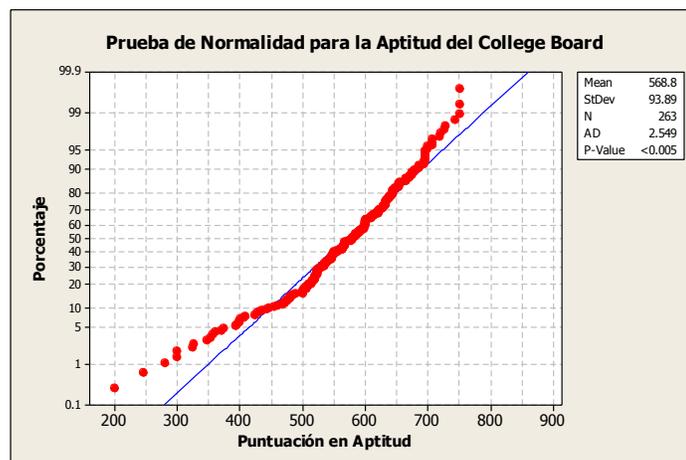
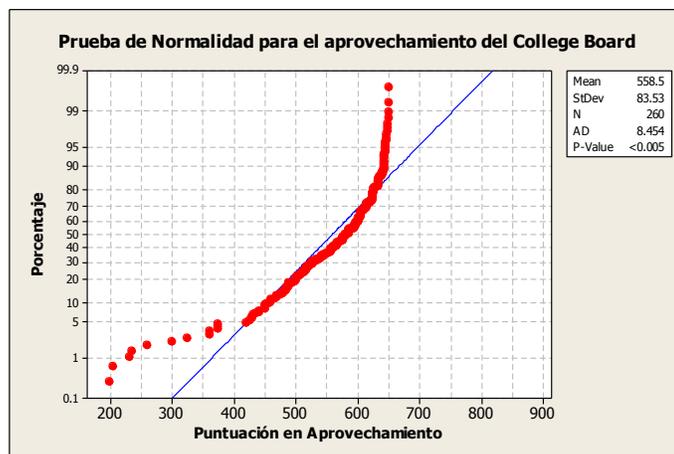
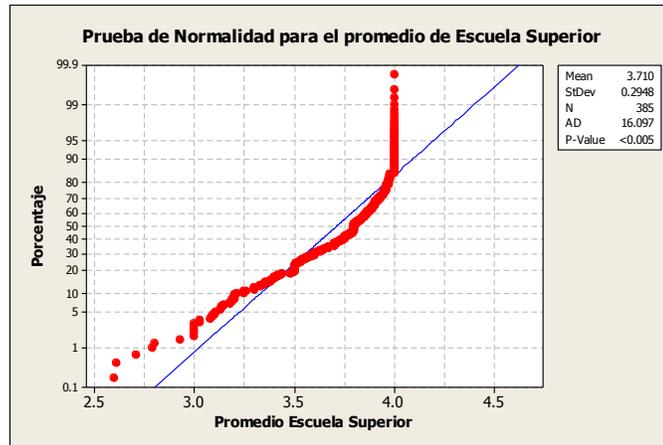
Consentimiento

He leído la información contenida en este documento y he escuchado la explicación dada por el investigador. Se me ha dado la oportunidad de hacer preguntas y estoy satisfecho con las respuestas recibidas. He recibido copia de este formulario. Mi firma en este documento certifica que asiento participar en este estudio.

Nombre del Participante _____
Firma del Participante _____ Fecha _____
Firma del Investigador _____ Fecha _____

Anexo 4

Gráficos de normalidad para las variables Promedio de Escuela Superior, Puntuación en aprovechamiento y puntuación en aptitud del College Board



Anexo 5

Gráficos de los diagnósticos del modelo de regresión

