

**USO DE TÉCNICAS MULTIVARIADAS Y MODELOS ESTADÍSTICOS
PARA EL ANÁLISIS DEL DESEMPEÑO ACADÉMICO DE LOS
ESTUDIANTES DE CÁLCULO I- UPRM**

Por

Oscar Y. Castrillón Velandia

Tesis sometida en cumplimiento parcial de los requerimientos para el grado de
MAESTRÍA EN CIENCIAS

en

MATEMÁTICAS (ESTADÍSTICA)

UNIVERSIDAD DE PUERTO RICO
RECINTO UNIVERSITARIO DE MAYAGÜEZ

Diciembre, 2007

Aprobado por:

Julio Quintana, Ph.D.
Presidente, Comité Graduado

Date

Edgardo Lorenzo, Ph.D.
Miembro, Comité Graduado

Date

Tokuji Saito, Ph.D.
Miembro, Comité Graduado

Date

Raúl Macchiavelli, Ph.D.
Representante de Estudios Graduados

Date

Julio Quintana, Ph.D.
Director del Departamento de Ciencias Matemáticas

Date

ABSTRACT

In this research a methodology based on multivariate analysis techniques is considered to characterize the students attending the course Calculus I (Fall semester, 2006-07) who participated in an exploratory study. This characterization was done with respect to their academic performance and the factors related to it. It is used the factorial analysis and the cluster analysis in the exploration of these possible profiles. Also a technique of analysis of panel data is applied to model the scores of the students in three out of four required partial examinations in the course. The results of this analysis show that some methodological characteristics of the professors and academic characteristics of the students are important predictors of their academic performance.

RESUMEN

En esta investigación se plantea una metodología basada en técnicas de análisis multivariado para caracterizar los estudiantes de Cálculo I (Primer semestre, 2006-07) que participaron en un estudio exploratorio. Esta caracterización se hizo con respecto a su rendimiento académico y los factores relacionados con él. Se utiliza el análisis factorial y el análisis de conglomerados en la exploración de los posibles perfiles.

También se aplica una técnica de análisis de datos en paneles para modelar las puntuaciones de los estudiantes en tres de los cuatro exámenes parciales requeridos en el curso. Los resultados de este análisis muestran que algunas características metodológicas de los profesores y características académicas de los estudiantes son variables explicativas importantes de su desempeño académico

Copyright © 2007

Por

Oscar Y. Castrillón Velandia

*A mis padres Maria Cristina y Oscar Iván por su constante
apoyo en mi formación profesional.*

A Claudia Liliana por estar tan cerca a pesar de la distancia.

AGRADECIMIENTOS

Al Dr. Julio Quintana por su colaboración en la realización de este de este documento.

A Santiago Velasco por su constante apoyo y sugerencias en el análisis de algunos resultados.

AL Dr. Raúl Macchiavelli por sus sugerencias metodológicas en el análisis del modelo con datos en paneles.

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

A mis amigos y amigas porque se convierten en mi familia.

Tabla de Contenido

	Página
ENGLISH ABSTRACT.....	ii
RESUMEN EN ESPAÑOL.....	iii
Lista de tablas.....	ix
Lista de figuras.....	x
1. Introducción.....	1
2. Revisión de literatura.....	3
2.1. Análisis factorial y análisis de conglomerados.....	3
2.2. Datos en panel.....	4
3. Metodología.....	15
3.1. Instrumento.....	15
3.2. Población y muestra.....	16
3.3. Procedimiento.....	18
4. Análisis Exploratorio.....	23
4.1. Introducción.....	23
4.2. Análisis Factorial.....	25
4.2.1. Validación del análisis factorial.....	25
4.2.2. Descripción de los factores.....	30
4.3. Análisis de conglomerados sobre los factores.....	33
4.3.1. Validación del análisis de conglomerados.....	33
4.3.2. Distribución de las variables en cada conglomerado.....	37
4.3.3. Descripción de los conglomerados.....	47
4.4. Selección de conglomerados por el nivel de correlación entre los factores y las puntuaciones obtenidas en el primer examen.....	57
4.4.1. Distribución de las variables en los nuevos conglomerados.....	60
4.4.2. Descripción de los nuevos conglomerados.....	66
5. Análisis de datos en paneles.....	71
5.1. Introducción.....	71
5.2. Procedimiento.....	72
5.2.1. Estimación del modelo I.....	72
5.2.2. Estimación del modelo II.....	74
5.2.3. Estimación del modelo III.....	77
5.3. Distribución de probabilidad para los estudiantes clasificados en cuatro conglomerados.....	79
6. Conclusiones.....	81
7. Limitaciones.....	87

BIBLIOGRAFÍA..... 88

Lista de Tablas

Tabla		Página
4-1	Descripción de las variables utilizadas en el análisis factorial.....	26
4-2	Porcentaje de varianza acumulado de acuerdo al número de factores.....	28
4-3	Componentes Rotados.....	29
4-4	Número de conglomerados sugeridos por los tres criterios de validación.....	34
4-5	Centroides de cada factor por conglomerado.....	35
4-6	Prueba de Kruskal- Wallis sobre la puntuación del examen.....	39
4-7	Coefficientes de correlación de Spearman.....	57
4-8	Clasificación por el Método de k-medioides.....	57
4-9	Centroides de los factores por conglomerados seleccionados.....	58
4-10	Prueba de Kruskal- Wallis sobre la puntuación del examen.....	62
5-1	Descripción del mejor subconjunto de Variables.....	73
5-2	Coefficientes estimados para el mejor subconjunto de variables utilizando el modelo I.....	73
5-8	Test Tipo I de Efectos Fijos.....	74
5-8	Test Tipo II de Efectos Fijos.....	74
5-5	Coefficientes estimados para el mejor subconjunto de variables utilizando el modelo II.....	75
5-6	Descripción del mejor subconjunto de Variables.....	74
5-7	Coefficientes estimados para el mejor subconjunto de variables utilizando el modelo III.....	78
5-8	Test Tipo III de Efectos Fijos	79

Lista de Figuras

Figura		Página
4-1	Gráfico de sedimentación.....	28
4-2	Partición óptima del conjunto de estudiantes en 8 conglomerados basándose entres criterios de validación.....	34
4-3	Descripción de los ocho conglomerados por los factores seleccionados.....	36
4-4	Proporción de estudiantes por facultad en los ocho conglomerados.....	37
4-5	Proporción de estudiantes por género en los ocho conglomerados.....	37
4-6	Proporción de estudiantes por el tipo de escuela en los ocho conglomerados.....	38
4-7	Promedio del puntaje del primer examen por conglomerado.....	38
4-8	Promedio del primer examen Vs su desviación por conglomerado.....	39
4-9	Proporción de estudiantes por respuesta ante el ítem interés por aprender matemáticas, por conglomerado.....	40
4-10	Proporción de estudiantes por respuesta ante el ítem interés por estudiar el curso, por conglomerado.....	40
4-11	Proporción de estudiantes por respuesta ante el ítem gusto por las matemáticas, por conglomerado.....	40
4-12	Proporción de estudiantes por respuesta ante el ítem capacidad para estudiar matemáticas, por conglomerado.....	40
4-13	Proporción de estudiantes por respuesta ante el ítem exigencias para estudiar matemáticas, por conglomerado.....	41
4-14	Proporción de estudiantes por respuesta ante el ítem prioridad que da al curso, por conglomerado.....	41
4-15	Proporción de estudiantes por respuesta ante el ítem curso debiera estar fuera del currículo, por conglomerado.....	41
4-16	Proporción de estudiantes por respuesta ante el ítem dificultad del examen, por conglomerado.....	41
4-17	Proporción de estudiantes por respuesta ante el ítem rapidez con la que el profesor cubrió los temas, por conglomerado.....	41
4-18	Proporción de estudiantes por respuesta ante el ítem tiempo de la clase para cubrir los temas, por conglomerado.....	41
4-19	Proporción de estudiantes por respuesta ante el ítem claridad con la que explica el profesor, por conglomerado.....	42
4-20	Proporción de estudiantes por respuesta ante el ítem opinión sobre darse de baja del curso, por conglomerado.....	42
4-21	Proporción de estudiantes por respuesta ante el ítem cantidad de ejercicios resueltos en clase, por conglomerado.....	42
4-22	Proporción de estudiantes por respuesta ante el ítem las otras materias absorbieron la mayor parte del tiempo, por conglomerado.....	42

4-23	Proporción de estudiantes por respuesta ante el ítem cantidad de temas para este examen, por conglomerado.....	42
4-24	Proporción de estudiantes por respuesta ante el ítem número de veces que ha tomado el curso, por conglomerado.....	42
4-25	Proporción de estudiantes por respuesta ante el ítem nota del curso anterior de matemáticas, por conglomerado.....	43
4-26	Proporción de estudiantes por respuesta ante el ítem repaso de las notas faltando poco tiempo, por conglomerado.....	43
4-27	Proporción de estudiantes por respuesta ante el ítem desarrollar ejercicios al terminar cada sección, por conglomerado.....	43
4-28	Proporción de estudiantes por respuesta ante el ítem visitas a la oficina del profesor, por conglomerado.....	43
4-29	Proporción de estudiantes por respuesta ante el ítem valoración del esfuerzo por obtener buena nota, por conglomerado.....	43
4-30	Proporción de estudiantes por respuesta ante el ítem estudiar en grupo, por conglomerado.....	43
4-31	Proporción de estudiantes por respuesta ante el ítem uso de exámenes de semestres anteriores, por conglomerado.....	44
4-32	Proporción de estudiantes por respuesta ante el ítem desarrollar ejercicios sin leer la teoría, por conglomerado.....	44
4-33	Proporción de estudiantes por respuesta ante el ítem tiempo para resolver el examen, por conglomerado.....	44
4-34	Proporción de estudiantes por respuesta ante el ítem practicar con ejercicios de la libreta, por conglomerado.....	44
4-35	Proporción de estudiantes por respuesta ante el ítem desempeño en matemáticas en la escuela superior, por conglomerado.....	44
4-36	Proporción de estudiantes por respuesta ante el ítem trabaja actualmente, por conglomerado.....	44
4-37	Proporción de estudiantes por respuesta ante el ítem repasar las notas de clase al terminar cada sección, por conglomerado.....	45
4-38	Proporción de estudiantes por respuesta ante el ítem tener beca , por conglomerado.....	45
4-39	Proporción de estudiantes por respuesta ante el ítem el profesor recordó su horario de oficina, por conglomerado.....	45
4-40	Proporción de estudiantes por respuesta ante el ítem tener libro de texto, por conglomerado.....	45
4-41	Proporción de estudiantes por respuesta ante el ítem ausencias, por conglomerado.....	45
4-42	Proporción de estudiantes por respuesta ante el ítem número de familiares cercanos con por lo menos un grado de bachillerato, por conglomerado.....	45
4-43	Proporción de estudiantes por respuesta ante el ítem formular preguntas en clase, por conglomerado.....	46
4-44	Proporción de estudiantes por respuesta ante el ítem cantidad de personas con las que estudia en grupo, por conglomerado.....	46

4-45	Proporción de estudiantes por respuesta ante el ítem memorizar la mayoría de ejercicios, por conglomerado.....	
4-46	Proporción de estudiantes por respuesta ante el ítem importancia del curso para su formación profesional, por conglomerado.....	46
4-47	Proporción de estudiantes por respuesta ante el ítem hacer ejercicios del prontuario faltando poco tiempo, por conglomerado.....	46
4-48	Proporción de estudiantes por respuesta ante el ítem asistencia al centro de apoyo, por conglomerado.....	46
4-49	Distribución de factores: Conglomerado 1.....	47
4-50	Distribución de factores: Conglomerado 2.....	48
4-51	Distribución de factores: Conglomerado 3.....	49
4-52	Distribución de factores: Conglomerado 4.....	51
4-53	Distribución de factores: Conglomerado 5.....	52
4-54	Distribución de factores: Conglomerado 6.....	53
4-55	Distribución de factores: Conglomerado 7.....	54
4-56	Distribución de factores: Conglomerado 8.....	55
4-57	Partición óptima del conjunto de estudiantes en cuatro conglomerados basándose en tres criterios de validación.....	58
4-58	Descripción de los cuatro conglomerados por los factores seleccionados.....	58
4-59	Distribución de los factores por conglomerado.....	59
4-60	Distribución de los ocho factores en los cuatro conglomerados	59
4-61	Proporción de estudiantes por facultad en los cuatro conglomerados.....	60
4-62	Proporción de estudiantes por género en los cuatro conglomerados.....	60
4-63	Proporción de estudiantes por el tipo de escuela en los cuatro conglomerados.....	61
4-64	Promedio del puntaje del primer examen por conglomerado.....	61
4-65	Promedio del primer examen Vs su desviación por conglomerado.....	62
4-66	Proporción de estudiantes por respuesta ante el ítem interés por aprender matemáticas, por conglomerado.....	63
4-67	Proporción de estudiantes por respuesta ante el ítem interés por estudiar el curso, por conglomerado.....	63
4-68	Proporción de estudiantes por respuesta ante el ítem gusto por las matemáticas, por conglomerado.....	63
4-69	Proporción de estudiantes por respuesta ante el ítem capacidad para estudiar matemáticas, por conglomerado.....	63
4-70	Proporción de estudiantes por respuesta ante el ítem exigencias para estudiar matemáticas, por conglomerado.....	64
4-71	Proporción de estudiantes por respuesta ante el ítem prioridad que da al curso, por conglomerado.....	64
4-72	Proporción de estudiantes por respuesta ante el ítem curso debiera estar fuera del currículo, por conglomerado.....	64
4-73	Proporción de estudiantes por respuesta ante el ítem dificultad del examen, por conglomerado.....	64
4-74	Proporción de estudiantes por respuesta ante el ítem rapidez con la que el profesor cubrió los temas, por conglomerado.....	64

4-75	Proporción de estudiantes por respuesta ante el ítem tiempo de la clase para cubrir los temas, por conglomerado.....	64
4-76	Proporción de estudiantes por respuesta ante el ítem claridad con la que explica el profesor, por conglomerado.....	65
4-77	Proporción de estudiantes por respuesta ante el ítem cantidad de ejercicios resueltos en clase, por conglomerado.....	65
4-78	Proporción de estudiantes por respuesta ante el ítem cantidad de temas para este examen, por conglomerado.....	65
4-79	Proporción de estudiantes por respuesta ante el ítem opinión sobre darse de baja del curso, por conglomerado.....	65
4-80	Proporción de estudiantes por respuesta ante el ítem las otras materias absorbieron la mayor parte del tiempo, por conglomerado.....	65
4-81	Proporción de estudiantes por respuesta ante el ítem número de veces que ha tomado el curso, por conglomerado.....	65
4-82	Proporción de estudiantes por respuesta ante el ítem nota del curso anterior de matemáticas, por conglomerado.....	66
4-83	Proporción de estudiantes por respuesta ante el ítem trabaja actualmente, por conglomerado.....	66
5-1	Distribución de probabilidad de la nota para un estudiante típico con 10 créditos en cada período.....	76
5-2	Distribución de probabilidad de la nota para un estudiante típico con 15 créditos en cada período.....	76
5-3	Distribución de probabilidad de la nota para un estudiante típico con 18 créditos en cada período.....	76
5-6	Proporción de la nota de los estudiantes para los dos tipos de análisis.....	80
5-7	Distribución de probabilidad estimadas puntuaciones del primer examen para los estudiantes clasificados en cuatro conglomerados utilizando el modelo II.....	80

1. INTRODUCCIÓN

Los cursos básicos de matemáticas en el Recinto Universitario de Mayagüez presentan tasas altas de fracasos y bajas. En particular, en los últimos 10 años el 46% de las notas obtenidas en los cursos que ofrece el departamento de matemáticas ha resultado en D, F o W (bajas) [38].

El siguiente estudio representa un aporte metodológico de análisis multivariado con el fin de explorar qué variables afectan el desempeño académico de los estudiantes de Cálculo I del Recinto Universitario de Mayagüez. Es importante destacar que los resultados y conclusiones del análisis no pueden extrapolarse a todos los individuos de población ya que la muestra usada no es probabilística aleatoria y los cuestionarios utilizados no fueron validados.

El estudio se realizó en dos etapas. En la primera se llevó a cabo un análisis exploratorio en el cual se utilizaron las técnicas: Análisis Factorial y Análisis de Conglomerados. En la segunda etapa se aplicó un modelo de análisis de datos en paneles para estudiar el efecto del cambio en el tiempo de variables cualitativas y cuantitativas que explican el rendimiento académico de los estudiantes que participaron en este estudio. Las variables que se midieron en este estudio fueron, en su mayoría, aquellas que están directamente relacionadas con el estudiante y que por su misma naturaleza son sensibles a cambios en el tiempo.

El trabajo se estructura de la siguiente manera. En el Capítulo 2 se expone la revisión bibliográfica, se mencionan algunos trabajos en ciencias sociales donde se utilizan las

técnicas de análisis factorial, análisis de conglomerados y análisis de datos en paneles. También se mencionan algunas definiciones y ventajas del análisis de datos en paneles. La metodología del estudio se describe en el Capítulo 3. En el Capítulo 4, llamado análisis exploratorio, se muestra la validación y resultados obtenidos al aplicar el análisis factorial y el análisis de conglomerados. El Capítulo 5 muestra los resultados de aplicar un análisis de datos en paneles implementando un modelo logit acumulativo con interceptos aleatorios; éstos se comparan con los obtenidos en el análisis exploratorio. Las conclusiones generales del estudio son presentadas en el Capítulo 6.

2. REVISIÓN DE LITERATURA

2.1 Análisis Factorial y Análisis de conglomerados

El análisis factorial es una técnica del análisis multivariado que permite reducir y clasificar un conjunto de variables explicativas reemplazándolas por un número menor de variables llamadas factores [24]. Así, un factor es una variable o conjunto de ellas no directamente observables pero que puede ser obtenido de las variables explicativas. Por ejemplo, del conjunto de variables explicativas del desempeño académico de un estudiante, algunas pueden estar midiendo en parte el mismo concepto subyacente, como la actitud hacia el estudio o el interés por un curso en particular.

Para realizar un análisis factorial se deben tener en cuenta los siguientes pasos [36]: Formulación del problema, análisis de la matriz de correlación, extracción de factores, determinación del número de factores, rotación de factores, interpretación de los factores y el cálculo de las puntuaciones factoriales. En el capítulo de *Análisis Exploratorio* se exponen con detalle las etapas mencionadas anteriormente.

El objetivo del análisis de conglomerados (o *cluster analysis*) es la clasificación (no supervisada) de individuos (u objetos) en clases disjuntas a partir de sus diferencias o similitudes, las cuales son caracterizadas en términos de distancias [28]. Así, los individuos que pertenecen a un mismo grupo tendrán características semejantes y los individuos de

grupos distintos tendrán características diferentes. Estos grupos se conocen en la literatura como conglomerados o “*clusters*” en inglés.

Algunos antecedentes de estudios realizados en el campo de las ciencias sociales muestran la utilidad del análisis factorial. Por ejemplo, Cruz [14] utiliza el análisis factorial para clasificar de mayor a menor el desarrollo de los departamentos de Perú, combinando los puntajes factoriales con análisis de conglomerados. Rúa [34] utiliza el análisis factorial para constatar la estructura subyacente dentro del conjunto de todas las asignaturas tomadas por los estudiantes de Ciencias Empresariales Internacionales impartida en la Universidad Pontificia Comillas de Madrid. Luego mediante un análisis de conglomerados, encontró ocho tipologías o patrones de comportamiento en relación con su rendimiento académico.

González [22] utiliza el análisis factorial para determinar las dimensiones y elementos que pueden conforma un modelo de indicadores de calidad universitaria a partir de la opinión de una muestra representativa de estudiantes.

El Instituto Nacional de Estadística e Informática de Perú utiliza el análisis factorial como parte del proceso para calcular el índice de desarrollo educativo [26].

2.2 Datos en Panel

El análisis de datos en paneles es una forma de datos longitudinales cada vez más utilizada entre los investigadores de las ciencias sociales y las ciencias del comportamiento [18]. Con frecuencia, la psicología, la sociología y las ciencias de la salud utilizan el análisis de datos en paneles para estudiar características de grupos de personas a las cuales se les lleva un seguimiento por un período de tiempo.

Se dice que un conjunto de datos es de panel cuando se tienen observaciones de series temporales sobre una muestra de unidades individuales [6]. Es decir, en este análisis se observa el mismo conjunto de individuos en distintos momentos en el tiempo. Formalmente, para una variable y_{it} se tiene $i = 1, \dots, N$ observaciones de corte transversal; éstas pueden ser países, ciudades, empresas, artículos, grupo de personas o incluso individuos y $t = 1, \dots, T$ observaciones de series temporales; caracterizando en cada período particular las unidades de secciones cruzadas.

Con observaciones repetidas de muchas secciones cruzadas, el análisis de panel permite estudiar la dinámica de cambio en períodos cortos de tiempo. La combinación de las series de tiempo con secciones cruzadas puede aumentar la calidad y cantidad de datos, cosa que sería imposible usando sólo una de estas dimensiones [18].

2.2.1 Modelo estadístico para datos de panel

Un modelo de regresión básico para datos de corte transversal viene dado por la expresión $y_i = \alpha + x_i' \beta + e_i$ donde $i = 1, 2, \dots, N$. En este modelo se asume que las unidades en estudio son homogéneas, por tanto las variaciones en la variable respuesta se deben a las diferencias de las características medidas para cada individuo y el azar. Cuando cada unidad transversal se observa durante T_i períodos, el modelo resultante será $y_{it} = x_{it}' \beta + e_{it}$, $i = 1, 2, \dots, N$, $t = 1, 2, \dots, T_i$. Este será el prototipo de un modelo con datos de panel [25]. Si $T_i = T$, es decir, si todas las unidades transversales se observan el mismo número de veces, se dirá que el panel de datos está completo o equilibrado “balanced”; en otro caso se dirá que el panel es desbalanceado “unbalanced”.

Para controlar la presencia de efectos inobservables individuales se supone que $e_{it} = \alpha_i + v_{it}$, donde α_i recoge la heterogeneidad transversal persistente no observada y v_{it} representa el término de perturbación clásico. Si se supone que el efecto α_i es un parámetro fijo entonces el modelo se considera de *efectos fijos*. Por el contrario, si el efecto α_i se supone aleatorio entonces el modelo se considera de *efectos aleatorios* [6].

Cuando también existen efectos temporales persistentes y no observados, se considera la descomposición $e_{it} = \alpha_i + \delta_t + v_{it}$, donde ahora δ_t representa dichos efectos temporales inobservables (específicos de cada período y no incluidos explícitamente entre los regresores).

2.2.2 Estimación del Modelo

El contexto básico para el análisis de datos en paneles es un modelo de regresión de la forma

$$y_{it} = \alpha_i + \beta' x_{it} + v_{it}, \quad \begin{array}{l} i = 1, \dots, N, \\ t = 1, \dots, T, \end{array} \quad (2-1)$$

Hay \mathbf{K} regresores en x_{it} sin incluir el término constante. El *efecto individual* es α_i , que se considera constante a lo largo del tiempo t y específico para la unidad de sección cruzada individual. Si las α_i son iguales para todas las unidades, mínimos cuadrados ordinarios proporcionan estimaciones consistentes y eficientes de α y β . Existen dos marcos básicos utilizados para generalizar este modelo. El enfoque de efectos fijos, considera α_i como un término constante específico de grupo en el modelo de regresión. El enfoque de efectos aleatorios especifica que α_i es un error específico de grupo, similar a e_{it} , excepto que para

cada grupo hay una única extracción muestral, que aparece en la regresión de forma idéntica en cada período.

Modelo de efectos fijos

Este modelo supone que las diferencias entre los promedios de las unidades pueden captarse mediante diferencias en el término constante. Por tanto en (2-1), cada α_i es un parámetro desconocido que debe ser estimado. Sean y_i y X_i las T observaciones de la i -ésima unidad, sea e_i el vector $T \times 1$ de errores asociado y $\mathbf{1}$ un vector $T \times 1$ de unos. Así (2-1) puede escribirse como

$$y_i = \mathbf{1}\alpha_i + X_i\beta + e_i$$

Reagrupando se tiene

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ & & \vdots & \\ 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} + \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \beta + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

o

$$y = [d_1 \ d_2 \ \cdots \ d_n \ X] \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + e$$

donde d_i es un vector de dimensión $nT \times 1$ que indica la i -ésima unidad. Sea la matriz $nT \times nD = [d_1 \ d_2 \ \cdots \ d_n]$. Entonces reuniendo las nT filas se obtiene

$$y = D\alpha + X\beta + e$$

Este modelo se conoce como modelo mínimo cuadrático con variables ficticias (MCVF). Si n es suficientemente pequeño, el modelo puede estimarse por mínimos cuadrados ordinarios, con K regresores en X y n columnas en D , como una regresión múltiple con $n + K$

parámetros. De forma equivalente, para obtener estimaciones del vector β puede realizarse la regresión de $y_{it} - \bar{y}_i$ sobre el vector de regresores $x_{it} - \bar{x}_i$ y estimar los α_i a partir de las relaciones $\hat{\alpha}_i = y_{it} - \bar{x}_i' \hat{\beta}$. Este estimador se denomina intra-grupos (within).

Efectos fijos de grupo y de tiempo

El enfoque de mínimos cuadrados con variables artificiales se puede extender, también, para que incluya un efecto temporal específico. Este puede formularse como

$$y_{it} = \alpha_i + \gamma_t + \beta' x_{it} + e_{it}$$

Este modelo se obtiene del anterior añadiendo $T - 1$ variables artificiales adicionales. Uno de los efectos temporales debe dejarse para evitar la colinealidad. Para estimar los parámetros se utiliza la transformación “within” similar al caso anterior; así, para estimar el vector β habría que realizar la regresión de $y_{it} - \bar{y}_i - \bar{y}_t + \bar{y}_{..}$ sobre el vector de regresores $x_{it} - \bar{x}_i - \bar{x}_t + \bar{x}_{..}$ posteriormente estimar los α_i a partir de las relaciones $\hat{\alpha}_i = (\bar{y}_i - \bar{y}_{..}) - (\bar{x}_i - \bar{x}_{..})' \hat{\beta}$ y los $\hat{\gamma}_t = (\bar{y}_t - \bar{y}_{..}) - (\bar{x}_t - \bar{x}_{..})' \hat{\beta}$.

Observaciones de los modelos anteriores

- El número de grados de libertad en el modelo de efectos fijos es NT-N-T-k o NT-N-k, según que se incluyan o no efectos temporales aparte de los individuales. Existe, por tanto, una importante pérdida de grados de libertad cuando N es elevado.

- El estimador “within” construido no permite estimar el efecto de variables invariantes en el tiempo ($x_{it} = x_i \forall t$) tales como el sexo, raza, religión, etc. ya que dichas variables desaparecen con la transformación en desviaciones respecto a la media.
- Este modelo es apropiado cuando el análisis se centra sobre un conjunto específico de individuos de N unidades, la inferencia que se haga será específica a dicho conjunto.

Modelo de efectos aleatorios

En lugar de suponer la heterogeneidad inobservable fija, se supone aleatoria. Considérese una reformulación del modelo

$$y_{it} = \alpha + \beta' x_{it} + u_i + e_{it}$$

donde hay K regresores además del término constante. El componente u_i es el error aleatorio que caracteriza a la i -ésima observación y es constante a lo largo del tiempo. Éstos se pueden interpretar como el conjunto de factores no incluidos en la regresión que son específicos a cada individuo. Se supone que $u_i \sim IID(0, \sigma_u^2)$, $e_{it} \sim IID(0, \sigma_e^2)$, siendo ambas variables aleatorias independientes entre sí, y se supone de nuevo que las variables x_{it} son independientes de las variables u_i y e_{it} .

El modelo de efectos aleatorios es apropiado cuando las N unidades transversales son una muestra aleatoria de una población mayor; en este caso se espera que el efecto individual se caracterice mejor por una variable aleatoria y las inferencias que se realicen serán respecto a la población y no respecto a la muestra aleatoria extraída.

Bajo la hipótesis de efectos aleatorios, el término de error compuesto $w_{it} = e_{it} + u_i$ tendrá la estructura

$$E[w_{it}w_{js}] = \begin{cases} \alpha_u^2 + \alpha_e^2 & i = j, t = s \\ \alpha_u^2 & i = j, t \neq s \\ 0 & i \neq j \end{cases}$$

por tanto, se incumple la hipótesis de que $E[ww'] = \sigma^2 I$. Entonces el estimador por mínimos cuadrados no será el mejor estimador lineal insesgado (MELI) y la matriz de covarianza no tendrá la forma $\sigma^2(X'X)^{-1}$.

Como antes, considerando la formulación del modelo en bloques de T observaciones para cada unidad i , $y_i = X_i\beta + e_i$, sea $\Omega = E[w_iw_j]$ donde $w_i = [w_{i1}, w_{i2}, \dots, w_{iT}]$, Entonces

$$\Omega = \begin{bmatrix} \sigma_e^2 + \sigma_u^2 & \sigma_u^2 & \sigma_u^2 & \dots & \sigma_u^2 \\ \sigma_u^2 & \sigma_e^2 + \sigma_u^2 & \sigma_u^2 & \dots & \sigma_u^2 \\ & \vdots & & & \vdots \\ \sigma_u^2 & \sigma_u^2 & \sigma_u^2 & \dots & \sigma_e^2 + \sigma_u^2 \end{bmatrix} = \sigma_e^2 I + \sigma_u^2 J,$$

donde J es una matriz de unos. Como las observaciones i y j son independientes, la matriz de varianzas y covarianzas de los errores para las nT observaciones es

$$V = \begin{bmatrix} \Omega & 0 & \dots & 0 \\ 0 & \Omega & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \Omega \end{bmatrix}$$

En este caso se debe utilizar el estimador por Mínimos Cuadrados Generalizados (MCG) para tener en cuenta la violación de la hipótesis de esfericidad. Para este caso, el estimador (MCG) se obtiene mediante la regresión de la variable $y_{it} - \lambda_i \bar{y}_i$ sobre el vector de regresores $x_{it} - \lambda_i \bar{x}_i$ donde

$$\lambda_i = 1 - \sqrt{\frac{\sigma_e^2}{T_i \sigma_u^2 + \sigma_e^2}}$$

En la práctica se utiliza el estimador por Mínimos Cuadrados Generalizados Factibles (MCGF), el cual usa estimaciones consistentes del parámetro λ_i . Como estimación de la varianza se utiliza σ_e^2 la varianza residual obtenida del estimador “within” multiplicada por el factor (NT-K)/(NTN-K) y la varianza σ_u^2 es estimada por $\hat{\sigma}_e^2 - \frac{\hat{\sigma}_b^2}{T}$ donde $\hat{\sigma}_b^2$ es la varianza residual del estimador “between” (este estimador se obtiene al realizar la regresión por MCO de \bar{y}_i sobre \bar{x}_i).

2.3 Modelos Lineales Generalizados

Los modelos lineales generalizados (MLG) son una extensión del Modelo Lineal General a una familia más general; fueron propuestos por Nelder y Wedderburn. Esta nueva familia unifica tanto a modelos con variables de respuesta categórica como numérica; y considera distribuciones como labinomial, poisson, hipergeométrica, binomial negativa y otras, ya no únicamente la distribución Normal. En estos modelos, la componente aleatoria no necesariamente es homocedástica, es decir no se requiere de un supuesto de homogeneidad de varianzas.

Un modelo lineal generalizado consiste de un predictor lineal $\eta = x' \beta$ y de una función g (llamada función de enlace), monótona y diferenciable, que describe la relación entre el valor esperado de la variable respuesta y el predictor lineal; $g(\mu_i) = x'_i \beta$. La variable respuesta Y es independiente y tiene una distribución de probabilidad en la familia exponencial la cual, de manera general, puede escribirse como

$$f(y) = \exp\left\{\frac{y\theta - \psi(\theta)}{\phi} + c(y, f(\phi))\right\}$$

donde $\psi(\cdot)$ y $c(\cdot)$ son funciones conocidas, θ es un parámetro de localización y ϕ es un parámetro de dispersión. La varianza de la variable respuesta depende de la media μ a través de una función de varianza V , de tal manera que: $\text{var}(y_i) = \phi V(\mu_i) / w_i$ donde w_i es un peso conocido para cada observación.

2.3.1 Modelo logit acumulativo multinomial

Algunas variables multinomiales representan por su propia naturaleza un orden; como las notas de un curso. Los modelos acumulativos tienen en cuenta la ordenación de las categorías, utilizan las probabilidades hasta una determinada categoría, $P(Y \leq j)$, en lugar de las probabilidades de cada categoría. En términos generales el modelo estaría dado por

$$P(Y \leq j|x) = F(\alpha_j - x' \beta)$$

donde α_j determina los límites de cada categoría de la variable respuesta. β es el vector de coeficientes. Distintas elecciones de F dan lugar a modelos acumulativos específicos. Así, para el modelo logístico acumulativo F es la distribución logística $F(x) = \frac{\exp(x)}{1 + \exp(x)}$. Este

modelo es un MLG con función de enlace $g = F^{-1}$.

2.3.2 Modelo “Logit” multinomial con Efectos Aleatorios

Las observaciones repetidas de un individuo tienden a ser similares, en el caso de variables categóricas, significa que los individuos tienden a estar en las mismas categorías de una observación a la siguiente. La similaridad se modela a través de un término de heterogeneidad común. Considérese el modelo

$$V_{it,j} = \alpha_{ij} + x'_{it,j} \beta$$

α_{ij} representa la heterogeneidad que es específica a cada individuo pero constante en el tiempo. Así, la probabilidad de que el i –ésimo individuo esté en la clase j en el tiempo t es

$$\pi_{it,j}(\alpha_i) = \frac{\exp(V_{it,j})}{\sum_{k=1}^c \exp(V_{it,k})} = \frac{\exp(\alpha_{ij} + x'_{it,j} \beta)}{\sum_{k=1}^c \exp(\alpha_{ik} + x'_{it,k} \beta)}, \quad j = 1, 2, \dots, c$$

donde $\alpha_i = (\alpha_{i1}, \dots, \alpha_{ic})'$.

Se observa que, por la forma de la ecuación, que el término de heterogeneidad que es constante sobre la alternativa j no es afectado por la probabilidad condicional. La estimación del modelo por máxima verosimilitud para el i –ésimo es

$$L(y_i | \alpha_i) = \prod_{t=1}^{T_i} \prod_{j=1}^c (\pi_{it,j}(\alpha_i))^{y_{it,j}} = \prod \left\{ \frac{\exp(\sum_{j=1}^c y_{it} (\alpha_{ij} + x'_{it,j} \beta))}{\sum_{j=1}^c y_{it} (\alpha_{ij} + x'_{it,j} \beta)} \right\}$$

Se supone que $\{\alpha_i\}$ está idéntica e independientemente distribuido con función de distribución G_a la cual generalmente es una normal para cada observación.

2.3.3 Quasi-Verosimilitud

Un modelo lineal generalizado supone una cierta distribución para la variable respuesta. Sin embargo, existen ocasiones en las cuales resulta confuso especificar la distribución exacta de esta variable. La quasi-verosimilitud permite hacer inferencias con cierto grado de robustez a partir de dos condiciones importantes. Primero, no se necesita imponer un supuesto distribucional y, segundo, sólo es preciso especificar la relación entre la media y la varianza.

Para un sujeto $i, i = 1, \dots, n$, sea y_i el resultado sobre el vector Y con $\mu_i = E(Y_i)$ y función de varianza $v(\mu_i)$, sea x_{ij} el valor de la j -ésima variable exploratoria. Para una función de enlace g , el predictor lineal es $\eta_i = g(\mu_i) = \sum_j \beta_j x_{ij} = \mathbf{x}'_i \boldsymbol{\beta}$. El estimador quasi-verosimilitud del parámetro $\boldsymbol{\beta}$ es la solución de las ecuaciones parciales

$$u(\boldsymbol{\beta}) = \sum_i \left(\frac{\partial \mu_i}{\partial \boldsymbol{\beta}} \right)' v(\mu_i)^{-1} (y_i - \mu_i) = 0$$

donde $\mu_i = g^{-1}(\mathbf{x}'_i \boldsymbol{\beta})$.

Para un individuo i , sea $y_i = (y_{i1}, \dots, y_{iT_i})'$ el vector de respuestas y $\boldsymbol{\mu}_i = (\mu_{i1}, \dots, \mu_{iT_i})'$, donde $\mu_{it} = E(Y_{it})$. Sea \mathbf{x}'_{it} un vector de dimensión $p \times 1$ de variables exploratorias para y_{it} . El predictor lineal del modelo es $\eta_{it} = g(\mu_{it}) = \mathbf{x}'_{it} \boldsymbol{\beta}$ para la función de enlace g . Suponiendo que y_{it} tiene una distribución de probabilidad en la familia exponencial de la forma

$$f(y_{it}; \boldsymbol{\theta}_{it}, \phi) = \exp \left\{ \frac{y_{it} \boldsymbol{\theta}_{it} - \psi(\boldsymbol{\theta}_{it})}{\phi + c(y_{it}, \phi)} \right\}$$

donde ϕ es conocida y $\boldsymbol{\theta}_{it}$ es un parámetro natural. Se deduce que $\mu_{it} = E(Y_{it}) = b'(\boldsymbol{\theta}_{it})$, $v(\mu_{it}) = \text{var}(Y_{it}) = b''(\boldsymbol{\theta}_{it})\phi$.

El método también supone una matriz de correlación $\mathbf{R}(\boldsymbol{\alpha})$ para Y_i , que depende del parámetro $\boldsymbol{\alpha}$, $\text{cor}(Y_{it}, Y_{is}) = \boldsymbol{\alpha}$ para cada par in Y_i . Sea $\mathbf{b}_i(\boldsymbol{\theta}) = (b(\boldsymbol{\theta}_{i1}), \dots, b(\boldsymbol{\theta}_{iT_i}))$ y sea \mathbf{B}_i una matriz diagonal con elementos $b''_i(\boldsymbol{\theta})$. Así la matriz de covarianza para Y_i es

$$\mathbf{V}_i = \mathbf{B}_i^{1/2} \mathbf{R}(\boldsymbol{\alpha}) \mathbf{B}_i^{1/2} \phi$$

Sea Δ_i una matriz diagonal con elementos $\partial\theta_{it}/\partial\eta_{it}$ para $t=1,\dots,T_i$ y sea $\mathbf{D}_i = \partial\mu_i/\partial\beta = \mathbf{B}_i\Delta_iX_i$, una matriz de dimensión $T_i \times p$ con elementos típicos expresando $\partial\mu_{it}/\partial\beta_j$ en la forma $(\partial\mu_{it}/\partial\theta_{it})(\partial\theta_{it}/\partial\eta_{it})(\partial\eta_{it}/\partial\beta_j)$. Así, como en el modelo anterior, para un individuo con una sola respuesta, el estimador del parámetro β es la solución de la ecuación

$$\sum_{i=1}^n D_i'V^{-1}[y_i - \mu_i(\beta)] = 0$$

y análogo al modelo anterior, el estimador del parámetro β es la solución de la ecuación

$$u(\beta) = \sum_i \left(\frac{\partial\mu_i}{\partial\beta} \right)' v(\mu_i)^{-1} (y_i - \mu_i) = 0$$

2.4 Ventajas y limitaciones de usar datos en paneles.

Hsiao [25] lista las siguientes ventajas de los datos en panel.

- a) Control de la heterogeneidad individual: los datos transversales y temporales no son capaces, por sí solos, de controlar la heterogeneidad inherente en el comportamiento de los individuos, empresas, regiones o países, corriéndose el riesgo de obtener estimaciones sesgadas cuando se utilizan datos de un tipo o de otro. Sin embargo a través del uso de los datos de panel puede controlarse estos efectos específicos, transversales o temporales, sean observables o no.
- b) Proporciona datos con mayor cantidad de información, con mayor grado de control de la variabilidad y con menor nivel de colinealidad entre los regresores; y también

- aumenta el número de grados de libertad y, por tanto, da lugar a una mayor eficiencia en las estimaciones.
- c) Ayudan a identificar y medir efectos que no son detectables con datos puros de corte transversal o de series temporales.
 - d) Son un medio adecuado para estudiar procesos dinámicos de ajuste ya que a partir de ellos se pueden analizar los cambios en el tiempo de las distribuciones transversales.

En términos generales Baltagi [7] menciona las siguientes limitaciones:

- a) Los procesos para la obtención y el procesamiento de la información estadística sobre las unidades individuales de estudio, cuando ésta se obtiene por medio de encuestas, entrevistas o utilizando algún otro medio de levantamiento de los datos. Ejemplos de este tipo de limitaciones son [29]: cobertura de la población de interés, porcentajes de no respuesta, preguntas confusas, distorsión deliberada de las respuestas, aunque estos problemas se presentan no sólo en análisis de datos por paneles sino en cualquier proceso donde haya encuestas.
- b) Corta dimensión temporal de los paneles. Los paneles típicos abarcan un corto intervalo temporal por cada individuo. Esto significa que para poder tener un volumen importante de puntos es necesario aumentar la muestra en el número de individuos. Esto no siempre es posible, ya que aumenta las probabilidades de retiro del panel.

En economía, Albarrán y González [3] mediante una estimación econométrica de datos en paneles, examinaron la relación entre el nivel de ingresos y el gasto en seguros de vida y de otro tipo en los 15 países de la Unión Europea, Estados Unidos y Japón. Esta técnica también se usa para estudiar la relación entre empresas y salarios de personas a través del tiempo. En las ciencias políticas, esta clase de análisis se usa para estudiar el comportamiento de entidades u organizaciones a través del tiempo. En un estudio para determinar si la implantación de una política educativa basada en un desayuno escolar aumentaba la tasa de retención en el sistema educativo en Cochabamba – Bolivia, se utilizó un modelo de efectos fijos de datos en panel [32].

Rockoff J. [33], utiliza el análisis de datos en paneles para determinar si la calidad de los maestros tiene algún impacto sobre los logros de los estudiantes en la escuela elemental.

En otro campo, Frees y Millar [19] utilizan los datos en paneles para predecir la venta de boletos de lotería en el estado de Wisconsin en Estados Unidos.

3. METODOLOGÍA

En este capítulo se describe de manera detallada la forma en que se realizó el estudio. Se mencionarán las características de los instrumentos que se utilizaron para la recolección de la información, los sujetos de estudio, dificultades en el tamaño de la muestra, los procedimientos y técnicas que formaron parte del análisis.

3.1 Instrumento

Se diseñaron tres cuestionarios, uno por cada período previo a cada examen parcial, que permitieran conocer algunas características y hábitos de estudio de los estudiantes matriculados en el curso de Cálculo I del Recinto Universitario de Mayagüez-UPRM. Cada cuestionario se entregó durante los tres días siguientes a la presentación de los tres primeros exámenes parciales del curso. El primer cuestionario (anexo 1) se dividió en dos partes, la primera comprendía preguntas sociodemográficas como departamento y facultad a la que pertenece, estado civil, número de hijos, etc. La segunda parte, un poco más extensa, mediante diferentes escalas tipo Likert¹ indagaba sobre el interés y gusto por el estudio de las matemáticas, tiempo y actividades realizadas para la preparación de examen y opinión sobre algunas características del profesor y del examen. Los otros dos cuestionarios sólo contenían preguntas referentes a los hábitos en la preparación del examen, opinión sobre las

¹ . La escala de Likert es una escala ordinal que utiliza ítems o frases para medir fenómenos sociales y sobre las que una persona da su opinión. Cada frase es presentada seguida de una escala de estimación (rating scale) que consiste en una graduación que va desde "totalmente de acuerdo" hasta "totalmente en desacuerdo" o variaciones de ésta, incluyendo grados intermedios, con respecto a la afirmación. [17], [4] y [11].

características del profesor y del examen; también contenían preguntas sobre la situación académica del estudiante cuya respuesta pudiese cambiar en el período previo a cada examen. La puntuación correspondiente de cada examen parcial fue suministrada directamente por los profesores.

3.2 Población y Muestra

La población objetivo estaba constituida por 520 estudiantes matriculados en el curso de Cálculo I del primer semestre académico del año 2007, distribuidos en 13 secciones, las cuales estaban asignadas a 7 profesores y a dos instructores (estudiantes del programa graduado de maestría en matemáticas). Este curso es evaluado periódicamente a través de exámenes departamentales, implicando que todas las secciones trabajan los mismos temas en el período previo al examen. Uno de los profesores a cargo de una sección trabaja una metodología y temática diferente, razón por la cual no se incluyó en el estudio.

En estudios longitudinales es difícil contar con personas que estén dispuestas a colaborar con este tipo de investigaciones pues se requiere llevar un registro de los individuos en estudio a través del tiempo. Para hacer esto, se solicitó a los estudiantes que identificaran cada cuestionario con su número de estudiante. Cada uno de ellos recibió una carta que explicaba la necesidad de identificar los cuestionarios; esta carta era firmada y devuelta. Aunque ningún estudiante se negó a identificar el cuestionario se observó, en el proceso de entrada de datos, que muchos de los cuestionarios no estaban identificados y que otros no habían sido devueltos, lo cual llevó a una disminución significativa de los datos.

Para el primer período, correspondiente a los días previos al examen I, se recogieron 265 formularios, para el segundo período, días previos al segundo examen, se obtuvieron 247

cuestionarios de los cuales 183 estudiantes dieron respuesta al primero y para el tercer período se obtuvieron 226, de estos, 166 respondieron el primero. Con la intersección de los tres cuestionarios se obtuvieron 125 estudiantes, es decir, 125 estudiantes dieron respuesta a los tres cuestionarios. Dado lo anterior, el estudio se realizó con una muestra no probabilística de 265 estudiantes voluntarios para el primer período (toma de datos correspondientes al primer examen parcial), 247 para el segundo período y 226 para el tercero.

Como se dijo anteriormente, con los cuestionarios II y III sólo se midieron características que, a juicio personal, puede cambiar con el tiempo así que éstos no contienen información sobre el tipo de escuela, facultad a la que pertenece, si tiene o no beca y algunas preguntas referentes al gusto de las matemáticas. Por esta razón en el análisis exploratorio sólo se utilizaron los datos recopilados en el primer cuestionario. La muestra definitiva para este análisis quedó constituida por 110 mujeres y 155 hombres. Ochenta y nueve (89) estudiantes pertenecen a la facultad de Artes y Ciencias, 172 a la facultad de Ingeniería (se incluyeron en este grupo los estudiantes de las facultades de Administración de Empresas y Ciencias Agrícolas) y 4 estudiantes son de la Escuela Superior. 157 estudiantes provienen de escuela pública y 105 de escuela privada, tres estudiantes no suministraron esta información.

Para el análisis de datos por paneles se utilizaron los datos de los 125 estudiantes que respondieron los tres cuestionarios, uno por cada período. Sin embargo, no se tuvo en cuenta los datos de un estudiante de Escuela Superior que tomó el curso, ya que sus características son muy diferentes a las de los demás estudiantes. Así, la muestra para este análisis está constituida por los 124, el 42% son mujeres, el 31% pertenece a la facultad de Artes y

Ciencias y el 69% restante a las facultades de Ingeniería, Empresas y Ciencias Agrícolas. El 71% tiene beca por parte del Recinto, y el 68.55% toma el curso por primera vez.

3.3 Procedimiento

El análisis se realizó en dos etapas. En la primera se llevó a cabo un análisis exploratorio; en el cual se utilizaron las técnicas: Análisis Factorial y Análisis de Conglomerados. En la segunda etapa modelamos, mediante un análisis de datos en paneles, la nota obtenida por los estudiantes en 3 de los 4 exámenes parciales requeridos en el curso.

Se aplicó un análisis factorial a los datos suministrados por los estudiantes que entregaron el primer cuestionario con el propósito de disminuir el número de variables iniciales a un conjunto de variables subyacentes o factores que tengan la misma potencia explicativa. Como resultado del análisis se obtuvieron 8 factores los cuales hacían referencia a la actitud que el estudiante tiene hacia el estudio de las matemáticas, a las características del profesor y del examen, al rezago académico, al esfuerzo y tiempo dedicado en la preparación del examen, al número de personas con las que estudia, al trabajo realizado durante las clases, a algunas características sociales y a la actitud que el estudiante tiene con respecto al curso de Cálculo I.

Con el fin de encontrar las posibles tipologías, se aplicó un análisis de conglomerados a los 265 estudiantes que respondieron el primer cuestionario, se utilizaron como variables descriptivas los 8 factores obtenidos en el análisis anterior. Los criterios de validación [9] aplicados al análisis de conglomerados determinaron que el número óptimo de topologías es

ocho. La descripción de los conglomerados o grupos encontrados se realizó a partir de la representación gráfica de los centroides (promedio de los factores en cada conglomerado) en sectores circulares y de la distribución porcentual de todas las variables en cada uno de los conglomerados.

Con el propósito de “refinar” la clasificación obtenida se escogió los factores que estaban más correlacionados con la puntuación del primer examen y se procedió a efectuar un segundo análisis de conglomerados. Los factores seleccionados, utilizando el coeficiente rho de Spearman [16], fueron la “actitud hacia las matemáticas”, las “características del profesor y del examen” y el “rezago académico”. Como resultado del análisis se seleccionaron 4 conglomerados.

En la segunda parte del estudio se aplicó un análisis de datos por paneles con los datos de los estudiantes que entregaron los tres cuestionarios. Por razones de tiempo y conocimiento del procedimiento GLIMMIX incorporado en el software estadístico SAS, no fue posible explorar los resultados con la información de todos los cuestionarios en cada período.

Debido a que la gran mayoría de variables predictoras o explicativas son categóricas, se categorizó la variables respuesta, puntuación del examen, en 5 categorías; así, la variable toma el valor 0 si la nota está en el intervalo de 0 a 59, 1 de 60 a 69, 2 de 70 a 79, 3 de 80 a 89 y 4 si es igual a 100 o mayor.

Se utilizó un modelo “*logit*” acumulativo multinomial con interceptos aleatorios donde se tuvo en cuenta el factor tiempo.

Para encontrar el mejor subconjunto de variables explicativas se utilizó el método de selección “forward” [1].

4. ANÁLISIS EXPLORATORIO

4.1 Introducción

Se plantea una metodología basada en técnicas de análisis multivariado que permita caracterizar los estudiantes de Cálculo I que participaron en la investigación. En la exploración de los posibles perfiles se utilizó el análisis factorial y el análisis de conglomerados.

La técnica de análisis factorial permitió reducir la dimensión original del conjunto de variables, de modo que dicho conjunto pueda sustituirse por un número menor de variables, denominadas factores, con los que se logra una explicación prácticamente similar al logrado por el conjunto de variables inicial o de partida. Es decir, se consigue transformar un conjunto de variables interdependientes en un conjunto de variables no correlacionadas (Factores). También podremos detectar la estructura subyacente en el conjunto de datos, donde el modelo matemático de esta técnica es similar al de regresión múltiple, en el que cada variable aparece como combinación lineal de una serie de factores no observables directamente.

El análisis de conglomerados establece la partición de un conjunto de observaciones, en este estudio son los estudiantes, en grupos, de modo que los estudiantes que pertenezcan a un mismo grupo tendrán características similares y estudiantes de grupos distintos tendrán características diferentes. Se trata de que haya homogeneidad dentro de cada grupo y heterogeneidad entre los grupos [27].

Antes de aplicar el análisis factorial se determinó, por tres métodos analíticos diferentes, el nivel de correlación que existe en el conjunto de variables y así decidir la conveniencia de su implantación. Este análisis permitió trabajar con un número reducido de variables, lo cual facilitó el proceso de clasificación².

Con los factores (grupos de variables más correlacionadas) obtenidos en el análisis factorial se procedió a realizar el análisis de conglomerados con el fin de obtener algunas clases o tipologías subyacentes. Obtenidos los grupos, se describieron los conglomerados a partir de los factores que tuvieron más influencia en la clasificación. En la formación de los grupos no se tuvo en cuenta la nota promedio de los estudiantes, facultad a la que pertenece, su género ni el tipo de escuela de donde proviene, pues la caracterización se hizo principalmente sobre estas variables permitiendo determinar si los hábitos, intereses y percepciones que tienen los estudiantes acerca de las matemáticas están relacionados con su desempeño académico.

Las variables que se utilizaron en el análisis fueron recopiladas mediante un cuestionario que se aplicó en los tres días siguientes a la presentación del primer examen. La mayoría de las variables fueron medidas en diferentes escalas tipo Likert. El resto de variables, exceptuando las características sociodemográficas, fue medido a través de escalas de respuesta dicotómicas, Sí/No. Dada las características de la muestra, no fue posible tener todos los estudiantes que estaban tomando el curso de Cálculo I en la muestra de estudio. La muestra que utilizamos en este estudio está formada por 265 estudiantes.

² Uno de los mayores problemas que se presenta en el análisis estadístico multivariado es la formación de conglomerados o grupos debido a la gran dimensionalidad de los datos [5].

4.2 Análisis Factorial

Queremos clasificar los estudiantes teniendo en cuenta la información contenida en las siguientes variables cuya breve descripción podemos ver en la Tabla 4-1. Llevar a cabo una clasificación con estas variables podría dificultar la interpretación de los grupos formados por el análisis de conglomerados (“clusters”). Por esta razón se decidió disminuir la dimensión aplicando un análisis factorial, evitando de esta manera la información redundante.

4.2.1 Validación del análisis factorial

El análisis factorial tiene sentido si existe una suficiente correlación entre las variables. Esto nos indica la presencia de factores comunes. Para comprobar la existencia de correlación entre las variables se utilizaron tres métodos analíticos: matriz de correlaciones, la prueba de esfericidad de Bartlett [24] y el índice de Kaiser-Meyer-Olkin (KMO) [21].

Dado que las variables se midieron usando escalas tipo Likert, se utilizó el coeficiente rho de Spearman, criterio no paramétrico, para determinar la correlación entre las variables. Los resultados de la prueba (Anejo 2) muestran que existe una correlación significativa entre algunos pares de variables, sugiriendo su agrupación.

La prueba de esfericidad de Bartlett contrasta la hipótesis de igualdad de la matriz de correlaciones con la matriz identidad (hipótesis nula) [15]. Así, el no poder rechazar la hipótesis nula implica que las correlaciones entre las variables serían nulas y no tendría sentido llevar a cabo el análisis factorial. La prueba consiste en una estimación de χ^2 a partir de una transformación de la matriz de correlaciones.

Tabla 4-1 Descripción de las Variables Utilizadas en el Análisis Factorial

Variab les	Descripción
P1	Tener beca
P8	Nota del curso anterior que tomó de matemáticas
P10	Consideración del desempeño en la Escuela Superior
P11	Número de veces que ha tomado el curso
P12	Importancia que tiene el curso
P13	Prioridad que da al curso
P14	Trabaja actualmente
P16	Promedio de horas de estudio en el centro de apoyo
P19	Promedio de horas de estudio fuera del centro de apoyo
P22	Ir a la oficina del profesor
P24	Ausencias
P25	Número de familiares cercanos con por lo menos un grado de bachillerato
P26	Cantidad de ejercicios hechos
P27	Exámenes anteriores
P28	Practicar con ejercicios de la libreta
P29	Desarrollar ejercicios sin leer la teoría
P30	Tener libro de texto
P31	Curso debiera estar fuera del currículo
P32	Formular preguntas en la clase
P33	Recordar el horario de oficina
P34	Darse de baja
P35	Cantidad de temas para este examen
P36	Tiempo para resolver el examen
P37	Tiempo de la clase para cubrir los temas
P38	Cantidad de ejercicios resueltos en clase
P39	Ejercicios al terminar cada sección
P40	Desarrollar ejercicios faltando poco tiempo para el examen
P41	Repasar las notas de clase
P42	Repasar las notas de clase faltando poco tiempo
P43	Estudiar en grupo los temas del examen
P44	Memorizar la mayoría de ejercicios
P45	Las otras materias absorbieron mayor parte del tiempo
P47	Créditos matriculados
P48	Cantidad de personas con las que estudia en grupo
P49	Gusto por las matemáticas
P50	Dificultad del examen
P51	Esfuerzo por obtener buena nota
P52	Capacidad para estudiar matemáticas
P53	Rapidez con la que el profesor cubrió los temas
P54	Claridad con la que explica el profesor
P55	Interés por estudiar el curso
P57	Exigencias en el estudio
P58	Interés por aprender matemáticas
P60	Edad

El valor obtenido para los datos en estudio es de 2611.878 con un valor p muy cercano a cero indicando el rechazo de la hipótesis nula.

El índice KMO compara los coeficientes de correlación de Pearson con los coeficientes de correlación parcial entre las variables. Se acepta comúnmente que [21]:

- Si $KMO < 0.5$ no se recomienda realizar un análisis factorial.
- Si $0.5 < KMO < 0.6$ el grado de correlación es medio y habría aceptación media de realizar un análisis factorial.
- Si $KMO > 0.7$ indica alta correlación entre las variables y, por tanto conviene aplicar un análisis factorial.

Para los datos en estudio, el índice obtenido fue de 0.720, lo que indica que es conveniente aplicar el análisis factorial.

El criterio que se ha tenido en cuenta para determinar el número de factores ha sido el de la gráfica de sedimentación en la cual se visualiza el porcentaje de varianza explicada por cada factor [28]. Los valores faltantes han sido imputados por la moda de cada variable.

Se observa, Figura 4-1, que después del sexto componente o factor, el porcentaje de varianza explicado es mínimo; sin embargo, queriendo ganar interpretabilidad, se han seleccionado ocho componentes.

Los porcentajes de varianza explicados por cada factor seleccionado y el porcentaje de varianza acumulado de acuerdo con el número de factores seleccionados figuran en la Tabla 4-2. El método utilizado ha sido el de componentes principales con una rotación varimax, que logra extraer de forma ortogonal el valor de la correlación de la varianza en el factor

(correlación cero entre los factores). El uso de esta rotación facilitó la interpretación de los factores.

Figura 4-1

Gráfico de sedimentación

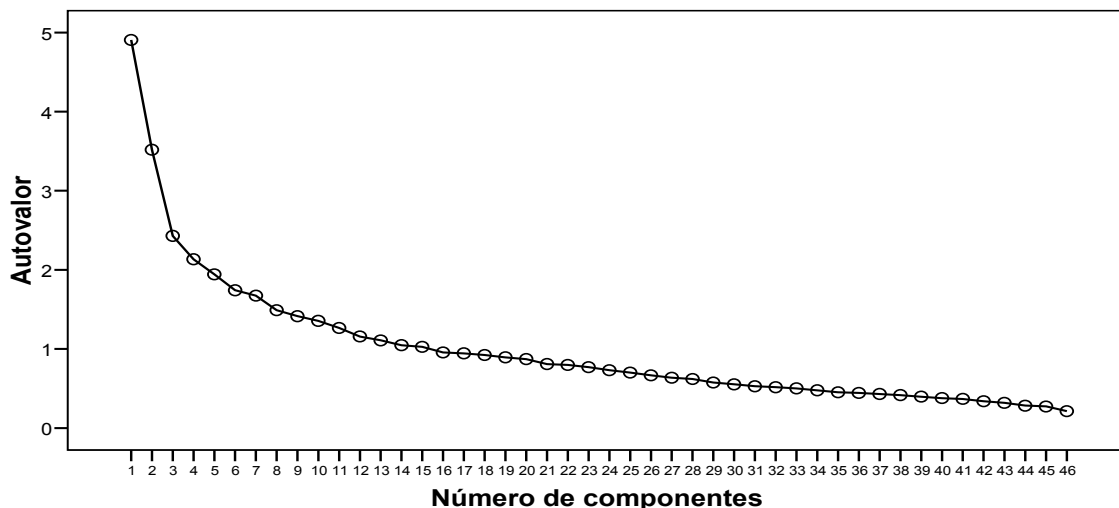


Tabla 4-2 Porcentaje de Varianza Acumulado de Acuerdo al Número de Factores

Varianza total explicada						
Factor	Autovalores iniciales			Suma de las saturaciones al cuadrado de la rotación		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	4.776	10.856	10.856	3.777	8.583	8.583
2	3.494	7.941	18.796	2.925	6.647	15.230
3	2.379	5.408	24.204	2.758	6.269	21.500
4	1.996	4.536	28.740	2.607	5.925	27.425
5	1.762	4.005	32.745	1.843	4.190	31.615
6	1.662	3.778	36.523	1.843	4.188	35.803
7	1.586	3.604	40.127	1.675	3.807	39.610
8	1.432	3.254	43.381	1.659	3.771	43.381

La Tabla 4-3 muestra las variables ordenadas en función del grado de explicación de la varianza total de cada uno de los factores en sentido decreciente incluyendo solo aquellos casos en los que la correlación es superior, en valor absoluto, a 0.3.

Tabla 4-3 Componentes Rotados

Variables	Componente							
	1	2	3	4	5	6	7	8
P58	0.729							
P55	0.697							
P49	0.646							
P52	0.638							
P57	0.566							
P13	0.556							
P31	-0.469							
P50		0.665						
P53		0.576						
P37		-0.541						
P54		-0.512				0.410		
P38		-0.504						
P35		0.467						
P34		0.451						
P45		0.407					0.325	
P11			0.758					
P8			-0.750					
P60			0.723					
P47			-0.443					
P14			0.405					
P39				0.652				
P26				0.639				-0.385
P16				0.579				
P51	0.457			0.485				
P19	0.303			0.466				
P42				-0.460				
P22				0.453				
P43					0.651			
P27					0.506			
P36		-0.368			-0.411			
P10					-0.382			
P29					0.340		0.304	
P28						0.661		
P41						0.598		
P33						0.524		
P24						-0.362		
P1							0.646	
P30							-0.629	
P25							-0.485	
P32							0.320	
P44								0.568
P40		0.366						-0.491
P48					0.368			0.468
P12					0.328			-0.362

4.2.2 Descripción de los factores

FACTOR 1

Hemos denominado a este factor “*Actitud hacia las matemáticas*”, pues las variables que están más correlacionadas con él, describen la percepción que el estudiante tiene ante la materia. Un valor elevado de este factor indica que el estudiante tiene mucho interés por aprender matemáticas y estudiar el curso. También que el estudiante le concede una prioridad alta a la materia y considera que su capacidad y exigencias para estudiar matemáticas son altas. Los valores negativos se deben a la forma en la cual se ha medido la variable por ejemplo; para la variable “curso debiera estar fuera del currículo” identificada como P31, el valor obtenido en el factor 1 es -.469 indica que al estudiante no le gustaría que el curso de Cálculo I estuviera fuera de su currículo de estudios.

FACTOR 2

Llamamos a este factor “*Características del profesor y del examen*”, pues las variables que están más correlacionadas con este factor, en su mayoría, tienen relación con la metodología utilizada por el profesor en la enseñanza del curso. También vemos algunas variables muy específicas sobre las características del examen. Un valor alto en este factor indica que el estudiante ha considerado el examen muy difícil, que la cantidad de temas evaluados fue demasiada y que el profesor cubrió los temas muy rápido, siendo poco claro en sus explicaciones.

FACTOR 3

A este factor lo hemos llamado “*Rezago académico*”. Las variables que están más correlacionadas con este factor hacen referencia a la edad e historial que el estudiante tiene sobre la asignatura. Valores altos en esta componente indican que la edad del estudiante está por encima del promedio de la edad de los estudiantes encuestados, que ha tomado varias veces la clase y que tiene pocos créditos matriculados.

FACTOR 4

Se interpreta como factor de “*Esfuerzo y tiempo dedicado*”. Las variables que pertenecen a este factor describen algunos hábitos que tienen los estudiantes en su preparación para el examen. Así, un valor alto de esta componente indica que el estudiante resuelve ejercicios al terminar cada sección, asiste con frecuencia al centro de apoyo, no repasa las notas de clases cuando falta poco tiempo para el examen, asiste con alguna frecuencia a la oficina del profesor y considera que ha hecho mucho esfuerzo para obtener una buena nota en el examen.

FACTOR 5

Llamamos a este factor “*Estudio en grupo*”. En este factor están correlacionadas las variables que describen la actividad principal que los estudiantes desarrollan al estudiar en grupo. Así, un valor alto de este factor indica que el estudiante estudia en grupo y lo hace con tres o más personas, también resuelve exámenes de semestres anteriores y piensa que el tiempo para resolver el examen fue poco; también considera que su desempeño en matemáticas en la escuela superior fue bajo.

FACTOR 6

Se interpreta como factor de “*Repaso del trabajo en clase*”. Un valor elevado en este factor implica que el estudiante practicó ejercicios de la libreta, tiene pocas o ninguna ausencia a clase, repasó las notas de clase y considera que el profesor ha sido claro en sus explicaciones. Se observa que este factor, implícitamente dice si el estudiante ha tomado notas durante la clase.

FACTOR 7

El factor 7 refleja un poco el “*Entorno social*” que rodea al estudiante. Puntuaciones altas en este factor indican que el estudiante posee una beca por parte de la universidad, tiene pocos o ningún familiar cercano que haya obtenido al menos un título de bachillerato, no tiene libro de texto y piensa que las otras clases le consumieron mucho tiempo.

FACTOR 8

Se interpreta como factor de “*Actitud hacia el curso*”. Un valor alto en este factor implica que el estudiante no considera importante este curso para su formación profesional, memoriza la mayoría de ejercicios, no resuelve ejercicios faltando poco tiempo para el examen y cuando estudia en grupo lo hace, en promedio, con más de tres personas.

4.3 Análisis de conglomerados sobre los factores

Se aplicó la técnica de análisis de conglomerados a los 265 estudiantes que entregaron el cuestionario I. Se utilizaron como variables descriptivas del desempeño académico los factores obtenidos a partir del análisis factorial, expuesto anteriormente. Es importante señalar que estas variables están normalizadas como resultado del análisis factorial.

4.3.1 Validación del análisis de conglomerados

Se aplicaron varias técnicas de análisis de conglomerados para decidir, de acuerdo a tres criterios de validación interna y al número de estudiantes por grupo, cuál sería la mejor partición. Estos criterios de validación fueron [8]: el índice de Dunn, el índice de Davies-Bouldin y el ancho promedio de silueta. Estos se aplicaron a cuatro algoritmos jerárquicos aglomerativos: “Ward”, “Complete”, “Single” y “Average”, a un algoritmo jerárquico divisivo denominado “Diana”, existente en la librería “*cluster*” del software R y al método de partición por k-medioídes (PAM), que también figura en la librería “*cluster*” de R. En cada algoritmo de agrupación se usó la distancia euclidiana.

Para calcular el índice de Dunn y de Davies-Bouldin, se utilizaron las funciones que escribió González en su tesis de maestría [23] “A comparison in cluster validation techniques”. El ancho promedio de silueta se obtuvo con la función “*silhouette*” de la librería “*cluster*”

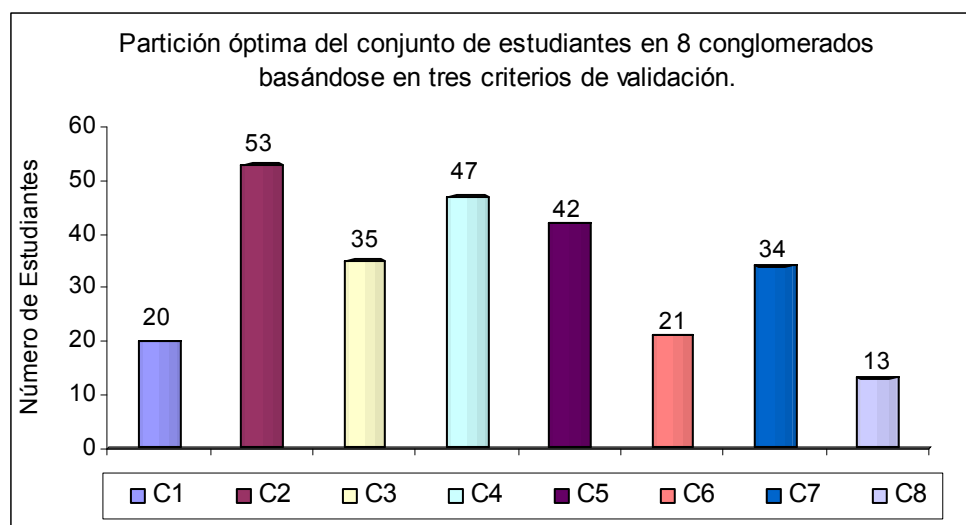
Las Tablas 1 y 2 del Anejo 3, muestran que los métodos de “single” y “average” no hacen una buena partición del conjunto ya que en ambos casos, la mayoría de las observaciones quedan asignadas a un sólo grupo. La Tabla 4-4 muestra el número de grupos de acuerdo a los tres criterios de validación utilizados.

Tabla 4-4

Número de conglomerados sugeridos por los tres criterios de validación			
Método	Dunn	Davies-Bouldin	Ancho de Silueta
Ward	2	7	2
Complete	5	5	2
Single	2	2	2
Average	2	5	2
Diana	8	6	5
Pam	8	8	8

Se decidió escoger finalmente los resultados que se obtuvieron al aplicar el método de partición alrededor de los medioides para determinar la clasificación, ya que es la partición que presenta mejor consistencia de acuerdo con los tres criterios de validación. Por tanto, el número óptimo de conglomerados para la explicación de las diferentes tipologías de los estudiantes es ocho. La Figura 4-2 muestra el número de individuos por conglomerado.

Figura 4-2



Una vez determinada la tipología (o clasificación) se calcularon los promedios de cada factor (centroides de cada conglomerado), los que se exponen en la Tabla 4-5.

Tabla 4-5 Centroides de Cada Factor por Conglomerado

FACTOR	CONGLOMERADOS							
	C1	C2	C3	C4	C5	C6	C7	C8
F1	-0.573	0.342	0.641	0.432	-1.044	-0.821	0.339	0.009
F2	0.396	0.288	-0.836	-0.332	-0.021	0.604	0.694	-1.056
F3	-0.743	0.274	-0.674	0.101	-0.383	0.346	0.558	0.695
F4	1.137	-0.324	-0.401	0.911	-0.759	0.188	0.176	-0.959
F5	0.816	0.245	-0.791	0.493	0.353	-0.468	-1.013	0.359
F6	-1.067	-1.008	0.362	0.693	0.267	1.045	-0.294	0.490
F7	0.500	0.120	0.063	-0.401	-0.380	0.640	-0.478	1.467
F8	0.818	-0.702	-0.319	0.311	0.099	-0.922	0.631	0.859

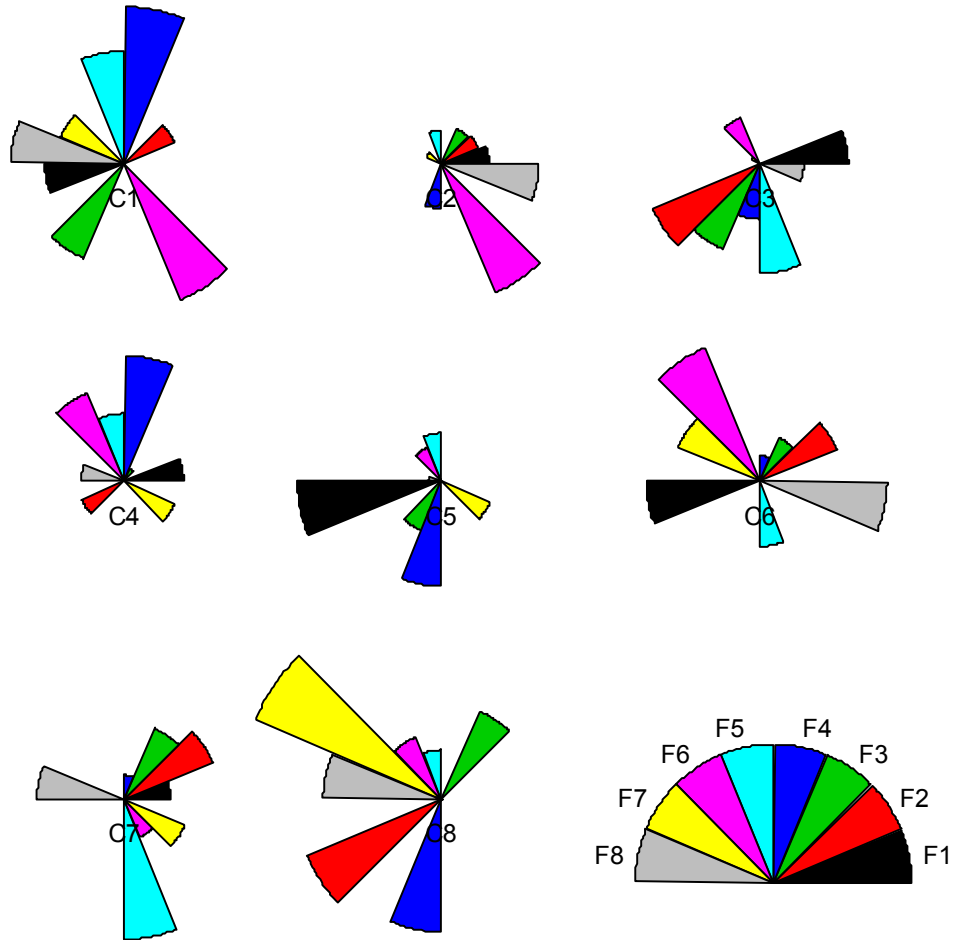
Para visualizar, las diferencias de cada conglomerado hemos construido una gráfica de sectores circulares [13]. Cada conglomerado es representado por una gráfica, los radios de cada sector circular son proporcionales al promedio de cada factor en el respectivo conglomerado. En la Figura 4-3, la dirección del sector circular indica el signo del promedio del factor que representa; así, si el sector está ubicado en la parte superior del semicírculo el promedio del factor es positivo y si está por debajo, es negativo.

Cada gráfica permite comparar, de manera sencilla, el valor del promedio de una variable (factor) en cada uno de los ocho conglomerados. Por ejemplo, se puede observar que los integrantes del conglomerado 3 tienen una puntuación alta positiva en el factor 1 mientras que los integrantes de los conglomerados 5 y 6 dan una puntuación totalmente opuesta.

Comprobamos mediante la prueba no paramétrica Kruskal-Wallis que para cada factor existe por lo menos un conglomerado que difiere significativamente de los demás. Anejo 4.

Figura 4-3

Descripción de los ocho conglomerados por los factores seleccionados



F1	<i>Actitud hacia las matemáticas</i>
F2	<i>Características del profesor y del examen</i>
F3	<i>Rezago académico</i>
F4	<i>Esfuerzo y tiempo dedicado</i>
F5	<i>Estudio en grupo</i>
F6	<i>Repaso del trabajo en clase</i>
F7	<i>Entorno social</i>
F8	<i>Actitud hacia el curso</i>

4.3.2 Distribución de las variables en cada conglomerado

A continuación se presenta una descripción gráfica de la proporción de estudiantes que pertenecen a la facultad de Artes y Ciencias, a la facultad de Ingeniería (se incluyeron en esta facultad 3 estudiantes de la facultad de Ciencias Agrícolas y 2 estudiantes de la facultad de Administración de Empresas) o si se encuentran en la escuela superior, en cada uno de los ocho grupos. Así mismo, se representa el número de mujeres y hombres, si provienen de escuela pública o privada y el promedio del primer examen parcial por conglomerado. Con la intención de lograr mayor objetividad en la formación de los conglomerados se decidió no incluir el género, el tipo de escuela donde procedían los estudiantes ni la puntuación obtenida en el examen, en el análisis factorial.

Figura 4-4

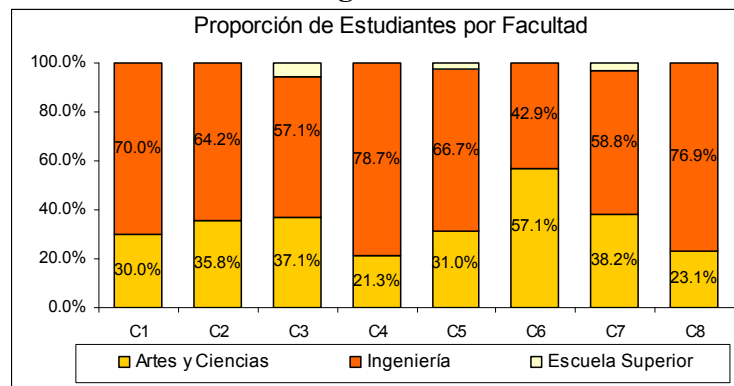


Figura 4-5

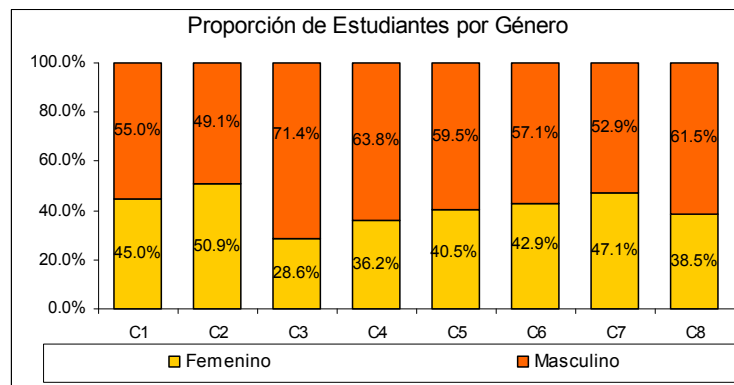


Figura 4-6

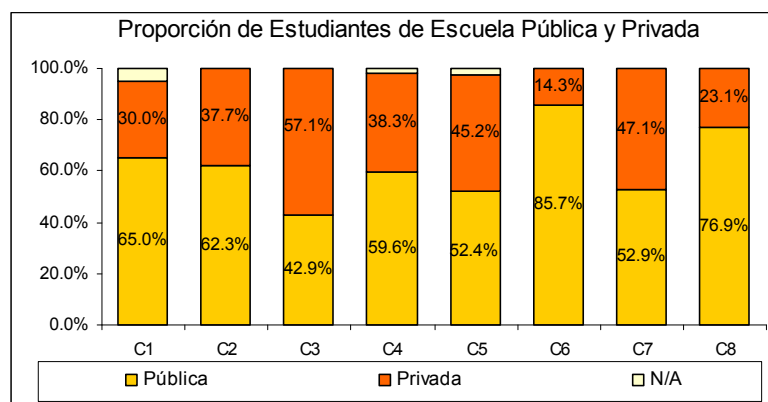
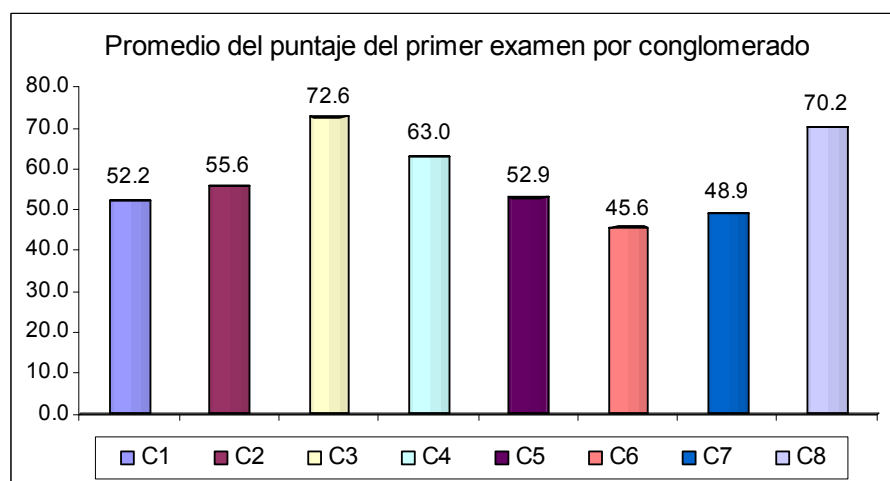


Figura 4-7



Se aplicó la prueba no paramétrica de Kruskal-Wallis, Tabla 4-6, para la comparación de medias para el promedio de la nota del primer examen por conglomerado, la cual nos indica que existe al menos una diferencia estadísticamente significativa entre los promedios de los ocho grupos.

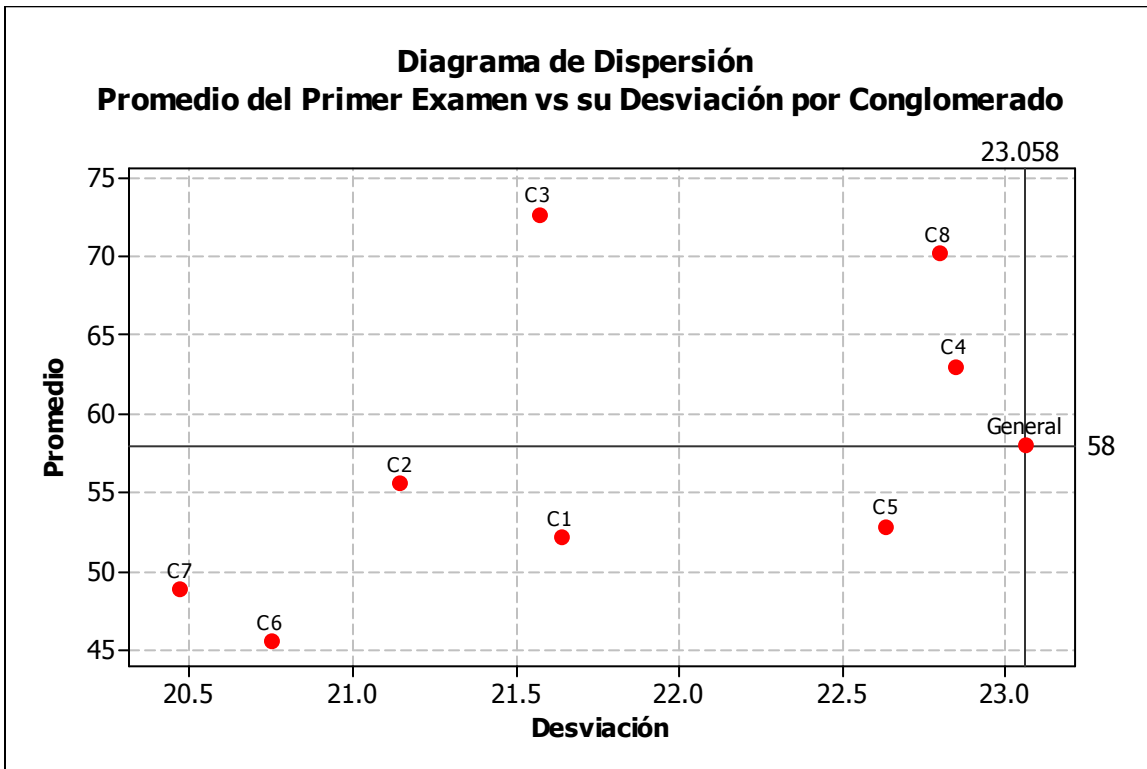
En la Figura 4-8, se puede observar el diagrama de dispersión de los conglomerados en función de su promedio y desviación estándar. El punto denominado *General* representa el promedio y la desviación estándar del grupo formado por los 256 estudiantes (hay nueve faltantes).

Tabla 4-6 Prueba de Kruskal-Wallis Sobre la Puntuación del Examen

Grupos	N	Median	Ave Rank	Z
C1	19	53.0	110.6	-1.10
C2	51	54.0	122.5	-0.65
C3	34	77.5	176.7	4.08
C4	45	67.0	146.1	1.75
C5	40	48.5	112.7	-1.47
C6	21	46.0	90.2	-2.48
C7	33	51.0	102.6	-2.15
C8	13	77.0	167.9	1.97
Overall	256	-	128.5	-

9 casos contienen valores perdidos
H = 33.60 DF = 7 Valor-p = 0.000

Figura 4-8



4.3.2.1 Representación gráfica de la distribución porcentual de los estudiantes por respuesta a los ítems relacionados con cada factor.

Con el fin de facilitar la descripción de los ocho conglomerados formados, se muestra a continuación, la proporción de estudiantes por respuesta ante los ítems que están más correlacionados con cada factor, por conglomerado.

Figura 4-9

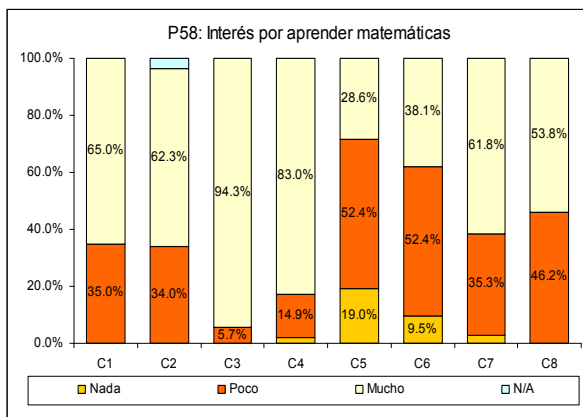


Figura 4-11

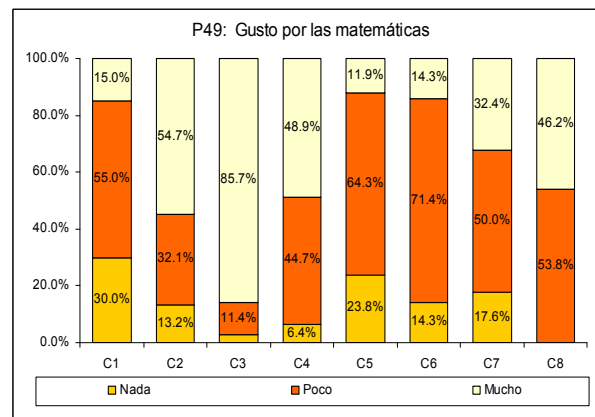


Figura 4-10

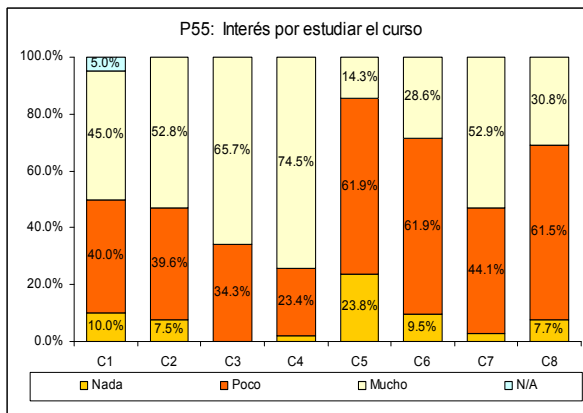


Figura 4-12

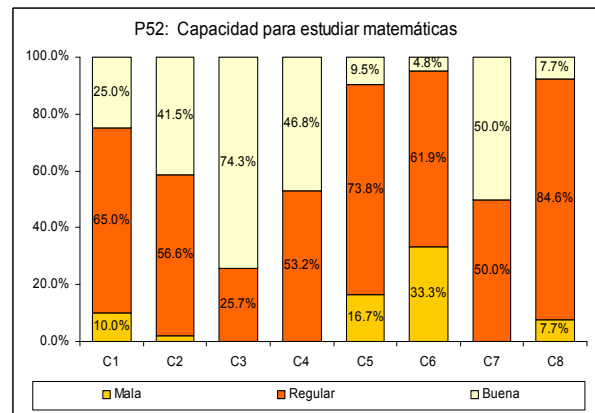


Figura 4-13

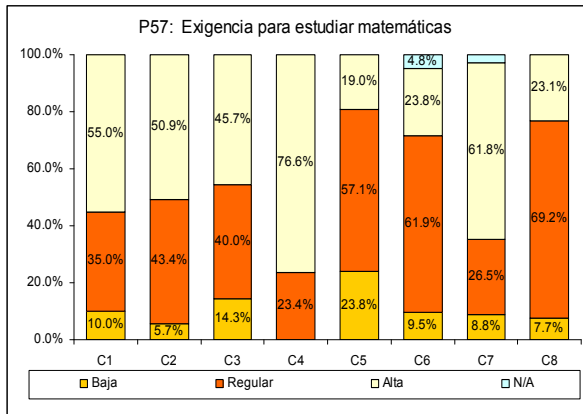


Figura 4-14

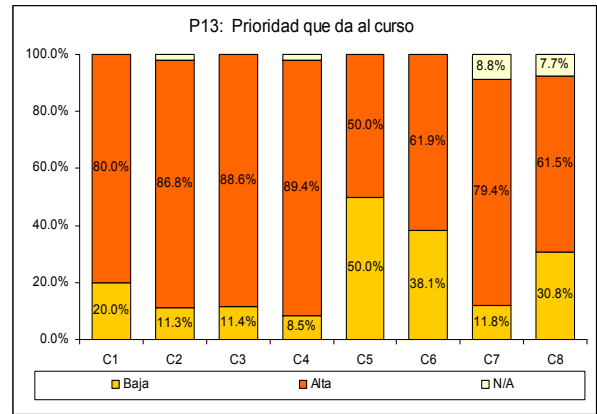


Figura 4-15

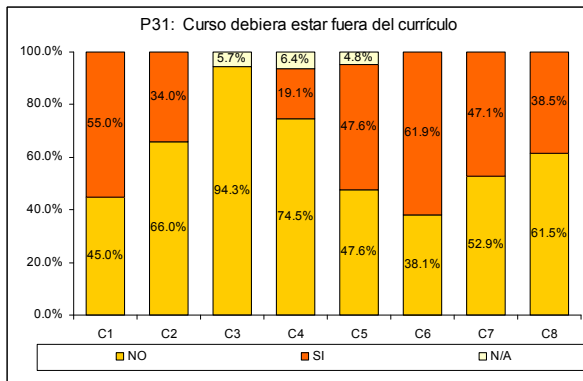


Figura 4-17

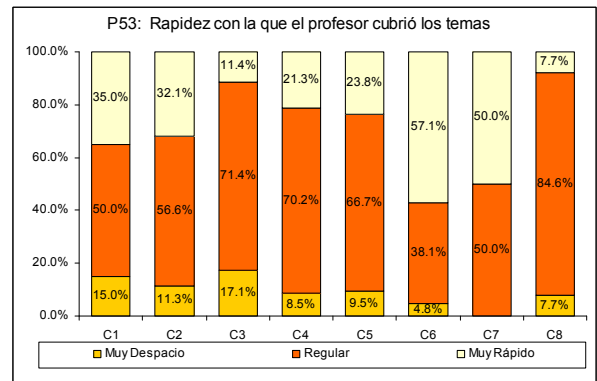


Figura 4-16

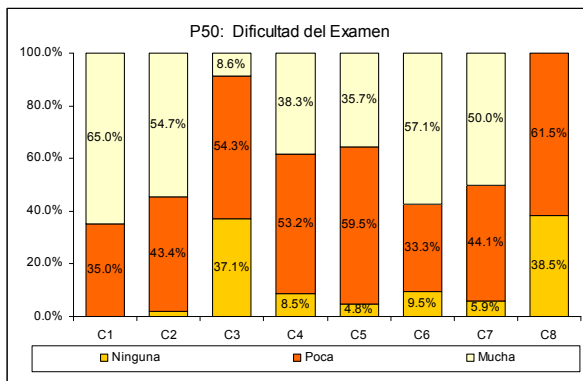


Figura 4-18

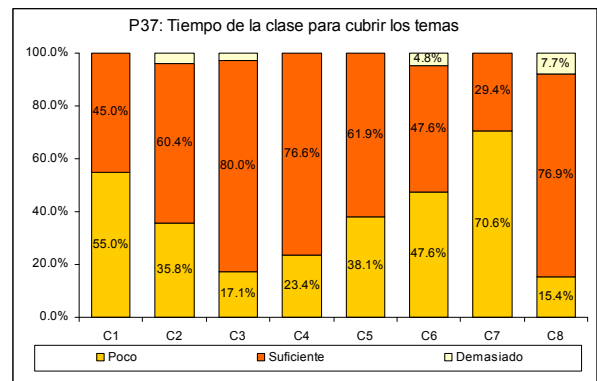


Figura 4-19

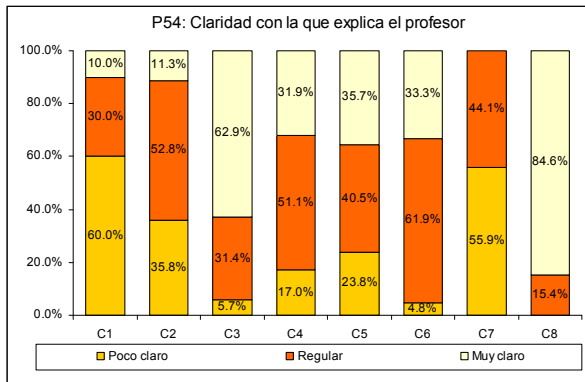


Figura 4-20

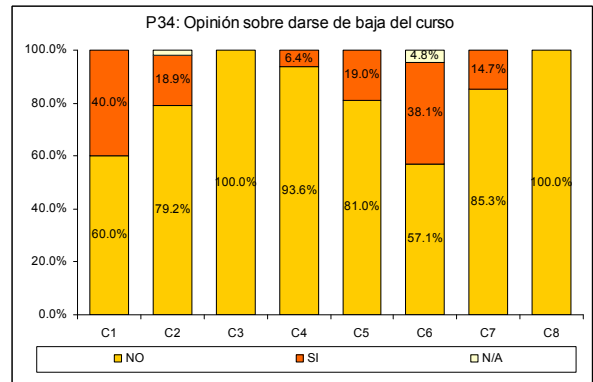


Figura 4-21

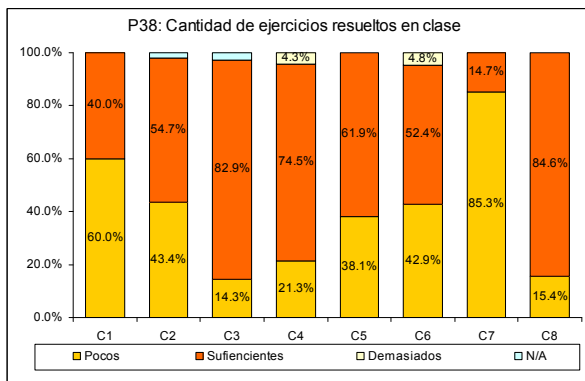


Figura 4-22

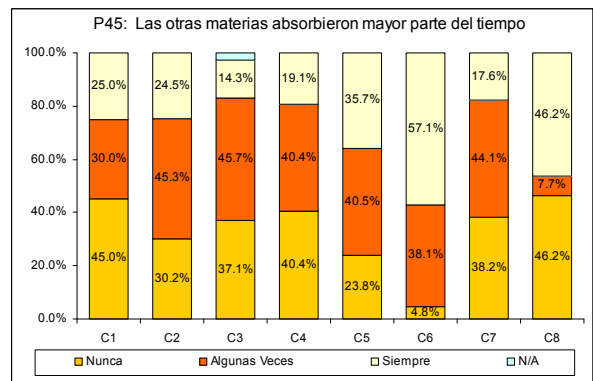


Figura 4-23

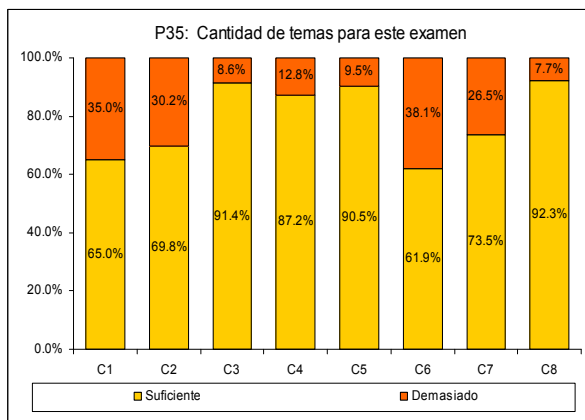


Figura 4-24

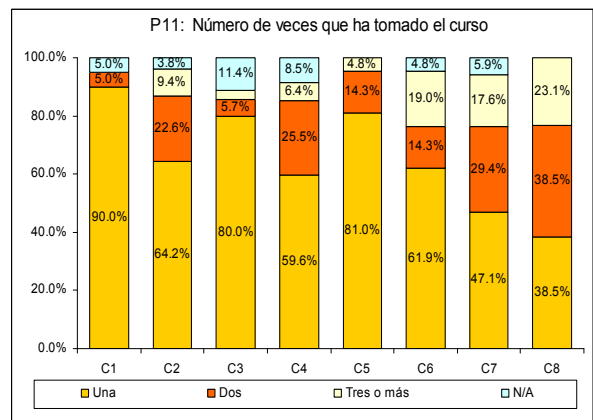


Figura 4-25

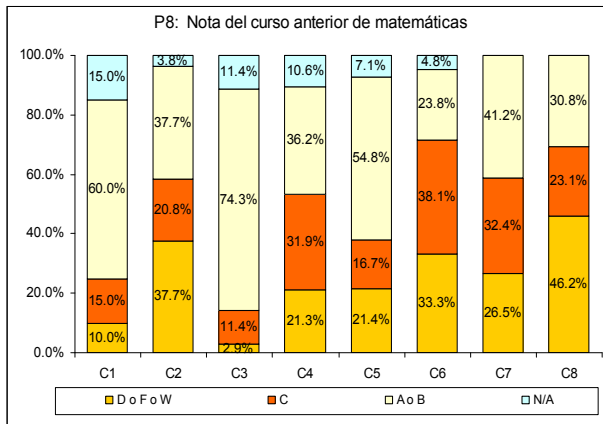


Figura 4-26

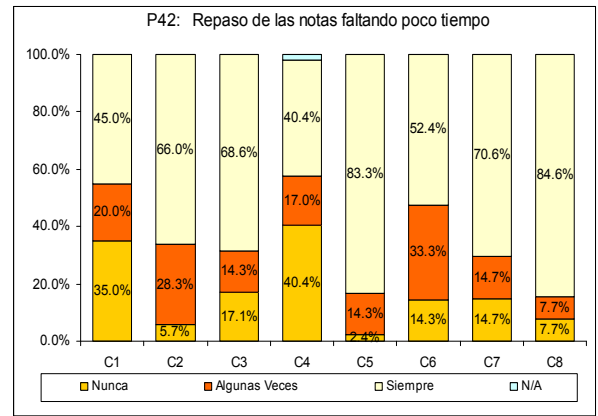


Figura 4-27

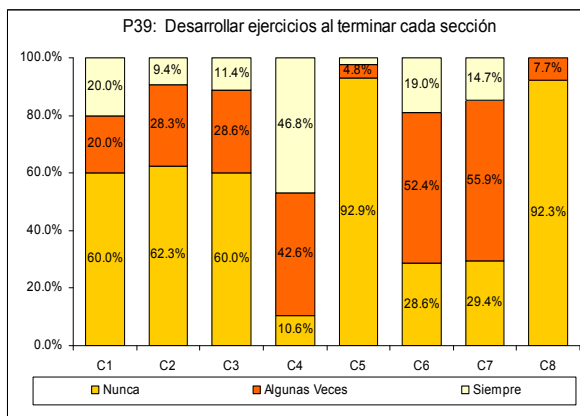


Figura 4-28

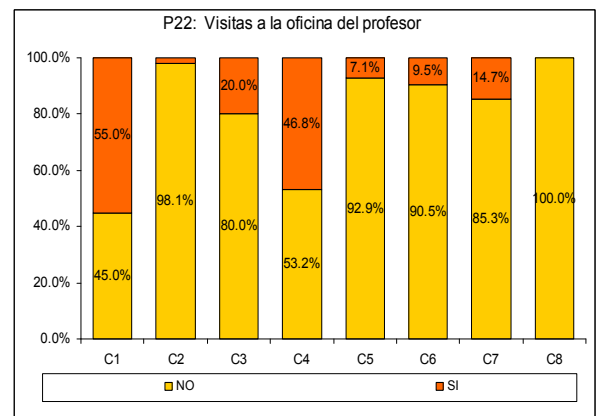


Figura 4-29

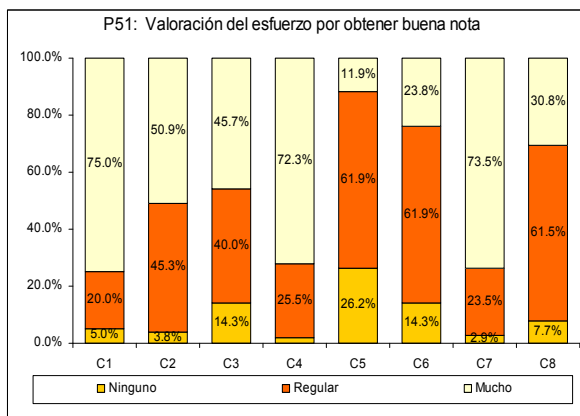


Figura 4-30

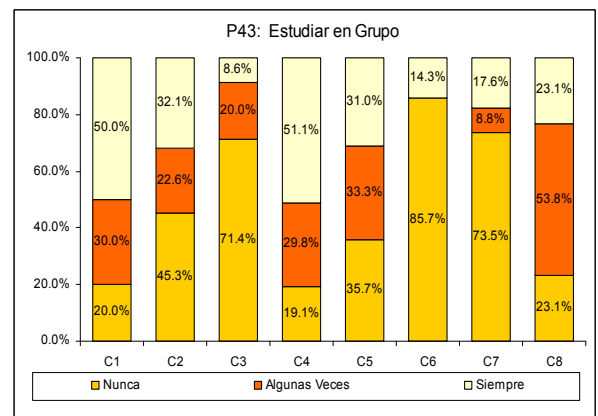


Figura 4-31

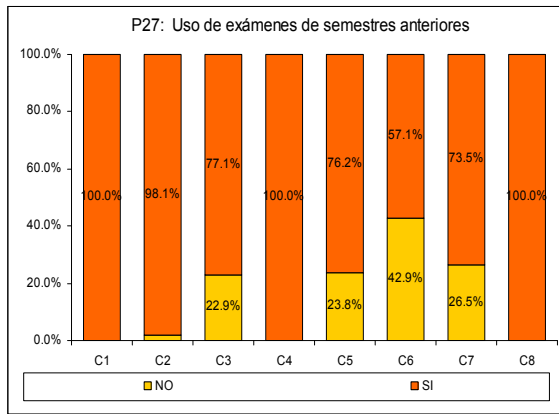


Figura 4-32

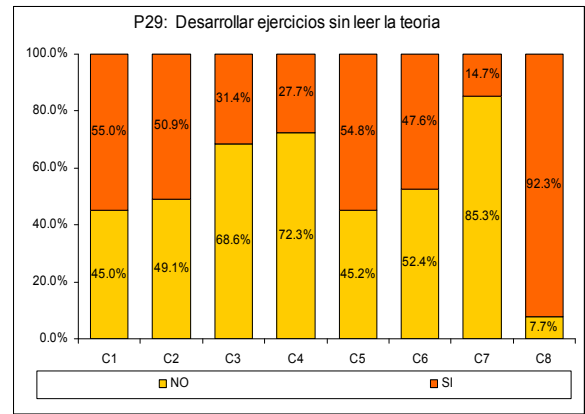


Figura 4-33

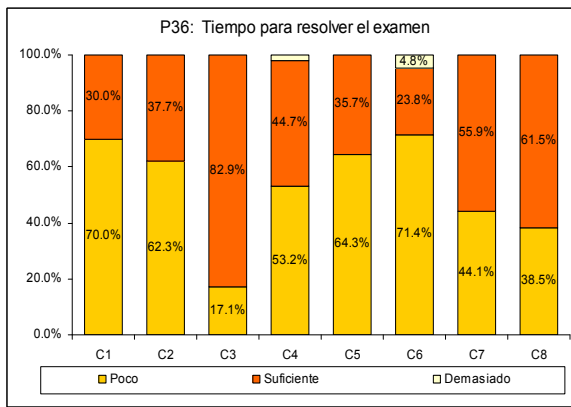


Figura 4-34

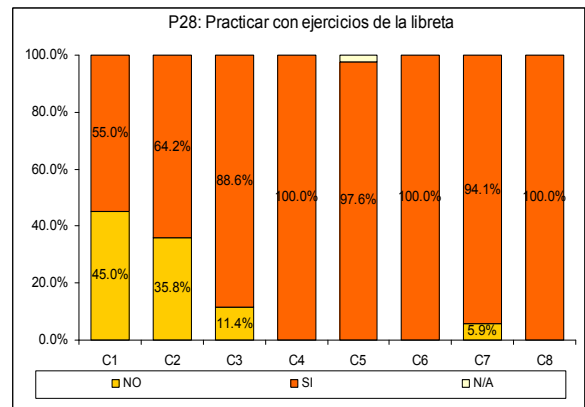


Figura 4-35

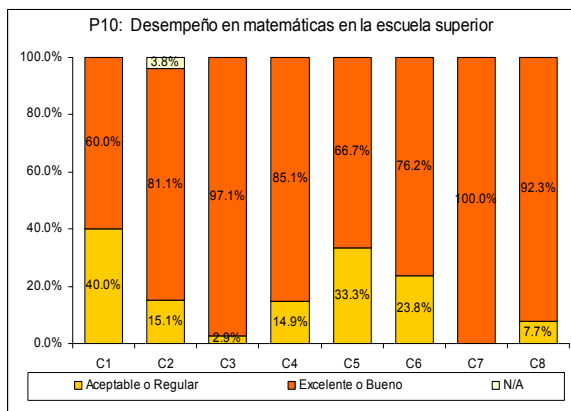


Figura 4-36

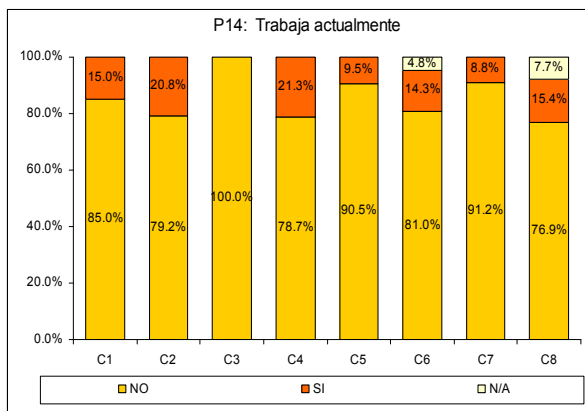


Figura 4-37

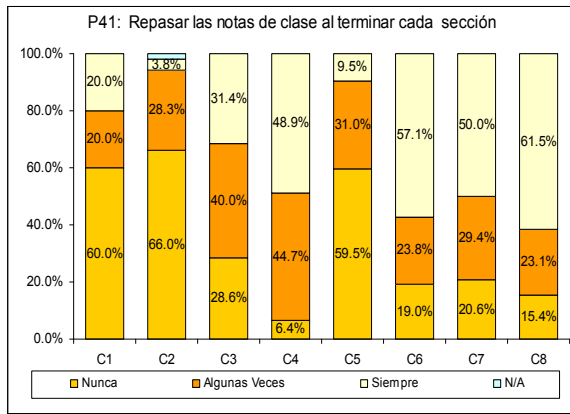


Figura 4-38

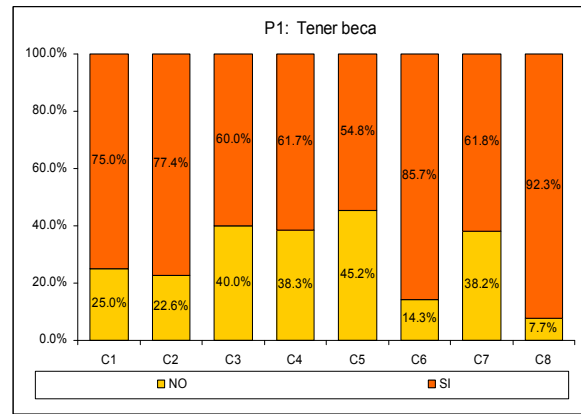


Figura 4-39

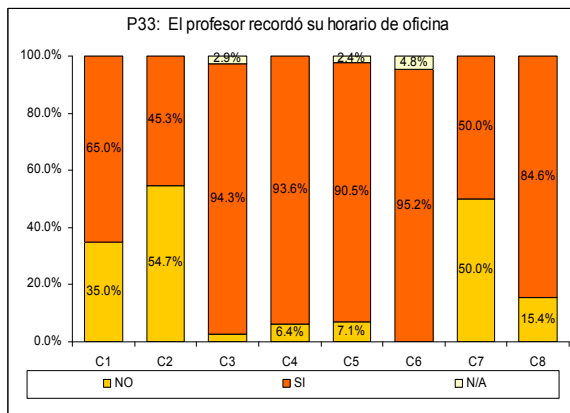


Figura 4-40

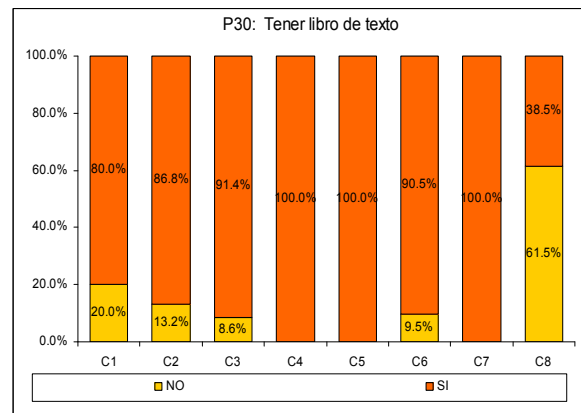


Figura 4-41

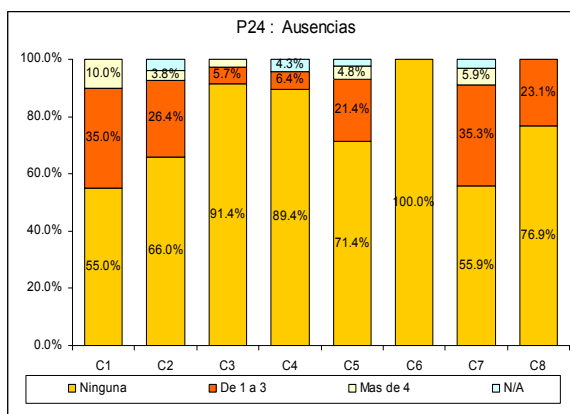


Figura 4-42

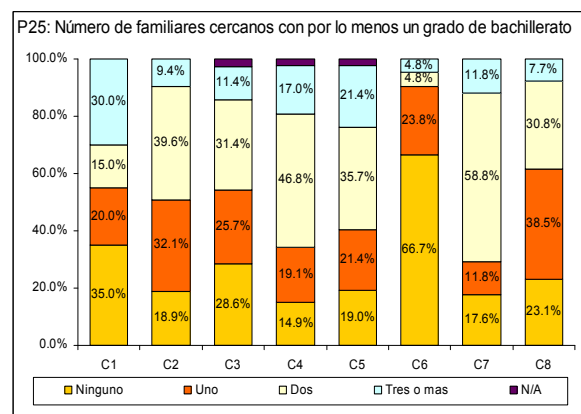


Figura 4-43

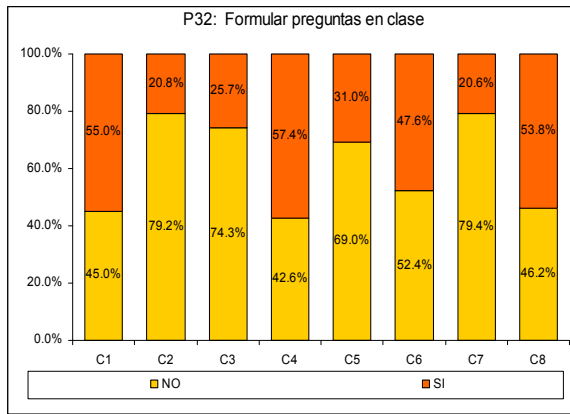


Figura 4-44

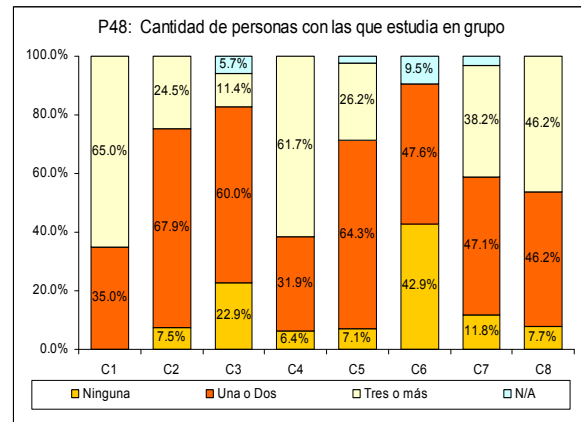


Figura 4-45

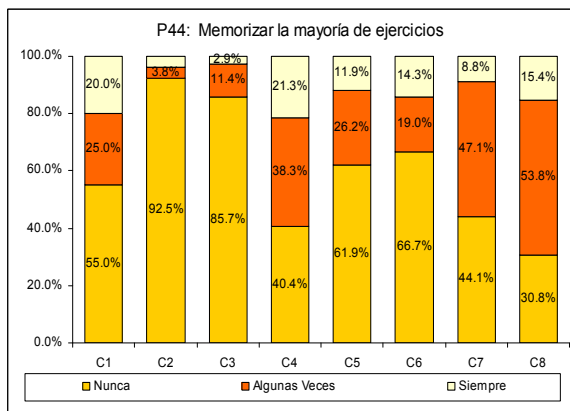


Figura 4-46

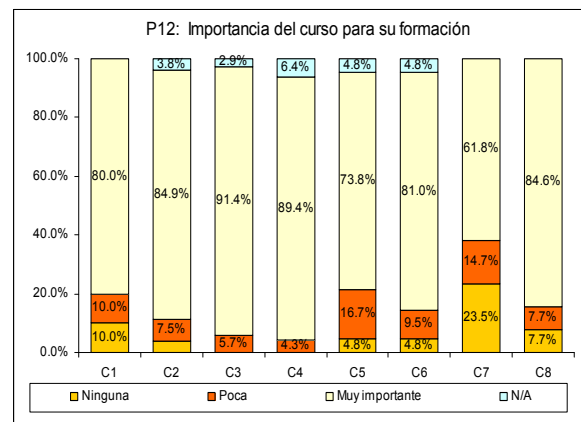


Figura 4-47

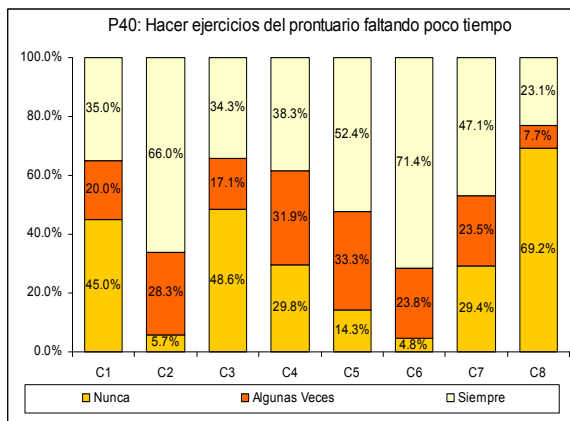
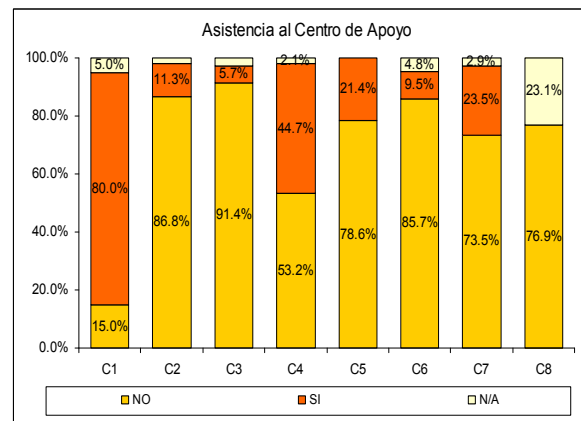


Figura 4-48

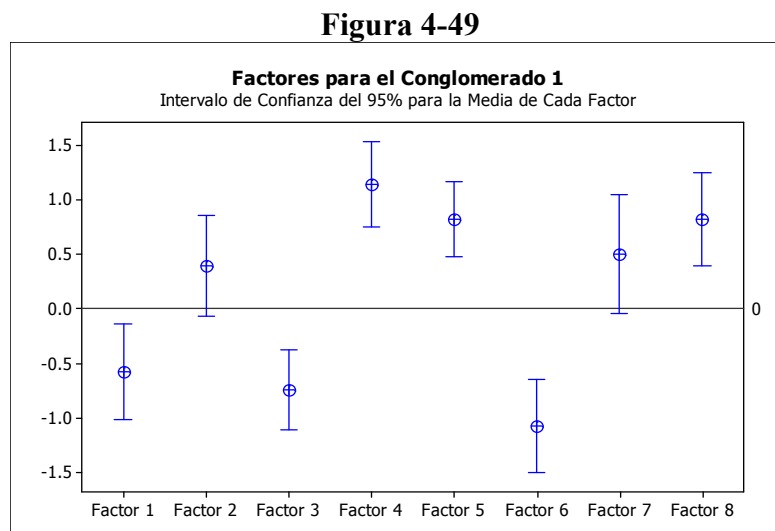


Aunque en el análisis factorial se tuvo en cuenta el promedio de horas que los estudiantes asistieron al Centro de Apoyo, se decidió que es más ilustrativo señalar si el estudiante asiste o no al centro de apoyo, ya que algunos valores atípicos (promedios muy altos) afectaban el promedio general del conglomerado dificultando la comparación entre éstos. Figura 4-48.

4.3.3 Descripción de los conglomerados

Para determinar la significancia que tiene cada variable (factor) en los conglomerados, se estimaron intervalos de confianza del 95% alrededor de la media. Nótese que los factores están normalizados con media cero y desviación estándar uno. Por tanto, si el intervalo de confianza contiene el valor cero para un factor en un conglomerado particular, diremos que ese factor no es significativo en dicho conglomerado. Otra herramienta que se utilizó para la interpretación fue la gráfica de sectores circulares; ésta ayuda a comparar una variable, no sólo entre los grupos, sino también dentro de los grupos. Se describen a continuación las características más importantes encontradas en los ocho conglomerados.

Conglomerado 1



Este conglomerado está formado por 20 estudiantes, el 68.4% proviene de escuela pública.

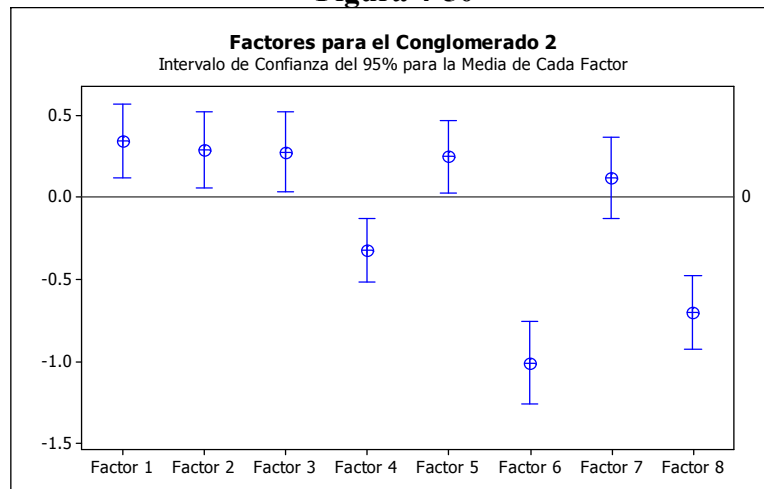
El promedio general para el primer examen fue de 52.2, el cual está ubicado 5.84 puntos por debajo del promedio general.

Este conglomerado se caracteriza principalmente por tener un puntaje bajo en el factor 6 de “*Repaso del trabajo en clase*”, el 45% no practica con ejercicios de la libreta, el 60% no repasa las notas al terminar cada sección, el 35% ha tenido de una a tres ausencias a clase y el 55% considera que el profesor ha sido poco claro en sus explicaciones. También observamos una baja puntuación en el factor 3 “*rezago académico*”. El 90% está tomando el curso por primera vez, el 60% tuvo una calificación de A o B en el curso anterior de matemáticas y el 15% de C. El valor negativo del factor 1 indica una mala actitud hacia las matemáticas; el 55% dice tener poco gusto por esta materia mientras que el 30% dice que no le gusta. Este grupo tiene una puntuación alta en el factor 4 de “*esfuerzo y tiempo dedicado*” El 75% considera que ha hecho mucho esfuerzo para obtener una buena nota en el examen. El 55% y 80%, los porcentajes más altos en todos los grupos, afirma haber ido a la oficina del profesor y al centro de apoyo respectivamente. El 60% nunca desarrolla ejercicios al terminar cada sección y el 20% lo hace algunas veces.

Conglomerado 2

Este conglomerado está integrado por 53 estudiantes, es el conglomerado con el mayor número de estudiantes. El promedio general del grupo es 55.61, ubicado 2.4 puntos por debajo del promedio general, que es el promedio más cercano a éste.

Figura 4-50

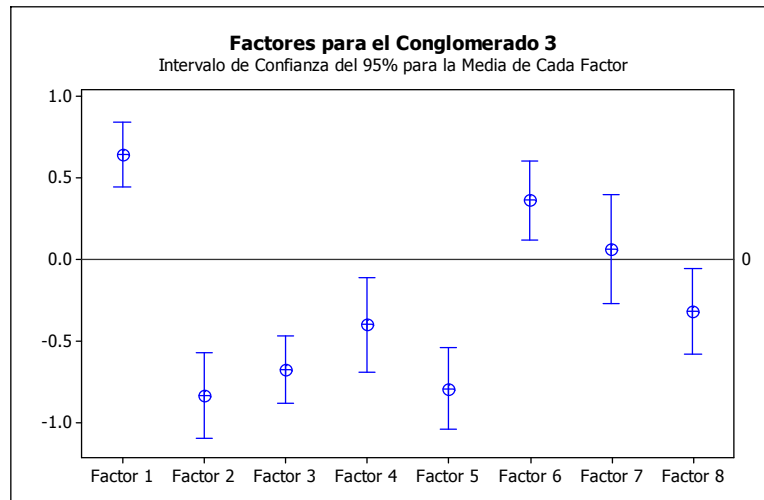


Este grupo se caracteriza principalmente, al igual que el conglomerado 1, por obtener una puntuación muy baja negativa en el factor “*Repaso del trabajo en clase*”. El 35.8% de los estudiantes no practica los ejercicios de la libreta, el 66% nunca repasa las notas de clase y el 28.3% lo hace algunas veces. El 66% dice no haber faltado a ninguna clase. Otro factor que sobresale en este grupo es “*Actitud hacia el curso*”. El 80% de los estudiantes considera que el curso es importante para su formación profesional, el 92.5%, el más alto en todos los conglomerados, afirman que nunca memorizan la mayoría de ejercicios y el 66% siempre resuelve ejercicios del prontuario faltando poco tiempo para el examen.

Conglomerado 3

Este grupo está integrado por 35 estudiantes, representan el 13.2% del total de la muestra. El 60.6% pertenece a la facultad de Ingeniería y el 39.4% a la facultad de Artes y Ciencias. El 42.9% provienen de escuela pública y el 57.1% de escuela privada. El promedio general del grupo, el más alto de todos los conglomerados, es de 72.65, ubicado 14.65 puntos por encima del promedio general.

Figura 4-51

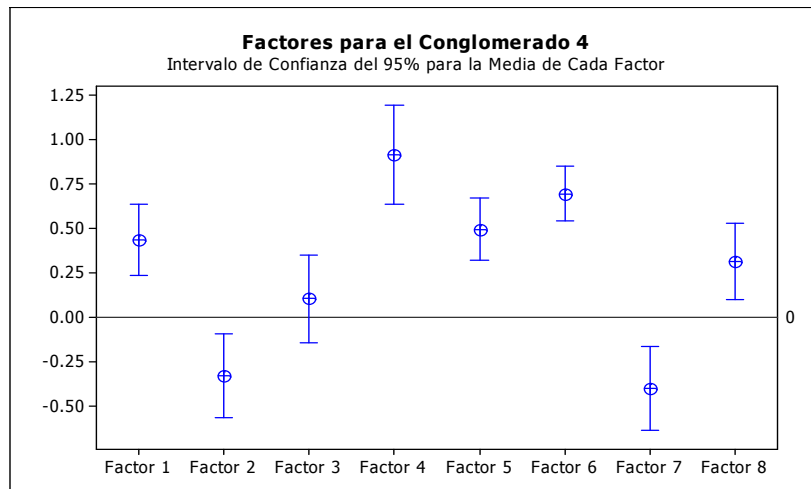


Este grupo se caracteriza por tener un puntaje alto en el factor 1 “*Actitud hacia las matemáticas*”. El 94.3% de los estudiantes dicen tener mucho interés por aprender matemáticas, el 65.7% (el segundo más alto dentro de los conglomerados) afirma tener mucho interés por estudiar el curso y el 85.7% siente mucho gusto por esta materia. El factor 2 “*Características del profesor y del examen*” también tiene un puntaje significativo en este grupo. El 54.3% encontró poca dificultad en el examen y el 37.1% ninguna dificultad, el 71.4% considera que la velocidad con la que el profesor cubrió los temas fue regular y el 17.1% que fue muy despacio, 62.9% considera que el profesor fue muy claro en la explicación de los temas. El factor 3, “*rezago académico*”, en este conglomerado tiene un valor bastante bajo al igual que en el conglomerado 1, sin embargo, la dispersión en este grupo es mucho menor (Anejo 6). El factor “*estudio en grupo*” tiene una puntuación baja. El 73% dice no haber estudiado en grupo ninguno de los temas del examen, el 82.9% considera que el tiempo para resolver el examen fue suficiente, 97.1% afirman que su desempeño en la escuela fue Excelente o Bueno y el 31.4% desarrolla ejercicios sin leer la teoría.

Conglomerado 4

Este grupo está formado por 47 estudiantes, el 59.6% viene de escuela pública. El promedio de este grupo es de 63, ubicado 5 puntos por encima del promedio general.

Figura 4-52

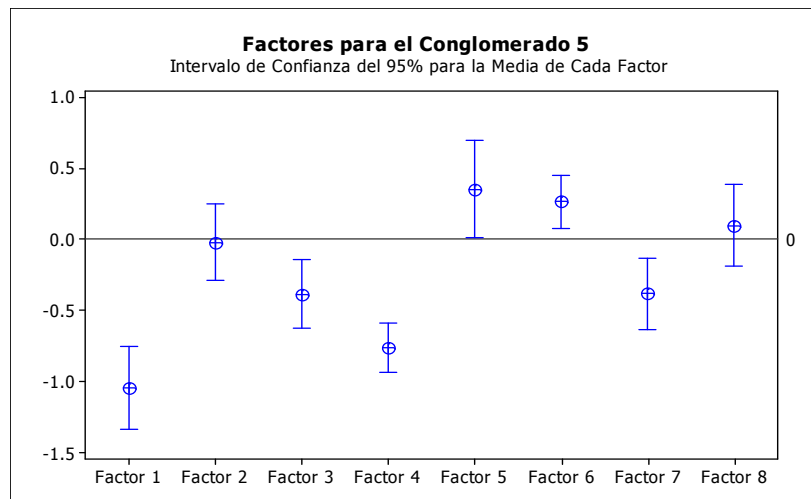


Los estudiantes de este conglomerado se caracterizan principalmente por tener un puntaje alto en el factor de “*esfuerzo y tiempo dedicado*”. El 42.6% resuelve algunas veces ejercicios al terminar cada sección y el 46.8% lo hace siempre. El 72.3% considera que ha hecho mucho esfuerzo para obtener una buena nota y 44.7% dice haber asistido al Centro de Apoyo. Los estudiantes de este grupo también se caracterizan por estudiar en grupo, el 51.1%, el porcentaje más alto de los conglomerados, dice haber estudiado los temas del examen siempre en grupo y el 30% lo hizo algunas veces. Por otra parte, referente al factor 6, el 60% tomó el curso por primera vez, el 44.7% algunas veces repasa las notas de clase al terminar cada sección y el 49% dice hacerlo siempre, 89.4% asistió a todas las clases previas al examen.

Conglomerado 5

Este conglomerado está integrado por 42 estudiantes, el 52.4% proviene de escuela pública. El promedio del examen 1 para este grupo fue de 52.9, ubicado 5.15 puntos por debajo del promedio general.

Figura 4-53

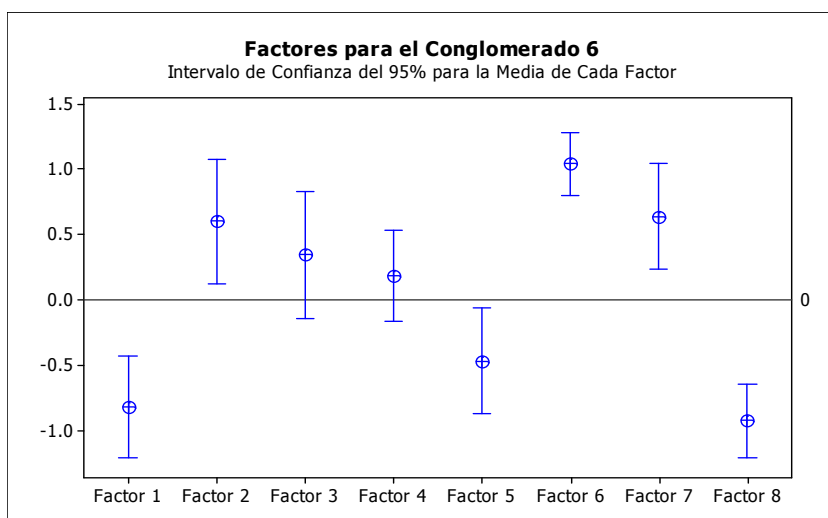


Se trata del grupo con la puntuación más baja en el factor 1 “*actitud hacia la matemática*”. El 52.4% dice tener poco interés por aprender matemáticas mientras que el 19% dice no tener alguno. El 14%, el puntaje más bajo de todos los conglomerados, tiene mucho interés por estudiar el curso. El 64.3% dice tener poco gusto por las matemáticas y el 23% no siente gusto en ésta área. El 73.8% dice que su capacidad para estudiar matemáticas es regular y el 16.7% afirma que su capacidad es mala. El 19% tiene altas exigencias para el estudio de las matemáticas. Otro de los factores con una puntuación baja es el factor de esfuerzo y tiempo dedicado. EL 24% no usó exámenes de semestres anteriores como material de preparación para el examen, el 64.3% afirma que el tiempo para resolver el examen fue poco. El 33.3% piensa que su desempeño en la escuela fue aceptable o regular y el 42% desarrolla ejercicios sin haber leído la teoría.

Conglomerado 6

El conglomerado está integrado por 21 estudiantes, el 85.7% proviene de la escuela pública, el promedio general de este grupo es de 45.6, el más bajo de todos los conglomerados, ubicado 12.4 puntos por debajo del promedio general.

Figura 4-54



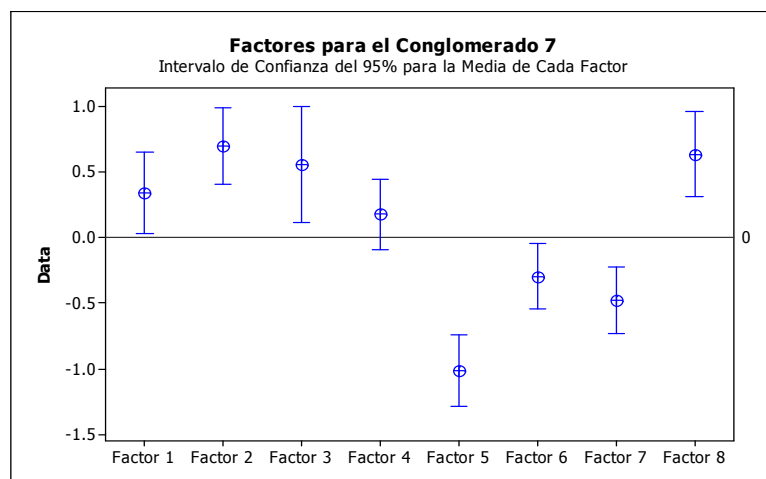
Este grupo se caracteriza principalmente por su bajo puntaje en la actitud hacia la matemática. El 70% de estos estudiantes tiene poco interés por estudiar el curso, el 71.4% dice tener poco gusto por esta materia. El 62% evalúa su capacidad e exigencias para estudiar matemáticas como regular. El factor 2, “*características del profesor y del examen*”, también sobresale en este grupo. El 57% considera que el examen estuvo muy difícil y que el profesor cubrió muy rápido los temas evaluados. El 62% considera que el profesor fue poco claro en sus explicaciones. El 38% del grupo, el porcentaje más alto entre los conglomerados, dice que se evaluó demasiado material. El 19% ha tomado tres o más veces el curso mientras que el 4.3% lo ha hecho exactamente dos veces. El 33.3% obtuvo D, F o W en su curso anterior de matemáticas. Este conglomerado también se caracteriza por su alta puntuación en el factor repaso del trabajo en clase. Todos los estudiantes afirman haber practicado con ejercicios de

la libreta, el 57% repasa siempre las notas de clase al terminar cada sección, los integrantes de este grupo asistieron a todas las clases. Se observa un puntaje elevado para el factor 7, “*entorno social*” y un puntaje bajo para el factor “*actitud hacia el curso*”. El 85.7% tiene una beca por parte de la universidad, 90% dice tener libro de texto, 66.7% no tiene un familiar cercano con, al menos, un grado universitario. El 71.4% siempre realiza ejercicios faltando poco tiempo para el examen y el 23.8% lo hace alguna vez. El 47.6% estudia con una o dos personas, el 43% no estudia en grupo. Se resalta del factor 5, que el 43% no uso exámenes de semestres anteriores en la preparación del examen.

Conglomerado 7

Este grupo está formado por 34 estudiantes, el 53.9% proviene de escuela pública. El promedio general de este grupo en el primer examen fue 48.9, el segundo más bajo, ubicado a 9.12 unidades por debajo del promedio general.

Figura 4-55

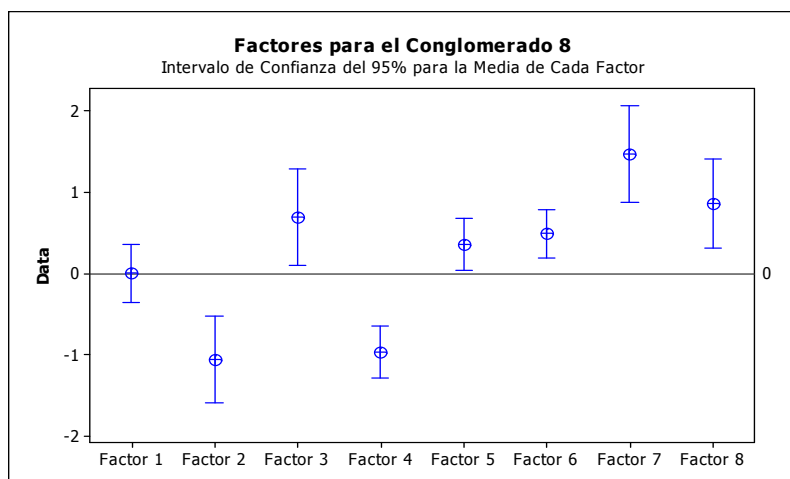


Este conglomerado se caracteriza por tener un puntaje muy bajo en el factor “*estudio en grupo*”, el 73.5% de los estudiantes estudió los temas del examen solo, el 26.5% no usó exámenes de semestres anteriores, el 44.1% considera que el tiempo para resolver el examen

fue poco, todos los estudiantes afirman que su desempeño en la Escuela Superior fue bueno o excelente. El 14.7% resuelve ejercicios sin leer la teoría. El segundo valor más bajo corresponde al factor de *entorno social*. El 61.8% no tiene una beca de la universidad, todos los estudiantes tienen libro de texto, el 82.4% tiene al menos un familiar cercano con por lo menos un grado de bachillerato. El 79.4% no pregunta en clase. El factor, características del profesor y del examen también sobresale en este grupo; el 50% encontró el examen muy difícil, este mismo porcentaje considera que el profesor cubrió los temas del examen muy rápido, el 70.6% consideró poco el tiempo de la clase para cubrir los temas y el 56% consideró que el profesor fue poco claro durante sus explicaciones. El 85.3% consideró poca la cantidad de ejercicios resueltos en clase. Algunas características encontradas en el factor 8 son: el 44% nunca memorizó ejercicios, el 47% hizo ejercicios del prontuario faltando poco tiempo para el examen, el 47% estudia con una o dos personas y el 38.2% lo hizo con tres o más, el 23.5% considera que el curso no tiene ninguna importancia para su formación profesional.

Conglomerado 8

Figura 4-56



Este conglomerado está formado por 13 personas, el 76.9% proviene de escuela pública. El promedio general para el primer examen fue de 70.2 ubicado 12.2 puntos por encima del promedio general.

Este conglomerado se caracteriza principalmente por tener un puntaje alto en el factor denominado “*entorno social*”. El 92.3% tiene beca por parte de la universidad, el 61.5% no tiene libro de texto, el 77% tiene al menos un familiar cercano con por lo menos un grado de bachillerato. También se caracteriza por evaluar positivamente las características del profesor y del examen y por el puntaje bajo en esfuerzo y tiempo dedicado (factor 4). El 38.5% no encontró dificultad en el examen y el 61.5% restante encontró poca dificultad. El 84.6% consideró que la rapidez con la que se cubrieron los temas fue regular, el 76.9% consideró que el tiempo de la clase para cubrir los temas fue suficiente, 84% de los estudiantes de este grupo afirma que el profesor fue muy claro en la explicación de los temas evaluados, 84.6% y 92.3% consideran que la cantidad de ejercicios resueltos en clase y la cantidad de temas evaluados respectivamente fueron suficientes. Ningún estudiante de este grupo ha pensado darse de baja. El 92.3% no desarrolla ejercicios al terminar cada sección, el 30% no resolvió ningún ejercicio propuesto en el prontuario, el 61.5% considera que su esfuerzo para obtener una buena nota fue regular, el 84.6% siempre repasó las notas de clase faltando poco tiempo para el examen, ningún estudiante fue a la oficina del profesor, el 77% no asistió al Centro de Apoyo y el 23% restante no respondió la pregunta.

4.4 Selección de conglomerados por el nivel de correlación entre los factores y las puntuaciones obtenidas en el primer examen

Con el objetivo de “refinar” la clasificación, se seleccionaron los factores que están más correlacionados con la puntuación del primer examen. Los resultados obtenidos motivaron a que se utilizaran los factores seleccionados como variables para caracterizar a cada estudiante y realizar un nuevo análisis de clasificación no supervisada.

La Tabla 4-7 muestra el coeficiente de correlación de Spearman con su respectivo valor p entre los factores y la puntuación del primer examen.

Tabla 4-7 Coeficientes de Correlación de Spearman

Rho de Spearman		Exa_1	F1	F2	F3	F4	F5	F6	F7	F8
Exa_1	Coefic	1.000	0.239	-0.444	-0.248	0.073	-0.026	0.020	0.067	-0.024
	p-value	.	0.000	0.000	0.000	0.246	0.682	0.754	0.285	0.699

De acuerdo a los resultados obtenidos, los factores que están más correlacionados con la nota del primer examen son: el factor 1 “*Actitud hacia las matemáticas*”, el factor 2 “*Características del profesor y del examen*” y el factor 3 “*Rezago académico*”.

El método que se usó en la clasificación fue el de k-mediodes (PAM), ya que fue el más consistente en el análisis anterior. De nuevo se usaron los índices de Dunn y Davies-Bouldin y el ancho promedio de silueta para validar el número óptimo de conglomerados. Los resultados de la validación pueden observarse en la Tabla 4-8. Nótese que el índice de Dunn sugiere escoger 5 conglomerados y los otros dos criterios sugieren escoger 4. Sin embargo, la diferencia entre el índice de 5 conglomerados y 4 es muy pequeña, así que se optó por tomar 4 conglomerados como número óptimo.

Tabla 4-8 Clasificación por el Método de k-medioides

Número de Conglomerados	Índice de Validación			Número de estudiantes por conglomerado									
	Dunn	DB	Silueta										
2	0.054	1.803	0.253	183	82								
3	0.027	1.471	0.228	87	100	78							
4	0.052	1.174	0.261	81	50	80	54						
5	0.053	1.195	0.251	71	48	69	45	32					
6	0.047	1.294	0.211	58	46	43	43	39	36				
7	0.048	1.325	0.216	27	32	42	54	42	40	28			
8	0.033	1.218	0.225	52	29	42	19	41	32	24	26		
9	0.058	1.135	0.234	38	21	47	20	36	23	27	36	17	
10	0.047	1.148	0.240	46	21	42	17	24	31	36	11	23	14

Figura 4-57

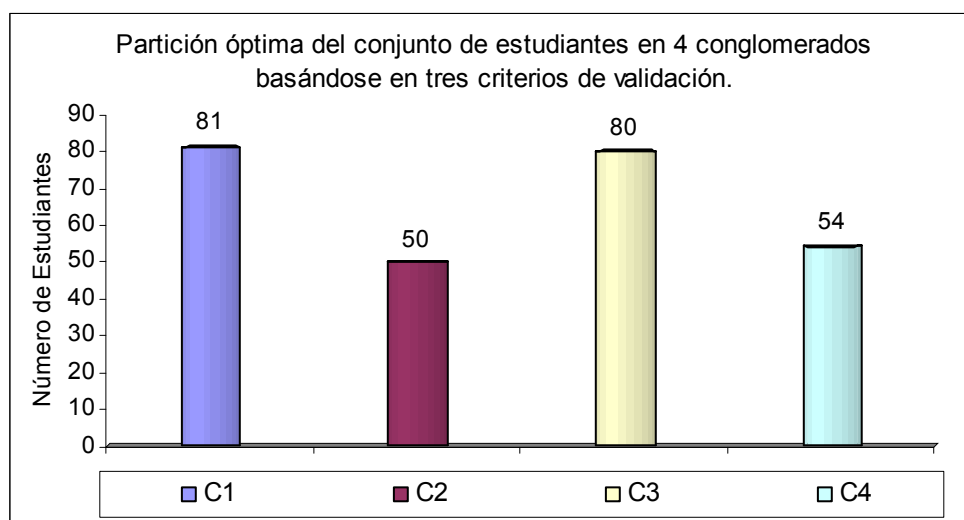


Tabla 4-9 Centroides de los Factores por Conglomerados Seleccionados.

Conglomerados	Centroides		
	F1	F2	F3
C1	0.576	0.835	-0.213
C2	-0.024	-0.044	1.544
C3	0.361	-0.977	-0.421
C4	-1.376	0.236	-0.487

Figura 4-58

Descripción de los cuatro conglomerados por los factores seleccionados

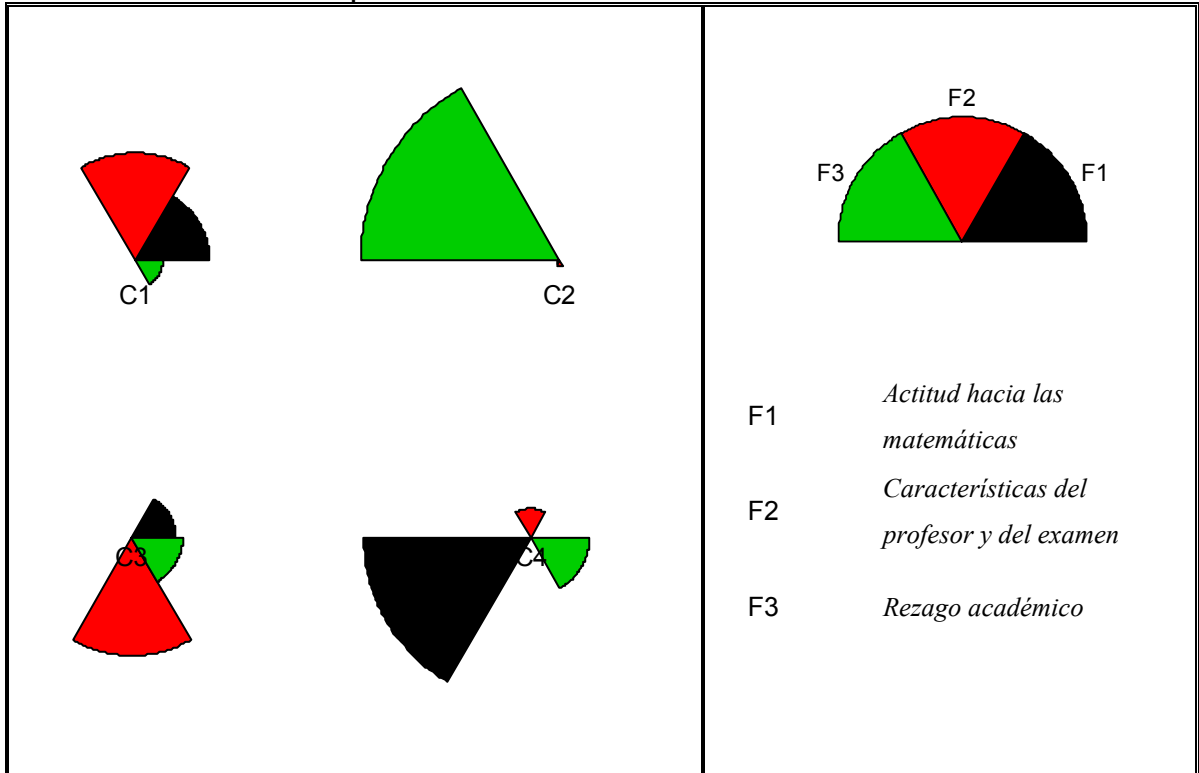
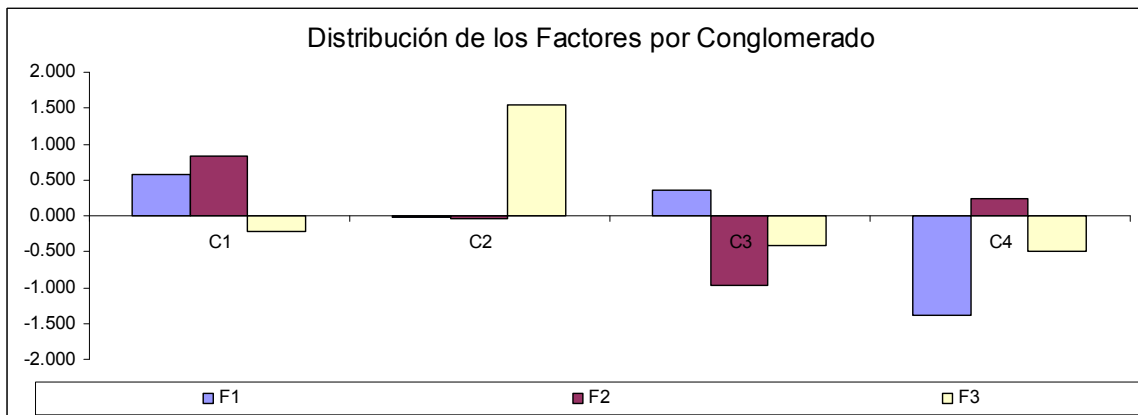
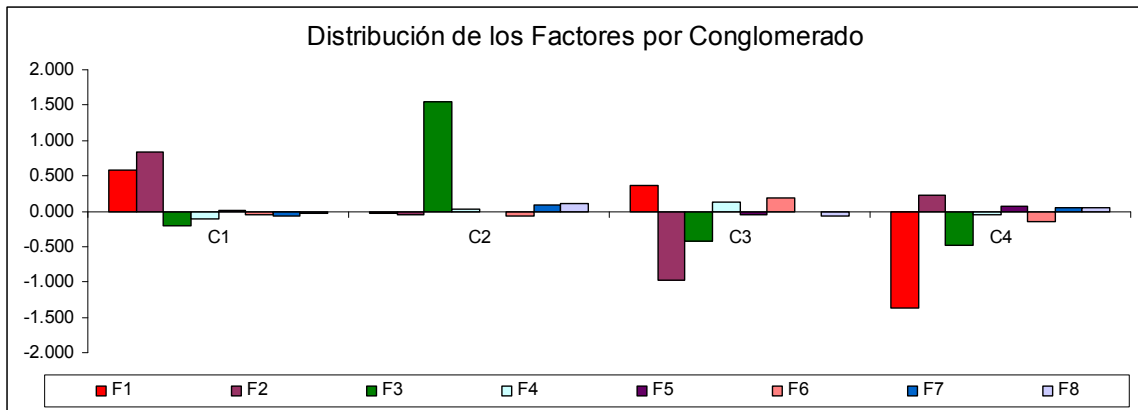


Figura 4-59



Se introdujeron los factores que no fueron tenidos en cuenta en la clasificación, en los cuatro conglomerados, para verificar que su influencia es prácticamente nula.

Figura 4-60



Lo anterior queda mejor expuesto al graficar intervalos de confianza del 95% para la media de todos los factores en cada conglomerado. Las gráficas pueden observarse en el Anejo 8.

4.4.1 Distribución de las variables en los nuevos conglomerados

Igual que en la clasificación anterior, presentamos a continuación una descripción porcentual de estudiantes por facultad, género y escuela de procedencia.

Figura 4-61

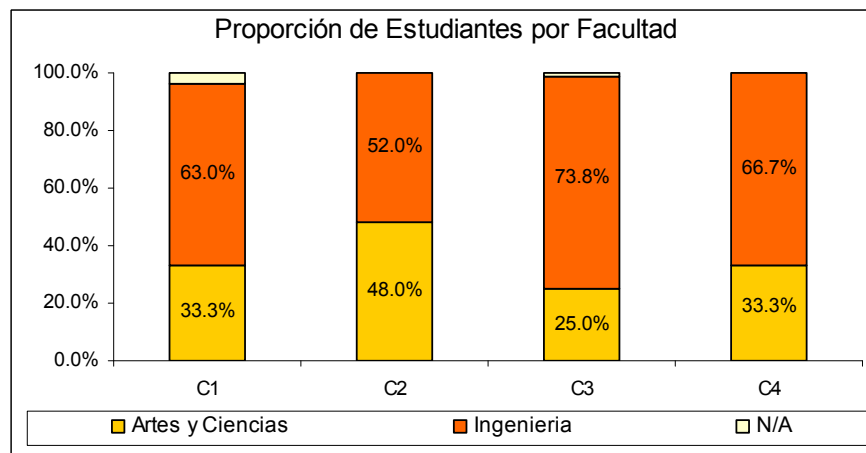


Figura 4-62

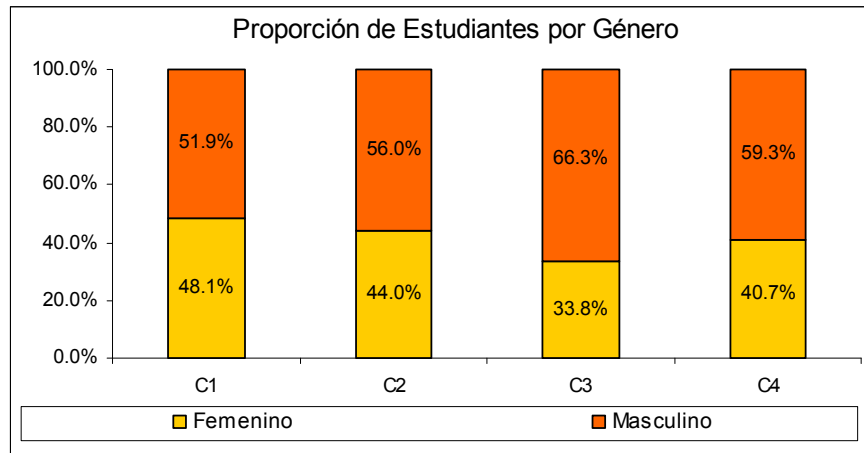


Figura 4-63

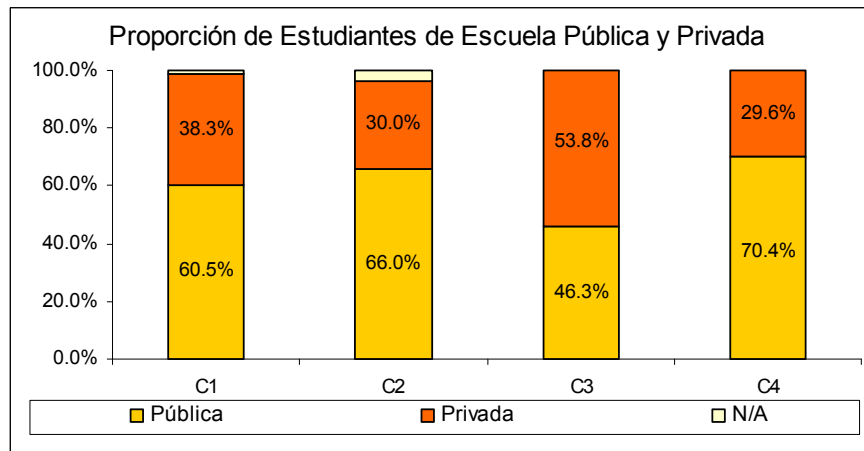
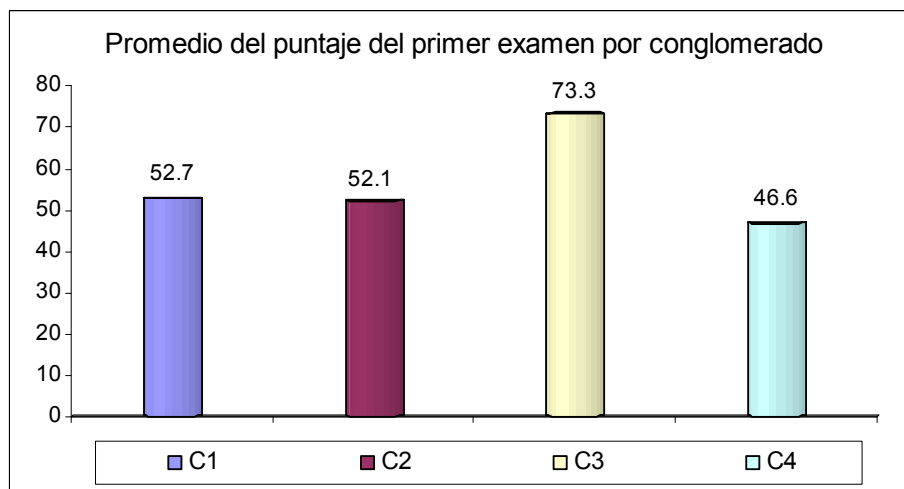


Figura 4-64



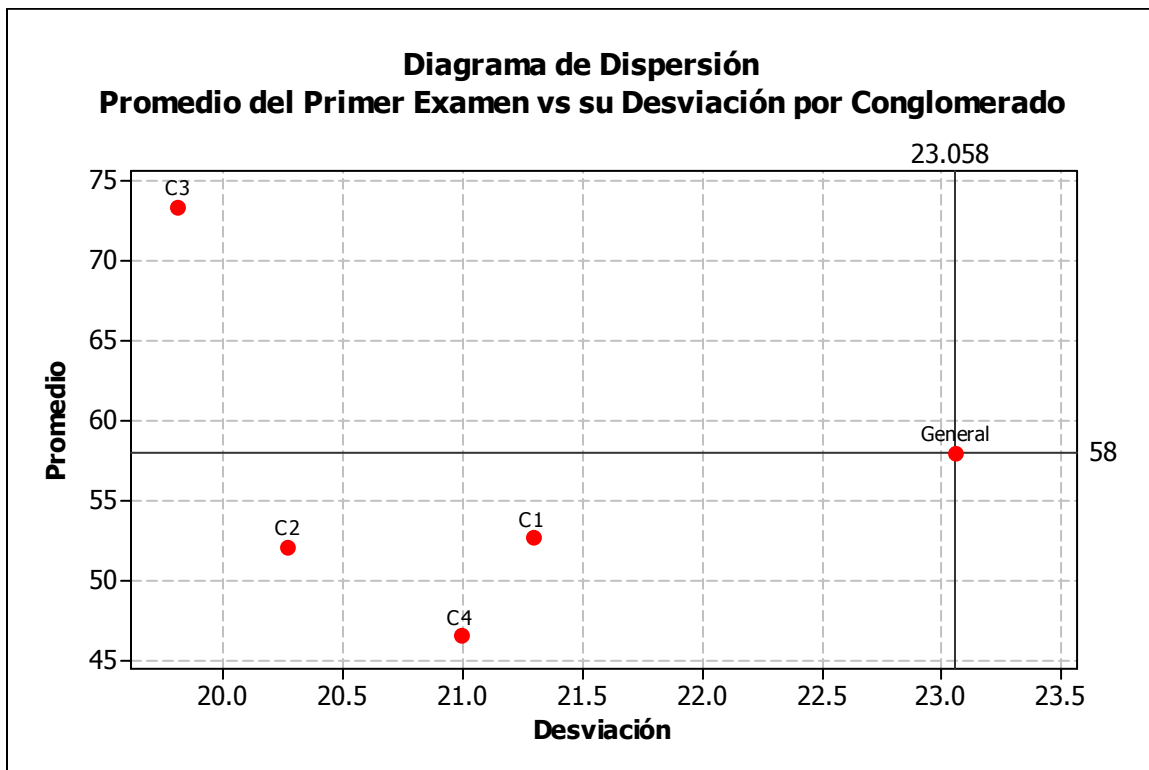
La prueba de Kruskal Wallis, Tabla 4-10 para la comparación de medias, indica que existe al menos una diferencia estadísticamente significativa entre los promedios de los ocho grupos.

Tabla 4-10 Prueba de Kruskal- Wallis Sobre la Puntuación del Examen

Conglomerados	N	Mediana	Rango Promedio	Z
1	79	53.0	112.8	-2.26
2	48	56.5	110.9	-1.83
3	77	78.0	178.9	7.14
4	52	46.5	93.9	-3.77
Total	256		128.5	
H = 53.25 DF = 3 P = 0.000				

En la Figura 4-65, puede observarse la desviación estándar de la nota y el promedio de cada conglomerado incluyendo el grupo formado por los 265 estudiantes llamado General.

Figura 4-65



4.4.1.1 Representación gráfica de la distribución porcentual de los estudiantes por respuesta a los ítems relacionados con cada factor

Con el fin de facilitar la descripción de los cuatro conglomerados, se muestra a continuación, la proporción de estudiantes por respuesta ante los ítems que están más correlacionados con cada uno de los tres factores, por conglomerado.

Figura 4-66

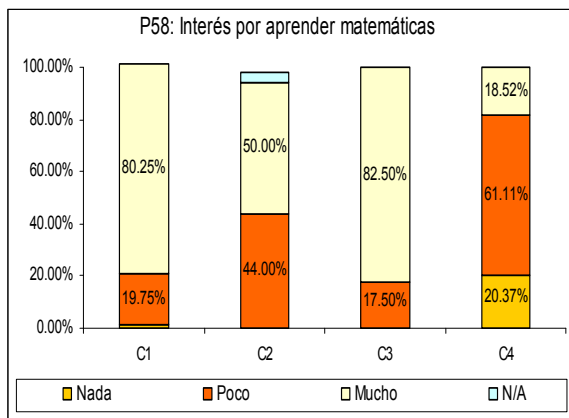


Figura 4-67

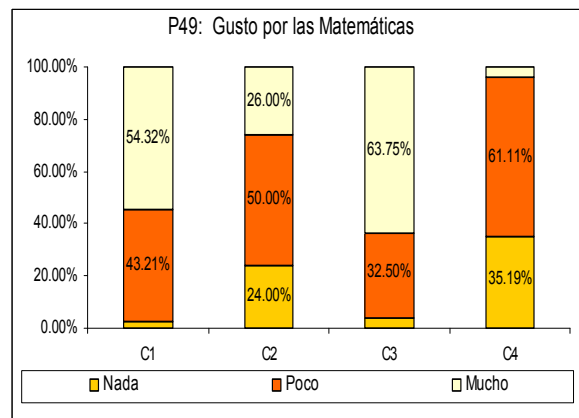


Figura 4-68

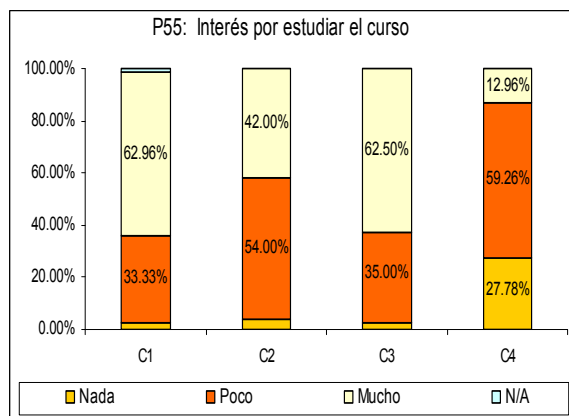


Figura 4-69

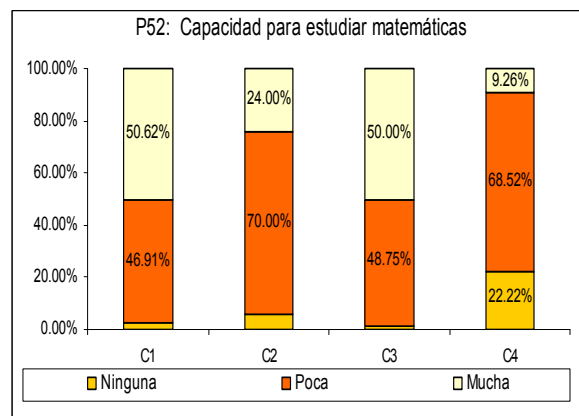


Figura 4-70

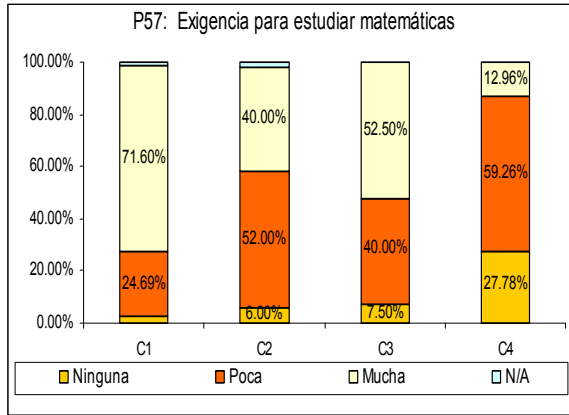


Figura 4-73

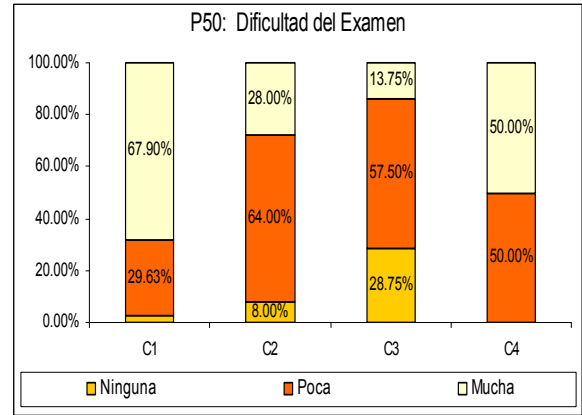


Figura 4-71

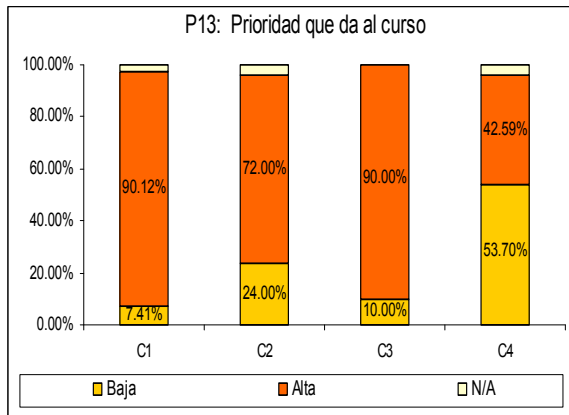


Figura 4-74

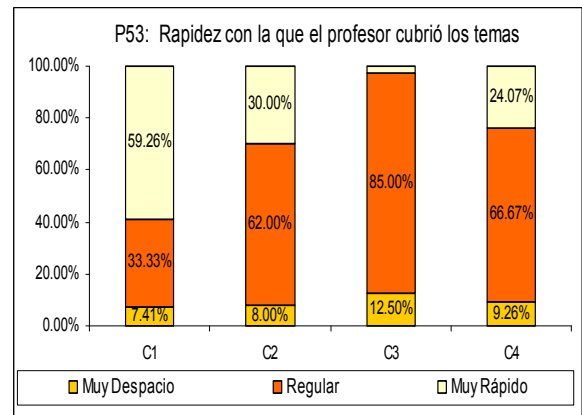


Figura 4-72

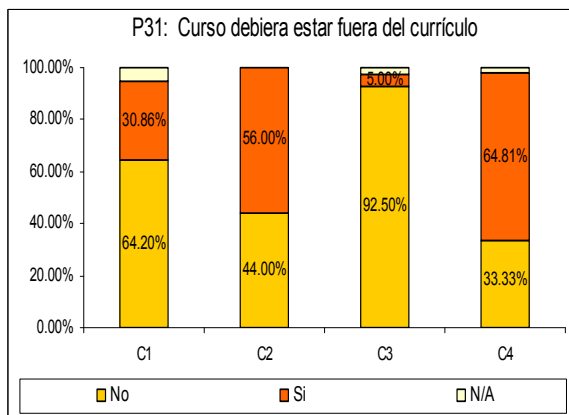


Figura 4-75

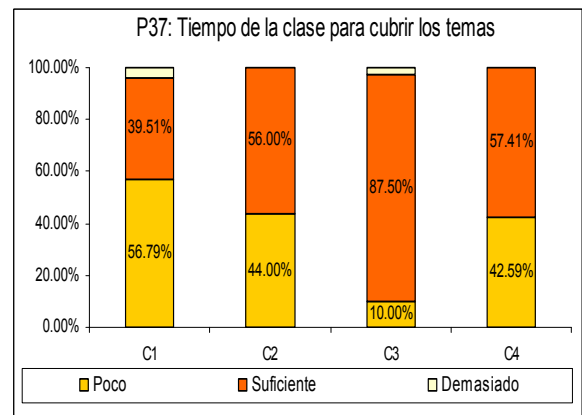


Figura 4-76

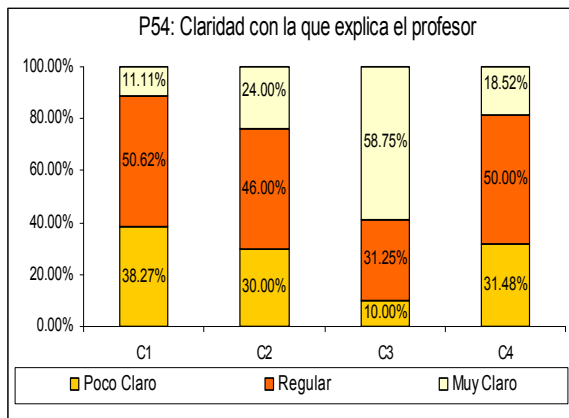


Figura 4-79

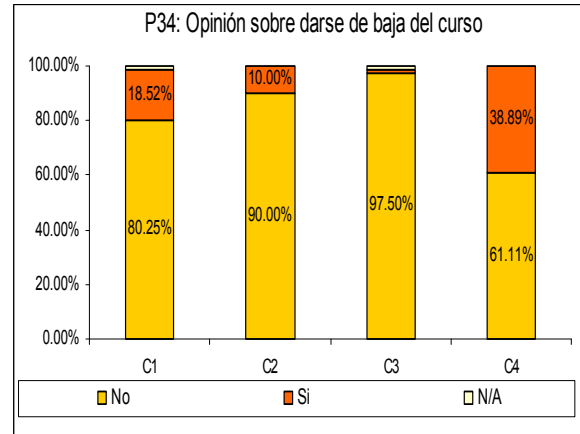


Figura 4-77

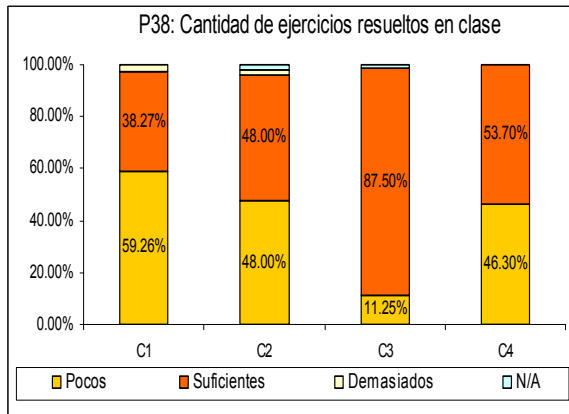


Figura 4-80

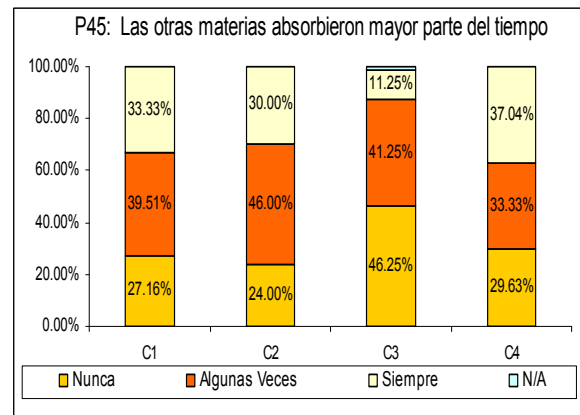


Figura 4-78

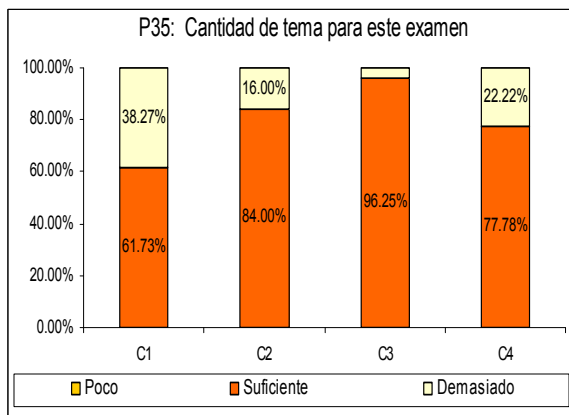


Figura 4-81

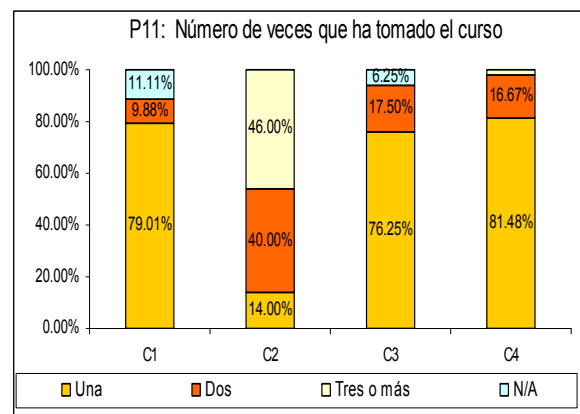


Figura 4-82

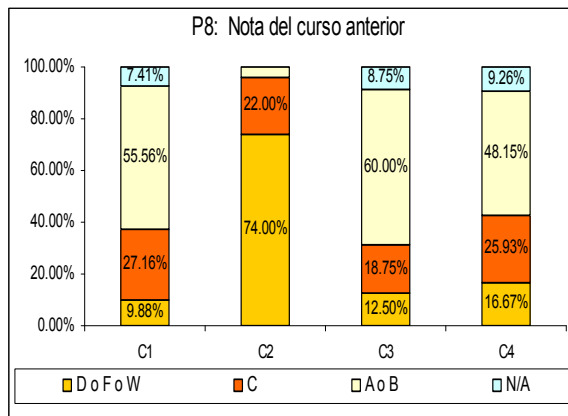
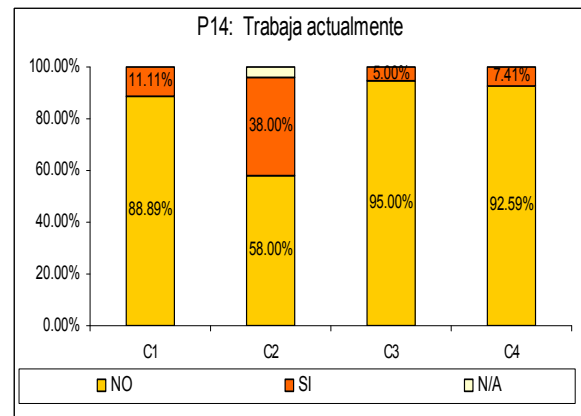


Figura 4-83



4.4.2 Descripción de los nuevos conglomerados

Conglomerado 1

Este conglomerado está integrado por 81 estudiantes, el 60.5% proviene de escuela pública y el 63% pertenece a la facultad de ingeniería. El promedio del grupo en el examen 1 fue de 52.7 ubicado 5.3 puntos por debajo del promedio general.

Este grupo se caracteriza principalmente por el puntaje obtenido en el factor “*Características del profesor y del examen*”. El 68% de los estudiantes dice que el examen estuvo muy difícil. El 61.7% considera que la cantidad de tema evaluado fue suficiente y el 38.3% piensa que fue demasiado. El 18.5% ha pensado en darse de baja del curso. El 27.2% dice que nunca las otras materias absorbieron la mayor parte del tiempo, el 39.5% dice que algunas veces y el 33.3% dice que siempre. El 59.3% afirma que el profesor cubrió los temas del examen muy rápido, el 33.3% considera que los temas se cubrieron con una velocidad regular. El 56.8% considera que hubo poco tiempo en la clase para cubrir los temas del examen y el 39.5% considera que fue suficiente. El 38.3% opina que el profesor fue poco claro durante la

explicación de los temas evaluados, el 50.6% considera que fue regular y el 11.1% considera que fue muy claro. El 59.3% considera que se resolvieron pocos ejercicios y el 38.3% considera que fueron suficientes.

Los estudiantes que pertenecen a este grupo también se caracterizan por tener una buena *actitud hacia el estudio de las matemáticas*. El 63% dice tener mucho interés por estudiar el curso y el 33.33% poco. El 54.32% siente mucho gusto por la matemáticas y el 43.2% siente poco gusto. El 50.6% considera que tiene mucha capacidad para estudiar matemáticas mientras que el 46.9% considera que tiene poca, el 71.6 afirma que se exige mucho en el estudio de esta área y el 90.1% da alta prioridad al curso de Cálculo I.

El 79% de los estudiantes que pertenece a este conglomerado está tomando el curso por primera vez y el 9.9% lo toma por segunda vez.

Conglomerado 2

Este conglomerado está integrado por 50 estudiantes, el 66% proviene de escuela pública y el 52% pertenece a la facultad de ingeniería. El promedio de la nota del primer examen en este grupo es de 52.1, ubicado 5.9 puntos por debajo del promedio general.

Este conglomerado se caracteriza principalmente por su alto puntaje en el factor “*Rezago académico*”. El 14% de los estudiantes está tomando por primera vez el curso, el 40% lo toma por segunda vez y 46% lo toma por tercera o más veces. El 74% obtuvo una calificación de D, F o W (se dio de baja) en el curso anterior que tomó de matemáticas, el 22% obtuvo C. El 38% trabaja en el período previo al primer examen (estos estudiantes laboraron durante las 4 semanas previas al primer examen). Este grupo tiene, con un promedio de 13.5 créditos, el menor número de créditos matriculados.

Conglomerado 3

Este conglomerado está formado por 80 estudiantes, el 53.8% proviene de escuela privada y el 73.8% pertenece a la facultad de ingeniería. El promedio del grupo en la nota del primer examen es de 73.3, ubicada 15.3 puntos por encima del promedio general.

Este conglomerado se caracteriza principalmente por su puntaje en el factor “*Características del profesor y del examen*”. El 28.8% no encontró ninguna dificultad en el examen, el 57.5% encontró poca y el 13.7% encontró mucha dificultad. El 96.3% dice que la cantidad de tema en el primer examen fue suficiente. El 97.5% no ha pensado en darse de baja del curso. El 46.3% afirma que las otras materias nunca absorbieron la mayor parte del tiempo, el 41.3% considera que alguna veces y el 11.3% considera que siempre. El 12.5% considera que el profesor cubrió los temas del examen muy despacio y el 85% considera que la velocidad con la que el profesor cubrió los temas fue regular. El 87.5% considera que fue suficiente el tiempo de la clase para cubrir los temas evaluados, mientras que el 10% considera que fue poco. El 58.76% considera que el profesor fue muy claro durante las explicaciones de los temas evaluados, el 31.3% considera que sus explicaciones fueron regulares y el 10% afirma que fue poco claro. El 87.5% considera suficientes la cantidad de ejercicios resueltos en clase y el 11.25% considera que se resolvieron pocos.

Este grupo también se caracteriza por la actitud que tiene hacia las matemáticas. El 82.5% dice tener mucho interés por aprender matemáticas. El 62.5% afirma que tiene mucho interés por estudiar el curso. El 63.8% siente mucho gusto por las matemáticas y el 32.5% siente poco agrado por esta área. El 50% considera que tiene una buena capacidad para estudiar matemáticas mientras que el 48.8% considera que tiene poca. El 52.5% afirma que se exige mucho para estudiar matemáticas y el 90% da una prioridad alta al curso de cálculo I.

El 76.3% de los estudiantes que pertenece a este conglomerado está tomando el curso por primera vez y el 17.5% lo toma por segunda vez.

Conglomerado 4

Este conglomerado esta formado por 54 estudiantes, el 70.4% proviene de escuela pública y el 66.7% pertenece a la facultad de ingeniería. El promedio de este grupo en la nota del primer examen es de 46.6, ubicado 11.4 puntos por debajo del promedio general.

Este conglomerado se caracteriza principalmente por tener una baja “*Actitud hacia las matemáticas*”. El 20.4% no tiene interés por aprender matemáticas, el 61.1% tiene poco. Sólo el 18.5% dice tener mucho interés por aprender esta materia. El 35.19% dice no tener gusto por las matemáticas y 61.1% dice tener mucho gusto por esta área. El 22.2% dice no tener ninguna capacidad para estudiar matemáticas, el 68.5% dice tener poca capacidad mientras que el 9.3% dice tener mucha. El 27.8% considera que no se exige para estudiar matemáticas, el 59.3% considera que se exige poco y el 13% considera que se exige mucho. El 53.7% de estos estudiantes da una baja prioridad al curso del Calculo I. El 64.8% considera que el curso debería estar fuera de su currículo.

Algunas características del factor “*Características del profesor y del examen*” son las siguientes: El 50% de los estudiantes encontró poca dificultad en el examen y el 50% restante consideró que el examen estaba muy difícil. El 22.2% piensa que se evaluaron demasiados temas y el 38.9% ha pensado en darse de baja del curso del Cálculo I. El 37% considera que las otras materias siempre absorbieron la mayor parte del tiempo. El 66.7% considera que el profesor cubrió los temas con una rapidez regular, el 24% considera que fue muy rápido y el 9.3% que fue muy despacio. El 57.4% afirma que hubo suficiente tiempo para cubrir los

temas del examen, mientras que el 42.6% considera que el tiempo de la clase para cubrir los temas fue poco. El 31.5% dice que el profesor fue poco claro durante las explicaciones de los temas evaluados, el 50% piensa que su explicación fue regular y el 18.5% que fue muy claro. El 46.3% considera que en la clase se hicieron pocos ejercicios mientras que 53.7% considera que fueron suficientes.

El 81.5% de los estudiantes que pertenece a este conglomerado esta tomando el curso por primera vez y el 16.8% lo toma por segunda vez.

5. ANÁLISIS DE DATOS EN PANELES

5.1 *Introducción*

El diseño de datos en panel obtiene información de los mismos individuos, sujetos o unidades de análisis en varios momentos del tiempo, lo cual permite observar cambios a nivel individual, situación que no es tenida en cuenta utilizando modelos transversales.

Se aplicó una técnica de análisis de datos por paneles para estudiar el efecto del cambio en el tiempo de variables cualitativas y cuantitativas que expliquen el rendimiento académico de los estudiantes en el curso de Cálculo I. Se utilizaron como variables predictoras las variables presentadas en la Tabla 2-1 además se adicionaron las variables “facultad” y “género”. La variable respuesta fue la nota obtenida en cada examen parcial, ésta se categorizó en 5 niveles, a saber: Si la nota toma un valor entre 0 y 59 se asigna a la categoría 0, si toma un valor entre 60 y 69 se asigna a la categoría 1, si toma un valor entre 70 y 79 se asigna a la categoría 2, si toma un valor entre 80 y 89 se asigna a la categoría 3 y si toma un valor entre 90 y 100 o más se asigna a la categoría 4.

Hasta el momento son pocos los programas estadísticos y econométricos que incluyen en sus procedimientos funciones para estimar modelos de datos por paneles con respuesta multinomial. Para este análisis se utilizó el procedimiento GLIMMIX de SAS [35].

5.2 Procedimiento

Se estimaron tres modelos logit acumulativos con interceptos aleatorios; en los dos primeros se utilizó la muestra obtenida por los 124 estudiantes que respondieron a los tres cuestionarios, en el tercer modelo se utilizó la información contenida en todos los cuestionarios. En el primer modelo, Modelo I, se estimaron los coeficientes de las variables sin tener en cuenta el efecto del tiempo sobre la variable respuesta; este factor fue tenido en cuenta en los Modelos II y III. La especificación matemática del modelo I para N individuos con T mediciones en el tiempo, J categorías en la variable respuesta y un matriz de x de mediciones de las variables explicativas es la siguiente:

$$\text{logit}[P(y_{it} \leq j | u_i)] = u_i + \alpha_j + x'_{it} \beta \quad j = 1, \dots, J-1; \quad i = 1, \dots, N; \quad t = 1, \dots, T$$

donde y_{it} denota la categoría respuesta del individuo i en el tiempo t , y β es un vector de parámetros de efectos fijos asociados a cada variable predictora.

5.2.1 Estimación del Modelo I

Para escoger el mejor subconjunto de variables explicativas se aplicó el método de *selección hacia adelante* (forward). En cada paso “entraba” al modelo la variable que tenía el menor valor p (correspondiente a la prueba F). El método continuaba hasta no encontrar una variable cuyo valor p fuera menor a 0.05. En la Tabla 5-1 se presentan las variables que resultaron significativas.

Las únicas variables “fijas” (que no cambian con el tiempo) que han permanecido en el modelo son P8 y P11; además, P8 es la variable más significativa del modelo, Tabla 5-3.

Tabla 5-1 Descripción del Mejor Subconjunto de Variables-Modelo I

VARIABLE	DESCRIPCIÓN	Niveles	Valores
P8	Nota del curso anterior que tomó de matemáticas	3	1 2 3
P45	Las otras materias absorbieron mayor parte del tiempo	3	1 2 3
P50	Dificultad del examen	3	1 2 3
P34	Darse de baja	2	1 2
P53	Rapidez con la que el profesor cubrió los temas	3	1 2 3
P54	Claridad con la que explica el profesor	3	1 2 3
P11	Número de veces que ha tomado el curso	3	1 2 3
P22	Ir a la oficina del profesor	2	1 2
P47	Créditos matriculados	Continua	

Tabla 5-2 Coeficientes Estimados para el Mejor Subconjunto de Variables Utilizando el Modelo I

Efecto	Examen	P8	P45	P50	P34	P53	P54	P11	P22	Coefic	Error Estándar
Intercepto	0									3.557	1.596
Intercepto	1									4.582	1.605
Intercepto	2									5.607	1.615
Intercepto	3									7.305	1.634
P8		1								1.200	0.465
P8		2								2.595	0.456
P8		3								0.000	.
P45			1							-1.442	0.369
P45			2							-0.871	0.334
P45			3							0.000	.
P50				1						-1.292	0.434
P50				2						-0.952	0.276
P50				3						0.000	.
P34					1					-1.034	0.434
P34					2					0.000	.
P53						1				-1.589	0.545
P53						2				-0.910	0.366
P53						3				0.000	.
P54							1			1.316	0.413
P54							2			0.445	0.290
P54							3			0.000	.
P47										-0.227	0.082
P11								1		0.806	0.663
P11								2		-0.324	0.643
P11								3		0.000	.
P22									1	0.639	0.301
P22									2	0.000	.

Tabla 5-3 Test Tipo I de Efectos Fijos

Efectos	DF	Valor F	Pr > F
P8	2	16.2	<.0001
P45	2	7.63	0.0006
P50	2	7.17	0.0010
P34	1	5.69	0.0179
P53	2	4.94	0.0080
P54	2	5.09	0.0070
P47	1	7.67	0.0061
P11	2	3.58	0.0298
P22	1	4.51	0.0349

5.2.2 Estimación del Modelo II

Al igual que en el modelo I, se realizó el método de *selección hacia adelante*, con el fin de obtener el mejor subconjunto de variables explicativas. Se observa, Tabla 5-4, que el efecto del tiempo es significativo.

Tabla 5-4 Test Tipo II de Efectos Fijos

Efectos	DF	Valor F	Pr > F
Tiempo	2	4.79	0.0092
P8	2	17.4	<.0001
P45	2	7.27	0.0009
P50	2	4.99	0.0076
P34	1	4.39	0.0373
P53	2	5.14	0.0066
P54	2	5.22	0.0062
P47	1	6.86	0.0095
P11	2	3.39	0.0356

Tabla 5-5 Coeficientes Estimados para el Mejor Subconjunto de Variables Utilizando el Modelo II

Efecto	Examen	T	P8	P45	P50	P34	P53	P54	P11	Coefic	Error Estándar
Intercepto	0									4.061	1.600
Intercepto	1									5.077	1.609
Intercepto	2									6.110	1.620
Intercepto	3									7.831	1.641
Time		1								-0.269	0.286
Time		2								-0.847	0.285
Time		3								0	.
P8			1							1.321	0.472
P8			2							2.739	0.465
P8			3							0	.
P45				1						-1.419	0.372
P45				2						-0.839	0.339
P45				3						0	.
P50					1					-1.062	0.442
P50					2					-0.823	0.280
P50					3					0	.
P34						1				-0.927	0.442
P34						2				0.000	.
P53							1			-1.671	0.554
P53							2			-0.908	0.369
P53							3			0	.
P54								1		1.363	0.422
P54								2		0.445	0.292
P54								3		0	.
P47										-0.218	0.083
P11									2	0.740	0.672
P11									3	-0.380	0.654
P11									4	0	.

Para visualizar el cambio de probabilidad de que un estudiantes obtenga una nota en particular a través del tiempo, en este caso en los tres períodos correspondientes a cada examen parcial, se estimaron dichas probabilidades para un estudiante *típico*, $u_i = 0$, donde todas las categorías de las variables explicativas se han fijado en el valor más a lto, es decir, el efecto de las variables explicativas es cero y el número de créditos ha sido fijado en 10, 15 y 18.

Figura 5-1 Distribución de Probabilidad de la Nota para un Estudiante Típico con 10 Créditos en Cada Período

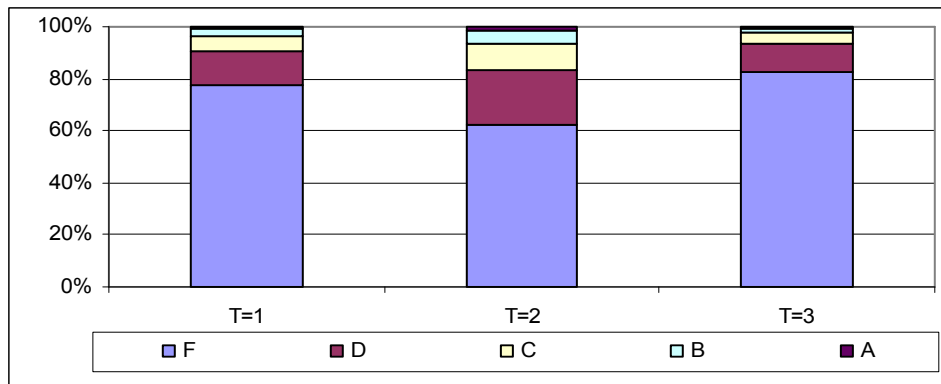


Figura 5-2 Distribución de Probabilidad de la Nota para un Estudiante Típico con 15 Créditos en Cada Período

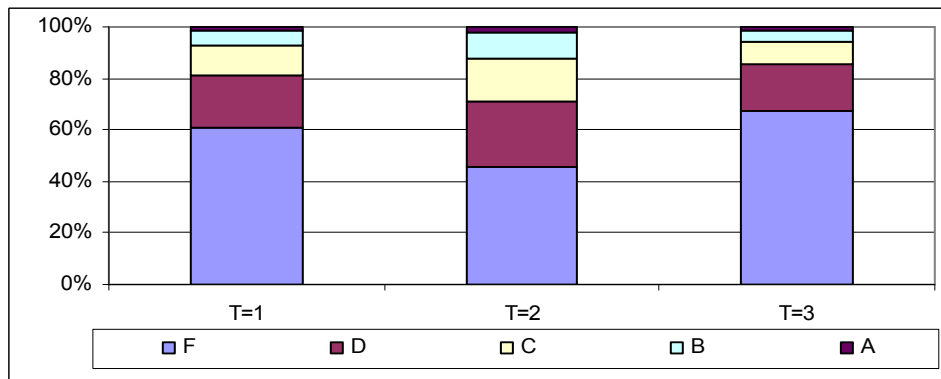
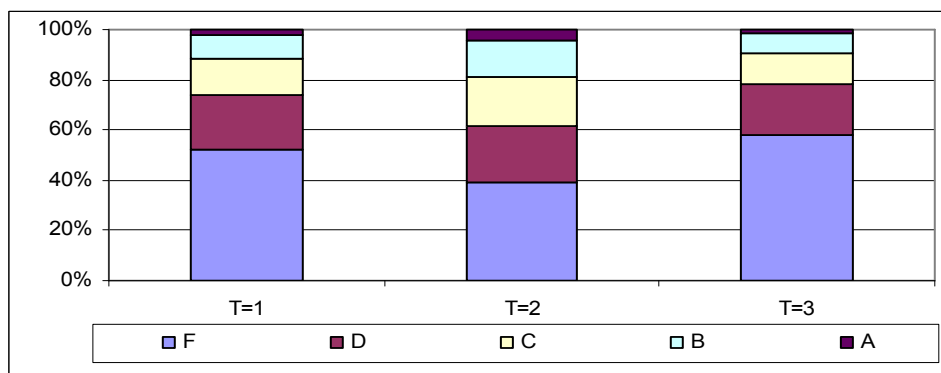


Figura 5-3 Distribución de Probabilidad de la Nota para un Estudiante Típico con 18 Créditos en Cada Período



Se observa que la probabilidad de obtener una mejor puntuación en cada examen parcial aumenta para los estudiantes que tienen un mayor número de créditos matriculados. Al fijar todas las categorías en su valor más alto en los tres períodos se puede apreciar el efecto del tiempo sobre las puntuaciones. Así, se observa que la probabilidad de obtener A o B en el segundo parcial es más alta que en el primer y tercer parcial.

5.2.2 Estimación del Modelo III

Las variables significativas que resultaron, después de realizar el método de *selección hacia adelante* son expuestas en la Tabla 5-6.

Tabla 5-6 Descripción del Mejor Subconjunto de Variables-Modelo III

VARIABLE	DESCRIPCIÓN	Niveles	Valores
P8	Nota del curso anterior que tomó de matemáticas	3	1 2 3
P45	Las otras materias absorbieron mayor parte del tiempo	3	1 2 3
P50	Dificultad del examen	3	1 2 3
P34	Darse de baja	2	1 2
P13	Prioridad que da al curso	2	1 2
P24	Ausencias	3	1 2 3
P54	Claridad con la que explica el profesor	3	1 2 3
P11	Número de veces que ha tomado el curso	3	1 2 3
P22	Ir a la oficina del profesor	2	1 2
P47	Créditos matriculados	Continua	

En este modelo la variable “Rapidez con la que el profesor cubrió los temas” (P53) no es significativa; si lo son las ausencias a clase (P24) y prioridad que el estudiante da al curso.

Tabla 5-7 Coeficientes Estimados para el Mejor Subconjunto de Variables Utilizando el Modelo III

Effect	Examen	Time	P8	P34	P45	P50	P54	P22	P24	P11	P13	Coeficie	Error Estandar
Intercepto	0											2.659	1.2489
Intercepto	1											3.6398	1.2546
Intercepto	2											4.6351	1.2612
Intercepto	3											6.03	1.2724
Time		1										0.1909	0.2586
Time		2										-0.4611	0.2549
Time		3										0	.
P8			1									1.1085	0.3599
P8			2									2.0751	0.3333
P8			3									0	.
P34				1								-1.2735	0.3344
P34				2								0	.
P45					1							-0.8209	0.2839
P45					2							-0.4172	0.2652
P45					3							0	.
P50						1						-1.6762	0.3604
P50						2						-1.0556	0.2207
P50						3						0	.
P54							1					1.179	0.3131
P54							2					0.5758	0.2321
P54							3					0	.
P47												-0.1652	0.06359
P22								1				0.6755	0.2522
P22								2				0	.
P24									1			-0.6325	0.2408
P24									2			0	.
P11										1		0.6293	0.4826
P11										2		-0.3237	0.4661
P11										3		0	.
P13											2	0.6824	0.302
P13											3	0	.

Tabla 5-8 Test Tipo III de Efectos Fijos

Effect	DF	F Value	Pr > F
Time	2	5.29	0.0056
P8	2	19.73	<.0001
P34	1	14.51	0.0002
P45	2	4.24	0.0154
P50	2	15.40	<.0001
P54	2	7.40	0.0007
P47	1	6.75	0.0099
P22	1	7.17	0.0079
P24	1	6.90	0.0091
P11	2	4.15	0.0168
P13	1	5.11	0.0246

5.3 Distribución de probabilidad para los estudiantes clasificados en cuatro conglomerados

Con el fin de comparar las distribuciones de probabilidades observadas y calculadas. Para obtener la distribución de probabilidad calculada se estimaron las probabilidades de obtener una nota en particular (F, D, C, B o A) para todos los individuos de cada conglomerado utilizando el modelo III. Cada estudiante se asignó a la categoría que tenía la probabilidad de ocurrencia más alta; luego se cálculo la proporción de estudiantes por categoría. Los resultados pueden observarse en las Figuras 5-4 y 5-5.

El modelo no asignó a ningún estudiante a la categoría D, por lo tanto, para visualizar los resultados de manera más clara se agruparon las categorías F Y D en la distribución de probabilidades observadas. Se observa que existen pocas diferencias entre las dos distribuciones.

Figura 5-4 Distribución de Probabilidad Calculada de las Puntuaciones del Primer Examen Para los Estudiantes Clasificados en Cuatro Conglomerados

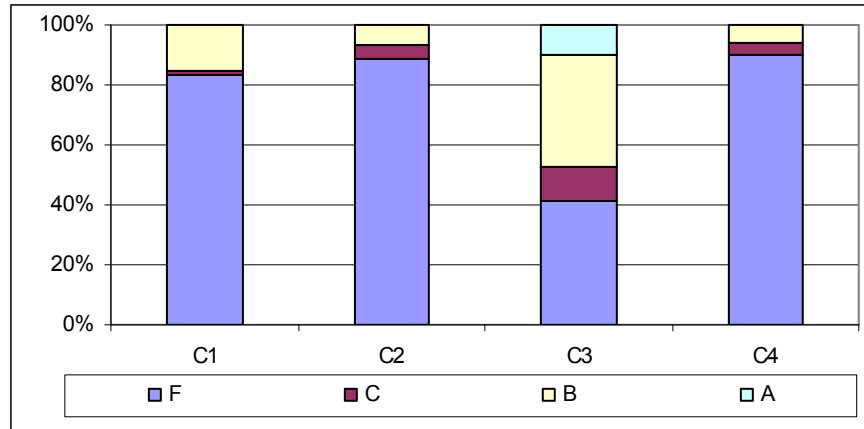
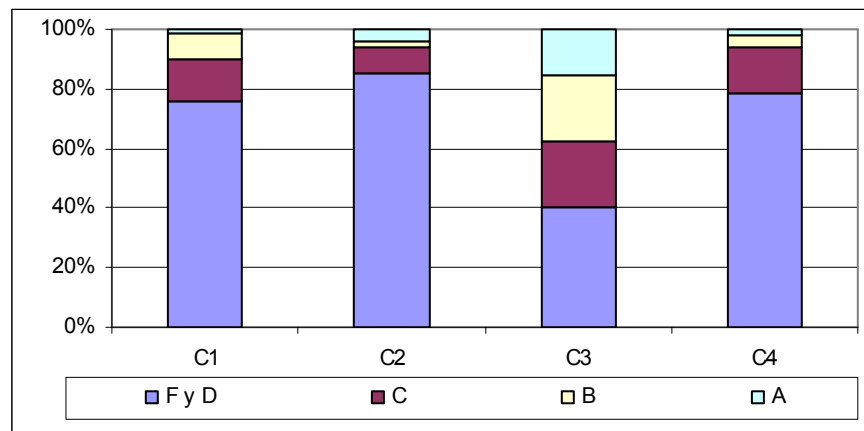


Figura 5-5 Distribución Porcentual Observada de las Puntuaciones del Primer Examen Para los Estudiantes Clasificados en Cuatro Conglomerados



6. CONCLUSIONES

Durante el desarrollo de este trabajo se ha llegado a las siguientes conclusiones:

Reducimos la dimensión original del conjunto de datos, a través del análisis factorial, a un conjunto de 8 variables subyacentes o factores que explicaron la configuración de las correlaciones encontradas. Estos factores describen características como actitud hacia las matemáticas, metodología del profesor, características del examen, rezago académico, esfuerzo y tiempo dedicado en la preparación para el examen, estudio en grupo, repaso del trabajo en clase, entorno social y la actitud hacia el curso.

Con las variables subyacentes encontradas en el análisis factorial se aplicó un análisis de clasificación no supervisado, cuya validación determinó que el número de óptimo de conglomerados claramente diferenciados es ocho. Este análisis permitió obtener una caracterización sobre las actitudes, hábitos y percepciones que los estudiantes que formaron parte de estudio tienen acerca de la enseñanza y aprendizaje de las matemáticas y en particular del curso del Cálculo I impartido en el Recinto Universitario de Mayagüez por el Departamento de Ciencias Matemáticas. La clasificación de los estudiantes en los ocho grupos permitió además obtener grupos más homogéneos con respecto a la nota del primer examen, variable no incluida en el análisis factorial.

Utilizando las gráficas de estrellas (modificado a sectores circulares), los intervalos de confianza del 95% para cada variable en cada uno de los ocho conglomerados y las gráficas

de barra que midieron la contribución de cada categoría de una variable en cada uno de los conglomerados, permitieron describir los ocho grupos, de los cuales podemos concluir:

Los grupos C3, C8 y C4 tienen, en orden descendente, los mejores promedios en el examen 1.

Los estudiantes que pertenecen a los grupos C3 y C4 sienten gusto por el estudio de las matemáticas, mientras que los estudiantes del grupo C8, en promedio, se mantienen neutrales en este factor. Los tres grupos de estudiantes valoran de manera positiva las características del examen y la metodología utilizada por el profesor. También se puede notar que estos estudiantes en general repasan las notas de clase y practican los ejercicios escritos en la libreta.

El grupo C8 está formado principalmente por estudiantes rezagados, la gran mayoría de los estudiantes tienen beca por parte del Recinto, la mayoría no tiene libro de texto y provienen de escuela pública. Lo anterior muestra que, posiblemente, son estudiantes de “pocos” recursos económicos. La principal actividad que hacen estos estudiantes durante la preparación del examen es memorizar y repasar los ejercicios de la libreta y de exámenes de semestres anteriores. No leen la teoría antes de desarrollar los ejercicios.

Los integrantes del grupo C3 se caracterizan por ser estudiantes con un buen desempeño en su curso. Durante la preparación del examen no estudian en grupo, a diferencia del grupo C4, no sólo tienen una buena actitud hacia el estudio de la matemática sino también hacia el curso. La mayoría nunca memoriza los ejercicios, consideran que el tiempo para resolver el examen fue suficiente. La nota en el curso anterior de matemáticas fue de C o más. El factor “entorno social” no es significativo en este grupo.

Los estudiantes del grupo C4 dedican mucho tiempo a la preparación del examen. Las otras características tienen, en promedio, valores bajos.

Los grupos C6 y C7 obtienen en orden ascendente las puntuaciones más bajas en el examen. La mayoría de estudiantes que integran el grupo C6 provienen de escuela pública, estos estudiantes muestran poco gusto e interés por las matemáticas, sin embargo tienen interés por estudiar el curso y dedican tiempo a practicar los ejercicios de la libreta y repasar las notas de clase al terminar cada sección. Evalúan de manera negativa las características del profesor y del examen. La mayoría de estudiantes toman el curso por primera vez y muchos no practicaron con exámenes de semestres anteriores.

Los estudiantes del grupo C7 se caracterizan por no estudiar en grupo, todos tienen libro de texto y más de la mitad no tienen beca. Evalúan de manera negativa las características del examen y del profesor. Muchos son estudiantes rezagados que muestran poco o ningún esfuerzo en la preparación del examen.

Los grupos C2, C1 y C5 obtuvieron un promedio por debajo del promedio general pero cercano a éste. Los estudiantes del grupo C2 dicen tener gusto e interés por las matemáticas y en particular por el curso, sin embargo el trabajo y esfuerzo en la preparación del examen es mínimo. El grupo C1 considera que ha hecho mucho esfuerzo en la preparación del examen, pero no realiza un trabajo continuo de repasar al terminar cada sección. Evalúa de manera negativa las características del profesor y del examen. Los estudiantes del grupo C5 no se preocupan por el curso, la mayoría no tiene gusto ni interés por el estudio de las matemáticas, el trabajo que realizan en clase es regular y el esfuerzo y tiempo dedicado por la mayoría es mínimo.

La prueba de correlación, utilizando el coeficiente rho de Spearman, indica que los factores que están más correlacionados con la nota del examen 1 son: la *“actitud hacia la matemática”*, las *“características del profesor y del examen”* y el *“rezago académico”*. Esta

prueba llevó a reducir más el número de conglomerados de ocho a cuatro, logrando una tipificación más concreta de la muestra de estudiantes.

Con respecto a la nueva re-agrupación de la muestra en cuatro conglomerados, se puede concluir que a la gran mayoría de estudiantes rezagados, clasificados en el grupo C2, les son indiferentes las características del profesor y del examen, tampoco tienen una actitud específica (gusto/disgusto) hacia el estudio de las matemáticas. Estos estudiantes tienen en general un bajo desempeño en el primer examen tendiendo con ello a permanecer en el rezago.

Los grupos C1 y C3 tienen prácticamente los mismos porcentajes de buena actitud hacia el estudio de las matemáticas, su principal diferencia se debe a la evaluación que hacen de las características del profesor y del tema evaluado. El grupo C3, con el mejor promedio, evalúa de manera positiva las características del profesor y del examen mientras que los estudiantes del grupo C1 evalúan de manera negativa estas características.

Dado que los estudiantes del grupo C3 obtuvieron el mejor promedio en la puntuación del examen y el promedio de los estudiantes del grupo C1 estuvo por debajo del promedio general, es posible que esas diferencias, en la nota promedio, se deban a las características del profesor. Aunque no se ha tenido en cuenta las habilidades y destrezas propias de cada estudiante.

Con relación al grupo C4 se puede observar que los estudiantes que tienen poco gusto e interés por las matemáticas obtuvieron las puntuaciones más bajas en el examen.

Con respecto al modelo aplicado a los datos en paneles es importante resaltar que ninguna de las variables que definían el factor “*actitud hacia las matemáticas*”, el cual estaba muy

correlacionado con la nota del primer examen, ha quedado en el modelo. Esto posiblemente se debe a que en la medida en que los estudiantes avancen en el curso las otras variables comienzan a tener más importancia en relación con el desempeño académico. Sin embargo, en el modelo II las ocho variables relevantes corresponden a los factores “*Rezago Académico*”, y “*Características del Profesor y del Examen*”.

Las variables que resultaron significativas en alguno de los tres modelos son:

- *Nota del curso anterior.* Hay una alta probabilidad de que el rendimiento futuro sea como el rendimiento del curso previo.
- *Dificultad del examen.* Los estudiantes con puntuaciones bajas en los exámenes tienden a opinar que hubo poca claridad en las explicaciones que recibieron.
- *Dificultad del examen.* Los estudiantes con alta probabilidad de obtener puntuaciones bajas tienden a opinar que los exámenes fueron difíciles.
- *Rapidez con la que el profesor cubrió los temas.* Los estudiantes con alta probabilidad de obtener puntuaciones bajas tienden a opinar que el profesor cubrió muy rápidamente los temas.
- *Prioridad que da al curso.* Los estudiantes con alta probabilidad de obtener puntuaciones bajas en sus exámenes tienden a asignarle baja prioridad al curso.
- *Número de veces que ha tomado el curso.* Los estudiantes con alta probabilidad de obtener puntuaciones bajas tienden a tomar dos o más veces el curso.
- *Número de créditos matriculados.* Los estudiantes con alta probabilidad de obtener puntuaciones bajas tienden a tener una carga académica baja (doce o menos créditos matriculados)

- *Darse de baja.* Los estudiantes con alta probabilidad de obtener puntuaciones bajas tienden a pensar en darse de baja del curso.
- *Las otras materias absorbieron la mayor parte del tiempo.* Los estudiantes con alta probabilidad de obtener puntuaciones bajas tienden a dedicar poco tiempo al curso.
- *Ausencias.* Los estudiantes con alta probabilidad de obtener puntuaciones bajas tienden a faltar a clases.
- *Visitar la oficina del profesor.* Los estudiantes con alta probabilidad de obtener puntuaciones altas tienden a asistir a horas de consulta con su profesor.

7. LIMITACIONES

- Los cuestionarios utilizados no fueron validados, principalmente por limitaciones de tiempo. Llevar a cabo una validación de los instrumentos utilizados requeriría un semestre académico.
- La muestra no fue aleatoria, por lo tanto, las inferencias que se pudieran hacer sobre los resultados de este estudio serían válidas para la población de estudiantes que voluntariamente hubieran participado en un estudio similar.
- El proceso de identificación de los cuestionarios que fue necesario para poder utilizar el análisis de dato por paneles pudo afectar la tasa de respuesta en los tres paneles de tiempo.
- La población de estudiantes matriculados en el curso de Cálculo I es poco constante en la asistencia a clase. Esto afectó la composición y el tamaño de la muestra en cada periodo de tiempo en que se recopilaron los datos.

BIBLIOGRAFÍA

- [1] Acuña, E. (2007), *Análisis de Regresión*. Departamento de Matemáticas, Universidad de Puerto Rico-Recinto Universitario de Mayagüez.
- [2] Agresti, A. (2002), *Categorical Data Analysis*. 2da Ed. John Wiley & Sons.
- [3] Albarran I. y González P. (2001), Estimación mediante datos de panel de la influencia del poder adquisitivo sobre el mercado asegurador en las principales economías occidentales. *Anales del Instituto de Actuarios Españoles*. No 7, 169-190
- [4] Ander, E. (1978), *Introducción a las Técnicas de Investigación Social*. Humanitas, Buenos Aires.
- [5] Friedman, J. (1997), On Bias, Variance, 0/1 Loss, and the Curse of Dimensionality. *Data mining and Knowledge Discovery*. Vol 1, No 1, 55-57.
- [6] Arellano M. y Bover, O. (1990), *La Econometría de Datos de Panel*, Investigaciones Económicas (Segunda Época), Vol XIV, 1:3-45
- [7] Baltagi, B. (2002), *Econometric Analysis of Panel Data*. 2da Ed. John Wiley & Son, New York.
- [8] Batistakis, Y., Halkidi, M. and Vazirgiannis, M. (2001), On Clustering Validation Techniques. *Journal of Intelligent Information System*, 17:2/3, 107-145.
- [9] Batistakis, Y., Halkidi, M. and Vazirgiannis, M. (2002), Cluster Validity Methods: Part I. *SIGMOD Record* 31(2), 40-45.

- [10] Cameron A. (1998). *Regression analysis of count data*. Ed. Cambridge University Press, United Kingdom.
- [11] Cañadas, I., Sánchez, A. (1998), Categorías de Respuesta en Escalas Tipo Likert. *Psicothema*, Vol. 10, nº 3, pp. 623-631
- [12] Corral, Ni. (2002), *Metas Académicas y Rasgos Cognitivos- Motivacionales de Estudiantes Universitarios*. Instituto de Ciencias de la Educación – Facultad de Humanidades, Argentina. <http://www1.unne.edu.ar/cyt/2002/09-Educacion/D-004.pdf>
- [13] Correa, J. y González, N. (2002), *Gráficos Estadísticos con R*. Universidad Nacional- Sede Medellín
- [14] Cruz, M. (1995), *Clasificación de los Departamentos en el Perú por Análisis Factorial y de Acumulación*. *Revista Apuntes*, No. 37, II semestre.
- [15] Dallas, E. (1998), *Applied Multivariate Methods for Data Analysis*, Ed. Duxbury Press, California.
- [16] Dowdy, S., Wearden, S., Chilko., D. (2004), *Statistics for Research*, 3ra ed. Wiley, New Jersey.
- [17] Elejabarrieta, F., Iñiguez, L. (1984), *Construcción de Escalas de Actitud Tipo Thurst y Likert*.
- [18] Frees W. (2004), *Longitudinal and Panel Data: Analysis and applications in the social sciences*. Ed. Cambridge university Press, New York.
- [19] Frees W. y Millar T. (2002), Forecasting lottery sales using panel data. *Tulossa International Journal of Forescasting*. <http://research.bus.wisc.edu/jfrees/>. Acedido en mayo de 2006.

- [20] Gardner, H. (1995), *Inteligencias Múltiples (La Teoría en la Práctica)*. Ed. Paidós, Barcelona.
- [21] Garson, D. (1998), *Factor Analysis*. <http://www2.chass.ncsu.edu/garson/pa765/factor.htm>. Consultado en febrero de 2007.
- [22] González López, Ignacio (2003), Determinación de los Elementos que Condicionan la Calidad de la Universidad: Aplicación Práctica de un Análisis Factorial. *Revista Electrónica de Investigación y Evaluación Educativa*, v. 9, n. 1.
- [23] González, M. (2005), *A Comparison in Cluster Validation Techniques*. Tesis
- [24] Hair, J., Ronald, R., Black, W. (2006), *Multivariate Data Analysis. 5ta Ed.* Prentice Hall, New Jersey.
- [25] Hsiao C. (2003), *Analysis of panel data*. 2da Ed. Cambridge University Press, United Kingdom.
- [26] Instituto Nacional de Estadística e Informática, Perú. (1995) *¿Cómo se Calcula el Índice de Desarrollo Educativo de la Niñez y la Adolescencia?* Boletín Informativo.
- [27] Kaufman, L., Rousseeuw P. (1990), *Finding groups in data: An Introduction to Cluster Analysis*. John Wiley & Sons, New York.
- [28] Lattin, J., Carroll, J., Green, P. (2003), *Analyzing Multivariate Data*. Thomson Learning, Canada.
- [29] Mayorga, M. (2000), *La técnica de Datos de Panel: Una Guía para su Uso e Interpretación*. [en línea] Banco Central de Costa Rica-División Económica-Departamento de Investigaciones Económicas. <http://www.bccr.fi.cr/ndie/Documentos/NT-05-2000.PDF>, Consultado en octubre de 2006.

- [30] Moral, J. (2006), *Predicción del Rendimiento Académico Universitario*. Perfiles Educativos V.28 No. 113. Méjico.
- [31] Mueller, C.W. (1978), *Factor Analysis; statistical methods and practical issues*. Beverly Hills, California: Sage.
- [32] Olivera M. (2004), *Efectos del Desayuno Escolar en la Tasa de Retención del Sistema Escolar Público*. Revista Latinoamericana de Desarrollo Económico, No. 3, P. 77-116.
- [33] Rockoff J. E. (2003), The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, 94(2), 247-252.
- [34] Rúa, A. (2001), *Búsqueda de Patrones de Rendimiento Académico Mediante Técnicas de Análisis Multivariante*. Aplicación a 1° E4. <http://150.214.55.100/asepuma/laspalmas2001/laspalmas/Doco23.PDF>. Consultado en noviembre de 2006.
- [35] SAS Inc. *SAS online documentation: The GLIMMIX PROCEDURE*, 2006
- [36] Salvador, M. , Gargallo, P. (2006), Análisis Factorial. [en línea] [5campus.com](http://www.5campus.com), **Estadística** <http://www.5campus.com/leccion/factorial>, Consultado en enero de 2007.
- [37] Segura, J., Solís, J., Segura, V. (2006), *Análisis de Regresión Logística para Datos Correlacionados Utilizando Tres Procedimientos del Sistema Estadístico SAS*. Revista Científica, FCV- LUZ, Vol. XVI, No 3, 282-287.
- [38] UPRM (2005), *Estudio de Calificaciones y Fracazos*. Oficina de Investigación Institucional y Planificación del Recinto Universitario de Mayagüez.

ANEJO 1: Cuestionario I



UNIVERSIDAD DE PUERTO RICO
RECINTO UNIVERSITARIO DE MAYAGÜEZ
DEPARTAMENTO DE MATEMÁTICAS



ESTUDIO DE ALGUNOS FACTORES QUE AFECTAN EL RENDIMIENTO
ACADÉMICO DE LOS ESTUDIANTES DE CÁLCULO I.

A continuación encontrará una serie de preguntas en la que se pretende medir algunas características que pueden influir en su rendimiento académico. **Responda las preguntas con mucha confianza y honestidad. Tenga en cuenta que no representa una nota para su curso de Cálculo I.**

1. ¿Recibe Beca o algún tipo de ayuda económica? ___ Sí ___ No. Si su respuesta es sí; indique el tipo de beca y número de horas a la semana que dedica a la actividad.

<i>Tipo de Beca o Ayuda</i>	<i>No. De horas</i>	<i>Tipo de beca o Ayuda</i>
<input type="radio"/> <i>Actividad Atlética</i>	_____	<input type="radio"/> <i>Ayuda suplementaria</i>
<input type="radio"/> <i>Banda y Orquesta</i>	_____	<input type="radio"/> <i>Pell</i>
<input type="radio"/> <i>Estudio Trabajo</i>	_____	<input type="radio"/> <i>AMS</i>

2. Indique la nota final obtenida en su curso anterior de matemáticas (si se dio de baja, favor de indicarlo).

<input type="checkbox"/> <i>A</i>	<input type="checkbox"/> <i>B</i>	<input type="checkbox"/> <i>C</i>	<input type="checkbox"/> <i>D</i>	<input type="checkbox"/> <i>F</i>	<input type="checkbox"/> <i>W</i>
-----------------------------------	-----------------------------------	-----------------------------------	-----------------------------------	-----------------------------------	-----------------------------------

3. Estudió en escuela: ___ *Pública* ___ *Privada*.

4. Considera que su desempeño en matemáticas en la Escuela Superior fue:

<input type="checkbox"/> <i>Excelente</i>	<input type="checkbox"/> <i>Bueno</i>	<input type="checkbox"/> <i>Aceptable</i>	<input type="checkbox"/> <i>Regular</i>	<input type="checkbox"/> <i>Deficiente</i>
---	---------------------------------------	---	---	--

5. Indique los cursos que ha tomado.

<input type="checkbox"/> <i>Prebásica</i>	<input type="checkbox"/> <i>Precálculo I</i>	<input type="checkbox"/> <i>Precálculo (5 créditos)</i>
<input type="checkbox"/> <i>Precálculo</i>	<input type="checkbox"/> <i>Precálculo II</i>	

6. Número de veces que ha tomado este curso (incluyendo ésta) _____

7. Califique de 1 a 5 la importancia de este curso para su formación profesional. _____

8. Qué prioridad da a la materia de Cálculo I.

<input type="checkbox"/> <i>Alta</i>	<input type="checkbox"/> <i>Regular</i>	<input type="checkbox"/> <i>Baja</i>
--------------------------------------	---	--------------------------------------

9. ¿Trabaja actualmente? (Diferente a estudio y trabajo)

- Sí, ¿Cuántas horas por semana? _____*
- No.*

	<i>Agosto 9-15</i>	<i>Agosto 16-22</i>	<i>Agosto 23-30</i>
10. Promedio de horas por semana que estudió para este examen en el <i>centro de apoyo</i>	_____	_____	_____
11. Promedio de horas por semana que estudió para este examen <i>fuera del centro de apoyo</i>	_____	_____	_____

12. ¿En el periodo previo a este examen visitó la oficina del profesor?

- Sí, ¿Cuántas veces? _____*
- No*

13. Indique el número de horas que ***no asistió*** a clase (del 8 al 30 de agosto). _____

14. Número de miembros de la familia inmediata que hayan obtenido por lo menos un grado de bachillerato. _____

15. ¿Cuántos ejercicios desarrolló del prontuario? *Nota: Había cerca de 300 ejercicios propuestos.*

<input type="checkbox"/> Ninguno	<input type="checkbox"/> De 61 a 120	<input type="checkbox"/> 181 a 240
<input type="checkbox"/> De 1 a 60	<input type="checkbox"/> De 121 a 180	<input type="checkbox"/> 241 a 300

Para este examen:	Sí	No
16. Practiqué usando exámenes de semestres anteriores.	<input type="checkbox"/>	<input type="checkbox"/>
17. Practiqué los ejercicios de la libreta.	<input type="checkbox"/>	<input type="checkbox"/>
18. Comencé a desarrollar los ejercicios sin leer la teoría.	<input type="checkbox"/>	<input type="checkbox"/>
19. Tengo el libro de texto	<input type="checkbox"/>	<input type="checkbox"/>
20. Le gustaría que este curso estuviera fuera de su currículo de estudios.	<input type="checkbox"/>	<input type="checkbox"/>

21. Hice preguntas durante la clase.	<input type="checkbox"/>	<input type="checkbox"/>
22. El profesor recordó su horario de oficina durante la enseñanza de los temas para este examen	<input type="checkbox"/>	<input type="checkbox"/>
23. He pensado en darme de baja de este curso	<input type="checkbox"/>	<input type="checkbox"/>

	<i>Poco(a)</i>	<i>Suficiente</i>	<i>Demasiado(a)</i>
24. Considera que la cantidad de tema para este examen fue:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25. Considera que el tiempo para resolver este examen fue:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
26. Considera que el tiempo de la clase para cubrir los temas que se evaluaron en este examen fue:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27. Considera que la cantidad de ejercicios que se resolvió en clase fue:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Para los siguientes enunciados marque con una X en el encasillado que mejor represente su situación.

Siempre	Casi Siempre	Algunas veces	Casi nunca	Nunca
S	CS	AV	CN	N

	S	CS	AV	CN	N
28. Desarrollé ejercicios del prontuario cuando se terminaba cada sección.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
29. Desarrollé ejercicios del prontuario faltando poco tiempo para el examen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
30. Repasé las notas de clase cuando se terminaba cada sección.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
31. Repasé las notas de clase cuando faltaba poco tiempo para el examen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
32. Los temas del examen los estudié en grupo.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
33. Memorice la mayoría de ejercicios	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
34. Considero que las otras materias absorbieron la mayor parte de mi tiempo.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
35. Los temas del examen los estudié solo.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

36. Número de créditos en los que está matriculado actualmente: _____
37. Cuando estudia en grupo lo hace con _____ (número de personas sin incluirse)
38. Indique el grado de gusto por las matemáticas.
NINGÚN GUSTO 1 2 3 4 5 6 7 8 9 *MUCHO GUSTO*
39. Evalúe la dificultad de este examen.
MUY DIFÍCIL 9 8 7 6 5 4 3 2 1 *NINGUNA DIFICULTAD*
40. Valora el esfuerzo que tú haces para sacar buena nota en el curso de Cálculo I.
NINGÚN ESFUERZO 1 2 3 4 5 6 7 8 9 *MUCHO ESFUERZO*
41. Valora tu propia capacidad para estudiar matemáticas.
MUY MALA 1 2 3 4 5 6 7 8 9 *MUY BUENA*
42. Valora la rapidez con la que el profesor cubrió los temas de este examen.
MUY RÁPIDO 9 8 7 6 5 4 3 2 1 *MUY DESPACIO*
43. Valora la claridad con la que el profesor explicó los temas de este examen.
POCO CLARO 1 2 3 4 5 6 7 8 9 *MUY CLARO*
44. Valora el interés que te tomas por estudiar este curso
NINGÚN INTERÉS 1 2 3 4 5 6 7 8 9 *MUCHO INTERÉS*
45. Valora la importancia de obtener una buena nota en este curso.
MUY IMPORTANTE PARA MÍ NADA 9 8 7 6 5 4 3 2 1 *IMPORTANTE PARA MÍ*
46. Valora la exigencias que te pones a ti mismo respecto al estudio de matemáticas.
EXIGENCIAS MUY BAJAS 1 2 3 4 5 6 7 8 9 *EXIGENCIAS MUY ALTAS*
47. Valora tu interés de aprender matemáticas.
NINGÚN INTERÉS 1 2 3 4 5 6 7 8 9 *MUCHO INTERÉS*

Número de estudiante: _____ Sección: _____ Edad: _____ Género: _____

Concentración: _____

Departamento _____ Facultad: _____

Estado civil: _____ Número de hijos: _____

ANEJO 2: Correlaciones de Spearman.

(*) Correlación significativa al 10%

(**) Correlación significativa al 5%

Varables	Exa_1	Sexo	Facultad	P9	P60	P16	P19	P47	P25
Exa_1	1.000	-0.023	.184(**)	0.100	-.255(**)	0.050	0.081	.171(**)	0.009
Sexo	-0.023	1.000	.128(*)	0.052	0.004	-0.007	-0.072	-0.069	0.078
Facultad	.184(**)	.128(*)	1.000	0.079	-.214(**)	0.035	0.014	-0.024	0.094
P9	0.100	0.052	0.079	1.000	-.154(*)	-.213(**)	-0.007	.195(**)	.322(**)
P60	-.255(**)	0.004	-.214(**)	-.154(*)	1.000	-0.021	0.049	-.261(**)	-0.114
P16	0.050	-0.007	0.035	-.213(**)	-0.021	1.000	.169(**)	0.046	-0.030
P19	0.081	-0.072	0.014	-0.007	0.049	.169(**)	1.000	0.017	-0.001
P47	.171(**)	-0.069	-0.024	.195(**)	-.261(**)	0.046	0.017	1.000	.126(*)
P25	0.009	0.078	0.094	.322(**)	-0.114	-0.030	-0.001	.126(*)	1.000
P8	.317(**)	-0.011	.150(*)	0.027	-.452(**)	0.099	0.026	.323(**)	0.075
P10	0.091	-.170(**)	-0.086	-0.005	-0.070	-.162(**)	0.110	0.051	0.046
P11	-0.044	0.016	-0.107	-0.119	.481(**)	-0.035	-0.016	-.240(**)	-0.066
P12	0.016	0.066	.136(*)	0.016	-.207(**)	-0.107	0.078	0.027	-0.065
P13	.221(**)	0.001	.270(**)	0.080	-.141(*)	0.007	.207(**)	0.071	0.109
P14	-0.110	0.066	-0.018	-0.084	.233(**)	0.027	0.035	-0.084	0.003
P22	.182(**)	.139(*)	0.110	0.039	-.153(*)	.289(**)	0.108	.152(*)	0.104
P24	-.169(**)	-0.094	-0.105	0.044	.224(**)	-0.056	0.006	-0.013	-0.019
P26	0.047	-.131(*)	-0.045	-0.068	.146(*)	.178(**)	.419(**)	0.005	-0.045
P1	0.044	-0.103	-0.046	-.339(**)	.165(**)	0.081	-0.010	0.043	-.245(**)
P27	.157(*)	-.140(*)	0.032	0.019	-0.101	.170(**)	.174(**)	0.037	0.082
P28	-0.026	0.042	-0.087	-0.034	0.064	-0.059	0.011	0.029	-0.044
P29	0.015	-0.023	0.036	-0.049	.131(*)	-0.065	-0.036	-0.026	-0.037
P30	-0.050	-0.079	-0.024	0.098	-.146(*)	0.070	0.011	-0.058	0.079
P31	-.258(**)	-0.117	-0.008	-0.051	.256(**)	-0.021	-0.045	-.222(**)	-0.090
P32	0.082	0.087	0.068	0.012	0.023	.165(**)	.138(*)	0.031	-0.027
P33	.132(*)	0.036	.131(*)	0.042	-0.106	0.026	-0.046	0.080	0.019
P34	-.282(**)	-0.114	-.132(*)	-0.090	-0.033	0.079	-0.019	0.031	-0.105
P35	-.140(*)	0.008	-0.019	-0.089	-0.039	0.017	0.066	-0.054	-0.037
P36	.171(**)	0.050	0.018	0.082	-.121(*)	-0.096	-0.053	-0.008	0.019
P37	.207(**)	0.009	0.093	0.010	-0.041	-0.044	-0.055	0.004	0.024
P38	.337(**)	0.034	0.061	-0.030	-0.099	-0.016	-0.025	0.054	-0.056
P39	-0.018	0.038	-0.041	-0.104	0.076	.190(**)	.352(**)	0.005	-0.016
P40	-.212(**)	-.134(*)	-0.071	0.120	.139(*)	-.136(*)	0.047	0.018	-0.042
P41	-0.071	-0.020	0.003	-0.074	0.096	0.040	.185(**)	0.048	-0.084
P42	-0.072	-0.049	-0.021	0.080	0.012	-0.120	-0.063	-0.040	0.057
P43	0.069	-0.071	.163(**)	0.072	-.122(*)	0.116	0.058	0.113	0.059
P44	-0.106	0.042	0.082	0.014	0.074	0.108	0.046	-0.087	-0.002
P45	-.291(**)	-0.016	-.171(**)	-0.121	0.097	0.068	0.072	0.074	-0.076
P48	-0.115	-0.065	-0.022	-0.064	0.072	.191(**)	-0.029	-0.047	0.053
P49	.267(**)	0.060	.128(*)	0.055	-.254(**)	-0.025	.172(**)	0.117	0.073
P50	-.384(**)	-0.056	0.017	-0.054	0.067	0.108	.125(*)	0.029	0.029
P51	0.107	-.121(*)	0.096	-0.019	-0.011	.282(**)	.310(**)	0.065	0.012
P52	.171(**)	0.071	0.040	0.043	-0.086	0.073	.179(**)	0.105	0.108
P53	-.193(**)	-0.007	-0.074	0.056	0.004	-0.012	0.037	0.030	0.009
P54	.286(**)	0.101	0.036	0.051	-0.062	-.126(*)	-0.122	0.047	-0.074
P55	.240(**)	0.061	0.121	0.064	-.175(**)	.147(*)	.242(**)	0.120	0.111
P57	.151(*)	-0.079	.199(**)	0.094	-.137(*)	.186(**)	.276(**)	0.109	0.113
P58	.189(**)	-0.048	0.078	-0.020	-.244(**)	0.076	.167(**)	.152(*)	0.066

Varables	P8	P10	P11	P12	P13	P14	P22	P24	P26
Exa_1	.317(**)	0.091	-0.044	0.016	.221(**)	-0.110	.182(**)	-.169(**)	0.047
Sexo	-0.011	-.170(**)	0.016	0.066	0.001	0.066	.139(*)	-0.094	-.131(*)
Facultad	.150(*)	-0.086	-0.107	.136(*)	.270(**)	-0.018	0.110	-0.105	-0.045
P9	0.027	-0.005	-0.119	0.016	0.080	-0.084	0.039	0.044	-0.068
P60	-.452(**)	-0.070	.481(**)	-.207(**)	-.141(*)	.233(**)	-.153(*)	.224(**)	.146(*)
P16	0.099	-.162(**)	-0.035	-0.107	0.007	0.027	.289(**)	-0.056	.178(**)
P19	0.026	0.110	-0.016	0.078	.207(**)	0.035	0.108	0.006	.419(**)
P47	.323(**)	0.051	-.240(**)	0.027	0.071	-0.084	.152(*)	-0.013	0.005
P25	0.075	0.046	-0.066	-0.065	0.109	0.003	0.104	-0.019	-0.045
P8	1.000	0.104	-.580(**)	0.055	0.037	-.269(**)	0.114	-0.120	-0.053
P10	0.104	1.000	0.003	-0.028	-0.010	0.002	-0.089	0.044	0.100
P11	-.580(**)	0.003	1.000	-.171(**)	-0.053	.183(**)	-0.037	0.051	0.048
P12	0.055	-0.028	-.171(**)	1.000	.320(**)	-0.059	-0.008	-0.091	-0.089
P13	0.037	-0.010	-0.053	.320(**)	1.000	-0.047	0.106	-0.100	.154(*)
P14	-.269(**)	0.002	.183(**)	-0.059	-0.047	1.000	0.029	0.120	0.020
P22	0.114	-0.089	-0.037	-0.008	0.106	0.029	1.000	-.130(*)	0.082
P24	-0.120	0.044	0.051	-0.091	-0.100	0.120	-.130(*)	1.000	-0.052
P26	-0.053	0.100	0.048	-0.089	.154(*)	0.020	0.082	-0.052	1.000
P1	0.017	-0.027	-0.031	-0.021	0.048	-0.056	-0.013	0.037	0.030
P27	0.041	-0.006	-0.002	0.073	.166(**)	0.066	0.031	0.007	.158(*)
P28	-0.037	0.045	0.052	-0.015	-0.081	0.022	0.013	-0.103	-.174(**)
P29	-0.122	-.129(*)	0.114	0.012	-.127(*)	0.088	-0.011	0.040	-0.069
P30	-0.040	0.005	-0.064	0.048	0.045	-0.073	0.087	-0.041	0.117
P31	-.221(**)	-0.016	.162(*)	-.269(**)	-.248(**)	0.052	-0.086	.188(**)	0.090
P32	-0.016	-0.066	0.036	0.039	0.097	.142(*)	0.114	-0.091	0.041
P33	.180(**)	0.026	-.131(*)	-0.026	-0.019	-0.063	0.092	-.160(*)	-0.004
P34	-0.039	-0.031	-0.103	-0.022	-.135(*)	0.072	-0.030	0.040	0.100
P35	0.016	-0.003	-0.023	0.063	-0.088	-0.011	0.014	0.093	0.044
P36	0.018	.200(**)	-0.044	-0.004	-0.013	-0.027	-0.033	-0.024	0.031
P37	0.037	0.023	-0.046	0.101	0.035	-0.089	-0.030	-0.107	0.034
P38	0.113	-0.005	-.131(*)	.188(**)	0.003	-0.049	-0.042	-.163(**)	0.067
P39	-0.106	0.087	.139(*)	0.100	.216(**)	0.074	.203(**)	-.135(*)	.455(**)
P40	-0.080	-0.058	0.019	0.019	-0.052	.175(**)	-0.092	0.047	.135(*)
P41	-0.047	.144(*)	.136(*)	0.054	0.082	0.019	.133(*)	0.034	0.068
P42	0.054	0.018	-0.009	-.140(*)	-0.049	0.024	-0.063	0.041	-.209(**)
P43	0.108	0.004	-.186(**)	.198(**)	.170(**)	0.060	0.006	-0.012	0.046
P44	-0.081	0.020	0.079	-0.050	-0.042	0.117	0.076	0.080	-0.113
P45	-0.050	-0.067	0.049	-0.103	-.177(**)	.121(*)	-0.097	0.009	-0.015
P48	-0.067	0.043	-0.007	0.041	0.027	-0.013	0.066	0.088	-0.035
P49	.199(**)	0.110	-.152(*)	.218(**)	.244(**)	-.145(*)	0.044	-.142(*)	-0.106
P50	0.010	-.142(*)	-0.122	-0.036	-0.006	0.075	0.084	0.105	0.090
P51	0.032	0.092	0.018	0.019	.263(**)	0.014	0.097	-0.017	.300(**)
P52	.212(**)	.151(*)	-.150(*)	0.068	.211(**)	-.131(*)	-0.003	-0.008	0.064
P53	-0.019	-0.112	-0.001	-0.100	-0.035	-0.016	-0.020	0.007	-0.007
P54	0.013	-0.025	-0.017	.137(*)	0.000	-0.117	-0.082	-.146(*)	-.137(*)
P55	.169(**)	0.062	-.144(*)	.165(**)	.381(**)	0.016	.140(*)	-.162(**)	0.119
P57	.165(**)	0.109	-.132(*)	.124(*)	.330(**)	0.030	.176(**)	-0.009	.220(**)
P58	.193(**)	0.120	-0.110	.210(**)	.395(**)	-0.059	0.093	-0.099	0.027

Varables	P1	P27	P28	P29	P30	P31	P32	P33	P34
Exa_1	0.044	.157(*)	-0.026	0.015	-0.050	-.258(**)	0.082	.132(*)	-.282(**)
Sexo	-0.103	-.140(*)	0.042	-0.023	-0.079	-0.117	0.087	0.036	-0.114
Facultad	-0.046	0.032	-0.087	0.036	-0.024	-0.008	0.068	.131(*)	-.132(*)
P9	-.339(**)	0.019	-0.034	-0.049	0.098	-0.051	0.012	0.042	-0.090
P60	.165(**)	-0.101	0.064	.131(*)	-.146(*)	.256(**)	0.023	-0.106	-0.033
P16	0.081	.170(**)	-0.059	-0.065	0.070	-0.021	.165(**)	0.026	0.079
P19	-0.010	.174(**)	0.011	-0.036	0.011	-0.045	.138(*)	-0.046	-0.019
P47	0.043	0.037	0.029	-0.026	-0.058	-.222(**)	0.031	0.080	0.031
P25	-.245(**)	0.082	-0.044	-0.037	0.079	-0.090	-0.027	0.019	-0.105
P8	0.017	0.041	-0.037	-0.122	-0.040	-.221(**)	-0.016	.180(**)	-0.039
P10	-0.027	-0.006	0.045	-.129(*)	0.005	-0.016	-0.066	0.026	-0.031
P11	-0.031	-0.002	0.052	0.114	-0.064	.162(*)	0.036	-.131(*)	-0.103
P12	-0.021	0.073	-0.015	0.012	0.048	-.269(**)	0.039	-0.026	-0.022
P13	0.048	.166(**)	-0.081	-.127(*)	0.045	-.248(**)	0.097	-0.019	-.135(*)
P14	-0.056	0.066	0.022	0.088	-0.073	0.052	.142(*)	-0.063	0.072
P22	-0.013	0.031	0.013	-0.011	0.087	-0.086	0.114	0.092	-0.030
P24	0.037	0.007	-0.103	0.040	-0.041	.188(**)	-0.091	-.160(*)	0.040
P26	0.030	.158(*)	-.174(**)	-0.069	0.117	0.090	0.041	-0.004	0.100
P1	1.000	-0.044	0.053	.130(*)	-.161(**)	-0.031	0.109	0.002	0.013
P27	-0.044	1.000	-.123(*)	0.102	-0.013	-0.065	0.120	0.039	0.023
P28	0.053	-.123(*)	1.000	0.033	-0.004	-0.045	0.076	0.113	-0.017
P29	.130(*)	0.102	0.033	1.000	-.156(*)	0.029	0.029	-0.027	0.002
P30	-.161(**)	-0.013	-0.004	-.156(*)	1.000	-0.096	-.148(*)	0.010	-0.049
P31	-0.031	-0.065	-0.045	0.029	-0.096	1.000	-0.031	0.021	.266(**)
P32	0.109	0.120	0.076	0.029	-.148(*)	-0.031	1.000	0.056	0.003
P33	0.002	0.039	0.113	-0.027	0.010	0.021	0.056	1.000	-.128(*)
P34	0.013	0.023	-0.017	0.002	-0.049	.266(**)	0.003	-.128(*)	1.000
P35	0.006	0.069	-0.001	0.003	-0.036	.202(**)	-0.007	-0.027	.215(**)
P36	-0.074	-0.095	-0.062	0.011	0.011	-0.025	0.005	0.010	-0.095
P37	0.032	0.057	-0.098	0.083	-0.087	-0.105	0.004	0.115	-.188(**)
P38	0.067	0.051	0.023	0.009	-0.013	-.159(*)	0.007	.247(**)	-.142(*)
P39	-0.008	0.107	0.030	-.147(*)	.164(**)	-0.052	.219(**)	0.024	0.015
P40	-0.015	0.002	-0.029	0.083	0.051	0.104	-0.072	-0.018	.164(**)
P41	0.071	0.090	.308(**)	-0.101	-0.055	-0.029	.201(**)	.133(*)	-0.031
P42	0.030	0.044	.122(*)	0.047	-0.005	0.094	-0.015	0.015	0.080
P43	-0.002	.278(**)	-0.033	0.068	0.026	-0.067	.159(**)	0.100	-0.009
P44	-0.042	0.004	0.096	0.018	-0.032	0.113	0.036	0.055	0.109
P45	0.083	-0.045	.164(**)	0.026	-0.037	0.077	0.044	-0.025	.221(**)
P48	0.047	.211(**)	0.016	-0.005	-0.050	0.026	.138(*)	0.007	-0.057
P49	0.012	0.095	-0.040	-0.050	-0.048	-.377(**)	0.067	0.019	-.155(*)
P50	-0.031	0.060	-0.039	-0.044	0.044	.159(*)	-0.001	-0.112	.238(**)
P51	0.050	.203(**)	-.125(*)	-.232(**)	-0.006	-0.065	.160(**)	-0.090	-0.016
P52	0.033	.121(*)	-0.006	-.145(*)	0.068	-.287(**)	0.031	0.021	-.186(**)
P53	0.015	-0.060	0.031	-0.054	-0.070	0.078	0.026	-0.102	0.026
P54	0.004	0.028	.149(*)	0.083	-0.078	-0.113	0.005	.362(**)	-.140(*)
P55	0.034	.140(*)	-0.042	-.162(**)	0.018	-.304(**)	.241(**)	-0.020	-.133(*)
P57	0.017	.171(**)	-0.031	-.193(**)	-0.027	-.145(*)	0.109	0.032	-0.100
P58	0.059	.198(**)	-0.046	-0.091	-0.049	-.316(**)	0.098	-0.009	-.158(*)

Varables	P35	P36	P37	P38	P39	P40	P41	P42	P43
Exa_1	-.140(*)	.171(**)	.207(**)	.337(**)	-0.018	-.212(**)	-0.071	-0.072	0.069
Sexo	0.008	0.050	0.009	0.034	0.038	-.134(*)	-0.020	-0.049	-0.071
Facultad	-0.019	0.018	0.093	0.061	-0.041	-0.071	0.003	-0.021	.163(**)
P9	-0.089	0.082	0.010	-0.030	-0.104	0.120	-0.074	0.080	0.072
P60	-0.039	-.121(*)	-0.041	-0.099	0.076	.139(*)	0.096	0.012	-.122(*)
P16	0.017	-0.096	-0.044	-0.016	.190(**)	-.136(*)	0.040	-0.120	0.116
P19	0.066	-0.053	-0.055	-0.025	.352(**)	0.047	.185(**)	-0.063	0.058
P47	-0.054	-0.008	0.004	0.054	0.005	0.018	0.048	-0.040	0.113
P25	-0.037	0.019	0.024	-0.056	-0.016	-0.042	-0.084	0.057	0.059
P8	0.016	0.018	0.037	0.113	-0.106	-0.080	-0.047	0.054	0.108
P10	-0.003	.200(**)	0.023	-0.005	0.087	-0.058	.144(*)	0.018	0.004
P11	-0.023	-0.044	-0.046	-.131(*)	.139(*)	0.019	.136(*)	-0.009	-.186(**)
P12	0.063	-0.004	0.101	.188(**)	0.100	0.019	0.054	-.140(*)	.198(**)
P13	-0.088	-0.013	0.035	0.003	.216(**)	-0.052	0.082	-0.049	.170(**)
P14	-0.011	-0.027	-0.089	-0.049	0.074	.175(**)	0.019	0.024	0.060
P22	0.014	-0.033	-0.030	-0.042	.203(**)	-0.092	.133(*)	-0.063	0.006
P24	0.093	-0.024	-0.107	-.163(**)	-.135(*)	0.047	0.034	0.041	-0.012
P26	0.044	0.031	0.034	0.067	.455(**)	.135(*)	0.068	-.209(**)	0.046
P1	0.006	-0.074	0.032	0.067	-0.008	-0.015	0.071	0.030	-0.002
P27	0.069	-0.095	0.057	0.051	0.107	0.002	0.090	0.044	.278(**)
P28	-0.001	-0.062	-0.098	0.023	0.030	-0.029	.308(**)	.122(*)	-0.033
P29	0.003	0.011	0.083	0.009	-.147(*)	0.083	-0.101	0.047	0.068
P30	-0.036	0.011	-0.087	-0.013	.164(**)	0.051	-0.055	-0.005	0.026
P31	.202(**)	-0.025	-0.105	-.159(*)	-0.052	0.104	-0.029	0.094	-0.067
P32	-0.007	0.005	0.004	0.007	.219(**)	-0.072	.201(**)	-0.015	.159(**)
P33	-0.027	0.010	0.115	.247(**)	0.024	-0.018	.133(*)	0.015	0.100
P34	.215(**)	-0.095	-.188(**)	-.142(*)	0.015	.164(**)	-0.031	0.080	-0.009
P35	1.000	-.152(*)	-.172(**)	-0.094	0.075	0.022	0.063	-0.059	0.025
P36	-.152(*)	1.000	0.101	0.104	0.008	-0.110	-0.087	-0.036	-.153(*)
P37	-.172(**)	0.101	1.000	.330(**)	-0.008	-0.002	-0.046	-0.084	0.112
P38	-0.094	0.104	.330(**)	1.000	0.006	-0.076	-0.008	-0.092	0.084
P39	0.075	0.008	-0.008	0.006	1.000	0.004	.373(**)	-.216(**)	-0.010
P40	0.022	-0.110	-0.002	-0.076	0.004	1.000	-0.063	.270(**)	-0.051
P41	0.063	-0.087	-0.046	-0.008	.373(**)	-0.063	1.000	-0.008	0.004
P42	-0.059	-0.036	-0.084	-0.092	-.216(**)	.270(**)	-0.008	1.000	-0.064
P43	0.025	-.153(*)	0.112	0.084	-0.010	-0.051	0.004	-0.064	1.000
P44	-0.047	-0.096	-0.075	-0.112	0.094	-0.058	.228(**)	0.044	.243(**)
P45	.160(**)	-.122(*)	-.128(*)	-0.108	-0.008	.176(**)	0.071	0.064	-0.091
P48	0.035	-.123(*)	0.082	-0.043	0.060	-0.008	0.075	0.044	.384(**)
P49	0.002	0.118	.140(*)	.123(*)	0.061	-0.101	0.001	0.011	0.114
P50	.317(**)	-.315(**)	-.239(**)	-.278(**)	0.097	.171(**)	0.019	-0.087	0.077
P51	0.048	-0.065	-0.047	-0.042	.379(**)	-0.069	.207(**)	-0.096	0.091
P52	-0.044	0.109	0.056	0.092	.139(*)	-0.073	0.096	-0.027	0.056
P53	.183(**)	-.137(*)	-.211(**)	-.166(**)	0.004	0.118	0.112	0.053	-0.012
P54	-0.106	.137(*)	.217(**)	.369(**)	-0.097	-.128(*)	0.107	-0.013	-0.047
P55	0.006	-0.001	0.049	0.068	.318(**)	-.126(*)	.174(**)	-0.044	.126(*)
P57	.148(*)	-0.007	-0.039	-0.020	.313(**)	0.006	.197(**)	-0.082	.191(**)
P58	0.046	0.043	0.099	0.090	.209(**)	-.126(*)	.217(**)	-.164(**)	0.113

Varables	P44	P45	P48	P49	P50	P51	P52	P53
Exa_1	-0.106	-.291(**)	-0.115	.267(**)	-.384(**)	0.107	.171(**)	-.193(**)
Sexo	0.042	-0.016	-0.065	0.060	-0.056	-.121(*)	0.071	-0.007
Facultad	0.082	-.171(**)	-0.022	.128(*)	0.017	0.096	0.040	-0.074
P9	0.014	-0.121	-0.064	0.055	-0.054	-0.019	0.043	0.056
P60	0.074	0.097	0.072	-.254(**)	0.067	-0.011	-0.086	0.004
P16	0.108	0.068	.191(**)	-0.025	0.108	.282(**)	0.073	-0.012
P19	0.046	0.072	-0.029	.172(**)	.125(*)	.310(**)	.179(**)	0.037
P47	-0.087	0.074	-0.047	0.117	0.029	0.065	0.105	0.030
P25	-0.002	-0.076	0.053	0.073	0.029	0.012	0.108	0.009
P8	-0.081	-0.050	-0.067	.199(**)	0.010	0.032	.212(**)	-0.019
P10	0.020	-0.067	0.043	0.110	-.142(*)	0.092	.151(*)	-0.112
P11	0.079	0.049	-0.007	-.152(*)	-0.122	0.018	-.150(*)	-0.001
P12	-0.050	-0.103	0.041	.218(**)	-0.036	0.019	0.068	-0.100
P13	-0.042	-.177(**)	0.027	.244(**)	-0.006	.263(**)	.211(**)	-0.035
P14	0.117	.121(*)	-0.013	-.145(*)	0.075	0.014	-.131(*)	-0.016
P22	0.076	-0.097	0.066	0.044	0.084	0.097	-0.003	-0.020
P24	0.080	0.009	0.088	-.142(*)	0.105	-0.017	-0.008	0.007
P26	-0.113	-0.015	-0.035	-0.106	0.090	.300(**)	0.064	-0.007
P1	-0.042	0.083	0.047	0.012	-0.031	0.050	0.033	0.015
P27	0.004	-0.045	.211(**)	0.095	0.060	.203(**)	.121(*)	-0.060
P28	0.096	.164(**)	0.016	-0.040	-0.039	-.125(*)	-0.006	0.031
P29	0.018	0.026	-0.005	-0.050	-0.044	-.232(**)	-.145(*)	-0.054
P30	-0.032	-0.037	-0.050	-0.048	0.044	-0.006	0.068	-0.070
P31	0.113	0.077	0.026	-.377(**)	.159(*)	-0.065	-.287(**)	0.078
P32	0.036	0.044	.138(*)	0.067	-0.001	.160(**)	0.031	0.026
P33	0.055	-0.025	0.007	0.019	-0.112	-0.090	0.021	-0.102
P34	0.109	.221(**)	-0.057	-.155(*)	.238(**)	-0.016	-.186(**)	0.026
P35	-0.047	.160(**)	0.035	0.002	.317(**)	0.048	-0.044	.183(**)
P36	-0.096	-.122(*)	-.123(*)	0.118	-.315(**)	-0.065	0.109	-.137(*)
P37	-0.075	-.128(*)	0.082	.140(*)	-.239(**)	-0.047	0.056	-.211(**)
P38	-0.112	-0.108	-0.043	.123(*)	-.278(**)	-0.042	0.092	-.166(**)
P39	0.094	-0.008	0.060	0.061	0.097	.379(**)	.139(*)	0.004
P40	-0.058	.176(**)	-0.008	-0.101	.171(**)	-0.069	-0.073	0.118
P41	.228(**)	0.071	0.075	0.001	0.019	.207(**)	0.096	0.112
P42	0.044	0.064	0.044	0.011	-0.087	-0.096	-0.027	0.053
P43	.243(**)	-0.091	.384(**)	0.114	0.077	0.091	0.056	-0.012
P44	1.000	0.062	.180(**)	-0.033	0.078	0.104	0.012	0.047
P45	0.062	1.000	-0.083	-0.113	.211(**)	-0.048	-.142(*)	0.097
P48	.180(**)	-0.083	1.000	-0.009	0.085	.155(*)	0.095	-0.080
P49	-0.033	-0.113	-0.009	1.000	-0.097	0.097	.457(**)	-0.012
P50	0.078	.211(**)	0.085	-0.097	1.000	.154(*)	0.017	.251(**)
P51	0.104	-0.048	.155(*)	0.097	.154(*)	1.000	.271(**)	0.118
P52	0.012	-.142(*)	0.095	.457(**)	0.017	.271(**)	1.000	-0.017
P53	0.047	0.097	-0.080	-0.012	.251(**)	0.118	-0.017	1.000
P54	-0.087	0.034	-0.109	.122(*)	-.345(**)	-0.092	-0.006	-.208(**)
P55	0.020	-0.002	0.054	.457(**)	0.054	.395(**)	.450(**)	0.054
P57	0.121	0.039	.155(*)	.201(**)	.177(**)	.408(**)	.291(**)	0.109
P58	-0.014	-0.043	0.101	.458(**)	-0.060	.311(**)	.325(**)	-0.007

Varables	P54	P55	P57	P58
Exa_1	.286(**)	.240(**)	.151(*)	.189(**)
Sexo	0.101	0.061	-0.079	-0.048
Facultad	0.036	0.121	.199(**)	0.078
P9	0.051	0.064	0.094	-0.020
P60	-0.062	-.175(**)	-.137(*)	-.244(**)
P16	-.126(*)	.147(*)	.186(**)	0.076
P19	-0.122	.242(**)	.276(**)	.167(**)
P47	0.047	0.120	0.109	.152(*)
P25	-0.074	0.111	0.113	0.066
P8	0.013	.169(**)	.165(**)	.193(**)
P10	-0.025	0.062	0.109	0.120
P11	-0.017	-.144(*)	-.132(*)	-0.110
P12	.137(*)	.165(**)	.124(*)	.210(**)
P13	0.000	.381(**)	.330(**)	.395(**)
P14	-0.117	0.016	0.030	-0.059
P22	-0.082	.140(*)	.176(**)	0.093
P24	-.146(*)	-.162(**)	-0.009	-0.099
P26	-.137(*)	0.119	.220(**)	0.027
P1	0.004	0.034	0.017	0.059
P27	0.028	.140(*)	.171(**)	.198(**)
P28	.149(*)	-0.042	-0.031	-0.046
P29	0.083	-.162(**)	-.193(**)	-0.091
P30	-0.078	0.018	-0.027	-0.049
P31	-0.113	-.304(**)	-.145(*)	-.316(**)
P32	0.005	.241(**)	0.109	0.098
P33	.362(**)	-0.020	0.032	-0.009
P34	-.140(*)	-.133(*)	-0.100	-.158(*)
P35	-0.106	0.006	.148(*)	0.046
P36	.137(*)	-0.001	-0.007	0.043
P37	.217(**)	0.049	-0.039	0.099
P38	.369(**)	0.068	-0.020	0.090
P39	-0.097	.318(**)	.313(**)	.209(**)
P40	-.128(*)	-.126(*)	0.006	-.126(*)
P41	0.107	.174(**)	.197(**)	.217(**)
P42	-0.013	-0.044	-0.082	-.164(**)
P43	-0.047	.126(*)	.191(**)	0.113
P44	-0.087	0.020	0.121	-0.014
P45	0.034	-0.002	0.039	-0.043
P48	-0.109	0.054	.155(*)	0.101
P49	.122(*)	.457(**)	.201(**)	.458(**)
P50	-.345(**)	0.054	.177(**)	-0.060
P51	-0.092	.395(**)	.408(**)	.311(**)
P52	-0.006	.450(**)	.291(**)	.325(**)
P53	-.208(**)	0.054	0.109	-0.007
P54	1.000	0.032	-0.050	0.071
P55	0.032	1.000	.408(**)	.447(**)
P57	-0.050	.408(**)	1.000	.447(**)
P58	0.071	.447(**)	.447(**)	1.000

ANEJO 3: Resultados de los criterios de validación

Clasificación por el Método Single							
Grupo	Índice de Validación			Número de estudiantes por grupo			
	Dunn	DB	Silueta				
2	0.439	0.641	0.234	264	1		
3	0.397	0.660	0.142	263	1	1	
4	0.369	0.667	0.103	262	1	1	1
5	0.369	0.812	0.104	258	1	1	4 1
6	0.366	0.748	0.092	258	1	1	3 1 1
7	0.359	0.730	0.079	257	1	1	1 3 1 1
8	0.366	0.710	0.076	256	1	1	1 1 3 1 1
9	0.366	0.704	0.065	255	1	1	1 1 3 1 1 1
10	0.360	0.708	0.028	254	1	1	1 1 1 3 1 1 1

Tabla. 1

Clasificación por el Método Average							
Grupo	Índice de Validación			Número de estudiantes por grupo			
	Dunn	DB	Silueta				
2	0.324	1.518	0.257	259	6		
3	0.324	1.148	0.199	259	2	4	
4	0.257	1.389	0.140	253	2	6	4
5	0.257	1.144	0.104	253	2	4	2 4
6	0.268	1.347	0.094	241	2	12	4 2 4
7	0.227	1.387	0.077	230	11	2	12 4 2 4
8	0.229	1.471	0.070	212	11	2	18 12 4 2 4
9	0.229	1.426	0.061	212	11	2	18 6 6 4 2 4
10	0.232	1.539	0.059	183	11	29	2 18 6 6 4 2 4

Tabla. 2

Clasificación por el Método de Ward											
Grupo	Índice de Validación			Número de estudiantes por grupo							
	Dunn	DB	Silueta								
2	0.197	3.523	0.10086	86	179						
3	0.174	3.071	0.05224	86	98	81					
4	0.180	2.674	0.07215	30	98	81	56				
5	0.180	2.434	0.07153	30	98	53	28	56			
6	0.180	2.327	0.06439	30	45	53	53	28	56		
7	0.180	2.193	0.05551	30	45	33	53	28	20	56	
8	0.181	2.150	0.06835	30	45	33	53	28	20	38	18
9	0.181	2.047	0.07989	30	45	33	53	28	20	24	14 18
10	0.181	1.934	0.08218	30	26	33	53	28	20	24	14 19 18

Tabla. 3

Clasificación por el Método Complete													
Grupo	Índice de Validación			Número de estudiantes por grupo									
	Dunn	DB	Silueta										
2	0.191	3.574	0.099	78	187								
3	0.201	3.135	0.087	17	187	61							
4	0.203	2.492	0.064	17	187	24	37						
5	0.212	2.095	0.065	13	187	24	37	4					
6	0.180	2.430	0.052	13	110	77	24	37	4				
7	0.180	2.222	0.062	13	75	35	77	24	37	4			
8	0.191	2.091	0.070	13	75	35	77	24	13	24	4		
9	0.199	2.043	0.071	13	75	35	73	24	13	24	4	4	
10	0.199	1.993	0.073	13	27	35	73	24	48	13	24	4	4

Tabla. 4

Clasificación por el Método de División - Algoritmo Diana													
Grupo	Índice de Validación			Número de estudiantes por grupo									
	Dunn	DB	Silueta										
2	0.141	3.520	0.075	131	134								
3	0.149	2.809	0.087	131	59	75							
4	0.150	2.452	0.086	51	59	80	75						
5	0.162	2.243	0.088	51	59	80	64	11					
6	0.163	2.161	0.071	51	59	41	39	64	11				
7	0.166	2.222	0.078	51	43	41	39	64	11	16			
8	0.174	2.186	0.072	51	43	41	39	10	54	11	16		
9	0.175	2.113	0.075	51	33	41	39	10	54	11	16	10	
10	0.177	2.075	0.063	51	33	11	39	30	10	54	11	16	10

Tabla. 5

Clasificación por el Método de k-medioides													
Grupo	Índice de Validación			Número de estudiantes por grupo									
	Dunn	DB	Silueta										
2	0.146	3.286	0.079	154	111								
3	0.121	2.771	0.086	113	77	75							
4	0.113	2.615	0.076	89	62	67	47						
5	0.122	2.375	0.084	80	55	53	44	33					
6	0.122	2.261	0.083	25	69	46	44	49	32				
7	0.122	2.087	0.083	24	63	38	40	24	33	43			
8	0.171	1.987	0.091	20	53	35	47	42	21	34	13		
9	0.171	1.923	0.087	21	48	28	37	20	36	21	28	26	
10	0.171	1.853	0.090	20	45	27	28	38	17	19	28	13	30

Tabla. 6

ANEJO 4: Prueba de comparación de medias. Comparación de centroides entre conglomerados

Kruskal-Wallis Test on Fact_1

Grupo	N	Median	Ave Rank	Z
1	20	-0.37750	86.9	-2.80
2	53	0.38400	157.3	2.58
3	35	0.66400	184.7	4.29
4	47	0.43800	165.6	3.22
5	42	-0.94350	58.3	-6.88
6	21	-0.71700	67.9	-4.06
7	34	0.42800	159.5	2.16
8	13	-0.09200	124.6	-0.40
Overall	265		133.0	

H = 96.29 DF = 7 P = 0.000
H = 96.29 DF = 7 P = 0.000 (adjusted for ties)

Kruskal-Wallis Test on Fact_2

Grupo	N	Median	Ave Rank	Z
1	20	0.5565	165.9	2.00
2	53	0.3440	154.8	2.31
3	35	-0.8330	68.5	-5.35
4	47	-0.3430	108.2	-2.45
5	42	-0.2335	129.5	-0.32
6	21	0.5670	177.6	2.78
7	34	0.5935	185.3	4.26
8	13	-1.3570	59.6	-3.54
Overall	265		133.0	

H = 72.67 DF = 7 P = 0.000
H = 72.67 DF = 7 P = 0.000 (adjusted for ties)

Kruskal-Wallis Test on Fact_3

Grupo	N	Median	Ave Rank	Z
1	20	-0.95650	69.7	-3.84
2	53	0.13400	158.8	2.74
3	35	-0.69600	76.4	-4.69
4	47	-0.01000	145.9	1.27
5	42	-0.45450	102.8	-2.78
6	21	0.48500	157.5	1.53
7	34	0.37300	171.3	3.12
8	13	0.69200	188.5	2.68
Overall	265		133.0	

H = 64.06 DF = 7 P = 0.000
H = 64.06 DF = 7 P = 0.000 (adjusted for ties)

Kruskal-Wallis Test on Fact_4

Grupo	N	Median	Ave Rank	Z
1	20	1.1950	215.9	5.03
2	53	-0.3930	111.0	-2.34
3	35	-0.6250	100.3	-2.71
4	47	0.9090	199.7	6.58
5	42	-0.8090	70.7	-5.74
6	21	0.2370	154.1	1.31
7	34	0.1940	154.2	1.73
8	13	-0.9610	53.8	-3.82
Overall	265		133.0	

H = 115.51 DF = 7 P = 0.000

H = 115.51 DF = 7 P = 0.000 (adjusted for ties)

Kruskal-Wallis Test on Fact_5

Grupo	N	Median	Ave Rank	Z
1	20	0.9085	202.8	4.24
2	53	0.1350	148.1	1.60
3	35	-0.7560	69.0	-5.31
4	47	0.4410	173.9	4.03
5	42	0.3825	158.1	2.31
6	21	-0.4020	97.1	-2.24
7	34	-1.0035	56.3	-6.25
8	13	0.5730	166.1	1.60
Overall	265		133.0	

H = 102.00 DF = 7 P = 0.000

H = 102.00 DF = 7 P = 0.000 (adjusted for ties)

Kruskal-Wallis Test on Fact_6

Grupo	N	Median	Ave Rank	Z
1	20	-0.8810	55.2	-4.72
2	53	-0.8680	58.7	-7.89
3	35	0.4220	160.7	2.29
4	47	0.6690	188.8	5.50
5	42	0.3720	149.6	1.53
6	21	1.0110	219.6	5.40
7	34	-0.2770	101.1	-2.60
8	13	0.4140	169.2	1.74
Overall	265		133.0	

H = 137.37 DF = 7 P = 0.000

H = 137.37 DF = 7 P = 0.000 (adjusted for ties)

Kruskal-Wallis Test on Fact_7

Grupo	N	Median	Ave Rank	Z
1	20	0.22700	162.3	1.78
2	53	0.17800	145.5	1.33
3	35	-0.01500	138.2	0.43
4	47	-0.52600	103.5	-2.91
5	42	-0.31300	105.0	-2.58
6	21	0.50700	186.4	3.33
7	34	-0.41550	96.5	-2.97
8	13	1.43800	229.0	4.63
Overall	265		133.0	

H = 55.31 DF = 7 P = 0.000

H = 55.31 DF = 7 P = 0.000 (adjusted for ties)

ruskal-Wallis Test on Fact_8

Grupo	N	Median	Ave Rank	Z
1	20	0.6545	195.6	3.80
2	53	-0.9350	76.5	-6.00
3	35	-0.4600	108.9	-2.00
4	47	0.1810	162.4	2.90
5	42	-0.1125	140.0	0.65
6	21	-0.9700	60.4	-4.53
7	34	0.6410	180.3	3.85
8	13	0.7070	196.8	3.08
Overall	265		133.0	

H = 93.72 DF = 7 P = 0.000

H = 93.72 DF = 7 P = 0.000 (adjusted for ties)

ANEJO 5: Pruebas de ji-cuadrado para las proporciones de estudiantes por facultad, género y escuela de procedencia en los ocho conglomerados.

Tabulated statistics: Grupos, Facultad

Rows: Grupos Columns: Facultad

	1	2	Missing	All
1	6	14	0	20
2	19	34	0	53
3	13	20	2	33
4	10	37	0	47
5	13	28	1	41
6	12	9	0	21
7	13	20	1	33
8	3	10	0	13
All	89	172	*	261

Cell Contents: Count

Pearson Chi-Square = 10.254, DF = 7, P-Value = 0.175
Likelihood Ratio Chi-Square = 10.254, DF = 7, P-Value = 0.175

Tabulated statistics: Grupos, Género

Rows: Grupos Columns: Sexo

	1	2	All
1	9	11	20
2	27	26	53
3	10	25	35
4	17	30	47
5	17	25	42
6	9	12	21
7	16	18	34
8	5	8	13
All	110	155	265

Cell Contents: Count

Pearson Chi-Square = 5.523, DF = 7, P-Value = 0.596
Likelihood Ratio Chi-Square = 5.605, DF = 7, P-Value = 0.587

Tabulated statistics: Grupos, Escuela

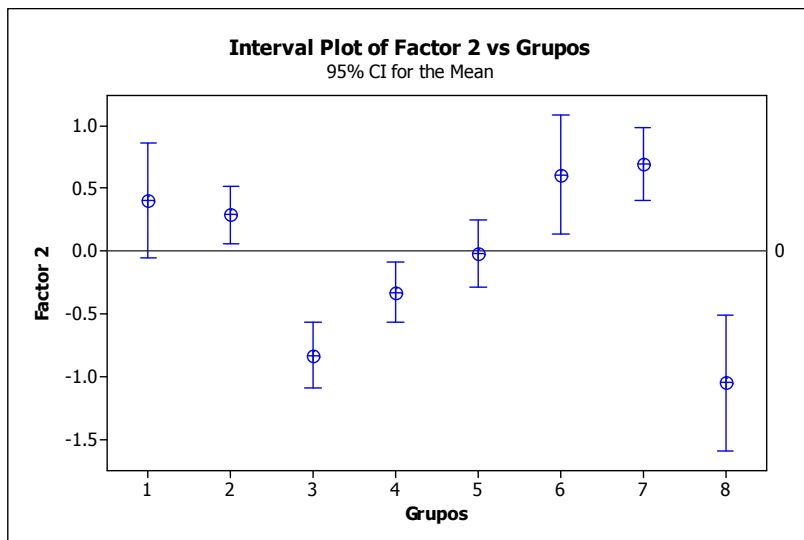
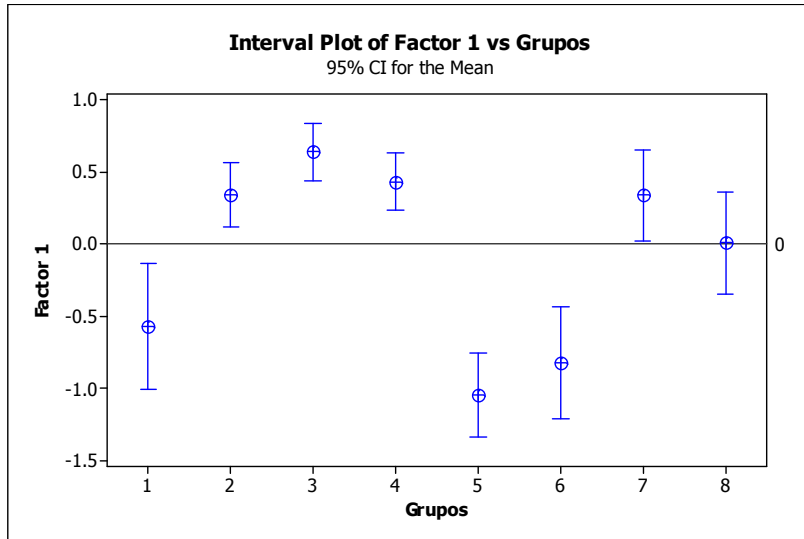
Rows: Grupos Columns: Escuela

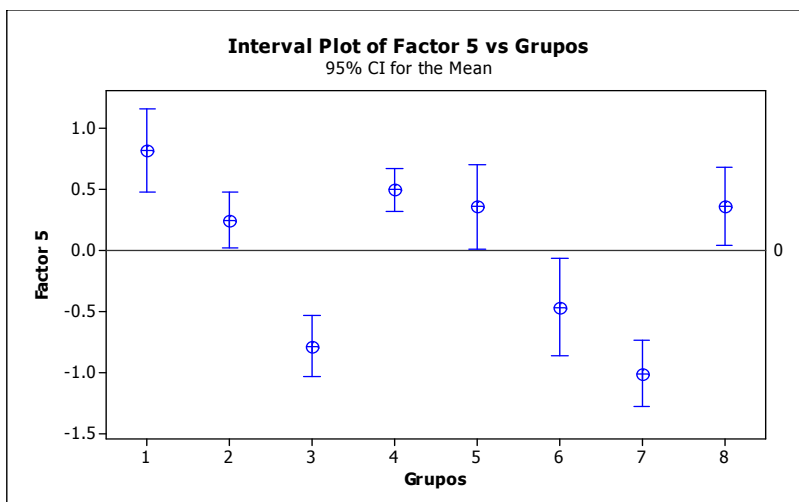
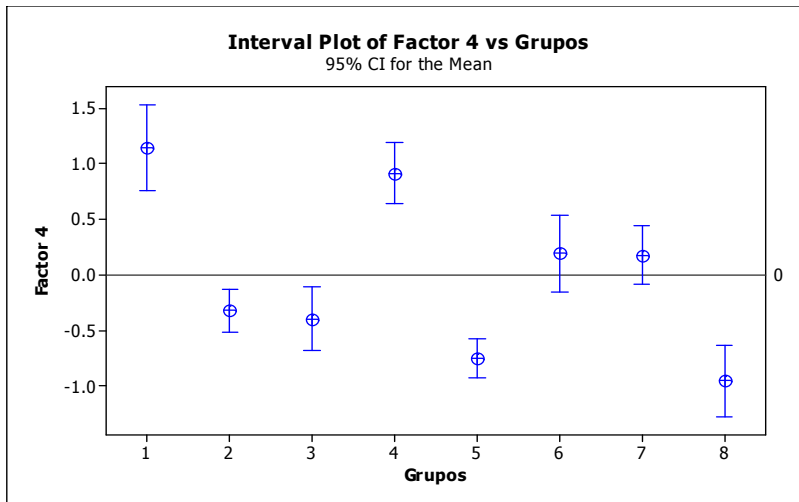
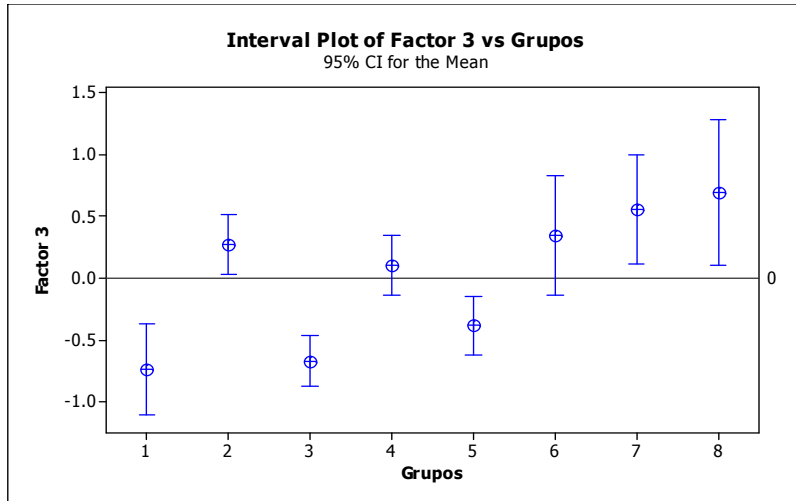
	1	2	Missing	All
1	13	6	1	19
2	33	20	0	53
3	15	20	0	35
4	28	18	1	46
5	22	19	1	41
6	18	3	0	21
7	18	16	0	34
8	10	3	0	13
All	157	105	*	262

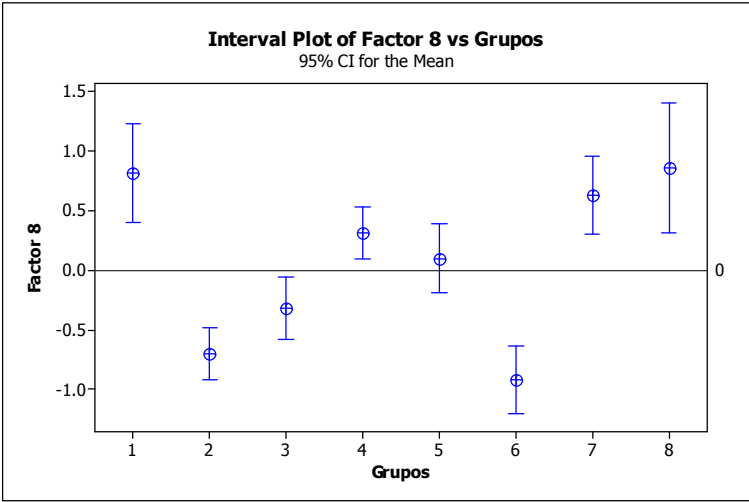
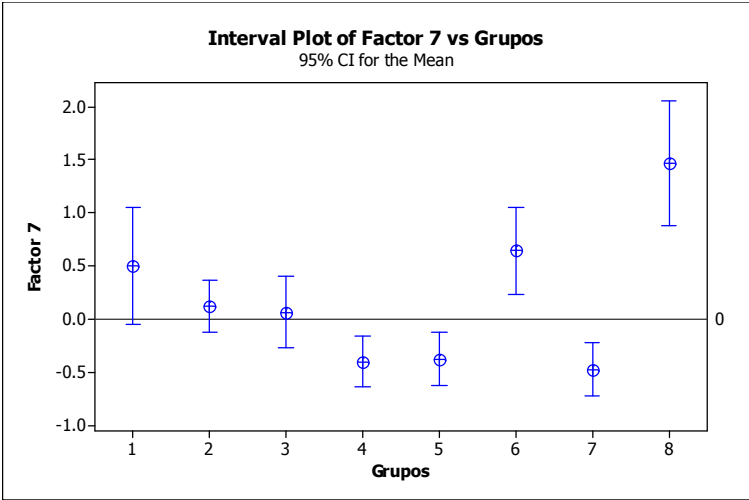
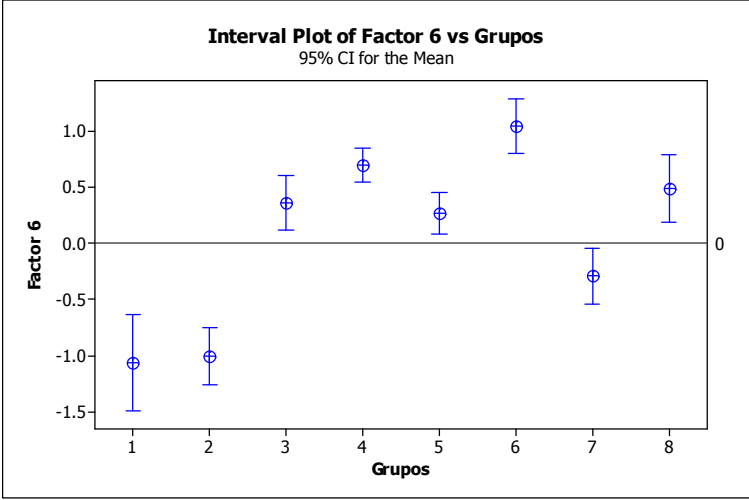
Cell Contents: Count

Pearson Chi-Square = 13.695, DF = 7, P-Value = 0.057
Likelihood Ratio Chi-Square = 14.582, DF = 7, P-Value = 0.042

ANEJO 6: Intervalos de confianza del 95% para la media de cada factor en los ocho conglomerados







ANEJO 7: Pruebas de ji-cuadrado para las proporciones de de estudiantes por facultad, género y escuela de procedencia en los ocho conglomerados.

Tabulated statistics: Grupos 4, Facultad

Rows: Grupos 4 Columns: Facultad

	1	2	Missing	All
1	27	51	3	78
2	24	26	0	50
3	20	59	1	79
4	18	36	0	54
All	89	172	*	261

Cell Contents: Count

Pearson Chi-Square = 7.035, DF = 3, P-Value = 0.071
 Likelihood Ratio Chi-Square = 6.967, DF = 3, P-Value = 0.073

Tabulated statistics: Grupos 4, Género

Rows: Grupos 4 Columns: Sexo

	1	2	All
1	39	42	81
2	22	28	50
3	27	53	80
4	22	32	54
All	110	155	265

Cell Contents: Count

Pearson Chi-Square = 3.595, DF = 3, P-Value = 0.309
 Likelihood Ratio Chi-Square = 3.622, DF = 3, P-Value = 0.305

Tabulated statistics: Grupos 4, Escuela

Rows: Grupos 4 Columns: Escuela

	1	2	Missing	All
1	49	31	1	80
2	33	15	2	48
3	37	43	0	80
4	38	16	0	54
All	157	105	*	262

Cell Contents: Count

Pearson Chi-Square = 10.298, DF = 3, P-Value = 0.016

Likelihood Ratio Chi-Square = 10.293, DF = 3, P-Value = 0.016

ANEJO 8: Intervalos de confianza del 95% para la media de cada factor en cada uno de los cuatro grupos.

