

ANÁLISIS MULTIVARIADO DE LAS CALIFICACIONES DE BACHILLERATO EN COSTA RICA EN EL AÑO 2006

Fabricio Bolaños Guerrero^{1*}

Recepción: 4 de abril de 2008 • Aprobación: 6 de junio de 2008

RESUMEN

En este trabajo se presenta un análisis multivariado de los resultados obtenidos por los estudiantes de último año de todos los colegios de Costa Rica. Se realizaron: un análisis de componentes principales y una clasificación por medio de un árbol de decisión de los 654 centros educativos. Además se tomaron los promedios de cada una de las 20 regiones educativas en las que está dividido el sector educativo nacional y a estos datos se les hizo un análisis similar.

Palabras clave: Análisis de componentes principales, clasificación jerárquica, examen de bachillerato, Costa Rica.

ABSTRACT

This paper presents a multivariate analysis of the results obtained by all high school seniors in Costa Rica. An analysis of the main components and a classification of the 654 institutions—using a decision tree—were carried out. Furthermore, the average of each one of the educational regions in which the national education sector is divided was also taken. This data was analysed in a similar way.

Key words: Main components analysis, hierarchical classification, bachillerato exam, Costa Rica.

* Profesor en la Sede del Pacífico de la Universidad de Costa Rica [fabriciobolanos@gmail.com] - fabricio.bolanos@ucr.ac.cr]

1. Introducción

En Costa Rica los estudiantes de último año de colegio deben realizar –para obtener el título de bachiller en secundaria– los exámenes nacionales de bachillerato; se realizan 6 pruebas en las áreas de español, estudios sociales, matemática, educación cívica, un idioma (inglés o francés) y una del área de ciencias (física, química o biología).

Los datos con que se cuentan se tomaron de 654 centros educativos (los individuos u objetos del análisis) medidos por las variables: tipo de colegio (público, privado o subvencionado), área educativa a la que pertenece la institución, cantidad de estudiantes que presentaron las pruebas, cantidad de estudiantes que

aprobaron el bachillerato, promedio de la institución en las materias de español, estudios sociales, matemática, educación cívica, idioma y ciencias. Además, se tiene una nota promedio de bachillerato que toma en cuenta no solo las notas de estos exámenes sino también la nota que les asigna el colegio mismo, mediante un promedio ponderado.

En vista de que los estudiantes escogen un idioma entre inglés y francés se ponderó el resultado cuando era posible, en otro caso se tomó solo el resultado de un único idioma cuando había sólo uno por colegio. Lo mismo se hizo con las materias científicas de biología, química y física.

Para aclarar la idea del trabajo se presentarán los datos de las 20 regiones educativas del país del año 2006.

Región Educativa	%	PROM EXAMEN	PROM BACH	Estudios					
				Español	Soc	Mate	Cívica	Idioma	Ciencias
SAN JOSE	66,60	74,59	80,99	76,48	77,63	65,68	73,80	82,52	71,17
DESAMPARADOS	55,95	71,27	78,26	74,13	75,99	60,28	73,27	74,25	70,31
PURISCAL	48,50	67,14	76,63	69,40	71,13	57,32	71,50	69,66	65,09
ALAJUELA	65,89	73,16	80,38	75,57	77,10	64,42	74,07	76,91	70,43
SAN RAMON	69,47	74,79	81,08	75,80	78,11	68,22	74,74	78,72	73,81
CARTAGO	55,22	71,91	78,46	74,60	76,74	61,60	74,41	73,94	70,32
TURRIALBA	45,17	68,40	77,95	72,62	74,79	52,58	72,05	71,24	66,22
HEREDIA	64,67	72,79	79,90	74,13	77,08	63,15	73,95	77,84	70,83
LIBERIA	61,24	71,88	79,83	72,52	76,03	62,42	73,17	75,37	71,18
NICOYA	41,35	67,16	75,95	64,48	70,72	57,30	72,97	68,06	69,14
SANTA CRUZ	41,45	67,77	76,04	67,75	73,27	54,09	73,32	72,95	64,49
CAÑAS	39,92	68,96	77,64	71,84	74,95	52,89	73,90	71,44	68,08
PUNTARENAS	51,26	68,14	77,38	68,54	71,20	57,96	71,80	71,90	67,43
AGUIRRE	62,88	71,63	80,20	69,53	77,21	59,18	75,19	77,99	65,97
PEREZ ZELEDON	55,98	70,45	79,26	73,26	76,63	57,83	73,67	71,15	69,90
COTO	49,37	68,02	77,40	68,20	72,04	59,23	71,63	69,63	67,70
LIMON	56,11	70,33	78,19	69,70	75,22	58,65	73,17	75,90	68,67
GUAPILES	56,51	71,00	79,10	72,50	75,83	61,69	74,79	70,62	70,25
SAN CARLOS	43,57	66,82	76,36	70,37	71,72	55,18	70,78	66,91	64,65
UPALA	31,42	62,71	74,15	65,54	68,43	50,54	68,12	60,56	62,21

Elaboración propia a partir de datos de MEP

Además se analizaron los datos por las 20 áreas educativas en las que está dividido el sector educativo nacional, para esto se tomó el promedio de los centros educativos que pertenecen a su respectiva ubicación geográfica.

En las siguientes secciones se presentan: antecedentes de las pruebas de bachillerato en Costa Rica; los resultados del análisis en el cual se toman en cuenta todos los 654 centros educativos del país. En la sección 4 se hace el análisis tomando en cuenta los resultados regionales y en la última sección se presentan las conclusiones y perspectivas de trabajo futuro.

2. Antecedentes

Las pruebas de bachillerato son pruebas sumativas que se aplican en Costa Rica a los egresados de la Educación Diversificada. Se aplicaron hasta 1972, pero se suspendieron desde 1973 hasta 1987; fueron reestablecidas por acuerdo, adoptado en enero de 1988, por el Consejo Superior de Educación. Su aplicación se reinició en diciembre de 1988.

Desde 1988 las pruebas comprendieron las asignaturas de Español (subdividida en Composición y Ortografía, y Gramática y Literatura), Estudios Sociales, Ciencias (a escoger entre Biología, Física y Química), Matemática e Idioma extranjero (a escoger entre Inglés y Francés), a las que se agregó en 2001 la asignatura de Educación Cívica.

Dichas pruebas han estado reguladas por diversos decretos ejecutivos: mediante Decreto Ejecutivo N° 28.457-MEP del 2 de febrero de 2000, publicado en La Gaceta N° 33 del 16 de febrero del 2000, el cual fue reformado por el Decreto Ejecutivo

N° 28729 del 24 de mayo del 2000, publicado en La Gaceta N° 131 del 7 de julio del 2000, se promulgó el Reglamento de Evaluación Educativa de los Aprendizajes. Posteriormente el Consejo Superior de Educación en la sesión N° 63-2000, celebrada el 20 de diciembre del 2000, acogió y aprobó en forma unánime y en firme diversas modificaciones a dicho reglamento, lo cual condujo a la promulgación de un nuevo Reglamento de Evaluación de los Aprendizajes y a la derogatoria del Decreto Ejecutivo N° 28457-MEP y sus reformas, por medio del Decreto Ejecutivo N° 29373-MEP, publicado en el Alcance 27 de La Gaceta N° 070-09 de abril de 2001. Este decreto a su vez fue derogado por el artículo 155 del Reglamento de Evaluación de los Aprendizajes, Decreto Ejecutivo N° 31635 de 4 de febrero de 2004, publicado en La Gaceta N° 30 de 12 de febrero de 2004. Actualmente este último reglamento es el que está vigente, pero sufrió una reforma en 2008 mediante el decreto 34449-MEP publicado en La Gaceta N° 74 del 17 de abril de 2008.

Para las siguientes secciones se debe de tener en cuenta la siguiente **descripción de variables**: Ya que las materias que se evalúan en las pruebas nacionales son Matemática, Español, Estudios Sociales, Cívica, Ciencias e Idioma tanto en los círculos de correlaciones como en los árboles de clasificación donde lo importante es “observar” el comportamiento de las variables; estas se redefinieron como Mate, Español, Estudios Sociales, Cívica, Ciencias e Idioma. La variable PROM BACH representa el promedio de todas las asignaturas de cada colegio en las pruebas y PROMOCION representa el porcentaje de aprobación del colegio.

3. Análisis de todos los colegios

Primero se hace un análisis en el que se incluyen todos los 654 colegios. Para éstos se hizo un análisis de componentes principales (ACP) y una clasificación jerárquica.

3.1. Análisis de componentes principales

Los datos se analizaron aplicando el método de Análisis en Componentes Principales a todos los colegios utilizando el software Statistica, tomando como variables activas únicamente los promedios de notas de las 6 materias, y se proyectó como variables suplementarias el porcentaje de promoción, el promedio de bachillerato y el promedio de los 6 exámenes realizados. El plano principal con la proyección de todos los individuos-colegios y el círculo de correlaciones de las variables se muestran a continuación:

Gráfico 1
Círculo de Correlación y Plano principal

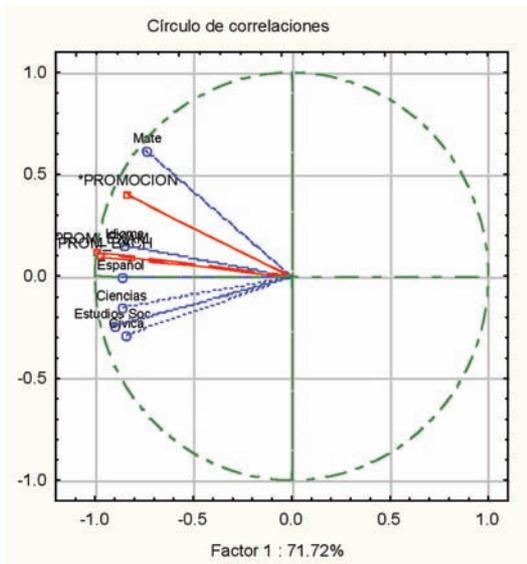
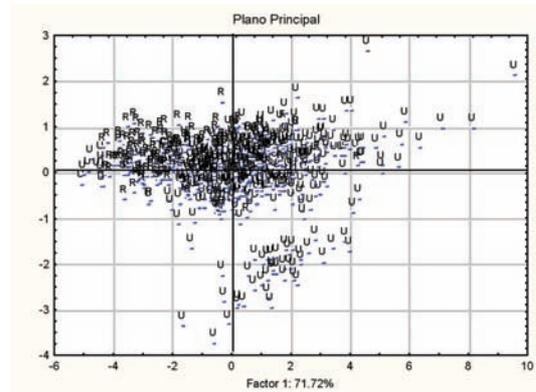


Figura 1
Círculo de correlaciones y plano principal del ACP de los colegios.

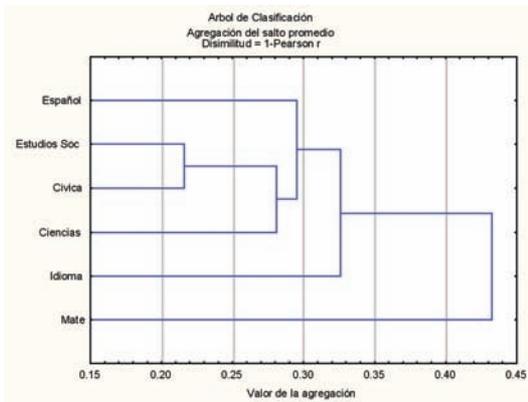


Del círculo de correlaciones podemos ver que todas las materias se comportan de la misma manera, donde los estudiantes alcanzan resultados satisfactorios con excepción de matemática que varía un poco de las demás variables, en la cual los resultados no son tan buenos como en las otras. Esto refleja un efecto talla [1,2] lo que significa que los colegios que son buenos en una materia son buenos en todas. En el plano principal se observan que la gran mayoría de los colegios privados (marcados con R) obtienen buenos resultados; los colegios públicos se representan con la letra U. Curiosamente se observa que existen algunos colegios públicos que superan a los privados (a la izquierda del plano principal), y al determinar cuales son estos colegios se observó que son los Colegios Científicos y los Colegios Humanísticos.

Por otro lado se observan los colegios que obtienen los peores resultados que corresponden a zonas rurales del país y la mayoría de ellos son del tipo Centro Integrado de Educación de Jóvenes y Adultos (CINDEA)

En el plano principal se observa un grupo de colegios públicos hacia abajo que tienen notas promedio pero obtienen bajos resultados en matemática, pues

Figura 2
Clasificación jerárquica de las variables de todos los colegios



se encuentran totalmente opuestos a la variable Mate, este grupo de colegios es el que hace que matemática se diferencie un poco de las demás materias.

3.2. Clasificación de los colegios y las variables

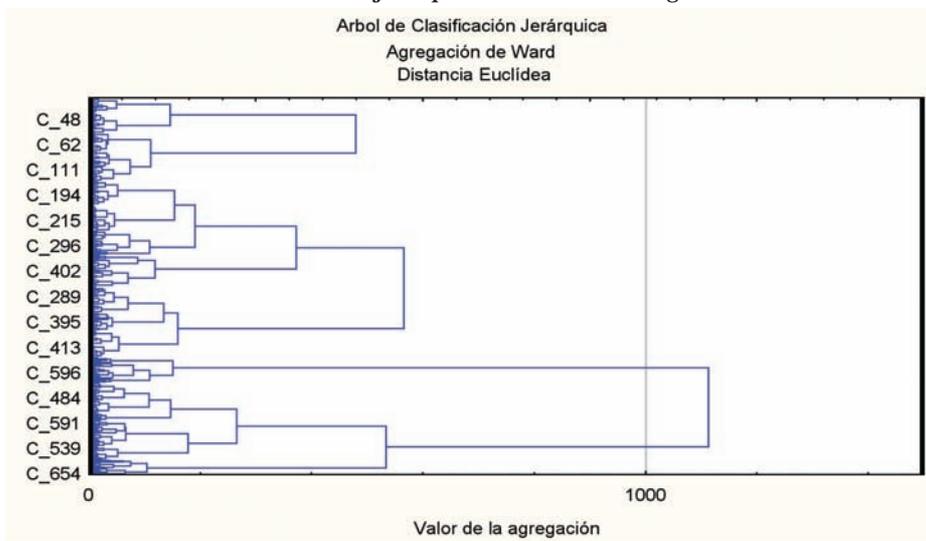
Al hacer la clasificación jerárquica de las variables basándose en la disimilitud definida

por 1 menos la correlación de Pearson y la agregación del salto promedio [2], se observa un resultado parecido entre educación cívica y estudios sociales, y la separación de matemáticas (ver Figura 2), además que se confirman los resultados observados en el círculo de correlaciones de la Figura 1. De aquí se pueden apreciar claramente dos grupos; uno la materia de matemática y otro conformado por las demás asignaturas; estas muestran comportamientos muy parecidos primero entre estudios sociales y cívica, lo cual es de esperar pues por lo general es el mismo docente, luego éstas con ciencias, español e idioma.

Al realizar la clasificación jerárquica de los colegios (ver Figura 3), se observan 4 clases las cuales son:

- los colegios que obtienen las mejores calificaciones,
- los que obtienen las peores calificaciones
- el grupo promedio que no tiene buenos resultados en matemática y
- el grupo promedio.

Figura 3
Clasificación jerárquica de todos los colegios



Análisis por regiones

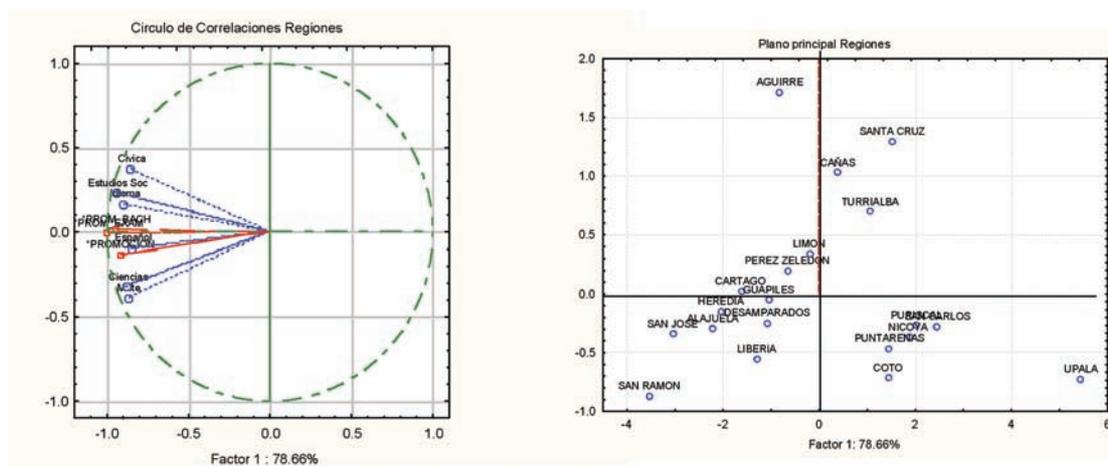
Para este análisis se tomaron los promedios de los colegios de cada región.

El Análisis de Componentes Principales aplicados a las 20 áreas educa-

tivas, con las mismas variables del análisis anterior, arrojó los siguientes resultados de la Figura 4.

En este caso, en el círculo de correlaciones podemos apreciar dos grupos de variables las materias más científicas

Figura 4
Círculo de correlación y plano principal por regiones



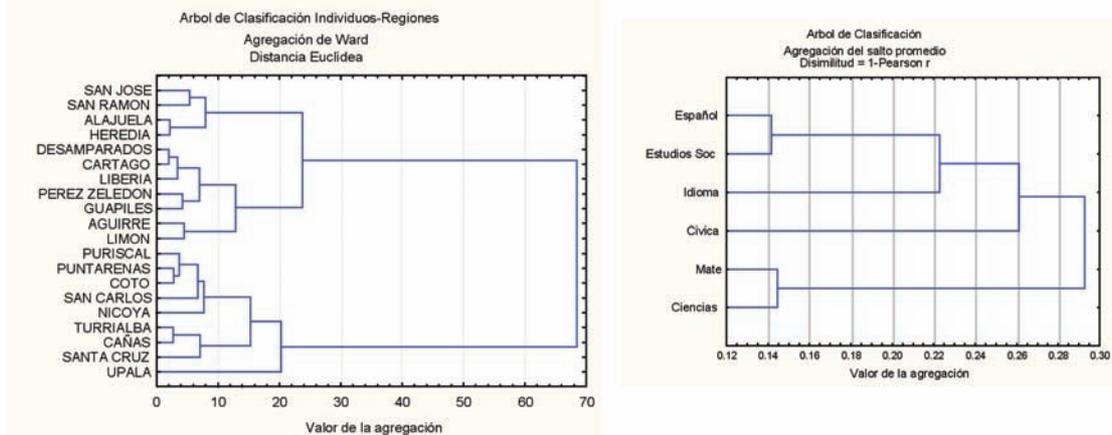
(ciencias y matemáticas) y las teóricas. Por otra parte en el plano principal se pueden notar las regiones del Gran Área Metropolitana que salen mejor en matemáticas y ciencias (San Ramón, San José, Alajuela, Heredia), destacándose de estas el área educativa de San Ramón que fue la que ese año obtuvo los mejores promedios nacionales por regiones lo que sería un análisis importante pero el mismo no es el objetivo del presente trabajo y las que son débiles en estas (Aguirre, Santa Cruz, Cañas y Turrialba) y el grupo que tiene notas bajas en Cívica, Estudios Sociales e Idioma (Upala, Puntarenas, Puriscal, San Carlos, Nicoya y Coto)

Al hacer el análisis de las variables haciendo uso del árbol de decisión se obtiene el resultado que se muestra en la Figura 5.

En este análisis se pueden apreciar tres grandes grupos de regiones: una la Gran Área Metropolitana, donde están las zonas que salieron mejor en los exámenes de bachillerato; otra que tuvo un comportamiento parecido a la primera donde se rezagaron de su ubicación metropolitana las zonas de Desamparados y Cartago, y la última donde se ubican las regiones más alejadas de la capital, Upala, Santa Cruz y Cañas, entre otras.

En cuanto a las variables, se pueden apreciar claramente dos grupos: las materias de matemática y ciencias por un lado, y otro conformado por las demás asignaturas; éstas muestran comportamientos muy parecidos, primero entre estudios sociales y español, luego éstos con idioma y cívica; resultado también bastante previsible ya que las cuatro son bastante memorísticas.

Figura 5
Clasificaciones jerárquicas de las regiones y de las variables



4. Conclusiones

De este trabajo se desprenden las siguientes conclusiones:

En primer lugar, la diferencia que se está presentando entre los colegios públicos y privados ya que estos están obteniendo mejores resultados a nivel del examen de bachillerato. Sin embargo, los colegios científicos y humanistas (que son públicos) obtienen los mejores resultados; esto es posiblemente debido a la selección que se hace de los estudiantes y al nivel de exigencia de los mismos.

También, la brecha que se está notando en los resultados entre los colegios de la Gran Área Metropolitana y las zonas más alejadas del centro del país, ya que las diferencias son bastante notorias.

Finalmente, el desempeño de los estudiantes en la asignatura de matemática es diferente de las demás materias lo que sugiere la existencia de algún problema en este campo a nivel nacional ya que dicho comportamiento no es exclusivo de las zonas más alejadas del país.

Se plantean posibles trabajos posteriores que reflejen la diferenciación entre

colegios públicos y privados a nivel regional y realizar otros análisis de los resultados de las diferentes regiones a través de un lapso de tiempo, comparando con los resultados de otros años, para determinar si el desempeño ha venido mejorando o empeorando para hacer llegar esta información a las direcciones regionales de educación con la intención de tomar las medidas correctivas de ser necesario.

5. Bibliografía

- Castillo, William; González, Jorge; Trejos, Javier (2007) *Análisis Multivariado de Datos*. Preedición, Universidad de Costa Rica, San José. 321 páginas.
- Diday, E.; Lemaire, J.; Pouget, J.; Testu, F. (1982). *Eléments d'Analyse des Données*. París: Dunod.
- Jackson, J. (1986). *Principal Component Analysis*. New York: John Wiley & Sons.
- Lebart, L.; Morineau, A.; Piron, M. (1995). *Statistique Exploratoire Multi-dimensionnelle*. París: Dunod.
- Ministerio de Educación Pública de Costa Rica (2006). *Base de datos de resultados de pruebas de bachillerato año 2006* [archivo digital]. San José.

Poder Ejecutivo de Costa Rica (2000). Decreto Ejecutivo N° 28.457-MEP del 2 de febrero de 2000. *La Gaceta*, N° 33 (16 de febrero).

Poder Ejecutivo de Costa Rica (2000). Decreto Ejecutivo N° 28.729-MEP del 24 de mayo de 2000. *La Gaceta*, N° 131 (7 de julio).

Poder Ejecutivo de Costa Rica (2001). Decreto Ejecutivo N° 29.373-MEP del 2 de febre-

ro de 2001. *La Gaceta*, N° 070 (9 de abril).

Poder Ejecutivo de Costa Rica (2004). Decreto Ejecutivo N° 31.635-MEP del 4 de febrero de 2004. *La Gaceta*, N° 30 (12 de febrero).

Poder Ejecutivo de Costa Rica (2008). Decreto Ejecutivo N° 34.449-MEP. *La Gaceta*, N° 74 (17 de abril).