

Estudio del rendimiento académico de una asignatura con relación a asignaturas de cursos anteriores

Luisa Zúnica, Pedro Blesa, Rosa Alcover, Jorge Más, José M. Valiente

Escuela Técnica Superior de Informática Aplicada.
Universidad Politécnica de Valencia
46022 Valencia

e-mail: lrzunica@eio.upv.es , pblesa@dsic.upv.es , ralcover@eio.upv.es , jmas@fis.upv.es , jvalient@disca.upv.es

Resumen

Se intenta estudiar la relación entre el rendimiento de una asignatura tipo de segundo curso con respecto al rendimiento de las asignaturas del curso anterior, con el fin de detectar si existe una relación más fuerte con la/s asignatura/s posibles prerequisites. No se detecta que sea significativo este efecto, pero en el análisis estadístico realizado se obtienen dos factores o componentes relevantes en el rendimiento de dicha asignatura de segundo curso. Estos factores se pueden interpretar como el rendimiento global del alumno, y su “destreza” en asignaturas propiamente informáticas.

1. Motivación y objetivos

Desde hace algún tiempo han proliferado en la Universidad española diversos estudios sobre rendimiento académico, consecuencia de la creciente percepción que tenemos los docentes de que “algo” no funciona en la Universidad. Las quejas iniciales surgidas en aquellas Facultades y Escuelas con bajas notas de acceso han dado paso a quejas en Centros tradicionalmente considerados de “élite”, donde ya se percibe, también, sensación de fracaso académico. El recurso fácil de achacar todas las culpas al alumnado se entremezcla con otras voces más autocriticas que ven deficiencias tanto en los Centros como en el profesorado y en los planes de estudio vigentes. Consecuencia, en buena medida, de estas preocupaciones, ha sido la contrarreforma de los planes de estudio, que están implantándose actualmente. Ya con la idea de acometer esta mencionada reforma con cierto conocimiento de

causa, el Consejo de Universidades recomendó la realización de este tipo de estudios para evaluar los planes de estudio vigentes [1].

En particular, en la Universidad Politécnica de Valencia se puso en marcha, hace ya una decena de años, el Proyecto de Innovación Educativa (P.I.E), con la intención de transformar las técnicas docentes. Acogidos a este Proyecto, un grupo de profesores de la UPV relacionados con la Escuela Técnica Superior de Informática Aplicada o ETSIA (antigua Escuela Universitaria de Informática) iniciamos un Proyecto [2] con el fin de elaborar una herramienta informática que permitiera extraer, de manera cómoda, sencilla, y sin necesidad de conocimientos informáticos avanzados, datos de rendimiento académico de la abundante base de datos de la que dispone la Universidad [3].

Mediante esta aplicación se han elaborado algunos estudios relativos a diferentes aspectos que pensamos tienen influencia sobre el rendimiento académico, factores tales como la nota de acceso de los alumnos a la Universidad, procedencia de los alumnos (COU, FP.....), organización de los estudios, y otros [4,5]. En la mayoría de ellos hemos estudiado, de manera especial, la problemática existente en primer curso, que nos parece especialmente grave. Pero también tenemos la sensación de que en segundo y tercer curso, donde el fracaso académico también es elevado, la problemática no es la misma que en primero, y pueden afectar factores distintos y de distinta manera.

En este trabajo pretendemos abordar esta última cuestión: analizar, de manera rigurosa, factores que inciden en el rendimiento académico en primer curso y en segundo curso tomando como ejemplo la titulación de Ingeniero Técnico

en Informática de Sistemas en la E.T.S. de Informática Aplicada de la UPV.

2. Estudio realizado

El primer problema que se presenta al intentar hacer estudios de rendimiento académico es el de definir de forma clara qué es el rendimiento académico. En este sentido, y a partir de otros estudios sobre la cuestión [6], definimos el rendimiento académico de un alumno en una asignatura en un curso como un parámetro comprendido entre 0 y 100 en el que se tiene en cuenta la calificación obtenida, la convocatoria en la que se aprobó la asignatura, y los años transcurridos desde la primera matriculación del alumno en la asignatura, según la fórmula:

$$R_{ij} = \frac{0.4^{n_{ij}-1} k_{ij} C_j}{C_j} * 40$$

donde R_{ij} es el rendimiento del alumno i en la asignatura j , n_{ij} es el año en el que el alumno i aprobó la asignatura j (referido al año en el que el alumno se matriculó en la asignatura j por primera vez), C_j es el número total de créditos de la asignatura j y k_{ij} es un coeficiente que depende tanto de la calificación obtenida por el alumno i en la asignatura j , como de la convocatoria en la que se aprobó, pudiendo tomar los valores siguientes:

Calificación	Convocat. Ordinaria	Convocat. Extraordinaria
<5	0	0
≥5 y <7	1,5	1
≥7 y <9	2	1,5
≥9	2,5	2

A partir de esta definición, se pueden definir rendimientos referidos a distintos ámbitos; para este estudio hemos utilizado el rendimiento de una asignatura j durante un periodo de tiempo P , entendido como la suma de los rendimientos, en la asignatura j , de todos los alumnos que durante el periodo P estuvieron matriculados en dicha asignatura:

$$R_j(P) = \frac{\sum_{i \in j(P)} 0.4^{n_{ij}-1} k_{ij}}{\sum_{i \in j(P)} i} * 40$$

El segundo problema que se nos presenta para llevar a cabo el estudio pretendido, es el de definir de forma clara la población a estudiar. En el caso de estudios sobre primer curso es generalizada la técnica de elegir una cohorte de alumnos que ingresan en una titulación un curso determinado. Pero el hacer un estudio sobre segundo curso plantea el problema de definir quiénes son alumnos de segundo curso, máxime en un Centro donde no existen incompatibilidades de matrícula entre asignaturas, y es común que muchos alumnos cursen, de forma simultánea, asignaturas de varios cursos.

Por ello, el procedimiento seguido fue el siguiente:

- Elegimos una asignatura de segundo curso (primer cuatrimestre) que, a nuestro juicio, fuera representativa en cuanto a resultados académicos; dicha asignatura fue Tecnología de Computadores (TCO)
- Seleccionamos todos aquellos alumnos que durante el curso 2000-2001 estuvieron matriculados en dicha asignatura, calculando el rendimiento obtenido por cada alumno en dicho curso.
- Buscamos, en los cinco cursos anteriores al 2000-2001, para cada uno de los alumnos señalados, el rendimiento obtenido en cada una de las asignaturas de primer curso (0 en caso de que la asignatura no hubiera sido aprobada, o su valor en caso contrario).
- Hicimos un estudio estadístico de toda la información obtenida.

Durante el curso 2000-2001 hubo matriculados en la asignatura TCO 295 alumnos. Al consultar las calificaciones obtenidas por dichos alumnos en las asignaturas de primer curso a 110 de estos alumnos les faltaban algunos datos, por lo que no pudieron ser incluidos en el análisis. Así pues, los alumnos estudiados fueron un total de 185.

3. Análisis Estadístico

En principio, un primer análisis que podría responder a nuestro objetivo consistiría en plantear directamente un modelo de regresión lineal múltiple que ligara el rendimiento obtenido por los alumnos en la asignatura Tecnología de Computadores (RTCO) con los correspondientes a la totalidad de las asignaturas de primer curso (variables explicativas). Sin embargo, este análisis presenta el problema de que las variables explicativas están muy relacionadas entre sí (por ejemplo, los alumnos buenos en general obtienen buenos rendimientos en la mayor parte de las asignaturas). En estas situaciones puede resultar estadísticamente imposible diferenciar a partir de los datos los efectos de las diferentes variables, pudiendo incluso resultar que ninguna de ellas aparezca como estadísticamente significativa, pese a que su efecto conjunto sí que lo es claramente.

En consecuencia, se ha optado por plantear, en primer lugar, un análisis de componentes principales (ACP) sobre las doce variables correspondientes a los rendimientos de las asignaturas de primer curso [7,8]. El objetivo de este primer análisis es el de condensar o resumir la información contenida en estas 12 variables en un conjunto reducido de nuevas variables, no observables directamente e incorrelacionadas entre sí. A estas nuevas variables se les denomina componentes. Una vez obtenidas estas componentes, se ha planteado un modelo de regresión tomando RTCO como variable dependiente y las nuevas variables (componentes), obtenidas en el análisis previo, como variables explicativas o independientes.

Para el análisis de datos se ha utilizado el programa *Statgraphics*. La siguiente tabla, en donde cada componente se identifica en la primera columna por su "Factor Number", muestra los resultados del ACP, apreciándose (en la cuarta columna) que las cinco primeras componentes explican el 70.464% de la variabilidad total generada por los rendimientos en los alumnos de las doce asignaturas de primer curso. Se puede observar como las tres primeras componentes son las más relevantes, y las únicas para las que los valores propios asociados son mayores que 1.

Factor Analysis

Factor Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	3,79314	31,610	31,610
2	1,87026	15,586	47,195
3	1,03054	8,588	55,783
4	0,930574	7,755	63,538
5	0,831145	6,926	70,464
6	0,747084	6,226	76,690
7	0,656801	5,473	82,163
8	0,575508	4,796	86,959
9	0,502572	4,188	91,147
10	0,431711	3,598	94,744
11	0,371239	3,094	97,838
12	0,259428	2,162	100,000

En la tabla siguiente se recogen las correlaciones de las doce variables estudiadas (que corresponden a los rendimientos de las doce asignaturas, y se indican en la primera columna, así, RALG es Álgebra) con los cinco factores, o componentes obtenidos:

	Factor 1	Factor 2	Factor 3
RAD1	0,6257	-0,60099	0,0005748
RAD2	0,461046	-0,42729	-0,022731
RALG	0,608918	0,299588	-0,229924
RAM1	0,672737	0,433403	0,0829508
RAM2	0,518438	0,615221	0,0705987
REC1	0,654986	-0,15231	-0,106565
RES1	0,64444	0,3198	-0,101065
RFCO	0,532464	-0,3688	0,0060651
RFFI	0,544929	0,080442	0,0367246
RINT	0,110721	0,040769	0,966623
RIPR	0,502559	-0,60229	0,083306
RMAD	0,638234	0,195028	0,0321449

	Factor 4	Factor 5
RAD1	0,025003	-0,130839
RAD2	0,561262	-0,025720
RALG	-0,021365	0,453123
RAM1	-0,150341	-0,146342
RAM2	-0,0059330	-0,181611
REC1	-0,172309	-0,239751
RES1	0,0458893	-0,395488
RFCO	-0,415436	0,413077
RFFI	0,581528	0,24955
RINT	0,0078179	0,0621313
RIPR	-0,186428	-0,200165
RMAD	-0,120271	0,25084

A partir de la tabla anterior se puede observar que para la primera componente todos los coeficientes

de las diferentes asignaturas son positivos y del mismo orden de magnitud (en torno a 0.5-0.6), salvo el coeficiente de la asignatura Inglés Técnico (RINT) que es del orden de 0.11. Esta primera componente o factor estaría indicando que los alumnos que son buenos estudiantes obtienen en general buenos rendimientos en todas las asignaturas. Por sus características especiales la asignatura Inglés Técnico (INT) se diferencia de las demás puesto que los alumnos pueden llegar a la ETSIA con diferentes niveles de inglés, y esto no está necesariamente relacionado con lo buen o mal estudiante que pueda ser un alumno.

Respecto a la segunda componente o factor, se observan coeficientes tanto positivos como negativos. Si nos fijamos en aquellos cuyo valor absoluto es mayor (RIPR, RAD1, RAD2, RFCO, RAM1, RAM2, RES1) pueden diferenciarse, para un mismo nivel promedio global recogido por el primer factor, dos grupos de estudiantes: aquellos que presentan en primer curso cierta destreza en las asignaturas relacionadas con la programación y materias propias de la Informática (Introducción a la Programación, Algoritmos y Estructuras de Datos 1 y 2, Fundamentos de Computadores) y aquellos alumnos que la presentan en las asignaturas más relacionadas con las matemáticas como pueden ser Análisis Matemático 1 y 2 y Estadística 1. Opinamos que el primer grupo de alumnos puede proceder de FP2 mientras que el segundo puede proceder de COU, aunque habría que hacer otros estudios para comprobar esta afirmación.

La tercera componente pone de manifiesto la diferencia entre el rendimiento en la asignatura Inglés Técnico (con un coeficiente de 0.97) y el del resto de las asignaturas de primer curso, con coeficientes marcadamente inferiores.

A efectos de claridad y con el fin de mostrar gráficamente las afirmaciones anteriores, en la figura 1 se han representado las correlaciones de las doce variables estudiadas con las dos primeras componentes o factores.

Nótese como, respecto a la primera componente, todos los puntos representados se sitúan a la derecha (correlaciones positivas). La segunda componente está diferenciando rendimientos en asignaturas propiamente informáticas (correlaciones negativas) de aquellos obtenidos en asignaturas más matemáticas (correlaciones positivas), quedando el rendimiento

de la asignatura Inglés Técnico en una situación intermedia.

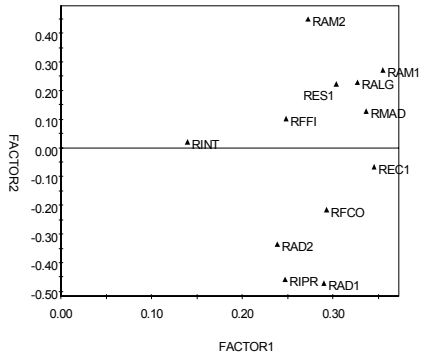


Figura 1. Gráfico de las correlaciones de las doce variables con las dos primeras componentes

Una vez obtenidas las componentes principales el siguiente paso ha consistido en plantear un modelo de regresión que permita analizar el efecto de éstas sobre el promedio del rendimiento en la asignatura de segundo curso TCO. El modelo es el siguiente: $E(RTCO) = \beta_0 + \beta_1 * FACTOR1 + \beta_2 * FACTOR2 + \beta_3 * FACTOR3 + \beta_4 * FACTOR4 + \beta_5 * FACTOR5$, donde los parámetros β_i recogen y cuantifican el efecto de las componentes (FACTOR_i) sobre el rendimiento medio en TCO. La siguiente tabla recoge los resultados de este análisis de regresión:

Multiple Regression Analysis

Dependent variable: RTCO

Parameter	Estimate	Standard Error	T	P-Value
CONSTANT	13,46	1,36	9,92	0,0000
FACTOR1	1,467	0,36	4,09	0,0001
FACTOR2	-2,319	0,73	-3,19	0,0017
FACTOR3	0,113	1,32	0,08	0,9317
FACTOR4	0,923	1,46	0,63	0,5285
FACTOR5	1,932	1,64	1,18	0,2394

R-squared = 13,6246

En el análisis se observa que únicamente las dos primeras componentes resultan estadísticamente significativas (valores en la columna *P-value* inferiores a 0.05). Además la estimación de los coeficientes de las dos primeras componentes (columna *Estimate*) presenta signos contrarios. Respecto al efecto de la primera componente, el signo positivo obtenido implica la conclusión, bastante esperable por cierto, de que los alumnos que al ser buenos estudiantes obtienen buenos rendimientos en primer curso también obtienen buenos rendimientos en la asignatura TCO de 2º curso.

En cuanto al efecto de la segunda componente, el signo negativo obtenido para β_2 indica que en general los alumnos con valores negativos en la misma obtienen mejores calificaciones en TCO. Estos alumnos serían aquéllos en los que predomina la destreza en las materias más propias de la Informática frente a la correspondiente a materias de carácter más matemático.

Sin embargo, el valor obtenido para el parámetro *R-squared* indica que el 13.6% de la variabilidad del rendimiento de los alumnos en la asignatura TCO viene explicado por las variables consideradas en el modelo planteado, lo que resulta un porcentaje bastante reducido. El modelo es, por tanto susceptible de mejora puesto que otros factores no contemplados en él ayudarían a explicar mejor dicha variabilidad.

4. Conclusiones

En su origen, este trabajo estaba enfocado principalmente a estudiar si existía una relación importante entre una determinada asignatura, y cualquier otra que fuese prerequisite, y diferente en comparación con la de otras asignaturas previas. Para ello se escogió una asignatura típica de segundo curso de nuestra escuela (TCO: Tecnología de Computadoras) y tomamos como población todos los matriculados en dicha asignatura en un curso, obteniendo el rendimiento académico de cada individuo en las asignaturas previas hasta los cinco cursos anteriores.

En esta situación, un modelo de regresión múltiple no es directamente aplicable a las variables manejadas (rendimiento de las diferentes asignaturas), dada la elevada relación entre las

variables explicativas. Por ello se ha planteado, un análisis de regresión tomando como variables explicativas las componentes principales obtenidas en un análisis previo (análisis de componentes principales).

Las dos primeras componentes resultan estadísticamente significativas, y tienen, en nuestra opinión, una interpretación clara. La primera componente estaría indicando que los alumnos que, en general, son buenos estudiantes obtienen también buenos rendimientos en la asignatura TCO. En cuanto a la segunda componente, se ha visto que separa claramente dos grupos de estudiantes: los alumnos que presentan ciertas habilidades en asignaturas relacionadas directamente con la informática, tales como programación y estructura de computadoras, y los que las muestran en las asignaturas relacionadas con las matemáticas. El primer grupo de alumnos podría proceder de formación profesional, y el segundo de COU. Esta segunda componente influye en el rendimiento de TCO en el sentido de que los alumnos que muestran una mayor destreza en las asignaturas informáticas obtienen también mejores rendimientos en TCO.

Dado el escaso porcentaje de variabilidad del rendimiento de TCO explicado por el modelo pensamos que este es susceptible de mejora puesto que otros posibles factores no contemplados en él ayudarían a explicar mejor dicha variabilidad. Es nuestra intención seguir con el análisis, y ofrecer en un futuro nuevas conclusiones.

Agradecimientos

A Hanane Tachafait y Jesús Lafuente por su esfuerzo en el desarrollo de la aplicación informática, y al Instituto de Ciencias de la Educación (ICE) de la Universidad Politécnica de Valencia por su financiación.

Referencias

- [1] B.O.E. de 17 de Enero de 1997
- [2] P.I.D. 9052 de la UPV. Libro de resúmenes de los P.I.D. de la UPV. 2001.
- [3] Más, Alcover, y otros. *Una herramienta informática para un estudio multidimensional del rendimiento académico en la EUI de la*

- UPV. Libro de resúmenes del VII CUIE. Huelva. 1999.
- [4] Benlloch, J.V; Bonet, E. y otros. *Estudio comparado del rendimiento de los alumnos de primer curso procedentes de COU frente a los alumnos procedentes de FP*. Libro de resúmenes de las IV Jornadas de enseñanza universitaria de Informática. Andorra. 1998.
- [5] Más, J. Meseguer, J.M^a *Estudio sobre la heterogeneidad de conocimientos básicos en alumnos de primer curso de universidades politécnicas.* Libro de resúmenes de las VI Jornadas de enseñanza universitaria de Informática. Alcalá de Henares. 2000.
- [6] González, R. M. *Rendimiento académico en la UPM: estudio longitudinal en primer ciclo, Vol. 1 y 2*. I.C.E. de la UPM, 1993.
- [7] Mardia, K., Kent, J. Y Bibby, J. (1979). *Multivariate Analysis*. Ed. Academic Press, London
- [8] Sharma, S. (1996). *Applied Multivariate Techniques*. Ed Wiley