VALIDACIÓN DE UN INSTRUMENTO PARA LA CALIFICACIÓN DE EXÁMENES DE MATEMÁTICAS

Testing an instrument for the marking of mathematical exams

Elena Mengual, Núria Gorgorió y Lluís Albarracín

Universitat Autònoma de Barcelona

Resumen

En este trabajo probamos empíricamente el modelo teórico propuesto por Gairín, Muñoz y Oller (2012) para la calificación de la prueba de matemáticas en las PAU. Concretamos el modelo en una guía para la corrección de dos problemas que entregamos a 4 profesores para que calificasen las respuestas a dichos problemas elaboradas por los alumnos. A partir del análisis de las respuestas de los alumnos que generan variabilidad en las calificaciones recogidas sugerimos criterios que permiten una reducción de la variabilidad de las calificaciones, identificando distintos factores a considerar para refinar el modelo.

Palabras clave*: evaluación, pruebas de acceso a la universidad, calificación, matemáticas.*

Abstract

In the following study we empirically test the theoretical model proposed by Gairín, Muñoz and Oller (2012) for grading the mathematics tests in university entrance exams. We specify this model in a guide to marking two problems, which were handed to 4 teachers to correct the answers submitted by the students. Based on the analysis of the students' answers, which give rise to a wide range in grades, we suggest criteria which may allow to reduce the variability of the final marks, by identifying different factors to refine our model.

Keywords*: assessment, university entrance examination, marking, mathematics.*

Mengual, E., Gorgorió, N. y Albarracín, L. (2013). Validación de un instrumento para la calificación de exámenes de matemáticas. En A. Berciano, G. Gutiérrez, A. Estepa y N. Climent (Eds.), *Investigación en Educación Matemática XVII* (pp. 367-381). Bilbao: SEIEM.

JUSTIFICACIÓN

El acceso a los estudios universitarios en España depende de la superación de las Pruebas de Acceso a la Universidad (PAU en adelante), comúnmente llamadas selectividad. En este contexto, en ocasiones un estudiante accede a una plaza de estudio deseada por tener unas centésimas más que otro aspirante (Escudero y Bueno, 1994). Dada su relevancia para determinar el futuro de los estudiantes sería deseable que las calificaciones obtenidas fueran consistentes con su nivel de aprendizaje. Sin embargo diferentes estudios han coincidido en la existencia de diferencias entre correctores y en la necesidad de mejorar el sistema de corrección para aumentar la fiabilidad de las pruebas (Escudero y Bueno, 1994; Grau, Cuxart y Martí-Recober, 2002).

La prueba de matemáticas de las PAU constituye una evaluación sumativa, dado que es la comprobación del grado en que el alumno ha alcanzado los conocimientos, destrezas y habilidades matemáticas al final de una etapa educativa. Estamos ante una evaluación nomotética criterial, ya que la valoración del sujeto viene dada por unos criterios de corrección que elaboran los armonizadores de las PAU, externos al entorno educativo del alumno. Además, este alumno es anónimo, lo que conlleva a que el corrector desconozca sus capacidades y sus posibilidades de desarrollo. Los criterios de evaluación deberían permitir valorar al alumnado de forma homogénea y determinar en qué medida ha desarrollado aprendizaje. Sin embargo, en el caso concreto de las pruebas de matemáticas de las PAU, la corrección y calificación no son tareas triviales, en particular, se ha constatado que los correctores no corrigen igual la primera que la última prueba (Casanova, 1998).

Este hecho influye en la variabilidad de las calificaciones que un mismo corrector asigna a respuestas equivalentes de diferentes alumnos a un mismo ejercicio, como han comprobado Gairín, Muñoz y Oller (2012). Además, estos autores, han constatado que distintos correctores califican de forma diferente soluciones equivalentes dadas por los alumnos. Desde esta perspectiva, presentan un modelo para la elaboración de unos criterios de corrección que debería permitir una reducción en la variabilidad de las calificaciones otorgadas por los correctores.

En este artículo presentamos una concreción del modelo teórico propuesto por dichos autores, lo concretamos en una guía para la corrección de unos problemas de las PAU y analizamos hasta qué punto dicha concreción permite una reducción de la variabilidad de las calificaciones.

EL MODELO DE GAIRÍN, MUÑOZ Y OLLER

El estudio de Gairín, Muñoz y Oller (2012) tiene como objetivo reducir el grado de subjetividad propio de la evaluación mediante la elaboración de un marco que permita definir unos criterios de evaluación más precisos y claros para la calificación de los exámenes de matemáticas de las PAU. Inicialmente, identifican diversos fenómenos no deseables en la corrección de esta prueba, como son: distintos correctores penalizan de forma diferente un mismo error; se penalizan los errores con independencia de las exigencias de enunciado del problema; o el corrector considera un error como muy grave y da por finalizada la corrección. A la vista de estos fenómenos, Gairín, Muñoz y Oller (2012) afirman que los correctores podrían evaluar en base a su experiencia y, por tanto, hacen una valoración de lo que el alumno logra en el ejercicio distinta de la que pueda hacer otro corrector.

Estos autores afirman que no se debe penalizar tareas que no se solicitan explícitamente en el ejercicio y que para poder determinar un valor numérico de la penalización de un error se necesita una jerarquización de las tareas que se llevan a cabo en la resolución del problema. Distinguen las tareas necesarias para resolver un ejercicio en principales, auxiliares específicas (en las que aparecen contenidos específicos de los contenidos a evaluar) y auxiliares generales (basadas en los conocimientos trabajados a lo largo de su formación matemática). En particular, determinan que existen dos tipos de preguntas en la prueba de matemáticas de las PAU en función de las tareas que se realizan para resolverlos –preguntas de categoría C1 y de categoría C2. Las de categoría C1 son aquellas cuyo objetivo principal es el dominio de una técnica de cálculo; para resolver este tipo de pregunta se necesita realizar tareas principales y auxiliares generales. En las preguntas de categoría C2, el objetivo principal es el dominio de un concepto matemático; para resolver este tipo de cuestiones se necesita realizar tareas principales, auxiliares específicas y auxiliares generales.

Gairín, Muñoz y Oller (2012) proponen el siguiente modelo de penalización de errores al que denominan "modelo de tercios":

TAREAS AUXILIARES GENERALES
Conjunto de errores ≤ 1/3 de la puntuación
Continuar proceso calificación

TAREAS AUXILIARES ESPECÍFICAS
Conjunto de errores ≤ 2/3 de la puntuación
Continuar proceso calificación

TAREAS PRINCIPALES
Conjunto de errores ≤ total puntuación
Puede finalizar el proceso de calificación

Este modelo permite que los errores cometidos en las tareas auxiliares no eclipsen el conocimiento matemático presente en la resolución, ya que el proceso de calificación debe continuar y el corrector debe estimar en qué grado el alumno ha alcanzado el objetivo principal de la pregunta. Ahora bien, si el error está relacionado con el objetivo principal, el corrector puede penalizar el ejercicio hasta con el total de la puntuación otorgada a ese problema. Si la resolución del problema lleva implícitos más de un objetivo principal es necesario asignar puntuaciones independientes a cada objetivo para poder jerarquizar las tareas en cada objetivo y aplicar el modelo.

La aplicación de este modelo debería ayudar a disminuir la variabilidad existente entre dos correcciones de un mismo ejercicio ya que, al poner estos límites el corrector debe dejar a un lado sus propios criterios. Estudios como los de Cuxart, Martí-Recober y Ferrer (1997) y Grau, Cuxart y Martí-Recober (2002), afirman que con unos criterios de corrección más fiables se obtiene una menor variabilidad en la las notas de matemáticas de las PAU. Estos estudios además concretan que los coordinadores deberían comunicar a los correctores qué es lo que quieren valorar en el examen y en cada pregunta en concreto (Cuxart, et al., 1997), algo que recoge el modelo de Gairín, Muñoz y Oller (2012).

OBJETIVOS DEL ESTUDIO

Partiendo del modelo de Gairín, Muñoz y Oller –en adelante modelo GMO–, en este trabajo pretendemos, a partir del la concreción de una guía de corrección basada en el modelo GMO

- Contrastar si la utilización de dicho instrumento por profesores con las características necesarias para ser corrector de la prueba de matemáticas de las PAU reduce la variabilidad de las calificaciones que asignan los correctores, comparada con la variabilidad mostrada en Gairín, Muñoz y Oller (2012).

- Identificar posibles causas que contribuyen a que se produzca variabilidad entre las calificaciones que otorgan distintos profesores a una misma pregunta.

Los resultados de este estudio permitirían refinar los criterios de calificación GMO y establecer el proceso por el cual, dada un ejercicio de matemáticas se establece la guía práctica para su corrección.

METODOLOGÍA

Presentamos un estudio experimental [3] de tipo interpretativo en el que en el que analizamos cuantitativa y cualitativamente las correcciones de 4 profesores que imparten clases en centros de secundaria y bachillerato, con distinto grados de experiencia, a las respuestas de pruebas de matemáticas de las PAU siguiendo el modelo de calificación GMO. Queremos ver de qué forma los profesores aplican el modelo, estudiar la variabilidad en las calificaciones obtenidas y comprobar las limitaciones de la concreción utilizada del modelo.

Disponemos de 76 exámenes de matemáticas de las PAU de Zaragoza de Septiembre del 2010. De las dos opciones, A y B, que se presentan al alumno para elegir tomamos las pruebas correspondientes a la opción A, dado que 67 la prefieren a la opción B que fue escogida sólo por 9. Cada opción contiene 4 preguntas y para nuestro estudio elegimos A3 y A4, ambas con varios apartados, que contienen preguntas tanto de categoría C1 como de categoría C2. Es decir, cada profesor corrige medio examen (5 puntos) y consideramos como datos sus correcciones y utilizamos las calificaciones de cada uno de los diferentes apartados.

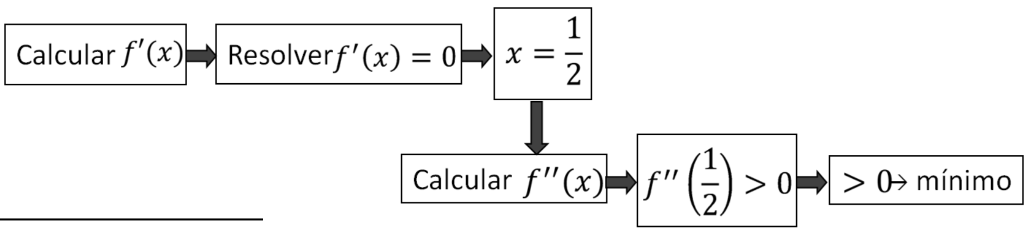
Instrumento para la recogida de datos

Una vez seleccionadas las preguntas a corregir se elabora la guía que concreta los criterios de corrección propuestos en el modelo GMO. A continuación, de los dos apartados que tiene la pregunta A3, mostramos el proceso de desarrollo de la guía para el apartado a).Este apartado es de categoría C2, lo cual implica que para su resolución se necesita el dominio de un concepto matemático y el desarrollo de tareas principales, auxiliares específicas y auxiliares generales. El resto de criterios para los demás ejercicios se desarrollan de forma análoga.

A3. Sea la función $f(x) = xlnx + (1-x)\ln(1-x) \;\; x\epsilon(0,1)$

a) Calcular sus extremos relativos. (1,5 puntos)

El objetivo principal de A3 a) es evidenciar el dominio de un concepto matemático, en este caso de punto relativo y su clasificación, es decir, determinar si es máximo o mínimo. Antes de definir las tareas que requiere dar respuesta a esta pregunta, identificamos los pasos que el alumno podría seguir hasta la solución, que quedan recogidos en el siguiente diagrama:



[3] Este estudio se corresponde con la fase inicial de la investigación que la primera autora está realizando en el marco del Máster de Investigación en Didáctica de las Matemáticas y las Ciencias de la Universitat Autònoma de Barcelona.

A partir de estos pasos, proponemos unos criterios de evaluación para este ejercicio que, una vez revisados por expertos, son los que se muestran a continuación:

TAREAS PRINCIPALES:
- Concepto:
  + Extremo relativo
  + Clasificación del extremo relativo
- Procedimentales: Procedimiento para el cálculo de extremos relativos.

TAREAS AUXILIARES:
- Específicas:
  + Cálculo de derivadas: f'(x) y f''(x)
  + Sustituir un punto en una función f''(x)
- Generales:
  + De tipo algebraico: al simplificar las derivadas, al resolver ecuaciones y/o inecuaciones.
  + De tipo aritmético: al sustituir el punto en la función.

PENALIZACIÓN (el apartado A3 a) vale 1,5 puntos):
- Como mucho 0.5 puntos de penalización por el CONJUNTO de los errores en tareas auxiliares generales.
- Como mucho 1 punto por el CONJUNTO de errores en tareas auxiliares generales y específicas.
- Se puede penalizar con la totalidad por errores en tareas principales.

Criterios de corrección para A3 a)

El resto de criterios correspondientes al resto de preguntas se desarrollaron de forma análoga y son los que se proporciona a los profesores participantes para que realice las correcciones.

ANÁLISIS DE DATOS

Análisis global de los datos

Una vez elaborados estos criterios de corrección para los ejercicios A3 y A4, los 4 profesores corrigieron en base a ellos las respuestas de los 67 alumnos. En la siguiente tabla se muestran la nota media que otorga cada uno de los correctores, considerando únicamente los dos ejercicios que estamos teniendo en cuenta.

Tabla 1. Nota media por prueba (M)

	M en media prueba
Corrector 1	1,8641791
Corrector 2	2,1380597
Corrector 3	1,95835821

Corrector 4	2,017164179

Vemos que cuando tenemos en cuenta sólo A3 y A4, la máxima diferencia entre los cuatro correctores es de 0,27 puntos aproximadamente. Gairín, Muñoz y Oller (2012) constataron un diferencia máxima entre dos correctores de aproximadamente dos puntos, en nuestro estudio, en medio examen la máxima diferencia entre los cuatro es 0,27. A pesar de lo incipiente de nuestro estudio, comparando nuestros resultados con los de estos autores, podemos pensar que la concreción y entrega de la guía para la corrección puede reducir considerablemente la variabilidad, ya que 0,27 sobre 5 es menor que 2 sobre 10. De manera análoga, para cada uno de los ejercicios, A3 y A4, a partir de las correcciones que los 4 profesores asignan a cada estudiante se calculan la media y la desviación típica de las puntuaciones de los alumnos en cada apartado de las preguntas. Viendo las diferencias entre los correctores decidimos considerar como lindar de lo aceptable una desviación típica del 15% de la puntuación total de cada apartado de A3 y A4. Hemos elegido el 15% en cada apartado porque de esta forma la desviación máxima entre correctores que consideramos adecuada para la nota final del examen no debería exceder de 1,5 puntos que es menor que los 2 puntos de variación recogidos por Gairín, Muñoz y Oller (2012). Si consideramos las posibles compensaciones entre las calificaciones de los diversos apartados, la desviación total debería ser considerablemente menor. Al respecto de lindar, no existe un valor consensuado para este margen, con lo que en esta primera aproximación escogemos el 15% para cada apartado porque esperamos que en todo el examen sea menor. Las tablas recogen estos resultados.

Tabla 2. Desviación típica (*DT*)

A3	Coinciden las notas al 100%	*DT* de las notas de < 15% puntuación ejercicio y ≠ 0	*DT* de las notas de > 15% puntuación ejercicio y ≠ 0	Desviación máxima
a)	40’37%	53’22%	2’41%	0,3129
b)	40’99%	54’09%	4’92%	0,3142

Tabla 3. Desviación típica (*DT*)

A4	Coinciden las notas al 100%	*DT* de las notas de < 15% puntuación ejercicio y ≠ 0	*DT* de las notas de > 15% puntuación ejercicio y ≠ 0	Desviación máxima
a)	73’33%	18’33%	8’33%	0,3714
b)	70’21%	23’41%	6’38%	0,3247
c)	74’58%	5’08%	20’34%	0,3247

Observamos que, los casos en que la desviación típica es superior a la que hemos definido como aceptable –el 15% de la puntuación total del ejercicio– varían entre un 2,41% en el caso de A3.a) y un 20,34% en el caso de A4.c). Para comprender las diferencias en las correcciones que generan dichos márgenes de variabilidad, hemos analizado las correcciones de los ejercicios. El análisis de dichas correcciones tiene como propósito estudiar si es posible detallar más los criterios de corrección para que disminuya la variabilidad.

Análisis detallado de las respuestas que generan variabilidad

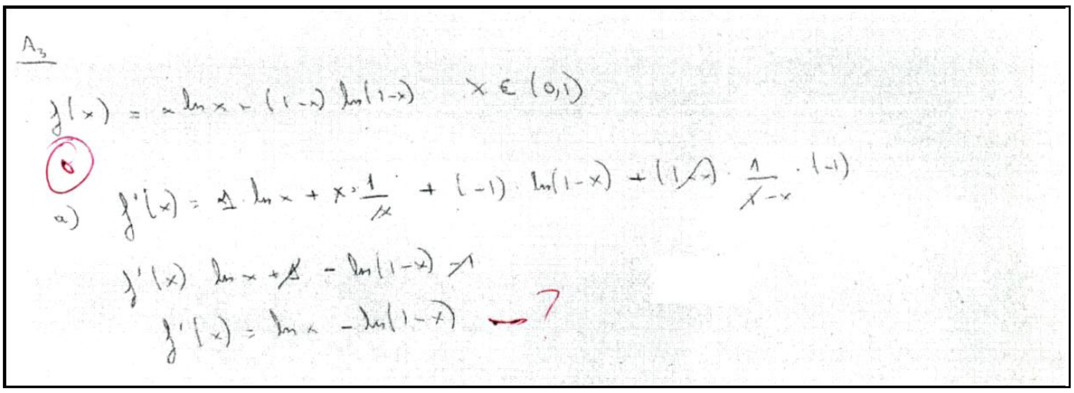
A continuación, mostramos qué sucede en los distintos apartados de las preguntas A3 y A4 y algunas dificultades que los correctores han puesto de manifiesto al calificar estas preguntas.

Apartado A3 a)

Su objetivo es ver si el alumno domina el concepto de extremo relativo (categoría C2). Las diferencias en la calificación de estos ejercicios se deben principalmente a que los alumnos no consiguen terminar el ejercicio, lo que implica que cada corrector debe tomar decisiones que no aparecen explícitas en los criterios de evaluación. El modelo GMO indica la máxima penalización posible en cada tarea en el caso de acabar el ejercicio. Sin embargo, no establece cómo puntuar una pregunta no terminada. El modelo GMO tampoco establece cómo se debe calificar al alumno que resuelve bien el problema pero que se equivoque al redactar las conclusiones.

Vemos, por ejemplo, que los correctores han calificado de forma distinta que el alumno sea capaz de calcular f'(x) y/o f''(x) cuando el alumno no ha completado la resolución del ejercicio:

CORRECTOR 1



A3

$f(x) = x \ln x + (1-x)\ln(1-x) \quad x \in (0,1)$

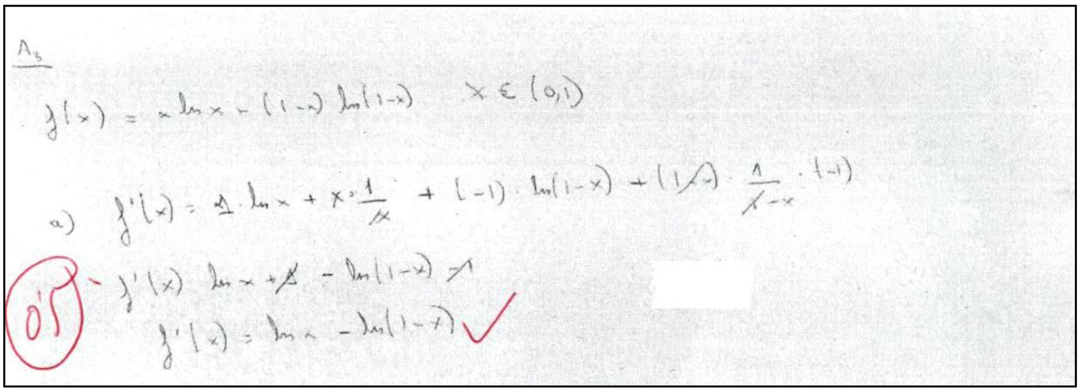
0

a) $f'(x) = 1\cdot\ln x + x\cdot\frac{1}{x} + (-1)\cdot\ln(1-x) + (1-x)\cdot\frac{1}{1-x}\cdot(-1)$

$f'(x)\ \ln x + 1 - \ln(1-x) - 1$

$f'(x) = \ln x - \ln(1-x)$?

CORRECTOR 2



A3

$f(x) = x \ln x + (1-x)\ln(1-x) \quad x \in (0,1)$

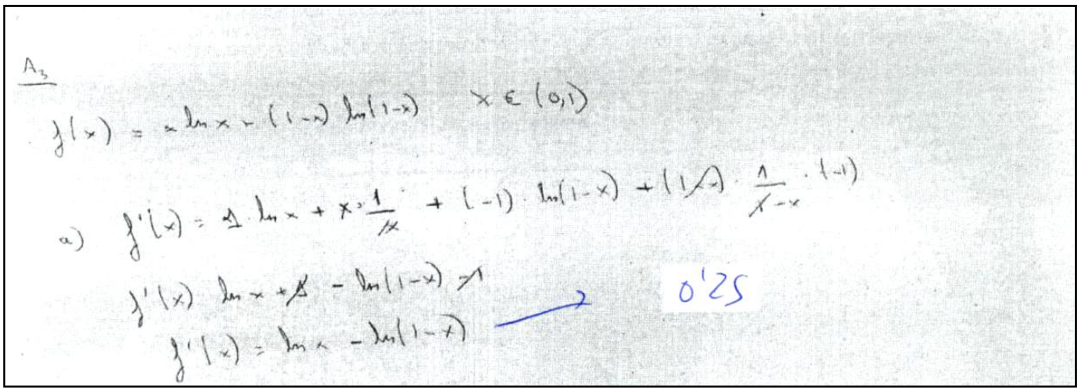
a) $f'(x) = 1\cdot\ln x + x\cdot\frac{1}{x} + (-1)\cdot\ln(1-x) + (1-x)\cdot\frac{1}{1-x}\cdot(-1)$

0'5

$f'(x)\ \ln x + 1 - \ln(1-x) - 1$

$f'(x) = \ln x - \ln(1-x)$

CORRECTOR 3


A3

$f(x) = x \ln x + (1-x) \ln(1-x) \quad x \in (0,1)$

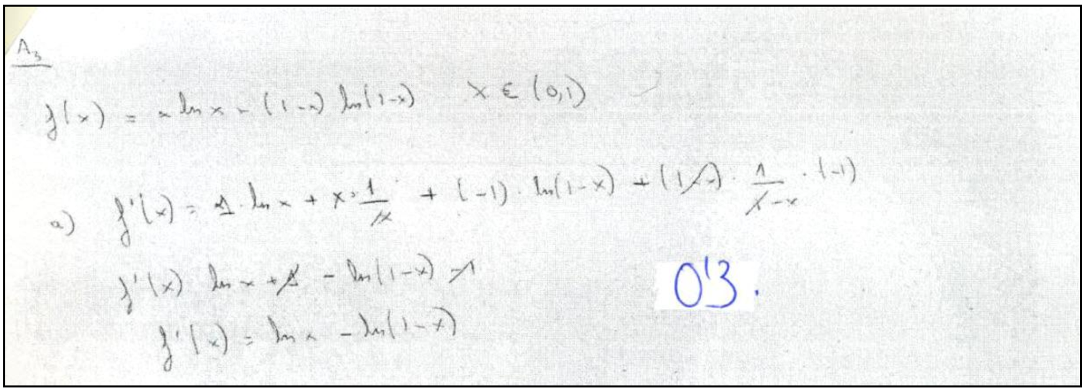
a) $f'(x) = 1 \cdot \ln x + x \cdot \frac{1}{\cancel{x}} + (-1) \cdot \ln(1-x) + (\cancel{1-x}) \cdot \frac{1}{\cancel{1-x}} \cdot (-1)$

$f'(x) \ \ln x + \cancel{1} - \ln(1-x) \ \cancel{-1}$

$f'(x) = \ln x - \ln(1-x)$

0'25

CORRECTOR 4


A3

$f(x) = x \ln x + (1-x) \ln(1-x) \quad x \in (0,1)$

a) $f'(x) = 1 \cdot \ln x + x \cdot \frac{1}{\cancel{x}} + (-1) \cdot \ln(1-x) + (\cancel{1-x}) \cdot \frac{1}{\cancel{1-x}} \cdot (-1)$

$f'(x) \ \ln x + \cancel{1} - \ln(1-x) \ \cancel{-1}$

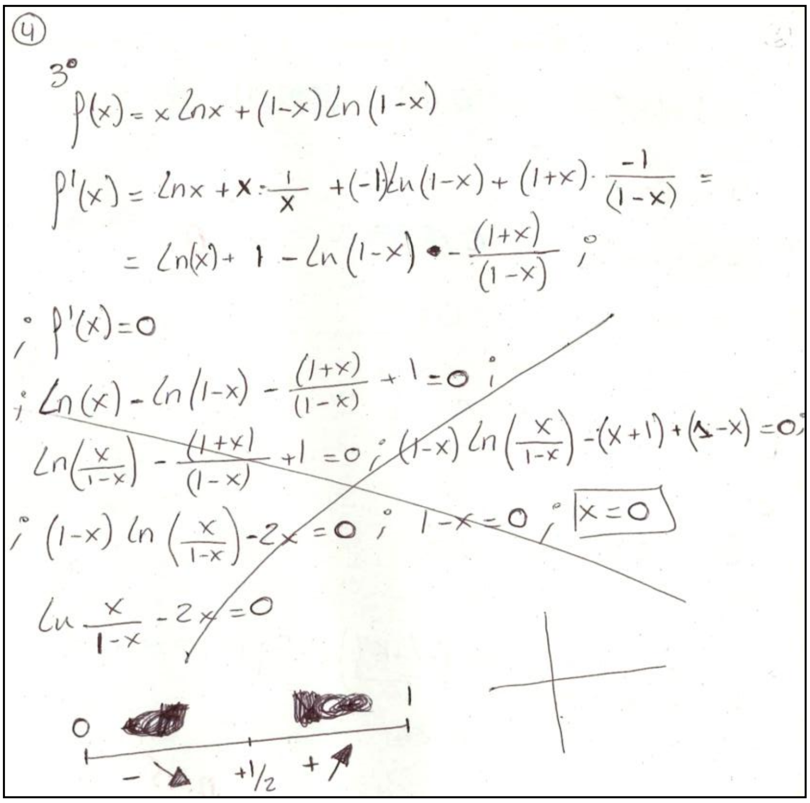
$f'(x) = \ln x - \ln(1-x)$

0'3.

<u>Apartado A3 b)</u>

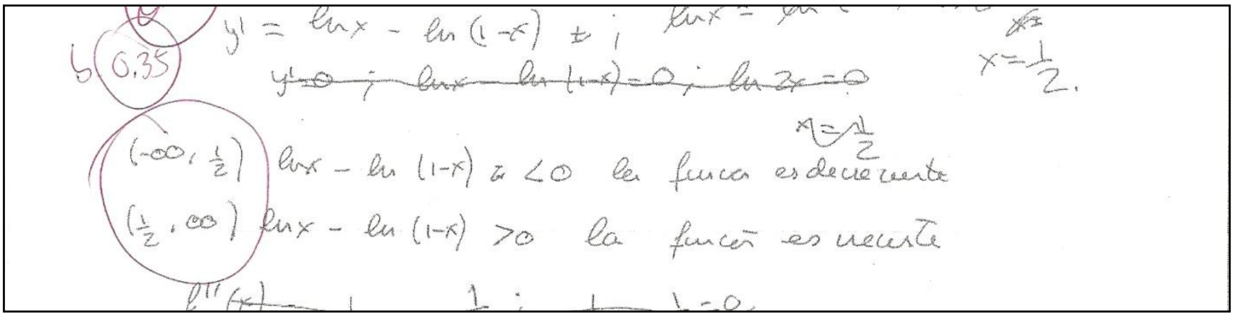
En este ejercicio los alumnos deben ser capaces de dar respuesta a dos tareas, calcular los intervalos de crecimiento y decrecimiento y calcular los puntos de inflexión (categoría C2). Algunas de las diferencias entre los correctores son consecuencia de que los alumnos no muestran el resultado de sustituir el punto en f'(x), hecho que genera dudas a los correctores acerca del cómo obtienen esa conclusión.

Otras veces, el alumno se limita a poner un dibujo pero sin justificar donde se ha apoyado para hacerlo:

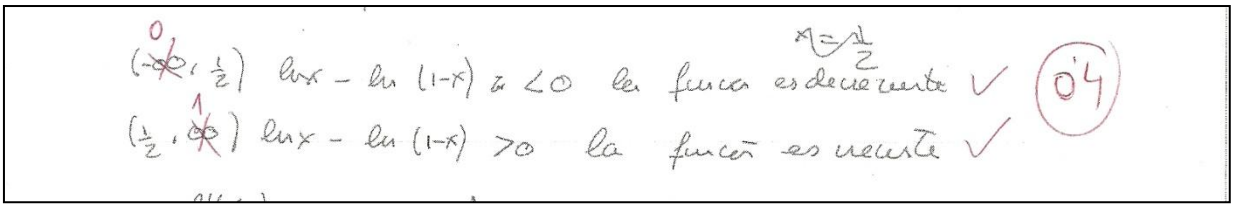

4
3°
$f(x) = x\ln x + (1-x)\ln(1-x)$
$f'(x) = \ln x + x \cdot \frac{1}{x} + (-1)\ln(1-x) + (1+x)\cdot\frac{-1}{(1-x)} =$
$= \ln(x) + 1 - \ln(1-x) - \frac{(1+x)}{(1-x)}$;
; $f'(x)=0$
; $\ln(x) - \ln(1-x) - \frac{(1+x)}{(1-x)} + 1 = 0$;
$\ln\left(\frac{x}{1-x}\right) - \frac{(1+x)}{(1-x)} + 1 = 0$; $(1-x)\ln\left(\frac{x}{1-x}\right) - (x+1) + (1-x) = 0$;
; $(1-x)\ln\left(\frac{x}{1-x}\right) - 2x = 0$; $1-x=0$; $\boxed{x=0}$
$\ln\frac{x}{1-x} - 2x = 0$
0 1
$-$ $+\frac{1}{2}$ $+$

Los correctores también han calificado de forma distinta que los alumnos escriban los intervalos fuera del dominio en el que está definida la función. f(x) está definida en el intervalo (0,1) y algunos alumnos escriben los intervalos (-∞,½) y (½,∞) en vez de (0,½) y (½,1).

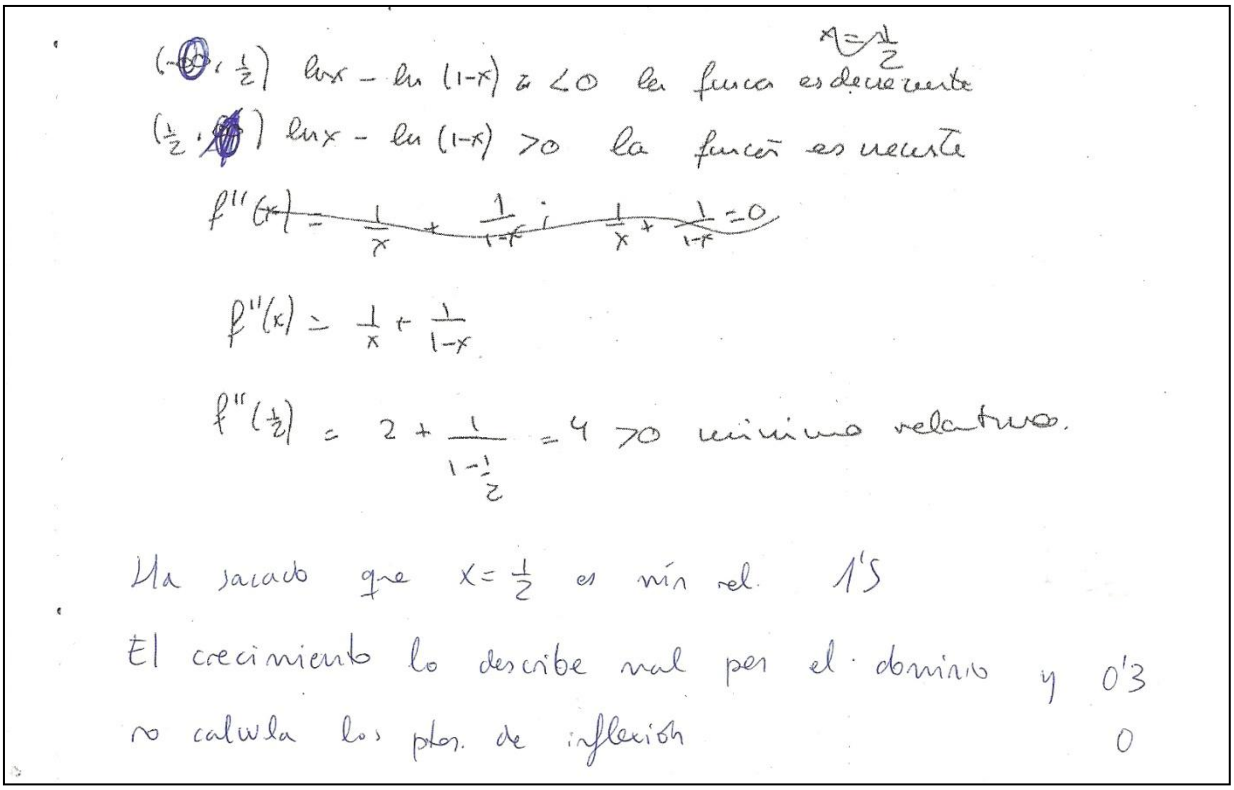
CORRECTOR 1


0,35
$y' = \ln x - \ln(1-x)$
~~$y'=0$; $\ln x - \ln(1-x) = 0$; $\ln 2x = 0$~~
$x = \frac{1}{2}$.
$x = \frac{1}{2}$
$(-\infty, \frac{1}{2})$ $\ln x - \ln(1-x) < 0$ la funcion es decreciente
$(\frac{1}{2}, \infty)$ $\ln x - \ln(1-x) > 0$ la funcion es creciente

CORRECTOR 2


$x = \frac{1}{2}$
0
$(\cancel{-\infty}, \frac{1}{2})$ $\ln x - \ln(1-x) < 0$ la funcion es decreciente ✓ 0'4
1
$(\frac{1}{2}, \cancel{\infty})$ $\ln x - \ln(1-x) > 0$ la funcion es creciente ✓

CORRECTOR 3


$x=\frac{1}{2}$

$(-\infty, \frac{1}{2})$ $\ln x - \ln(1-x) < 0$ la funcion es decreciente

$(\frac{1}{2}, \infty)$ $\ln x - \ln(1-x) > 0$ la funcion es creciente

$\cancel{f''(x) = \frac{1}{x} + \frac{1}{1-x} ; \frac{1}{x} + \frac{1}{1-x} = 0}$

$f''(x) = \frac{1}{x} + \frac{1}{1-x}$

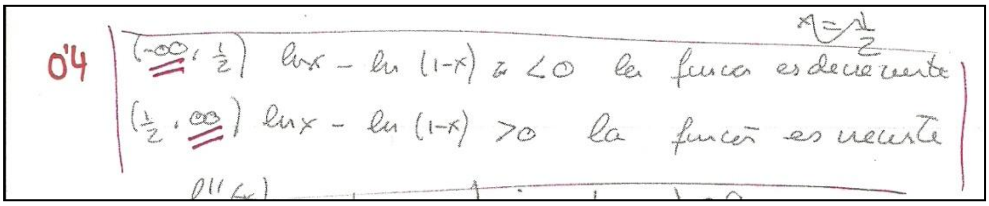
$f''(\frac{1}{2}) = 2 + \frac{1}{1-\frac{1}{2}} = 4 > 0$ minimo relativo.

Ha sacado que $x=\frac{1}{2}$ es mín rel. 1'5

El crecimiento lo describe mal por el dominio y 0'3

no calcula los ptos. de inflexión 0

CORRECTOR 4


0'4

$x=\frac{1}{2}$

$(-\infty, \frac{1}{2})$ $\ln x - \ln(1-x) < 0$ la funcion es decreciente

$(\frac{1}{2}, \infty)$ $\ln x - \ln(1-x) > 0$ la funcion es creciente

Apartado A4 a)

En este apartado se pide al alumno que defina cuál es el plano determinado por tres puntos(categoría C2). Los correctores tienen en cuenta cuando el alumno sabe que esto se consigue con un punto y dos vectores linealmente independientes determinados por estos tres puntos. La solución a este ejercicio es:

$$\pi \equiv x + y + z - 1 = 0$$

La mayor discrepancia en la puntuación se origina cuando el alumno resuelve bien el ejercicio pero no escribe la igualdad a cero. Nuestros datos muestran que no sólo los correctores dan un peso distinto sino que, además, dos de los correctores califica de distinta manera en distintos ejercicios este mismo error:

EXAMEN 1

CORRECTOR 1

4

a) $A(1,0,0)$ $B(0,1,0)$ $C(0,0,1)$

0,75

$\vec{AB} = (-1,1,0)$

$\vec{AC} = (-1,0,1)$

$$\begin{vmatrix} x-1 & y-0 & z-0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{vmatrix} = x-1-(-1(z-0)-1(y-0)) = x-1+z+y$$

$\pi: x+y+z-1 = ?$

b. 0

CORRECTOR 2

4

a) $A(1,0,0)$ $B(0,1,0)$ $C(0,0,1)$

1

$\vec{AB} = (-1,1,0)$

$\vec{AC} = (-1,0,1)$

$$\begin{vmatrix} x-1 & y-0 & z-0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{vmatrix} = x-1-(-1(z-0)-1(y-0)) = x-1+z+y$$

$\pi: x+y+z-1$

b) 0

CORRECTOR 3

4

a) $A(1,0,0)$ $B(0,1,0)$ $C(0,0,1)$

$\vec{AB} = (-1,1,0)$

$\vec{AC} = (-1,0,1)$

$$\begin{vmatrix} x-1 & y-0 & z-0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{vmatrix} = x-1-(-1(z-0)-1(y-0)) = x-1+z+y$$

$\pi: x+y+z-1=0 \rightarrow$ 0'9

0'9

CORRECTOR 4

4

a) $A(1,0,0)$ $B(0,1,0)$ $C(0,0,1)$

0'9

$\vec{AB} = (-1,1,0)$

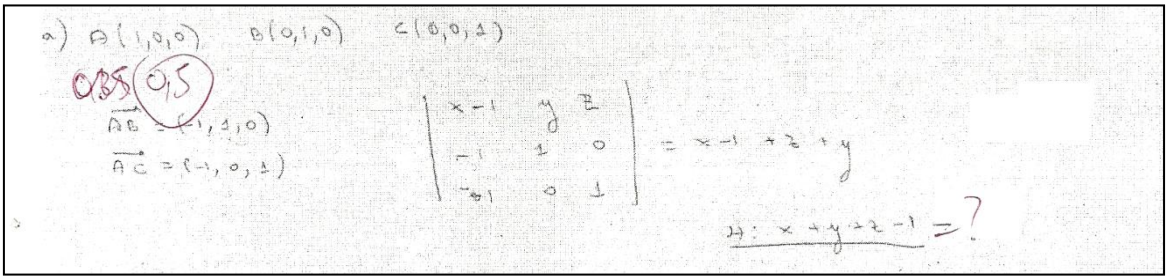
$\vec{AC} = (-1,0,1)$

$$\begin{vmatrix} x-1 & y-0 & z-0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{vmatrix} = x-1-(-1(z-0)-1(y-0)) = x-1+z+y$$

$\pi: x+y+z-1$

EXAMEN 2

CORRECTOR 1



a) $A(1,0,0)$ $B(0,1,0)$ $C(0,0,1)$

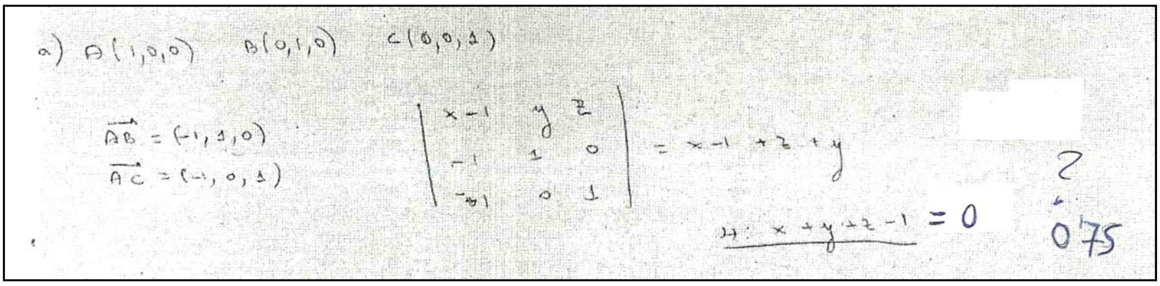
0,25 0,5

$\vec{AB}=(-1,1,0)$
$\vec{AC}=(-1,0,1)$

$\begin{vmatrix} x-1 & y & z \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{vmatrix} = x-1+z+y$

$\pi: x+y+z-1 = ?$

CORRECTOR 3



a) $A(1,0,0)$ $B(0,1,0)$ $C(0,0,1)$

$\vec{AB}=(-1,1,0)$
$\vec{AC}=(-1,0,1)$

$\begin{vmatrix} x-1 & y & z \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{vmatrix} = x-1+z+y$

$\pi: x+y+z-1 = 0$

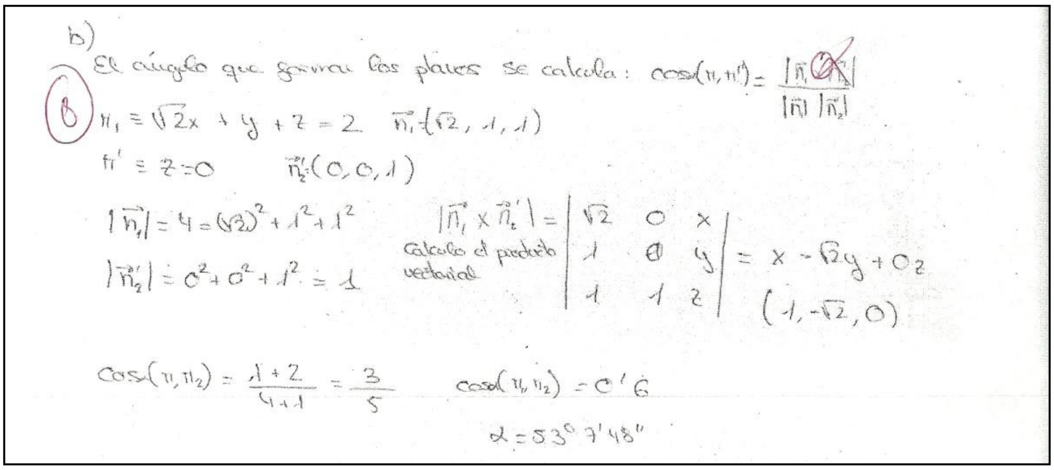
2

0'75

Apartado A4 b)

El objetivo principal de esta pregunta es verificar si el alumno domina la técnica para el cálculo del ángulo que determinan dos planos. Las diferencias entre las calificaciones se deben a que unos correctores tienen en cuenta cuando el alumno sabe que para calcular este ángulo necesita los vectores normales a dichos planos y los calculen bien. En esta ocasión, el alumno no es capaz de resolver el problema con éxito pero al menos sabe que para resolver esta pregunta necesita conocer cuáles son los vectores normales, dato que aparece en las tareas auxiliares.

CORRECTOR 1



b)

El ángulo que forman los planos se calcula: $\cos(\pi,\pi') = \dfrac{|\vec{n_1} \cancel{\times} \vec{n_2}|}{|\vec{n_1}|\,|\vec{n_2}|}$

0

$\pi_1 \equiv \sqrt{2}x + y + z = 2$ $\vec{n_1}(\sqrt{2},1,1)$

$\pi' \equiv z = 0$ $\vec{n_2'}(0,0,1)$

$|\vec{n_1}| = 4 = (\sqrt{2})^2+1^2+1^2$

$|\vec{n_2'}| = 0^2+0^2+1^2 = 1$

$|\vec{n_1} \times \vec{n_2'}| = \begin{vmatrix} \sqrt{2} & 0 & x \\ 1 & 0 & y \\ 1 & 1 & z \end{vmatrix} = x - \sqrt{2}y + 0z$

Calculo el producto vectorial

$(1,-\sqrt{2},0)$

$\cos(\pi,\pi_2) = \dfrac{1+2}{4+1} = \dfrac{3}{5}$ $\cos(\pi,\pi_2) = 0'6$

$\alpha = 53^\circ 7' 48''$

CORRECTOR 2

b)

El ángulo que forman los planos se calcula: $\cos(\pi,\pi') = \dfrac{|\vec{n}_1 \times \vec{n}_2|}{|\vec{n}|\,|\vec{n}_2|}$

$\pi_1 \equiv \sqrt{2}x + y + z = 2 \quad \vec{n}_1 (\sqrt{2}, 1, 1)$

$\pi' \equiv z = 0 \quad \vec{n}_2' (0, 0, 1)$

$|\vec{n}_1| = 4 = (\sqrt{2})^2 + 1^2 + 1^2$

$|\vec{n}_1 \times \vec{n}_2'| = \begin{vmatrix} \sqrt{2} & 0 & x \\ 1 & 0 & y \\ 1 & 1 & z \end{vmatrix} = x - \sqrt{2}y + 0z$

Calculo el producto vectorial

$|\vec{n}_2'| = 0^2 + 0^2 + 1^2 = 1$

$(1, -\sqrt{2}, 0)$

0'5

$\cos(\pi,\pi_2) = \dfrac{1+2}{4+1} = \dfrac{3}{5} \quad \cos(\pi,\pi_2) = 0'6$

$\alpha = 5\cancel{3}^\circ\, 7'48''$

CORRECTOR 3

b)

El ángulo que forman los planos se calcula: $\cos(\pi,\pi') = \dfrac{|\vec{n}_1 \times \vec{n}_2|}{|\vec{n}|\,|\vec{n}_2|}$

$\pi_1 \equiv \sqrt{2}x + y + z = 2 \quad \vec{n}_1 (\sqrt{2}, 1, 1)$

$\pi' \equiv z = 0 \quad \vec{n}_2' (0, 0, 1)$

$|\vec{n}_1| = 4 = (\sqrt{2})^2 + 1^2 + 1^2$

$|\vec{n}_1 \times \vec{n}_2'| = \begin{vmatrix} \sqrt{2} & 0 & x \\ 1 & 0 & y \\ 1 & 1 & z \end{vmatrix} = x - \sqrt{2}y + 0z$

Calculo el producto vectorial

$|\vec{n}_2'| = 0^2 + 0^2 + 1^2 = 1$

$(1, -\sqrt{2}, 0)$

$\cos(\pi,\pi_2) = \dfrac{1+2}{4+1} = \dfrac{3}{5} \quad \cos(\pi,\pi_2) = 0'6$

$\alpha = 53^\circ\, 7'48''$

0

CORRECTOR 4

El ángulo que forman los planos se calcula: $\cos(\pi,\pi') = \dfrac{|\vec{n}_1 \times \vec{n}_2|}{|\vec{n}|\,|\vec{n}_2|}$

$\pi_1 \equiv \sqrt{2}x + y + z = 2 \quad \vec{n}_1 (\sqrt{2}, 1, 1)$

$\pi' \equiv z = 0 \quad \vec{n}_2' (0, 0, 1)$

$|\vec{n}_1| = 4 = (\sqrt{2})^2 + 1^2 + 1^2$

$|\vec{n}_1 \times \vec{n}_2'| = \begin{vmatrix} \sqrt{2} & 0 & x \\ 1 & 0 & y \\ 1 & 1 & z \end{vmatrix} = x - \sqrt{2}y + 0z$

Calculo el producto vectorial

$|\vec{n}_2'| = 0^2 + 0^2 + 1^2 = 1$

$(1, -\sqrt{2}, 0)$

$\cos(\pi,\pi_2) = \dfrac{1+2}{4+1} = \dfrac{3}{5} \quad \cos(\pi,\pi_2) = 0'6$

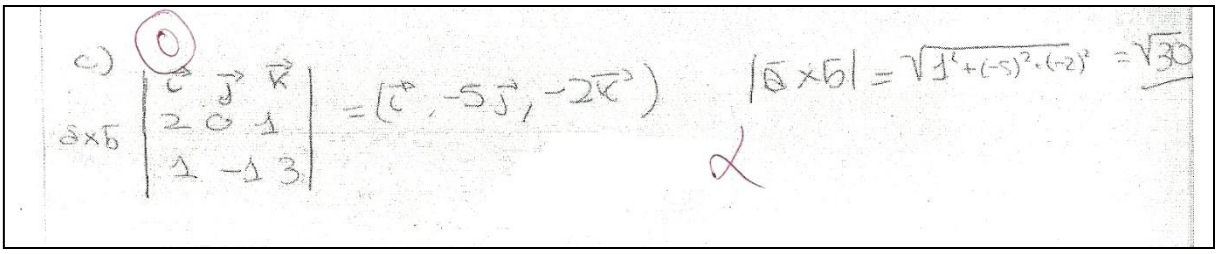
$\alpha = 53^\circ\, 7'48''$

0'5.

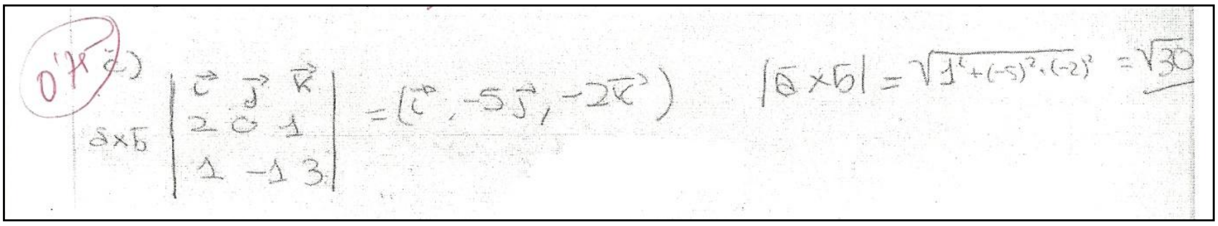
Apartado A4 c)

En este apartado el alumno solo debe realizar un cálculo, el de un producto vectorial (categoría C1). Sin embargo, hay más diferencias en este apartado que en los anteriores (a y b), debido a que el alumno utiliza una mala notación o se equivoca en la cuenta y los correctores aplican de distinta manera los criterios de corrección:

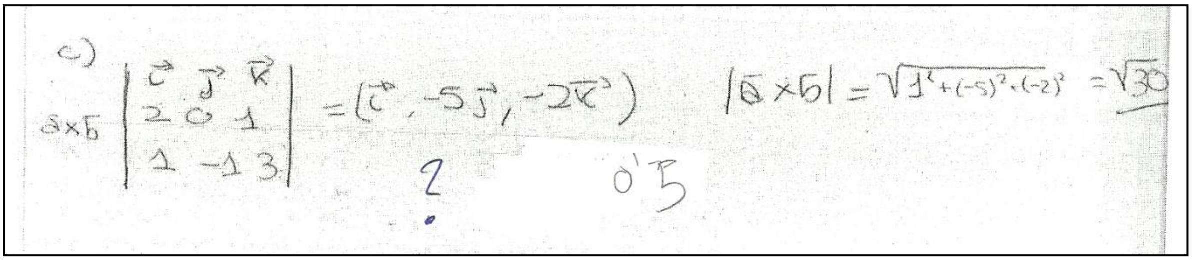
CORRECTOR 1



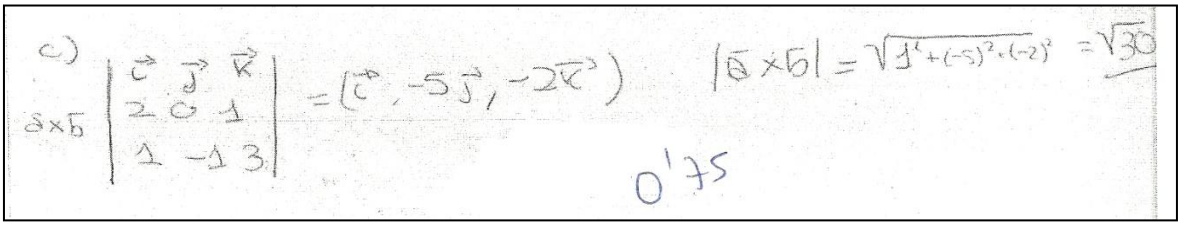
CORRECTOR 2



CORRECTOR 3



CORRECTOR 4



DISCUSIÓN, CONCLUSIONES Y PROSPECTIVA

En nuestro estudio concretamos el modelo de Gairín, Muñoz y Oller (2012) para la corrección y calificación de dos preguntas de la prueba de selectividad de Zaragoza en 2010. A partir del modelo teórico que proponen, generamos una guía de corrección que entregamos a 4 profesores para que califiquen dichas preguntas. A partir del análisis de las respuestas de los alumnos que generan variabilidad en las calificaciones recogidas identificamos los siguientes factores a considerar para la mejora del modelo GMO:

1. El alumno interrumpe la respuesta de una pregunta de clase C2 en cualquier punto del proceso de resolución.
2. Mala justificación de una solución o una mala expresión final de ésta.
3. Resultados de cálculos no justificados.

4. Errores conceptuales que no afectan a la resolución pero que muestran un determinado desconocimiento.

El análisis de las respuestas de los alumnos que los correctores calificaban de forma discrepante nos ha permitido observar factores no incluidos en el modelo GMO, como la valoración de una pregunta a medio contestar o los cálculos no justificados, que permiten todavía una variabilidad no deseada en las correcciones. Consideramos que el modelo ha demostrado su pertinencia y potencialidad y que la inclusión de soluciones que incorporen las problemáticas detectadas puede hacer ganar en robustez al modelo.

Gairín, Muñoz y Oller (2012) detectan casos en que distintos correctores asignan puntuaciones medias con diferencias que representan un 20% del total. En nuestro estudio comprobamos que una amplia mayoría de las calificaciones se sitúan por debajo del 15% del total. De esta forma, consideramos que la corrección a partir de la guía que proponemos como concreción del modelo muestra indicios de fiabilidad mayores que la corrección a partir de los criterios utilizados en la prueba real. Sin embargo, somos conscientes que no podemos aventurarnos todavía a realizar una afirmación contundente ya que el número de correctores participantes en la prueba no es suficiente para otorgar a los datos fiabilidad estadística. Por ello, nuestro estudio debería ampliarse abriendo la participación a un mayor número de correctores, centrándonos en aquellas respuestas de los alumnos que generan mayor variabilidad. En este punto, siguiendo a Muñiz y Fonseca-Pedrero (2008) deberíamos conseguir un instrumento en el cual los criterios de corrección no fuesen ni demasiado escuetos ni excesivamente precisos.

Dado un ejercicio de las PAU, consideramos que la forma en la que se explicitan los criterios requiere de un conocimiento profundo no sólo de los contenidos tratados, sino también de los errores que pueden cometer los alumnos. Asimismo, esta tarea no debe relegarse a correctores no expertos, dado que la concreción del modelo requiere que quien elabore la guía para la corrección separe claramente las diferentes tareas que debe realizar el alumno. Por ello, la elaboración de la guía requiere un nivel de análisis que no debe exigirse a una persona que no haya corregido previamente exámenes de matemáticas del nivel tratado. Por otra parte, resulta claro que la corrección de cualquier prueba tiene una componente de variación de error atribuible al evaluador y no al estudiante, debido a que en la corrección existen variables difíciles de medir.

Referencias

Álvarez Méndez, J. M. (2000). *Evaluar para aprender, examinar para excluir*. Madrid: Morata.

Casanova, M. A. (1998). Evaluación, concepto, tipología y objetivos. En M.A. Casanova. *La evaluación educativa*. México: SEP-Muralla (Cap. 3, pp. 67-102). Disponible en http://www.reformasecundaria.sep.gob.mx/espanol/pdf/evaluacion/casanova/casanova3.pdf

Cuxart, A., Martí, M., y Ferrer, F. (1997). Algunos factores que inciden en el rendimiento y la evaluación en los alumnos de las Pruebas de Aptitud de Acceso a la Universidad. *Revista de Educación, 314*, 63-88

Escudero, T., y Bueno, C. (1994). Investigaciones y Experiencias: Examen de Selectividad. El estudio de un tribunal paralelo. *Revista de Educación, 304*, 281-298

Gairín, J.M., Muñoz, J.M., y Oller, A.M. (2012). Propuesta de un modelo para la calificación de exámenes de matemáticas. En A. Estepa, A. Contreras, J. Deulofeu, M. C. Penalva, F. J. García y L. Ordóñez (Eds.), *Investigación en Educación Matemática XVI* (pp. 261-274).

Grau, R. Cuxart, A., y Martí-Recober, M. (2002). La calidad en el proceso de corrección de pruebas de acceso a la universidad: variabilidad y factores. *Revista de Investigación Educativa*, *20*(1), 209-223

Muñiz, J., y Fonseca-Pedrero, E. (2009). Construcción de instrumentos de medida para la evaluación universitaria. *Revista de Investigación en Educación*, *5*, 13-25.