

Las correcciones en Matemáticas en las Pruebas de Acceso a la Universidad

Mathematics corrections in University Entrance Examination

ROSA NORTES MARTÍNEZ-ARTERO¹

mrosa.nortes@um.es

ANDRÉS NORTES CHECA

anortes@um.es

FRANCISCO LOZANO PATO

flozano@um.es

Universidad de Murcia, España

Resumen:

En las correcciones de Pruebas de Acceso a la Universidad en Matemáticas Aplicadas a las Ciencias Sociales hay un coordinador y varios correctores, profesores de secundaria y universidad, especialistas en Matemáticas. Para comprobar si hay diferencias en la corrección de las pruebas se eligen dos evaluadores que corrigen grupos distintos de exámenes y en las reclamaciones de los alumnos se toma un grupo de los exámenes reclamados, corregidos por varios evaluadores, y vueltos a corregir por un evaluador distinto de los anteriores. Ante estudios que indican que hay diferencias en los resultados, se constata que en el examen de junio de 2014 en la universidad de Murcia no existen diferencias significativas en los resultados de los dos evaluadores que corrigen grupos distintos, ni tampoco cuando distintos evaluadores corrigen el mismo grupo y es de nuevo corregido por otro profesor. En el análisis de las cuestiones de las dos op-

Abstract

For the corrections of University entrance examinations, in this case in the area of Mathematics Applied to Social Science, there is one coordinator and several markers, Secondary and University teachers, who are Mathematics specialists. In order to check whether there are differences in the exams corrections, two evaluators are chosen to correct different groups of examinations and, when students appeal for revision, one set of these exams, which have already been corrected by several evaluators, is corrected again by another evaluator different from the previous ones. In view of studies that show that there are differences in the results, we have verified that in the University of Murcia's June 2014 examinations there are no significant differences in the results of the two evaluators who correct different sets, nor even when other evaluators correct the same set and this is again corrected by a new teacher. In the analysis of the questions,

1 Dirección para correspondencia (correspondence address):

Rosa Nortes Martínez-Artero. Dpto. de Didáctica de las Ciencias Matemáticas y Sociales. Facultad de Educación. Universidad de Murcia. Campus Universitario de Espinardo. 30100 Murcia (España).

ciones que puede elegir el alumno, las de probabilidad y de contraste de hipótesis son las de mayor dificultad en su resolución. En los cuatro grupos corregidos y en el total de la prueba, los alumnos suspenden.

Palabras clave:

Matemáticas, prueba de acceso a la universidad, alumnos, errores.

from the two options that the students can choose, those of probability and contrast of hypothesis have proved to be the most difficult to solve. The students failed in the four groups corrected and in the total examination.

Key words:

Mathematics; university entrance examination; students; errors.

Résumé:

Pendant les corrections des examens d'entrée à l'Université de Mathématiques Appliquées aux Sciences Sociales il y a un coordinateur et plusieurs correcteurs, des professeurs dans le secondaire et à l'université, des spécialistes en Mathématiques. Pour vérifier s'il y a des différences dans la correction des examens, on choisit deux évaluateurs qui corrigent des groupes différents d'examens et pour la correction des réclamations des élèves, on prend un groupe d'examens réclamés, corrigés par différents évaluateurs, et corrigés à nouveau par un évaluateur différent aux autres. Parmi les études qui indiquent qu'il y a des différences dans les résultats, on constate que dans les examens de juin 2014 à l'université de Murcia, il n'existe pas de différences significatives dans les résultats des deux évaluateurs qui corrigent des groupes différents, ni non plus quand les différents évaluateurs corrigent le même groupe et qui est à nouveau corrigé par un autre professeur. Dans les analyses des questions de deux options que peut choisir l'élève, celles de probabilité et de contraste d'hypothèses sont celles où il y a de plus grandes difficultés dans la résolution. Dans les quatre groupes corrigés et dans le total de l'examen, les élèves échouent.

Mots clés:

Mathématiques; examen d'entrée à l'université; élèves; erreurs.

Fecha de recepción: 7-11-2014

Fecha de aceptación: 22-12-2014

Marco teórico

Los resultados en Matemáticas de las PAU (Pruebas de Acceso a la Universidad) han sido estudiados por numerosos autores. Unos artículos están dedicados a la clasificación de los ejercicios (Ruiz de Gauna, Sarasua y García, 2011), otros dedicados a algunos resultados en las cuestiones de estadística (García y García, 2005; Ramos, Espinel y Ramos, 2009), también a las habilidades matemáticas evaluables (Boal, Bueno, Levis y Sein-Echaluce, 2008), a la calidad en el proceso de evaluación, propuesta y validación de un modelo (Grau, Cuxart y Marti-Recober, 2002; Gairin, Muñoz y Oller, 2012; Mengual, Gorgorió y Albarracín, 2013), o también a destacar algunos errores (Nortes y Nortes, 2010; Mallart, 2014), por citar algunos representativos.

Lorenzo, Argos, Hernández y Vera (2014) se hacen una serie de preguntas para ver si la prueba de acceso cumple con las funciones inicialmente establecidas: ¿Sirve para homologar y certificar los conocimientos adquiridos en Secundaria? ¿Sirve para evaluar la madurez? ¿Sirve para la evaluación externa del sistema? ¿Sirve como valoración de la capacidad del alumnado para seguir con éxito enseñanzas universitarias? ¿Sirve para ubicar adecuadamente al alumnado, ordenarlo y darle prioridad en la elección?, llegando a la conclusión de que solo se cumple de forma efectiva esta última, por lo que proponen hacer solo una prueba específica para ingresar en estudios con límite de plazas.

Sin embargo, esta solución puede perjudicar a aquellos alumnos que inicialmente pensaban realizar unos estudios y al ver su imposibilidad cambian de titulación. Así, por ejemplo, hay alumnos que en su día optaron por estudiar la titulación del Grado de Matemáticas y al no aprobar las asignaturas al cabo de uno o dos años abandonan y realizan el Grado de Maestro de Primaria.

Huidobro, Méndez y Serrano (2010) consideran que las PAU tienen como finalidad valorar las calificaciones de los alumnos obtenidas en el bachillerato, sus conocimientos, competencias y su madurez académica, al tiempo que establecer un orden entre los alumnos que desean matricularse en unos estudios con límite de acceso y por ello Grau, Cuxart y Martí-Recober (2002) ven la necesidad de que la corrección de las pruebas sea fiable y que se lleven a cabo pruebas piloto que informen del nivel de dificultad y sobre la variabilidad que origina la corrección de los exámenes.

En las PAU, los correctores califican siguiendo los criterios de evaluación marcados por el coordinador y en el protocolo de la prueba ya se especifica la puntuación que corresponde a cada problema. Posteriormente en la sesión de reparto de exámenes se indican los criterios específicos, incluidas las penalizaciones por errores cometidos. A este respecto Grau, Cuxart y Martí-Recober (2002) tras un estudio en donde se lleva a cabo la corrección triple de cada examen, se asignan al azar los exámenes a los correctores, se simultanean en la corrección exámenes originales y fotocopias, con la intervención de un número suficiente de exámenes, con la intervención de un número suficiente de correctores y con la participación de coordinadores como correctores de fotocopias, llegan a las conclusiones, entre otras, de que un aumento de la concordancia en la corrección de Matemáticas ayuda a reducir la discrepancia

entre correcciones y que los coordinadores deberían esforzarse en saber comunicar a los correctores que es lo que quieren medir o valorar en el examen y en cada pregunta en concreto.

Mallart (2014) analiza la resolución de problemas de la prueba de Matemáticas II de PAU en Cataluña y en los criterios generales de evaluación se insiste en que el alumno ha de explicar el porqué de sus respuestas, llegando a valorarse con cero, cuestiones con resultado correcto pero que el proceso no ha sido explicado.

Gairin, Muñoz y Oller (2012) proponen un modelo para la calificación de exámenes de Matemáticas analizando el caso especial de las PAU junio 2010 de la asignatura Matemáticas Aplicadas a las Ciencias Sociales (MACS) en Zaragoza, en donde los coordinadores de cada asignatura elaboran unos criterios específicos de corrección, aunque se constata la existencia de diferencias significativas entre las calificaciones de distintos correctores cuando califican un mismo examen (corrector A = 5.74 y corrector B = 3.86) y lo achacan a que existe una fuerte componente de subjetividad en la valoración de las respuestas de los alumnos y por tanto se producen diferencias en las calificaciones otorgadas por distintos correctores.

Gairin, Muñoz y Oller (2013), indican que los criterios de evaluación dados por los coordinadores son de "puntuación por acierto", no dándose información detallada de las posibles soluciones o errores que puedan cometer los alumnos y mucho menos se cuantifican o ponderan. En su estudio analizan las actuaciones de 6 correctores que calificaron un total de 409 exámenes, 205 de Matemáticas Aplicadas a las Ciencias Sociales (3 correctores) y 204 de Matemáticas II, (3 correctores) viendo que las respuestas de los alumnos contenían una gran cantidad de errores y con una calificación media de 3.75 y 4.19, respectivamente.

Mallart (2014) ante una muestra de 104 exámenes de Matemáticas II en la convocatoria de junio de 2012 en Cataluña, y tras su corrección, llega a la conclusión de que los alumnos asimilan correctamente los procedimientos mecánicos de resolución aunque manifiestan ausencia de rigor y precisión en los cálculos efectuados.

En Nortes y Nortes (2010) tras analizar una muestra de 134 exámenes de MACS correspondientes a la convocatoria de septiembre PAU 2009, se constata que las puntuaciones más altas se obtienen cuando hay que aplicar directamente una fórmula, mientras que en aquellas cuestiones en donde el enunciado es largo y hay que interpretarlo co-

rrectamente, es donde las puntuaciones son más bajas y más errores se cometen.

Para corroborar los resultados de los estudios previos, partiendo de la Prueba de Matemáticas Aplicadas a las Ciencias Sociales (MACS) en la convocatoria de junio de 2014 en la Universidad de Murcia se hace un doble planteamiento para un estudio pormenorizado de corrección de las cuestiones que componen cada opción.

Objetivo

El objetivo de este estudio es doble:

1. Ver si existen diferencias significativas entre los resultados de dos correctores de Matemáticas Aplicadas a las Ciencias Sociales (MACS) en dos muestras distintas en las dos opciones de la prueba.
2. Ver si existen diferencias significativas en una muestra que es evaluada doblemente por distintos profesores inicialmente y por un evaluador en uno de los grupos de revisión en las dos opciones de la prueba.

Método

Diseño

El estudio realizado lo componen dos fases. En la primera se analizan los resultados otorgados por dos profesores universitarios correctores de Matemáticas Aplicadas a las Ciencias Sociales (MACS) en la convocatoria de junio de 2014 en la Universidad de Murcia. Cada profesor corrigió un número determinado de exámenes recibido del coordinador.

En la segunda fase, se han tomado los exámenes correspondientes a reclamaciones efectuadas por los alumnos, que ha llevado a una segunda calificación de los mismos, por lo que hay una primera puntuación otorgada por distintos evaluadores y una segunda calificación otorgada por el profesor que revisó y volvió a evaluar dicho examen.

Muestra

De los 1745 exámenes de MACS correspondientes a la prueba de junio de 2014 se han encargado de su corrección 12 evaluadores que han corregido un número próximo a 145 exámenes cada uno, evaluadores que son profesores de secundaria y de universidad especialistas en Matemáticas. La muestra la constituyen 291 exámenes de primera evaluación correspondientes a los dos evaluadores y 141 exámenes presentados a la revisión con las dos calificaciones, la inicial efectuada por distintos profesores y la posterior efectuada por un mismo profesor, distinto de los anteriores. Su distribución es la que aparece en la tabla 1.

Tabla 1. Distribución muestra.

	Fase I	Fase 2
Profesor 1	147 exámenes	
Profesor 2	144 exámenes	141 exámenes
Varios prof.		141 exámenes

Procedimiento

En una sesión el coordinador de la asignatura MACS reunió a los profesores evaluadores y les entregó el protocolo del examen corregido, la distribución de la puntuación en cada uno de los problemas y la penalización por errores cometidos.

Cada profesor evaluó su “paquete” e introdujo las calificaciones que posteriormente fueron publicadas junto al resto de asignaturas y la calificación total de la Prueba.

Los alumnos que no estaban de acuerdo con su puntuación presentaron reclamación en tiempo y forma, llamando al coordinador para corregirlos a algunos de los 12 evaluadores iniciales. Con los exámenes de los reclamantes se hicieron grupos de aproximadamente 140 exámenes y uno de esos nuevos “paquetes” fue corregido por uno de los dos profesores iniciales (profesor 2) que llevó a cabo una segunda evaluación.

Instrumento

La prueba de Matemáticas corregida es la siguiente:

OPCIÓN A

CUESTIÓN A1. Discutir el siguiente sistema por el método de Gauss, según los valores del parámetro a , siendo a un número real distinto de 0. (2 puntos)

$$ax + y - 2az = 1$$

$$ax - y = 2$$

$$ax + y + (a-1)z = 3a-1$$

Resolverlo para $a = 1$. (0,5 puntos).

CUESTIÓN A2. El coste de fabricación de un modelo de teléfono móvil viene dado por la función $C(x) = x^2 + 10x + 325$, donde x representa el número de teléfonos móviles fabricados. Supongamos que se venden todos los teléfonos fabricados y que cada uno se vende por 80 euros.

- Determinar la función beneficio (definida como ingreso menos coste) que expresa el beneficio obtenido en función de x . (0,5 puntos).
 - ¿Cuántos teléfonos deben fabricarse para que el beneficio sea máximo? ¿A cuánto asciende dicho beneficio? (1,5 puntos).
 - ¿Para qué valores de x se tienen pérdidas (beneficios negativos)? (0,5 puntos).
-

CUESTIÓN A3. Hallar las siguientes integrales indefinidas:

a) $\int (x^5 - 2x + 3)dx$ (0,75 puntos) b) $\int (2e^x + 5)dx$ (0,75 puntos).

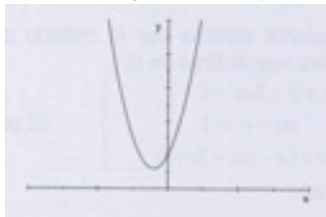
CUESTIÓN A4. Dados dos sucesos A y B de un mismo experimento aleatorio, calcule $P(A)$ y $P(B)$ sabiendo que son independientes y que $P(A^c)=0,6$ y $P(A \cup B)=0,7$. (2 puntos).

CUESTIÓN A5. El peso (en gramos) de los pollos que llegan a un matadero sigue una distribución normal con desviación típica de 315 g. Sabiendo que una muestra de 64 pollos ha dado un peso medio de 2750 g, hallar un intervalo de confianza para el peso medio con un nivel de confianza del 97 %. (1,5 puntos).

OPCIÓN B

CUESTIÓN B1. Una fábrica de tintas dispone de 1000 kg del color A, 800 kg del color B y 300 kg del color C, con los que fabrica dos tipos de tinta, una para la etiqueta de un refresco y otra para un cartel. Cada bote de tinta de la etiqueta necesita 10 kg del color A, 5 kg del color B y 5 kg del color C y el de tinta del cartel requiere 5 kg de A y 5 kg de B. Obtiene un beneficio de 30 euros por cada bote de tinta para etiquetas y de 20 euros por cada uno de tinta para carteles. Si vende todos los botes fabricados, ¿cuántos botes de cada tipo de tinta debe fabricar para maximizar su beneficio?, ¿cuál es el beneficio máximo? (3 puntos).

CUESTIÓN B2. La siguiente gráfica corresponde a la función $f(x)=x^2+4x+a$, siendo a un número real. Calcular a para que el área encerrada por la gráfica, el eje OX y las rectas $x=0$ y $x=3$ valga 57. (1,5 puntos).



CUESTIÓN B3. Hallar las derivadas de las siguientes funciones:

a) $f(x) = \frac{e^{2x}}{x+2}$ (0,75 puntos) b) $g(x) = \ln(x^2-5x)$ (0,75 puntos).

CUESTIÓN B4. En una población el 60 % de los individuos toma diariamente leche y el 40 % toma diariamente yogur. Además, el 30 % de los individuos toma leche y yogur diariamente.

¿Cuál es la probabilidad de que un individuo tome a diario leche pero no yogur? (0,5 puntos).

¿Cuál es la probabilidad de que tome a diario leche o yogur? (0,75 puntos).

Si un individuo toma diariamente leche, ¿qué probabilidad hay de que también tome a diario yogur? (0,75 puntos).

CUESTIÓN B5. El tiempo de espera para recibir un tratamiento médico es, en promedio, de 30 días. Después de tomar medidas para intentar reducirlo, para una muestra de 80 pacientes el tiempo medio de espera es de 27 días. Suponiendo que el tiempo de espera sigue una distribución normal con una desviación típica igual a 8, plantear un test para contrastar que las medidas no han mejorado la situación frente a que sí lo han hecho. ¿Cuál es la conclusión a la que se llega con un nivel de significación del 5 %? (2 puntos).

Los datos han sido tratados estadísticamente con el paquete Systat.

Resultados

Globales

En la tabla 2 se presentan las medias de cada evaluador y el resultado total de la prueba y en la tabla 3 los porcentajes de suspensos, aprobados, notables y sobresalientes por evaluador.

Tabla 2. Resultados globales.

GLOBALES					
	P1	P2	C1	C2	MACS
Casos	147	144	141	141	1745
Mínimo	0	0	0	0	0
Máximo	10	9.10	7.70	7.45	10
Media	3.95	4.07	3.29	3.31	4.16
DT	2.38	2.15	1.63	1.59	2.26
Intervalo	62.6 %	66.0 %	70.2 %	70.2 %	

- Los alumnos suspenden con el profesor 1 (P1) y con el profesor 2 (P2) con una diferencia en la media de 12 centésimas.
- En ambos profesores el intervalo ($\bar{x} - \sigma$, $\bar{x} + \sigma$) recoge el 62.6 % y el 66 % de las evaluaciones, respectivamente, siendo en una distribución normal de 68.2 %.
- La media de los alumnos que pidieron segunda corrección (C2) se encuentra muy próxima al resultado de la primera corrección (C1), la diferencia es 23 milésimas superior en la segunda y el intervalo ($\bar{x} - \sigma$, $\bar{x} + \sigma$) en ambas evaluaciones es del 70.2 %.

Tabla 3. Calificaciones muestras.

CALIFICACIONES				
	P1	P2	C1	C2
Suspenseo	70.7 %	66.7 %	87.2 %	85.8 %
Aprobado	14.3 %	20.1 %	10.6 %	12.1 %
Notable	12.2 %	12.5 %	2.1 %	2.1 %
Sobresal.	2.7 %	0.7 %	0 %	0 %

- Dos tercios o más de los alumnos suspenden en cada evaluación.
- El segundo profesor (P2) establece mas notas intermedias que el primero (P1).
- De los alumnos reclamantes (C1), el 87.2 % había suspendido.
- Los dos profesores iniciales (P1 y P2) no tienen diferencias en el porcentaje de notables, al igual que en los alumnos con segunda corrección (C2) respecto de la primera (C1).
- Ningún alumno reclamante (C1) tenía sobresaliente y ninguno lo obtuvo (C2).

Por Cuestiones y Evaluador

Se presenta en la tabla 4 los porcentajes de las respuestas dadas por los alumnos a cada cuestión clasificados por cuestiones “en blanco”, “nada bien” o cero, “algo bien” y “todo bien” o puntuación máxima. Sombreados se presentan algunos resultados interesantes.

Tabla 4. Porcentajes de exámenes por tipo de respuesta y cuestión.

PROFESOR 1										
Cuestiones	A1	A2	A3	A4	A5	B1	B2	B3	B4	B5
En blanco	2.0	6.1	10.2	8.1	6.1	4.1	31.6	4.1	1.0	9.2
Nada bien	14.3	24.5	20.4	10.2	6.1	16.3	28.6	30.6	41.8	62.2
Algo bien	63.3	59.2	34.7	77.6	32.7	56.1	17.3	41.8	42.9	24.5
Todo bien	20.4	10.2	34.7	4.1	55.1	23.3	22.5	23.5	14.3	4.1
N.º Casos	49	49	49	49	49	98	98	98	98	98
PROFESOR 2										
Cuestiones	A1	A2	A3	A4	A5	B1	B2	B3	B4	B5
En blanco	1.8	5.3	3.5	8.8	7.0	4.6	20.7	2.3	4.6	9.2
Nada bien	0	36.8	8.8	14	8.8	8.0	23.0	10.3	21.8	41.4
Algo bien	87.7	57.9	70.2	71.9	45.6	75.9	42.5	70.1	71.3	47.1
Todo bien	10.5	0	17.5	5.3	38.6	11.5	13.8	17.2	2.3	2.3
N.º Casos	57	57	57	57	57	87	87	87	87	87

- La cuestión mejor valorada por el Profesor 1 es la A5 con el 55.1 % “todo bien”, seguida de la A3 con el 34.7 %.
- La cuestión mejor valorada por el Profesor 2 es la A5 con el 38.6 % “todo bien”, seguida de la A3 con el 17.5 %.
- La cuestión peor valorada es la B5 tanto por el Profesor 1 (71.4 %) como por el Profesor 2 (50.6 %) porcentajes de que la dejan en blanco o tienen cero.
- La cuestión B2 es la que más alumnos dejan en blanco tanto en el Profesor 1 (31.6 %) como en el Profesor 2 (20.7 %).
- La cuestión A2 ningún alumno la hace bien en el Profesor 2.

Por Contenidos

Se establecen los resultados atendiendo a los bloques: Álgebra, Análisis y Probabilidad/Muestreo y dentro de ellos: Discusión de sistemas,

programación lineal, función beneficio, área encerrada entre dos curvas, integrales, derivadas, probabilidad de sucesos, otras probabilidades, intervalos de confianza y contraste de hipótesis. Se presentan en las tablas 5 hasta la 14.

Los errores se analizan en los 147 exámenes evaluados por el Profesor 1.

ÁLGEBRA

A1. Discutir un sistema por el método de Gauss, según los valores del parámetro a y resolverlo para $a = 1$. (2,5 puntos).

Tabla 5. Resultados cuestión A1.

MATRICES, GAUSS, SISTEMAS				
	P1-A1	P2-A1	C1-A1	C2-A1
Media	1.19	1.28	0.82	1.09
D. T.	0.92	0.71	0.82	0.63
Casos	49	57	38	38

Error en el Método de Gauss: 42.9 %.

Error en la discusión: 20.4 %.

Error para $a=1$: 22.5 %.

No lo hace por Gauss: 2.1 %.

Esta cuestión consiste en discutir un sistema por el método de Gauss, fue aprobada por el grupo del Profesor 2, casi aprobada en el grupo del Profesor 1 y con calificaciones bajas por los Correctores (C1) y tras la revisión la subida de más de un punto (C2). El error mayoritariamente cometido en el proceso del método de Gauss es al operar con las filas tratando de reducir a una ecuación con una incógnita.

B1. Problema de programación lineal de una fábrica de tintas. (3 puntos).

Tabla 6. Resultados cuestión B1.

PROGRAMACIÓN LINEAL				
	P1-B1	P2-B1	C1-B1	C2-B1
Media	1.34	1.42	1.27	0.29
D. T.	1.10	1.01	0.97	0.91
Casos	98	87	103	103

Error planteamiento: 16.3 %.

Error representación: 37.8 %.

De ellos representan mal la recta $5x=300$ como horizontal: 13.3 %.

Error en el cálculo de los vértices: 20.4 %.

En el problema de programación lineal las calificaciones en todos los casos están próximas al aprobado (1.5), tanto en los alumnos de primera corrección como en los de segunda. El error que cometen mayoritariamente es en la representación, de ellos la tercera parte consideran la representación de una recta vertical como horizontal, y eso lleva a muchos al error de la obtención incorrecta de los vértices.

ANÁLISIS

A2. Función beneficio en la fabricación de móviles. Número de teléfonos para el beneficio sea máximo y para qué valores de x se obtienen pérdidas. (2,5 puntos).

Tabla 7. Resultados cuestión A2.

BENEFICIOS, INGRESOS Y COSTES				
	P1-A2	P2-A2	C1-A2	C2-A2
Media	1.17	0.71	0.74	0.72
D. T.	0.98	0.76	0.86	0.79
Casos	49	57	38	38

No expresa bien la función beneficio: 34.7 %.

No deriva bien $B(x)$: 6.1 %.

No obtiene bien el intervalo de beneficio: 14.3 %.

Por las puntuaciones obtenidas el Profesor 1 ha dado mejores puntuaciones que el resto de correctores (P2, C1 y C2). Posiblemente el Profesor 2 ha sido más riguroso, pues los alumnos reclamantes (C1) tienen una puntuación más alta que el grupo completo (P2). En la segunda corrección (C2) baja la media. Los errores aquí son conceptuales al equivocarse en la función beneficio y en el intervalo de beneficio, y procedimentales al derivar.

B2. Dada la gráfica de una función con un parámetro "a". Calcular a para que el área encerrada valga 57. (1,5 puntos).

Tabla 8. Resultados cuestión B2.

ÁREA ENCERRADA				
	P1-B2	P2-B2	C1-B2	C2-B2
Media	0.47	0.61	0.43	0.45
D. T.	0.64	0.64	0.61	0.54
Casos	98	87	103	103

Error planteamiento integral: 20.4 %.
 Error en obtención de primitiva: 13.3 %.
 Error en regla Barrow: 8.2 %.
 Error en obtención de a: 3.1 %.

Estando la media en 0.75 todos los grupos suspenden, las calificaciones más altas corresponden al profesor 2, y el error principal se comete al plantear la integral y en comprender los datos del enunciado. Hay errores graves como el de un alumno que hace: $f(x)=x^2+4x+a \Rightarrow f'(x)=2x+4+a \Rightarrow f''(57)=2+a \Rightarrow a=-2$. Y otro: $f(x)=x^2+4x+a \Rightarrow f'(x)=2x+4$ de donde $x=2$.

A3. Hallar dos integrales indefinidas. (1,5 puntos).

Tabla 9. Resultados cuestión A3.

INTEGRALES				
	P1-A3	P2-A3	C1-A3	C2-A3
Media	0.87	0.76	0.75	0.82
D. T.	0.64	0.49	0.50	0.46
Casos	49	57	38	38

Error integral: 32.7 %.
 Error falta constante: 36.7 %.
 Error simplificación: 4.1 %.

En esta cuestión al tratarse de derivadas inmediatas es un ejercicio sencillo y las puntuaciones en los cuatro grupos están por encima de la media (0.75), siendo los errores cometidos al aplicar el algoritmo y a la falta de inclusión de la constante, habiendo algún alumno que simplifica mal. En la segunda corrección sube la puntuación.

B3. Hallar las derivadas de dos funciones. (1,5 puntos).

Tabla 10. Resultados cuestión B3.

	DERIVADAS			
	P1-B3	P2-B3	C1-B3	C2-B3
Media	0.69	0.90	0.74	0.69
D. T.	0.59	0.52	0.57	0.54
Casos	98	87	103	103

Error en cálculo de derivada: 82.7 %.

Error en simplificación: 20.4 %.

Aquí las derivadas son inmediatas, si bien un porcentaje muy alto se equivoca al desconocer la derivada de una función exponencial y/o logarítmica. Al simplificar se obtienen muchos errores graves. La evaluación del Profesor 2 es la más alta de todas, superando el aprobado (0.75). En la segunda corrección disminuye la puntuación.

Probabilidad/Muestreo

A4. Dados dos sucesos calcular probabilidades sabiendo que son independientes. (2 puntos).

Tabla 11. Resultados cuestión A4.

	PROBABILIDAD. SUCESOS INDEPENDIENTES			
	P1-A4	P2-A4	C1-A4	C2-A4
Media	0.47	0.49	0.44	0.35
D. T.	0.38	0.48	0.27	0.22
Casos	49	57	38	38

Error al considerar que $P(A \cap B) = 0$: 36.7 %.

Error al considerar $P(A \cap B) = 1 - P(A \cup B)$: 6.1 %.

Error al considerar $P(A) \times P(B) = P(A \cup B)$: 4.1 %.

La probabilidad de sucesos independientes, muchos alumnos la confunden con la probabilidad de sucesos incompatibles, siendo el error más repetido. De ahí que al ser un resultado que se necesitaba para las preguntas posteriores las puntuaciones en todos los casos no llegan a 0,5, habiendo descendido la evaluación en la segunda corrección.

B4. Problema de probabilidad con tres apartados, probabilidad de uno sí y otro no, probabilidad de uno u otro y probabilidad de uno habiéndose producido el otro. (2 puntos).

Tabla 12. Resultados cuestión B4.

OTRAS PROBABILIDADES				
	P1-B4	P2-B4	C1-B4	C2-B4
Media	0.73	0.71	0.49	0.49
D. T.	0.76	0.60	0.65	0.60
Casos	98	87	103	103

Error al interpretar mal $P(L \cap Y^c)$: 19.4 %.

El principal error consiste en interpretar mal los datos del enunciado y traducir erróneamente. Se comete el error repetido de obtener como probabilidad un número superior a la unidad y la falta de rigor. Ningún grupo aprueba, siendo las notas de los alumnos reclamantes bajas y su corrección posterior inferior.

A5. Hallar un intervalo de confianza para el peso medio de pollos a partir de una muestra. (1,5 puntos).

Tabla 13. Resultados cuestión A5.

INTERVALO DE CONFIANZA				
	P1-A5	P2-A5	C1-A5	C2-A5
Media	1.15	1.05	1.19	1.15
D. T.	0.52	0.56	0.45	0.44
Casos	49	57	38	38

No calcula bien intervalo: 18.4 %.

Intervalo al revés: 10.2 %.

Se equivoca al sustituir: 8.2 %.

Intervalo mal: 2 %.

Al tratarse de aplicar una fórmula que los alumnos tienen aprendida, los resultados son altos, de notable, y los errores son de cálculo mayoritariamente.

B5. Test de hipótesis unilateral. Hipótesis nula igual a 30 e hipótesis alternativa menor que 30. (2 puntos).

Tabla 14. Resultados cuestión B5.

CONTRASTE DE HIPÓTESIS				
	P1-B5	P2-B5	C1-B5	C2-B5
Media	0.29	0.30	0.13	0.09
D. T.	0.58	0.55	0.31	0.26
Casos	98	87	103	103

Error al considerarlo bilateral: 56.1 %.

Error al considerarlo unilateral contrario: 10.2 %.

Mal hipótesis: 7.1 %.

El error mayoritario es el de traducción del problema. Empiezan por no establecer bien las hipótesis, la región de aceptación y de rechazo y sobre todo que se trata de un contraste unilateral y los alumnos mayoritariamente lo consideran bilateral. De ahí que los resultados sean muy bajos, los más bajos de todas las cuestiones de la prueba. En la segunda corrección disminuyen.

Por Evaluador y Opción

Para contrastar las medias entre cuestiones de las dos opciones se toman aquellas que tienen la misma valoración en ambas, que son C3 y C4, junto a la puntuación total. En tabla 15 se presentan las medias y la probabilidad, sombreadas cuando son diferencias significativas y en negrita a favor de quien es.

Tabla 15. Contraste de medias por opción.

	Profesor 1			Profesor 2		
	C3	C4	CAL	C3	C4	CAL
Opción A	0.87	0.46	4.83	0.76	0.49	4.29
Opción B	0.69	0.73	3.51	0.90	0.71	3.93
Prob.	.093	.002	.001	.121	.018	.309
	Corrector 1			Corrector 2		
	C3	C4	CAL	C3	C4	CAL
Opción A	0.75	0.44	3.94	0.82	0.35	4.13
Opción B	0.74	0.49	3.06	0.69	0.49	3.01
Prob.	.895	.622	.004	.184	.180	.000

- Hay diferencias significativas en la cuestión 4 de probabilidad corregidas por los profesores 1 y 2, con mejor resultado en la opción B.
- Hay diferencias significativas entre en las calificaciones por opción en tres de las cuatro paquetes evaluados, siempre con mejores resultados en la opción A.

Por Opción y Evaluador

Se toman las medias de las cinco cuestiones por opción, comparando las calificaciones entre los dos profesores de la primera fase y determinando la probabilidad existente que se presentan en la tabla 16. Posteriormente en la tabla 17 se presentan los resultados de las correcciones de la fase 2.

Tabla 16. Contraste de medias por evaluador y opción.

OPCIÓN A						
Cuestión	A1	A2	A3	A4	A5	CAL A
P1	1.19	1.17	0.87	0.47	1.15	4.83
P2	1.28	0.71	0.76	0.49	1.05	4.29
Prob.	.578	.008	.342	.863	.376	.195
Mín – Máx	0 – 2.5	0 – 2.5	0 – 1.5	0 – 2	0 – 1.5	.1 – 10
Media	1.24	0.92	0.81	0.48	1.09	4.54
OPCIÓN B						
Cuestión	B1	B2	B3	B4	B5	CAL B
P1	1.34	0.47	0.69	0.73	0.29	3.51
P2	1.42	0.61	0.90	0.71	0.30	3.92
Prob.	.611	.145	.011	.813	.832	.215
Min – Max	0 – 3	0 – 1.5	0 – 1.5	0 – 2	0 – 2	0 – 10
Media	1.38	0.53	0.79	0.72	0.94	3.71

- Hay diferencias significativas en los resultados de la cuestión A2, mayor puntuación en el profesor 1, aunque no se llega al aprobado.
- Hay diferencias significativas en la opción B3, mayor puntuación en el profesor 2, y que además aprueban esta cuestión.
- En las ocho cuestiones restantes no hay diferencias significativas entre las calificaciones otorgadas por los profesores 1 y 2. (CAL A)
- El Profesor 1 tiene mejor calificación en la opción A, mientras que

en la opción B (CAL B) es el profesor 2, aunque estas diferencias no son significativas.

- La nota media otorgada por los dos evaluadores en la opción A supera en ocho décimas a la otorgada en la opción B.

Tabla 17. Contraste de medias por corrector de exámenes reclamados y opción.

OPCIÓN A						
Cuestión	A1	A2	A3	A4	A5	CAL A
C1	0.82	0.74	0.75	0.44	1.19	3.95
C2	1.09	0.71	0.82	0.35	1.15	4.13
Prob.	.110	.912	.521	.127	.707	.570
Min – Max	0 – 2.5	0 – 2.5	0 – 1.5	0 – 2	0 – 1.5	1.5 – 7.5
Media	0.95	0.73	0.79	0.40	1.17	4.04
OPCIÓN B						
Cuestión	B1	B2	B3	B4	B5	CAL B
C1	1.27	0.43	0.74	0.49	0.13	3.06
C2	1.29	0.45	0.69	0.49	0.09	3.01
Prob.	.868	.791	.553	.933	.304	.830
Min – Max	0 – 3	0 – 1.5	0 – 1.5	0 – 2	0 – 2	0 – 7.5
Media	1.28	0.44	0.72	0.49	0.11	3.03

- No hay diferencias significativas en ninguna de las diez cuestiones ni tampoco en las calificaciones finales entre los dos correctores de la misma muestra.
- En la opción B mejor nota, por muy poco, a favor de C1, mientras que en la opción A es al revés.
- La nota media otorgada en las dos correcciones de reclamación, la opción A supera en un punto (de 0 a 10) a la de la opción B.

Discusión

Ruiz de Gauna, Sarasua y García (2011, p. 23) indican que “las Matemáticas Aplicadas a las Ciencias Sociales es la asignatura con mayor índice de suspensos en el conjunto de las PAU, tanto a nivel de la Universidad del País Vasco, como del conjunto de España” y en nuestro estudio comprobamos la cantidad de suspensos en donde en el total de las PAU

junio 2014 en la Universidad de Murcia la nota media es de 4.16 y en particular en las dos muestras estudiadas 3.95 y 4.07, respectivamente. En otro estudio Gairin, Muñoz y Oller (2013) con 419 alumnos y 3 evaluadores obtuvieron de media 3.75 y en Nortes y Nortes (2010) la media fue de 4.5 en una muestra de 134 alumnos.

En Matemáticas Aplicadas a las Ciencias Sociales en el presente estudio de cada tres alumnos examinados dos suspenden y de cada siete alumnos que piden revisión de examen seis suspenden, aumentando la nota media tras la revisión en 23 milésimas.

El que la diferencia entre dos correctores distintos al corregir distintas muestras sea de 12 centésimas y que la diferencia entre cada muestra y la media global de la prueba sea de 21 y 9 centésimas, respectivamente, da a entender que los criterios de corrección han sido precisos y que los evaluadores se han ajustado a ellos.

En Gairin, Muñoz y Oller (2012) para una nota de 3.86 obtenían que el 6 % de alumnos estaban en el intervalo [8-10] y el 50 % en el intervalo [0-4], y en el presente estudio el Profesor 1 para una nota media de 3.95 los porcentajes son 5.4 % y 57.1 % y el Profesor 2 para una nota media de 4.07 los porcentajes son 4.9 % y 49.3 %, respectivamente, lo que indica que estas tres evaluaciones distan mucho de las del segundo corrector al que citaban los anteriores autores en su estudio que para una media de 5.74 los intervalos eran de 12 % y 15 %, respectivamente, lo que nos lleva a suponer que el segundo corrector no se ajustó a los criterios de corrección.

La cuestión B2 de calcular el área encerrada por una curva no fue contestada por el 31.6 % y por el 20.7 % en los paquetes de los dos profesores iniciales que sumado a los que obtuvieron calificación de cero hace que más del 60 % en un caso y más del 40 % en el otro caso no tuvieran ninguna puntuación positiva en dicha cuestión. Por el contrario en la cuestión A5 de calcular un intervalo de confianza es el 55 % y el 38 % en cada caso los que tienen una puntuación máxima. Su correspondiente B5 por el contrario es la que más ceros tienen (62.2 % y 44.4 %, respectivamente).

Hay altos porcentajes de errores en la discusión del sistema al aplicar el método de Gauss y en el planteamiento del problema de programación lineal, en donde la representación y en concreto la confusión de una recta paralela a un eje de coordenadas y no al que correspondía, hace que el recinto no sea el correcto. En la segunda cuestión el no

escribir bien la función beneficio y el planteamiento de la integral en el área encerrada entre dos rectas son los errores más destacados. Una derivada inmediata y una integral inmediata, si se sabe el procedimiento, es un algoritmo sencillo, pero si se desconoce, difícilmente se podrá resolver bien. Con la cuestión de probabilidad es un hecho que se repite año tras año, la confusión entre probabilidad de sucesos incompatibles y de sucesos independientes hace que sea el error más grave así como la interpretación de “tome leche pero no yogur” que originó casi un 20 % de errores. En la última cuestión el calcular el intervalo y considerar el contraste bilateral en lugar de unilateral, son los errores más repetidos.

Respecto a la cuestión 5 sobre muestreo hay una gran diferencia entre las puntuaciones obtenidas. Mientras que en una opción contestan mayoritariamente en el cálculo de un intervalo de confianza con una puntuación de notable, en la otra opción de contraste de hipótesis unilateral los alumnos obtiene las notas más bajas de las diez cuestiones analizadas. Resultados que tienen su correspondencia con otros estudios como el de García y García (2005) en donde la cuestión de contraste de hipótesis en una muestra de las PAU junio 2002 en Gran Canaria, tan solo lo resolvió bien el 18 %.

Analizando en su conjunto la prueba, en donde la opción A es elegida por la tercera parte de los alumnos obteniendo una nota superior en un punto a la opción B, y clasificando cada cuestión en Estándar (E) o Difícil (D), siguiendo a Ruíz de Gauna, Sarasua y García (2011) en donde consideran una dificultad E si la comprensión del texto es sencilla, el ejercicio es identificable con su técnica de resolución, no dando lugar a cálculos largos y complicados; y consideran una dificultad D si el texto es difícil de comprender, las técnicas de resolución no son fácilmente identificables y los cálculos pueden llevar al alumno a cometer errores, podríamos indicar que A1, A3, A5, B1 y B3 son de dificultad E, mientras que el resto son de dificultad D, teniendo la opción A tres estándar y dos difícil al contrario que la opción B.

Conclusiones

En las PAU, el reparto de exámenes a corregir es de forma anónima y al azar. El que entre dos correctores distintos que corrigen dos muestras distintas la diferencia entre las calificaciones no sea significativa explica,

por una parte, que la distribución de puntuación ha sido detallada por parte del coordinador y por otra, que los criterios de corrección se han aplicado detalladamente, contestando así al objetivo 1.

Tampoco existen diferencias significativas cuando corrigen los mismos exámenes evaluadores distintos, pues en las reclamaciones, entre la primera puntuación y la segunda no existen diferencias significativas entre ellas, contestando así al objetivo 2.

Resultados estos que no confirman los obtenidos por Gairin, Muñoz y Oller (2012) en donde constatan la existencia de diferencias significativas entre las calificaciones de distintos correctores y según nuestro estudio no hacen necesaria la propuesta de efectuar pruebas piloto que informen del nivel de dificultad de la prueba, tal como indican Grau, Cuxart y Martí-Recober (2002).

La cuestión de probabilidad, con la puntuación más baja, nos indica que las capacidades de entender y expresarse con claridad no están adquiridas por parte de los estudiantes y como dicen Boal, Bueno, Levis y Sein-Echaluce (2008, p. 21) los alumnos de la PAU en Matemáticas “continúan siendo dependientes de las explicaciones de alguien que les traduzca la información matemática”.

Lo que si se ha detectado en este estudio es la existencia de una opción que incluía una cuestión sobre contraste de hipótesis con calificaciones muy bajas, lo que puede suponer que ese contenido al estar situado al final de la programación de la asignatura no ha sido estudiado suficientemente o bien a un malentendido entre coordinador de la asignatura y profesores de 2.º de Bachillerato al programar los contenidos.

Como aspectos positivos, están, por una parte, el detallado desglose de puntuación en cada cuestión realizado por el coordinador y, por otra, el seguimiento por parte de los evaluadores en la corrección que ha llevado a no encontrar diferencias significativas entre dos evaluadores al corregir distintas muestras y entre evaluadores distintos al corregir la misma muestra. Como aspectos negativos, que cada cuestión tenga distinta puntuación según la opción elegida en tres de las cinco cuestiones, pues de tener la misma puntuación se podría haber evitado lo ocurrido en la cuestión 5 en donde la diferencia de puntuación entre las dos opciones es de medio punto, mientras que el grado de dificultad es bastante superior en una que en otra.

El resultado de algunas cuestiones de este estudio lo anticipan Mallart (2014) y Nortes y Nortes (2010) en que dicen que los alumnos tie-

nen asimilados los procesos mecánicos de resolución de un problema, pero cuando hay que hacer un planteamiento como en las cuestiones de probabilidad o en el contraste de hipótesis las dificultades y los errores aumentan considerablemente.

Tal como indican Gairin, Muñoz y Oller (2013, p. 42) “como quienes preguntan, quienes responden y quienes califican las respuestas tiene visiones personales diferentes sobre un mismo concepto matemático”, se pone de manifiesto la necesidad de trabajar sobre los criterios de calificación proporcionados a los correctores y extraer conclusiones de los resultados para poder mejorarlos en las Pruebas de Acceso a la Universidad.

Referencias bibliográficas

- Boal, N., Bueno, C., Leris, M. D. y Sein-Echaluce, M. L. (2008). Las habilidades matemáticas evaluadas en las pruebas de acceso a la universidad. Un estudio en varias universidades públicas españolas. *Revista de Investigación Educativa*, 26 (1), 11-23.
- Gairin, J.M., Muñoz, J.M. y Oller, A.M. (2013). Anomalías en los procesos de identificación de errores en las pruebas escritas de matemáticas en las P.A.U. *Campo Abierto*, 32 (2), 27-50.
- Gairin, 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. Ordoñez (Eds.). *Investigación en Educación Matemática XVI* (pp 261-274). Jaén: SEIEM.
- García, I y García, J. A. (2005). Algunos resultados sobre la actuación de los alumnos en las cuestiones de estadística en la PAU. En Actas de las XI Jornadas para el Aprendizaje y la Enseñanza de las Matemáticas (pp. 733-738). Canarias: Consejería de educación, cultura y deporte (Gobierno de Canarias).
- Grau, A., Cuxart, A. y Marti-Recober, M. (2002). La calidad en el proceso de corrección de las pruebas de acceso a la universidad: variabilidad y factores. *Revista de Investigación Educativa*, 20 (1), 209-223.
- Huidobro, J. A., Méndez, A. y Serrano, M. L. (2010). Del Bachillerato a la Universidad: las Matemáticas en las carreras de ciencias y tecnología. *Aula Abierta*, 38 (1), 71-80.
- Lorenzo, M., Argós, J., Hernández, J. y Vera, J. (2014). El acceso y la entrada del estudiante a la universidad: situación y propuestas de mejora facilitadoras del tránsito. *Educación XXI*, 17 (1), 15-38.
- Mallart, A. (2014). La resolución de problemas en la prueba de Matemáticas de acceso a la Universidad: procesos y errores. *Educatio siglo XXI*, 32 (2), 233-254.
- Mengual, E., Gorgorigo, N. y Albarracín, L. (2013). Validación de un instrumento para la calificación de exámenes de Matemáticas. En A. Berciano, G. Gutierrez, A. Estepa y N. Climent (Eds.), *Investigación en Educación Matemática XV* (pp. 367-381).

- Nortes, R. y Nortes, A. (2010). Resolución de problemas de matemáticas en las pruebas de acceso a la universidad. Errores significativos. *Educatio Siglo XXI*, 28 (1), 317-342.
- Ramos, C. E., Espinel, M. C. y Ramos, R. M. (2009). Identificación de los errores en los contrastes de hipótesis de los alumnos de Bachillerato. *Suma*, 61, 35-44.
- Ruíz de Gauna, J. (2010). La enseñanza de las matemáticas del bachillerato, los libros de texto y las pruebas de acceso a la UPV-EHU (1970-2008). (Tesis doctoral no publicada). Donostia: Universidad del País Vasco.
- Ruíz de Gauna, J. Sarasua, J. y García, J. M. (2011). Una tipología y clasificación de los ejercicios de matemáticas de selectividad. *Épsilon*, 28 (2), 21-38.

