

Influencia del tipo de evaluación en la calificación. Un análisis estadístico aplicado a la materia Matemáticas II de ADE

Influència del tipus d'avaluació en la qualificació. Una anàlisi estadística aplicada a la matèria Matemàtiques II d'ADE

Influence of the type of assessment on grading. A statistical analysis applied to the subject Mathematics II in ADE

J. Antonio Seijas-Macías

Departamento de Economía, Facultade de Economía e Empresa
Universidade da Coruña
Campus de Elviña, 15071 A Coruña, España
<https://orcid.org/0000-0002-6056-3257>
antonio.smacias@udc.gal

Recibido: 28/07/2023 | Aceptado: 11/12/2023 | Publicado: 01/2024

Resumen

En el presente trabajo realizamos un análisis estadístico para determinar en qué medida la elección del proceso de realización de la evaluación afecta a las calificaciones que obtiene el alumnado de la materia. Nuestro trabajo se ha centrado en la asignatura de Matemáticas II del grado en Administración y Dirección de Empresas (ADE). Dicha materia se imparte en el segundo cuatrimestre del curso de primero del grado en ADE. Hemos considerado los últimos cuatro años en los cuales se han utilizado tres sistemas evaluativos diferentes. El año 2018-2019 todo el proceso de evaluación fue presencial, y el 2019-2020, por el confinamiento, todo el proceso de evaluación fue mediante pruebas online (virtual). El año 2020-2021 combinamos pruebas presenciales y pruebas online dentro de la evaluación generando un sistema semipresencial; y finalmente el año 2021-2022 supuso el regreso al sistema del año 2018-2019, siendo todo presencial. El análisis estadístico de los datos obtenidos muestra que las calificaciones obtenidas son estadísticamente diferentes en los tres procesos. Además, el número de personas que participan en la evaluación se ve afectado y los mejores índices de participación corresponden al sistema semipresencial.

Palabras clave: Procesos de evaluación, Resultados, Matemáticas ADE, Test Kruskal-Wallis, Test Chi-Cuadrado.

Resum

En aquest treball fem una anàlisi estadística per determinar en quina mesura l'elecció del procés de realització de l'avaluació afecta les qualificacions que obté l'alumnat de la matèria. El nostre treball s'ha centrat en l'assignatura Matemàtiques II del grau en Administració i Direcció d'Empreses (ADE). Aquesta matèria s'imparteix al segon quadrimestre del curs de primer del grau a ADE. Hem considerat els darrers quatre anys en què s'han utilitzat tres sistemes avaluatius diferents. L'any 2018-2019 tot el procés d'avaluació va ser presencial, tot seguit, l'any 2019-2020, a causa de la situació de confinament, tot el procés d'avaluació va ser mitjançant proves en línia (virtual), l'any 2020-2021 vam combinar proves presencials i proves en línia dins de l'avaluació generant un sistema semipresencial; i finalment l'any 2021-2022 va suposar el retorn al sistema de l'any 2018-2019, tot sent presencial. L'anàlisi estadística de les dades obtingudes mostra que les qualificacions obtingudes són estadísticament diferents en els tres processos. D'altra banda, també el nombre de persones que participen a l'avaluació es veu afectat. Els resultats mostren que els millors índexs de participació corresponen al sistema semipresencial.

Paraules clau: Processos d'avaluació, Resultats, Matemàtiques ADE, Test Kruskal-Wallis, Test Chi-Quadrat.

Seijas-Macías, J. A. (2024). Influencia del tipo de evaluación en la calificación. Un análisis estadístico aplicado a la materia Matemáticas II de ADE. *RIDU. Revista d'Innovació Docent Universitària*, 16, 36-48. <https://doi.org/10.1344/RIDU2024.16.4>

© 2024 El autor. Este artículo es de acceso abierto sujeto a la licencia Reconocimiento 4.0 Internacional de Creative Commons, la cual permite utilizar, distribuir y reproducir por cualquier medio sin restricciones siempre que se cite adecuadamente la obra original. Para ver una copia de esta licencia, visite: <https://creativecommons.org/licenses/by/4.0/>

Abstract

In this paper, we carry out a statistical analysis to figure out the extent to which the choice of assessment process affects the grades obtained by students in the subject. Our work has focused on the subject of Mathematics II of the degree in Business Administration and Management (ADE). This subject is taught in the second semester of the first year of the ADE degree. We have considered the last three years in which three different evaluation systems have been used. In the year 2018-2019 the entire assessment process was face-to-face; then, in the year 2019-2020, due to the confinement situation, the entire assessment process was utilizing online tests (virtual); in the year 2020-2021 we combined face-to-face and online tests within the assessment, generating a blended system; and finally, the year 2021-2022 meant the return to the 2018-2019 system, with everything being face-to-face. The statistical analysis of the data obtained shows that the grades obtained are statistically different in the three processes. On the other hand, the number of people taking part in the assessment is also affected. The results show that the best participation rates correspond to the blended learning system.

Keywords: Evaluation processes, Results, Mathematics ADE, Kruskal-Wallis Test, Chi-Squared Test.

1. Introducció

La asignatura de Matemáticas II se imparte en el grado en Administración y Dirección de Empresas (ADE) y en el programa de simultaneidad de los grados en ADE y Derecho de la Facultad de Economía y Empresa de la Universidade da Coruña (UDC). Esta asignatura se integra en el plan de estudios de la titulación en el primer año (segundo cuatrimestre) en ADE y en el segundo año (segundo cuatrimestre) en el programa de simultaneidad. El sistema de evaluación consta de dos partes diferenciadas: un proceso de evaluación continua y un examen final. El equilibrio entre ambos sistemas se basa en la regla 70-30, donde ninguna de las dos partes debe superar el 70% de la cualificación final, ni ser inferior al 30%. El sistema de evaluación continua consiste en la realización de diversas pruebas (test, exámenes cortos, boletines de ejercicios, ...) durante el cuatrimestre. Por su parte el examen final se produce en dos oportunidades: una primera oportunidad al finalizar la docencia del cuatrimestre y una segunda oportunidad en el mes de julio (antes de la finalización del curso académico).

En este trabajo, hemos analizado la presencia de diferencias significativas en los resultados de evaluación obtenidos por el alumnado en la materia de Matemáticas II en los años académicos 2018-2019, 2019-2020, 2020-2021 y 2021-2022. La elección de estos cuatro años responde al hecho de que durante los mismos se ha utilizado sistemas de evaluación y docencia claramente diferenciados motivados por la situación de pandemia que ha afectado al estado español desde el año 2020.

El análisis de los datos se ha basado en los valores numéricos de la cualificación final reflejada en las actas de la asignatura de los años estudiados.

2. El impacto de las TIC sobre la docencia y la evaluación en Matemáticas para Economía y Empresa

La introducción del Espacio Europeo de Educación Superior (EEES) a finales de la primera década del siglo XXI ha supuesto un cambio importante en el paradigma de la enseñanza y de la evaluación de las materias en la universidad. La transformación del sistema universitario propiciada por el EEES implicó la necesidad de adaptar los sistemas evaluativos (Zaragoza et al., 2009; Sánchez, 2010). La evaluación ha adquirido un nuevo enfoque formativo frente al enfoque evaluativo anterior (Vallés et al., 2011; Fraile et al., 2013). En este sentido, el sistema de evaluación continua se ha erigido como aquel que proporciona los mejores resultados (Arribas, 2012).

El cambio en el sistema evaluativo se ha visto propiciado por la introducción de las TIC dentro del proceso de enseñanza y evaluación del estudio presencial (Rodríguez, 2008). La mayoría de las herramientas

informáticas aplicadas a la evaluación se ha centrado en el uso del software para el diseño de pruebas tipo test, de respuesta cerrada, sin entrar en los planteamientos pedagógicos necesarios (Rodríguez, 2008). En esta línea, se proponen la realización de pruebas no sólo en base a la tecnología disponible, sino conforme a criterios pedagógicos admitidos (García et al., 2002). En la misma línea, Barberà (2016) realiza un análisis sobre los procesos evaluativos y las aportaciones de la tecnología a los mismos. Un amplio estudio de la evaluación desde una perspectiva histórica a lo largo del siglo XX se realiza en Escudero (2003).

Los resultados de la evaluación de las materias de matemáticas y estadística en las titulaciones de ciencias sociales, en particular en las de ámbito económico: economía, empresariales, administración y dirección de empresas, finanzas, auditoría, etc. ha sido objeto de amplios estudios en la literatura especializada: Alacid et al. (2017) se centran en los resultados en base a la destreza del alumnado de primer curso. Por su parte, Gómez-Déniz et al. (2015) realizan un estudio sobre los factores determinantes en las notas de matemáticas empresariales y concluyen que la media de la nota aumenta más con la edad y la nota de acceso. Boncompte et al. (2010) también realizan un perfil del alumnado de primer curso, concluyendo la necesidad de tener grupos lo más homogéneos posible, y la implantación de la matrícula semestral frente al sistema actual de matrículas anuales. Más recientemente, Esparza y Márquez (2020), mantienen el enfoque en la heterogeneidad del alumnado de estas titulaciones. Kaspriková (2012) realiza un análisis estadístico sobre los exámenes tipo test de la materia de matemáticas para economistas e identifica la existencia de dos grupos diferenciados de estudiantes en base a su rendimiento académico en la materia. No obstante, en estos trabajos no se analiza el efecto que la realización de pruebas online puede tener sobre los resultados finales de la materia estudiada.

La realización de pruebas de evaluación a distancia se ha difundido a raíz de la implantación de las TIC en el proceso formativo. Existe una amplia literatura sobre los tipos de evaluación y sus posibles efectos sobre el rendimiento del alumnado. En este punto, las investigaciones se han centrado en destacar la diferencia en rendimientos entre enseñanza presencial y enseñanza a distancia (online). Rane y Mackenzie (2020) realizan una comparativa entre los resultados de los estudiantes en la realización de test online o pruebas presenciales tradicionales en un curso de ingeniería financiera. Aunque sus conclusiones son limitadas, la realización de pruebas online puede favorecer la mejora resultados en particular en grupos de docencia grandes. En la misma línea, Matusek y Greganova (2019) presentan un estudio sobre una comparativa en los resultados de los exámenes de Matemáticas en la Facultad de Economía y Gestión, entre dos cursos de Matemáticas con diferente orientación. En su trabajo, aplican un análisis estadístico de diferencia de medias para distribuciones normales, y concluyen la presencia de diferencias significativas en los resultados de ambas asignaturas.

La crisis del coronavirus ha propiciado una mayor difusión de la enseñanza online y de la necesidad de realizar las pruebas de evaluación mediante métodos no presenciales (Rapanta et al., 2020). Ante esta situación, algunos autores se han cuestionado si los resultados de la evaluación se habían visto afectados por el sistema de evaluación empleado, llegando a la conclusión de la presencia de significativas diferencias a favor de los exámenes online (Klůfa, 2021). Este hecho contrasta con algunos estudios anteriores que no encuentran diferencias significativas en los resultados obtenidos por los estudiantes a la hora de realizar los exámenes mediante métodos tradicionales y exámenes online (Yates y Beaudrie, 2009; Kirtman, 2009).

En línea con los trabajos de Matusek y Greganova (2019), Rane y MacKenzie (2020), Rapanta et al. (2020) y Klůfa (2021), en este trabajo hemos realizado un análisis comparativo para comprobar si la utilización de sistemas de evaluación online, durante la crisis del coronavirus, ha supuesto alguna variación significativa en los resultados de la evaluación en la materia de Matemáticas II del grado en ADE de la Universidad de A Coruña. Aunque somos conscientes de que hay diversos aspectos que pueden explicar la diferencia de resultados en los años analizados: alumnado diferente, diferencias en los criterios de ponderación de las pruebas, diferencias en los niveles de conocimientos del alumnado, notas de acceso, ... Creemos que, aunque de una forma parcial, gran parte de las diferencias en los años del período analizado se pueden haber visto motivadas de forma muy sustancial por la aplicación de un sistema de evaluación

online que, ha supuesto una ruptura muy importante del esquema establecido en sistema de valoración del trabajo del alumnado.

3. Estudio de los datos de calificaciones de Matemáticas II

La asignatura de Matemáticas II se imparte en el grado de ADE y en el programa de simultaneidad de ADE + Derecho de la Facultad de Economía y Empresa de la Universidad de A Coruña. La materia se imparte en el segundo cuatrimestre del Primer Curso en ADE y, en el mismo cuatrimestre, pero en el Segundo Curso en ADE + Derecho. Los contenidos de la materia engloban el estudio de las funciones de varias variables (vectoriales y reales) y del cálculo diferencial en una primera parte. La segunda parte se centra en el estudio de la Optimización de funciones reales: sin restricciones, con restricciones de igualdad y finalmente, el estudio de la programación lineal.

Las personas matriculadas tienen dos oportunidades para superar la materia: en junio (1ª oportunidad) y en julio (2ª oportunidad). La superación de la materia se consigue mediante una calificación global igual o superior a 5 puntos que se obtienen mediante una serie de pruebas de evaluación continua y un examen final (cuya ponderación oscila entre el 40% y 60%). En ningún caso, el examen final puede superar el 70% de la calificación global de la materia¹. En este sentido, el alumnado no puede renunciar a la evaluación continua y acudir con un examen final del 100%.

La materia de Matemáticas II se ha tenido que adaptar, desde el año 2020, a las circunstancias especiales derivadas de la situación de pandemia que ha afectado a España y al conjunto del Mundo durante estos últimos años. Esto ha supuesto que los sistemas de evaluación y de docencia se modificaran respecto a los habituales en las universidades presenciales.

El año académico 2018-2019 se siguió un sistema tradicional presencial. Las clases eran presenciales y el alumnado debía asistir a ellas, para ello se realizaba un control de asistencia que, si no se cumplía, conllevaba la pérdida de la evaluación continua. La evaluación continua ponderaba el 50% de la calificación final y consistía en pruebas de tipo test y pruebas escritas de respuesta breve, todas ellas realizadas presencialmente en las aulas de la facultad. Por su parte, el examen final se realizaba en las fechas marcadas de forma presencial.

El año académico 2019-2020 comenzó con los mismos criterios del año anterior, pero debido al confinamiento de la población y la suspensión de las actividades académicas presenciales en las universidades, se realizó una rápida adaptación a un sistema completamente online (o virtual). La docencia se impartió de forma telemática y se eliminó el requisito de asistencia. La evaluación continua fue realizada online mediante pruebas tipo test y boletines de ejercicios con entrega en un período de tiempo pactado (3 horas). Por su parte, el examen final se realizó en las fechas pactadas pero su realización fue online. Dicho examen constaba de una parte tipo test y de una parte de ejercicios de desarrollo que se realizaban en un periodo máximo de 2 horas. La ponderación entre ambas partes fue del 50% para cada una de ellas.

El año académico 2020-2021 permitió retomar, de forma parcial, las actividades presenciales y se realizó mediante un sistema mixto donde se combinaba docencia presencial y online. Las pruebas de evaluación continua se realizaron de forma presencial (pruebas prácticas) y de forma online (pruebas tipo test). Por su parte, el examen final se realizó de forma presencial. En este año, no se exigieron requisitos de asistencia y se modificó la ponderación entre las partes de la evaluación: el examen final representó el 40% de la cualificación y las pruebas de evaluación continua ponderaban el 60%. No obstante, en la segunda oportunidad los criterios se modificaban y se realizaba una nueva ponderación de la evaluación continua que pasaba a ser el 50% de la calificación final, siendo el examen el otro 50%.

¹ Esta situación se ha modificado para el curso 2022/23, aunque dicho curso no es objeto de análisis de este trabajo.

Finalmente, el año académico 2021-2022 supuso la vuelta a la normalidad. Todas las clases y pruebas se realizaron de forma presencial (aunque aún se mantuvieron unas horas de docencia online) pero ahora las pruebas eran todas presenciales, tanto exámenes como pruebas de evaluación continua. Al igual que en los años precedentes no se exigieron requisitos de asistencia. Este año se invirtieron las ponderaciones y así el examen final tenía una ponderación del 60% sobre la calificación final, mientras que la evaluación continua era el 40% restante. En este caso en la segunda oportunidad, se introdujo una nueva prueba de evaluación continua que representaba el 40% y permitía mejorar la calificación obtenida en la primera oportunidad.

En todos estos años se han personalizado las pruebas tipo test y todas las pruebas realizadas online.

Para el análisis de los datos de nota media, solo hemos considerado las calificaciones del alumnado que ha participado de forma activa en la asignatura, esto es, ha realizado pruebas que ponderan en la evaluación un total igual o superior al 20%. Por tanto, se han eliminado las personas con la cualificación de No Presentado. Tampoco se han considerado, los posibles casos de personas que convalidan la materia. En la Tabla 1 podemos ver un resumen de las cifras correspondientes a los años estudiados.

	Alumnado	Media	Mediana	Desv. Típica	Máximo	IQR ²
2018/2019	254	4.9652	5.18	2.0822	9.33	3.052
2019/2020	291	5.3130	5.81	2.4002	9.95	3.593
2020/2021	299	3.7425	3.94	2.3796	8.98	4.195
2021/2022	323	3.4655	3.73	2.4621	9.74	4.545

Tabla 1: Matemáticas II (ADE – UDC)

Los datos muestran que durante estos cuatro años ha aumentado el número de personas que participan de forma activa en la materia, especialmente, durante el año 2019/2020, pero la evolución de las notas medias ha sido más errática. Por su parte, la desviación típica se ha mantenido en valores relativamente altos, lo que indica que hay una gran variabilidad en las calificaciones obtenidas por el alumnado.

Las gráficas (Figura 1) muestran la distribución de las calificaciones durante los años estudiados. Se observa que dichas distribuciones presentan diferencias claras entre ellas. Este hecho nos lleva a considerar si las diferencias observadas en la nota media para los cuatro años analizados han sido significativas. También hay que destacar que la distribución ha pasado de ser una distribución unimodal en los cursos 2018/19 y 2020/21 a una distribución claramente bimodal en los dos cursos siguientes.

Para analizar la presencia de dichas diferencias se puede utilizar el análisis ANOVA de un factor que permite examinar las diferencias en las medias de tres o más grupos (McDonald, 2014). Los requisitos para poder realizar un ANOVA son:

1. Observaciones aleatorias e independientes.
2. Distribución normal de los datos.
3. Homocedasticidad (varianza constante entre grupos).

Las cuatro muestras analizadas son aleatorias y claramente independientes. En cuanto al segundo requisito, en nuestro caso, una simple observación de las gráficas permite sospechar la no normalidad de los datos. No obstante, es necesario confirmar esta suposición con un análisis más profundo.

En la Figura 2 hemos representado los gráficos tipo qq-plot de las calificaciones correspondientes a los diferentes años. Los gráficos muestran que las distribuciones de calificaciones se alejan de la normal (especialmente, para los años 2019/2020 y 2020/2021).

² IQR: Recorrido Intercuartílico.

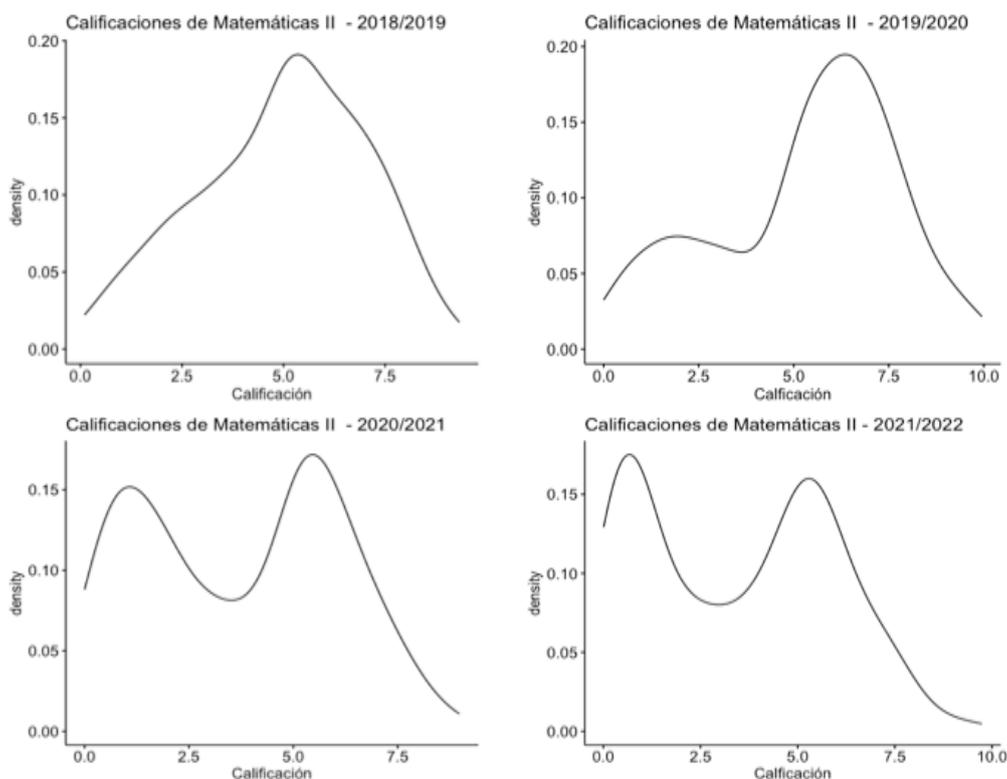


Figura 1: Evolución de las calificaciones de Matemáticas II

Con el objetivo de confirmar o descartar la normalidad hemos utilizado el test de Shapiro-Wilk (Shapiro y Wilk, 1965) (véase Tabla 2). Los resultados de dicho test son concluyentes y nos permiten descartar el supuesto de normalidad en la distribución de las calificaciones de Matemáticas II.

En base a estos datos, debemos rechazar la hipótesis de normalidad y, por tanto, no utilizar el análisis de ANOVA un factor para comparar las medias de las calificaciones de Matemáticas II.

Dada la no normalidad de los datos a analizar optamos por utilizar un método alternativo de estudio de la diferencia de medias.

El test de Kruskal-Wallis (Kruskal y Wallis, 1952) es un test no-paramétrico alternativo al test ANOVA de un factor que extiende el test de dos muestras de Wilcoxon (Wilcoxon, 1945) a situaciones donde tenemos más de dos grupos de datos.

Se recomienda utilizar este test cuando los datos no verifican las hipótesis del test ANOVA de un factor. Aunque el test de Kruskal-Wallis no requiere que los datos sigan una distribución normal, si asume que los datos presentan homocedasticidad.

Hemos utilizado el test de Levene (Levene, 1960) (dada la no normalidad de los datos) y el resultado ha sido (ver Tabla 3) que no podemos descartar la presencia de heterocedasticidad. A pesar de este diagnóstico, hemos utilizado la regla (Dean y Voss, 1999) de que la varianza máxima no supere en más de tres veces la varianza mínima (ver Tabla 3) para considerar que no hay un problema de heterocedasticidad. Dado que la ratio entre la varianza máxima y la mínima es $1.4 < 3$, consideramos que no hay heterocedasticidad, y hemos optado por utilizar el test de Kruskal-Wallis para analizar las diferencias de medias.

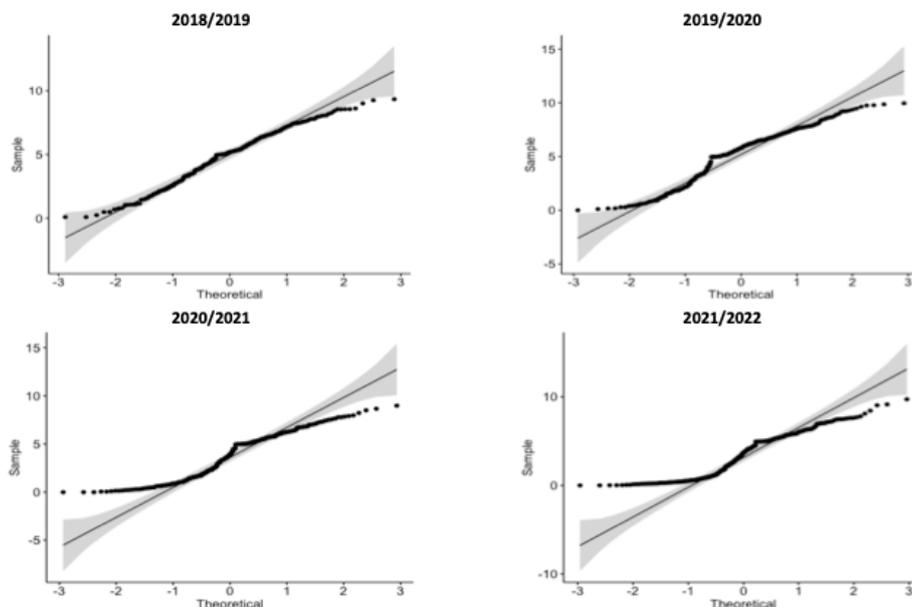


Figura 2: QQ-plot de datos de calificaciones de Matemáticas II

	W	p-valor
2018/2019	0.98071	0.001597
2019/2020	0.95116	2.869e-08
2020/2021	0.93659	5.203e-10
2021/2022	0.92467	1.093e-11

Tabla 2: Test de Normalidad Shapiro-Wilk

	F valor	Pr(>F)
Test de Levene	9.9597	1.746e-06
	Varianza	
2018/2019	4.3357	
2019/2020	5.7610	
2020/2021	5.6629	
2021/2022	6.0662	

Tabla 3: Análisis de Heterocedasticidad

3. Análisis de los resultados

Para analizar la calificación media hay que considerar que la distribución de notas es bimodal para los dos últimos cursos del período analizado. En este sentido, hay que tener en cuenta el hecho de que los gráficos de tipo boxplot no pueden representar de forma adecuada este tipo de distribuciones y se hace necesario realizar correcciones sobre las gráficas a utilizar (Choondpradub y McNeil, 2005; Martínez, 2015).

En esta situación hemos optado por realizar un análisis gráfico de las medias utiliza el paquete gráfico Beanplot (Kampstra, 2008). Este paquete permite comparaciones visuales de datos univariantes entre grupos y observar anomalías como distribuciones bimodales o medidas duplicadas. En la Figura 3, observamos el beanplot correspondiente a los cuatro cursos considerados. En el gráfico, la línea horizontal azul representa el valor de la media aritmética de cada año. Por su parte los gráficos de densidad de cada año muestran como a lo largo del período analizado la distribución de las calificaciones se ha desplazado de una distribución unimodal en el curso 2018/2019 a una distribución claramente bimodal en el año 2021/2022.

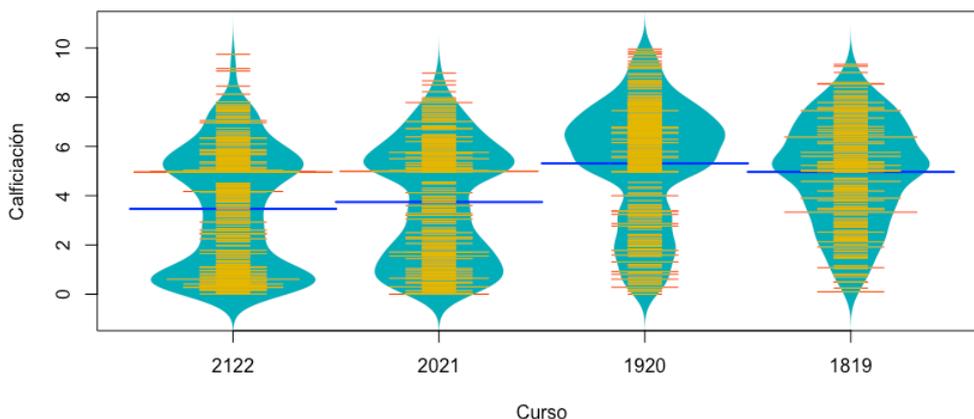


Figura 3: Beanplot datos de calificaciones de Matemáticas II

El análisis de la cualificación media muestra diferencias significativas en los cuatro años académicos estudiados. El mayor valor corresponde al año 2019/2020 donde todo el proceso de evaluación y, casi todo el proceso de docencia se tuvo que realizar de forma no presencial (online). Por el contrario, el año con los menores valores de nota media corresponde al año 2021/2022 donde el proceso de evaluación totalmente presencial, aunque se mantenía un cierto nivel mínimo de docencia online, junto con el año anterior donde la docencia y la evaluación combinaba actividades presenciales y no presenciales (online). El año 2018/2019, último año totalmente presencial, presenta valores intermedios (véase Figura 3).

También el grado de variabilidad ha ido en aumento, los valores mínimos corresponden al año presencial, mientras que la mayor variabilidad se produce en los años 2021-22 y 2020-21.

El test de Kruskal-Wallis confirma la presencia de diferencias significativas entre los cuatro grupos de calificaciones (ver Tabla 4). Por su parte, la comparación por cursos muestra que las diferencias también son significativas cuando comparamos los cuatro años por pares, siendo los años 2020/2021 y 2021/2022 los más próximos.

	Chi-cuadrado	p-valor	
Test de Kruskal-Wallis	120.84	2.2e-16	
Comparación Pares	21/22	20/21	19/20
	2020/2021	0.068	--
	2019/2020	2e-16	2.8e-14
	2018/2019	8.1e-13	7.43e-09
			0.025

Tabla 4: Diferencia de valores medios

A continuación, hemos analizado cuál ha sido la evolución de resultados en los cuatro años analizados.

Para ello hemos considerado el total de personas matriculadas y las hemos asignado a tres grupos: los que superan la materia (calificación final igual o superior a 5.0), lo que no superan la materia (calificación final inferior a 5.0) y los que no participan de forma activa que se califican con “No presentado”. En la Tabla 5, tenemos los valores globales.

	Superada	No Superada	No Presentado
2018/2019	151/411 (36,73%)	103/411 (25,06%)	157/411 (38,20%)
2019/2020	205/417 (49,16%)	86/417 (20,62%)	126/417 (30,22%)
2020/2021	138/350 (39,43%)	161/350 (46%)	51/350 (14,57%)
2021/2022	133/369 (36,04%)	190/369 (51,49%)	46/369 (12,47%)

Tabla 5: Resultados de Matemáticas II

El curso 2019/2020 ha supuesto un aumento del número de personas que superan la materia, obteniendo el máximo de cuatro cursos analizados, lo que ha supuesto que se reduzcan los suspensos y el de no presentado. No obstante, el año 2021/2022 rompe la tendencia, de forma que presenta los peores valores de personas que superan la materia, aunque muy similar al del curso 2018/2019, pero con la menor cifra de no presentado del período analizado. Después del año 2019/2020, el número de personas que no se presentaba a la evaluación de la materia ha caído de forma importante. No obstante, este menor índice de absentismo sobre el proceso de evaluación no se ha visto acompañado por un aumento en el número de personas que aprueban la materia, sino que ha ido a aumentar el número de personas que no superan la materia y marcaba su máximo en el curso 2021/2022.

En un análisis un poco más pormenorizado de los datos, observamos que el año 2018/2019 (totalmente presencial) es el año donde el número de no presentados ha sido más elevado, por el contrario, los años siguientes donde había actividades online (bien sea de evaluación como de docencia) suponen una fuerte caída en el número de No Presentado. No obstante, esta fuerte caída no se ha transformado en un mayor aumento del número de personas que superan la materia (con excepción del año 2019/2020, donde toda la materia funcionó online). Así vemos que en los años 2020/2021 y 2021/2021 presentan los valores más altos de personas que No Superan la materia, acompañado de los menores valores de personas que la Superan o, como ya comentamos, No Presentado.

Podríamos pensar en que los buenos resultados del año 2019/2020 llevaron a una mayor tasa de participación en la materia, aunque por desgracia, no ha mejorado los resultados.

En la Figura 4 podemos ver estas diferencias. Los colores azules representan que los valores observados son mayores que los valores esperados si los datos fuesen obtenidos de forma aleatoria. El color rojo indica que dichos valores son menores que los esperados. El año 19/20 presenta valores mayores en la categoría “aprobados” y menores en la categoría “suspense”. Por su parte, el año 21/22 muestra valores mayores en la categoría “suspense” y mucho menores en la categoría “no presentado”, esta misma situación, aunque con diferencias menores ya se había producido en el año 20/21. En general, el número de aprobados se ha mantenido en 3 de los años analizados, y sólo ha tenido diferencias positivas destacables en el año 19/20 que como ya hemos comentado es el año donde toda la actividad docente fue online. En los años posteriores, 20/21 y 21/22, el primero con un sistema semipresencial de docencia y evaluación y el segundo, donde toda la evaluación vuelve a ser presencial y se mantiene una parte pequeña de docencia online, muestra un mayor número de suspensos que en el año 18/19, y un menor número de “no presentado”.

Por último, hemos realizado un análisis de tablas de contingencia. Mediante el test de Chi-cuadrado (Yates, 1934) examinamos si las filas y columnas de la tabla de resultados están significativamente asociadas. Los datos muestran que las filas y columnas de la tabla de resultados están significativamente asociadas (Tabla 6).

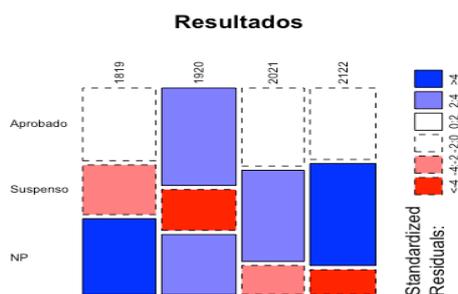


Figura 4: Mosaico de datos de los Resultados de Matemáticas II

	Chi-cuadrado	p-valor	Grados de libertad
Test de Contingencia	161	2.2e-16	6
Valores esperados	Superada	No superada	No presentado
2018/2019	166.58	143.46	100.96
2019/2020	169.01	145.56	102.43
2020/2021	141.86	122.17	85.97
2021/2022	149.56	128.80	90.64

Tabla 6: Análisis de Tablas de Contingencia de Chi-cuadrado

La Figura 5 muestra la contribución de cada celda al valor total de la Chi-cuadrado a través de los residuos de Pearson. En la gráfica de la izquierda se muestran la contribución de cada categoría al valor de la Chi-cuadrado en términos absolutos. Los residuos positivos aparecen en azul y los negativos en rojo. Por su parte, la gráfica de la izquierda muestra la contribución relativa. El tamaño del círculo es proporcional a la contribución.

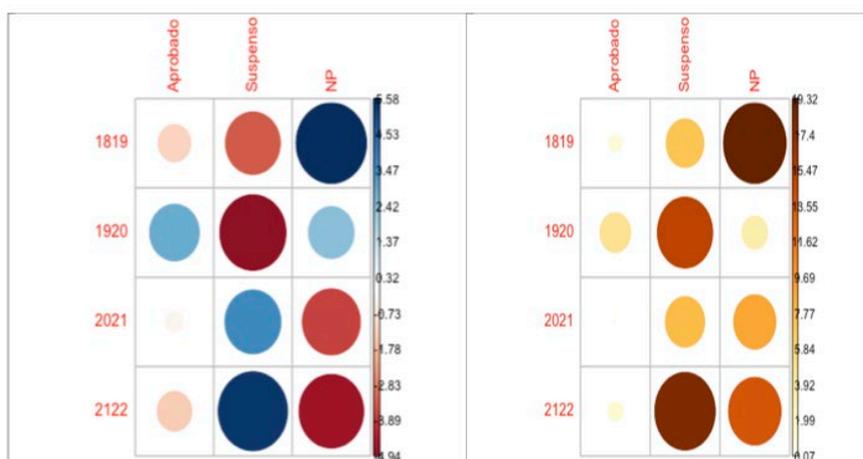


Figura 5: Análisis de Contingencia Resultados de Matemáticas II

Los datos muestran que la columna “aprobada” está fuertemente asociada con el año 2019/2020, mientras que la columna “suspensa” se asocia al año 2021/2022. Por su parte la columna “No presentado” presenta dos situaciones claras: una asociación significativa con el año 2018/2019 (positiva) y con el año 2021/2022 (negativa). Por su parte las mayores contribuciones al valor de la Chi-cuadrado vienen dados por la categoría: No Presentado-18/19 (19,32%), Suspense-21/22 (18,05%), Suspense-19/20 (15.13%), NP-

21/22 (13,65%). Por tanto, los datos muestran que el número de personas que no superan la materia es especialmente significativo en el año 2021/2022, igual que el número de “no presentados”, que en dicho año es muy bajo. Por el contrario, en el año 2018/2019 el número de “no presentados” es especialmente alto.

4. Conclusiones

En este trabajo hemos comparado la evolución de las calificaciones en la materia de Matemáticas II para ADE y el programa de simultaneidad de ADE + Derecho en los años académicos 2018/2019, 2019/2020, 2020/2021 y 2021/2022. En estos años, por la situación sanitaria mundial, nos hemos visto obligados a utilizar estrategias docentes y evaluadoras. El año 2018/2019 utilizamos los medios habituales de docencia y evaluación presencial de una asignatura impartida en una universidad presencial. El año 2019/2020, el confinamiento afectó a casi todo el curso académico y obligó a que la mayor parte de la docencia y todo el proceso evaluador fuese realizado de forma telemática (docencia online). El año 2020/2021 se realizó con un proceso mixto donde se combinaron métodos online y presenciales. Finalmente, el año 2021/2022 se restauró el proceso de evaluación presencial (similar al año 2018/2019) y aunque la mayor parte de la docencia fue presencial también se mantuvo una pequeña proporción de docencia online.

Los datos muestran que existen diferencias significativas entre los resultados obtenidos en los diferentes años. Las calificaciones medias fueron más altas durante el año 2019/2020, pero el año 2020/2021 registraron valores menores, inclusive, al año con docencia totalmente presencial. Por otro lado, en el año 2020/2021 se produjo un fenómeno que se intuía en los años anteriores: la distribución de las calificaciones muestra una estructura bimodal, habiendo dos puntos de mayor concentración de las calificaciones, en torno al valor de 5 y en torno al valor de 2. Estos mismos datos, con mayor grado de profundidad se producen también en el año 2021/2022.

En cuanto a los resultados obtenidos podemos resaltar los siguientes puntos: El año 2018/2019 (presencial) se caracteriza por un valor medio alto de personas que Superan la materia, un valor medio bajo de personas que “No Superan” y un valor muy alto de personas que optaban por no participar de forma activa en la materia, lo que conlleva el valor más alto de “No Presentado” de la serie. Durante el año de evaluación online aumenta de forma importante el número de personas que superan la materia, pero la reducción en el número de No Presentados es pequeña. Por su parte, el año con docencia semipresencial, presenta valores muy altos en el número de personas que No Superan la materia y una fuerte reducción en el número de No Presentados, el menor valor de la serie. Estos datos sugieren que el alto número de aprobados del año 2019/2020, actuó como una especie de “efecto llamada” para que aumentase de forma importante el número de presentados a las pruebas en el año siguiente; pero esto no redundó en mejores resultados, de hecho, supuso un fuerte aumento del número de personas que no superan la materia. Por último, el año 2021/2022, donde casi se volvió a la normalidad del año 2018/2019, presenta una repetición aumentada de los datos del año 2020/2021, continua la reducción del número de no presentados, pero a costa de un aumento del número de suspensos frente al número de aprobados.

En el futuro se abren dos líneas de investigación a seguir: estudiaremos cómo evoluciona esta tendencia utilizando para ello los resultados del año 2022/2023, donde ya se ha retomado de forma completa la docencia y evaluación presencial, tal y como se realizaba en 2018/2019 y, al mismo tiempo, ver si el efecto postpandemia puede haber influido en los resultados.

Referencias

Alacid, V., Caballero, M. D., y Lafuente, M. (2017). Aproximación al nivel de destreza matemática de los alumnos de primer curso del grado en ADE. *Anales de ASEPUMA*, 25, 1-21.

- Arribas, J. M. (2012). El rendimiento académico en función del sistema de evaluación empleado. *RELIEVE. Revista Electrónica de Investigación y Evaluación Educativa*, 18(1), 1-15. <https://doi.org/10.7203/relieve.18.1.2974>
- Barberà, E. (2016). Aportaciones de la tecnología a la e-Evaluación. *RED. Revista de Educación a Distancia*, 50(4), 1-10. <http://dx.doi.org/10.6018/red/50/4>
- Boncompte, M., Sarrasí, F. J., Bermúdez, L., Castañer, A., Marín-Solano, J., Navas, J., y Pons, M. A. (2010). Estudio del perfil del alumnado de primer curso en la Facultad de Economía y Empresa de la Universidad de Barcelona. En: Vicerreitoria de Formació e Innovació Educativa (Eds.), *La docencia en el nuevo Escenario del Espacio Europeo de Educación Superior* (pp. 357-360). Colección Formación e Innovación Educativa na Universidade.
- Choonpradub, C., y McNeil, D. (2005). Can the box plot be improved? *Songklanakarin Journal of Science and Technology*, 27(3), 649-657.
- Dean, A., y Voss, D. (1999). *Design and analysis of experiments*. Springer-Verlag.
- Escudero, T. (2003). Desde los Tests hasta la investigación evaluativa actual. Un siglo, el XX, de intenso desarrollo en la evaluación en educación. *RELIEVE. Revista Electrónica de Investigación y Evaluación Educativa*, 9(1), 11-43. <https://doi.org/10.7203/relieve.9.1.4348>
- Esparza, R., y Márquez, D. (2020). Implicacions dels coneixements de matemàtiques en els estudiants de nou accés a graus universitaris d'economia i empresa. *RIDU. Revista d'Innovació Docent Universitària*, 12, 95-117. <https://doi.org/10.1344/RIDU2020.12.10>
- Fraille, A., López-Pastor, V., Castejón, J., y Romero, R. (2013). La evaluación formativa en docencia universitaria y el rendimiento académico del alumnado. *Aula Abierta*, 41(2), 23-34. <http://hdl.handle.net/11162/97223>
- García, J., Pérez, M. A., Rodríguez, B., y Sánchez, M. C. (2002). Evaluar en la red. *Revista de Teoría de la Educación. Educación y cultura en la sociedad de la información*, 3. <http://hdl.handle.net/11162/81572>
- Gómez-Déniz, E., García-Artiles, M. D., y Dávila-Cárdenas, N. (2015). Estudio de los factores determinantes de las notas de matemáticas empresariales. *Anales de ASEPUMA*, 23, 1-18. <http://hdl.handle.net/10553/56677>
- Kampstra, P. (2008). Beanplot: A boxplot alternative for visual comparison of distributions. *Journal of Statistical Software*, 28(1), 1-9. <https://doi.org/10.18637/jss.v028.c01>
- Kaspriková, N. (2012). Statistical evaluation of examination Tests in mathematics for economists. *Journal on Efficiency and Responsibility in Education and Science*, 5(4), 203-211. <https://doi.org/10.7160/eriesj.2012.050403>
- Kirtman, L. (2009). Online versus In-Class courses: An examination of differences in learning outcomes. *Issues in Teacher Education*, 18(2), 103-116.
- Klůfa, J. (2021). Evaluation of distance exams in mathematics. *The 15th International Days of Statistics and Economics*, Prague, September 9-11, 499-508.
- Kruskal, W., y Wallis, W. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260), 583-621. <https://doi.org/10.2307/2280779>
- Levene, H. (1960). Robust tests for equality of variances. En: Olkin, I, Hotelling, H, et al. (Eds.), *Contributions to Probability and Statistics: Essays in Honour of Harold Hotelling* (pp. 278-292), Stanford University Press.
- Martínez, E. Z. (2015). Description of continuous data using bar graphs: a misleading approach. *Revista da Sociedade Brasileira de Medicina Tropical*, 48(4), 494-497. <http://dx.doi.org/10.1590/0037-8682-0013-2015>

- Matusek, V., y Greganova, R. H. (2019). Comparison of exam results in Mathematics at Faculty of Economics and Management, Slovak University of Agriculture in Nitra. *MERAA. Mathematics in Education, Research and Applications*, 5(2), 78-83. <https://doi.org/10.15414/meraa.2019.05.02.78-83>
- McDonald, J. H. (2014). *Handbook of Biological Statistics*. Sparky House Publishing.
- Rane, V., y MacKenzie, C. A. (2020). Evaluating students with online testing modules in engineering economics: A comparison of student performance with online testing and with traditional assessments. *The Engineering Economist*, 65(3), 213-235. <https://doi.org/10.1080/0013791X.2020.1784336>
- Rapanta, C., Botturi, L., Goodyear, P., Guàrdia, L., y Koole, M. (2020). Online University Teaching during and after the Covid-19 Crisis: Refocusing teacher presence and learning activity. *Postdigital Science and Education*, 2, 923-945. <https://doi.org/10.1007/s42438-020-00155-y>
- Rodríguez, M. J. (2008). Aplicación de las TIC a la evaluación de alumnos universitarios. *Revista de Teoría de la Educación. Educación y cultura en la sociedad de la información*, 6(2). <https://doi.org/10.14201/eks.18185>
- Sánchez, M. P. (Coord.) (2010). *Técnicas docentes y sistemas de Evaluación en Educación Superior*. Narcea.
- Shapiro, S. S., y Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3-4), 591-611. <https://doi.org/10.2307/2333709>
- Vallés, C., Ureña, N., y Ruiz, E. (2011). La evaluación formativa en docencia universitaria. Resultados globales de 41 estudios de caso. *REDU. Revista de Docencia Universitaria*, 9(1), 135-158. <https://doi.org/10.4995/redu.2011.6184>
- Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics*, 1(6), 80-83. <https://doi.org/10.2307/3001968>
- Yates, F. (1934). Contingency tables involving small numbers and the chi-squared test. *Supplement to the Journal of the Royal Statistical Society*, 1(2), 217-235. <https://doi.org/10.2307/2983604>
- Yates, R. W. y Beaudrie, B. (2009). The impact of online assessment on grades in Community College Distance Education Mathematics Courses. *The American Journal of Distance Education*, 23(2), 62-70. <https://doi.org/10.1080/08923640902850601>
- Zaragoza, J., Luis-Pascual, J. C., y Manrique, J. C. (2009). Experiencias de innovación en docencia universitaria: resultados de la aplicación de evaluación formativa. *REDU. Revista de Docencia Universitaria*, 7(4). 1-33. <https://doi.org/10.4995/redu.2009.6232>