



FACULTAD
DE CIENCIAS
ECONÓMICAS



Universidad
Nacional
de Córdoba

REPOSITORIO DIGITAL UNIVERSITARIO (RDU-UNC)

Uso de indicadores psicométricos para la optimización de cuestionarios. Experiencia en la evaluación de los estudiantes de Estadística I

Norma Patricia Caro, María Inés Ahumada, Verónica Arias

Ponencia presentada en Memorias de las 1as. Jornadas Virtuales de Aulas Abiertas y 2as.
Jornadas de Aulas Abiertas realizado en 2019 en la Facultad de Ciencias Económicas -
Universidad Nacional de Córdoba. Córdoba, Argentina



Esta obra está bajo una [Licencia Creative Commons Atribución – No Comercial – Sin Obra
Derivada 4.0 Internacional](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Uso de indicadores psicométricos para la optimización de cuestionarios. Experiencia en la evaluación de los estudiantes de Estadística I

Eje temático: Relato de Experiencia – Evaluación

Materia / Comisión de referencia: Estadística I

Norma Patricia Caro; María Inés Ahumada y Verónica Arias
pacaro@eco.unc.edu.ar; ahumada.mi@gmail.com; veroarias1@gmail.com

Facultad de Ciencias Económicas de la Universidad Nacional de Córdoba

RESUMEN:

La evaluación de los estudiantes es un aspecto importante en el proceso educativo y el contexto cambiante en el que éste último se desarrolla motiva a los docentes a buscar estrategias innovadoras para hacer más eficiente dicha tarea.

A partir del año 2016 se comenzó a evaluar a los alumnos desde la virtualidad, haciendo uso intensivo de los cuestionarios Moodle. La finalidad de tal innovación fue complementar las evaluaciones escritas tradicionales con otras instancias evaluativas intermedias. Los cuestionarios utilizados son analizados periódicamente en base a las opiniones de los alumnos y a los indicadores psicométricos brindados por la propia plataforma Moodle.

El objetivo del presente trabajo es compartir la experiencia, tanto de la implementación de instancias evaluativas intermedias en un contexto de masividad, como de análisis y mejora de los cuestionarios mediante los indicadores provistos por la propia herramienta.

Se logra, con esta experiencia, un impacto positivo, motivando a los alumnos a dar mayor continuidad al proceso de aprendizaje y obteniendo mejoras cualitativas en sus rendimientos finales. Asimismo, los indicadores psicométricos reportados por el módulo de estadísticas de Moodle permiten detectar fácilmente las debilidades del instrumento, que derivan en múltiples estrategias para realizar ajustes e incrementar la calidad de este.

Palabras Clave: Cuestionarios – Fiabilidad – Evaluación formativa – Estadística – Moodle

1. CONTEXTO

La experiencia de enseñanza-aprendizaje que se presenta, fue desarrollada dentro de los cursos de Estadística previstos en el Ciclo de Formación Básica Común del actual plan de estudios de La Facultad de Ciencias Económicas de la Universidad Nacional de Córdoba. Se trata de la asignatura Estadística I, que cuenta con 250 estudiantes en promedio, por año y abarca conceptos básicos de Estadística Descriptiva, Introducción a la Probabilidad, Modelos de probabilidad y Distribuciones en el muestreo. La Estadística es una de las disciplinas que puede ser aplicada al estudio metódico de problemas de los más diversos campos de las ciencias económicas. Se apoya en el uso de software estadístico, indispensable para el manejo de bases de datos.

A partir del año 2016 se planificó incorporar cuatro instancias de evaluaciones individuales, presentadas a los alumnos como “Autoevaluaciones” no obligatorias, con recompensa de puntaje, que se adiciona a la nota de cada prueba parcial. Estas instancias se pensaron también como herramientas para motivar al alumno a apropiarse de los contenidos, de manera paulatina a lo largo del cursado y evitar que el momento dedicado al estudio se acote a los días previos a cada parcial.

2. DESCRIPCIÓN DE LA EXPERIENCIA

El equipo docente recorrió las siguientes etapas para la elaboración de un instrumento de evaluación: Preparación, Aplicación, Valoración y Devolución. El desafío fue repensar esas instancias pero en el contexto de la virtualidad. La etapa de “Preparación” implicó investigar en las alternativas de construcción del “Banco de Preguntas” en la plataforma Moodle, definir la creación de las categorías que agrupen las preguntas conforme a los contenidos de la asignatura. Es importante la organización del banco de preguntas porque permite la reutilización de las mismas. Se exploró también la posibilidad de crear cuestionarios a partir de una mezcla aleatoria de preguntas de diferentes categorías, dando como resultado exámenes diferentes sobre un mismo conjunto de temas predefinidos. La etapa de “Aplicación” del instrumento, requiere también prever aspectos operativos vinculados a decidir la manera en la que el alumno responderá el cuestionario. Por la cantidad de alumnos que cursan habitualmente la asignatura y la escasez de computadoras disponibles dada la masividad, se realiza una inscripción previa de aquellos que deseaban responder el cuestionario. Dado el carácter de “evaluación sumativa” otorgado a estos cuestionarios, los alumnos realizaron la actividad en aulas informáticas de la Facultad, de manera individual. Se restringió la posibilidad de responderlo, únicamente en la Facultad, aprovechando las medidas de seguridad informática del apartado de restricciones extras para los intentos que tiene Moodle. Las etapas de valoración y devolución son amplias y fácilmente aplicables en los cuestionarios de Moodle, ya que cuenta con herramientas para diseñar la retroalimentación que puede ser global para cada ítem o individual para cada respuesta posible al ítem. Por ejemplo, ante una pregunta con cinco alternativas de respuesta (sólo una correcta), es posible incorporar un comentario global sobre el tema que se evalúa en el ítem o pregunta que será visualizado por el alumno, independientemente de la alternativa que seleccionó. El esmero en completar de manera clara esta información representa un aporte sumamente valorado por el estudiante, porque le permite comprender mejor cuál fue su error, o bien reforzar el concepto o contenido del ítem cuando responde correctamente la pregunta. El acceso a esta devolución, sobre el resultado del cuestionario respondido, se habilita inmediatamente después de que el alumno lo finalice.

Por último, luego de concluida la evaluación, se puede volver a consultar en cualquier momento el cuestionario con el objetivo de “aprender del error” y prepararse mejor para la instancia del parcial obligatorio para aprobar el curso.

3. RESULTADOS

La implementación de los cuestionarios llevó a una segunda etapa, que fue medir la eficiencia del instrumento de evaluación, es decir medir qué tan confiables resultaron y si sirven realmente para evaluar y discriminar a los estudiantes en términos de su aprendizaje.

3.1 Revisión del cuestionario

Las sucesivas aplicaciones de los cuestionarios permiten avanzar sobre nuevas etapas de revisión, apelando a las herramientas que proporciona la propia plataforma Moodle, en este caso se usó el reporte de estadísticas del cuestionario. A través del cálculo de indicadores psicométricos se evalúa el cuestionario en dos aspectos. Uno de ellos está referido a características del mismo en su conjunto, como instrumento de evaluación y el otro aspecto apunta a analizar individualmente cada una de las preguntas que lo conforman. A través de la interpretación de estos indicadores psicométricos que se presentan en el “Reporte de estadísticas de examen”, se pretende aprovechar el potencial de este recurso, proponiendo ajustes de diseño fundamentados en análisis cuantitativos objetivos.

En los Cuadros 1 y 2 se presentan los resultados para cada ítem o pregunta para un cuestionario del año 2016 y uno mejorado del año 2017, respectivamente, resaltando los valores no adecuados en cada caso. Estos reportes son generados de manera automática para cada cuestionario que haya sido respondido por un grupo de alumnos en una misma instancia. En esta experiencia sólo tienen la posibilidad de resolverlo una vez (por el carácter de evaluación sumativa). En el reporte se muestra una descripción del comportamiento individual de cada ítem, en el contexto global del cuestionario. Luego, analizando la valoración de comportamiento del ítem, es posible identificar una posible acción a realizar en cada pregunta con el objetivo de mejorar la calidad del cuestionario.

Entre los indicadores más relevantes, se mencionan:

El **Índice de Facilidad o índice de Nivel de Dificultad** es la puntuación promedio de los alumnos en el ítem, es igual al porcentaje de aprobación de los ítems sobre el total de intentos del mismo. Los índices de facilidad entre 30% y 70% suelen aportar diferencias importantes entre el nivel de conocimiento, habilidad y preparación entre los alumnos.

El **Índice de discriminación y Eficiencia de discriminación** mide la correlación entre las calificaciones ponderadas en la pregunta y las del resto del examen. Revela qué tan efectivo es el ítem para clasificar o separar a los alumnos que poseen un mayor grado de conocimiento de los que poseen un grado menor. Para el índice de discriminación los valores adecuados van entre el 30% y el 50%, y si son mayores al 50% son muy buenos. Para la eficiencia discriminativa, más del 50% es un valor óptimo.

Cuadro 1: Comportamiento individual de cada ítem- Cuestionario 2016

Ítem/Pregunta	Índice de dificultad (30% - 70%)	Índice de discriminación (más del 30%)	Eficiencia discriminativa (más del 50%)
1	25,40	13,71	20,29
2	28,57	16,20	21,96
3	7,94	7,92	16,79
4	84,13	41,60	60,95
5	47,62	49,83	60,30
6	44,44	15,07	18,36
7	52,38	15,45	17,75
8	71,43	32,54	40,91
9	26,98	2,49	3,57
10	42,86	43,32	54,63
11	39,68	24,95	32,01

Por ejemplo, para el ítem N°4 se podría revisar y aumentar el nivel de dificultad,-reformulando el enunciado o por medio de las alternativas propuestas. Notar que preguntas como el ítem N°3, que resultó con bajo nivel de dificultad y al mismo tiempo con baja eficiencia discriminativa, también debería revisarse. Realmente las alternativas para optimizar cada ítem son amplias y requieren de cierta dedicación para que efectivamente los ajustes redunden en mejoras de los indicadores psicométricos.

A continuación se muestran los resultados obtenidos a partir de los ajustes aplicados a los ítems del cuestionario anterior, que se refieren a una mejor formulación de las preguntas.

Cuadro 2: Comportamiento individual de cada ítem- Cuestionario 2017

Ítem/Pregunta	Índice de dificultad (30% - 70%)	Índice de discriminación (más del 30%)	Eficiencia discriminativa (más del 50%)
1	65,63	63,83	67,86
2	53,13	57,34	69,24
3	59,38	77,49	88,61
4	62,5	82,82	91,06
5	40,63	51,96	69,84
6	46,88	58,81	74,83
7	66,67	91,75	96,99
8	15,63	35,01	73,70
9	59,38	81,98	93,16
10	65,63	88,91	95,52

3.2 Estadísticas globales del cuestionario

El reporte estadístico incluye también información general del cuestionario, en el que se hace una valoración global en la que influyen aspectos como la cantidad de alumnos que respondieron, la extensión del instrumento en cuanto a la cantidad de ítems que lo conforman e incluso la ponderación que se le asigna a los ítems al realizar el diseño del mismo. En el Cuadro 3 se muestran de manera comparativa, los resultados de ambos años.

Cuadro 3: Información general sobre el cuestionario - Años 206 y 2017

Nombre del cuestionario	Grupo 3 Año 2016	Grupo 3 Año 2017	Indicador de Referencia
Nombre del curso	Estadística I	Estadística I	
Número de primeros intentos	63	29	
Número total de intentos	63	32	
Promedio de los primeros intentos	42,86%	54%	
Promedio de todos los intentos	42,86%	54%	promedio de 50% a 75%
Calificación media de los últimos intentos	42,86%	57%	
Calificación media de los mejores intentos	42,86%	57%	
Calificación media (de todos los intentos)	45,45%	70%	
Desviación estándar (para todos los intentos)	19,55%	36%	
Asimetría de la distribución de puntuaciones (para todos los intentos)	-0,2388	-0.56	
Curtosis de la distribución de puntuaciones (para todos los intentos)	-0,9296	-1.36	
Coefficiente de consistencia interna (para todos los intentos)	56,10%	92%	>80%
Ratio de error (para todos los intentos)	66,26%	29%	<50%
Error estándar (para todos los intentos)	12,96%	11%	entre 5% y 8%

A continuación, se interpretan los principales estadísticos obtenidos:

El **promedio de intentos** fue del 42,86% y 54% respectivamente. Este valor es la media aritmética de las calificaciones obtenidas por los alumnos que finalizaron el cuestionario. Los valores fuera de los límites de referencia sugieren repensar la evaluación (Valores de referencia: 50% - 75%).

La **calificación media** (de todos los intentos) fue del 45,45% en 2016 y 70% en 2017. Bajo este título Moodle reporta el valor de la Mediana de las calificaciones. Es decir, es el valor central de la distribución. Entonces su interpretación es que la mitad de los alumnos obtuvieron una calificación inferior al 70% y el 50% restante, alcanzaron una calificación mayor que 70%, por ejemplo para el valor de 2017.

La **asimetría** de la distribución de puntuaciones (para todos los intentos), se trata de una medida de la asimetría de la distribución de calificaciones. Los valores positivos indican presencias de unos pocos valores de calificaciones altas y los coeficientes negativos indican presencia de pocos valores bajos y calificaciones concentradas en valores altos de la escala. En esta oportunidad el valor es de -0,23 y -0,56.

La **curtosis** de la distribución de puntuaciones (para todos los intentos). La Curtosis (0,92 y -1,36) es una medida de forma e indica que tan aplanada es la distribución. Un coeficiente de curtosis mayor que 1 puede indicar que el cuestionario no está discriminando muy bien entre los alumnos muy buenos (o los muy malos) y aquellos que se ubican en el promedio.

El **coeficiente de consistencia interna** Alpha de Cronbach (para todos los intentos). Es deseable alcanzar valores mayores que 0,70 (Acock, 2013), ya que un valor bajo indica, o bien que algunas de las preguntas no son muy buenas para discriminar entre alumnos con diferentes niveles de conocimiento del tema que se evalúa y por esto las diferencias entre las puntuaciones totales están más bien asociadas al azar; o bien puede revelar que algunas de las preguntas están funcionando con una calidad diferente al restos de los ítems, provocando que el examen en su conjunto no sea homogéneo. Se acepta que si el valor del coeficiente es inferior a 64%, el examen completo es

insatisfactorio, y en nuestro caso se obtuvo, en la primera ocasión, en 2016, un 56,1%. Luego de revisar el instrumento y de corregirlo, se llega al 92%, al año siguiente.

El **Ratio de error** (para todos los intentos) arroja un valor del 66,26% en la primera ocasión y luego disminuye al 29%, en la segunda. La tasa de error debe ser baja y se debe a efectos aleatorios en lugar de diferencias auténticas en la habilidad entre los estudiantes.

3.3 Valoración de los cuestionarios por parte de los alumnos

También se invitó a los estudiantes a manifestar los aspectos positivos y negativos de los cuestionarios, a través de una encuesta, lo que se ve reflejado en las figuras 1 y 2.

Figura 1. Palabras representativas de aspectos positivos y principales frases

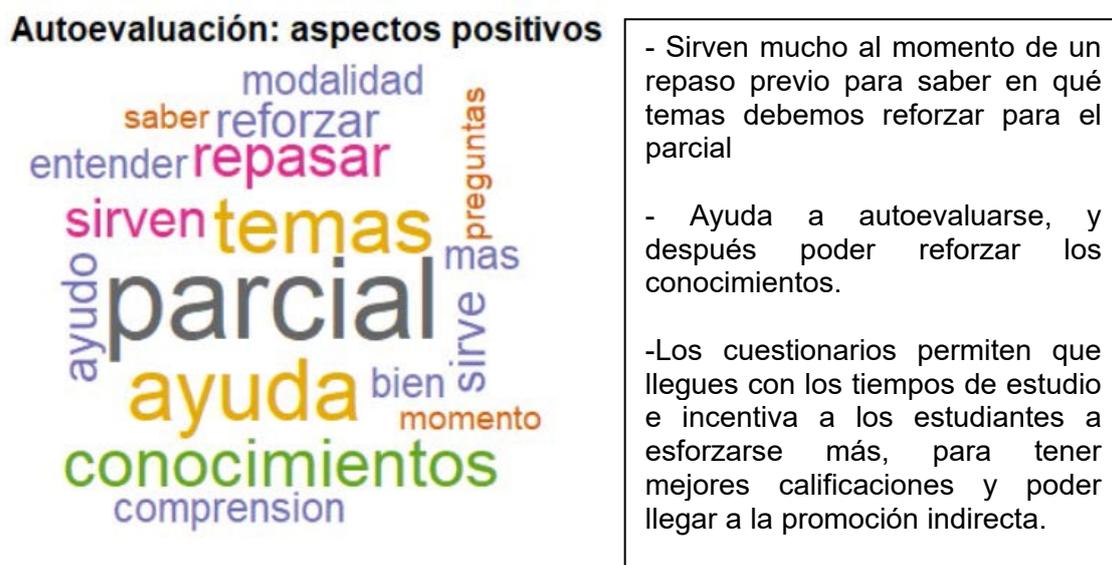
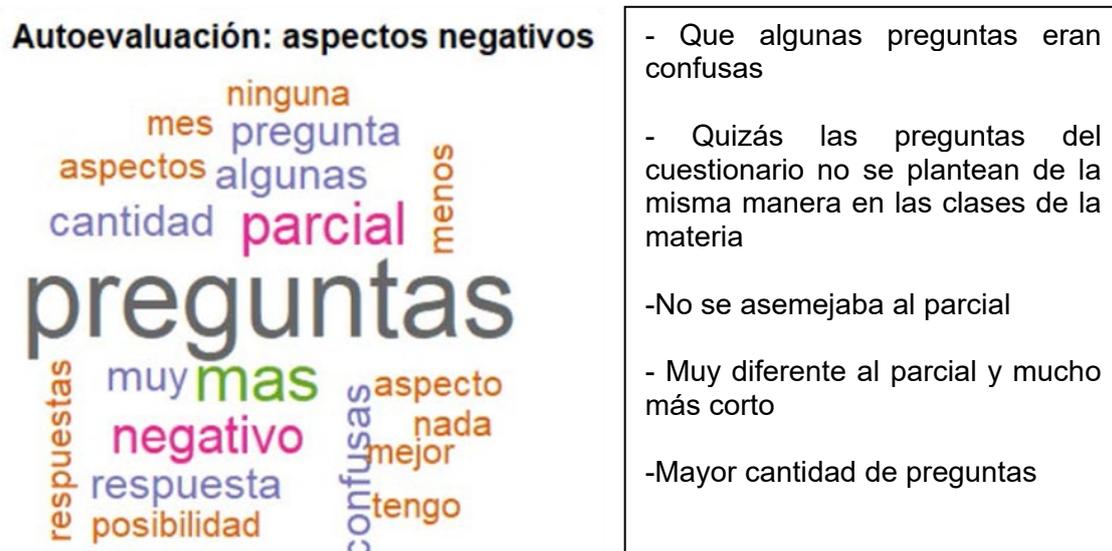


Figura 2. Palabras representativas de aspectos negativos y principales frases



4. CONCLUSIONES Y SUGERENCIAS

Se encontraron evidencias sobre el impacto en el proceso de aprendizaje que tiene la aplicación de los cuestionarios de Moodle debido a que los estudiantes manifestaron satisfacción en la realización de dichas actividades previas a las evaluaciones parciales, lo que motiva a continuar mejorando la propuesta de enseñanza y medir los resultados en los exámenes finales. Si los resultados de estas evaluaciones no son satisfactorias, tal vez no se deba sólo a conocimientos no aprendidos sino a un instrumento no confiable para evaluarlos. Dentro de las buenas prácticas de enseñanza está el mejoramiento y el diseño de los cuestionarios para evaluar en la virtualidad, considerando un análisis de medidas psicométricas similares a las presentadas en este trabajo para fundamentar objetivamente tales mejoras. Se logró aprovechar el enorme potencial latente en los cuestionarios que se presentan a los alumnos como una instancia de autoevaluación, en el sentido de que se espera que sean un reflejo de cómo se están preparando para las instancias de evaluación y de esta manera aprender la asignatura.

5. REFERENCIAS

- Blanco, M., & Ginovart, M. (2012). Los cuestionarios del entorno Moodle: su contribución a la evaluación virtual formativa de los alumnos de matemáticas de primer año de las titulaciones de Ingeniería. *RUSC. Universities and Knowledge Society Journal*, 9(1).
- Brink, R., & Lautenbach, G. (2011). Electronic assessment in higher Education. *Educational Studies*, 37(5), 503-512.
- Crews, T. B., & Curtis, D. F. (2011). Online course evaluations: Faculty perspective and strategies for improved response rates. *Assessment & Evaluation in Higher Education*, 36(7), 865-878.
- Daly, C., Pachler, N., Mor, Y., & Mellar, H. (2010). Exploring formative e-assessment: using case stories and design patterns. *Assessment & Evaluation in Higher Education*, 35(5), 619-636.
- Delgado García, A. M., & Oliver Cuello, R. (2006). La evaluación continua en un nuevo escenario docente. *RUSC. Universities and Knowledge Society Journal*, 3(1).
- Ferrão, M. (2010). E-assessment within the Bologna paradigm: evidence from Portugal. *Assessment & Evaluation in Higher Education*, 35(7), 819-830.
- Graff, M. (2003). Cognitive style and attitudes towards using online learning and assessment methods. *Electronic Journal of e-learning*, 1(1), 21-28.
- Iglesias Rodríguez, A., Olmos Migueláñez, S., Torrecilla Sánchez, E. M., & Mena Marcos, J. J. (2014). Evaluar para optimizar el uso de la plataforma moodle (studium) en el departamento de didáctica, organización y métodos de investigación. *Tendencias pedagógicas*.
- Llorente Cejudo, M. (2007). Moodle como entorno virtual de formación al alcance de todos. *Comunicar*, 15(28).
- Martín Galán, B., & Rodríguez Mateos, D. (2012). La evaluación de la formación universitaria semipresencial y en línea en el contexto del EEES mediante el uso de los informes de actividad de la plataforma Moodle. *RIED. Revista iberoamericana de educación a distancia*, 15(1).
- Stegmann, C., Huertas, M. A., Juan, Á. A., & Prat, M. (2008). E-learning de las asignaturas del ámbito matemático-estadístico en las universidades españolas: oportunidades, retos, estado actual y tendencias. *RUSC. Universities and Knowledge Society Journal*, 5(2), 1-14.
- San Martín, E. (2004). Elementos de Psicometría- Teoría de Medición, Teoría Clásica de Test y teoría de Respuesta al Ítem. *CLATSE VI Sexto Congreso Latinoamericano de Sociedades Estadísticas. Departamento de Estadística - Pontificia Universidad Católica de Chile*.