



FACULTAD  
DE CIENCIAS  
ECONÓMICAS



Universidad  
Nacional  
de Córdoba

# REPOSITORIO DIGITAL UNIVERSITARIO (RDU-UNC)

## Evaluando a los estudiantes de Estadística con cuestionarios del entorno Moodle

Norma Patricia Caro, María Inés Ahumada

Ponencia presentada en XXXII Jornadas Nacionales de Docentes de Matemática de Facultades de Ciencias Económicas y afines realizado en 2017 en la Universidad Nacional de Entre Ríos.  
Paraná, Entre Ríos, 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/)

## Evaluando a los estudiantes de Estadística con Cuestionarios del entorno Moodle

CARO, Norma Patricia - AHUMADA, María Inés  
Facultad de Ciencias Económicas, Universidad Nacional de Córdoba  
[npatriciacar@gmail.com](mailto:npatriciacar@gmail.com) - [ahumada.mi@gmail.com](mailto:ahumada.mi@gmail.com)

**Especialidad:** Educación Matemática

**Palabras Clave:** Cuestionarios, Moodle, Evaluación formativa, Estadística, Análisis psicométrico

### Resumen

La Evaluación de los estudiantes es un aspecto importante en el proceso educativo y el contexto cambiante en el que se desarrolla tal proceso motiva a los docentes a buscar estrategias innovadoras para mejorar esta tarea. En ese marco el módulo de cuestionarios disponible en el entorno Moodle representa una herramienta muy potente en comparación con otras metodologías de evaluación tradicionales.

Desde 2005, en un curso presencial de Estadística I de las carreras de la Facultad de Ciencias Económicas-UNC, se fueron combinando recursos del entorno virtual, tornando la propuesta de enseñanza-aprendizaje en una experiencia de aplicación de la metodología Blended-learning. Un recurso de alto impacto fue el uso de software para procesamientos de bases de datos reales. La implementación requería contar con un Aula Virtual para proveer al alumno todo el material requerido (Tutoriales, bases de datos, etc.)

Cada año se intentaba medir el impacto de la combinación de elementos de la enseñanza virtual, en cuanto a su contribución al logro aprendizajes significativos, pero no se conseguía detectar una mejora significativa concreta. En 2015 se migró el contenido del AV al entorno Moodle, permitiendo en 2016 realizar la primera implementación de instancias de evaluación de tipo sumativas durante el cursado de la asignatura. Los resultados obtenidos revelan mejoras significativas en los principales indicadores, en comparación con los anteriores. En virtud esta primera experiencia, se continúa trabajando para mejorar y ajustar los instrumentos de evaluación de Moodle, basados en el análisis de los propios Reportes de examen que provee la plataforma.

### 1 Introducción

El Blended Learning que, en un contexto de enseñanza virtual, se traduce como formación combinada o enseñanza mixta, se trata de una modalidad de estudios que incluye formación no presencial y presencial.

Con el tiempo se volvió más relevante la selección y aplicación de las TICs, destacando que el equipo docente en los primeros años trabajó intensamente en la elección del software estadístico y en la elaboración de los materiales a entregar a los alumnos para emplearlo (Tutoriales, Instructivos y bases de datos). De manera simultánea se avanzó sobre la conformación del Aula Virtual, en una plataforma propietaria, la cual permitió desplegar la propuesta de enseñanza-aprendizaje. Cada año se exploraban, se alternaban nuevas herramientas disponibles en la plataforma: Noticias, Foros, Encuestas, publicación de archivos, entre otras.

En 2015 se reemplazó la plataforma del Entorno Virtual, migrando el Aula una plataforma Moodle. En 2016 se incorporaron algunos cambios en las técnicas de evaluación, sostenidos en el enfoque de una adecuada evaluación formativa, acorde al contexto institucional en el que se desarrolla la asignatura y con el objetivo de incidir en una mejora del aprendizaje que pudiera ser significativamente tangible.

La evaluación de los alumnos es un tema sensible y su finalidad es medir los logros en los resultados de aprendizaje previsto, lo que está relacionado con procedimientos idóneos que deberían seguirse en los procesos evaluativos (Blanco y Ginovart, 2012). La necesidad de generar dispositivos de evaluación confiables constituye una ventaja estratégica para cualquier docente que desee incorporar tecnologías de la información y comunicación (TIC) en su práctica académica. Asimismo, la gestión en tiempo real de todos los resultados es un resultado que contribuye en la práctica docente.

Blanco y Ginovart (2012) plantean *que varios estudios han puesto de relieve el papel cada vez más importante de las tecnologías de la información y la comunicación en el campo de la evaluación (Delgado y Oliver, 2006; Graff, 2003, Steegmann et al, 2008), hasta el punto de que la evaluación virtual ha pasado a ser un nuevo elemento del ámbito educativo (Brinck y Lautenbach, 2011; Crews y Curtis, 2011; Daly et al, 2010; Ferrao, 2010).* La evaluación como tema crítico puede ser abordado de distintas maneras, pudiendo clasificar a los cuestionarios en dos grupos: formativos y evaluativos.

**Cuadro 1.** Tipos de cuestionarios

| Cuestionarios Formativos  | Cuestionarios Evaluativos o sumativos   |
|---|---|
| El objetivo es aprender. El alumno tiene la oportunidad de intentar varias veces resolver el mismo examen o la misma pregunta, pues el objetivo es, que si se identifica que ignora algo, lo aprenda al resolver el examen. | El objetivo es valorar si se otorgan o no los créditos del tema/curso. El profesor usará la calificación del examen como un elemento importante para la calificación del curso. |

Esta experiencia está enfocada en un curso masivo de Estadística I, en la Facultad de ciencias Económicas de La Universidad Nacional de Córdoba. A partir del año 2015 el contenido del Aula Virtual de la asignatura con la cual se venía trabajando, fue migrado al entrono Moodle y en 2016 se aplicaron cuestionarios, propuestos a los alumnos como autoevaluaciones de proceso. Los cuestionarios en Moodle no sólo han demostrado ser útiles para llevar a cabo evaluaciones sino que pueden ser modificados y adaptados según las necesidades del proceso de enseñanza aprendizaje. Tales ajustes pueden ser propuestos con fundamentos basados en el análisis de los índices psicométricos que se obtienen a partir de la misma plataforma Moodle.

## 2 Descripción de la experiencia

La presente experiencia de enseñanza-aprendizaje se encuentra contextualizada dentro de los cursos de Estadística previstos en el ciclo Formación Básica Común del actual plan de estudios (Año 2009). Se trata de la asignatura Estadística I, que abarca conceptos básicos de Estadística Descriptiva, Introducción a la Probabilidad llegando a incluir temas de distribuciones en el muestreo. La Estadística es una de las disciplinas que toma los principales conceptos de álgebra elemental, lógica simbólica y cálculo diferencial e integral para utilizarlos en el desarrollo y producción de herramientas de análisis de datos, que podrán ser aplicadas al estudio metódico de problemas de los más diversos campos de la realidad. De estos últimos interesan particularmente los relacionados con la economía, la administración y la contabilidad. Asimismo, se apoya en el uso de software estadístico, indispensables para el manejo de bases de datos.

En la asignatura, la cantidad de estudiantes inscriptos para cursar en el periodo analizado, desde 2005 a 2016, fue de 231 alumnos en promedio por semestre, siendo el 85% de los mismos los que registraron alguna actividad en el cursado. El Cuadro 1 resume indicadores básicos que se encontraron de cada curso.

Cuadro 2. Indicadores básicos

| Año         | Turno        | Cant. Inscriptos | Cant. Con Actuación | Regular    | Regular (% Alumnos con Actuación) | Libre     | Libre (% Alumnos con Actuación) |
|-------------|--------------|------------------|---------------------|------------|-----------------------------------|-----------|---------------------------------|
| 2005        | Noche        | 255              | 214                 | 130        | 61%                               | 84        | 39%                             |
| 2006        | Noche        | 256              | 192                 | 113        | 59%                               | 79        | 41%                             |
| 2007        | Noche        | 219              | 164                 | 84         | 51%                               | 80        | 49%                             |
| 2008        | Noche        | 211              | 170                 | 87         | 51%                               | 83        | 49%                             |
| 2009        | Noche        | 182              | 143                 | 84         | 59%                               | 59        | 41%                             |
| 2010        | Noche        | 169              | 135                 | 82         | 61%                               | 53        | 39%                             |
| 2011        | Tarde        | 242              | 207                 | 133        | 64%                               | 74        | 36%                             |
| 2012        | Tarde        | 213              | 176                 | 107        | 61%                               | 69        | 39%                             |
| 2013        | Tarde        | 220              | 187                 | 103        | 55%                               | 84        | 45%                             |
| 2014        | Tarde        | 294              | 253                 | 159        | 63%                               | 94        | 37%                             |
| 2015        | Tarde        | 245              | 222                 | 148        | 67%                               | 74        | 33%                             |
| <b>2016</b> | <b>Tarde</b> | <b>270</b>       | <b>243</b>          | <b>174</b> | <b>72%</b>                        | <b>69</b> | <b>28%</b>                      |

A partir del año 2016 se planificó incorporar cuatro instancias de evaluaciones individuales, presentadas a los alumnos como "Autoevaluaciones" no obligatorias, con recompensa de puntos que se suman al puntaje de los correspondientes parciales, que sí son obligatorios para obtener la regularidad de la materia. 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 restrinja a los días previos a cada parcial.

## 2.1 Diseño del Instrumento de evaluación

El equipo docente recorrió las etapas sugeridas para la elaboración de un instrumento de evaluación: Preparación, Aplicación, Valoración y Devolución. El desafío fue repensar esas instancias en el entorno virtual.

La etapa de "Preparación" implicó investigar en las alternativas de construcción del "Banco de Preguntas", definir la creación de las categorías que agrupen las preguntas conforme a los contenidos de la asignatura. La importancia que tiene darle una estructura al banco de preguntas es permitir la reutilización de las mismas, ya que el banco de preguntas es escalable, año a año se pueden sumar preguntas a las categorías de preguntas creadas, o crear nuevas categorías y también modificar las preguntas para mejorarlas, ampliarlas, etc. El hecho de que las preguntas sean reutilizables, permite crear cuestionarios a partir de una mezcla aleatoria de preguntas de diferentes categorías, dando como resultado exámenes diferentes sobre un mismo conjunto de categorías predefinidas<sup>1</sup>.

La etapa de "Aplicación", requiere también prever aspectos operativos vinculados a decidir la manera en la que el alumno responderá el cuestionario. Por la cantidad de alumnos que cursaban se realiza una inscripción previa de aquellos que deseaban participar, para organizar la disponibilidad de aulas y de computadoras. Dado el

<sup>1</sup> documento: [https://docs.moodle.org/all/es/Gestionando\\_preguntas](https://docs.moodle.org/all/es/Gestionando_preguntas)

carácter evaluativo otorgado a estos cuestionarios, los alumnos realizaron la actividad en aulas informáticas de la Facultad, de manera individual y se aprovecharon las medidas de seguridad con las que cuenta el recurso:

- Se indicó una dirección IP fija pública (la de la Facultad) desde la cual sólo se podía responder el cuestionario.
- Se empleó una clave de seguridad para iniciar la evaluación que se otorga al alumno al ingresar al aula.
- Se eligió un orden aleatorizado para las preguntas de respuestas múltiples.

Las etapas de valoración y devolución son amplia y fácilmente aplicables en los cuestionarios de Moodle, ya que en su diseño es posible ponderar el peso de cada pregunta y también es de suma utilidad la posibilidad de incluir una "retroalimentación" a cada respuesta (correcta o incorrecta) y/o una "retroalimentación" global. El esmero en completar de manera clara esta información representa un aporte al alumno para que pueda comprender cuál fue su error o reforzar el concepto o contenido cuando responden bien una pregunta.

Por último, cada alumno luego de concluida la evaluación puede volver a consultar en cualquier momento su cuestionario con el objetivo de coadyuvar al estudio y a la preparación para la instancia del parcial.

### 3 Resultados y discusión

Para el año 2016, las conclusiones obtenidas a partir de los resultados de dichas experiencias evidencian una mejora en la motivación de los alumnos para participar en las actividades de evaluaciones que se propusieron y, en principio, sobre rendimiento académico medido a partir de las notas de parciales y exámenes finales. Cabe aclarar que tales conclusiones son preliminares aún aceptado la existencia de otros múltiples factores no contemplados en los análisis cuantitativos que se presentan aquí con fines descriptivos y que se podrían incorporar a futuro para mejorar el estudio inferencial del impacto de la metodología sobre el aprendizaje.

Asimismo, es esperable y razonable pensar que los alumnos que realizan las Autoevaluaciones obtengan mejores notas en promedio, que aquellos que no participan en la actividad. Sin embargo sería deseable encontrar tal diferencia significativa en las notas del Examen Final. El ANOVA refleja que aún en la instancia de la primera vez que se presenta a rendir el final (Nota Examen 1), la nota promedio de los alumnos que participan en tres o cuatro autoevaluaciones es significativamente mayor que las notas promedio de aquellos que no hacen la actividad o se prestan hasta en dos instancias.

*Cuadro 3: Análisis de la Varianza*

| Fuente | Suma de cuadrados | G. de libertad | Cuadrados Medios | F    | p-valor |
|--------|-------------------|----------------|------------------|------|---------|
| Modelo | 34,95             | 2              | 17,47            | 3,72 | 0,0265  |
| Error  | 713,99            | 152            | 4,7              |      |         |
| Total  | 748,94            | 154            |                  |      |         |

El Test DGC (Alfa=0.05 – PCALT=0.8710) con un nivel de significación del 5% arroja diferencias significativas en el promedio de notas en el examen final de los alumnos que respondieron 3 o más cuestionarios con respecto al promedio de notas de los que realizaron menos de 2 o menos evaluaciones.

---

### 3.1 Revisión del cuestionario

Este resultado positivo motiva al equipo docente a continuar y avanzar sobre una nueva etapa de revisión de los cuestionarios, apelando a elementos de la Psicometría, en dos aspectos un referido al cuestionario en línea en su conjunto como instrumento de evaluación y el otro aspecto apunta a analizar individualmente los ítems lo conforman. Por ello se comenzó planteando la necesidad de ganar experiencia en realizar una correcta interpretación de los indicadores que devuelve la plataforma Moodle, en particular el "Reporte de estadísticas de examen", con el objetivo de aprovechar el potencial de éste recurso, proponiendo ajustes de diseño fundamentados en análisis cuantitativos objetivos.

De manera ejemplificativa se muestra el reporte del Primer cuestionario aplicado el 08/04/2016:

#### Información sobre el cuestionario

Nombre del cuestionario 1° Grupo - 14 hs Aula D2 - BATERÍAS D 2016

Nombre del curso Estadística I - División Caro

Abrir cuestionario viernes, 8 de abril de 2016, 14:00

Cerrar cuestionario viernes, 8 de abril de 2016, 16:00

Abierto para 2 horas

Número de primeros intentos 35

Número total de intentos 35

Promedio de los primeros intentos 72%

Promedio de todos los intentos 72%

Calificación media de los últimos intentos 72%

Calificación media de los mejores intentos 72%

Calificación media (de todos los intentos) 80%

Desviación estándar (para todos los intentos) 16%

Asimetría de la distribución de puntuaciones (para todos los intentos) -0,80

Curtosis de la distribución de puntuaciones (para todos los intentos) 0,48

Coefficiente de consistencia interna (para todos los intentos) 42%

Ratio de error (para todos los intentos) 76%

Error estándar (para todos los intentos) 12%

Cuadro 4: Comportamiento individual del ítem

| Q# | Tipo de pregunta | Nombre de la pregunta                        | Intentos | Índice de dificultad | Desviación estándar | Calificación aleatoria estimada | Peso estimado | Peso efectivo | Índice de discriminación | Eficiencia discriminativa |
|----|------------------|--|----------|----------------------|---------------------|---------------------------------|---------------|---------------|--------------------------|---------------------------|
| 1  | Opción múltiple  | Conceptos básicos                            | 35       | 68.57%               | 47.10%              | 20.00%                          | 10%           | 16.18%        | 49.10%                   | 62.43%                    |
| 2  | Opción múltiple  | Tipo de variable y Escala de medición        | 35       | 77.14%               | 42.60%              | 25.00%                          | 10%           | 14.89%        | 45.49%                   | 59.06%                    |
| 3  | Opción múltiple  | EQUIPAMIENTO TÉCNICO Y CALIDAD               | 35       | 71.43%               | 45.83%              | 20.00%                          | 10%           | 13.55%        | 23.92%                   | 31.86%                    |
| 4  | Opción múltiple  | Empresas exportadoras                        | 35       | 85.71%               | 35.50%              | 20.00%                          | 10%           | 13.47%        | 48.13%                   | 66.27%                    |
| 5  | Opción múltiple  | Pymes Industriales Familiares                | 35       | 62.86%               | 49.02%              | 20.00%                          | 10%           | 8.62%         | -12.48%                  | -14.68%                   |
| 6  | Opción múltiple  | Innovación                                   | 35       | 40.00%               | 49.71%              | 20.00%                          | 10%           | 6.47%         | -20.96%                  | -28.87%                   |
| 7  | Emparejamiento   | Rubro de Actividad e Inversión en Tecnología | 35       | 100.00%              | 0.00%               | 12.50%                          | 10%           | 0.00%         |                          |                           |
| 8  | Opción múltiple  | Interpretando "Frecuencias"                  | 35       | 25.71%               | 44.34%              | 20.00%                          | 10%           | 12.79%        | 20.12%                   | 37.82%                    |
| 9  | Opción múltiple  | Interpretando Frecuencias                    | 35       | 97.14%               | 16.90%              | 25.00%                          | 10%           | 7.95%         | 38.24%                   | 100.00%                   |
| 10 | Opción múltiple  | Interpretando Frecuencias                    | 35       | 94.29%               | 23.55%              | 25.00%                          | 10%           | 6.06%         | 4.68%                    | 7.55%                     |

Este completo reporte que se incluye como submenú dentro de *Administración del examen > Resultados > Estadísticas*, es generado de manera automática por cada Cuestionario que haya sido respondido por un grupo

de alumnos. Muestra en la primera parte un análisis estadístico psicométrico referida al Cuestionario de manera global y luego se dispone una descripción del comportamiento individual de cada ítem, cada uno con enlaces profundizar el análisis minucioso de una pregunta en particular. Estas estadísticas están diseñadas para usarse sólo con exámenes sumativos en donde los alumnos tuvieron un único intento para resolverlo.

A continuación se realiza la interpretación de los indicadores (Estadísticas de examen) obtenidos para el cuestionario presentado como ejemplo, siguiendo algunas referencias que aporta la propia ayuda de Moodle:

## **3.2 Estadísticas del cuestionario**

### **3.2.1 Promedio de los intentos: 72%**

Este valor es la media aritmética de las calificaciones obtenidas por los alumnos que finalizaron el cuestionario, que en este caso al momento de diseñar el cuestionario se optó por una escala de 0 a 100, pero Moodle ofrece varias alternativas. Para exámenes discriminantes con retroalimentación diferida, instituciones expertas de referencia, sugieren que en la aplicación de este tipo de recurso los exámenes eficientes deben obtener una calificación promedio de 50% a 75%. Los valores fuera de estos límites, sugieren repensar la evaluación.

### **3.2.2 Calificación media (de todos los intentos): 80%**

Bajo este título Moodle reporta el valor de la **Mediana** de las calificaciones. La Mediana es el valor de la calificación que divide en dos partes iguales a las 35 calificaciones, dispuestas en orden según su magnitud. Es decir, es el valor central de la distribución. Entonces su interpretación es que la mitad de los alumnos obtuvieron una calificación inferior al 80% y el 50% restante, alcanzaron una calificación mayor que 80%.

### **3.2.3 Desviación estándar (para todos los intentos): 16%**

Es una medida de la dispersión de las calificaciones alrededor del Promedio. La sugerencia es intentar obtener valores entre 12% y 18%. Desvíos estándar inferiores a 12% sugieren que las calificaciones están demasiado amontonadas entorno al promedio.

### **3.2.4 Asimetría de la distribución de puntuaciones (para todos los intentos): -0,80**

Se trata de una medida de la asimetría de la distribución de calificaciones. Si se obtiene "Cero" implica que la distribución de las calificaciones es perfectamente simétrica, mientras que los valores positivos indican presencias de unos pocos valores de calificaciones altas ('cola' a la derecha) y los coeficientes negativos indican presencia de pocos valores bajos y calificaciones concentradas en valores altos de la escala.

Las recomendaciones para este indicador es intentar obtener un valor de sesgo de -1 a 1. Si fuera un valor negativo muy elevado, puede estar indicando una falta de discriminación entre los alumnos, a los que les va mejor que al promedio. Asimismo, valores positivos muy altos (mayores a 1) pueden revelar falta de discriminación cerca de la calificación umbral para pasar/reprobar.

### 3.2.5 Curtosis de la distribución de puntuaciones (para todos los intentos): 0,48

La Curtosis es una medida de forma e indica que tan aplanada es la distribución, es decir aporta una idea acerca de la deformación vertical de la misma. Para este indicador se busca obtener un valor en el rango entre 0 y 1. Ya que 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 son “alumnos promedio”.

### 3.2.6 Coeficiente de consistencia interna (para todos los intentos): 42%

Este coeficiente (Alpha de Cronbach) aporta una medición global, o también llamada “*medición compuesta*” sobre todos los ítems (o *componentes de medición*), como una estimación (inferior) de la fiabilidad de una prueba psicométrica. Aquí es deseable alcanzar valores mayores, 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 está funcionando con una calidad diferente al restos de los ítems, provocando que el examen en su conjunto no sea homogéneo.

En la práctica es casi imposible obtener consistencia interna mayor al 90%, por ello la mayor parte la bibliografía que trata el tema indica que un valor de 75% o más es satisfactorio. Se acepta que si el valor del coeficiente es inferior a 64%, el examen completo es insatisfactorio, y en nuestro caso se obtuvo un 42% por tanto se deberían considerar revisión del instrumento para adoptar medidas correctivas.

### 3.2.7 Ratio de error (para todos los intentos): 76%

Esta tasa de error se relaciona con el coeficiente de consistencia interna anterior, según la bibliografía de referencia, de acuerdo con la tabla que se muestra a continuación:

Cuadro 5: Relación entre Coeficiente de Consistencia Interna y la Tasa de Error

|                                     |     |    |    |    |    |    |    |    |
|-------------------------------------|-----|----|----|----|----|----|----|----|
| Coeficiente de Consistencia Interna | 100 | 99 | 96 | 91 | 84 | 75 | 64 | 51 |
| Tasa de Error                       | 0   | 10 | 20 | 30 | 40 | 50 | 60 | 70 |

La tasa de error estima el porcentaje de la desviación estándar que debida a efectos aleatorios en lugar de diferencias auténticas de la habilidad entre los estudiantes. De manera que cuestionarios con valores de tasa de error superiores al 50%, no deberían considerarse satisfactorios. Así, se muestra que el ratio de error para el cuestionario bajo análisis fue igual a 76%, evidenciando la necesidad de revisarlo. Tal valor implica que un cuarto de la desviación estándar se debe a diferencias en habilidad de los alumnos y que el resto son efectos aleatorios.

### 3.2.8 Error estándar (para todos los intentos): 12%

El Error estándar es la desviación estándar multiplicada por la Tasa de error, y dividida por 100. Es una medida de la incertidumbre en la calificación de cualquier alumno dado. Hipotéticamente, si ese mismo estudiante resolviera otro cuestionario equivalente en la misma cátedra, se esperaría que su calificación estuviera dentro de más-menos un error estándar de la calificación anterior. Las interpretaciones aceptadas para este indicador plantean que cuanto menor sea su valor, mejor es cuestionario. Pero la experiencia demuestra que lograr que el



error estándar sea inferior al 5% o 6%. Mientras que si el error estándar es mayor al 8%, es probable que una proporción sustancial de los alumnos se encuentren erróneamente calificados.

### 3.3 Estadísticas de las preguntas o ítems

Luego de la vista general, Moodle reporta el análisis detallado de todas las preguntas o ítems del cuestionario. A su vez individualmente se puede ir profundizando el análisis, ítem por ítem, es decir indagando sobre el comportamiento individual de los *“componentes de medición”* que conforman las mediciones compuestas del reporte general. (Cuadro 4).Entonces, para cada ítem, se detalla el número de orden en el que se ubican dentro del cuestionario y el *“Tipo y Nombre de la Pregunta”* según cómo fueron definidas en el Banco de preguntas. Luego, la cantidad de intentos efectuados por ítem y en las columnas subsiguientes se muestran los indicadores de interés para valorar el aporte individual del ítem

#### 3.3.1 Índice de Facilidad / Nivel de Dificultad

Es simplemente 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.

**Cuadro 6:** Índice de Facilidad: Valores de referencia.

| Índice de Facilidad (%) | hasta 5  | 6 - 10      | 11 - 20 | 20 - 34               | 35 - 64                          | 66 - 80        | 81 - 89 | 90 - 94   | 95- 100              |
|-------------------------|--|-------------|---------|-----------------------|----------------------------------|----------------|---------|-----------|----------------------|
| Interpretación          | Extremadamente difícil, o algo está mal en la pregunta | Muy difícil | Difícil | Moderadamente difícil | Correcta para el alumno promedio | Bastante fácil | Fácil   | Muy fácil | Extremadamente fácil |

Como posible regla general, los índices de facilidad entre 30% y 70% suelen aportar diferencias importantes entre el nivel de conocimiento, habilidad y preparación entre los alumnos. Es importante que el docente tenga claro cuál es el propósito de la evaluación y las características de los estudiantes, para poder plantear niveles de dificultad deseados acordes.

#### 3.3.2 Desviación estándar

Es una medida de la dispersión de las calificaciones respecto a la media, mide la magnitud de cuánto puede discriminar la pregunta. Ante un índice de facilidad muy alto (o muy bajo), la dispersión será grande, sin embargo una buena desviación estándar no implica necesariamente una buena discriminación. Según la tabla de referencia un valor de 1/3 del máximo de la pregunta (33%) en la tabla de arriba generalmente es insatisfactorio

#### 3.3.3 Puntaje esperado aleatoriamente

Este es el promedio de calificación que se esperaría que los estudiantes obtuvieran si respondieran al azar. Sólo disponibles para las preguntas de opción múltiple y válido para cuestionarios de retroalimentación diferida. Son deseables valores por debajo del 40%.

#### 3.3.4 Ponderación deseada y efectiva

La ponderación deseada o peso de la pregunta, Moodle la expresa como un porcentaje del puntaje general. Y la ponderación efectiva es una estimación de la ponderación o peso que realmente tiene la pregunta para contribuir

a la dispersión total de las calificaciones. Vemos en el Cuadro 5 que algunas ponderaciones efectivas resultaron mayores o menores que las deseadas, demostrado para cuándo son mayores que la pregunta tiene una mayor participación en la dispersión de las calificaciones de lo que deseado al momento de diseñar el cuestionario.

### 3.3.5 Índice de discriminación y Eficiencia de discriminación

Esta es la correlación entre las calificaciones ponderadas en la pregunta y las del resto del examen. Revela que tan efectivo es el ítem para clasificar o separar a los alumnos más capaces de los menos capaces. Los resultados pueden interpretarse según el Cuadro 6 siguiente:

**Cuadro 7:** índice de discriminación de los ítems.

| Índice (%)     | valores negativos              | 0 - 19                      | 20 - 29                  | 30 - 50                    | 50 o superior               |
|----------------|--------------------------------|-----------------------------|--------------------------|----------------------------|-----------------------------|
| Interpretación | Ítem probablemente<br>inválido | Discriminación<br>muy débil | Discriminación<br>débil. | Discriminación<br>Adecuada | Discriminación<br>Muy buena |

### 3.3.6 Eficiencia de discriminación

El indicador intenta estimar que tan bueno es el índice de discriminación en relación con la dificultad de la pregunta. Un ítem muy fácil o muy difícil no discrimina entre alumnos con habilidades diferentes, debido a que la mayoría de ellos obtendrán el mismo puntaje para esta pregunta y viceversa. En el Cuestionario aparecen ítems con valores negativos que deberían ser revisados.

## 4 Conclusiones finales

Se encontraron evidencias sobre el impacto que tiene la aplicación de los cuestionarios de Moodle sobre el rendimiento de los alumnos, medido a partir de la nota del examen final, hecho que motiva a continuar mejorando la propuesta de enseñanza. Pero el análisis exhaustivo de los instrumentos aplicados muestra falencias importantes en el diseño de los mismos, a partir de la interpretación del Reporte de estadísticas de examen. Revelando la existencia de un enorme potencial para mejorar 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 que en definitiva les permiten aprobar 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).
- Steegmann, 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*.