

# Explorando la dificultad de los ítems en las pruebas de estudiantes de grado de Arquitectura: Un análisis en el contexto del curso PMEC

**Palabras clave:** Diseño de ítems, Pruebas de selección múltiple, PMEC - FADU - UDELAR, Teoría de Respuesta al ítem, Modelos lineales y no lineales Generalizados Mixtos.

## Abstract

El trabajo se centra en el armado de una base de datos para profundizar en el análisis de la Teoría de Respuesta al Ítem (TRI), utilizando diversas librerías especializadas del lenguaje R. El objetivo principal es trabajar con datos de calificaciones de los estudiantes del curso de Principios Matemáticos de la Estabilidad de las Construcciones (PMEC), junto con datos sociodemográficos.

La base de datos creada incluye variables como Id, Grupo, Edad, Sexo, Dedicación antes y posterior a las pruebas, y las repuestas dadas por cada estudiante en dos pruebas con 10 preguntas cada una de selección múltiple. Estas variables representan información relevante para el análisis y la evaluación del desempeño de los estudiantes en el curso PMEC de la carrera de Arquitectura de la Facultad de Arquitectura, Diseño y Urbanismo (FADU), de la Universidad de la República.

La construcción de la base de datos se llevó a cabo mediante la aplicación de "reshape2" para transformar los datos a diferentes formatos, "dplyr" para manipular y transformar los datos, y "openxlsx" para leer y escribir archivos de Excel y la aplicación de funciones auxiliares previamente definidas. Además, se realizaron otras transformaciones con el fin de normalizar columnas y cálculos de sumas, para preparar los datos y así aplicar modelos específicos de la TRI.

En el campo de la Teoría de Respuesta al Ítem (IRT), se destacan los Modelos Lineales y No Lineales Generalizados Mixtos (GLMM y GNLMM) por su flexibilidad para incorporar características tanto del ítem como de la persona. Por lo tanto, en una segunda etapa, se aplicaron diferentes funciones de paquetes de R para analizar los ítems y las propiedades de las personas. Algunas de las funciones utilizadas son: "glmer" del paquete "lme4", "eirm" del paquete "eirm", "RM" del paquete "eRm", "rasch" del paquete "ltm" y "tam" del paquete "TAM".

En varios modelos, fue necesario profundizar en la aplicación de las técnicas de optimización más adecuadas para obtener mejores ajustes a los datos. Por ejemplo, de los diferentes paquetes y técnicas utilizados para ajustar los modelos de la TRI, en el paquete lme4, que incluye la función glmer fue necesario utilizar el paquete optimx para acelerar el proceso de optimización y ajustar de manera adecuada el modelo.

La aplicación de estos modelos ha demostrado su utilidad al identificar, en cada etapa de evaluación del curso PMEC, los ítems que presentan propiedades más difíciles en comparación con aquellos que poseen diferentes características. Gracias a este exhaustivo análisis, se ha logrado determinar hasta qué punto cada paquete es capaz de ajustar los modelos IRT.

Cabe destacar que cada paquete ofrece opciones y funcionalidades distintas, y su elección depende de las necesidades específicas del análisis. Se realizó una evaluación detallada para determinar cuál modelo es más recomendable en cada caso, considerando factores como la complejidad de las propiedades de los ítems y las dificultades asociadas. Estas decisiones son cruciales para obtener resultados precisos y confiables en el análisis de datos educativos y proporcionar recomendaciones fundamentadas para mejorar el proceso de evaluación.

## Bibliografía

Bolt, D. M., Kim, J.-S., Blanton, M., Knuth, E. (2016). Chapter 2: Applications of Item Response Theory in Mathematics Education Research. *Journal for Research in Mathematics Education*. Monograph, 15, 31–52.

De Boeck, P. (Eds.), Wilson, M. (Eds.). (2004). *Explanatory Item Response Models: A Generalized Linear and Nonlinear Approach*. Statistics for Social and Behavioral Sciences. Springer Science & Business Media.

Gil, O., Pantaleón, C., Patiño, L., Perchman, M., Seijo, R., Zino, L. (2006). Una propuesta para la Cátedra de Matemática de la Facultad de Arquitectura. Montevideo.

Gil, O., Nuin, A.L. (2017) Transformación de la enseñanza en un curso de Matemática de primer año para estudiantes de Arquitectura (2013-2016), comunicación presentada en JIES, Jornadas de Investigación en Educación Superior, Montevideo, Uruguay.

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of educational measurement*, 50(1), 1-73.

Mair, P., and Hatzinger, R. (2007). CML based estimation of extended Rasch models with the eRm package in R. *Psychology Science*, 49, 26-43.

Reynolds, K., Khorramdel, L., von Davier, M. (2022). Can students' attitudes towards mathematics and science be compared across countries? Evidence from measurement invariance modeling in TIMSS 2019. *Studies in Educational Evaluation*, 74.

Rjimen, F., Tuerlinckx, F., De Boeck, P., & Kuppens, P. (2003). A nonlinear mixed model framework, for Item Response theory. *Psychological Methods*, 8(2), 185 – 205.

Rose, N., Nagy, G., Nagengast, B., Frey, A., Becker, M. (2019). Modeling Multiple Item Context Effects With Generalized Linear Mixed Models. *Frontiers in Psychology*, 10.

Explorando la dificultad de los ítems en las pruebas de estudiantes de grado de  
Arquitectura: Un análisis en el contexto del curso PMEC

---

Wittmann, E., Weyland, U., Seeber, S., Warwas, J., Striković, A., Krebs, P., Pohley, M., Wilczek, L. (2022). Test sensitivity in assessing competencies in nursing education. *Empirical Research in Vocational Education and Training*, 14(1), 1–26.