

Diferencia de la Dificultad como procedimiento para detectar funcionamiento diferencial del ítem (DIF) en SABER 5° y 9°

Martha Ligia Cuevas mcuevas@icfes.gov.co Víctor H. Cervantes vcervates@icfes.gov.co
Subdirección de Estadística – Instituto Colombiano para la Evaluación de la Educación (ICFES)



Introducción

El funcionamiento diferencial del ítem se presenta cuando evaluados comparables en la puntuación total de una prueba, pero provenientes de grupos diferentes, responden de forma distinta a los ítems individuales de la misma (Camilli & Sheprd, 1994). El DIF se relaciona con los siguientes conceptos:

- Validez.
- Interpretaciones iguales de las puntuaciones por género, zona, sector y nivel socioeconómico.
- Equidad.
- Funcionamiento Diferencial del Ítem (DIF o FDI).
- Impacto. Sesgo
- Métodos para detectar DIF.

Diferencia de la dificultad

Este procedimiento consiste en comparar las dificultades de los ítems para cada grupo, usualmente después de haber sido calibrados con un modelo de un parámetro en el marco de la teoría de respuesta al ítem, a través de la siguiente prueba estadística:

$$t = \frac{b_1 - b_2}{\sqrt{SE_{b_1} + SE_{b_2}}}$$

Donde:

b_1 es la dificultad para el grupo 1.

b_2 es la dificultad para el grupo 2.

SE_{b_1} es el error estándar de la dificultad para el grupo 1.

SE_{b_2} es el error estándar de la dificultad para el grupo 2.

Algunos estudios han mostrado que la detección de DIF a través de este procedimiento puede verse afectada por variables como la razón de tamaños de muestra (diferencia en el tamaño entre los grupos que se están comparando), el desajuste del modelo, el porcentaje de ítems con DIF y el impacto (diferencias reales en la distribución del rasgo entre los grupos) (Berrío, 2008).

Objetivos

- Identificar algunas variables que pueden afectar la diferencia de la dificultad.
- Encontrar los mejores puntos de corte para la diferencia de la dificultad para la detección de DIF, tomando en cuenta algunas variables que los afectan

¿Cómo funciona la diferencia de la dificultad bajo diferentes condiciones teniendo en cuenta su potencia y tasas de error tipo I?

Metodología

Estudio de simulación.

Variables independientes

- Tamaño de muestra: $n = 7500, 8500, 26000$ y 33000 .
- Razón de tamaños entre los grupos: 1, 3, 4, 6, 10, 20, 150 y 250.
- Longitud de la prueba: $K = 18$ y 26 .
- Magnitud de DIF: No DIF, dos ítems con DIF (pequeño y grande), y cuatro ítems con DIF (pequeño y grande).
- 270 condiciones experimentales.
- Réplicas por condición: $R = 200$.
- Dos ítems con DIF uniforme y dos ítems con DIF no uniforme cuando la condición tenía cuatro ítems.
- Variables fijas: dificultad y discriminación.

Variables dependientes y modelos

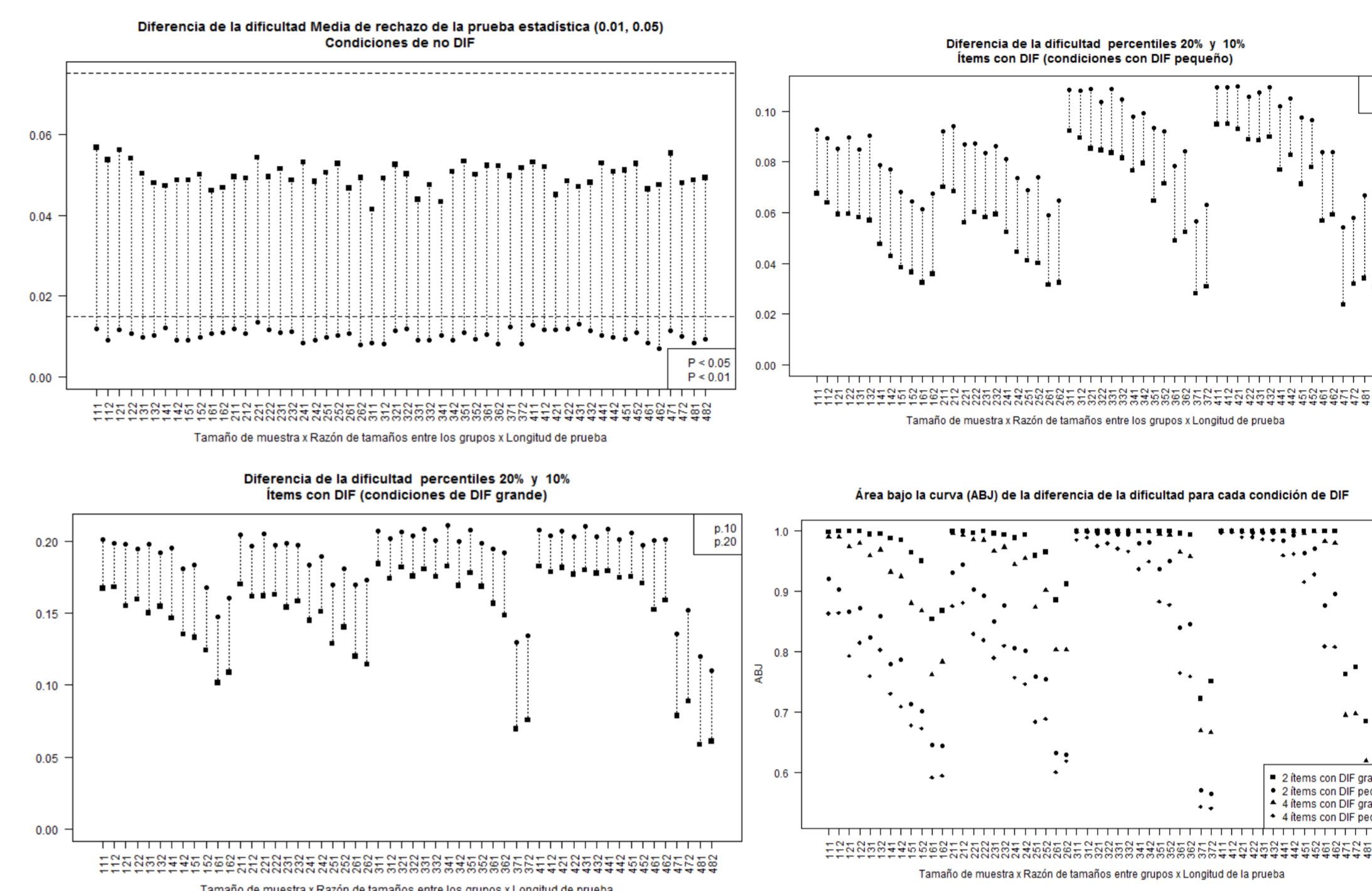
- Diferencia de la dificultad (medida de tamaño del efecto).
- Modelos lineales para definir cuáles variables afectaban la medida de tamaño del efecto.

Puntos de Corte

- Curvas operador receptor (COR o ROC).
- Para cada condición.
- Para cada combinación de variables con efecto sobre la diferencia de la dificultad.
- Selección de los mejores puntos de corte en forma eficiente, es decir, manteniendo buenas tasas de error tipo I y potencia.

Resultados

Las variables razón de tamaños, tamaño de DIF y longitud de la prueba tuvieron efecto sobre la diferencia de la dificultad ($p < 0.05$) como puede apreciarse en las siguientes figuras.



Discusión

- Debido a las altas tasas de error tipo I fue necesario llevar a cabo una etapa de purificación, en donde los ítems detectados con DIF en una primera fase no fueron tenidos en cuenta para la estimación de las habilidades en una segunda fase. Las figuras corresponden al segundo análisis.
- Aunque algunos puntos de corte fueron obtenidos para niveles en los que variaban la razón, la magnitud del DIF y la longitud de prueba, estos puntos de corte pudieron ser simplificados sin perder potencia ni control sobre las tasas de error Tipo I para algunas condiciones.
- Es necesario contar con puntos de corte específicos de las medidas de tamaño del efecto para condiciones particulares.
- Los resultados concuerdan con otros estudios como el de Berrío (2008) en cuanto a que se aprecia un efecto de la razón de tamaños.

Referencias

- Berrío, A. I. (2008). *Efecto de la Razón de Tamaños y Desajustes al Modelo en la Detección de Ítems con Funcionamiento Diferencial mediante Procedimientos Basados en IRT (Diferencia de Dificultad y de Lord*. Tesis de Maestría para optar por el de Psicología, Universidad Nacional de Colombia, Bogotá, Colombia. título de Magíster en Psicología, Departamento
- Camilli, G. & Shepard, L. A. (1994). *Methods for Identifying Biased Test Items*. 4. United States: SAGE Publications.



Las opiniones, tesis y argumentos expresados son de propiedad exclusiva de los autores y no representan el punto de vista del Instituto Colombiano para la Evaluación de la Educación ICFES.

Prosperidad para todos