

# Comprensión de Análisis Cuantitativo de Resultados (en Educación)

Primer Ciclo de Clases FEVED - Noviembre 2025

Bladimir Padilla    Ariel Aloe

University of Iowa

# Contenidos

- Antecedentes
- Criterios para comprender (producir y consumir) análisis cuantitativos de resultados
- Ejemplos de análisis cuantitativo: modelos y resultados comunes en contextos educativos

# Marco de trabajo

- Propósito: Reconocer características básicas del análisis cuantitativo de resultados (educativos)
- Aprendizajes esperados:
  - Comprender elementos básicos del análisis cuantitativo de resultados
  - Identificar criterios para producir/consumir resultados cuantitativos
  - Evaluar críticamente los componentes de un análisis cuantitativo
- Reglas de participación:
  - Se puede participar en cualquier momento escribiendo en el chat o levantando la mano (virtual)
  - La participación es altamente valorada

# **Antecedentes**

# Glosario: Análisis Cuantitativo

Ha sido definido de múltiples maneras: enfoque sistemático para examinar fenómenos mediante datos numéricos y métodos estadísticos; proceso que utiliza técnicas matemáticas, estadísticas o computacionales para analizar datos numéricos y generar conocimientos; proceso que recopila y evalúa datos medibles y verificables para comprender el comportamiento y predecir resultados.

En esta sesión, análisis cuantitativo se refiere a la aplicación de técnicas estadísticas sobre el análisis de datos (especialmente numéricos) en contextos educativos.

# Glosario: Resultados Educativos

Los resultados cuantifican un producto, logro o desempeño en contextos educativos. Pueden incluir puntajes de pruebas estandarizadas, tasas de graduación, niveles de competencia, entre otros.

En esta sesión, resultados educativos se refiere a medidas cuantitativas que reflejan el producto de alguna medición de desempeño académico, habilidades, competencias o logros en contextos educativos. También, lo usaremos para referirnos a cualquier variable que contenga los valores de esa medición.

# ¿Por qué es importante comprender análisis cuantitativos?

- **Toma de decisiones informadas:** Las políticas y prácticas educativas deben basarse en evidencia rigurosa.
- **Evaluación crítica:** No todos los estudios tienen la misma calidad metodológica.
- **Alfabetización estadística:** La producción masiva de investigación educativa requiere habilidades para interpretar datos.
- **Comunicación efectiva:** Entre investigadores, comunidades educativas y tomadores de decisiones.

# Componentes de un análisis (estudio) cuantitativo

Pregunta de investigación → Diseño → Recolección de datos →  
Análisis → Interpretación → Diseminación

# **Criterios**

# Comprender el propósito del estudio

## Qué hace el estudio?

- *Describir:* Analizar el rendimiento promedio en SIMCE Matemáticas por curso
- *Explorar:* Analizar la relación entre género y rendimiento en SIMCE Matemáticas
- *Explicar:* Analizar el impacto de un programa de reforzamiento en los puntajes de SIMCE Matemáticas
- *Predecir:* Analizar la capacidad predictiva de las horas de estudio sobre el rendimiento en SIMCE Matemáticas

**Implicación:** El propósito guía el diseño del estudio y el análisis e interpretación de los resultados.

# Identificar el tipo de estudio

## Diseños principales:

- *Descriptivo*: Caracterización de fenómenos/muestras
- *Correlacional*: Relaciones entre variables sin manipulación
- *Cuasi-experimental*: Grupos de comparación sin asignación aleatoria
- *Experimental*: Asignación aleatoria de estudiantes a condiciones
- *Longitudinal vs. Transversal*: Seguimiento en el tiempo vs. medición única

**Implicación:** El tipo de estudio determina qué tipo de inferencias son posibles.

# Comprender las hipótesis y preguntas de investigación

## Elementos a identificar:

- **Claridad:** ¿Están claramente formuladas?
- **Especificidad:** ¿Son precisas y testables?
- **Alineación:** ¿Se corresponden con el propósito/diseño del estudio?

**Implicación:** Las hipótesis guían la selección de técnicas estadísticas y la interpretación de resultados.

# Evaluar el diseño del estudio y la metodología

**Metodología:** plan general que nos lleva desde las preguntas de investigación hacia sus respuestas, incluyendo actividades como: muestreo, mediciones, y modelamiento de datos.

## Aspectos clave:

- **Validez interna:** ¿Los resultados reflejan relaciones reales o existen variables confundentes?
- **Validez externa:** ¿Los resultados son generalizables (generalizabilidad vs transportabilidad)?
- **Sesgos:** ¿Existen sesgos de selección, medición o análisis?
- **Control de variables:** ¿Se controlaron factores relevantes?

# Evaluar la muestra y el muestreo

## Preguntas esenciales:

- ¿Cuál es el tamaño de la muestra? ¿Es adecuado para el análisis propuesto?
- ¿Qué método de muestreo se utilizó? (probabilístico vs. no probabilístico)
- ¿Cuál es la tasa de respuesta? ¿Existe sesgo de no respuesta?
- **● ¿La muestra es representativa de la población objetivo?**
- **● ¿Cómo se manejaron los datos faltantes?**

**Consideración:** Muestras grandes no garantizan calidad si el muestreo es deficiente.

# Examinar las variables y medidas

## Tipos de variables más comunes:

- **Dependientes/Resultados:** Su variación es el foco principal del análisis (ej: puntaje SIMCE)
- **Independientes/Predictoras:** Explican/predicen la variabilidad de la variable dependiente (ej: nivel socioeconómico)

Luego tenemos:

- **Covariables:** Sirven para *controlar* la relación entre otras variables (ej: edad)
- **Moderadores:** Moderan la relación entre predictores y resultado (ej: administración de colegio sobre género)
- **Mediadores:** Crean relaciones indirectas entre predictores y resultado (ej: motivación entre género y rendimiento)

# Revisar las estadísticas descriptivas

## Información fundamental:

- **Tendencia central:** Media, mediana, moda
- **Dispersión:** Desviación estándar, rango, varianza
- **Distribución:** Asimetría, curtosis, histogramas

**Utilidad:** Esta información revela patrones, datos atípicos y el comportamiento general de las variables antes del *análisis inferencial*.

# Juzgar las propiedades y supuestos de los análisis estadísticos

## Supuestos comunes:

- **Normalidad:** Los datos siguen una *distribución normal*
- **Homocedasticidad:** La varianza es constante entre grupos/predictores
- **Independencia:** Las observaciones son independientes

**Importante:** En la práctica, estos supuestos rara vez se cumplen, pero podemos usar métodos robustos para superar su violación.

# Interpretar los resultados en su contexto

## Elementos de interpretación:

- **Contexto teórico:** ¿Los resultados tienen sentido conceptualmente?
- **Contexto práctico:** ¿Qué implicaciones tiene esto para las prácticas/políticas educativas?
- **Contexto metodológico:** ¿Los resultados son consistentes con el diseño del estudio?

**Precaución:** Las palabras importan. La interpretación no puede ir más allá de lo que permite el diseño (correlación vs causalidad).

# Considerar la significancia estadística y práctica de los resultados

## Significancia estadística ( $p < 0.05$ ):

- Probabilidad de obtener resultados tan extremos o más extremos que los observados, asumiendo que la hipótesis nula es verdadera.
- Depende fuertemente del tamaño de la muestra, la magnitud del efecto y la variabilidad de los datos.

## Significancia práctica:

- Tamaño del efecto ( $d$  de Cohen,  $r$  de Pearson,  $\eta^2$ )
- Comunicación/interpretación: ¿Una diferencia de 5 puntos en SIMCE es relevante?

**Problema:** Ambos tipos de significancia suelen confundirse.

# Juzgar la validez de los resultados

**Validez:** Grado en que la evidencia y la teoría respaldan la interpretación de los resultados y las decisiones pretendidas.

**Evidencias de validez:** - **Validez de constructo:** ¿Las medidas capturan lo que pretenden medir? - **Validez de contenido:** ¿El contenido de la medida es representativo del dominio? - **Validez de criterio:** ¿Las medidas se correlacionan con criterios externos relevantes?

En medición educativa, la validez es un proceso continuo y perfectible.

# Juzgar la confiabilidad de los resultados

**Confiabilidad:** Grado en que los resultados son consistentes bajo condiciones similares.

## Perspectivas de confiabilidad:

- **Confiabilidad de formas alternativas:** Consistencia de resultados entre dos formas equivalentes de un test
- **Confiabilidad test-retest:** Consistencia de resultados de una administración del test a otra
- **Confiabilidad de consistencia interna:** Consistencia de cada ítem con los otros ítems del test

Alfa de Cronbach ( $\alpha$ ) es una medida común de consistencia interna con fuertes supuestos psicométricos.

# Evaluar las limitaciones del estudio

## **Limitaciones comunes:**

- Diseño que no permite inferencia causal
- Muestra no representativa o de conveniencia
- Medidas de auto-reporte con sesgos
- Datos faltantes significativos

# Evaluar las implicaciones y recomendaciones

## Preguntas críticas:

- ¿Las implicaciones se derivan lógicamente de los resultados?
- ¿Las recomendaciones son prácticas y factibles?
- ¿Se consideran las limitaciones al formular recomendaciones?
- ¿Se identifican áreas para investigación futura?
- ¿Las conclusiones son consistentes con la literatura existente?

**Advertencia:** Desconfíe de recomendaciones basadas en estudios con diseños deficientes.

# Ejemplos: Modelos

# Regresión Lineal

**¿Qué es?** Modelo estadístico (*representación matemática simplificada de un fenómeno*) que examina la relación entre una variable dependiente continua y una o más variables independientes.

**Ecuación básica:**  $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \epsilon$

**Interpretación de resultados:**

- **Coeficientes o pendientes ( $\beta$ )**: Cambio esperado en Y por cada unidad de cambio en X
- **Intercepto ( $\beta_0$ )**: Valor esperado de Y cuando todas las X = 0
- **R<sup>2</sup>**: Proporción de varianza en Y explicada por el modelo (0 a 1)
- **Error residual ( $\epsilon$ )**: Diferencia entre los valores observados y los predichos por el modelo

# Regresión Lineal

**Ejemplo:** Predecir puntaje SIMCE Matemáticas ( $Y$ ) a partir de NSE ( $X_1$ , escala 1-5) y horas de estudio semanal ( $X_2$ ).

**Ecuación del modelo:**

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$$

**Ecuación estimada:**

$$\hat{Y} = 200 + 15X_1 + 3X_2$$

**Interpretación de coeficientes:**

- $\beta_0 = 200$ : Puntaje SIMCE esperado cuando  $X_1 = 0$  y  $X_2 = 0$  (intercepción)
- $\beta_1 = 15$ : Por cada nivel adicional de NSE, el puntaje SIMCE aumenta en 15 puntos, **manteniendo constantes las horas de estudio**
- $\beta_2 = 3$ : Por cada hora adicional de estudio semanal, el puntaje SIMCE aumenta en 3 puntos, **manteniendo constante el NSE**

# Análisis Factorial Exploratorio y Confirmatorio

## Análisis Factorial Exploratorio (AFE):

- Explora la estructura subyacente de un conjunto de variables observadas (ítems/indicadores)
- Identifica factores latentes y su estructura (número de ítems en cada factor)
- **Interpretación:** Carga factorial ( $\lambda$ ) pendiente ( $\hat{\beta}$ ), intercepto del ítem ( $\mu$ ) intercepto ( $\hat{\beta}_0$ ), error único ( $e$ ) error residual.

## Análisis Factorial Confirmatorio (AFC):

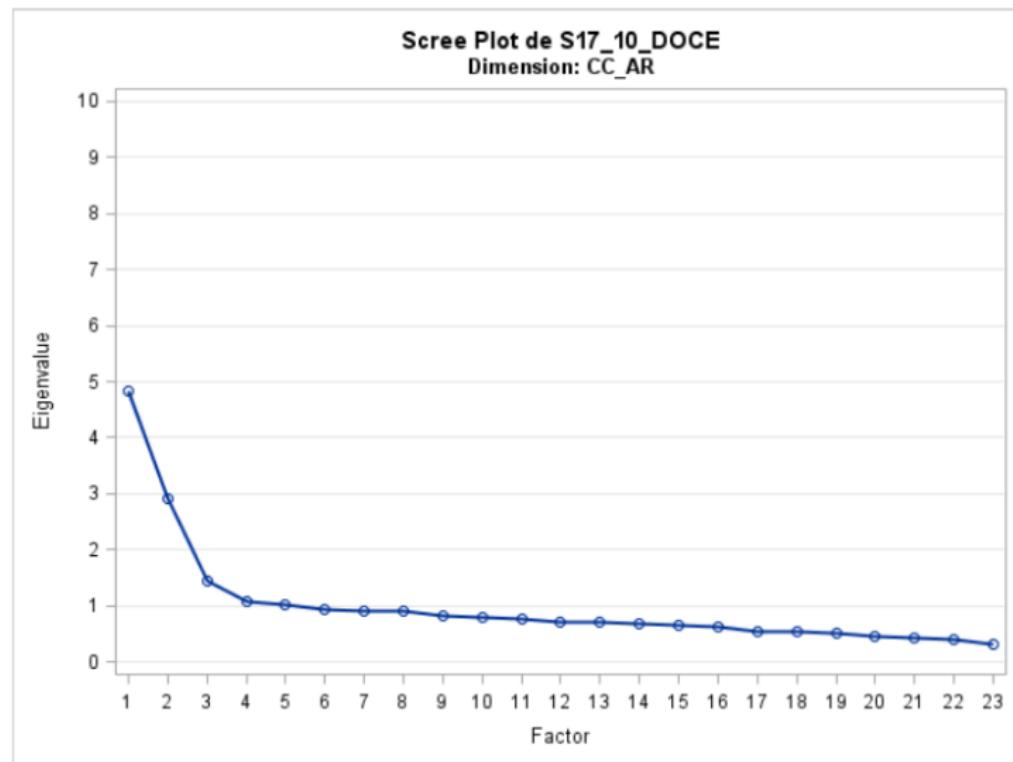
- Prueba una estructura factorial específica basada en teoría
- **Interpretación:** similar a AFE, pero con énfasis en ajuste global y local del modelo.

## Ejemplo

IDPS clima de convivencia escolar. AFE identifica dimensiones subyacentes: ambiente de seguridad, ambiente de respeto, ambiente seguro. AFC provee evidencia sobre el ajuste de esta estructura de 3 factores a los datos.

# Ejemplo: AFE

## B2. Resultados factorial exploratorio (*Scree test*)



Extraído de  
Informe Técnico  
IDPS 2017  
(p. 135).

# Ejemplo: AFC

TABLA 6 ANÁLISIS FACTORIAL CONFIRMATORIO DE SEGUNDO ORDEN PARA DIMENSIONES DE CUESTIONARIO DE PADRES Y/O APODERADOS

Indicador	Dimensión	Grado	N.º de ítems	N.º de factores	RMSEA	CFI	TLI
Clima de convivencia escolar	Ambiente de respeto	4º básico	21	6	0,064	0,951	0,942
		8º básico	24	6	0,054	0,942	0,933
		II medio	24	6	0,051	0,945	0,936
	Ambiente organizado	4º básico	9	1	0,225	0,911	0,882
		8º básico	9	1	0,232	0,896	0,861
		II medio	9	1	0,23	0,87	0,827
	Ambiente seguro	4º básico	10	2	0,092	0,966	0,956
		8º básico	10	2	0,073	0,979	0,972
		II medio	10	2	0,064	0,979	0,973
Participación y formación ciudadana	Participación	4º básico	15	4	0,068	0,959	0,948
		8º básico	15	4	0,07	0,961	0,952
		II medio	15	4	0,069	0,957	0,946

Extraído de Informe Técnico IDPS 2017 (p. 63).

# Ejemplo

## Factor: Ambiente de Respeto:

Ítems ( $X_i$ ) de la batería de preguntas p06 en cuestionario de estudiantes 2M en 2018:

- $X_1$ : "Los estudiantes de mi curso nos respetamos entre nosotros."
- $X_2$ : "Mis profesores se preocupan por tratar bien a los estudiantes."
- $X_6$ : "En mi establecimiento todas las personas se tratan con respeto."
- $X_7$ : "Los adultos del establecimiento se dicen las cosas con respeto."

# Ejemplo

## Modelo AFC para cada ítem

Cada individuo  $s$  responde a los ítems  $i$  como función de su nivel de percepción de un ambiente de respeto ( $\theta_s$ ):

$$\begin{aligned} X_{s1} &= \mu_1 + \lambda_1 \theta_s + e_{s,1}; & e_{s,1} &\sim N(0, \psi_1^2) \\ X_{s2} &= \mu_2 + \lambda_2 \theta_s + e_{s,2}; & e_{s,2} &\sim N(0, \psi_2^2) \\ X_{s6} &= \mu_3 + \lambda_6 \theta_s + e_{s,6}; & e_{s,6} &\sim N(0, \psi_6^2) \\ X_{s7} &= \mu_4 + \lambda_7 \theta_s + e_{s,7}; & e_{s,7} &\sim N(0, \psi_7^2) \end{aligned}$$

# Teoría de Respuesta al Ítem (TRI)

**¿Qué es?** Modelos psicométricos que relacionan las respuestas a ítems individuales con la habilidad latente de quienes son examinados. También conocidos como AFC para ítems categóricos (o IFA aplicado a contextos educativos).

## Parámetros clave (Modelo 2PL):

- **Dificultad ( $b_i$ )**: Nivel de habilidad necesario para tener 50% de probabilidad de respuesta correcta
- **Discriminación ( $a_i$ )**: Qué tan bien el ítem diferencia entre niveles de habilidad

## Ventajas sobre Teoría Clásica de Tests:

- Estimaciones de habilidad independientes de los ítems específicos
- Permite pruebas adaptativas (CAT)

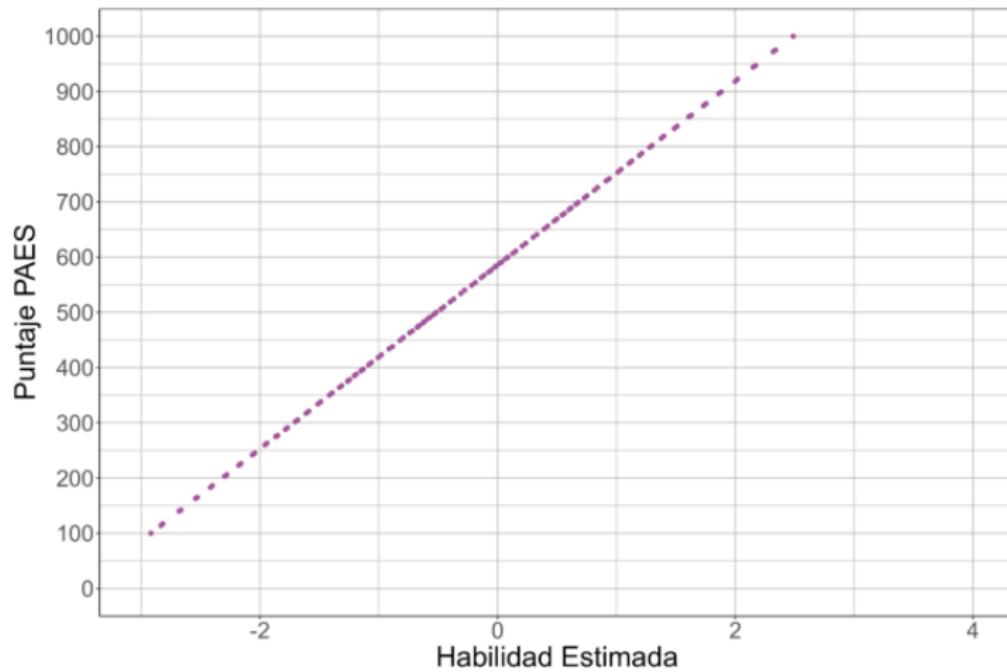
## Ejemplo: Modelo Rasch (1PL) en PAES 2024

El puntaje de un estudiante  $s$  en la prueba de Competencia Matemática M1 está determinado por su habilidad latente  $\theta_s$  (capturado por su patrón de respuestas) y las características de los ítems en el test (en PAES, dificultad).

El puntaje informado va entre 100 y 1000, y es una transformación lineal de la habilidad estimada  $\theta_s$ .

# Ejemplo: Modelo Rasch (1PL) en PAES 2024

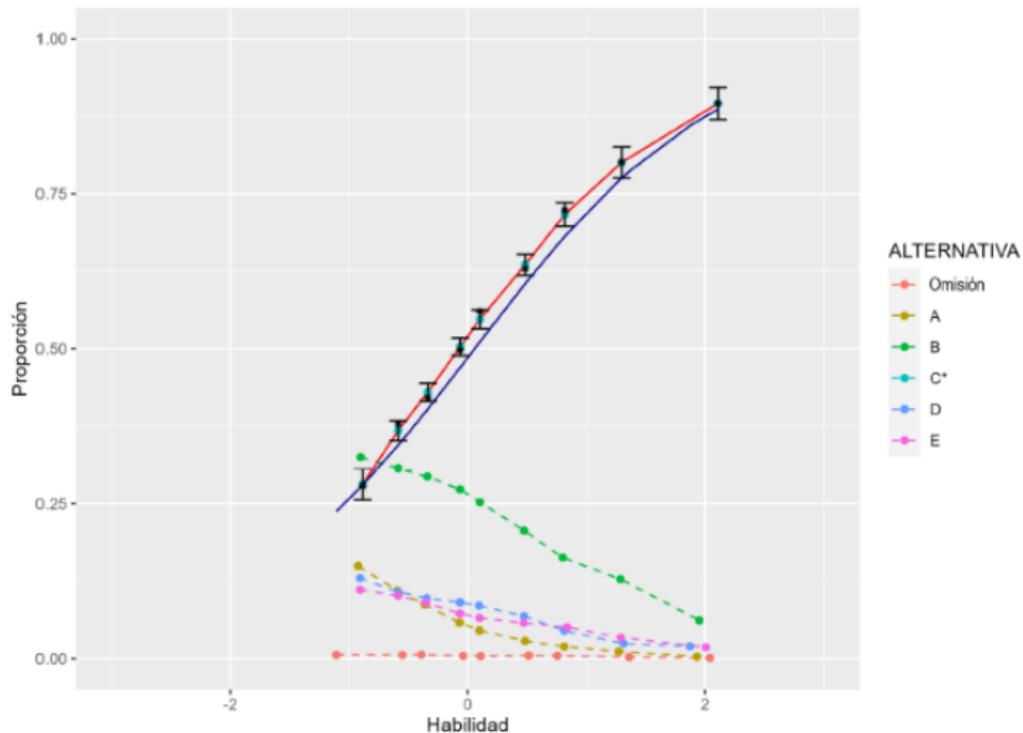
## Puntaje PAES para la prueba de Competencia Lectora



Extraído de  
Informe de  
Resultados PAES  
Regular Admisión  
2024 (p. 50).

Figura 40: Puntajes PAES versus estimación de habilidad en Competencia Lectora.

# Ejemplo: Modelo Rasch (1PL) en PAES 2024



Extraído de  
Informe de  
Resultados PAES  
Regular Admisión  
2024 (p. 89).

Figura 100: Ejemplo de curva característica de ítem (línea continua-corresponde a la clave) creciente en todo el rango de habilidad. Las curvas decrecientes corresponden a las alternativas incorrectas (distractores).

# **Ejemplos: Resultados Educativos**

# Pruebas estandarizadas SIMCE (Sistema de Medición de la Calidad de la Educación)

## Características:

- Evaluación censal nacional en Chile (4°, 6°, 8° básico, II medio)
- Mide aprendizajes del currículo nacional vigente en Lectura y Matemáticas
- Incluye cuestionarios de contexto (estudiantes, padres, profesores, directores)
- Clasificación en tres niveles: insuficiente, elemental, adecuado

## Análisis típicos:

- Tendencias temporales de resultados
- Brechas de equidad (NSE, género, ruralidad)
- Relaciones entre factores escolares y resultados

# IDPS (Indicadores de Desarrollo Personal y Social)

## Características:

- Complementan SIMCE desde 2014
- Miden dimensiones no académicas: autoestima, convivencia escolar, participación, vida saludable
- Basados en cuestionarios de contexto

## Análisis típicos:

- Similares a los test estandarizados SIMCE

**Relevancia:** Contribuyen a la medición integral de calidad educativa.

# Desafíos en el análisis de pruebas SIMCE e IDPS

## Potenciales problemas metodológicos:

- **Nivel de análisis:** ¿Individual o agregado (estudiantes vs escuelas)?
- **Validez de constructo:** ¿Los IDPS capturan adecuadamente lo que pretenden medir?
- **Unidimensionalidad:** ¿El rendimiento en matemáticas solo está vinculado con el conocimiento matemático?

# PAES (Prueba de Acceso a la Educación Superior)

## Características:

- Reemplazó a la PSU luego de las pruebas de transición en 2022
- Pruebas obligatorias: Competencia Lectora y Competencia Matemática M1
- Pruebas electivas: Matemática M2, Ciencias, e Historia
- Puntaje: entre 100 y 1000 por prueba

## Análisis relevantes:

- Poder predictivo del rendimiento universitario
- Equidad en el acceso (ej: brechas por tipo de establecimiento)
- Validez predictiva para el rendimiento académico

# Críticas al uso de PAES

## Debates actuales:

- **Sesgo socioeconómico:** Estudiantes de mayor NSE obtienen mejores puntajes
- **Validez diferencial:** ¿Predice igual en todas las carreras?
- **Reducciónismo:** ¿Una prueba puede capturar las competencias para la educación superior?
- **Preparación:** Preuniversitarios pueden generar inequidad

# Consideraciones finales

- **Pensamiento crítico:** Evaluar estudios cuantitativos requiere examinar múltiples dimensiones: diseño, muestra, medidas, análisis e interpretación.
- **El diseño importa:** El tipo de estudio determina qué preguntas se pueden responder (descripción, correlación, causalidad).
- **Dos tipos de significancia:** Estadística y práctica.
- **El contexto es clave:** La interpretación de resultados debe considerar el contexto teórico, metodológico y práctico.

# Preguntas y comentarios

Bladimir Padilla (gbpadilla@uiowa.edu) & Ariel Aloe

**University of Iowa**