

## ANÁLISIS ESTADÍSTICO DE UN TEST DE CONOCIMIENTOS PREVIOS DE MATEMÁTICAS PARA INGRESANTES UNIVERSITARIOS

Nélida H. Pérez, María A. Mini y Julio Benegas  
Universidad Nacional de San Luis, Argentina  
[nperez@unsl.edu.ar](mailto:nperez@unsl.edu.ar) - [jbenegas@unsl.edu.ar](mailto:jbenegas@unsl.edu.ar)

### Resumen

En este trabajo presentamos un análisis estadístico del Test de Conocimientos Previos de Matemáticas (TCPM) diseñado para medir el estado inicial de destrezas y conocimientos básicos en matemáticas de los alumnos ingresantes a carreras científico- tecnológicas de la Facultad de Ciencias Físico, Matemáticas y Naturales de la Universidad Nacional de San Luis.

El objetivo de la investigación está centrado en observar el diagnóstico utilizado, con miras a una eventual utilización posterior.

Para determinar la bondad de la prueba realizamos un análisis de la calidad, discriminación e índice de dificultad de los ítems, así como de la validez y confiabilidad del diagnóstico, para este análisis estadístico empleamos los programas TestGraf y SPSS.

El test se aplicó a 698 estudiantes ingresantes a la Universidad en el ciclo lectivo 2002. De la investigación pudimos inferir que el diagnóstico resultó: *difícil* para la población de aplicación; de *confiabilidad aceptable*, y de *buena calidad de ítems*, con variada dificultad y aceptable discriminación.

### Introducción

La Universidad Nacional de San Luis (UNSL) no escapa del preocupante problema de la preparación inicial de los estudiantes que pretenden ingresar a las distintas carreras de ciencias e ingeniería. Diversas acciones se han programado desde hace tiempo para tratar de mejorar el desempeño estudiantil y bajar el porcentaje de fracaso y deserción en el primer año de vida universitaria.

Una de esas acciones consiste en determinar el nivel de conocimientos y las habilidades con que los estudiantes arriban a la Universidad, con el objetivo de poder identificar, tan pronto como sea posible, las falencias de conocimientos, y su importancia relativa, en la población en riesgo. A tal efecto se elaboró un diagnóstico de respuestas múltiples para medir los conocimientos matemáticos básicos de los alumnos ingresantes. Luego de aplicada la prueba creímos imprescindible realizar un pormenorizado análisis de resultados, de manera de poder capitalizar la experiencia, tanto en esta como en otras instituciones. A tal efecto nos propusimos dos objetivos:

1°. Determinar la calidad general del diagnóstico utilizado, realizando un análisis estadístico pormenorizado de su aplicación.

2°. Evaluar el nivel de conocimientos matemáticos básicos de los ingresantes y relacionarlo con el sistema educativo regional. Considerando que el 90% de población encuestada es de la provincia de San Luis, y constituye la primera promoción que ha tenido una implementación completa de la nueva ley de educación general básica.

Este segundo objetivo es objeto de un estudio complementario y separado del presente trabajo.

### Construcción del diagnóstico

El Test de Conocimientos Previos de Matemática (TCPM) se diseñó para medir el estado inicial de conocimientos básicos y aptitudes en matemáticas de los alumnos

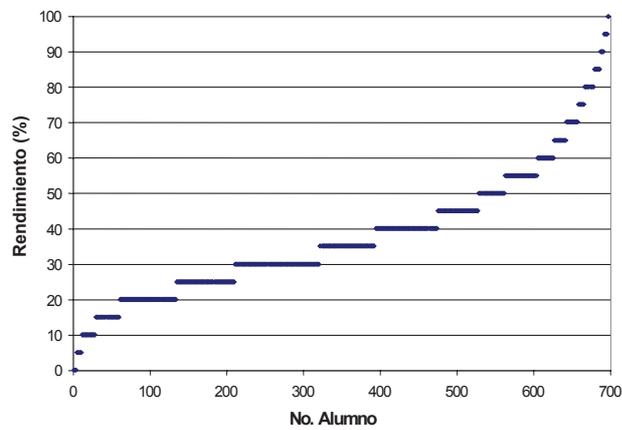
ingresantes a las carreras científico-tecnológicas de nuestra universidad. Los temas incluidos corresponden a una selección tomada del currículum del Tercer nivel de la Educación General Básica (EGB3) y del nivel Polimodal de la escuela pública Argentina.

El TCPM contiene 20 ítems, con cuatro opciones de respuesta posible cada uno. Las preguntas y sus distractores fueron propuestos por un equipo de docentes con experiencia en la enseñanza de la matemática en el ámbito universitario y en diversos programas de interacción con los profesores de EGB3 y Polimodal. (En ANEXO, prueba completa con ítems agrupados por tema).

**Análisis estadístico del diagnóstico**

La Figura 1 representa los resultados obtenidos de la aplicación del test a una muestra de 698 estudiantes (aproximadamente el 80% de la población total). Los encuestados son alumnos ingresantes (sin instrucción universitaria previa al diagnóstico) a las carreras de: Licenciatura y Prof. en Ciencias de la Computación, Ingeniería Electrónica, Licenciatura y Prof. en Matemáticas y Licenciatura y Prof. en Física.

Se observa que el rendimiento general es bajo con un valor porcentual de la media para toda la muestra de 38 %, con una desviación estándar de 45%. Desde el punto de vista del diseño de pruebas objetivas, este valor medio es demasiado bajo, mientras que la amplia desviación estándar es adecuada para discriminar entre poblaciones de distinto rendimiento.



**Figura 1:** Rendimiento (en %) en el TCPM de los ingresantes 2002

De la Figura 1 se desprende que alrededor de 220 alumnos apenas alcanzan el valor estadístico de contestar al azar (25%), mientras un total de 560 alumnos (80 % del total) tienen sólo la mitad o menos de las preguntas contestadas correctamente. Dada la escasa complejidad de los ítems de este diagnóstico, la primera conclusión es que esta población estudiantil carece del conocimiento y de las habilidades matemáticas mínimas para afrontar un curso universitario inicial de cálculo o álgebra.

**a) Características generales de los ítems**

Los parámetros estadísticos que caracterizan el diagnóstico están sintetizadas en la Tabla y han sido obtenidos a través de la aplicación de los programas estadísticos TestGraf [5] y SPSS[7]. Los valores de *Coeficiente Biserial*, *Discriminación* y *Dificultad* se refieren a parámetros calculados por el programa TestGraf [5] de análisis gráfico de diagnósticos de respuestas múltiples.

La mayoría de los ítems resultaron difíciles. Solo en 5 ítems la población encuestada supera el 50% de rendimiento, mientras que en ocho ítems se obtuvo rendimiento

comparable con la respuesta azarosa (cuatro opciones por pregunta). En todos los casos la desviación estándar,  $\sigma$ , es próxima a la mitad del rango total, lo cual indica una buena dispersión de los valores obtenidos.

Tema	Ítem	Media	Desvío Estándar	Coef. Biserial	Discrimi-nación	Dificultad
<b>ARITMÉTICA:</b> <i>Números Operaciones Proporciones</i>	<b>1</b>	<b>0.72</b>	<b>0.45</b>	<b>0.36</b>	<b>0.87</b>	<b>-0.18</b>
	<b>2</b>	0.61	0.49	0.43	0.73	-0.04
	<b>3</b>	0.49	0.50	0.42	0.62	0.21
	<b>4</b>	0.40	0.49	0.46	0.79	0.57
	<b>5</b>	0.26	0.44	0.31	0.5	1.68
	<b>18</b>	0.69	0.46	0.26	0.45	-0.44
<b>EXPRESIONES ALGEBRAICAS</b>	<b>6</b>	0.17	0.38	0.51	1.24	1.35
	<b>7</b>	0.19	0.39	0.57	1.07	1.25
	<b>8</b>	0.46	0.50	0.36	0.68	0.66
	<b>9</b>	0.13	0.34	0.23	2.04	2.16
<b>GEOMETRÍA</b>	<b>10</b>	0.35	0.48	0.45	0.75	0.83
	<b>11</b>	0.29	0.45	0.27	0.47	1.83
	<b>17</b>	0.33	0.47	0.51	0.87	0.77
<b>ECUACIONES de PRIMER GRADO</b>	<b>12</b>	0.66	0.47	0.37	0.84	-0.19
	<b>14</b>	0.55	0.50	0.33	0.55	0.24
<b>OPERACIONES CON POLINOMIOS y RAÍCES de ECUACIONES</b>	<b>13</b>	0.24	0.43	0.38	0.65	1.35
	<b>15</b>	0.27	0.45	0.24	0.55	1.65
	<b>16</b>	0.28	0.45	0.47	0.84	0.93
<b>TRIGONOMETRIA</b>	<b>19</b>	0.27	0.44	0.52	0.88	0.93
	<b>20</b>	0.22	0.41	0.38	0.61	1.54

Tabla: Resultados estadísticos de los 20 ítems del TCPM (ver texto).

El coeficiente de *punto biserial* representa la correlación, sobre toda la muestra, entre el ítem y el resultado global del test. Valores del coeficiente punto biserial de 0,20 o mayores se consideran aceptables. En nuestro caso tienen un coeficiente biserial menor de 0,30 solo ítems que son muy fáciles o muy difíciles, es decir aquellos que se espera no tengan una buena correlación con el resultado global del test.

La *discriminación* se refiere a cuan efectivamente un ítem distingue entre sujetos de bajo y alto rendimiento. El parámetro discriminación, está determinado por la pendiente de la curva de la opción correcta en la zona de probabilidad media.

El valor del índice *dificultad* está relacionado con el rendimiento global de los alumnos que alcanzan el valor medio en ese ítem. Se mide en unidades de la desviación estándar, a partir del rendimiento medio de la población. El desempeño o rendimiento así determinado constituye el eje horizontal de las figuras 2 y 3.

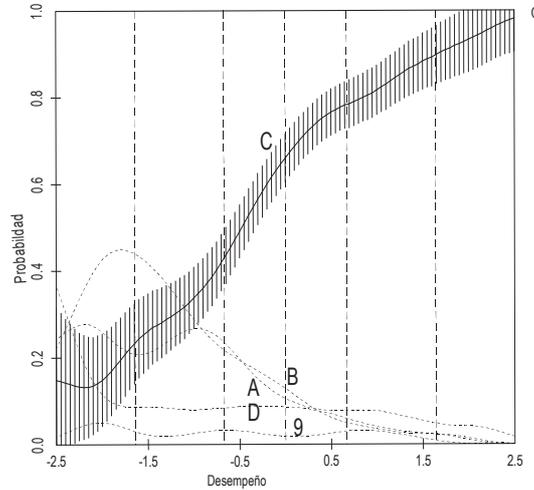
En la Tabla podemos observar que sólo 4 ítems tienen dificultad negativa (es decir son fáciles para esta población), mientras que en todos los restantes es positiva.

Para ejemplificar lo anterior y mostrar los atributos del software usado para un eventual uso en docencia o similar al presente, presentamos un análisis más detallado de los ítems 2 y 9.

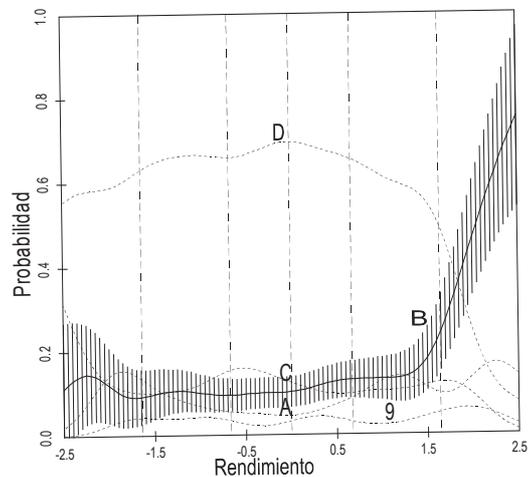
En la Figura 2 están representadas las curvas de probabilidad de elegir cada opción del ítem 2 en función del rendimiento o desempeño global del estudiante. El eje horizontal mide el desempeño en unidades de la desviación estándar, con el cero u origen en el valor medio de la población (es decir 38% de repuestas contestadas correctamente). A las izquierda de las líneas verticales de trazos se encuentran los estudiantes con rendimiento por debajo del 5%, 25%, 50%, 75% y 95 %, respectivamente

Se observa que la curva de la respuesta correcta (opción C) tiene valores de probabilidad importantes, aún para poblaciones de modesto desempeño global. Es por tanto una pregunta fácil, tal cual lo denotan los respectivos índices en la Tabla.

Además la curva de la Opción C es monótona creciente, con buena pendiente en casi todo el rango y por lo tanto separa (discrimina) adecuadamente las poblaciones de distinto rendimiento. De igual modo el análisis de las curvas de los distractores (opciones incorrectas) nos indica que estas opciones son atractivas solo para los



**Figura 2:** Curvas de probabilidad de elección de cada una de las opciones de la Pregunta No.2. La opción correcta (C) se muestra con la banda de confianza.



**Figura 3:** Curvas de probabilidad de la pregunta No.9

alumnos de rendimiento inferior al 25%, mientras que la curva señalada con “9”, que identifica a los alumnos que no responden la pregunta, tiene probabilidad prácticamente nula a lo largo de todo el eje de rendimiento.

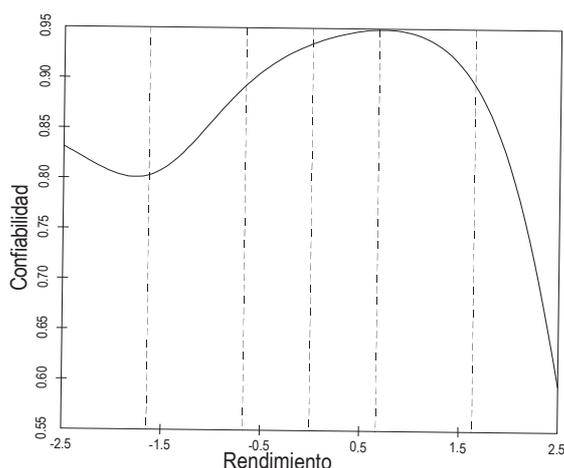
En la Figura 3 se representan las curvas de rendimiento del ítem 9.

Este es el ítem de mayor dificultad del test, medido tanto en función del desempeño promedio (sólo el 13 % lo contestó bien), como por el índice de dificultad. Se ve que sólo los alumnos con rendimiento en el 5% superior de la población responden el ítem correctamente con alguna probabilidad significativa. El ítem discrimina bien, pero sólo en esa zona de alto rendimiento, mientras que para el resto de la población no discrimina.

Un análisis por opción en este ítem revela que la opción D es la preferida por los estudiantes en casi todo el rango de rendimiento. Si bien a efectos del diagnóstico la pregunta no parece de calidad suficiente, puesto que solo discrimina en el 5% superior de la población, desde el punto de vista de la instrucción es importante, ya que la elección de un distractor particular revela conductas de trabajo o errores sistemáticos de la población.

Es importante destacar que al repetir este análisis en el resto de los ítems, la opción “no contesta”, identificada con el número “9” en las gráficas, ocurre solamente en los alumnos con rendimiento total inferior al 25%, lo que nos indica que las preguntas sin respuesta son producto de la ausencia de conocimiento en el tema involucrado y no de cansancio u otras causas. Por otro lado algunos distractores no son significativamente elegidos por los estudiantes (de cualquier nivel de rendimiento). Estos distractores deberían ser mejorados para aumentar la calidad global del diagnóstico.

A partir de un análisis similar realizado sobre todos los ítems, se encuentra que el diagnóstico en general tiene una adecuada variabilidad de dificultad de los ítems, valores aceptables de discriminación y buenos coeficientes punto biserial, lo cual permite afirmar que calidad global de los ítems del TCPM es aceptable.



**Figura 4:** Confiabilidad del TCPM para la población estudiada, según TestGraf (ver texto)

### b) Validez y confiabilidad del diagnóstico

Validez y confiabilidad son dos características que definen la calidad de un diagnóstico. En general la validez se establece a través del uso y la opinión de docentes y alumnos.

Las preguntas y sus distintas opciones de respuesta fueron propuestas por docentes dedicados a la enseñanza en cursos de introducción a la matemática universitaria y a

la formación y perfeccionamiento de docentes secundarios. Se tuvieron en cuenta pruebas anteriores y la experiencia del cuerpo de docentes. Posteriormente un grupo de docentes, distinto al que propuso las preguntas, también experimentado en la enseñanza de la matemática universitaria básica, analizó la pertinencia de las preguntas y sus respectivas opciones de respuesta. Este proceso de selección y corrección condujo a la versión del TCPM que se informa en este trabajo.

La confiabilidad de un diagnóstico es una medida de cuan consistentemente el test reproducirá el mismo resultado bajo las mismas condiciones. Las técnicas para establecer la confiabilidad son de tipo estadístico. Las más utilizadas en educación son la determinación del coeficiente  $\alpha$  de Crombach [4] y del coeficiente denominado KR-20, que corresponde a la fórmula 20 de un trabajo de Kuder y Richarson [6]. El coeficiente de confiabilidad varía entre 0 y 1. En pruebas estandarizadas se considera aceptable un valor cercano a 0.8, mientras que en pruebas construidas por los docentes se considera aceptable un valor mínimo de 0.60.

El análisis de confiabilidad del Test de Conocimientos Previos en Matemática se realizó mediante el programa estadístico SPSS [7], que determinó un coeficiente de confiabilidad  $\alpha=0,73$ . También se realizó un análisis por mitades del diagnóstico (split-half), donde aplicando la corrección de Spearman-Brown [3], se obtiene una confiabilidad de **0,74**. Estos resultados aseguran una aceptable confiabilidad del diagnóstico.

TestGraf [2] tiene un método quizás más completo para calcular la confiabilidad, ya que lo hace de manera dinámica, en función del rendimiento estudiantil.

En la Figura 4 se observa que la confiabilidad es máxima para estudiantes de rendimiento alrededor de 75% (4,5 puntos de un total de 10, para el presente ejemplo) y se mantiene buena para rendimientos entre 25% y 80%, aproximadamente, indicando su adaptabilidad para estudios como el presente.

### Conclusiones

Nuestra metodología de análisis, apoyada con la utilización conjunta de los paquetes estadísticos SPSS y TestGraf, nos permitieron alcanzar las siguientes conclusiones sobre el TCPM. (Test de Conocimientos Previos en Matemática, aplicado a una muestra de 698 estudiantes ingresantes a carreras científicas y de ingeniería de la Universidad Nacional de San Luis).

- a) El diagnóstico resultó difícil para la población que fue aplicado, que alcanzó una media de 38% de rendimiento, con un error estándar de 45% .
- b) La confiabilidad de diagnóstico, medido según el coeficiente  $\alpha$  de Crombach es buena, indicando una aceptable consistencia interna del instrumento. Las curvas dinámicas determinadas por TestGraf indican un máximo de la confiabilidad en la zona media de rendimiento.
- c) Los ítems del diagnóstico son en general de buena calidad, con valores del coeficiente punto biserial aceptables y una buena discriminación.
- d) Del análisis dinámico de las curvas de probabilidad de las opciones de cada ítem, en función del rendimiento colectivo, surge que algunos distractores pueden ser mejorados, ya que no son elegidos significativamente por los alumnos.

- e) Los resultados generales del diagnóstico indican que puede usarse apropiadamente para determinar el estado inicial de conocimientos básicos de matemáticas a los efectos de programar la instrucción.

### **Bibliografía**

- Otero M.R, Fanaro M. A. Y Elichiribehety I., “El conocimiento matemático de los estudiantes que ingresan a la universidad” *Revista Latinoamericana de Investigación en Matemática Educativa*, 4(3), p. 267-287, 2001.
- Landazábal Ma.C., Bilbao F., Otero J. y Caballero C., “Formación inicial y rendimiento en Física del primer curso universitario”. *Revista de Educación (Madrid)*, en prensa, 2003.
- Aubrecht G and Aubrecht J. “Constructing objective test” *Am. J. of Phys.* 51, 613-620, 1983.
- Baranger, D. “Construcción y Análisis de Datos”, Ed. Universitaria, UNM, 1992.
- TestGraf, J.O. Ramsay, McGill University, Canada, 1995. Disponible en <ftp://ego.psych.mcgill.ca/pub/ramsay/testgraf/>.
- Loewenthal K. M. “An Introduction to Psychological Test and Scales”, 2<sup>nd</sup> edition, Psychological Press, Taylor & Francis, 2001.
- SPSS, “Statistical Package for the Social”, SPSS Inc. Chicago, IL. USA, 1999.