

---

# Propuesta para la construcción de un Índice Socioeconómico para los estudiantes que presentan las pruebas SABER PRO

**Proposal of a socio-economic index for students who present SABER PRO tests**

Edwin Javier Cuellar Caicedo<sup>a</sup>  
ecuellar@contratista.icfes.gov.co

Stalin Guerrero<sup>b</sup>  
sguerrero@contratista.icfes.gov.co

Daniela Lopez<sup>c</sup>  
daniela.lopez@usantotomas.edu.co

---

## Resumen

El artículo presenta una propuesta para la construcción de un índice de nivel socioeconómico de estudiantes de educación superior, en el contexto de pruebas estandarizadas. Para ello, se utiliza la información del Instituto Colombiano para la Evaluación de la Educación (ICFES), correspondiente a 245.461 estudiantes que participaron en el examen SABER PRO en el año 2012, junto con la información del cuestionario sociodemográfico. Una revisión de la literatura permite definir las variables y dimensiones que conformarán el nivel socioeconómico de los estudiantes. La metodología utilizada para este estudio es el Análisis de Componentes Principales (ACP), el cual nos permite reducir la dimensionalidad de la información contenida en las variables. Dado que la mayoría de variables son categóricas, se propone usar el método de asignación óptima (De Leeuw & Mair 2007) con el fin de atribuirles valores que permitan analizar dichas variables de manera cuantitativa. Se encuentra que el índice construido se comporta bien frente a algunas variables de interés y tiene una relación positiva con los puntajes de las pruebas de las competencias genéricas de SABER PRO.

**Palabras clave:** Índice Socioeconómico, Asignación Óptima, Análisis de Componentes Principales (ACP), Métodos de Gifi.

## Abstract

---

<sup>a</sup>Estadístico, Unidad de Investigaciones, ICFES.

<sup>b</sup>Estadístico, Subdirección de estadísticas, ICFES.

<sup>c</sup>Estudiante de estadística, Universidad Santo Tomás.

This paper presents a proposal for the construction of an index to measure the socioeconomic status of students of higher education, in the context of standardized tests. In this study, we used information of Colombian Institute for the Evaluation of Education (ICFES) corresponding to 245.461 students who participated in SABER PRO tests in 2012 and also the information of sociodemographic questionnaire. A review of the literature allows to define the variables and components that compose the socioeconomic status of students. Principal Component Analysis (PCA) is used to summarize information and reducing the dimensionality of data. Since there are several categorical variables in the socioeconomic infomation, the method of optimal allocation (De Leeuw & Mair 2007) is proposed in order to assign weights to analyze quantitatively those variables. The results show a good performance of the index against some variables of interest and a positive relation is found between socioeconomic index and scores in SABER PRO tests.

**Keywords:** Socio-economic Index, Principal Component Analysis, Optimal Allocation , Gifi methods.

## 1. Introducción

En las últimas décadas Colombia ha sufrido múltiples procesos de transformación social, económica y política, que entre otras cosas, han impulsado al sistema de educación superior (MEN 2015). Conocer el contexto socioeconómico en el que se desenvuelven los estudiantes universitarios de Colombia constituye un factor muy importante para tener un mejor panorama de la calidad de la educación superior del país. El Examen de Estado de Calidad de la Educación Superior, SABER PRO, es una prueba estandarizada externa para evaluar la calidad de la educación superior y es obligatoria desde 2009<sup>1</sup>. Esta prueba evalúa cinco competencias genéricas (lectura crítica, razonamiento cuantitativo, comunicación escrita, inglés y competencias genéricas) y competencias específicas comunes para distintos grupos de referencia<sup>2</sup>.

La presente investigación busca medir el nivel socioeconómico de los estudiantes universitarios, con el fin de proveer un insumo que sirva para los procesos de evaluación y caracterización de los estudiantes de la educación superior en Colombia.

En la revisión de literatura se encontró que son escasos los estudios que se tienen en este campo en Colombia, lo que hace relevante la construcción un índice compuesto por un conjunto de factores económicos, sociales y culturales, conocidos como variables no observables o latentes (Díaz & Barrios 2002), pero que pueden

<sup>1</sup>Ley 1324 del 13 de julio del 2009 y Decreto 3963 de 2009.

<sup>2</sup>Los grupos de referencia son conjuntos de programas académicos que han sido construidos para la interpretación de resultados de la evaluación, y se basan en la clasificación de áreas de conocimiento y núcleos básicos de conocimiento del SNIES (Sistema Nacional de Información de la Educación Superior), así como en la clasificación de formaciones de la UNESCO.

ser aproximadas a través de variables observables, como los ingresos familiares mensuales, la dotación del hogar, el contexto familiar, entre otras.

La base fundamental para la construcción de un índice es la definición de las variables que lo componen. En la literatura es muy común encontrar investigaciones en educación que proponen índices socioeconómicos para estudiantes de educación básica y media (Gil-Flores 2013, Paz-Navarro et al. 2009, Porcel et al. 2010, Saz 2006), y que se usan como referencia en la presente investigación.

El nivel socioeconómico de los estudiantes es un factor de gran importancia en el contexto de la educación y se relaciona con su desempeño académico (Acevedo & Jaramillo 2007, Armenta et al. 2008, Caso-Niebla & Hernandez 2007, Contreras et al. 2008). La ocupación y el nivel educativo de los padres son variables que suelen usarse en este contexto. (Vargas 2013, Saz 2006). De igual forma, en la definición del nivel socioeconómico, además de las variables que estiman el nivel económico, se han incluido aquellas que reflejan la posesión de un capital cultural (Jabnou 2009, Saz 2006, Hernández & González 2011).

## 2. Metodología estadística

### 2.1. Análisis de Componentes Principales (ACP)

La metodología de ACP permite reducir la dimensión de los datos, preservando el máximo de información posible. De esta manera, se puede reducir un sistema de dimensión  $q$  a un sistema de dimensión menor, a través de la generación de nuevas variables (componentes) resultantes de la combinación lineal (sumas ponderadas) de las variables originales.

En un estudio realizado sobre  $n$ -individuos mediante  $p$ -variables  $X_1, \dots, X_p$ , es posible encontrar nuevas variables notadas por  $Y_k$  que sean combinaciones lineales de las variables originales  $X_j$ , y sujetas a ciertas condiciones. En este sentido se determina la primera componente principal  $Y_1$ , la cual contiene la mayor cantidad de variabilidad total contenida en los datos. De esta forma:

$$Y_1 = \gamma_{11}X_1 + \gamma_{12}X_2 + \dots + \gamma_{1p}X_p$$

donde las ponderaciones  $\gamma_{11}, \dots, \gamma_{1p}$  se escogen de tal forma que maximicen la razón de la varianza de  $Y_1$  a la varianza total; con la restricción:  $\sum_{j=1}^p \gamma_{1j}^2 = 1$ .

Por otra parte, el criterio sugerido para seleccionar el número de componentes se basa en la variabilidad que se quiere mantener. Dado que la suma de las varianzas originales, es la traza de la matriz de varianzas y covarianzas  $\mathbf{S}$  y es igual a la suma

de los valores propios de  $\mathbf{S}$ , cada componente principal explica una proporción de la variabilidad total, la cual se calcula de la siguiente manera:

$$\frac{l_k}{tra(\mathbf{S})}$$

Este cociente se denomina la proporción de la variabilidad total explicada por la  $k$ -ésima componente.

El uso del ACP para la construcción de índices en diferentes ámbitos, ha presentado ventajas sobre otras metodologías, debido a su fácil implementación y a que se pueden obviar los supuestos de normalidad, homocedasticidad y linealidad (Estevez 2002).

## 2.2. HOMALS

Este método hace parte de los denominados métodos Gifi y también se conoce como análisis de homogeneidad. Este método cuenta con múltiples funciones aplicables al análisis de datos multivariados. Sin embargo, para efectos de este estudio, sólo se tuvo en cuenta como una herramienta de asignación óptima, que permite básicamente asignar un puntaje cuantitativo a las variables categóricas de tipo nominal u ordinal por medio de un procedimiento iterativo (De Leeuw & Mair 2009).

Siguiendo la notación de (De Leeuw & Mair 2009), para  $i = 1, \dots, n$  individuos en  $j = 1, \dots, m$  variables categóricas con  $K_j$  niveles diferentes de respuesta, se construye un bloque de matrices  $G = [G_1, \dots, G_m]$ , donde  $G_j$  son matrices indicadoras dummy de dimensión  $n \times k_j$ . Además se define la matriz diagonal binaria  $M_j$  de dimensión  $n \times n$  para cada variable  $j$ . Los elementos de la diagonal  $(i, i)$  toman el valor 0 cuando el individuo  $i$  tiene un valor faltante en la variable  $j$ , o 1 en otro caso.

Después se consideran las matrices desconocidas  $\mathbf{X}$  ( $n \times p$ ) y  $Y_j$  ( $k_j \times p$ ) las cuales contienen las puntuaciones de los objetos y de las categorías respectivamente. Para encontrar dichos valores, se establece la siguiente función de pérdida:

$$\sigma(\mathbf{X}; Y_1, \dots, Y_m) = \frac{1}{m} \sum_{j=1}^m tr((\mathbf{X} - G_j Y_j)' M_j (\mathbf{X} - G_j Y_j))$$

Ambas matrices deben minimizar simultáneamente esta función por medio de la normalización de  $u' M_j \mathbf{X} = 0$  y  $X' M_j \mathbf{X} = I$ . Desde un punto de vista teórico, se concluye que la función de pérdida representa la suma de cuadrados de  $(\mathbf{X} - G_j Y_j)$ , la cual involucra los puntajes obtenidos para los individuos y las categorías.

En el lenguaje de programación R<sup>3</sup>, se encuentra disponible la función *homals* del paquete *homals* (De Leeuw & Mair 2007), la cual realiza este proceso iterativo, de manera que asigna los puntajes óptimos a cada categoría de las variables.

### 3. Construcción conceptual

En la construcción del Índice de Nivel Socioeconómico (INSE) se utilizó la información suministrada por el ICFES, correspondiente a 245.461 estudiantes de Colombia que participaron en el Examen de Calidad de la Educación Superior, SABER PRO y la información del cuestionario sociodemográfico, compuesto por 64 preguntas que buscan conocer algunos aspectos relacionados con su contexto familiar, social y económico. Cabe resaltar que estos estudiantes se encontraban próximos a culminar su ciclo de formación superior, con el requisito de haber aprobado al menos el 75 % de los créditos académicos de su respectivo programa de pregrado para el año 2012.

El INSE se construye a partir de tres dimensiones: *potencial económico del hogar*, *capital cultural* y *dotación de la vivienda*. Para la dimensión *potencial económico del hogar*, se tienen en cuenta tres variables *proxy* relacionadas con los ingresos económicos del hogar, estas variables son: *valor de la matricula*, *ingreso familiar mensual* y *SISBEN*<sup>4</sup>, todas ellas ordinales.

En cuanto a la dimensión *capital cultural*, se tienen en cuenta cuatro variables: educación de la madre y educación del padre, ocupación del padre y ocupación de la madre. Las variables de educación presentan categorías de respuesta como: primaria completa, primaria incompleta, secundaria incompleta, secundaria completa, técnico incompleto, técnico completo, universitario incompleto, universitario completo, posgrado, entre otros. Estas son consideradas variables ordinales y son de gran importancia si se tiene en cuenta que la educación de los padres es una variable que tiene una correlación alta con la educación de los hijos. Esto debido a las dificultades de movilidad y a la herencia intra-generacional (Nuñez n.d.). En cuanto a la ocupación de los padres, entre las posibles categorías de respuesta se encuentran: estudiante, jubilado, hogar, empresario, trabajador independiente, entre otras. Las anteriores variables se consideran importantes debido a la relación que existe entre el capital cultural y el desempeño académico. (Coleman 1988) (Gil-Flores 2013).

La última dimensión corresponde a *dotación de la vivienda* y es construida a partir de la información de la tenencia o no de ciertos elementos que podrían ser importantes o marcar la diferencia entre las condiciones socioeconómicas de

---

<sup>3</sup>Es un software libre y con entorno de desarrollo integrado(IDE) para el lenguaje de programación R , para computación estadística y gráficos.

<sup>4</sup>Sistema de Potenciales Beneficiarios para programas sociales

vivienda de un estudiante u otro. En estudios similares a este se han utilizado elementos de dotación o disponibilidad de espacios como una aproximación a la carencia de elementos físicos necesarios para el aprendizaje del alumno (Piñeros & Rodriguez 1998) (Ravela 2005).

Para la aplicación del ACP es necesario hacer uso únicamente de variables numéricas, sin embargo, las variables definidas dentro de las dimensiones del índice corresponden a variables categóricas. Por esta razón, la transformación de las mismas en numéricas se presenta como un requisito fundamental para el análisis y construcción del índice. Se decide aplicar el método de asignación óptima a todas las variables, con el fin de dar un puntaje cuantitativo a cada categoría, teniendo en cuenta que algunas son de tipo nominal u ordinal. Este procedimiento se lleva a cabo usando el lenguaje de programación R con ayuda de la función *homals*.

Con las variables categóricas transformadas a numéricas, se aplica el ACP a cada una de las dimensiones, con el fin de resumir dichas variables en componentes que retengan la mayor cantidad posible de información (varianza máxima). Con el fin de retener la cantidad de componentes necesarias, se realiza el Análisis Paralelo de Horn (1965), bajo el cual, para cada dimensión, se encuentra que solo se debe retener una componente. Por último, el índice corresponde a la primera coordenada del ACP realizado con las 3 escalas obtenidas de cada una de las dimensiones.

## 4. Resultados

Bajo el esquema presentado anteriormente, se analiza la información y se realizan los procedimientos correspondientes. Se encuentra que para cada una de las dimensiones es suficiente con el uso de una única componente para la construcción de la escala. Esto indica que la variabilidad e información contenida en la primera componente es suficiente para representar la información que contienen las variables que participan en la construcción de cada una de las dimensiones. Para confirmar el uso de una única componente se hace uso del *Análisis paralelo de Horn* (Horn 1965).

En la Tabla 1 se muestran las funciones de densidad obtenidas para cada una de las dimensiones: *potencial económico del hogar, capital cultural y dotación de la vivienda*.

Además las coordenadas de las variables para cada dimensión se presentan en la Tabla 2, Tabla 3 y Tabla 4.

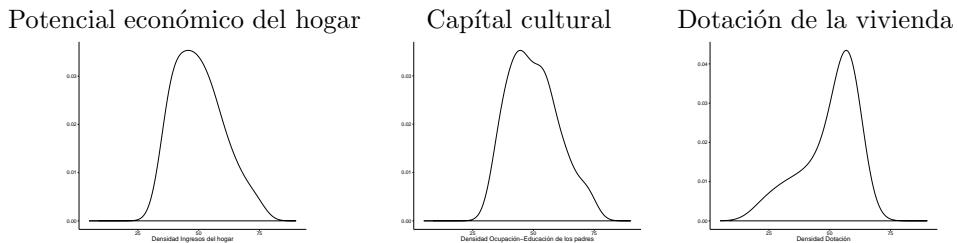


Tabla 1: Función de densidad para cada dimensión

	Dim.1	Dim.2	Dim.3
Valor matricula	0.77	-0.53	0.35
Ingreso familiar mensual	0.84	-0.06	-0.54
Nivel SISBEN	0.75	0.62	0.25

Tabla 2: Coordenadas por variable para la dimensión *Potencial económico del hogar*

	Dim.1	Dim.2	Dim.3	Dim.4
Nivel de educación de la madre	0.73	-0.36	-0.42	-0.41
Nivel de educación del padre	0.70	-0.52	0.29	0.39
Nivel de ocupación de padre	0.65	0.41	0.58	-0.27
Nivel de ocupación de la madre	0.65	0.55	-0.42	0.31

Tabla 3: Coordenadas por variable para la dimensión *Capital cultural*

	Dim.1	Dim.2	Dim.3	Dim.4
Internet	0.86	-0.14	-0.10	-0.49
Servicio de televisión	0.62	0.78	0.06	0.07
Telefonía	0.70	-0.27	0.64	0.18
Computador	0.76	-0.22	-0.51	0.33

Tabla 4: Coordenadas por variable para la dimensión *Dotación de la vivienda*

Dentro de los resultados obtenidos para el análisis por dimensión, se encuentra que la variabilidad explicada por la primera componente para cada dimensión *potencial económico del hogar*, *capital cultural* y *dotación de la vivienda* es de 62 %, 47 % y 55 %, respectivamente. En la Tabla 5 se encuentra la contribución de cada dimensión al INSE construido.

	Dim.1	Dim.2	Dim.3
Potencial económico del hogar	41.64	0.54	57.82
Capital cultural	26.45	60.51	13.04
Dotación de la vivienda	31.91	38.95	29.14

Tabla 5: Contribución de las dimensiones al *Índice socioeconómico*

De la Tabla 5 podemos concluir que el *potencial económico del hogar* es la dimensión que mayor contribución tiene dentro del INSE, seguida por la *dotación de la vivienda* y el *capital cultural*.

En la Figura 4, se encuentra la gráfica de densidad del INSE obtenido.

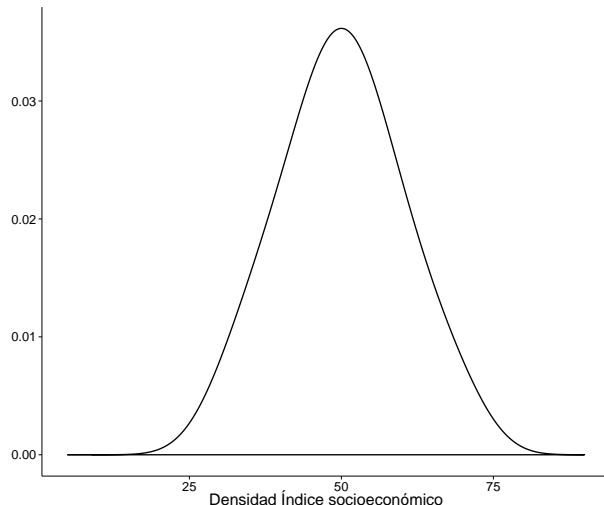


Figura 1: Distribución Índice Socioeconómico

De la Figura 4 es importante mencionar que no se presenta asimetría, el índice parece ser quasi-simétrico. Este comportamiento podría deberse a que las condiciones socioeconómicas de los estudiantes de educación superior en Colombia son más altas que las del total de la población colombiana, caso en el que se esperaría un índice con asimetría positiva. En la Tabla 6 se encuentran los estadísticos descriptivos que corresponden al INSE obtenido.

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
24.69	43.14	49.98	50.00	56.76	74.72

Tabla 6: Estadísticos descriptivos INSE

Con el fin de evaluar la validez externa del índice construido, es necesario analizar el comportamiento del INSE frente a algunas variables de interés como lo son el estrato socioeconómico y el origen de la institución educativa.

El estrato socioeconómico<sup>5</sup>, es una variable que se encuentra dentro del cuestionario pero que no fue introducida en la construcción del índice. En la Figura 4 se presenta la gráfica de densidad por estrato y se observa que niveles altos<sup>6</sup> del estrato están asociados con puntajes del INSE altos, mientras que los niveles bajos del estrato se asocian con puntajes del INSE bajos.

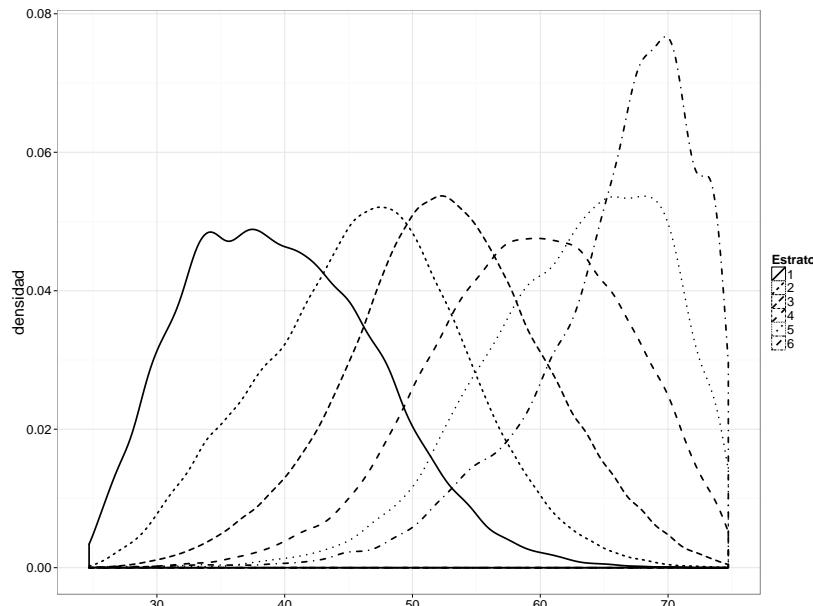


Figura 2: INSE por estrato socioeconómico

En la Figura 4 se presenta el comportamiento del INSE entre estudiantes de instituciones de educación superior de origen oficial y no oficial, encontrándose que los estudiantes de instituciones de origen no-oficial se asocian con puntajes del INSE relativamente más altos.

<sup>5</sup>Los estratos socioeconómicos son una herramienta que utiliza el Estado colombiano (Ley 142 de 1994, Artículo 102) para clasificar los inmuebles residenciales de acuerdo con los lineamientos del DANE, el cual tiene en cuenta el nivel de pobreza de los propietarios, la dotación de servicios públicos domiciliarios, la ubicación (urbana, rural), asentamientos indígenas, entre otros.

<sup>6</sup>Dentro del estrato, el nivel mínimo es 1 y el máximo es 6

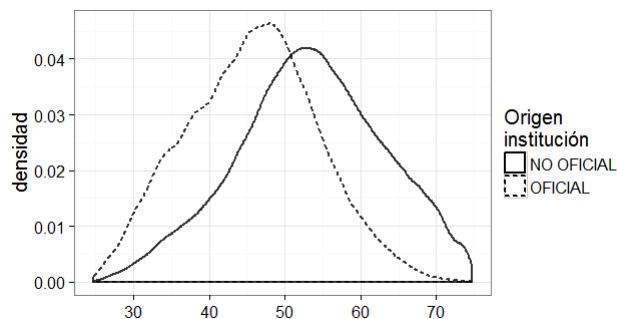


Figura 3: INSE por orígen

Debido a que en el nivel de estudios superiores algunos estudiantes ya no están dentro de un hogar sino que son la cabeza de hogar, el contraste con la variable dicotómica que corresponde a la respuesta de si el estudiante es o no cabeza de hogar se encuentra en la Figura 4

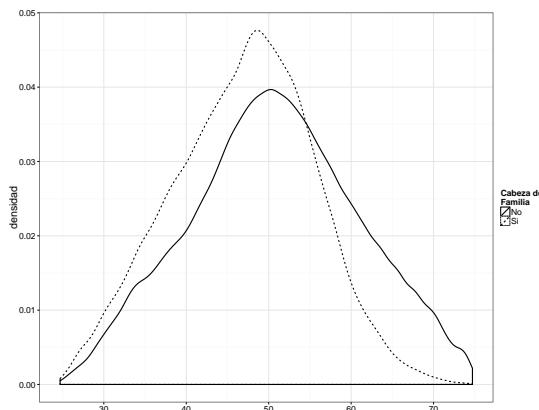


Figura 4: INSE - Estudiante es Cabeza de Familia

En la Figura 4 se puede observar que aquellos estudiantes que son cabeza del hogar presentan puntajes del INSE más bajos que aquellos que aún dependen de sus padres o de alguna otra persona. La presentación de los 3 resultados anteriores dan cuenta de un buen comportamiento del índice frente a variables de interés y validan su construcción.

Para finalizar, es importante entender si existe alguna relación entre el Nivel Socioeconómico (NSE)<sup>7</sup> y los puntajes obtenidos por los estudiantes. Como se muestra en la Tabla 7, la relación entre nivel socioeconómico y puntaje en las pruebas es

<sup>7</sup>Denominamos Nivel Socioeconómico a la clasificación del INSE por medio de sus cuartiles,

similar a través de las diferentes pruebas de competencias genéricas: *inglés, lectura crítica, razonamiento cuantitativo, comunicación crítica y competencias ciudadanas.*

Prueba	bajo	medio-bajo	medio-alto	alto
Inglés	9.570	9.798	10.027	10.933
Comunicación escrita	9.859	10.006	10.100	10.327
Lectura crítica	9.655	9.884	10.030	10.421
Competencias ciudadanas	9.702	9.908	10.041	10.366
Razonamiento cuantitativo	9.693	9.886	10.003	10.389

Tabla 7: Puntajes prueba de inglés por INSE y grupo de referencia

Como ilustración, en la Tabla 8 se encuentran los puntajes promedio, para los grupos de referencia universitarios, obtenidos para la prueba de inglés por NSE. Este comportamiento es muy similar al encontrado en las demás competencias genéricas.

	Grupo de referencia	Bajo	Medio-bajo	Medio-alto	Alto
1	Administración	9.568	9.725	9.937	10.904
2	Arquitectura	9.716	9.967	10.207	11.103
3	Bellas Artes	9.920	10.452	10.733	11.685
4	Ciencias Agropecuarias	9.614	9.849	10.083	10.717
5	Ciencias Militares	9.936	10.273	10.181	10.557
6	Ciencias Naturales	10.116	10.508	10.814	11.904
7	Ciencias Sociales	9.580	9.749	9.979	11.174
8	Comunicación	9.861	10.056	10.265	11.149
9	Derecho	9.644	9.706	9.860	10.626
10	Economía	9.885	10.241	10.539	11.842
11	Educación	9.612	9.893	10.182	10.872
12	Humanidades	9.966	10.421	10.944	12.239
13	Ingeniería	9.980	10.217	10.424	11.296
14	Medicina	10.354	10.569	10.620	11.366
15	Recreación	9.513	9.664	9.883	10.327
16	Salud	9.605	9.739	9.868	10.218

Tabla 8: Puntajes prueba de Inglés por NSE y grupo de referencia

La relación entre NSE y puntaje en las pruebas de inglés es directa, a mayor NSE se presenta un puntaje más alto en la prueba de Inglés. Esta relación se mantiene en los diferentes grupos de referencia universitarios.

---

de esta forma, los estudiantes que obtienen un INSE que se encuentra en el primer cuartil serán asignados al grupo con un NSE Bajo, el segundo cuartil al grupo con un NSE medio-bajo y de esta misma forma se asignan los dos grupos restantes.

## 5. Conclusiones

La metodología que se presenta en este artículo puede ser replicable en estudios en los que se requiera la construcción de un índice de nivel socioeconómico, en particular, en estudios sobre desempeño académico en educación superior.

La construcción de un índice socioeconómico para estudiantes de educación superior demanda el uso y planteamiento de variables adicionales a las utilizadas comúnmente en el contexto de los estudiantes de educación básica y media. Se espera que futuras investigaciones contribuyan a estos análisis y planteamientos bajo esta realidad.

La validez externa del INSE es verificada a través del comportamiento del índice en relación a variables de interés, como el estrato socioeconómico, el origen de la institución educativa y si el estudiante es o no cabeza de familia. Con respecto a estas tres variables, el INSE presenta el comportamiento esperado, de acuerdo a las características de las subpoblaciones que definen las variables de interés.

La simetría observada para el INSE de los estudiantes de SABER PRO contrasta con la encontrada en los análisis de las condiciones socioeconómicas de los estudiantes en SABER 11 y que presenta una asimetría positiva. Este comportamiento podría ser explicado por el efecto de la cobertura en la educación y que se asocia con barreras económicas entre diferentes niveles de la educación (según cifras del Ministerio de Educación Nacional<sup>8</sup>, la tasa de cobertura de la educación superior es cercana al 50 %).

## Referencias

- Acevedo, S. & Jaramillo, A. (2007), ‘Perfil socioeconómico de los estudiantes de pregrado EAFIT’.
- Armenta, N., Pacheco, C. & Pineda, E. (2008), ‘Factores socioeconómicos que intervienen en el desempeño académico de los estudiantes universitarios de la facultad de ciencias humanas de la universidad autónoma de baja california’, *Revista IIPSI* 1(1), 153–165.
- Caso-Niebla, J. & Hernandez, L. (2007), ‘Variables que inciden en el rendimiento académico de adolescentes mexicanos’, *Revista latinoamericana de psicología* 39(3), 487–501.
- Coleman, J. S. (1988), ‘Social capital in the creation of human capital’, *The University of Chicago Press* 94, 95–120.

<sup>8</sup>Ver [http://www.mineducacion.gov.co/sistemasdeinformacion/1735/articles-212350\\_Estadisticas\\_de\\_Educacion\\_Superior\\_.pdf](http://www.mineducacion.gov.co/sistemasdeinformacion/1735/articles-212350_Estadisticas_de_Educacion_Superior_.pdf)  
Consultado por última vez el 15 de febrero de 2016.

- Contreras, K., Caballero, C., Palacio, J. & Pérez, A. (2008), 'Factores asociados al fracaso académico en estudiantes universitarios de barranquilla (colombia)', *Psicoloía desde el Caribe* **22**, 110–135.
- Díaz, S. D. & Barrios, G. H. (2002), 'Eficiencia escolar y diferencias socioeconómicas: a propósito de los resultados de las pruebas de medición de la calidad de la educación en chile.', *Educação e Pesquisa* **28**(2), 25–39.
- De Leeuw, J. & Mair, P. (2007), 'Homogeneity analysis in R: The package homals', *Department of Statistics, UCLA*.
- De Leeuw, J. & Mair, P. (2009), 'Gifi methods for optimal scaling in r: The package homals', *Journal of Statistical Software, forthcoming* pp. 1–30.
- Estevez, J. F. (2002), 'La construcción de un Índice cuantitativo sobre educación superior utilizando la técnica de análisis de componentes principales.', *Revista de la Educación superior*. **31**(121), 138–153.
- Gil-Flores, J. (2013), 'Medición del nivel socioeconómico familiar en el alumnado de educación primaria', *Revista de educación* **362**, 298–322.
- Hernández, E. & González, M. (2011), 'Modelo de ecuación estructural que evalúa las relaciones entre el estatus cultural y económico del estudiante y el logro educativo', *Revista electrónica de investigación educativa* **13**(2), 188–203.
- Horn, J. L. (1965), 'A rationale and test for the number of factors in factor analysis.', *Psychometrika* **30**(2), 85–179.
- Jabnoun, N. (2009), 'Economic and cultural factors affecting university excellence', *Quality Assurance in Education* **17**(4), 416–429.
- MEN (2015), Colombia un país que avanza hacia el mejoramiento de las oportunidades de acceso a la educación superior., Technical report, Ministerio de Educación Nacional.
- Nuñez, J. (n.d.), Evidencia e interpretación de los problemas sociales en colombia. Requisito parcial para optar al título de DOCTOR EN CIENCIAS SOCIALES Y HUMANAS, PONTIFICIA UNIVERSIDAD JAVERIANA.
- Paz-Navarro, L., Roldán, R. & González, M. (2009), 'Funcionamiento familiar de alumnos con bajo rendimiento escolar y su comparación con un grupo de rendimiento promedio en una preparatoria de la universidad de guadalajara', *Revista de Educación y Desarrollo* **10**, 5–15.
- Piñeros, L. & Rodriguez, A. (1998), Los insumos escolares en la educación secundaria y su efecto sobre el rendimiento académico de los estudiantes: un estudio en Colombia, Technical Report 20934, El Banco Mundial.
- Porcel, E., Dapoz, G. & López, V. (2010), 'Predicción del rendimiento académico de alumnos de primer año la facena (unne) en función de su caracterización socioeducativa.', *Revista Electrónica de Investigación Educativa* **12**(2).

- Ravela, P. (2005), Estudio de los factores institucionales y pedagógicos que inciden en los aprendizajes en escuelas primarias de contextos desfavorecidos en uruguay, Technical report, ANEP/MECAEP/UMRE.
- Saz, M. M. A. (2006), 'Influencia del nivel socioeconómico y cultural en el rendimiento de los estudiantes de tercero básico y graduandos del año 2006', *Dirección General de Evaluación e Investigación Educativa*.
- Vargas, G. M. G. (2013), 'Factores asociados al rendimiento académico en estudiantes universitarios desde el nivel socioeconómico: Un estudio en la universidad de costa rica', *Revista Electrónica Educare* 17(3), 57–87.