

## Factores asociados al rendimiento académico en estudiantes universitarios

Jaime A. Gutiérrez-Monsalve<sup>1, 2\*</sup>, Juan Garzón<sup>2</sup> y Angela M. Segura-Cardona<sup>3</sup>

(1) Escuela de Graduados, Programa de Doctorado en Epidemiología y Bioestadística, Universidad CES, G.I. en Educación Superior GIES, Medellín- Colombia (correo-e: jgutierrez@uco.edu.co).

(2) Facultad de Ingenierías, Grupo de Investigación en Ingenierías Multidisciplinar (GIMU), Universidad Católica de Oriente UCO Rionegro - Colombia. (correo-e: fgarzon@uco.edu.co).

(3) Escuela de Graduados, Universidad CES, G.I. en Epidemiología y Bioestadística Medellín-Colombia. (correo-e: asegura@ces.edu.co).

\* Autor a quien debe ser dirigida la correspondencia

Recibido Mar. 13, 2020; Aceptado May. 12, 2020; Versión final Jul. 28, 2020, Publicado Feb. 2021

---

### Resumen

Se presenta una metodología para determinar las variables asociadas al rendimiento académico universitario (RA) de una cohorte de estudiantes del año 2014 en una universidad privada del noroeste de Colombia. Se utilizó un discriminante canónico y una regresión logística para establecer la asociación entre variables pedagógicas, institucionales y sociodemográficas con el bajo RA utilizando únicamente información de los sistemas administrativos de la universidad. El análisis discriminante permitió clasificar el 100% de los estudiantes con bajo RA a partir de las variables institucionales y sociodemográficas. Las variables pedagógicas no mostraron discriminar adecuadamente, a excepción de la facultad de derecho. Desde la regresión logística se encontraron asociaciones significativas del bajo RA con la trayectoria del estudiante, el hecho de tener beca, la repetencia y el número de asignaturas canceladas. Se concluye que el análisis discriminante y la regresión logística permiten clasificar y establecer los factores que explican el bajo RA universitario.

*Palabras clave: rendimiento académico; factores predictores; regresión logística; analítica de la educación*

## Factors associated to academic performance in university students

### Abstract

This study present a methodology to determine factors associated to academic performance (AP) in a 2014 cohort of students from different faculties at a northwest private university in Colombia. A canonical discriminant and a logistic regression are used to identify the association between pedagogical, institutional, and sociodemographic variables with low AP by using only information obtained from the UCO's administrative systems. Discriminant analysis successfully classified 100% of students with low AP through institutional and sociodemographic variables. Pedagogical variables did not show adequate classification, except for faculty of law. Logistic regressions showed significant associations between student degree path, having a scholarship, the repetition indicator, and the indicator of canceled subjects per semester. It is concluded that discriminant analysis and logistic regression allow to classify and establish the factors that explain the low AP.

*Keywords: academic performance; predictive factors; logistic regression; learning analytics*

## INTRODUCCIÓN

El rendimiento académico (RA) en un concepto complejo en si mismo, se puede definir como el valor atribuido a los resultados de aprendizaje de los estudiantes universitarios en un área temática determinada comparado con el nivel de conocimientos esperado en sus pares. Es considerado también como un indicador de la calidad de la enseñanza universitaria y su medida agrupa factores pedagógicos, institucionales, sociodemográficos y psicosociales (Montero Rojas & Villalobos Palma, 2007). Las calificaciones son los indicadores de logro más utilizados por los programas académicos para estimar el RA universitario de sus estudiantes. Es así como cada universidad y según su normativa determina los criterios evaluativos para dar una valoración del RA considerando el promedio ponderado en cada una de las asignaturas, el número de créditos matriculados, la cantidad de asignaturas aprobadas, indicadores de eficacia, eficiencia y productividad en un periodo específico de tiempo (Ridgell y Lounsbury, 2004). En la Tabla 1 se muestra una síntesis de los indicadores de valoración del rendimiento académico más utilizados en universidades latinoamericanas (García, 2014).

La valoración del RA universitario establece la relación entre lo que el estudiante aprende y lo que alcanza en el proceso de enseñanza – aprendizaje como producto del trabajo académico en las diferentes actividades en las que se desempeñó ante sus profesores (Rodríguez, Fita, y Torrado, 2004). Dicha valoración puede ser imprecisa, arriesgada y polémica, razón por la cual el abordaje del RA se limita a una generalización conceptual, pues el contexto, las características de los estudiantes y, las metodologías evaluativas empleadas influyen en la generalización de la medición (Garbanzo Vargas, 2013).

A diferencia de los porcentajes de graduación y deserción, el rendimiento académico puede medirse en etapas tempranas de la formación universitaria, es una medida continua en el tiempo que puede registrarse y gestionarse en cualquier momento del proceso formativo y ha sido considerada como una variable explicativa de la deserción universitaria (Ortiz-Lozano, Rua-Vieites, Bilbao-Calabuig, y Casadesús-Fa, 2020). La utilidad en la gestión académico-administrativa de los estudios sobre el RA se basa en identificar, definir y estudiar variables asociadas que puedan ser modificadas, intervenidas, controladas por la misma universidad con el fin de proveer a los tomadores de decisiones de insumos para mejorar los resultados de aprendizaje, la calidad del servicio educativo y controlar la deserción universitaria (Arora y Singh, 2017).

En términos prácticos y con el fin de precisar la valoración del RA por factores, resulta habitual identificar el RA con los resultados de aprendizaje, clasificándolo en dos categorías: El RA inmediato, determinado por las calificaciones obtenidas por el estudiante en su proceso formativo hasta obtener su titulación profesional y, el RA mediato que hace referencia a los resultados de aprendizaje obtenidos a partir de la aplicación de la formación recibida en el proceso de incorporación al mundo laboral de los graduados universitarios (Díaz y Arias, 1999).

Desde el factor pedagógico, el RA se comprende a partir de la eficacia de todas las estrategias de enseñanza utilizadas por los profesores, los métodos de evaluación, el material didáctico, así como los procedimientos que se siguen por parte de los profesores con el propósito de cubrir los objetivos del curso y sus contenidos (Page et al., 1990). Desde el factor institucional se incluyen todas las características estructurales y funcionales de cada universidad y/o programa académico, así como el grado de influencia de éstas en las particularidades propias de los estudiantes: el programa de elección, número de estudiantes por curso, número máximo de créditos que se le permite matricular, horario de los cursos, así como la disponibilidad de becas o subsidios económicos (Latiesa, 1992). El factor sociodemográfico concentra las variables que permiten clasificar a los estudiantes por rasgos diferenciales. Define las pautas sociales propias de cada cultura donde se desenvuelve el estudiante. En ellas se incluyen el estrato socioeconómico, sexo, sitio de residencia, colegio de procedencia, nivel educativo de los padres, situación laboral, acceso a recursos de información y consulta, rendimiento académico previo (Castaño, Gallón, Gómez, y Vásquez, 2006; Page et al., 1990).

Tabla 1. Algunos indicadores de medición del RA universitario

Indicador	Descripción del indicador
1	Asignaturas cursadas desde el momento del ingreso al programa académico
2	Asignaturas aprobadas desde el momento de ingreso al programa académico
3	Indicador de eficacia calculado como el cociente entre 2 y 1
4	Indicador de productividad media: cociente entre el número de asignaturas aprobadas desde el ingreso y los años desde el ingreso al programa académico
5	Promedio ponderado de calificaciones o notas (con o sin aplazos)
6	Índice que multiplica el indicador 4 y el 5
7	Índice que multiplica el indicador 3 y el 5
8	Número de aplazamientos acumulados durante el curso del programa académico
9	Asignaturas aprobadas respecto a asignaturas según duración teórica

El factor Psicosocial considera las variables que describen las interrelaciones que se dan entre el estudiante y la sociedad. Desde la dimensión no cognitiva variables como apertura a la experiencia, niveles de extroversión, estabilidad emocional, niveles de neurotismo, amabilidad y empatía (Chamorro-Premuzic y Furhnam, 2006), el autocontrol (Pérez-Villalobos, Cobo-Rendón, Sáez, y Díaz-Mujica, 2018), la autoeficacia y el autoconcepto (Mandel, 2011; Stegers - Jager, Cohen - Schotanus, y Themmen, 2012), así como la motivación, la ansiedad y la aptitud intelectual son variables importantes (Garbanzo Vargas, 2007). Desde el punto de vista de la calidad educativa universitaria, el RA es uno de los indicadores más importantes a la hora de medir los resultados académicos y valorar la calidad de las universidades y sus programas. Tanto los gobiernos como los entes regulatorios de la Educación Superior (ES) en latinoamérica y el mundo exigen que las universidades implementen indicadores de RA como un requisito para ofertar el programas académicos en condiciones de calidad (Dias Sobrinho, 2007; Zimmerman, Bandura, & Martinez-Pons, 1992).

En Colombia el Ministerio de Educación Nacional (MEN) a partir del Sistema de Acreditación Nacional dicta los lineamientos, guías, documentos, técnicas e indicadores que le permiten a los programas académicos y universidades ofertar sus propuestas curriculares en el país (Ministerio de Educación Nacional, 2019). El MEN ha adoptado una serie de indicadores que permiten medir el resultado académico inmediato y mediato de los estudiantes en los programas académicos y que incluyen entre otras variables; el desarrollo académico, pedagógico y didáctico; impacto de los egresados en el medio social y académico a partir de los resultados de aprendizaje; la relación entre lo que el estudiante aprende y lo que alcanza en el proceso de enseñanza-aprendizaje; la coherencia y objetividad de los sistemas de valoración y evaluación de los resultados de aprendizaje, así como las actitudes, conocimientos, habilidades y capacidades adquiridas en el proceso académico. También se presentan indicadores de seguimiento, políticas y estrategias para garantizar la permanencia y graduación de los estudiantes universitarios (Mejía, Barajas, & Fajardo, 2018).

En latinoamérica estos indicadores de calidad académica han sido adaptados también por países como Brasil, Chile, Costa Rica, Ecuador, El Salvador, México, Nicaragua, Panamá, Paraguay, Perú, República Dominicana y Uruguay todos ellos agrupados en la Red Iberoamericana para el Aseguramiento de la Calidad en la Educación (RIACES), donde siguiendo lineamientos similares a los de Colombia, los ES les exigen a las universidades la medición y el cumplimiento de criterios de calidad para el funcionamiento y la oferta de los programas académicos (Villanueva, 2004).

Ante las exigencias de los ES, las universidades de latinoamérica han desarrollado sistemas de información que les permiten medir las variables necesarias para acreditar su calidad y asegurar su oferta académica. En términos generales los sistemas de información administrativa permiten obtener información verificada sobre el estado de matrícula del estudiante, su identidad, el programa matriculado y su modalidad, sitio de residencia, asignaturas matriculadas por semestre, información y evaluación de los profesores, RA representado en calificaciones de tipo cuantitativo/cualitativo, promedio acumulado por semestre, edad, subsidios o descuentos otorgados, créditos cursados/aprobados por semestre, colegio de procedencia, fechas de matrícula y graduación, número de veces que un estudiante repite una asignatura, así como el curriculum vigente en el plan de estudios del estudiante. Toda esta información se debe reportar regularmente a los ES de cada país como el MEN en Colombia con el fin de certificar la calidad de los programas. Sin embargo este acervo de información ¿permitirá explicar el RA inmediato de los estudiantes universitarios a partir de los resultados de aprendizaje?, o ¿será necesario invertir dinero en otras fuentes de información como encuestas o instrumentos de medición?

Lo novedoso de este trabajo radica en que sería el primero en Colombia que utilice únicamente las variables derivadas de la información administrativa de los programas académicos para lograr explicar los factores que se asocien con el RA universitarios teniendo en cuenta diferentes áreas del conocimiento, sin tener que incurrir en gastos adicionales para buscar mayor información. Por lo expuesto anteriormente se propone establecer los factores explicativos del RA de estudiantes universitarios de pregrado tomando como estudio de caso una universidad privada del noroeste de Colombia (UCO) en una cohorte que inicio estudios de pregrado en el primer semestre de 2014 (2014-1). Para el desarrollo de la investigación se utilizará únicamente los datos derivados de los sistemas de información administrativa desde un enfoque de minería de datos en educación (Baradwaj & Pal, 2012).

## **METODOLOGÍA**

Se seleccionaron los estudiantes de la cohorte de inicio del semestre 2014-1 con información recolectada de 399 participantes de una universidad privada del noroeste de Colombia (UCO) (679868.04 m N; 459607.51 m E (WGS84)). Los estudiantes se distribuyeron en ocho áreas del conocimiento (Ciencias Agropecuarias (C.Agro.), Ciencias de la Educación (C.Educ.), Ciencias de la Salud (C.Salud.), Ciencias Económicas y Administrativas (C.Econ. y Adm.), Ciencias Sociales (C.Soc.), Derecho e Ingeniería).

Dentro del factor pedagógico (F.Pedagógico) se tomaron las calificaciones de los estudiantes a sus profesores en aspectos como pedagogía (Pedag.Prof.), didáctica de enseñanza y aprendizaje (Didact.Prof.), gestión académica y evaluación del docente (Gest.Prof.), percepción general de los estudiantes sobre sus profesores (Cal.Prof), identidad y filosofía institucional (Fil.Inst.Prof.), edad de los profesores (Edad.Prof.), salario medio de los profesores (Sal.Prof.) y estudiantes promedio por curso por profesor (Est.Prof.).

Dentro del factor institucional (F.Inst.) se consideraron variables como el número de veces que el estudiante matricula una misma asignatura (Veces), la trayectoria o avance por semestre del estudiantes en su programa académico (Nivel), el número de asignaturas matriculadas por el estudiante al semestre (Asign.Cur.), el número de asignaturas canceladas por el estudiante al semestre sin que se afecte su nota promedio semestral (Asign.Canc.), si el estudiante cuenta con algún tipo de subsidio o beca (Subsidio) y, el porcentaje de descuento que se le otorga al estudiante cuando formaliza su matrícula (%.Ayudas).

Dentro del factor sociodemográfico (F.Sociodemograf.) se consideraron variables como la edad, el sexo (femenino, masculino), si el estudiante cursó su secundaria en un colegio privado (Col.Privado) y la región de procedencia (Altiplano: municipios de El Retiro, Guarne, La Ceja, La Unión, Marinilla y Rionegro; Paramo: Abejorral, Argelia, Nariño, Sonsón; Embalses: El Peñol, Granada, Guatapé, San Carlos, San Rafael, San Vicente; Bosques: Cocorná, San Francisco, San Luis; V.Aburrá: Bello, Copacabana, Caldas, Envigado, Itagüí, La Estrella, Medellín, Girardota, Sabaneta; O.Antioquia (O.Ant.): Otros municipios del departamento de Antioquia-Colombia no incluidos; Otros Colombia (O.Col.): municipios de Colombia diferentes a los especificados en el departamento de Antioquia).

Dentro de los sistemas de información administrativa de la universidad no se encontró registro de ninguna variable del factor psicosocial. Dado que los criterios de medición en las calificaciones pueden variar entre programas de diferentes áreas del conocimiento, se utilizó como variable de medición del RA el promedio ponderado semestral del estudiante, el cuál se estipula en una escala de 1 a 5, donde 5.0 es la máxima calificación. Igualmente esta variable se operativizó a partir de quintiles. De acuerdo a la disponibilidad de recursos, la universidad solo puede atender por bajo RA al 20% de la población estudiantil.

Para determinar la calificación derivada del promedio ponderado semestral que clasifique a un estudiantes con bajo RA, se utilizó un discriminante canónico y se asignaron las reglas de clasificación que permitieran identificar el 100% de estudiantes con bajo RA para los programas académicos agrupados en cada una de las áreas de conocimiento. Todo el componente estadístico se desarrollo en software R (Team R, 2013).

### Modelo logístico

Con las variables que mostraron efectos significativos se generó un modelo logístico para determinar la probabilidad de que un estudiante se encuentre con bajo o normal rendimiento académico durante la trayectoria su trayectoria académica. En el modelo logístico la variable dependiente  $Y_i$  (definida como el evento de que un estudiante tenga bajo RA se calcula de acuerdo con la expresión (Hosmer, Lemeshow y Sturdivant, 2013) (1):

$$Y_i = \left( \frac{\Pi_i}{1-\Pi_i} \right) \quad (1)$$

Donde  $\Pi_i$  es la probabilidad de que en el estudiante  $i$  se produzca bajo rendimiento académico y es el valor de la variable dependiente en el caso  $i$  para los coeficientes de regresión  $\alpha$  y  $\beta$ .

$$Y_i = \alpha + \beta X_i \quad (2)$$

Expresado en forma de regresión, el modelo logístico puede representarse como:

$$\Pi_i = \frac{1}{1 + e^{-(\alpha + \beta X_i)}} \quad (3)$$

Las razones de riesgo para bajo RA (OR) indican la probabilidad de que un estudiante tenga bajo RA respecto a que tenga normal RA de acuerdo al valor que asuma la variable independiente  $X_i$ , constituyen una medida de asociación e interpretación del efecto de cada una de las variables independientes sobre el RA. De esta manera el OR puede estimarse a partir del modelo de regresión logística como sigue (Nizami & Prasad, 2017):

$$OR_i = \log \frac{\Pi_i}{1 - \Pi_i} = \frac{1}{\frac{1 + e^{-(\alpha + \beta X_i)}}{e^{-(\alpha + \beta X_i)}}} = \log \left( \frac{1}{e^{-(\alpha + \beta X_i)}} \right) = \alpha + \beta X_i \quad (4)$$

Por último se calcularon los OR ajustados con el fin de identificar los riesgos de bajo RA teniendo en cuenta la trayectoria de los estudiantes por nivel. Todas las pruebas de hipótesis se probaron con una significancia de 0.05 (Pregibon, 1981).

## RESULTADOS Y DISCUSIÓN

En primer lugar se clasifican a los estudiantes de una universidad privada del noroeste de Colombia (UCO) con bajo y normal rendimiento académico, luego se presentará la regresión logística que permita estimar los OR para la UCO y todas las áreas del conocimiento estudiadas, por último se hará un análisis de trayectorias del RA por nivel de formación.

### *Clasificación de los estudiantes según el rendimiento académico*

Con el fin de establecer cuáles son las variables que discriminan entre los estudiantes de bajo y normal RA, se utilizó un discriminante canónico para toda la universidad (UCO) y luego se realizó para cada una de las áreas de conocimiento tal y como se muestra en las Figuras 1 y 2. Los estudiantes con bajo RA se clasificaron de acuerdo a la capacidad que mostró el modelo para clasificar correctamente el 100% de ellos, de esa manera se determinó la calificación promedio semestral por debajo de la cual un estudiante es considerado con bajo RA para cada una de las áreas. Para este análisis se tuvo en cuenta la restricción de que la universidad solo puede atender el 20% de su población por bajo RA.

En ese sentido los estudiantes con calificaciones promedio semestrales por debajo de 3.34 de un total de 5.0 en ciencias agropecuarias deben ser considerados con bajo RA, en ciencias de la educación con calificaciones por debajo de 3.7, ciencias de la salud 3.42, ciencias económicas y administrativas 3.5, ciencias sociales 3.7, derecho 3.4 e ingeniería 3.2. De este modo ingeniería y ciencias agropecuarias son las que menores calificaciones necesitan para generar alertas sobre estudiantes con bajo RA, por su parte las ciencias sociales y ciencias de la educación son las áreas del conocimiento con calificaciones más altas para clasificar a un estudiante con bajo RA, esto sugiere que las mediciones del RA son específicas para cada área del conocimiento y no debe generalizarse para toda la UCO.

De acuerdo a la longitud de los vectores en el panel derecho de la estructura de Figura 1 y 2, las variables del factor pedagógico fueron las que menor discriminación aportaron, sin embargo en derecho, valoraciones altas en componentes pedagógicos (Pedag.Prof.), didáctica (Didac.Prof.), gestión académica y evaluativa (Gest.Prof.), identidad y filosofía institucional (Fil.Inst.Prof.) aportaron significativamente para discriminar a los estudiantes con normal RA. Por su parte las variables sociodemográficas de los profesores: edad (Edad.Prof) y salario (Sal.Prof.) discriminaron únicamente en ciencias de la educación, ciencias de la salud, ciencias económicas y administrativas e ingenierías hacia el bajo RA. El número de estudiantes por profesor (Est.Prof.) es una variable importante para la UCO, pues se cree que a mayor número de estudiantes por profesor menor será el RA del estudiante, en este sentido esta variable discriminó para bajo RA en ingeniería, ciencias sociales y ciencias de la salud. Sin embargo en ciencias de la educación, ciencias económicas y administrativas y derecho el tener más estudiantes por profesor discriminó hacia un RA normal, lo que evidencia resultados particulares por área de conocimiento.

En la Figura 1 se muestra el discriminante canónico de las variables que explican y estructuran los factores que influyen en el bajo rendimiento académico de los estudiantes de A. toda la Universidad y, los programas B. Ciencias Agropecuarias, C. Ciencias de la Educación, D. Ciencias de la Salud. Las cajas en las puntuaciones canónicas representan a los estudiantes con Bajo y Normal rendimiento académico. En la estructura se presentan en flechas las variables de cada factor. La dirección de las flechas indica que la variable tiene efectos sobre la caja de BAJO o NORMAL rendimiento según este se encuentre hacia arriba o hacia abajo. Bajo RA indica la nota de bajo RA en una escala de 1 a 5.

Del factor institucional variables como mayor repetencia (Veces), mayor número de asignaturas canceladas por semestre (Asign.Canc), no contar con beca estudiantil (Subsidio), estar en un nivel inferior en la trayectoria formativa (Nivel) y un menor porcentaje de ayudas económicas aportaron para discriminar a los estudiantes con bajo RA en la mayoría de áreas del conocimiento (Figura 1 y 2).

En las variables del factor sociodemográfico se vieron aportes diferenciados para cada una de las áreas de conocimiento: ser hombre se identificó como una variable que discrimina hacia el bajo RA en ciencias agropecuarias, ciencias de la educación, ciencias económicas y administrativas, ciencias sociales y derecho, sin embargo en ciencias de la salud el ser hombre discrimina hacia RA normal y en ingenierías no tuvo efectos.

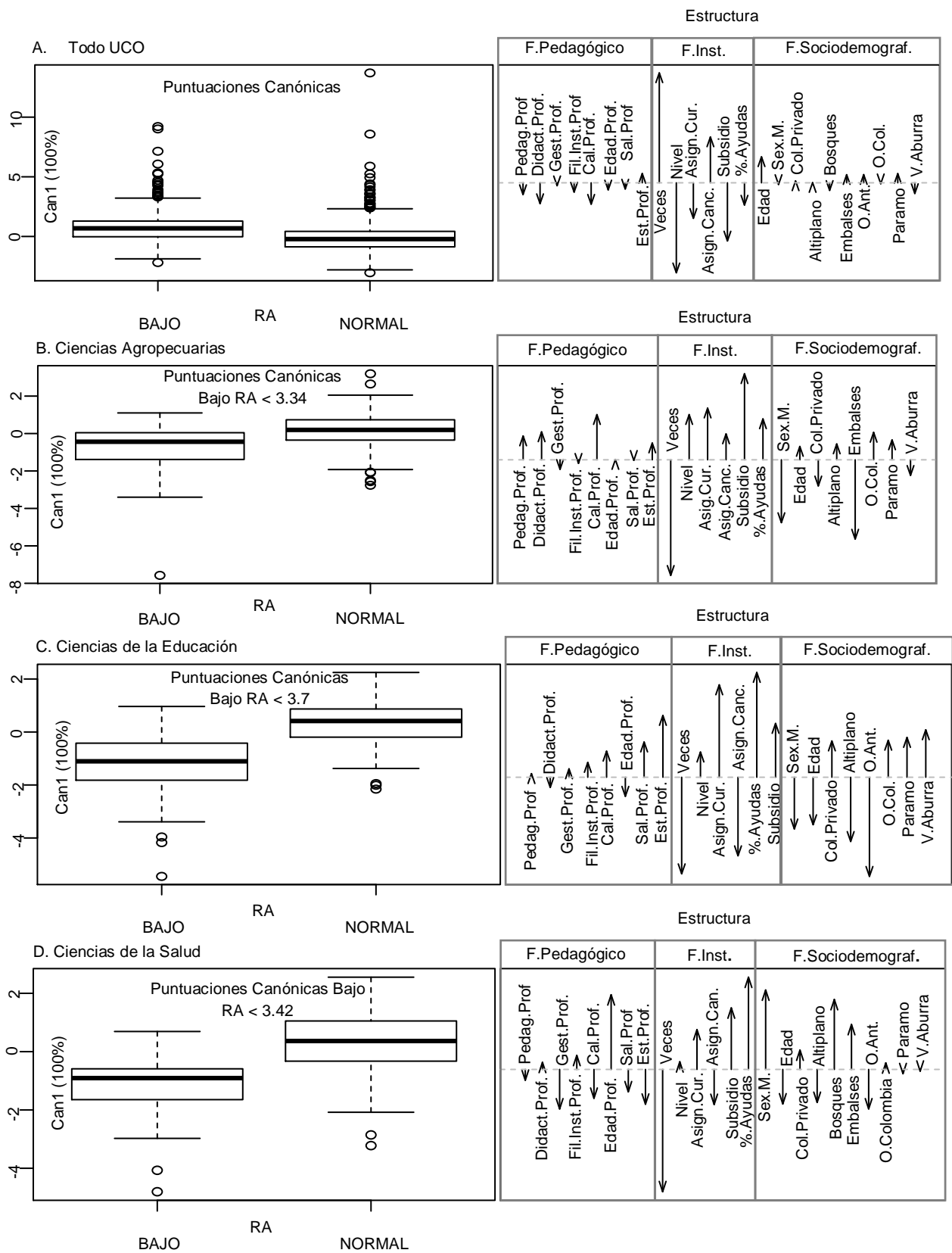


Fig.1: Discriminante canónico de las variables que explican y estructuran los factores que influyen en el bajo rendimiento académico de los estudiantes

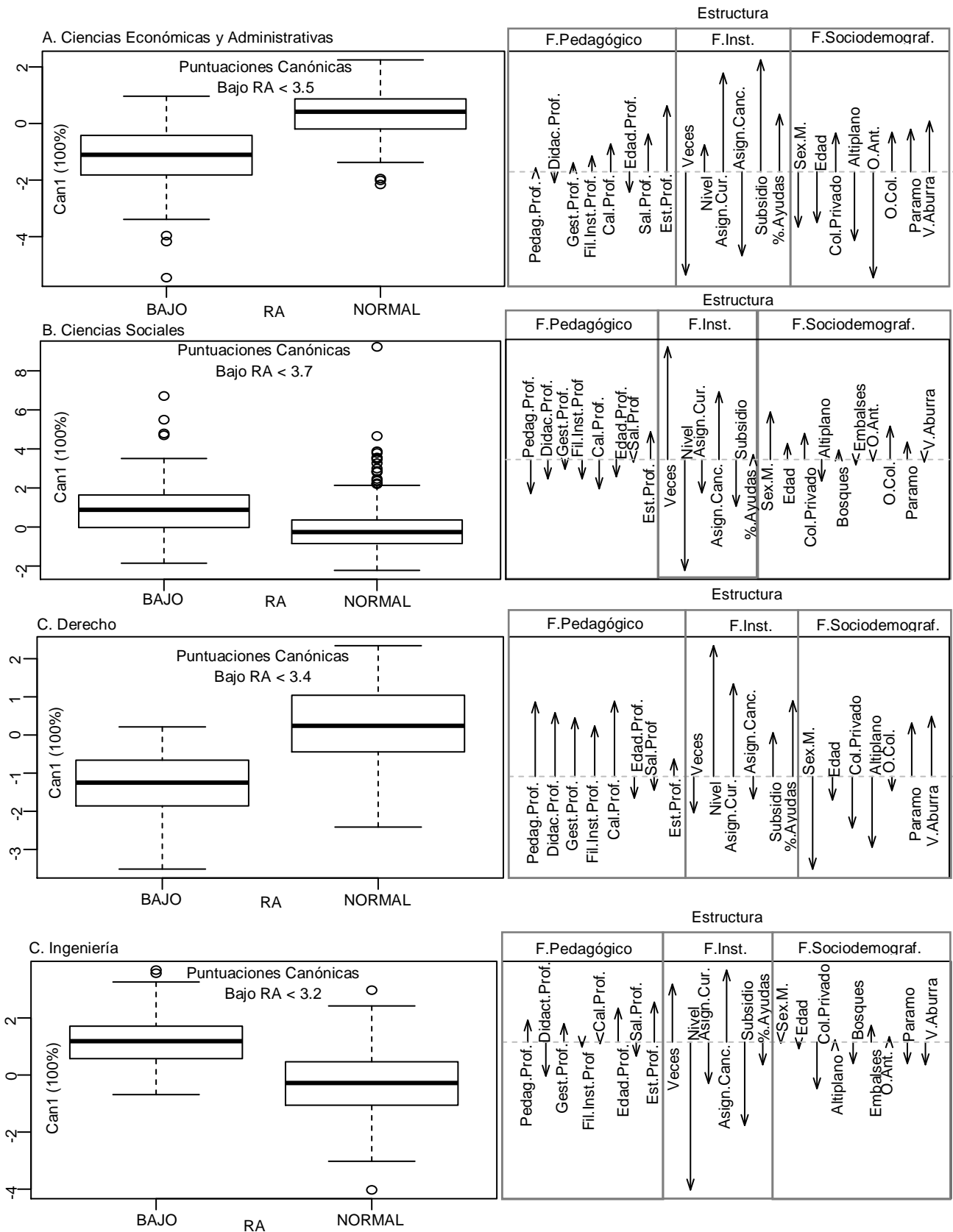


Fig.2: Discriminante canónico de las variables que explican y estructuran los factores que influyen en el bajo rendimiento académico en diferentes áreas de conocimiento

En el caso de la edad, tener más años de vida discrimina hacia el bajo RA en todas las áreas del conocimiento a excepción de ingeniería y ciencias agropecuarias donde no se observa una discriminación importante. En las variables región de residencia la discriminación fue muy particular e individual, por ejemplo los estudiantes provenientes del altiplano parecen discriminar hacia el bajo RA en ciencias de la educación, ciencias de la salud, ciencias económicas y administrativas, y derecho; por su parte los estudiantes provenientes del Valle de Aburrá discriminan para RA Normal para ciencias de la educación, ciencias económicas y administrativas, derecho e ingeniería (Figura 1 y 2).

Halpin en 1990 utilizó un discriminante canónico para determinar aquellas variables que lograban explicar la persistencia en estudiantes universitarios, en su modelo en la primera componente, logró concentrar el 74.5% de la varianza explicada con los factores prácticas pedagógicas, el acompañamiento estudiantil en las facultades, el desarrollo académico e intelectual de los estudiantes y el nivel de interacción entre el estudiante y la facultad, mientras que en la segunda componente que incluyó la motivación y el compromiso con las metas académicas del estudiante el resto de la varianza. Las variables estudiadas por este autor se agrupan en los factores pedagógico, institucional y sociodemográfico expuestas (Halpin, 1990).

Diseth y Martinsen en 2003, utilizaron un discriminante canónico para identificar los factores que mejor predicen el fracaso académico de un curso en psicología, con este modelo, en las dos primeras dimensiones canónicas encontraron que variables asociadas a la trayectoria académica pasada, variables psicosociales tales como la necesidad de conocer, el estilo explorador y asimilador de los estudiantes así como la motivación al logro y la motivación para reponerse al fracaso fueron las variables que explicaron entre un 70% y un 90% la varianza (Diseth & Martinsen, 2003). En este estudio se utiliza un discriminante para clasificar a los estudiantes con bajo RA e identificar las variables que generan mayor discriminación, las cuáles en el modelo explicativo pueden dar luces sobre los factores que afectan el bajo RA.

#### *Factores explicativo del bajo rendimiento académico*

A partir de una regresión logística optimizada por el criterio de máxima verosimilitud de Akaike (AIC) se calcularon medidas de asociación ajustadas a partir de razones de probabilidad (OR). En la Tabla 2 y 3, se muestran los OR ajustados para cada una de las áreas de conocimiento, así como la sensibilidad, la especificidad y el área bajo la curva ROC (AUC) para cada uno de los modelos desarrollados.

En la Tabla 2 se muestra las razones de riesgo ajustadas (OR) e intervalo de confianza del OR al 95% para el bajo RA en todas las variables analizadas en las diferentes áreas del conocimiento. El número entero en la primera línea indica el OR promedio, en la segunda línea se presenta el intervalo de confianza y el asterisco indica efecto significativo con un valor-P de Wald (\*\*\*) menor a 0.001, \*\* entre 0.01 y 0.1, \* entre 0.01 y 0.05).

En la Tabla 3 se muestra el número entero en la primera línea indica el OR promedio media, en la segunda línea se presenta el intervalo de confianza y el asterisco indica efecto significativo con un valor-P de Wald (\*\*\*) menor a 0.001, \*\* entre 0.01 y 0.1, \* entre 0.01 y 0.05). Se anexa la sensibilidad y la especificidad del modelo así como el área bajo la curva ROC (AUC). Las líneas punteadas indica que no hay suficientes estudiantes para calcular los respectivos OR.

Con un punto de corte del 20%, la especificidad (capacidad de identificar correctamente a los estudiantes con bajo RA) de todos los modelos desarrollados fue superior al 66%, resaltando el desempeño del modelo de ciencias de la educación e ingenierías con especificidades por encima del 80%. De acuerdo al área bajo la curva (AUC) ROC, todos los modelos reportan valores por debajo de 0.9, lo que significa que los modelos no predicen el azar, pero deben ser mejorados si lo que se quiere es generar un modelo predictivo (Tabla 3).

De acuerdo al valor-P de Wald (0.05 de significancia), ninguna variable del factor pedagógico mostró una asociación significativa con el bajo RA. El factor institucional fue el que mayor número de variables con asociación significativa presentó, es así como altos índices de repitencia (Veces) y cancelar más asignaturas al semestre presentaron asociaciones significativas en la UCO, ciencias agropecuarias y ciencias sociales; la trayectoria del estudiante en niveles avanzados (nivel) fue una variable protectora del bajo RA en todas las áreas menos en ciencias de la educación y ciencias de la salud; también la variable subsidio presentó efectos protectores en la UCO y ciencias de la educación. Por su parte en el factor sociodemográfico, el ser hombre estuvo asociado de manera significativa con el bajo RA en toda la UCO y ciencias agropecuarias; mientras que ser menor de 18 años solamente fue un factor protector en la facultad de ingenierías (Tabla 2 y 3). El modelo logístico reportado en este estudio permitió ajustar los riesgos de bajo RA a partir de las variables que resultaron ser significativas con los datos arrojados por el sistema de información administrativa de los estudiantes de la UCO.



Tabla 2: razones de riesgo ajustadas (OR) e intervalo de confianza del OR al 95% para el bajo RA en todas las variables analizadas en las diferentes áreas del conocimiento.

A. Factor Pedagógico								
Variable	UCO	C. Agro.	C. Educ.	C. Salud.	C. Econ y Adm.	C. Soc.	Derecho	Ingeniería
Pedag. Prof.	1.04 (0.95,1.09)	1.05 (0.94,1.17)	0.89 (0.74,1.07)	1.16 (0.94,1.43)	1.23 (0.99,1.51)	0.95 (0.87,1.03)	1.07 (0.9,1.26)	1.15 (0.96,1.39)
Didac. Prof.	0.96 (0.95,0.98)	0.92 (0.79,1.06)	1.03 (0.77,1.37)	1.03 (0.88,1.21)	0.9 (0.77,1.05)	1.05 (0.97,1.14)	0.73 (0.38,1.43)	0.94 (0.86,1.02)
Gest. Prof.	1.02 (0.97,1.06)	1.03 (0.93,1.14)	0.89 (0.73,1.07)	26.61 (0.53,2)	1.04 (0.97,1.12)	1.04 (0.97,1.12)	1.04 (0.88,1.23)	0.96 (0.8,1.15)
Fil. Inst. Prof.	0.97 (0.93,1.02)	1.04 (0.9,1.2)	0.85 (0.58,1.23)	0.84* (0.71,1)	1.04 (0.84,1.28)	0.98 (0.91,1.06)	1.14 (0.66,1.97)	0.88 (0.74,1.06)
Cal. Prof.	0.95* (0.91,0.99)	0.86 (0.74,1)	0.85 (0.64,1.13)	0.91 (0.64,1.31)	0.83* (0.73,0.95)	0.9* (0.84,0.96)	1.06 (0.75,1.49)	0.97 (0.85,1.11)
Edad. Prof.	0.98 (0.97,1)	0.97 (0.92,1.01)	0.99 (0.91,1.07)	0.95 (0.85,1.06)	0.99 (0.94,1.04)	0.99 (0.96,1.02)	1.05 (0.98,1.13)	0.94 (0.9,0.99)
Est. Prof.	1.002 (0.99,1.01)	0.99 (0.98,1.01)	1.00 (0.98,1.02)	1.02 (1,1.05)	0.99 (0.98,1.01)	1.01 (1.00,1.01)	1.00 (0.98,1.02)	0.99 (0.98,1.01)
B. Factor Institucional								
Variable	UCO	C. Agro.	C. Educ.	C. Salud.	C. Econ y Adm.	C. Soc.	Derecho	Ingeniería
Veces	3.8*** (2.71,5.41)	5.13*** (1.73,15.2)	3.34 (0.46,24.2)	4.88 (0.3,79.8)	2.14 (0.73,6.31)	3.00*** (1.82,4.93)	10.2 (0.85,123)	3.76 (1.48,9.57)
Nivel	0.74*** (0.71,0.78)	0.68*** (0.55,0.83)	0.84 (0.69,1.03)	0.65 (0.4,1.07)	0.59*** (0.49,0.71)	0.76*** (0.71,0.82)	0.47*** (0.31,0.72)	0.31*** (0.22,0.43)
Asign. Cur.	0.99 (0.94,1.03)	1.06 (0.95,1.18)	0.91 (0.61,1.35)	0.67* (0.47,0.97)	0.97 (0.83,1.14)	0.91* (0.84,0.99)	1.24(0.86,1.78)	0.97 (0.8,1.18)
Asign. Canc.	1.18** (1.05,1.33)	1.03 (0.84,1.27)	1.68 (0.9,3.14)	1.37 (0.51,3.65)	0.98 (0.69,1.41)	1.54** (1.17,2.03)	1.1 (0.39,3.12)	1.09 (0.81,1.48)
Subsidio	0.52*** (0.42,0.64)	0.83 (0.49,1.39)	0.26*** (0.1,0.68)	0.61 (0.21,1.81)	2.14 (0.73,6.31)	0.71 (0.49,1.03)	0.19 (0.02,1.68)	0.65 (0.37,1.12)
Edad. Prof.	0.98 (0.97,1)	0.97 (0.92,1.01)	0.99 (0.91,1.07)	0.95 (0.85,1.06)	0.99(0.94,1.04)	0.99 (0.96,1.02)	1.05 (0.98,1.13)	0.94 (0.9,0.99)
Est. Prof.	1.002 (0.99,1.01)	0.99 (0.98,1.01)	1.00 (0.98,1.02)	1.02 (1,1.05)	0.99 (0.98,1.01)	1.01 (1,1.01)	1.00 (0.98,1.02)	0.99 (0.98,1.01)

Tabla 3: razones de riesgo ajustadas (OR) e intervalo de confianza del OR al 95% para el bajo RA en todas las variables analizadas en las diferentes áreas del conocimiento.

A. Factor Sociodemográfico								
Variable	UCO	C. Agro.	C. Educ.	C. Salud.	C. Econ y Adm.	C. Soc.	Derecho	Ingeniería
Edad: menor 18 años	0.85 (0.67,1.07)	1.35 (1.1,1.66)	1.74 (0.65,4.64)	0.49 (0.12,1.98)	0.97 (0.54,1.76)	1.00 (0.62,1.60)	0.38 (0.12,1.16)	0.5* (0.27,0.9)
Sex. M.	1.4** (1.18,1.66)	1.38* (1.16,1.64)	1.85 (0.71,4.83)	0.16 (0.03,0.9)	1.3 (0.79,2.15)	1.36 (0.99,1.87)	1.38 (0.52,3.65)	0.99 (0.56,1.75)
Col. Privado	1.15 (0.87,1.52)	0.95 (0.43,2.01)	1.35 (0.51,3.57)	1.54 (0.43,5.56)	1.32 (0.60,2.94)	0.74 (0.47,1.08)	0.9 (0.27,3.03)	3.44 (0.93,12.5)
Altiplano	0.75 (0.45,1.25)	1.04 (0.83,1.3)	1.28 (0.29,5.68)	1.02 (0.25,4.16)	1.72 (0.94,3.15)	0.93 (0.62,1.39)	2.89 (0.79,10.6)	0.96 (0.48,1.93)
Bosques	0.41 (0.17,1.01)	0.6 (0.28,1.28)	-----	-----	-----	0.78 (0.27,2.27)	-----	0.84 (0.13,5.66)
Embalse	0.64 (0.34,1.21)	0.96 (0.62,1.5)	0.34 (0.02,4.83)	0.28 (0.03,3.13)	1.45 (0.23,8.99)	0.91 (0.47,1.76)	-----	2.13 (0.64,7.09)
O. Ant.	1.04 (0.56,1.92)	1.47 (0.97,2.21)	6.03 (1.23,29.58)	1.14 (0.16,8.1)	1.96 (0.58,6.69)	0.91 (0.47,1.76)	-----	0.94 (0.26,3.49)
O. Col	1.22 (0.69,2.16)	1.56 (1.11,2.19) *	1.79 (0.33,9.68)	0.86 (0.13,5.55)	-----	1.52 (0.86,2.7)	1.55 (0.24,10.0)	1.35 (0.52,3.52)
Paramo	0.86 (0.47,1.59)	1.25 (0.84,1.88)	0.58 (0.05,7.4)	0.65 (0.04,10.7)	2.41 (0.73,7.95)	1.99 (1.05,3.79)	-----	1.61 (0.35,7.51)
V. Aburra	1	1	1	1	1	1	1	1
Especif. %	70.25%	69.66%	81.82%	77.14%	79.17%	66.28%	76.47%	84.82%
Sensib. %	67.401%	66.38%	79.19%	75.68%	69.00%	74.95%	71.63%	68.38%
AUC	0.75	0.74	0.87	0.86	0.82	0.76	0.84	0.86

### Trayectoria del bajo RA en diferentes niveles o semestres de formación

En el factor institucional, una de las variables de mayor interés es el nivel de los estudiantes. En este caso el nivel es un indicador del avance del proceso de formación en el programa académico. En la UCO la mayoría de programas tienen diez niveles o semestres de duración. En la Figura 3 se muestra la trayectoria del OR ajustado para bajo RA. En la Figura 3 se muestra OR ajustado para bajo RA en la trayectoria del estudiante por niveles de formación en diferentes áreas de formación. Las barras indican el intervalo de confianza al 95% del OR ajustado. La línea punteada es la referencia y se encuentra a un OR de 1.

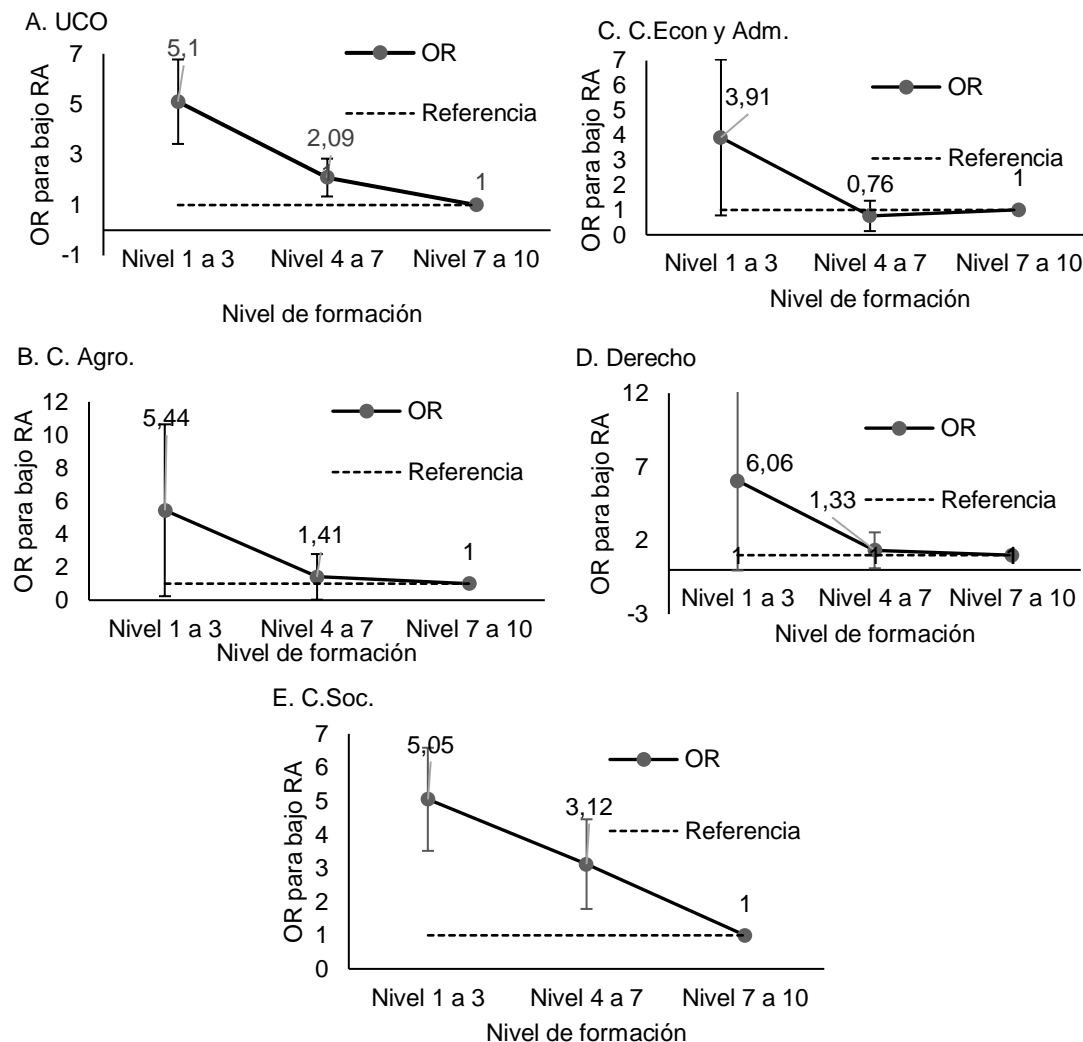


Fig.3: OR ajustado para bajo RA en la trayectoria del estudiante por niveles de formación en diferentes áreas de formación.

Es así como para toda la UCO los estudiantes que se encuentran entre el nivel 1 a 3 tienen 5.1 más probabilidad de tener bajo RA respecto a los estudiantes que se encuentran en niveles 7 a 10; en ciencias agropecuarias el riesgo de tener bajo RA en los niveles 1 a 3 es de 5.44, en ciencias económicas y administrativas es de 3.91, en derecho de 6.06 y en ciencias sociales de 5.05. Las demás áreas del conocimiento no mostraron asociación significativa del RA con el nivel. Estos resultados sugieren que los niveles inferiores de formación son los más críticos a la hora de intervenir el bajo RA.

Es interesante notar que a diferencia de lo que se creía, el colegio de egreso de los estudiantes no fue un factor de riesgo para el bajo RA, es así como Iregui, Melo y Ramos en 2007 hacen un estudio sobre la eficiencia de la educación en Colombia y comparan el RA de los estudiantes egresados de colegios públicos y privados, en tal sentido los autores concluyen que los colegios privados pueden beneficiarse de condiciones de entorno más favorables dado que atienden a estudiantes de mayores ingresos, sin embargo cuando los colegios públicos se asumen en entornos equivalentes a los privados, no existe diferencias estadísticas entre el RA de los estudiantes procedentes de ambos tipos de colegio, caso que parece ocurrir en este estudio (Iregui, Melo, & Ramos, 2007). En el caso del sexo, Lin y compañeros en el 2014 reportaron asociaciones significativas entre el sexo, el RA y los indicadores de permanencia en estudiantes de ingenierías de una universidad de Qatar en favor de las mujeres (Lin y Imbric, 2014), esto se observó en toda la UCO y en

ciencias agrarias. A nuestro conocimiento, no se encuentran reportes sobre la asociación entre el RA considerando las regiones de residencia de los estudiantes de la UCO y, además que estratifique su estudio por áreas de conocimiento.

En Holanda en 2009, el programa de ingeniería electrónica de la universidad tecnológica de Eindhoven utilizó diferentes modelos con el fin de predecir la probabilidad de deserción de estudiantes de primer semestre del programa a partir de la información contenida en el sistema de admisiones de la universidad. En este estudio, los modelos con mejor capacidad predictiva fueron el logístico y el árbol de clasificación y decisión (CART) que logró una sensibilidad del 80% y una especificidad del 79%, estos resultados fueron similares a los obtenidos en los modelos desarrollados, sin embargo al igual que ellos, la capacidad predictiva es muy limitada al partir únicamente de las variables administrativas (Dekker, Pechenizkiy, y Vleeshouwers, 2009). Aunque los porcentajes de sensibilidad y especificidad en los modelos logísticos no varían significativamente con los obtenidos en este trabajo, todavía es necesario incorporar otro tipo de variables de naturaleza psicosocial para poder lograr unos porcentajes de especificidad más adecuados, una AUC por encima de 0.9 y con esto poder desarrollar un modelo predictivo del bajo RA que genere alertas tempranas para el oportuno acompañamiento de los estudiantes universitarios.

## CONCLUSIONES

De acuerdo al trabajo presentado y a los resultados obtenidos, se pueden plantear las siguientes conclusiones principales:

(1) El análisis discriminante canónico permitió en este estudio clasificar a los estudiantes que presentaron rendimiento académico normal de aquellos clasificados con bajo rendimiento académico, asignando las calificaciones semestrales que los ponen en riesgo de tener bajo RA e identificar las variables que más los discriminan. Desde el factor institucional las variables que generaron mayor discriminación para bajo RA fueron el nivel, el indicador de repitencia (veces) y no tener beca. En las variables del factor sociodemográfico el sexo masculino y la edad fueron las más representativas. En cuanto a las variables del factor pedagógico, únicamente en derecho lograron discriminar a los estudiantes con bajo RA.

(2) A partir de una regresión logística se estimaron los niveles de asociación del bajo RA con todas las variables del factor pedagógico, institucional y sociodemográfico para cada una de las áreas de conocimiento de la UCO. La validación cruzada o especificidad de los modelos obtenidos estuvo por encima del 70% y la sensibilidad por encima del 67%. El AUC de la curva ROC se ubicó por encima de 0.74 por lo que el modelo se debe ajustar con variables del factor psicosocial y personal para ser predictivo, lo cuál supondría costos adicionales.

(3) De acuerdo a las razones de riesgo ajustadas (OR), la variable de repetición de asignaturas al semestre (Veces), estar en niveles inferiores en la trayectoria de formación en el programa académico, tener beca o subsidio y, ser de sexo masculino son variables que en la mayoría de áreas del conocimiento se asociaron con bajo RA.

(4) Para mejorar la capacidad predictiva del modelo de RA en la UCO es necesario incorporar otro tipo de variables de tipo psicosocial, sin embargo esto implicaría mayores inversiones para medirlas en todos los estudiantes de la universidad.

## AGRADECIMIENTOS

Los autores agradecen a la Universidad Católica de Oriente y a la Universidad CES.

## REFERENCIAS

- Arora, N., y Singh, N., *Factors Affecting the Academic Performance of College Students*, I-Manager's Journal of Educational Technology, 14(1), 47 (2017).
- Baradwaj, B. K., y Pal, S., *Mining Educational Data to Analyze Students' Performance*, ArXiv Preprint ArXiv:1201.3417 (2012).
- Castañó, E., Gallón, S., Gómez, K., y Vásquez, J., *Análisis de los Factores Asociados a la Deserción y Graduación Estudiantil Universitaria*. <https://doi.org/10.17533/udea.le.n65a2639>, Lecturas de Economía, (65) (2006).
- Chamorro-Premuzic, T., y Furrnam, A. *Intellectual Competence and the Intelligent Personality: A Third Way in Differential Psychology*. <https://doi.org/10.1037/1089-2680.10.3.251>, Review of General Psychology, 10(3), 251 (2006).
- Dias Sobrinho, J., *Acreditación de la educación superior en América Latina y el Caribe*. <http://hdl.handle.net/2099/7538>, La Educación Superior en el Mundo (2007).

- Díaz, M., y Arias, J., *La evaluación del rendimiento inmediato en la enseñanza universitaria*, Revista de Educación, 320, 353–377 (1999).
- Diseth, Å. G. E., y Martinsen, Ø., *Approaches to Learning, Cognitive Style, and Motives as Predictors of Academic Achievement*. <https://doi.org/10.1080/01443410303225>, Educational Psychology, 23(2), 195–207 (2003).
- Garbanzo Vargas, G. M., *Factores Asociados al Rendimiento Académico en Estudiantes Universitarios, una Reflexión desde la Calidad de la Educación Superior Pública*. DOI: 10.15517/REVEDU.V3111.1252, Educación, 31(1) (2007).
- Garbanzo Vargas, G. M., *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 (2013).
- García, A. M., *Rendimiento Académico y Abandono Universitario: Modelos, Resultados y Alcances de la Producción Académica en la Argentina*, Revista Argentina de Educación Superior, 8(6), 9-38 (2014).
- Halpin, R. L., *An Application of the Tinto model to the Analysis of Freshman Persistence in a Community College*. <https://doi.org/10.1177/009155219001700405>, Community College Review, 17(4), 22–32 (1990).
- Hosmer Jr, D. W., Lemeshow, S., y Sturdivant, R. X., *Applied Logistic Regression* (Vol. 398), John Wiley y Sons (2013).
- Iregui, A. M., Melo, L., y Ramos, J., *Análisis de Eficiencia de la Educación en Colombia*, Revista de Economía Del Rosario, 10(1), 21–41 (2007).
- Latesa, M., *La Deserción Universitaria: Desarrollo de la Escolaridad en la Enseñanza Superior: Éxitos y Fracasos*, (Vol. 124). CIS (1992).
- Lin, J. J., y Imbrie, P. K., *Modeling Retention and Graduation of Engineering Students of Difference Sexes*, 6th First Year Engineering Experience Conference, TX, August 7, 8 (2014).
- Mejía, A., Barajas, L., y Fajardo, H. *Documento Metodológico, MIDE Universitario 2018*. Bogotá DC (2018).
- Ministerio de Educación Nacional., *Decreto 1330, por el cual se sustituye el capítulo 2 y se suprime el Capítulo 7 del Título 3 de la Parte 5 del Libro 2 del decreto 1075 de 2015 - Único Reglamentario del Sector Educación* (2019).
- Montero Rojas, E., y Villalobos Palma, J., *Factores Institucionales, Pedagógicos Psicosociales y Sociodemográficos Asociados al Rendimiento Académico en la Universidad de Costa Rica: Un Análisis Multinivel*, Revista Electrónica de Investigación y Ecaluación Educativa (RELIEVE) (2007).
- Nizami, N., y Prasad, N., *Multinomial Logistic Regression Analysis. In Decent Work: Concept, Theory and Measurement* (pp. 315–319). Springer (2017).
- Ortiz-Lozano, J. M., Rúa-Vieites, A., Bilbao-Calabuig, P., y Casadesús-Fa, M., *University Student Retention: Best Time and Data to Identify Undergraduate Students at Risk of Dropout*. <https://doi.org/10.1080/14703297.2018.1502090>, Innovations in Education and Teaching International, 57(1), 74–85 (2020).
- Page, M., Moreal, B., Calleja, y otros cuatro autores, *Hacia un Modelo Causal del Rendimiento Académico*, Madrid: Cide (1990).
- Pérez-Villalobos, M. V., Cobo-Rendón, y otros dos autores, *Revisión Sistemática de la Habilidad de Autocontrol del Estudiante y su Rendimiento Académico en la Vida Universitaria*. <http://dx.doi.org/10.4067/S0718-50062018000300049>, Formación Universitaria, 11(3), 49–62 (2018).
- Pregibon, D., *Logistic Regression Diagnostics*, The Annals of Statistics, 9(4), 705-724 (1981).
- Richardson, M., Abraham, C., y Bond, R., *Psychological Correlates of University Students' Academic Performance: A Systematic Review and Meta-Analysis*, Psychological Bulletin, 138(2), 353–387 (2012).
- Ridgell, S. D., y Lounsbury, J. W., *Predicting Academic Success: general intelligence, "big five" personality traits, and work drive*. College Student Journal, 38(4), 607–618 (2004).
- Rodríguez, S., Fita, E., y Torrado, M., *El Rendimiento Académico en la Transición Secundaria-Universidad*, Revista de Educación, 334(1), 391–414 (2004).
- Team, R. C., *R foundation for statistical computing*. Vienna, Austria, 3(0) (2013).
- Villanueva, E., *La Acreditación en América Latina: el Caso de Argentina en la RIACES y en el MERCOSUR*. Revista Iberoamericana de Educación, 35(1), 99–112 (2004).
- Stegers-Jager, K. M., Cohen-Schotanus, J., y Themmen, A. P. N., *Motivation, learning strategies, participation and medical school performance*. <https://doi.org/10.1111/j.1365-2923.2012.04284.x>, Medical Education, 46(7), 678–688 (2012).
- Zimmerman, B. J., Bandura, A., y Martínez-Pons, M., *Self-motivation for Academic Attainment: The role of Self-efficacy Beliefs and Personal Goal Setting*. <https://doi.org/10.3102/00028312029003663>, American Educational Research Journal, 29(3), 663–676. (1992)