

Modelización del rendimiento matemático mediante análisis discriminante

Antonio Humberto Closas

Universidad Nacional del Nordeste, Argentina, *hclosas@hotmail.com*

Edgardo Alberto Arriola

Universidad Tecnológica Nacional, Argentina, *earriola2006@ca.frre.utn.edu.ar*

Mariela Rosana Amarilla

Universidad Tecnológica Nacional, Argentina, *profe.amarilla@ca.frre.utn.edu.ar*

Ethel Carina Jovanovich

Universidad Tecnológica Nacional, Argentina, *carijovanovich@ca.frre.utn.edu.ar*

Resumen: *El objetivo de este estudio consiste en proponer un modelo estadístico que permita explicar de qué manera se relacionan algunas variables personales y contextuales con el rendimiento matemático de estudiantes de Ingeniería. La muestra está compuesta por 142 jóvenes matriculados en distintas carreras que se imparten en un centro educativo tecnológico de la región nordeste de Argentina. El diseño metodológico es de corte transversal, línea cuantitativa y estilo correlacional. Los resultados obtenidos en la fase empírica permitieron contrastar que el modelo propuesto se ajusta a los datos muestrales y resulta de utilidad para clasificar el rendimiento en Matemáticas de nuevos individuos.*

Palabras clave: *Desempeño académico, Factores personales y contextuales, Estudiantes universitarios, Modelado estadístico.*

Modelling of mathematical performance using discriminant analysis

Abstract: *The aim of this study is to propose a statistical model that allows explaining how some personal and contextual variables are related to the mathematical performance of engineering students. The sample consisted of 142 young people enrolled in different courses of study taught in a technological educational center in the northeast region of Argentina. The methodological design is cross-sectional, quantitative and correlational. The results obtained in the empirical phase allowed us to verify that the proposed model fits the sample data and is useful for classifying the mathematics performance of new individuals.*

Key words: *Academic performance, Personal and contextual factors, University students, Statistical modelling.*

1. INTRODUCCIÓN

En la región nordeste de Argentina, entre las dificultades que se presentan al principio de la etapa académica, en distintas áreas de conocimiento, se encuentran los altos índices de

suspensos, los cuales en buena medida se deben a inconvenientes en la adaptación de los estudiantes a las exigencias curriculares que demanda el nivel educativo universitario.

A esta realidad, se le suma el hecho de que los jóvenes tampoco realizan aquellas actividades que les permitiría lograr aprendizajes acordes a las circunstancias, todo lo cual genera que muchos de ellos decidan no continuar los estudios iniciados o elijan cambiar de carrera.

El escenario que se origina debido a los bajos resultados educativos, por los suspensos, la no promoción o la deserción, claramente conlleva un importante costo social y económico tanto para la familia como para el Estado; lo cual es muy relevante, puesto que la educación constituye un sector estratégico que incide directa e invariablemente en el progreso de la sociedad y representa uno de los pilares principales para el desarrollo de cualquier Nación (Devincenzi et al., 2017).

También, la problemática descrita ha sido observada por los autores de este trabajo en virtud de la experiencia de muchos años en el proceso de enseñanza de Matemáticas, y de otras asignaturas del ciclo básico común, que se imparten en el centro educativo del cual procede la muestra, así como en distintos ámbitos académicos de su área de incumbencia.

De acuerdo con Delors (1997), el bajo rendimiento educativo es un fenómeno tan preocupante en el plano humano, moral y social que muy a menudo ocasiona exclusiones que estarán presentes en los jóvenes durante toda su vida. Por el contrario, la efectividad en el proceso educativo es un factor de bienestar a distintos niveles, propicia el crecimiento de los países, la viabilidad de las instituciones educativas, como también la prosperidad social y personal.

Si bien el rendimiento de los estudiantes se encuentra relacionado con variables de índole personal o individual que, de un modo u otro, participan y lo generan, existen otras variables de tipo contextual o ambiental que deberían tenerse en cuenta a efectos de conformar un conjunto de determinantes que permitan esclarecer mediante un enfoque integral las razones que ocasionan el desempeño académico.

En un interesante trabajo, Pintrich (2003) aboga por la necesidad de desarrollar modelos holísticos y de llevar a cabo investigaciones empíricas en las que explícitamente se estudien las interacciones e interrelaciones entre los distintos componentes que participan (motivacionales, cognitivos y relativos al contexto de aprendizaje).

En esta línea, se conocen diversos trabajos que analizan, mediante técnicas estadísticas multivariadas, de qué manera y en qué medida las variables atinentes al individuo (aptitudinales, emocionales, cognitivas, etc.), como aquellas que dependen del entorno (familiares, socioculturales, institucionales, etc.), se relacionan con los resultados educativos de los estudiantes (Doménech et al., 2004; González-Pienda et al., 2003; Kassarnig et al., 2018; Reynolds y Walberg, 1992; Schnitzler et al., 2021).

En este marco, el objetivo principal del presente estudio consiste en proponer un modelo, elaborado mediante la técnica estadística denominada análisis discriminante, que explique de qué manera se relacionan ciertas variables personales y contextuales con el rendimiento académico de estudiantes de Ingeniería, en el ámbito de una asignatura del área de Matemáticas.

En términos generales, un modelo es una representación abstracta y compacta de un sistema que permite describir y analizar su comportamiento (Voght, 1993). En este sentido, la modelización estadística es un recurso metodológico que resulta de utilidad para abordar el tratamiento de un fenómeno, como es el bajo rendimiento académico, puesto que da lugar a conceptualizar y comunicar sus dinámicas de una forma coherente.

Una característica central que posee el modelo que se plantea en este estudio es que posibilita explicar los datos recogidos, así como predecir observaciones futuras, lo cual contribuye no solo a comprender el problema, sino también a tomar decisiones con el objeto de prevenir las dificultades cognitivas en el área de conocimiento y, por ende, mejorar el desempeño educativo de los individuos de la población de la que fue seleccionada la muestra.

El modelo que será presentado en el marco de este estudio ha sido previamente contrastado en forma empírica. Para ello, fue necesario realizar una serie de estudios cuantitativos, propios de la técnica estadística que se utiliza, los cuales hicieron posible la estimación y validación de los coeficientes de la ecuación que mejor explique los datos muestrales (Hair et al., 2019). El trabajo de campo (aplicación de test y cuestionarios *ad hoc*), que permitió recoger los datos con los que se realizaron los diversos análisis estadísticos, tuvo lugar en el marco de la asignatura Análisis Matemático I (AMI). Esta materia es común a las carreras de Ingeniería (Sistemas de Información -ISI-, Electromecánica -IEM- y Química -IQ-) que se desarrollan en la Facultad Regional Resistencia (FRRe) de la Universidad Tecnológica Nacional (UTN), Argentina.

El análisis discriminante es una técnica de dependencia que permite crear un modelo predictivo capaz de establecer la pertenencia de objetos a determinados grupos. El modelo está compuesto por una función discriminante, en el supuesto de que haya dos grupos (o, un conjunto de funciones discriminantes, si hubiera más de dos grupos), basada en combinaciones lineales de las variables predictoras, o independientes, que proporcionan la mejor discriminación posible entre los grupos previamente establecidos (en nuestro caso, aprobados y suspensos) por medio de la variable dependiente (en este estudio, el rendimiento matemático). En líneas generales, las funciones se generan a partir de una muestra de observaciones para los que se conoce el grupo de pertenencia; posteriormente, las funciones pueden ser aplicadas a nuevos casos que dispongan de mediciones para las variables predictoras, pero de los que se desconozca el grupo de pertenencia (Uriel y Aldás, 2005).

En vista de lo que antecede y en atención al objetivo propuesto, en esta investigación serán dos las variables que formarán parte del factor explicativo personal: a) el *autoconcepto académico*, y b) las *estrategias de aprendizaje*. A su vez, las variables que integrarán el factor explicativo contextual serán: c) los *aspectos sociofamiliares*, y d) el *clima de clase*. En tanto que, la variable dependiente, o que se desea explicar, del modelo discriminante que más adelante se pondrá a consideración será el *rendimiento académico* en el contexto de la asignatura AMI.

Existen diversas definiciones sobre el constructo motivacional autoconcepto académico; así por ejemplo podría decirse que es aquella variable en la que la implicación activa del sujeto en su proceso de aprendizaje se incrementa cuando se percibe autoeficiente (Sánchez León, 2023). En tanto que para Arrivillaga et al. (2023), el autoconcepto académico se refiere a la imagen que el individuo construye sobre su desempeño y nivel de éxito en el ámbito educativo.

La variable cognitiva *estrategias de aprendizaje*; por una parte, implica una secuencia de actividades u operaciones mentales dirigidas a facilitar el aprendizaje y, por otra, incluye procesos de toma de decisión por parte de los estudiantes de carácter consciente e intencional, ajustados al objetivo que pretenden conseguir (Beltrán, 2003).

El abordaje de la variable contextual *aspectos sociofamiliares* se realizará desde la perspectiva macrosociológica y estará centrada en tres cuestiones concretas: el estatus socioeconómico, el contexto sociocultural y los rasgos de la zona donde habita el estudiante.

La variable del entorno del alumno *clima de clase*, se define como la atmósfera general de la sala de clase que es percibida por los estudiantes; en ocasiones, incluye también la percepción de los profesores y, excepcionalmente, de otros miembros de la comunidad educativa.

Los indicadores de la variable *rendimiento académico* serán las calificaciones obtenidas por los estudiantes en tres instancias de evaluaciones parciales escritas teórico-prácticas correspondientes a la asignatura AMI.

A continuación, se encuentra la metodología implementada (participantes, procedimiento de aplicación, características de la investigación, instrumentos de medición utilizados, aspectos del análisis discriminante). Más adelante, se ofrecen los resultados obtenidos, su discusión y por último las conclusiones a las que se ha llegado tras la ejecución de este estudio.

2. MATERIALES Y MÉTODO

2.1. Participantes

La muestra estuvo compuesta por 5 grupos-clase (2 de ISI, 2 de IEM y 1 de IQ) de la asignatura AMI, los cuales totalizaron 142 jóvenes (45 mujeres y 97 hombres), pertenecientes a las tres carreras de Ingeniería que se imparten en la sede central de la FRRe-UTN.

La edad de los estudiantes que respondieron los ítems de la encuesta tuvo una media (M) de 19.75 años y desviación estándar (DE) de 1.42. Los datos que identifican a la muestra utilizada en este trabajo se ilustran en la Tabla 1.

El procedimiento empleado para extraer la muestra consistió en la combinación de tres métodos: a) estratificado, b) por conglomerados, y c) aleatorio simple (Lohr, 2019). El primer método fue de utilidad para definir las especialidades de Ingeniería que participarían en la investigación, el segundo a efectos de asumir que los clústeres estarían conformados por grupos-clase, y el tercero debido a que estos últimos (también llamados comisiones de estudio) pudieran ser elegidos al azar.

Tabla 1

Detalles relativos a la muestra empleada en la etapa empírica del estudio.

Turno	Carrera	Alumnos	Edad
Tarde y Noche	Ingeniería en Sistemas de Información	$n = 40$ jóvenes (10 mujeres, 30 hombres)	$Mín. = 18$ $Máx. = 24$ $M = 20.08$ $DE = 1.31$
Tarde y Noche	Ingeniería Electromecánica	$n = 54$ jóvenes (6 mujeres, 48 hombres)	$Mín. = 18$ $Máx. = 24$ $M = 19.93$ $DE = 1.49$
Mañana	Ingeniería Química	$n = 48$ jóvenes (29 mujeres, 19 hombres)	$Mín. = 18$ $Máx. = 23$ $M = 19.27$ $DE = 1.33$

2.2. Procedimiento de aplicación

La aplicación de los tests y cuestionarios fue realizada por los docentes responsables en cada uno de los cinco clústeres que integran la muestra, al comienzo de la clase y durante un intervalo promedio de 20 minutos.

Concluido el trabajo de campo y el ordenamiento de la información obtenida, se procedió a la construcción de la matriz de datos en formato electrónico (en esta primera instancia se utilizó el

programa MS Excel), así como a su posterior control estadístico general mediante el paquete IBM SPSS Statistics 26 (George y Mallery, 2020).

2.3. Características de la investigación

La elaboración de un modelo de causa-efecto, como el que será propuesto en este trabajo, requiere de diferentes variables y de distintos análisis estadísticos, en virtud de lo cual se detallan en forma breve en la Tabla 2, en el marco del generoso espectro metodológico disponible, los criterios y métodos que serán considerados y que caracterizan la presente investigación.

Tabla 2

Características de la investigación

Criterios	Métodos
Dado que el estudio está basado en las respuestas que se brindan al tema objeto de consulta tal como se presenta en su contexto real y en una única instancia.	No experimental Transversal
Teniendo presente el objetivo que se pretende conseguir y el modo de reunir la información para lograrlo.	Explicativo Encuesta
En atención a las características de los instrumentos de medición que se aplican.	Test Cuestionario
Si se tiene en cuenta el marco de recogida de los datos y la forma en que se analizan.	De campo Cuantitativa
Debido al interés por estudiar las asociaciones entre las variables que participan y la manera de relacionarlas mediante el modelado estadístico.	Correlacional Funcional

2.4. Instrumentos de medición

Con el objeto de reunir los datos de las variables explicativas, en este estudio se utilizó un instrumento de medida conformado por: a) prueba *autoconcepto académico* del test Autoconcepto Forma 5, elaborado por García y Musitu (2014), y b) tres cuestionarios diseñados especialmente para esta ocasión. El primero de ellos se empleó en la medición de la variable personal (de tipo cognitiva) *estrategias de aprendizaje en Matemática*, el segundo para evaluar la variable contextual (de perspectiva macrosociológica) *aspectos sociofamiliares*, y el tercero con la intención de medir la otra variable de contexto (de carácter educacional) *elementos del clima de clase*.

Los datos de la variable por modelar estuvieron conformados por el promedio de las notas alcanzadas por los alumnos encuestados en tres instancias de evaluaciones parciales escritas teórico-prácticas, relativas al actual régimen de promoción y modalidad de cursado de la asignatura AMI. Esta variable es de tipo continua, sus valores oscilan entre 1 (uno) y 10 (diez) puntos. Sin embargo, al momento de aplicar la técnica análisis discriminante se calcularon los promedios de las notas parciales, los cuales fueron luego recodificados como 1 (valores inferiores a 6) y 2 (valores iguales o superiores a 6).

Cada una de las pruebas que se emplearon para recoger los datos de las variables explicativas (fuentes de información primaria) fueron evaluadas inicialmente en forma cualitativa por un equipo de profesores del Área de Matemáticas del Departamento de Materias Básicas (FRRe-UTN), respecto de dos aspectos característicos en este tipo análisis: a)

indicadores de validez subjetivos o individuales, y b) indicadores de validez factorial o estructural.

Luego de aplicar las encuestas y elaborar la base de datos en formato electrónico, se realizaron análisis estadísticos descriptivos (ítems, puntuación, media $-M-$ y desviación estándar $-DE-$) y estudios del área de la psicometría; *consistencia interna* mediante los coeficientes α (alfa) de Cronbach (Cronbach, 1951) y ω (omega) de McDonald (McDonald, 1970), junto con sus respectivos intervalos de confianza al 95% (límite inferior $-LI-$ y límite superior $-LS-$).

La consistencia interna, es una información de mucha utilidad; puesto que, además de permitir ponderar la fiabilidad (cuantía en que las medidas de las pruebas están libres de errores casuales o aleatorios) de las dimensiones del instrumento utilizado (ver Tabla 3), es considerada un indicador indirecto de la validez de constructo de la prueba y actualmente se evalúa a través de los coeficientes α de Cronbach y ω de McDonald (Vizioli y Pagano, 2022). El primero por ser un estadístico tradicional y el segundo con la intención de atenuar la utilización inadecuada del coeficiente α debido a posibles incumplimientos de sus supuestos paramétricos. Si bien, ambos coeficientes son conceptualmente semejantes, ω es más robusto y apropiado para estimar la consistencia interna cuando las respuestas a los ítems propuestos están planteadas a través de opciones valoradas en forma ordinal, como sucede en el cuestionario aplicado en esta investigación (Béland et al., 2018; Ventura León y Caycho-Rodríguez, 2017).

Tabla 3

Estadísticos descriptivos y de fiabilidad de los instrumentos utilizados.

Instrumento	Descriptivos				Fiabilidad					
	Ítems	Puntuación	M	DE	α	IC al 95%		ω	IC al 95%	
						LI	LS		LI	LS
Autoconcepto académico	6	Mín. = 0.10 Máx. = 8.92	5.03	2.06	.90	.87	.92	.91	.88	.93
Estrategias de aprendizaje	4	Mín. = 1.50 Máx. = 5.00	3.38	0.60	.65	.51	.77	.66	.53	.79
Aspectos sociofamiliares	3	Mín. = 1.00 Máx. = 5.00	2.89	0.91	.78	.66	.87	.83	.74	.93
Elementos del clima de clase	3	Mín. = 1.67 Máx. = 5.00	3.85	0.62	.61	.45	.75	.66	.51	.82

Respecto de la estimación puntual de los coeficientes α de Cronbach y ω de McDonald, se puede señalar que todos se encuentran por encima del valor .60, considerado suficiente en primeras fases de la investigación o estudios exploratorios (Huh et al., 2006; Nunnally, 1978), así como aceptable en instrumentos de medida con menos de 10 ítems (Loewenthal, 1996). Los valores de α y ω fueron calculados mediante el paquete JASP Team (2023).

En líneas generales, en virtud de la revisión cualitativa y del análisis de fiabilidad basado en las estimaciones obtenidas de los coeficientes α y ω , puede afirmarse que las pruebas analizadas poseen un desempeño psicométrico adecuado en el contexto de esta investigación.

2.5. Aspectos del análisis discriminante

A efectos de lograr el objetivo central de esta investigación se implementaron estudios cuantitativos que pertenecen al área de la estadística multivariante. En particular, en el contexto de las técnicas explicativas o de dependencia, como fuera anticipado, se ha elegido el análisis

discriminante (véase Uriel y Aldás, 2005). Se trata de un procedimiento de reducción de datos mediante el cual es posible obtener funciones capaces de separar dos o más grupos de objetos, eventos o individuos (variable dependiente (D): no métrica –de tipo categórica–), tomando como base un conjunto de medidas sobre los mismos representadas por una serie de variables (independientes (X_i): métricas –escala de intervalos o escala de razón–).

La expresión funcional del análisis discriminante (AD) que permite diferenciar los datos muestrales es la que se observa en la ecuación (1):

$$D = \beta_0 + \beta_1X_1 + \beta_2X_2 + \dots + \beta_iX_i + \dots + \beta_pX_p \quad (1)$$

Esta función discriminante se utiliza en caso de que los coeficientes que se obtengan (β_i), luego de analizar los datos, sean no estandarizados. De lo contrario (coeficientes estandarizados), debe emplearse una expresión similar a la ecuación (1), pero sin el término constante (β_0).

Ahora bien, una vez que se obtienen los coeficientes estandarizados, los mismos resultan de utilidad para identificar cuáles son las variables que contribuyen en mayor medida a discriminar las observaciones muestrales. En cambio, los coeficientes no estandarizados se emplean para la construcción de una ecuación que permite obtener la ubicación de los *centroides* (valores medios de cada grupo), así como para clasificar los individuos en los grupos definidos a priori.

En virtud de lo que antecede, la aplicación de la técnica AD permitirá, una vez obtenida la ecuación que mejor se ajuste a la realidad objeto de estudio, clasificar el rendimiento académico de los estudiantes en los dos niveles, aprobados y suspensos, establecidos de antemano (rol explicativo del modelo). También, si fuera de interés, podría emplearse el modelo para caracterizar o diferenciar el rendimiento, en cada grupo, de nuevos individuos pertenecientes a la población de la que proviene la muestra seleccionada (finalidad predictiva del modelo).

Cabe señalar que los supuestos previos del AD son los mismos que los del análisis de regresión múltiple. En especial, debe cumplirse que la distribución de las variables independientes sea normal multivariante. No obstante, existe clara evidencia, tanto matemática como empírica, de que las pruebas con una sola variable dependiente son altamente robustas bajo la violación de los principios de normalidad y homocedasticidad, excepto cuando las muestras son muy pequeñas y desiguales (Hair et al., 2019; Rencher, 2002; Tabachnick y Fidell, 2013).

En el análisis multivariante, sostiene Harris (2001), la tendencia está en considerar que en muestras grandes (en general, mayores a 30 observaciones) sus métodos son suficientemente robustos como para ser insensibles a ligeras desviaciones de los supuestos estadísticos. Este hecho, permite que las técnicas que poseen esta característica sean más flexibles y, por ende, menos restrictivas a la hora de pensar en su posible aplicación, lo que es importante cuando los estudios utilizan evidencias empíricas, como ocurre en esta investigación.

3. RESULTADOS Y DISCUSIÓN

Esta sección se inicia con la presentación de los resultados de ANOVAs univariados, para las cuatro variables independientes de este estudio, los cuales se muestran en la Tabla 4. En la misma, a continuación de lambda (λ) de Wilks, se encuentra el estadístico de prueba F de Fisher, junto con su grado de significación, a través del cual fue posible contrastar la hipótesis nula de igualdad de medias entre los grupos de la variable rendimiento matemático (Tabachnick y Fidell, 2013).

Tabla 4

Prueba de igualdad de medias entre grupos.

Variable	λ	F	Valor p
Autoconcepto académico	.88	17.74	.00
Estrategias de aprendizaje	.89	16.63	.00
Aspectos sociofamiliares	.94	7.86	.01
Elementos del clima de clase	.97	4.12	.04

Se puede ver que para un nivel de significación $\alpha = .05$, las cuatro variables independientes analizadas permiten rechazar la hipótesis nula. Cada una de ellas indica que existe diferencia significativa entre los dos grupos de alumnos que forman parte de la muestra (en todos los casos el estadístico F tiene asociado un valor $p < .05$).

En el test de homocedasticidad del modelo se utilizó el estadístico M de Box = 8.87, con el objeto de contrastar la hipótesis nula de que las matrices de covarianzas para cada grupo de la variable dependiente proceden de la misma población (Hair et al., 2019). En este caso, debido a que M tiene asociada una F de Fisher = .85 con un valor $p = .58$, es posible sostener que no existen evidencias suficientes para rechazar la hipótesis nula, por lo cual se concluye que ambos grupos de alumnos (aprobados y suspensos) presentan la misma variabilidad; uno de los supuestos en los que se fundamenta la técnica, junto con la independencia de las observaciones y la normalidad multivariante de las variables discriminantes.

La prueba λ de Wilks se emplea para determinar si el modelo que se propone es válido (contraste de significación global). Para ello, se calculó $\lambda = .78$ y su valor transformado χ^2 de Pearson = 33.28, al cual le corresponde un valor $p = .00$, motivo por el que se puede decir que se rechaza la hipótesis nula de que los grupos comparados tienen promedios iguales en el conjunto de las cuatro variables discriminantes (Rencher, 2002). Luego, con un nivel de confianza del 95%, se tiene evidencia suficiente para sostener que el modelo posee validez estadística, debido a que el valor p resultó menor que .05.

Los estadísticos M de Box y λ de Wilks, fueron convertidos en valores que siguen distribuciones F de Fisher y χ^2 de Pearson, respectivamente, con el objeto de que sea posible llevar a cabo la prueba de hipótesis inherente a cada uno de ellos.

La Tabla 5 contiene una versión estandarizada (cada una de las variables clasificadoras poseen $M = 0$ y $DE = 1$) de los coeficientes de la función canónica discriminante. Estos coeficientes tienen la característica de ser independientes de la métrica original de las variables predictoras, lo cual evita que haya problemas de interpretación que pueden presentarse cuando los ítems poseen distinta escala de medida, tal como sucede en esta investigación.

Tabla 5

Coefficientes estandarizados de la función canónica discriminante.

Variable	$V_1 =$ Autoconcepto académico	$V_2 =$ Estrategias de aprendizaje	$V_3 =$ Aspectos sociofamiliares	$V_4 =$ Elementos del clima de clase
Magnitud	.64	.48	.37	-.56

La magnitud de los coeficientes estandarizados es un indicador de la importancia que tiene cada variable en el cálculo de la función discriminante. En este sentido, se observa que la variable *Autoconcepto académico* tiene una influencia superior ($\beta_1 = .64$) a la ejercida por las demás variables independientes, a la hora de explicar o predecir el grupo de pertenencia de los estudiantes (Aprobados –A–, Suspensos –S–).

El signo positivo del coeficiente de las variables discriminantes V_1 , V_2 y V_3 , significa que los alumnos con buen rendimiento matemático han optado por seleccionar valores medios o altos a la hora de responder los ítems que respectivamente las conforman.

Por el contrario, el signo negativo del coeficiente de V_4 ($\beta_4 = -.56$) está indicando que los estudiantes con buen rendimiento en Matemáticas puntuaron bajo las respuestas a los ítems que integran esta variable latente.

A partir de la ecuación (1) y utilizando los coeficientes que aparecen en la Tabla 5, la función discriminante relativa al presente estudio es la que se aprecia en la ecuación (2):

$$D_{A-S} = .64 V_1 + .48 V_2 + .37 V_3 - .56 V_4 \quad (2)$$

En la Tabla 6 se observa, para las cuatro variables independientes, la *media* por separado para el grupo de aprobados, $n_1 = 39$, y suspensos, $n_2 = 103$, así como la *magnitud* de los coeficientes no estandarizados de la función discriminante.

Tabla 6

Valores medios y coeficientes no estandarizados de la función discriminante.

Variable	$V_1 =$ Autoconcepto académico	$V_2 =$ Estrategias de aprendizaje	$V_3 =$ Aspectos sociofamiliares	$V_4 =$ Elementos del clima de clase
Aprobados	6.14	3.69	3.13	3.72
Suspensos	4.61	3.26	2.79	3.89
Magnitud	.33	.83	.41	-.90

En caso de utilizar los coeficientes no estandarizados, la función discriminante resultará como se muestra en la ecuación (3) (advertir la presencia del término constante $\beta_0 = -2.18$, el cual es exclusivo para este tipo de función discriminante –ver ecuación (1)–; es decir, no estandarizada):

$$D_{A-S} = -2.18 + .33 V_1 + .83 V_2 + .41 V_3 - .90 V_4 \quad (3)$$

La evaluación de la ecuación (3) en los valores medio de cada variable explicativa en los grupos A y S (definidos a priori en la variable *Rendimiento matemático*), dará la posibilidad de obtener la ubicación de los respectivos centroides (cabe señalar que el programa SPSS brinda estos valores, aunque igualmente se realiza la explicación a efectos de mostrar una interesante utilidad de los coeficientes no estandarizados).

$$D_A = -2.18 + .33 \times 6.14 + .83 \times 3.69 + .41 \times 3.13 - .90 \times 3.72 = .84 \quad (4)$$

$$D_S = -2.18 + .33 \times 4.61 + .83 \times 3.26 + .41 \times 2.79 - .90 \times 3.89 = -.32 \quad (5)$$

Se observa que para el grupo de alumnos *aprobados* se obtiene, en promedio, una puntuación positiva en la función discriminante ($D_A = .84$). En cambio, en el grupo de estudiantes *suspensos*, se obtiene una puntuación media negativa ($D_S = -.32$).

Los centroides –puntuaciones medias de cada grupo a través de la ecuación (3)– resultan de gran utilidad para interpretar los valores de la función discriminante. Así, por ejemplo, si se desconociera el rendimiento académico de un alumno, pero se dispone de los datos de las variables independientes, se podría calcular la puntuación discriminante por medio de la ecuación (3) (finalidad predictiva del modelo discriminante). Posteriormente, a partir del valor que se obtenga, se procedería a clasificar o asignar el individuo al grupo de cuyo centroide se encuentre más próximo. Esto es, si el valor calculado por medio de la ecuación (3) resulta positivo el alumno debería ser clasificado en el grupo de aprobados; por el contrario, si fuera negativo, el estudiante corresponde que sea asignado al grupo de suspensos.

Si bien hasta el momento se ha logrado obtener el modelo discriminante, ecuación (3), así como identificar cuáles son las variables que poseen más poder para explicar el grupo de pertenencia de los individuos, ecuación (2); la mayor utilidad de una función discriminante radica en su capacidad para clasificar nuevos casos.

En efecto, luego de lograda la función discriminante y obtener los centroides, se podría emplearla para efectuar una clasificación de los mismos casos utilizados para estimar los coeficientes de la función (rol explicativo del modelo discriminante).

Con el propósito de contrastar el grado de eficacia de la función canónica discriminante desde el punto de vista de la clasificación, se presenta la Tabla 7. En ella es posible distinguir el número de sujetos correcta e incorrectamente clasificados mediante el procedimiento previamente descrito. La regla de clasificación que ofrece SPSS discrimina sin error 124 estudiantes (28 aprobados y 96 suspensos, se encuentran sobre la diagonal principal de la tabla), que sobre el total de 142 sujetos representan el 87.32%.

Tabla 7

Resultados de la clasificación.

Rendimiento matemático	Aprobados	Suspensos	Total
Aprobados	28 (71.79 %)	11 (28.21 %)	39 (100.00 %)
Suspensos	07 (06.80 %)	96 (93.20 %)	103 (100.00 %)

Se puede sostener, en virtud de los valores alcanzados, que el procedimiento en su conjunto (función discriminante + regla de clasificación) posee una eficacia correcta, motivo por el cual se considera que el diseño metodológico implementado ha sido adecuado y produjo los resultados esperados.

4. CONCLUSIONES

Lo primero que se debe señalar es que la técnica estadística aplicada, perteneciente al área del análisis multivariante, ha permitido lograr el objetivo planteado; esto es, elaborar un modelo con capacidad razonable para *discriminar* cuándo los sujetos deben pertenecer a un grupo o al otro (aprobados y suspensos, definidos de antemano por medio de la variable dependiente), a partir de un conjunto de medidas sobre los individuos (representadas por una serie de variables explicativas propias del sujeto y de su entorno socioeducativo).

Desde el punto de vista de los resultados alcanzados, las variables personales $V_1 = \text{Autoconcepto académico}$ y $V_2 = \text{Estrategias de aprendizaje}$, unidas a las variables contextuales $V_3 = \text{Aspectos sociofamiliares}$ y $V_4 = \text{Elementos del clima de clase}$, tanto en forma individual como en conjunto, son de utilidad para discriminar entre las medias de los grupos de la variable dependiente. La contribución que cada una de ellas realiza, a efectos de diferenciar entre aprobados y suspensos, es diferente; en este sentido, el predictor que más aporta es el *Autoconcepto académico*.

Los coeficientes no estandarizados de las variables V_1 , V_2 y V_3 , son los que, en sentido directo y en el orden indicado, mayores importancias tienen en el cálculo de la función discriminante. El signo positivo que cada uno posee significa que aquellos estudiantes con buen rendimiento en la asignatura AMI eligieron valores medios o altos a la hora de puntuar los ítems de las pruebas respectivas. El signo negativo del coeficiente –de magnitud significativa– de la variable V_4 , señala que los alumnos con buen rendimiento matemático optaron por puntuaciones bajas al responder las preguntas del cuestionario correspondiente.

Las variables explicativas fueron contrastadas empíricamente, a través de los estadísticos F de Fisher y χ^2 de Pearson, resultando todas relevantes a efectos de explicar o predecir los resultados académicos en el contexto de este estudio.

Asimismo, el modelo que se propone permite la formulación de medidas educativas y de apoyo personalizadas, en función de sus variables explicativas, las cuales pueden mejorar el rendimiento académico en Matemáticas y contribuir al éxito de los estudiantes en sus estudios universitarios.

Es necesario tener presente que los datos con los que se llevó a cabo esta investigación pertenecen a estudiantes matriculados en una materia determinada que integra el plan de estudio de ciertas carreras que se ofrecen en un centro académico específico. Debido a ello, resulta necesario ser muy cuidadosos a la hora de extender o extrapolar los resultados y conclusiones logradas a otras poblaciones educativas no representadas en la muestra que permitió implementar este estudio.

No obstante, se considera que la metodología implementada y la aplicación de la técnica análisis discriminante fueron decisiones correctas y un paso adelante en el estudio de la problemática abordada, como también un aporte a la comunidad académica y científica interesada en el tema objeto de estudio.

Finalmente, se estima que este trabajo puede servir como referencia para futuras investigaciones que se realicen en la misma línea, quizás con los matices que el escenario de aplicación demande, las que sin duda serán de utilidad para enriquecer el modelo funcional que en este desarrollo ha sido propuesto.

REFERENCIAS BIBLIOGRÁFICAS

- Arrivillaga, F., García, M. L. y Maldonado, N. P. (2023). El autoconcepto académico en matemáticas: ruta hacia una categorización a través del método de análisis conceptual. *Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, 13(26), 1-22.
- Béland, S., Cousineau, D., & Loye, N. (2017). Utiliser le coefficient omega de McDonald à la place de l'alpha de Cronbach. *McGill Journal of Education*, 52(3), 791-804.
- Beltrán, J. A. (2003). Estrategias de Aprendizaje. *Revista de Educación*, 3(332), 55-73.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.

- Delors, J. (1997). *La educación encierra un tesoro. Informe a la UNESCO de la Comisión Internacional sobre la Educación para el Siglo XXI*. Ediciones UNESCO.
- Devincenzi, G. H., Rohde, G. A., Giraudó, M. B. V., Bonaffini, M. L. y Bernaola, G. A. (2017). El análisis del rendimiento académico a través de modelos matemáticos y estadísticos. *Tordesillas Revista de Investigación Multidisciplinar*, 13(1), 59-72.
- Doménech, F., Jara, P. y Rosel, J. (2004). Percepción del proceso de enseñanza/aprendizaje desarrollado en Psicoestadística I y su incidencia en el rendimiento. *Psicothema*, 16(1), 32-38.
- García, F. y Musitu, G. (2014). *AF5. Autoconcepto Forma 5* (4a ed.). TEA.
- George, D. y Mallery, P. (2020). *IBM SPSS Statistics 26 Step by Step* (16th ed.). Routledge.
- González-Pienda, J. A., Núñez, J. C., Álvarez, L., Rocés, C., González-Pumariega, S., González, P., Muñiz, R., Valle, A., Cabanach, R. G., Rodríguez, S. y Bernardo, A. (2003). Adaptabilidad y cohesión familiar, implicación parental en conductas autorregulatorias, autoconcepto del estudiante y rendimiento académico. *Psicothema*, 15(3), 471-477.
- Hair, J. F., Black, W. C., Babin, B. J. y Anderson, R. E. (2019). *Multivariate Data Analysis* (8th ed.). Cengage Learning.
- Harris, R. J. (2001). *A Primer of Multivariate Statistics* (3rd ed.). Psychology Press.
- Huh, J., Delorme, D. E. y Reid, L. N. (2006). Perceived third-person effects and consumer attitudes on preventing and banning DTC advertising. *Journal of Consumer Affairs*, 40(1), 90-116.
- JASP Team (2023). JASP (Version 0.17.2.1) [Computer software]. <https://jasp-stats.org/>
- Kassarnig, V., Mones, E., Bjerre-Nielsen, A., Sapiezynski, P., Dreyer Lassen, D. y Lehmann, S. (2018). Academic performance and behavioral patterns. *EPJ Data Science*. 7(10), 1-16.
- Loewenthal, K. M. (1996). *An introduction to psychological tests and scales*. UCL Press Limited.
- Lohr, S. L. (2019). *Sampling: Design and Analysis* (3rd ed.). Chapman and Hall/CRC.
- McDonald, R. P. (1970). Theoretical foundations of principal factor analysis, canonical factor analysis, and alpha factor analysis. *British Journal of Mathematical and Statistical Psychology*, 23(1), 1-21.
- Nunnally, J. C. (1978). *Psychometric Theory* (2nd ed.). McGraw-Hill.
- Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology*, 95(4), 667-686.
- Rencher, A. C. (2002). *Methods of Multivariate Analysis* (2nd ed.). John Wiley & Sons.
- Reynolds, A. J. y Walberg, H. J. (1992). A Process Model of Mathematics achievement and attitude. *Journal for Research in Mathematics Education*, 23(4), 306-328.
- Sánchez León, A. F. (2023). Self-concept and academic performance of university students. *University, Science and Technology*, 27(118), 61-68.
- Schnitzler, K., Holzberger, D., & Seidel, T. (2021). All better than being disengaged: Student engagement patterns and their relations to academic self-concept and achievement. *European Journal of Psychology of Education*, 36(3), 627-652.
- Tabachnick, B. G. y Fidell, L. S. (2013). *Using Multivariate Statistics* (6th ed.). Pearson.
- Uriel, E. y Aldás, J. (2005): *Análisis Multivariante Aplicado*. Thomson.
- Ventura León, J. y Caycho-Rodríguez, T. (2017). El coeficiente omega: un método alternativo para la estimación de la confiabilidad. *Revista Latinoamericana de Ciencias Sociales, Niñez y Juventud*, 15(1), 625-627.
- Vizioli, N. y Pagano, A. (2022). De alfa a omega: estimación de la confiabilidad ordinal. Una guía práctica. *Revista Costarricense de Psicología*, 41(2), 119-136.

Closas, A. H., Arriola, E. A., Amarilla, M. R. y Jovanovich, E. C.

Voght, W. P. (2005). *Dictionary of Statistics and Methodology* (3rd ed.). Sage.