

Articulo-Ipince Daniel-Miñope Miguel.docx

 Universidad Peruana Union

Detalles del documento

Identificador de la entrega

trn:oid:::29566:464613200

Fecha de entrega

4 jun 2025, 3:08 p.m. GMT-5

Fecha de descarga

4 jun 2025, 3:11 p.m. GMT-5

Nombre de archivo

Articulo-Ipince Daniel-Miñope Miguel.docx

Tamaño de archivo

174.1 KB

18 Páginas

5745 Palabras

32.570 Caracteres

15% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Small Matches (less than 8 words)

Exclusions

- 16 Excluded Sources

Top Sources

- 12%  Internet sources
- 4%  Publications
- 12%  Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Top Sources

- 12% Internet sources
- 4% Publications
- 12% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Internet	repositorio.uptc.edu.co	<1%
2	Submitted works	Universidad Cesar Vallejo on 2022-12-22	<1%
3	Internet	cybertesis.unmsm.edu.pe	<1%
4	Submitted works	Universidad Cesar Vallejo on 2016-04-01	<1%
5	Submitted works	Universidad Continental on 2019-03-02	<1%
6	Submitted works	Universidad Tecnica De Ambato- Direccion de Investigacion y Desarrollo , DIDE o...	<1%
7	Internet	documents.mx	<1%
8	Internet	repositorio.utp.edu.co	<1%
9	Internet	www.researchgate.net	<1%
10	Internet	repositorio.upn.edu.pe	<1%
11	Internet	www.interciencia.net	<1%

12	Internet	www.scielo.sa.cr	<1%
13	Submitted works	Centro Educativo Particular Mixto Reina del Mundo on 2014-07-14	<1%
14	Submitted works	Universidad Pedagógica y Tecnológica de Colombia on 2018-08-22	<1%
15	Submitted works	Universidad Santo Tomás on 2015-05-19	<1%
16	Submitted works	Universidad del Pacífico - Escuela de Negocios on 2023-06-06	<1%
17	Internet	dspace.ucuenca.edu.ec	<1%
18	Internet	redi.unjbg.edu.pe	<1%
19	Internet	repositorio.ucsm.edu.pe	<1%
20	Publication	Jose Manuel Vazquez. "<![CDATA[Application of statistical models to the efficient ...	<1%
21	Internet	inscricao.faculdadeitop.edu.br	<1%
22	Internet	repositorio.upt.edu.pe	<1%
23	Publication	Judith Cavazos-Arroyo, Alejandro Melchor-Ascencio. " The influence of greenfluen...	<1%
24	Submitted works	Pontificia Universidad Católica del Perú on 2018-10-12	<1%
25	Submitted works	Universidad Anahuac México Sur on 2015-08-02	<1%

26	Internet	redie.ens.uabc.mx	<1%
27	Internet	rideca.cs.buap.mx	<1%
28	Internet	rsr.akvo.org	<1%
29	Internet	www.jove.com	<1%
30	Internet	www.uajournals.com	<1%
31	Submitted works	Mondragon Unibertsitatea on 2023-03-28	<1%
32	Submitted works	Universidad Internacional de la Rioja on 2015-07-22	<1%
33	Submitted works	Universidad TecMilenio on 2025-03-10	<1%
34	Internet	editorial.uicui.edu.mx	<1%
35	Internet	gala.gre.ac.uk	<1%
36	Internet	www.revistacomunicar.com	<1%
37	Publication	Diego Danilo Rodríguez Rosero, Ruber Erlinton Ordoñez Ortega, Mario Eduardo H...	<1%
38	Submitted works	Mondragon Unibertsitatea on 2024-12-13	<1%
39	Submitted works	Universidad Cesar Vallejo on 2024-11-29	<1%

40	Submitted works	Universidad Nacional de Trujillo on 2025-01-03	<1%
41	Submitted works	Universidad de Cantabria on 2023-11-24	<1%
42	Internet	aunarcali.edu.co	<1%
43	Internet	de.slideshare.net	<1%
44	Internet	es.slideshare.net	<1%
45	Internet	periodicos.ufpe.br	<1%
46	Internet	publicaciones.ucuenca.edu.ec	<1%
47	Internet	puertomaderoeditorial.com.ar	<1%
48	Internet	revzoilomarinello.sld.cu	<1%
49	Internet	silوtips	<1%
50	Submitted works	unsaac on 2025-05-30	<1%
51	Internet	www.haaj.org	<1%

Factores que influyen en el rendimiento académico en estudiantes del tercer ciclo de una universidad privada en Lima

Factors influencing academic performance in third cycle students of a private university in Lima.

Daniel Alberto Ipince Antunez, Miguel Angel Miño Gomez, Javier Linkolk López
Gonzales

Resumen

Objetivo: Determinar los factores que influyen en el rendimiento académico de los estudiantes del tercer ciclo de una universidad privada en Lima. **Metodología:** Enfoque cuantitativo, con diseño no experimental, de corte transversal y nivel explicativo. La muestra fue de 234 estudiantes, seleccionados mediante muestreo aleatorio simple. Se utilizaron técnicas de recolección documental y encuesta. **Resultados:** El modelo completo explicó el 13.6% de la variabilidad del rendimiento académico, solo un número reducido de variables predictoras presentó efectos estadísticamente significativos. El modelo *stepwise* mostró mejor ajuste ($AIC = 308.797$) en comparación con el modelo completo, manteniendo una capacidad predictiva adecuada ($Pseudo R^2 = 0.101$). **Conclusión:** Existe influencia significativa de los factores en el rendimiento académico en estudiantes del tercer ciclo de una universidad privada en Lima. **Palabras clave:** Rendimiento académico, regresión logística ordinal, factores, educación superior, análisis multivariado.

Abstract

Objective: To determine the factors that influence the academic performance of students in the third cycle of a private university in Lima. **Methodology:** Quantitative approach, with non-experimental design, cross-sectional and explanatory level. The sample consisted of 234 students, selected by simple random sampling. Documentary collection and survey techniques were used. **Results:** The complete model explained 13.6% of the variability of academic performance, only a small number of predictor variables showed statistically significant effects. The stepwise model showed better fit ($AIC = 308.797$) compared to the full model, maintaining an adequate predictive capacity ($Pseudo R^2 = 0.101$). **Conclusion:** There is a significant influence of the factors on academic performance in third cycle students of a private university in Lima. **Keywords:** Academic performance, ordinal logistic regression, factors, higher education, multivariate analysis.

Introducción

El rendimiento académico en la educación superior ha sido, en las últimas décadas, una de las preocupaciones centrales para universidades e investigadores en todo el mundo, debido a su papel decisivo en la formación profesional y en la mejora continua de la calidad educativa. A nivel internacional, estudios como los de Gutiérrez, Garzón y Segura (2021) en Colombia, Martínez et al. (2020) en México, y Barahona (2014) en Chile, identificaron múltiples factores, como los pedagógicos, institucionales, sociodemográficos y personales, que incidieron significativamente en el desempeño de los estudiantes universitarios, utilizando modelos estadísticos como la regresión logística y el análisis discriminante.

En el contexto peruano, esta preocupación también cobró relevancia, especialmente en universidades privadas, donde la diversidad de trayectorias estudiantiles, los distintos niveles de preparación previa y el entorno socioeconómico desafiaron la obtención de un rendimiento académico homogéneo. En una universidad privada de Lima, se evidenció que los estudiantes del tercer ciclo presentaban niveles de desempeño variados, lo que planteó la necesidad de identificar con mayor precisión los factores que influyen en sus logros académicos. Bajo esta premisa, la presente investigación partió de la siguiente interrogante: ¿Cuáles son los factores que influyen en el rendimiento académico en estudiantes del tercer ciclo de una universidad privada en Lima?

La justificación del estudio se fundamenta en tres dimensiones clave. Desde una perspectiva práctica, mejorar el rendimiento académico conlleva beneficios individuales y colectivos. Factores personales como la motivación, la disciplina y la salud emocional, junto a elementos contextuales como el entorno familiar, el acceso a servicios de apoyo y la percepción del clima institucional, influyen en el éxito estudiantil. Diversos autores como Gutiérrez, Garzón y Segura (2021) y Tinto (1993) han señalado que la combinación de estos factores puede determinar el abandono, la permanencia o el éxito en la trayectoria universitaria.

Diversos estudios han abordado el fenómeno del rendimiento académico mediante modelos de regresión. Martínez et al. (2020) utilizaron un modelo predictivo en estudiantes de Medicina, hallando que repetir escuela, estudiar menos de 15 horas por semana y el puntaje de ingreso eran variables significativas. Heredia, Rodríguez y Vilalta (2012) emplearon regresión logística ordinal en estudiantes de Ingeniería, encontrando que el promedio en asignaturas científicas del primer año era un predictor confiable del rendimiento futuro. Otros estudios relevantes son los de Barahona (2014), que halló la influencia del género y la satisfacción con la carrera; Pérez y Romero (2018), que usaron

regresión logística y múltiple para correlacionar rendimientos entre áreas; y Vitola (2015), quien identificó variables socioeconómicas y familiares como determinantes en Matemáticas. Todos ellos coinciden en que el rendimiento académico no depende de un único factor, sino de una red compleja de variables que actúan en conjunto o de manera mediada.

El objetivo fue determinar los factores que influyen en el rendimiento académico en estudiantes del tercer ciclo de una universidad privada en Lima. Asimismo, como objetivos específicos, se tuvo en cuenta identificar los factores contextuales y personales que influyen en el rendimiento académico.

En función del problema planteado y de los antecedentes teóricos revisados, se formuló la hipótesis general que orientó la presente investigación: existe influencia significativa de los factores en el rendimiento académico en estudiantes del tercer ciclo de una universidad privada en Lima.

Metodología

El presente estudio de investigación es un estudio que presentó un diseño no experimental debido a que las variables no fueron manipuladas de manera deliberada, asimismo, fue de tipo básica porque los aportes y hallazgos contribuyeron a la teoría del constructo de la variable, también se estableció en un enfoque cuantitativo porque utilizó la estadística para obtener los resultados y arribar nuevos postulados, además fue de nivel explicativo porque permitió dar respuesta mediante el análisis de regresión logística para predecir los factores que influyeron en el rendimiento académico y, tuvo corte transversal porque el estudio se realizó en un solo momento o tiempo único para la recolección de datos (Hernández & Mendoza, 2018). La variable independiente se representó mediante los factores contextuales y personales en estudiantes. Asimismo, la variable dependiente se estableció por el rendimiento académico en estudiantes.

Para determinar el tamaño de la muestra en una población finita, se utilizó fórmulas basadas en la estimación de proporciones. En el estudio, se parte de una población de $N = 600$ estudiantes, y se asume que la proporción (p) es de 0.50 (lo que maximiza la variabilidad), un nivel de confianza del 95% (lo que implica un valor Z de 1.96) y se desea una precisión (error muestral) $d = 5\%$, lo que es común en estudios educativos (Hernández & Mendoza, 2018; Cochran, 1977).

Dado que la población fue finita ($N = 600$), se aplicó la corrección para población finita. Este procedimiento se fundamentó en los métodos de muestreo probabilístico (Krejcie & Morgan, 1970; Cochran, 1977) y es ampliamente utilizado en investigaciones con

población conocida. Además, el muestreo aleatorio simple se aplicó para asegurar que cada estudiante de la población tenga la misma probabilidad de ser seleccionado, lo que incrementa la representatividad de la muestra (Hernández & Mendoza, 2018). Luego de ello, se utilizó la fórmula de Cochran para una muestra infinita y aplicando la corrección para población finita, se determinó que, para una población de 600 estudiantes, con una precisión del 5% ($d = 0.05$), una proporción de 50% ($p = 0.50$) y un nivel de confianza del 95% ($Z = 1.96$), el tamaño de muestra adecuado es de aproximadamente 234 estudiantes.

Para la recolección de datos de los 234 estudiantes universitarios se ha elegido la técnica documental y encuesta, para ello, los documentos de las calificaciones de los estudiantes fueron provistos de las diferentes asignaturas en el ciclo académico 2023 – I, las cuales se obtuvieron de la base de datos del área de estudios generales, además, se aplicó una encuesta a cada estudiante vía correo sobre los aspectos contextuales y personales

Ponsot et al., (2009) como se citó en Heredia, Rodríguez & Vilalta (2014) sostienen que la regresión logística de acuerdo a su expresión más simple, posee una respuesta de tipo binaria, es decir, propone que el logaritmo de “razón de probabilidad” (*odds*, según su denominación en inglés) es resultado del cociente entre la probabilidad de éxito y fracaso para un ensayo de Bernoulli, por ello como principio describe una función lineal bajo los parámetros denominados predictora lineal. Asimismo, en el caso particular de la regresión logística ordinal, se empleó una función de unión para establecer una relación de modo lineal con las variables explicativas a razón de probabilidad acumulada hasta la categoría “i” en la variable ordinal; y la probabilidad que tome un valor mayor en dicha categoría (Agresti, 2007; Hosmer & Lemeshow, 2000).

Por otro lado, en la regresión logística ordinal, las funciones mayores empleadas son de tipo: *Logit* y *Cloglog* (Agresti, 2007). De acuerdo con McCullagh (1980) propuso que la unión con *Logit*, es adecuada para el análisis de datos ordinales, teniéndose en cuenta una distribución de frecuencias uniforme, con la extensión de las categorías existentes; mientras que la unión *Cloglog* se utiliza para el análisis de datos categóricos, cuyo mayor valor son las categorías con mayor probabilidad.

En el presente estudio, los valores utilizados para la variable ordinal se encontraron representados con las calificaciones de logro, propuestas según el Ministerio de Educación del Perú (MINEDU): “En inicio; En proceso; Logro esperado y Logro destacado”; siendo para ello preciso emplear el Logit, como el más adecuado. Por ello, en vista que los datos a presentar son variables numéricas y categóricas, es

característica de un estudio multivariado. Para el procesamiento de las variables y su análisis, se utilizó el *Software Rstudio*.

Además, dicho análisis se realizó empleando la estadística descriptiva de cada variable continua, determinando las medidas de tendencia, variabilidad y asimetría, así como los diagramas de caja. Además, cada variable categórica se relacionó con la variable respuesta mediante la Prueba Chi - Cuadrado de Independencia. Navarro et al., (2014) mencionan que, cuando la variable respuesta posee más de dos (2) categorías, el modelo logístico se denomina polinómico. Al respecto, en la implementación de un modelo logístico polinómico se debe seleccionar un nivel para la variable respuesta; teniendo en cuenta el nivel de referencia con el que se comparan los demás niveles. Es decir, si se cambian los niveles de referencia, no se alteran los modelos, sin embargo, cambian las interpretaciones de los parámetros. Además, para el modelo logístico polinómico con variable respuesta $Y = \{0, 1, 2\}$; K variables explicativas representadas por $X = X_1, \dots, X_K$ y 0; como categoría de referencia, se tienen en cuenta para modelos logísticos.

El estudio se aplicó en una universidad privada de Lima con autorización institucional. La recolección de datos se realizó de forma virtual y automatizada. Además, se cumplió rigurosamente con las normas APA, garantizando el respeto a los derechos de autor y los principios éticos de autonomía y justicia en la investigación científica.

Resultados

De acuerdo con la tabla 1, la mayoría de los predictores presentaron coeficientes no significativos ($p > 0.05$), como es el caso de la fila 1 (coef. = 0.030, $p = 0.495$), que mostró un efecto mínimo sobre la probabilidad de obtener una mejor respuesta en la pregunta 9. Variables de las filas 1 a 15, 17 a 21, 23, 25, 26 y 27 presentaron coeficientes cercanos a cero o moderados, sin evidencia estadística suficiente para rechazar la hipótesis nula. Se identificaron también predictores con p-valores cercanos al umbral de significancia, como las filas 2 (0.720, $p = 0.093$), 16 (0.614, $p = 0.064$) y 20 (-0.500, $p = 0.102$), que mostraron tendencias a tener un efecto, aunque sin alcanzar significancia al 5%, por lo que su influencia debe considerarse con cautela. En contraste, algunos predictores sí fueron estadísticamente significativos ($p < 0.05$). La fila 22 (coef. = 0.598, $p = 0.019$) indicó un efecto positivo moderado, con un *odds ratio* (OR) ≈ 1.82 , lo que sugiere que un incremento en esta variable aumentó las probabilidades de obtener una mejor categoría en la pregunta 9. Más relevante aún fue la fila 24 (coef. = 1.584, $p = 0.014$), con un OR ≈ 4.88 , que reflejó un efecto sustancial. Finalmente, los coeficientes

de las filas 28 (10.196), 29 (12.255) y 30 (15.009), con valores p de 0.036, 0.012 y 0.002, respectivamente, fueron altamente significativos, lo que sugiere un impacto extremadamente fuerte. Los *odds ratios* correspondientes, $\exp(10.196)$, $\exp(12.255)$ y $\exp(15.009)$, fueron muy elevados, lo que confirma que estas variables fueron predictoras clave en el modelo.

De acuerdo con la tabla 2, la mayoría de los intervalos de confianza al 95 % incluyeron el valor cero, lo que indica que no se pudo rechazar la hipótesis nula de efecto nulo para esos predictores. Por ejemplo, los intervalos correspondientes a las preguntas 1, 2, 3, 4, 5, 6, 8, 13, 14, 16 a 19, 21, 22, 24 a 29, 31, 33 a 35, mostraron efectos pequeños o inciertos, con rangos que abarcaron tanto valores negativos como positivos. Esta evidencia sugiere que dichos predictores no tuvieron una influencia estadísticamente significativa sobre la probabilidad de que la respuesta a la pregunta 9 se ubicara en una categoría superior. Sin embargo, se identificaron dos predictores estadísticamente significativos, cuyos intervalos no incluyeron el cero. El intervalo 30 (0.101, 1.105) indicó un efecto positivo moderado, lo que implica que un aumento en esa variable se asoció con una mayor probabilidad de obtener una mejor calificación en la pregunta 9. Por su parte, el intervalo 32 (0.334, 2.881) evidenció un efecto más fuerte y consistente en la misma dirección.

De acuerdo con la tabla 3, se observó que la mayoría de los *odds ratios* (OR) obtenidos para los predictores analizados presentaron intervalos de confianza al 95 % que incluyeron el valor 1, lo cual indica que no se pudo comprobar estadísticamente un efecto significativo de estas variables, de manera individual, sobre la probabilidad de que la respuesta a la pregunta 9 se ubicara en una categoría superior. Específicamente, variables como las preguntas 1, 3, 4, 5, 6, 8, 13, 14, 16 al 29, 31, 33, 34 y 35 mostraron *odds ratios* con efectos leves o moderados, tanto positivos como negativos, pero con rangos amplios e intervalos que abarcaron el valor nulo. Esta condición impidió rechazar la hipótesis nula y, por ende, afirmar una influencia estadísticamente significativa. En contraste, se identificaron dos predictores que sí alcanzaron significación estadística. La pregunta 30 presentó un *odds ratio* de 1.818, con un intervalo de confianza de 1.107 a 3.019, lo cual implica que, por cada unidad de incremento en esta variable, las probabilidades de que un estudiante obtuviera una mejor categoría en la pregunta 9 aumentaron en un 82 %, con un nivel de confianza del 95 %. Más notable aún fue la pregunta 32, cuyo *odds ratio* fue de 4.874 (IC: 1.397 a 17.835), lo que indica que un incremento en esta variable se asoció con un aumento de casi 4.9 veces en las *odds ratios* de alcanzar una calificación más alta. Este hallazgo la posiciona como el predictor

con mayor impacto dentro del modelo. En suma, aunque varios predictores evidenciaron tendencias positivas, como en los casos de las preguntas 2, 24 y 31, sus intervalos de confianza incluyeron el valor 1, lo que impidió afirmar un efecto significativo. Por lo tanto, solo las variables correspondientes a las preguntas 30 y 32 mostraron una influencia clara, positiva y estadísticamente comprobada, convirtiéndose en factores relevantes para explicar el incremento en la percepción de logro reflejada en la respuesta a la pregunta 9.

De acuerdo con la tabla 4, los valores del Factor de Inflación de la Varianza (VIF) oscilaron entre 1.416 (para la pregunta 35) y 3.767 (para la pregunta 1). La mayoría de los predictores presentaron valores de VIF comprendidos entre 1.4 y 2.5, lo cual indica un nivel de multicolinealidad bajo a moderado. Solo algunas variables, como la pregunta 1 (VIF = 3.767), la pregunta 6 (VIF = 3.099) y la pregunta 22 (VIF = 2.330), mostraron valores algo más elevados, pero, aun así, se mantuvieron por debajo del umbral crítico de 5, ampliamente aceptado en la literatura. Estos resultados permiten concluir que no se detectaron problemas de multicolinealidad severa entre los predictores del modelo, por lo tanto, la correlación existente entre las variables independientes no compromete la estabilidad ni la interpretación de los coeficientes estimados.

De acuerdo con la tabla 5, en cuanto a la observación 1 resultó ser la más influyente dentro del modelo, al presentar un valor de 0.730. Este valor indicó que, de haberse eliminado dicha observación, los coeficientes del modelo habrían variado significativamente, lo que permitió suponer que presentaba características atípicas o valores extremos en alguna variable. Le siguieron en influencia las observaciones 34 y 68, con valores de 0.577 y 0.568 respectivamente, las cuales también generaron un impacto considerable en la estimación de los parámetros. Asimismo, las observaciones 97, 17 y 26, con valores comprendidos entre 0.510 y 0.562, se ubicaron en el límite de ser consideradas puntos influyentes, ya que su exclusión podría haber alterado la dirección o magnitud de algunos coeficientes. Por último, las observaciones 116, 42, 61 y 117, aunque con menor grado de influencia, también habrían generado variaciones en las estimaciones del modelo.

De acuerdo con lo mostrado en la tabla 6, se compararon tres modelos de regresión logística ordinal en función de las preguntas incluidas como predictores y sus respectivas medidas de ajuste: AIC, BIC y el Pseudo R^2 de McFadden. El modelo completo incluyó las 27 preguntas independientes consideradas en el cuestionario. Si bien este modelo explicó el 13.6 % de la variabilidad de la variable dependiente (pregunta 9), presentó valores elevados tanto en el AIC (337.465) como en el BIC

(423.949), lo que reflejó un alto nivel de complejidad. Por otro lado, el modelo reducido solo consideró cinco predictores: las preguntas 30, 32, 2, 16 y 24. Este modelo mostró una mejora sustancial en los indicadores de ajuste ($AIC = 314.433$, $BIC = 337.495$), lo que indicó mayor parsimonia. No obstante, su capacidad explicativa fue la más baja ($Pseudo R^2 = 0.071$), lo que implicó una pérdida considerable en la variación explicada respecto al modelo completo. Finalmente, el modelo generado mediante el procedimiento *stepwise* incluyó siete predictores seleccionados automáticamente: las preguntas 2, 24, 28, 29, 30, 32 y 33. Este modelo obtuvo el AIC más bajo (308.797) y un BIC de 337.625, lo que señaló el mejor equilibrio entre ajuste y simplicidad. Aunque su $Pseudo R^2$ fue de 0.101, menor que el del modelo completo, superó al del modelo reducido y conservó una capacidad explicativa razonable. En conjunto, se concluyó que el modelo *stepwise* representó la opción más adecuada para los datos analizados, al mantener un compromiso óptimo entre eficiencia, parsimonia y poder explicativo.

En la figura 2 se observó que el Área Bajo la Curva (AUC) fue de 0.787, lo cual indicó que el modelo presentó un rendimiento moderado a bueno en términos de capacidad predictiva. Para contextualizar este valor, se debe considerar que un AUC de 0.5 representa un modelo sin capacidad de discriminación, equivalente a una predicción aleatoria. Valores entre 0.7 y 0.8 se interpretan como aceptables, entre 0.8 y 0.9 como buenos, y superiores a 0.9 como excelentes. Por tanto, el valor obtenido de 0.787 permitió concluir que el modelo fue capaz de clasificar correctamente, es decir, de distinguir entre casos positivos y negativos, en aproximadamente el 79 % de las ocasiones, lo que se considera un desempeño adecuado para fines prácticos y de análisis.

Discusión

El presente estudio tiene como objetivo identificar y analizar los factores contextuales y personales que influyen en el rendimiento académico de los estudiantes del tercer ciclo en una universidad privada en Lima, mediante la aplicación de un modelo de regresión logística ordinal. Los resultados obtenidos permiten señalar que, si bien el modelo completo, que incluye una amplia gama de variables contextuales y personales, logra explicar aproximadamente el 13.6 % de la variación en el rendimiento académico, solo un número reducido de predictores muestra asociaciones estadísticamente significativas.

En particular, los análisis revelan que variables como Pgta30 y Pgta32 resultan determinantes en la probabilidad de que un estudiante obtenga una categoría superior

51 en la variable de rendimiento académico (Pregunta 9). La significación de estos predictores, evidenciada por intervalos de confianza que no incluyen el valor neutro y *odds ratios* claramente superiores a 1 ($\exp(0.598) \approx 1.82$ para Pgta30 y $\exp(1.584) \approx 4.88$ para Pgta32), sugiere que aspectos relacionados con la experiencia académica y el uso de servicios de apoyo institucional, tal como se ha interpretado en otros estudios (por ejemplo, Martínez et al., 2020; Heredia, Rodríguez y Vilalta, 2012), pueden ejercer un impacto directo y significativo en el rendimiento. Estos hallazgos coinciden con la perspectiva de Hosmer, Lemeshow y Sturdivant (2013), quienes destacan la importancia de identificar predictores clave mediante modelos de regresión logística, al permitir una mejor comprensión de la relación entre factores individuales y el éxito académico.

Por otro lado, la incorporación de múltiples variables en el modelo completo, si bien incrementa el poder explicativo (Pseudo $R^2 = 0.136$), también añade complejidad, lo que se refleja en valores elevados del AIC y BIC (337.465 y 423.949, respectivamente). Este fenómeno ha sido descrito en investigaciones previas (Burnham y Anderson, 2004; Kuhn y Johnson, 2013), que señalan que un modelo más complejo puede aumentar ligeramente la capacidad explicativa, pero a costa de perder parsimonia y estabilidad en la estimación de los parámetros. En este sentido, el modelo *stepwise*, que incluyó un número intermedio de predictores, mostró un mejor equilibrio entre ajuste y simplicidad (AIC = 308.797; BIC = 337.625; Pseudo $R^2 = 0.101$), lo que indica que una selección cuidadosa de variables puede proporcionar una herramienta predictiva más robusta sin sobrecargar el modelo con información redundante.

La evaluación de la multicolinealidad a través del VIF arrojó valores generalmente bajos, entre 1.416 y 3.767, lo cual indica que la correlación entre los predictores no fue excesiva y, por tanto, no comprometió la interpretación de los coeficientes. Este hallazgo contribuye a la validez interna del modelo, al permitir que los efectos individuales sean interpretados con mayor confiabilidad (Agresti, 2007). No obstante, es importante señalar que, aunque algunos predictores no mostraron efectos estadísticamente significativos de forma individual, ello no implica necesariamente que sean irrelevantes en el contexto global del rendimiento académico. Estudios como los de Tinto (1993) y Navarro et al. (2014) han advertido que los factores contextuales y personales pueden interactuar de manera compleja, ejerciendo influencias indirectas o mediadas en los resultados académicos. Por ejemplo, variables relacionadas con el apoyo familiar o la participación en actividades extracurriculares podrían no tener un efecto directo significativo, pero sí aportar al modelo mediante mecanismos de mediación o moderación de otras variables.

Asimismo, la curva ROC, con un valor de AUC de 0.787, evidencia que el modelo tuvo una capacidad moderada para discriminar entre los distintos niveles de rendimiento académico. Aunque un AUC cercano a 0.8 se considera aceptable, este resultado indica que aún existe margen de mejora en la precisión predictiva del modelo, posiblemente mediante la incorporación de nuevas variables o mediante el uso de modelos más complejos, como los análisis de ecuaciones estructurales (SEM), que permiten explorar relaciones interdependientes entre factores (Field, 2013).

Los hallazgos del estudio tienen importantes implicancias prácticas. En primer lugar, la identificación de predictores clave como Pgta30 y Pgta32 sugiere que las intervenciones dirigidas a fortalecer la experiencia educativa, como mejorar el acceso y la calidad de los servicios de apoyo académico. Además, la comparación entre modelos destaca la importancia de mantener un equilibrio entre complejidad y parsimonia, ya que un modelo excesivamente complejo puede dificultar su interpretación y aplicación en contextos reales, mientras que un modelo demasiado simplificado podría no captar la totalidad de los factores que inciden en el éxito académico.

Conclusiones

El estudio permitió determinar que el rendimiento académico de los estudiantes del tercer ciclo en una universidad privada de Lima está influenciado por un conjunto limitado de factores contextuales y personales. Aunque el modelo completo explicó hasta un 13.6 % de la variabilidad en el rendimiento, solo dos predictores (Pgta30 y Pgta32) mostraron efectos estadísticamente significativos, lo cual resalta la importancia de ciertos aspectos vinculados a la experiencia educativa y el uso de servicios de apoyo académico.

Se identificó que, entre los factores contextuales, aquellos relacionados con la percepción sobre el entorno institucional y los servicios de apoyo (por ejemplo, Pgta32) tuvieron un impacto significativo en el rendimiento académico. Estos resultados coinciden con estudios previos que destacan la relevancia del acompañamiento académico como elemento clave para el éxito educativo.

Respecto a los factores personales, solo una de las variables (Pgta30) mostró influencia significativa sobre el rendimiento, sugiriendo que ciertos componentes individuales, como el compromiso o la autogestión del aprendizaje, podrían tener un papel relevante. Sin embargo, muchos factores personales evaluados no presentaron significancia estadística individual, lo que indica la necesidad de estudios futuros que analicen posibles interacciones o efectos mediados.

Referencias bibliográficas

- Agresti, A. (2007). Regresión logística. *Introducción al análisis de datos categóricos*. (2ª ed., págs. 99-136). <http://dx.doi.org/10.1002/9780470114759.ch5>
- Agresti, A. (2010). *Analysis of ordinal categorical data*. Wiley.
- Artunduaga M. (2008). *Variables que influyen en el rendimiento académico*. (Tesis doctoral). España; Universidad Complutense de Madrid.
- Barahona, P. (2014). Factores determinantes del rendimiento académico de los estudiantes de la Universidad de Atacama. *Estudios Pedagógicos*. Vol. 40, No. 1. Chile., pp. 25-39. <https://scielo.conicyt.cl/pdf/estped/v40n1/art02.pdf>
- Burnham, K. P., & Anderson, D. R. (2004). *Model selection and multimodel inference: A practical information-theoretic approach*. Springer Science & Business Media.
- Dávila, N., García, M. D., Pérez, J. M., & Gómez, E. (2015). Un modelo de regresión logística asimétrico que puede explicar la probabilidad de éxito en el rendimiento académico. *Revista de Investigación Educativa*, 33(1), 27–45. <https://doi.org/10.6018/rie.33.1.178481>
- Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? *The Journal of Machine Learning Research*, 15(1), 3133-3181.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Sage.
- García, M. V., Alvarado, J. M., & Jiménez, A. (2000). *La predicción del rendimiento académico: regresión lineal versus regresión logística*. *Psicothema*. 2000, v. 12, Supl. N° 2, p. 248-252 <https://redined.educacion.gob.es/xmlui/handle/11162/4702>
- Gutiérrez, J. A., Garzón, J., & Segura, A. M. (2021). Factores asociados al rendimiento académico en estudiantes universitarios. *Formación universitaria*, 14 (1), 13-24. <https://dx.doi.org/10.4067/S0718-50062021000100013>
- Heredia, J. J., Rodríguez, A. G., & Vilalta, J. A. (2012). Empleo de la regresión logística ordinal para la predicción del rendimiento académico. *Revista Investigación Operacional*, Vol. 33, Núm. 3, pp. 252-267. <http://www.invoperacional.uh.cu/index.php/InvOp/article/view/415>
- Heredia, J., Rodríguez, A., & Vilalta, J. (2014). Predicción del rendimiento en una asignatura empleando la regresión logística ordinal. *Estudios Pedagógicos*, 40(1), 145–162. <https://doi.org/10.4067/s0718-07052014000100009>

- Hernández, R. & Mendoza, C. (2018). *Metodología de la investigación. Las rutas cuantitativa, cualitativa y mixta*. Ciudad de México, México: Editorial Mc Graw Hill Education, Año de edición: 2018, ISBN: 978-1-4562-6096-5, 714 p.
- Hosmer, D. W., & Lemeshow, S. (2000). *Regresión logística aplicada*. Segunda edición, John Wiley & Sons, Inc., United States. doi: 10.1002/0471722146
- Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression*. Wiley.
- Ibarra, M., & Michalus, J. C. (2010). Análisis del rendimiento académico mediante un modelo logit. *Revista Ingeniería Industrial*. Vol. 9 Núm. 2: 47-56. ISSN 0717-9103. ISSN Online 0718-8307. <https://revistas.ubiobio.cl/index.php/RI/article/view/56>
- Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*. Springer.
- Martínez, J., Ferrás, Y., Bermudez, L., Ortiz, Y., & Pérez, E. (2020). Regresión logística y predicción del bajo rendimiento académico de estudiantes en la carrera Medicina. *Revista Electrónica Dr. Zoilo E. Marinello Vidaurreta*, 45(4). <http://revzoilomarinello.sld.cu/index.php/zmv/article/view/2230.%3B>
- Navarro, E., Verbel, A., Robles, D., & Hurtado, K. (2014). Regresión Logística Ordinal Aplicada a la Identificación de Factores de Riesgo para Cáncer de Cuello Uterino. *Ingeniare*, 17, 87–105. <https://doi.org/10.18041/1909-2458/ingeniare.17.582>
- Pérez, J. M., & Romero, T. (2018). Análisis del rendimiento académico mediante regresión logística y múltiple. *Revista Electrónica De Conocimientos, Saberes Y Prácticas*, 1(2), 33–42. <https://doi.org/10.30698/recsp.v1i2.10>
- Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4), 427-437. <https://doi.org/10.1016/j.ipm.2009.03.002>
- Valera, J., Sinha, S., Varela, J. & Ponsot, E. (2009). Una explicación del rendimiento estudiantil universitario mediante modelos de regresión logística. *Revista Visión Gerencial*, 8 (2), pg:415-427. ISSN 1317-8822
- Villota, L. (2015). Regresión logística: una aplicación en la identificación de variables que inciden en el rendimiento académico, en el área de matemáticas. *Revista Educación Y Desarrollo Social*, 9(1)

Figuras y tablas

Figura 1.

Modelos logísticos polinómicos

$$\ln\left(\frac{P(Y = 1 | X)}{P(Y = 0 | X)}\right) = \beta_1 + \beta_{11}X_1 + \dots + \beta_{1K}X_K = \beta_1 + \sum_{i=1}^K \beta_{1i}X_i = Z_1 \quad (1.6)$$

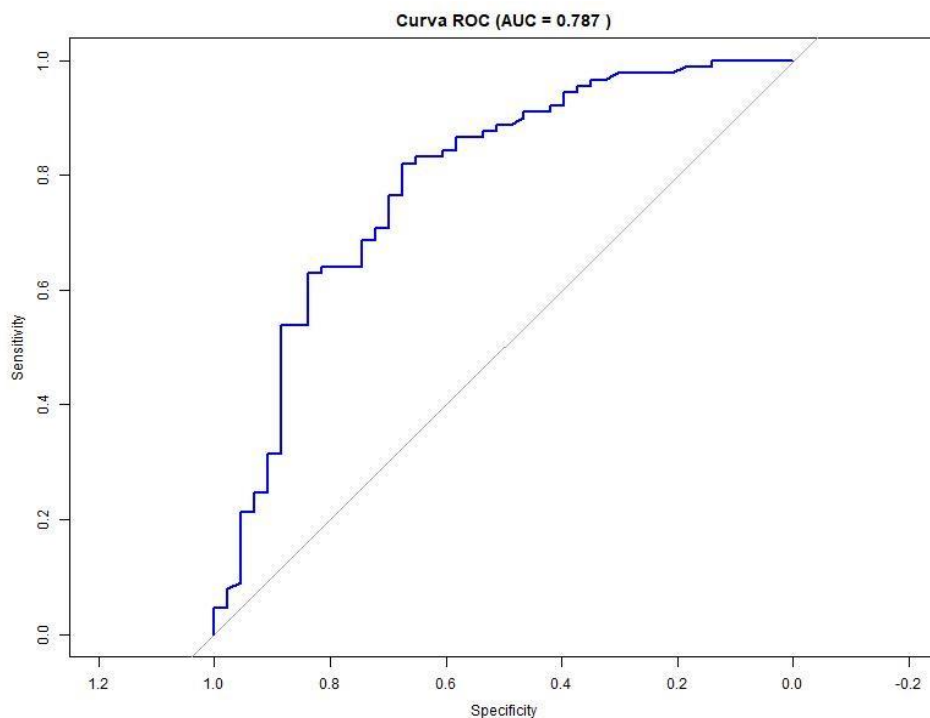
y

$$\ln\left(\frac{P(Y = 2 | X)}{P(Y = 0 | X)}\right) = \beta_2 + \beta_{21}X_1 + \dots + \beta_{2K}X_K = \beta_2 + \sum_{i=1}^K \beta_{2i}X_i = Z_2 \quad (1.7),$$

Nota. Adaptado de Navarro et al., (2014). Ordinal Logistic Regression.

Figura 2.

Curva ROC



Nota. La curva ROC (Receiver Operating Characteristic) evalúa la capacidad del modelo para distinguir entre dos categorías (por ejemplo, “positivo” vs. “negativo”). En el eje vertical se representa la sensibilidad (tasa de verdaderos positivos) y en el eje horizontal se representa la 1, especificidad (tasa de falsos positivos). Mientras más se aproxime la curva al vértice superior izquierdo (punto de sensibilidad = 1 y falsos positivos = 0), mejor será el poder discriminatorio del modelo.

Tabla 1.
Coeficientes y Estadísticos

Value	Std. Error	t value	p value
0.030	0.044	0.683	0.495
0.720	0.429	1.680	0.093
0.021	0.105	0.202	0.840
0.021	0.086	0.240	0.810
0.126	0.363	0.346	0.729
-0.816	0.794	-1.028	0.304
0.132	0.225	0.588	0.557
0.617	0.424	1.456	0.146
-0.034	0.203	-0.165	0.869
0.025	0.256	0.096	0.924
0.248	0.212	1.170	0.242
0.466	0.534	0.874	0.382
0.000	0.196	0.001	0.999
0.241	0.681	0.354	0.723
-0.627	0.645	-0.971	0.332
0.614	0.331	1.852	0.064
0.315	0.460	0.686	0.493
0.235	0.265	0.886	0.376
0.217	0.295	0.736	0.462
-0.500	0.306	-1.637	0.102
-0.628	0.460	-1.367	0.172
0.598	0.255	2.343	0.019
0.779	0.559	1.392	0.164
1.584	0.646	2.451	0.014
0.377	0.257	1.467	0.142
0.047	0.235	0.202	0.840
0.406	0.269	1.510	0.131
10.196	4.865	2.096	0.036
12.255	4.897	2.502	0.012
15.009	4.956	3.028	0.002

Nota. Modelo Ordinal y Función de Enlace. Se ha empleado un modelo de *odds* proporcionales (logit acumulativo) para explicar la probabilidad de que la respuesta a la pregunta 9 (con categorías ordenadas: malo, regular, bueno, muy bueno) se incremente a medida que varían las respuestas de las demás preguntas del cuestionario. Coeficientes: Cada coeficiente representa el cambio en el logit acumulativo (la razón logarítmica de las probabilidades) por cada unidad de incremento en la variable predictora correspondiente. Un coeficiente positivo indica que a medida que aumenta la respuesta en esa pregunta, aumenta la probabilidad de que la respuesta a la pregunta 9 sea en una categoría superior; un coeficiente negativo implica lo contrario. Significación Estadística: Se utiliza el valor p (*p value*) para determinar si el efecto estimado es significativamente diferente de cero. Generalmente, se considera significativo a un nivel del 5% ($p < 0.05$). En este conjunto, la mayoría de los coeficientes presentan valores p mayores a 0.05, lo que sugiere que, individualmente, esos predictores no tienen un efecto estadísticamente comprobado sobre la variable dependiente. No obstante, algunos coeficientes sí resultan significativos.

Tabla 2.
Intervalos de Confianza para los Coeficientes

Pgtas	2.5 %	97.5 %
1	-0.057	0.118
2	-0.116	1.569
3	-0.184	0.230
4	-0.148	0.189
5	-0.590	0.842
6	-2.389	0.737
8	-0.309	0.577
13	-0.210	1.458
14	-0.434	0.367
16	-0.481	0.527
17	-0.169	0.665
18	-0.576	1.522
19	-0.376	0.392
21	-1.104	1.586
22	-1.903	0.638
24	-0.033	1.270
25	-0.581	1.227
26	-0.285	0.758
27	-0.354	0.807
28	-1.105	0.097
29	-1.536	0.272
30	0.101	1.105
31	-0.317	1.885
32	0.334	2.881
33	-0.124	0.887
34	-0.412	0.514
35	-0.116	0.941

Nota. Si el intervalo incluye el cero: No se puede afirmar con un 95% de confianza que exista un efecto (positivo o negativo) de ese predictor sobre la probabilidad de obtener una categoría superior en la pregunta 9. Si el intervalo no incluye el cero: Se considera que el predictor tiene un efecto estadísticamente significativo; si el intervalo es enteramente positivo, el efecto es positivo, y si es enteramente negativo, el efecto es negativo.

Tabla 3.
Odds Ratios y sus Intervalos de Confianza

Variable	OR	CI_lower	CI_upper
Pgta1	1.031	0.945	1.125
Pgta2	2.054	0.890	4.800
Pgta3	1.021	0.832	1.258
Pgta4	1.021	0.863	1.208
Pgta5	1.134	0.554	2.322
Pgta6	0.442	0.092	2.090
Pgta8	1.142	0.734	1.780
Pgta13	1.854	0.811	4.296
Pgta14	0.967	0.648	1.443
Pgta16	1.025	0.618	1.695
Pgta17	1.281	0.845	1.944
Pgta18	1.594	0.562	4.583
Pgta19	1.000	0.686	1.480
Pgta21	1.273	0.331	4.883
Pgta22	0.534	0.149	1.893
Pgta24	1.847	0.967	3.561
Pgta25	1.371	0.559	3.412
Pgta26	1.265	0.752	2.134
Pgta27	1.242	0.702	2.241
Pgta28	0.606	0.331	1.101
Pgta29	0.533	0.215	1.312
Pgta30	1.818	1.107	3.019
Pgta31	2.179	0.729	6.584
Pgta32	4.874	1.397	17.835
Pgta33	1.458	0.884	2.428
Pgta34	1.049	0.663	1.671
Pgta35	1.501	0.890	2.562

Nota. *Odds Ratio* (OR): Indica cuánto se multiplican las probabilidades (*odds*) de obtener una categoría superior en la pregunta 9 por cada incremento unitario en el predictor. Si $OR > 1$, un aumento en el predictor se asocia con mayores probabilidades de una respuesta mejor. Si $OR < 1$, el aumento en el predictor se asocia con menores probabilidades. Intervalo de Confianza (IC) al 95%: Si el IC incluye el valor 1, no podemos afirmar con 95% de confianza que el efecto sea diferente de nulo.

Tabla 4.
Valores de VIF (Multicolinealidad)

Variable	VIF
Pgta1	3.767
Pgta2	1.553
Pgta3	2.161
Pgta4	1.513
Pgta5	1.465
Pgta6	3.099
Pgta8	1.508
Pgta13	1.470
Pgta14	1.551
Pgta16	1.918
Pgta17	1.751
Pgta18	1.674
Pgta19	1.648
Pgta21	1.835
Pgta22	2.330
Pgta24	1.482
Pgta25	1.544
Pgta26	1.478
Pgta27	1.436
Pgta28	1.723
Pgta29	1.706
Pgta30	1.549
Pgta31	2.108
Pgta32	1.464
Pgta33	1.501
Pgta34	1.660
Pgta35	1.416

Nota. La tabla de VIF (Variance Inflation Factor) evalúa la multicolinealidad entre los predictores. Un VIF de 1 indica que no existe correlación entre esa variable y el resto, mientras que valores mayores a 1 indican que la varianza del coeficiente se "infla" debido a la correlación con otros predictores. Generalmente, se consideran problemáticos los VIF superiores a 5 (o 10 en algunas aplicaciones), ya que sugieren una alta redundancia entre variables.

Tabla 5.
Top 10 Observaciones con Mayor Influencia (DFBETAS)

Observacion	Influencia
1	0.730
34	0.577
68	0.568
97	0.562
17	0.522
26	0.510
116	0.441
42	0.438
61	0.421
117	0.418

Nota. La tabla 5 de Top 10 Observaciones con Mayor Influencia (DFBETAS) nos indica cuáles registros del conjunto de datos tienen un impacto relativamente alto en la estimación de los coeficientes del modelo. En este contexto, el valor de "influencia" se calcula como la norma euclidiana de los DFBETAS, lo que nos permite cuantificar cuánto cambiarían los coeficientes si eliminamos esa observación.

Tabla 6.
Comparación de Modelos de Regresión Logística Ordinal

Modelo	Preguntas	AIC	BIC	Pseudo_R2
Completo	Pgta1, Pgta2, Pgta3, Pgta4, Pgta5, Pgta6, Pgta8, Pgta13, Pgta14, Pgta16, Pgta17, Pgta18, Pgta19, Pgta21, Pgta22, Pgta24, Pgta25, Pgta26, Pgta27, Pgta28, Pgta29, Pgta30, Pgta31, Pgta32, Pgta33, Pgta34, Pgta35	337.465	423.949	0.136
	Pgta30, Pgta32, Pgta2, Pgta16, Pgta24			0.071
Reducido	Pgta2, Pgta24, Pgta28, Pgta29, Pgta30, Pgta32, Pgta33	308.797	337.625	0.101