

UNIVERSIDAD TÉCNICA DE AMBATO



FACULTAD DE INGENIERÍA EN SISTEMAS, ELECTRÓNICA E INDUSTRIAL

MAESTRÍA EN MATEMÁTICA APLICADA

COHORTE 2021

Tema: “Modelo matemático para establecer los factores asociados a las calificaciones escolares utilizando regresión logística en los estudiantes del primer ciclo del Instituto Superior Tecnológico Tres de Marzo de la Provincia Bolívar”

Trabajo de Titulación, previo a la obtención del Título de Cuarto Nivel de Magister en
Matemática Aplicada

Modalidad del Trabajo de Titulación: Proyecto de Desarrollo

Autor: Licenciado, Cristóbal Martín Guambuguete Rea

Director: Ingeniero, Fabian Rodrigo Salazar Escobar, PhD.

Ambato – Ecuador

2023

A la Unidad Académica de Titulación de la Facultad de Ingeniería en Sistemas,
Electrónica e Industrial.

El Tribunal receptor del Trabajo de Titulación, presidido por: la Ingeniera Elsa Pilar Urrutia Urrutia, Magister e integrado por los señores: Ingeniera Clara Augusta Sánchez Benítez, Magister e Ingeniera Edith Elena Tubón Núñez, Magister designados por la Unidad Académica de Titulación de la Facultad de Ingeniería en Sistemas, Electrónica e Industrial de la Universidad Técnica de Ambato, para receptor el Trabajo de Titulación con el tema: “Modelo matemático para establecer los factores asociados a las calificaciones escolares utilizando regresión logística en los estudiantes del primer ciclo del Instituto Superior Tecnológico Tres de Marzo de la Provincia Bolívar” elaborado y presentado por el señor Licenciado Cristóbal Martín Guambuete Rea, para optar por el Título de cuarto nivel de Magíster en Matemática Aplicada; una vez escuchada la defensa oral del Trabajo de Titulación, el Tribunal aprueba y remite el trabajo para uso y custodia en las bibliotecas de la UTA.

Ing. Elsa Pilar Urrutia Urrutia, Mg.
Presidente y Miembro del Tribunal

Ing. Clara Augusta Sánchez Benítez, Mg.
Miembro del Tribunal

Ing. Edith Elena Tubón Núñez, Mg.
Miembro del Tribunal

AUTORÍA DEL TRABAJO DE TITULACIÓN

La responsabilidad de las opiniones, comentarios y críticas emitidas en el Trabajo de Titulación presentado con el tema: “Modelo matemático para establecer los factores asociados a las calificaciones escolares utilizando regresión logística en los estudiantes del primer ciclo del Instituto Superior Tecnológico tres de marzo de la provincia Bolívar”, le corresponde exclusivamente a: Licenciado Cristóbal Martin Guambuete Rea, Autor bajo la dirección del Ingeniero Fabian Rodrigo Salazar Escobar PhD, Director del Trabajo de Titulación; y el patrimonio intelectual a la Universidad Técnica de Ambato.

Licenciado, Cristóbal Martin Guambuete Rea

c.c.: 0201977253

AUTOR

Ingeniero, Fabian Rodrigo Salazar Escobar PhD.

c.c.: 1708019268

DIRECTOR

DERECHOS DE AUTOR

Autorizo a la Universidad Técnica de Ambato, para que el Trabajo de Titulación, sirva como un documento disponible para su lectura, consulta y procesos de investigación, según las normas de la Institución.

Cedo los Derechos de mi Trabajo de Titulación, con fines de difusión pública, además apruebo la reproducción de este, dentro de las regulaciones de la Universidad Técnica de Ambato.

Licenciado, Cristóbal Martin Guambuete Rea

c.c.: 0201977253

ÍNDICE GENERAL

Portada.....	i
A la Unidad Académica de Titulación	ii
AUTORÍA DEL TRABAJO DE TITULACIÓN	iii
DERECHOS DE AUTOR	iv
ÍNDICE GENERAL	v
ÍNDICE DE TABLAS.....	viii
ÍNDICE DE FIGURAS	ix
AGRADECIMIENTO	x
DEDICATORIA.....	xi
RESUMEN EJECUTIVO.....	xii
CAPÍTULO I.....	14
EL PROBLEMA DE INVESTIGACIÓN	14
1.1 Introducción.....	14
1.2 Justificación	17
1.3 Objetivos.....	18
1.3.1 General.....	18
1.3.2 Específicos.....	18
CAPÍTULO II.....	19
ANTECEDENTES INVESTIGATIVOS	19
2.1 Estado del Arte	19

2.2	Marco teórico.....	34
2.2.1	Desertores	34
2.2.2	Regresión Logística	34
2.2.3	Modelo Matemático.....	35
2.2.4	Modelo Logit	35
2.2.5	Variables de respuesta y predictoras	36
2.2.6	P valor.....	37
2.2.7	Correlación	38
2.2.8	Odd Ratio.....	39
CAPÍTULO III		41
MARCO METODOLÓGICO		41
3.1	Ubicación.....	41
3.2	Equipos y materiales.....	42
3.3.	Tipo de investigación.....	42
3.3.1	Investigación bibliográfica	42
3.3.2	Investigación seccional o transversal	42
3.3.3	Investigación cuantitativa	43
3.3.2	Investigación explicativa o causal	43
3.4.	Prueba de hipótesis	43
3.5.	Población o muestra.....	44
3.6.	Recolección de información	44
3.7.	Procesamiento de información y análisis estadístico	45
3.8.	Variables.....	47

CAPÍTULO IV	51
RESULTADOS Y DISCUSIÓN	51
4.1 Análisis descriptivo de las variables.....	51
4.2 Estadísticos descriptivos.....	56
4.3 Modelación matemática – Regresión Logística.....	57
CONCLUSIONES Y RECOMENDACIONES	66
5.1 Conclusiones.....	66
5.2 Recomendaciones	67
Bibliografía.....	69
ANEXOS	83
Anexo 1.....	83

ÍNDICE DE TABLAS

Tabla 1 Variables predictoras y de respuesta	37
Tabla 2 Interpretación OR	40
Tabla 3 Factores asociados a las calificaciones escolares	47
Tabla 4 Variables utilizadas por cada uno de los autores	48
Tabla 5 Variables con sus respectivos cofactores	50
Tabla 6 Análisis de los Odds Ratios	63

ÍNDICE DE FIGURAS

Figura 1 Factores determinantes asociados a la deserción	34
Figura 2 Coeficientes de correlación y diagramas de dispersión	38
Figura 3 Calificaciones de los estudiantes.....	51
Figura 4 Estado Civil.....	52
Figura 5 Nivel de formación del padre.....	53
Figura 6 Nivel de formación de la madre	53
Figura 7 Ingreso total del hogar.....	54
Figura 8 Cantidad de miembros del hogar.....	55
Figura 9 Estadísticos descriptivos del Dataset	56
Figura 10 Valores Odds Ratios de las variables intervinientes	61

AGRADECIMIENTO

En primer lugar, agradezco a mi Señor Jesucristo porque de él viene la ciencia y la sabiduría el cual me ha permitido por su gracia alcanzar este nuevo objetivo.

A mi familia que juntamente con Dios son un pilar fundamental en mi vida profesional ya que siempre me han apoyado en las buenas y en las malas a lo largo de mi vida académica.

A mis hijos, Deivis, Dilan, Gael, motivo de que día a día quiera superarme y ser de ejemplo para ellos ya que con su amor me anima a seguir adelante

A mi director, quien con su conocimiento y calidad de persona ha sido el eje principal para llevar a cabo esta investigación quien con paciencia me ha guiado a través de este camino.

Al Instituto Superior Tecnológico Tres de Marzo y las autoridades quienes me abrieron las puertas y confiaron esta ardua labor que ha sido posible gracias a la Universidad Técnica de Ambato que con el mejor contingente ha plasmado el conocimiento de la Maestría en Matemática Aplicada lo que me ha permitido subir un peldaño más en mi vida profesional.

Martin

DEDICATORIA

Este proyecto de investigación va dedicado a mis hijos, Deivis, Dilan, y Gael que hoy en día son tan sólo niños, pero un día sé que leerán esta dedicatoria, para ustedes mis niños hermosos espero que se sientan orgullosos de su papi, estoy seguro de que un día serán mejores que yo y esto es tan solo una muestra de que son lo más bonito que Dios me ha regalado me siento feliz por ustedes, con mucho amor su padre.

Martin

UNIVERSIDAD TÉCNICA DE AMBATO
FACULTAD DE INGENIERÍA EN SISTEMAS, ELECTRÓNICA E
INDUSTRIAL
MAESTRÍA EN MATEMÁTICA APLICADA
COHORTE 2021

TEMA:

MODELO MATEMÁTICO PARA ESTABLECER LOS FACTORES ASOCIADOS A LAS CALIFICACIONES ESCOLARES UTILIZANDO REGRESIÓN LOGÍSTICA EN LOS ESTUDIANTES DEL PRIMER CICLO DEL INSTITUTO SUPERIOR TECNOLÓGICO TRES DE MARZO DE LA PROVINCIA BOLÍVAR

MODALIDAD DE TITULACIÓN: Proyecto de Desarrollo

AUTOR: Licenciado, Cristóbal Martin Guambuguete Rea

DIRECTOR: Ingeniero, Fabian Rodrigo Salazar Escobar, PhD.

FECHA: Jueves, 19 de enero de 2023.

RESUMEN EJECUTIVO

El rendimiento académico representa uno de los elementos más importantes del ámbito educativo, ya que generalmente es considerado como un indicador de que el proceso de enseñanza y aprendizaje se está llevando de manera eficiente. Por consiguiente, el bajo rendimiento académico representa una preocupación a nivel mundial, debido a que esto a su vez genera desmotivación, ausentismo, deserción, pérdidas económicas al estado, entre otros. En Ecuador para las instituciones de educación superior el tema del bajo rendimiento académico constituye un grave problema. Una de las instituciones del Ecuador que presenta dicha problemática es el “Instituto Superior Tecnológico Tres de Marzo de la Provincia Bolívar”, la cual cuenta con las carreras tecnológicas como: Tecnología Superior en Electricidad, Tecnología Superior en Mecánica Automotriz, Tecnología Superior en Electrónica. Es importante resaltar, que en la presente investigación titulada: “Modelo matemático para establecer

los factores asociados a las calificaciones escolares utilizando regresión logística en los estudiantes del primer ciclo del Instituto Superior Tecnológico Tres de Marzo de la Provincia Bolívar”, tiene como objetivo implementar los factores socioculturales asociados a las calificaciones escolares utilizando regresión logística multivariante en alumnos del primer nivel del Instituto mencionado, mediante la aplicación metodológica descriptiva correlacional de carácter no experimental; para lo cual, se trabajó con la base de datos de 489 estudiantes del primer ciclo, pertenecientes a los años del 2019, 2020 y 2021, que fue desarrollada conforme a la guía de registro de matriculados de los Institutos públicos y privados proporcionados por la Senescyt; Se obtuvo como principales resultados, que los cofactores del nivel de formación de la madre, cantidad de miembros en el hogar mayor a nueve y estado civil unión libre, son los que incrementan la posibilidad de reprobar el ciclo académico. Para concluir, se determina que el modelo predice el 84.9% de los casos posibles para lo cual, considera las variables más significativas mediante el uso de los Odd Ratios como en el caso del cofactor estado civil unión libre, el cual aumenta la probabilidad de reprobar en un 200,12%.

Descriptor: Correlación, descriptores, educación, factores, logit, modelo matemático, odd ratio, p valor, probabilidad.

CAPÍTULO I

EL PROBLEMA DE INVESTIGACIÓN

1.1 Introducción

Actualmente, los centros educativos han experimentado grandes cambios en las últimas décadas, relacionados en los factores, referente de la heterogeneidad, multiculturalidad, diversidad de capacidades, diversidad de contextos familiares, así como otros elementos relacionados con la ciudadanía y las experiencias vitales, que marcan cognitiva, conductual y emocionalmente a los estudiantes.

Uno de los problemas a nivel mundial que ha incidido en la deserción escolar es la pandemia del covid-19, transformando la educación en un factor complejo de continuar ante el cierre de las instituciones educativas. Esto generó inestabilidad en el proceso de estudio, aplicando estrategias inmediatas como la educación en casa y la educación virtual (Pachay-López & Rodríguez-Gámez, 2021).

En este sentido, siguiendo las ideas de los citados autores la deserción escolar no debe ser catalogada como un problema menor o secundario; por el contrario, es un tema recurrente en la política educativa a nivel mundial. Asimismo, la concienciación de que la educación es esencial para el desarrollo vital de las personas realza su importancia en otras esferas de la sociedad.

La deserción escolar constituye además un problema grave en el sistema educativo de nivel superior, técnica, tecnológica o universitaria, ya que tiene una incidencia negativa sobre los procesos políticos, económicos, sociales y culturales del desarrollo nacional (Armas Sopalo, 2020). El ingreso a la universidad ha sido un problema recurrente para los estudiantes al graduarse del colegio, este proceso

se lleva a cabo con un instrumento desarrollado en base a las aptitudes y destrezas llamado Examen de Admisión para la Educación Superior, “EAES”, el mismo que los alumnos deben rendir de forma obligatoria para acceder a una educación superior, además de ser necesario para un buen desenvolvimiento. Según la Secretaría de Educación Superior, Ciencia, Tecnología e Innovación “Senescyt”, en su informe elaborado en septiembre del 2019 informó que el 51% de estudiantes que rinde la prueba de ingreso a la universidad ingresa a la carrera que desea, a partir del 2020, dicha prueba combina el resultado de la calificación obtenida en el colegio, con la obtenida en la prueba, de la siguiente manera: 40% Nota de grado y 60% Nota EAES (Senescyt, 2020).

En Ecuador para las instituciones de educación superior el tema de la deserción estudiantil constituye un grave problema, ya que el nivel de deserción constituye el indicador “Tasa de retención” basado en el modelo de evaluación institucional para los institutos superiores técnicos y tecnológicos para su proceso de acreditación, la cual fue establecido por el Consejo de Aseguramiento de la Calidad de la Educación Superior (CACES, 2020). De allí, se genera una gran importancia de establecer una alternativa que les permita a las instituciones de educación superior el poder identificar los casos de estudiantes próximos a retirarse e implementar acciones preventivas para evitarlo.

Una de las instituciones del Ecuador que presenta el problema de deserción es el Instituto Tres de Marzo, el cual se encuentra ubicado en el cantón de Chimbo de la provincia de Bolívar cuenta con las carreras tecnológicas: Tecnología Superior en Electricidad, Tecnología Superior en Mecánica Automotriz, Tecnología Superior en Electrónica, Tecnología Pecuaría, Tecnología en Procesamiento de Alimentos, Tecnología en Acuicultura; con una matrícula de estudiantes de 489. Es

importante resaltar, que en el presente proyecto se espera predecir los factores asociados a las calificaciones de estudiantes que cursan una carrera de nivel tecnológico superior pero que permita visualizar el proceso de disminución de la tasa de deserción, la cual es de vital importancia en las instituciones de educación superior, ya que representa un indicador de la calidad educativa evaluado por las instancias superiores, con el estudio de las variables socioeconómicas de los estudiantes que influyen en la probabilidad de que los estudiantes abandonen sus estudios en la institución

Con este fin, se propone la utilización de los modelos predictivos y técnicas de minería de datos para determinar patrones de comportamiento de los estudiantes, que establezcan su condición de potencial desertor, asociándole un índice de deserción como probabilidad de abandono del sistema educativo. En este orden de ideas, el objetivo del presente estudio es desarrollar un modelo matemático que permita conocer los factores asociados a las calificaciones por medio de regresión logarítmica en el primer nivel del Instituto Tres de Marzo.

1.2 Justificación

La presente tesis propone el desarrollo de un modelo matemático utilizando regresión logística multivariante para establecer los factores asociados a las calificaciones escolares en los estudiantes del primer ciclo del Instituto Superior Tecnológico Tres de Marzo de la Provincia Bolívar. Este modelo permitirá identificar variables académicas, físicas y actitudinales asociadas al fracaso académico y al bajo rendimiento, así como revisar las variables socioeconómicas que influyen en la probabilidad de que los estudiantes tengan un bajo rendimiento académico.

La importancia de esta investigación radica en que mejorará la toma de decisiones en la gestión académica, ya que el modelo predictivo proporcionará información valiosa sobre las variables que influyen en las calificaciones de los estudiantes. De esta manera, se podrán implementar cambios para optimizar la calidad de la enseñanza y el aprendizaje, y prevenir el fracaso académico. Además, el modelo predictivo también ayudará a identificar las áreas donde se necesitan más recursos para mejorar el rendimiento académico, optimizando los recursos disponibles y mejorando la eficiencia en su uso.

Finalmente, la implementación de un modelo predictivo matemático utilizando regresión logística multivariante también cumple con las exigencias de acreditación, al proporcionar una herramienta sólida y confiable para evaluar y hacer seguimiento al rendimiento académico de los estudiantes. En definitiva, el desarrollo del modelo matemático es de gran importancia para mejorar la calidad de la educación, formación de los estudiantes y la toma de dediciones asertivas del Instituto Superior Tecnológico Tres de Marzo de la Provincia Bolívar.

1.3 Objetivos

1.3.1 General

Implementar un modelo predictivo matemático utilizando regresión logística multivariante, considerando las variables que influyen en calificaciones de alumnos del primer nivel del Instituto Superior Tecnológico Tres de Marzo.

1.3.2 Específicos

- Identificar las variables académicas, físicas y actitudinales asociadas al fracaso académico y al bajo rendimiento en trabajos elaborados por la comunidad científica.
- Revisar las variables socioeconómicas de los estudiantes que influyen en la probabilidad de que los estudiantes tengan un bajo rendimiento académico en sus estudios en la institución, mediante una regresión logística multivariante.
- Diseñar un modelo predictivo matemático utilizando regresión logística multivariante, considerando las variables que influyen en calificaciones escolares.

CAPÍTULO II

ANTECEDENTES INVESTIGATIVOS

2.1 Estado del Arte

A nivel mundial existen investigaciones similares referente a la deserción estudiantil, que demuestran que es posible determinar diferentes técnicas de minería de datos, como la regresión logística para identificar a estudiantes con más probabilidad a la deserción escolar.

La deserción es una preocupación creciente en las instituciones de educación superior debido a sus altos índices, pero también a su impacto socioeconómico. (Cerpa, 2015). Este fenómeno se ha determinado como tema de estudio en los últimos años en medio de un panorama educativo que muestra un importante aumento en la cobertura de la educación superior, esfuerzo que obtendrá el impacto esperado a medida que los estudiantes logran sus objetivos de conllevar con éxito sus estudios. En una investigación aplicada a estudiantes de la carrera de Ingeniería Civil en la Universidad Católica de la Santísima Concepción, Chile determinó, a través de la regresión logística las variables que explican mejor la deserción son: los ingresos familiares y el rendimiento académico de los alumnos; sin embargo, se plantea la necesidad de representar la relación entre ingresos y deserción a la luz de los resultados de este estudio (Villa & Omar, 2010)

Concerniente a Latinoamérica, la deserción está poco estudiada y no tiene una base estadística a nivel nacional (Díaz C. J., 2009) sin embargo, el Ministerio de Educación de Chile ha expresado su preocupación por el impacto del fenómeno de la deserción temprana; en vista a este problema Christian J. Díaz realizó un análisis

de la deserción de estudiantes de ingeniería de la cohorte del 2004 al 2007 en Chile, tomando en consideración a 267 de ellos además, analizó los factores de deserción más comunes para ulterior a la investigación sugerir medidas para reducir las tasas de deserción universitaria. El autor aplicó el modelo de Kaplan-Meier y el modelo de riesgos proporcionales de Cox para determinar los factores que afectan en la deserción. Como resultado de este estudio muestra que los estudiantes tienen mayor probabilidad de desertar en el primer semestre y esa probabilidad va bajando conforme van avanzando de nivel; además, los estudiantes de escuela primaria fiscal constituyen la mayoría de todas las inscripciones universitarias y tienen las tasas más altas de abandono escolar en todo el sistema universitario.

En la investigación de los autores (Tapasco Alzate, Ruiz Ortega, Osorio García, & Ramírez Ramírez, Deserción estudiantil: incidencia de factores institucionales relacionados con los procesos de admisión. , 2019), indican que mediante un modelo de regresión logística multivariados y de técnica de optimización aplicado a un proceso de admisión permiten arrojar que algunas variables tienen incidencias estadísticamente significativas en el riesgo de la deserción escolar.

Además, en otra investigación relacionada con la deserción estudiantil realizada en la Universidad Nacional de Colombia, donde la problemática ha sido abordada desde un enfoque metodológico cuantitativo y cualitativo para dimensionar cuantitativamente el alcance de la deserción e identificar grupos que podrían considerarse en situación de riesgo. Con respecto a la continuidad de sus estudios, se han aplicado metodologías estadísticas descriptivas y predictivas donde se ha reconocido que la variable del rendimiento en el primer año, es la principal predictora del desempeño futuro, ligado con el estudio de los padres y la situación socio económica del estudiante (Goldenhersch, Coria, & Saino, 2011).

La deserción estudiantil es un proceso que se presenta al momento que el estudiante abandona de forma voluntaria o forzadamente una carrera de educación tecnológica, en la investigación realizada por (Martínez B. H., 2015) determina algunas variables que afectan a la deserción como: aspectos emocionales, nivel socio- económico, bajo rendimiento, del cual deben existir los condicionantes que les impedirán dejar los estudios.

Otra forma de analizar la deserción escolar (Vergaray, 2021), es mediante la aplicación de un modelo predictivo en las universidades, para la detección temprana de estudiantes con alto riesgo de deserción académica, es viable y muy beneficioso, para que las universidades a través de sus gestores académicos puedan aplicar estrategias más focalizadas para reducir sus índices de deserción académica.

Citando a (Cuji, Gavilanes, & Sanchez, 2017), este artículo describe el desarrollo de un modelo de predicción de la deserción de los estudiantes, para predecir la probabilidad de que un estudiante abandone los estudios utilizando una técnica de clasificación basada en árboles de decisión, para lo cual se tomó como muestra a 25 estudiantes. La metodología se basó en Knowledge Discovery in Database (KDD) y el algoritmo CART de R estudio; además, construyó un árbol de decisión de 4 reglas con una profundidad de 4 niveles, obteniendo de la investigación que las variables notas y el curso o nivel al que pertenecen dichos estudiantes tuvieron un mayor impacto en la deserción.

Finalmente, contextualizando a Elkin Castaño, Santiago Gallón, Karoll Gómez y Johanna Vásquez (2008), en su artículo científico de investigación “Análisis de los factores asociados a la deserción estudiantil en la Educación Superior: un estudio de caso”, realizan el análisis de los determinantes de deserción usando modelos de

riesgos en tiempo discreto. Se aplicó a una muestra de 348 estudiantes, la metodología basada en el modelo Prentice y Gloeckler permite dar seguimiento a los estudiantes para estimar una posible deserción a través del tiempo, desde su inicio hasta el final. Los resultados obtenidos determinan que el número de créditos aprobados es inversamente proporcional a la deserción; a partir del modelo se puede desarrollar políticas encaminadas a la permanencia de los estudiantes en el sistema educativo.

1. Según Pérez y otros (2017) en el trabajo de investigación titulado “Deserción en estudiantes de la carrera de Enfermería matriculados en el período 2010-2015”, de la Universidad Técnica de Manabí en el cual el objetivo fue analizar las causas de repitencia y deserción de los estudiantes aplicando la metodología de la investigación descriptiva, de eje transversal y retrospectivo. Se elaboró un cuestionario de preguntas relacionadas a la repitencia y deserción de los estudiantes universitarios la misma que fue aplicada a una muestra de 60 estudiantes de una población total de 146 pertenecientes a la carrera de enfermería del año lectivo mencionado, dando como resultado que el sexo femenino pertenecientes a zonas rurales son los que más desertan y las principales causas de este hecho es la ausencia a clases, docentes poco explícitos al momento de dar clases, no seguir los sílabos para impartir los contenidos, poca tutoría a los estudiantes, exámenes con preguntas fuera del contenido impartido y excesivos trabajos autónomos. Finalmente, se concluyó que las principales causas para la no graduación de esta corte de enfermería son las anteriormente mencionadas y afecta a las mujeres el sector rural.
2. Para Castrillón (2020), en su artículo científico con el tema: “Predicción de las principales variables que conllevan al abandono estudiantil por medio de

técnicas de minería de datos”, realiza el estudio de las principales variables que provocan el abandono estudiantil haciendo uso del algoritmo de clasificación bayesiana en la metodología. Además, diseñó una encuesta para la recolección de información tomando como referencia estudios previos dividida en 5 categorías como las: académicas, personales, familiares, sociales y económicas; la misma que fue aplicada a una población de 410 alumnos y mediante un análisis estadístico multivariado con algoritmo J48 haciendo uso de la plataforma Weka dando como resultado, con una efectividad mayor al 80% las variables que influyen al abandono estudiantil son: expectativas insatisfechas, pedagogía de los profesores, motivación del programa, sentimiento de frustración, importancia del programa, y procrastinación. La investigación concluye que la información es útil para llevar a cabo estrategias que mitiguen este problema de abandono estudiantil en las universidades.

3. Conforme a lo expuesto por Vinueza (2021) en su investigación “Diseño de un modelo matemático para estimar la deserción estudiantil mediante técnicas de análisis multivariado en una institución de educación superior tecnológica”, realizó un análisis descriptivo de las variables de investigación de esta forma, para alcanzar el objetivo de diseñar el modelo matemático se usó la metodología KDD para el descubrimiento del conocimiento, los resultados más importantes muestran que el 60,18% de los estudiantes son mujeres, el 89,75% son solteros, el 32,50% trabajan y el 1,8% son divorciados. El promedio de la renta económica es de \$560.21, la repetición y profesión, a la que pertenece el alumno, fue la única importante. Se encontró que la variable significativa repetición muestra que los estudiantes asocian el fracaso escolar con el deseo de desertar. Según la variable probabilidad de repetición, la probabilidad de

abandonar la universidad aumenta en un 7 % cuando un estudiante repite asignatura. Se determinó que la variable significativa “ocupación” muestra que los estudiantes de la carrera gastronomía desertan con más frecuencia, para la profesión ganadera los estudiantes tienen 58% y 68% menos deserción. Se implementó un modelo de predicción para evaluar la deserción y se observó que el modelo de regresión logística clasificó correctamente el 83% de los datos de entrenamiento y el 79% de los datos de prueba.

4. Para Darío Alejandro Viteri Castro (2011) en su tesis de titulación con el tema “Estudio sobre la deserción estudiantil en la Pontificia Universidad Católica del Ecuador - Matriz, en los niveles 1ro, 2do y 3ero de todas las Facultades y Escuelas del primer semestre del año académico”, realizó la investigación para determinar las causas de abandono estudiantil en la universidad para obtener el índice de deserción; la población que se utilizó fueron los estudiantes de niveles diferentes como: primero, segundo y tercero de las facultades de la PUCE, mediante la diferencia de estudiantes matriculados y desertados se obtuvo el índice de deserción. Con el 0.05 de alfa y 95% de confiabilidad, se obtuvo un índice del 8.02% que corresponde a estudiantes desertados, finalmente se concluyó que de cada 10 estudiantes desertados 7 son por causas económicas, personales o académicas.
5. Karla Rafaela Palacios Alvear (2021) en su trabajo de titulación con el tema “Comparativa de modelos de clasificación para inferir la probabilidad de deserción estudiantil en la Facultad de Ciencias Químicas de la Universidad de Cuenca”, realizó la siguiente investigación para determinar las siguientes variables que influyen en el abandono estudiantil en la universidad; en el área

que se trabajó fue en los estudiantes para determinar la probabilidad y el motivo por la que los estudiantes optan por la deserción estudiantil, mediante el uso de herramientas de software, dando los siguientes resultados: de cada 138 matriculados, continúan 118 y 20 estudiantes abandonan sus estudios en la primera etapa de proceso de curso, en conclusión que el modelo matemático se aleja en un 20% de la realidad debido a que existen menor cantidad de estudiantes desertores de los que se predice.

6. Lo determinado por Dolores Jano Salagre (2006) en su investigación titulado como “Determinación de los factores que afectan al rendimiento académico en la educación superior”, pretende determinar los factores que influyen en el rendimiento estudiantil, para llevar a cabo usó el ECTS que es el esfuerzo de un estudiante para preparar una determinada materia como uno de los factores. Los resultados se obtuvieron a través de la aplicación de una encuesta semanal haciendo uso de la estadística descriptiva y mediante técnicas multivariantes se explicó que las habilidades y los conocimientos previos sobre un estudiante tienen un impacto significativo en el rendimiento académico mas no así las variables sociodemográficas como el sexo, edad o situación laboral.
7. Para Jimmy Reyes y otros (2017) en el trabajo de investigación “Una aplicación del modelo de regresión logística en la predicción del rendimiento estudiantil”, en el programa de afiliados, estudiantes de primer año Universidad de Antofagasta; el análisis contempla tres criterios rigurosos para clasificar a los estudiantes como exitosos en el primer semestre de sus carreras. Se aplica un modelo de regresión logística, se comparan los resultados con los obtenidos del análisis discriminante y se analizan los índices de coincidencia y predictibilidad de estos. Los resultados también muestran que la varianza

porcentual del rendimiento explicada por el modelo es igual a 22.86% en este caso, con base en las variables.

8. Según Alberto Vélez y Natalí González (2005) en su trabajo con el tema “Factores asociados al rendimiento académico en estudiantes de medicina”, se realizó una evaluación de factores para determinar qué es lo que está perjudicando en el desarrollo de los estudiantes en medicina del primer semestre, la búsqueda se realizó en la población estudiantil para obtener los factores que posteriormente serían analizados para determinar la predicción y la asociación de variables mediante el uso del modelo de regresión logística. Tomando como referencia una muestra de 80 estudiantes, como resultados se dio a conocer que la mayoría de las mujeres de edad promedio entre los 17 y 18 años proceden de Bogotá de colegios monolingües, mixtos o privados y tienen rasgos de consumo de cigarrillo y alcohol, violencia intrafamiliar, pero no hay rasgos de drogas psicoactivas, por lo tanto, se concluye que a pesar de que existen factores que impiden al ingreso académico se va a evaluar el rendimiento para evitar el fracaso.
9. Emma V. Barreno Vereau (2012) en su estudio “Análisis comparativo de modelos de clasificación en el estudio de la deserción Universitaria”, en la Universidad de Lima proporciona un método para comparar modelos de regresión logística y árboles de clasificación usando una variable explicativa específica como punto de referencia para determinar si es probable que un estudiante en específico abandone los estudios. Para aplicar los modelos de clasificación se utilizó el software comercial Minitab 16 y el software libre para obtener modelos de regresión logística y árboles de clasificación. Entre las principales conclusiones se puede señalar que los dos modelos de clasificación

obtenidos (regresión logística y árboles de clasificación brindaron similares resultados, en ambos casos con un porcentaje de clasificación exitosa de por lo menos 94% en el análisis de la deserción universitaria, considerando determinadas variables explicativas

10. Según Elizabeth Guadalupe Chong González (2017) en su investigación con el tema “Factores que inciden en el rendimiento académico de los estudiantes de la Universidad Politécnica del Valle de Toluca”, con todos los alumnos universitarios examinaron puntos como: el contexto familiar y universitario, el apoyo que sus familias les brinda en todo el proceso académico; mediante investigaciones hemerográficas y bibliográficas se obtuvo mucha información para la elaboración de estadísticos de desempeño de los estudiantes. Los resultados obtenidos se analizaron en varias secciones como: rendimiento escolar, contexto familiar y demográficos dando los siguientes promedios: de los 348 entrevistados el 79% tienen una familia completa, el 15% solo tienen a la madre o el padre, el 5% son independientes y el 1% proviene de una familia extensa, en esta investigación se concluyó que los estudiantes consideran que la preparación de un profesor mejora su rendimiento académico y también las ganas de superar de cada uno de ellos para evitar que muchos alumnos abandonen los estudios.

11. Para Arismendi y Fuentes (2018) en su trabajo “Modelo de regresión logística como alternativa para medir la probabilidad de deserción temprana en la universidad de los llanos periodo 2015-2 – 2008-1”, comprobó que obtener puntajes altos en la prueba Saber 11, ser mujer, no haber completado la escuela secundaria, tener educación y tener padres viviendo juntos reducen la probabilidad de abandonar la escuela, según el modelo estimado. Del mismo

modo, graduarse en una escuela privada aumenta la probabilidad de abandonar los estudios, sin embargo, los valores de los predictores globales son relativamente bajos, lo que sugiere que se pueden incluir otras variables para explicar la probabilidad de salida anticipada. Para la metodología se usó software de análisis estadístico econométrico tales como Gretel y Minitab 18, que con un p valor menor al 0,05 incluye a las variables que son significativas al modelo y para ratificar lo mencionado se aplica el estadístico de Wald (W) donde la significancia se afina a un p valor menor al 0,005. El estudio confirmó que la Universidad de los Llanos tiene una alta tasa de deserción y, lo que es más importante se ha podido identificar los cursos con las tasas de deserción más alta. Dados estos resultados se espera que se fortalezcan los programas y enfoques de prevención de deserción para reducir estos números.

12. Para Loaiza y Rojas (2015) en su obra “Rendimiento Académico en Matemáticas un estudio vía Regresión Logística”, modelaron el rendimiento académico mediante regresión logística haciendo uso de variables sociales y personales del nivel secundario del instituto; la metodología usada fue a modo de monografía de forma cronológica en diferentes etapas como: revisión literaria, diseño metodológico, recolección de información, análisis, validación de datos y determinación del modelo. Los resultados obtenidos muestran que el 75,5% tienen un espacio destinado al estudio, un 50% menciona que tienen ayuda para la solución de trabajos matemáticos con 2,2 horas promedio para hacer tareas, el 61,7% estudia además de trabajar, y el promedio de la evaluación matemática es 6,9 además, se obtuvo una correlación de las variables predictoras con la variable dependiente con un p valor menor al 0,05. Se concluyó que el modelo logístico explica solo los factores personales y los

factores como ambiente escolar, los aspectos orgánicos solo aportan de forma negativa a la aprobación de la materia de matemáticas.

13. Lo planteado por Prieto (2015) en el trabajo desarrollado “Uso de Regresión Logística para predecir deserción estudiantil temprana”, usó la minería de datos y machine learning para determinar ciertos patrones en los factores académicos que inciden en la deserción; como parte de la metodología se usó la matriz de confusión para la clasificación binaria con el fin de evaluar a los modelos que fueron presentados en la investigación donde los verdaderos positivos representa a los desertores y los verdaderos negativos a los no desertores. Los resultados fueron obtenidos mediante la aplicación del método hold out donde el Recall debe ser mayor al 0,6 y tener alta precisión en este caso 0,82 lo que representa falsos positivos mínimos. Se concluyó que a mayores créditos (+12) y puntaje promedio de 4-5 la probabilidad de desertar disminuye y aumenta cuando los promedios y créditos mencionados son bajos.

14. El trabajo de Pérez, Escobar, Toledo, Gutiérrez y Reyes (2018), con título “Modelo de predicción de la deserción estudiantil de primer año en la Universidad Bernardo O’ Higgins” tuvo por objetivo diseñar un modelo predictivo de retención de los estudiantes de primer año de la Universidad Bernardo O’ Higgins - UBO (Santiago de Chile), identificando que las variables de entrada ya sean académicas, sociales o familiares, se revelan significativas para el análisis. Posteriormente el análisis bivariado permitió determinar diecisiete variables, significativamente asociadas con la deserción estudiantil; logrando a su vez identificar las relaciones de dependencia con el abandono de estudios. El modelo multivariado predijo en un 86,4 % la conducta de abandono

señalando siete variables categóricas independientes que, al final se develan como factores relevantes del modelo predicción.

15. Según Tuñón, I. y Poy, S. (2016) en el trabajo de investigación “Factores asociados a las calificaciones escolares como proxy del rendimiento educativo”, en el Estado argentino se ha llevado adelante diferentes políticas de inclusión educativa, cabe preguntarse sobre los logros educativos. ¿Cuáles son los factores sociodemográficos, socioeconómicos y socioculturales asociados a las calificaciones escolares como proxy del rendimiento educativo en asignaturas como Matemática y Lengua en la educación primaria y secundaria? El estudio permitió reconocer desigualdades de género y de trayecto educativo, Para ello se realizaron análisis a través de modelos de regresión logística con base en los microdatos de la Encuesta de la Deuda Social Argentina para el año 2011.

16. Tapasco, Ruiz, Osorio y Ramírez (2019) en su trabajo “Deserción estudiantil: incidencia de factores institucionales relacionados con los procesos de admisión”, estudiaron la incidencia que distintas variables institucionales del proceso de admisión tienen sobre el riesgo de deserción estudiantil. La población del estudio fueron los estudiantes matriculados en la Universidad de Caldas para las cohortes 2010-2012. Se utilizaron modelos logísticos multivariados y técnicas de optimización. Los modelos obtenidos arrojan que algunas variables del proceso de admisión, susceptibles de ser intervenidas institucionalmente, como el ‘puntaje de admisión’, ‘ingreso por segunda opción’, ‘programas de cupos especiales’ y ‘costo de matrícula’ tienen una incidencia estadísticamente significativa en el riesgo de deserción, influencia que varía entre facultades. En conclusión, para atenuar la deserción estudiantil

e incrementar el poder predictivo del desempeño futuro del estudiante, se pueden contemplar acciones como el repensar el ingreso por segunda opción, el sostener las políticas de ingreso por méritos académicos y de acción afirmativa y el modificar las ponderaciones actualmente asignadas a las áreas evaluadas en el proceso de admisión.

17. Según Camacho, Montalvo y Galezo (2019) en su investigación “Determinantes de la deserción estudiantil en estudiantes universitarios”, la deserción estudiantil es un problema significativo para el sistema educativo. Su trabajo se enfocó en un estudio descriptivo, correlacional que caracteriza a los 16.927 estudiantes de pregrado matriculados en la Universidad de Cartagena en diversas cohortes. Se establecieron los posibles determinantes en el riesgo de deserción y patrones temporales de deserción. Los datos se consultaron en la plataforma de SPADIES. La muestra de 1547 estudiantes desertores de la Universidad de Cartagena fue proporcional a cada programa y se utilizó un instrumento de 22 preguntas, elaborado por los investigadores. Se encontró que en los últimos años la tasa de deserción en la institución había disminuido y estaba en correspondencia con el comportamiento de la tasa promedio de deserción a nivel nacional. Según el estudio, los hombres y los estudiantes de los programas de la modalidad a distancia presentaban más riesgo de desertar. Además, quienes ingresan con una clasificación baja en las pruebas Saber, 11 presentaron un mayor riesgo de deserción en los primeros períodos de estudio.
18. En el trabajo Tatiana Fernández-Martín y otros (2019) denominado “Un análisis multinomial y predictivo de los factores asociados a la deserción universitaria”, el fenómeno de la deserción por su complejidad e impacto educativo y social ha sido ampliamente estudiado para entender las causas

específicas. Dentro de esta línea de investigación, el objetivo de este estudio fue desarrollar modelos explicativos y predictivos de la deserción estudiantil en los programas académicos de grado del Instituto Tecnológico de Costa Rica, con base en diferentes variables registradas en el Sistema de Indicadores de Gestión Institucional (SIGI). Entre los hallazgos se destaca que la probabilidad de deserción se relaciona con variables sociodemográficas, programa de estudio, beneficios obtenidos al ingresar, historial académico y rendimiento en el primer semestre de estudio. Estos resultados son el primer paso hacia la construcción de un modelo predictivo más robusto de la deserción estudiantil, que se espera contribuya a la toma de decisiones preventivas en esta universidad.

19. Jeylin Meybelin (2018) en su trabajo titulado “Análisis del rendimiento académico mediante regresión logística y múltiple”, usó regresión logística y múltiple para determinar el comportamiento y la relación entre las materias de matemáticas y los puntajes de lengua extranjera y literatura para analizar el rendimiento académico de los estudiantes de la Academia Nacional Santa Teresita en La Batía, Muelle de los Bueyes en el área de las matemáticas. Este es un estudio de método mixto que utiliza métodos estadísticos y entrevistas públicas con docentes para comprender mejor lo que significa analizar el desempeño escolar en este contexto, dando como resultado que los modelos de regresión múltiple estiman las puntuaciones en matemáticas cuando se controlan por lengua y literatura, así como por lenguas extranjeras.
20. Para Valera Jorge y otros (2009) en el artículo científico publicado como “Una explicación del rendimiento estudiantil universitario mediante modelos de regresión logística”, los resultados mostraron que los estudiantes con un

promedio de calificaciones bueno o aceptable en la escuela secundaria tenían más probabilidades de tener un mejor desempeño académico en el primer semestre que los estudiantes con un promedio de calificaciones bajo; en cambio, con respecto al rendimiento académico en el segundo semestre, los datos muestran que la mejor oportunidad en este semestre es para un estudiante con un buen promedio de calificaciones en el primer semestre.

2.2 Marco teórico

2.2.1 Desertores

En relación con la educación, los desertores son aquellos estudiantes que no han podido concluir sus estudios. Existe una gran cantidad de investigaciones en relación con este tema donde el factor común de la ocurrencia de este va desde factores personales hasta institucionales. Se puede clasificar en 4 tipos, como se puede observar en la figura 1 (Vries, Leon, Romero, & Hernandez, 2011).

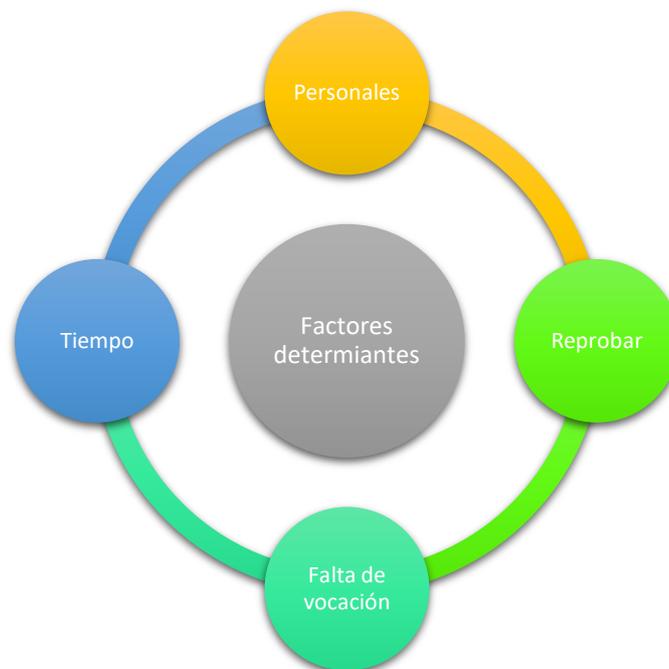


Figura 1
Fuente: El Autor (2023)
Figura que muestra los Factores determinantes asociados a la deserción

2.2.2 Regresión Logística

Es el método adecuado para modelizar variables de respuesta binaria, es decir 1 o 0, si o no, bueno o malo, pierde o aprueba, deserta o no, etc. La regresión logística permite usar variables categóricas y también continuas, de esta manera permite dar una interpretación biológica a cada uno de sus parámetros. Está

compuesta por una o más variables de entrada X_1, X_2, \dots, X_n y una variable dicotómica de respuesta generalmente representada como Y , o salida de la función. Además, tiene la siguiente probabilidad: $P(Y=1/X)$ para el cual X es vector de P covariables (Field, Miles, & Field, 2012).

2.2.3 Modelo Matemático

Son desarrollos que hacen uso de las fórmulas matemáticas para relacionar distintas variables, restricciones o parámetros. El modelo matemático simplifica fórmulas matemáticas, ecuaciones o funciones para estudiar un fenómeno o relacionar dos o más variables entre sí. Dentro del área de las matemáticas la rama que estudia dichos modelos es conocida como “teoría de los modelos”. Además, dichos modelos son utilizados para comprender fenómenos físicos, naturales o sociales; para hacer hipótesis y predecir valores de variables en el tiempo. (Roldán, 2022).

2.2.4 Modelo Logit

El modelo Logit forma parte de los Modelos Lineales Generalizados, este modelo es un caso particular donde la distribución es binomial y la función de enlace hace referencia al logaritmo de razón de probabilidad; para lo cual se hace uso del modelo lineal con forma de la siguiente ecuación (Paladino, 2022).

$$y = \beta_0 + \beta_1 x + \dots + \beta_i x$$

Donde:

y : variable dependiente.

x : variable independiente.

β_0, β_1 : parametros poblacionales.

β_0 : término constante – término de intercepto.

β_1 : *parametro pendiente*.

Considerando la ecuación anterior cabe recalcar que la distribución de probabilidad no es normal, por lo que no se podría utilizar la fórmula directa; por lo tanto, es necesario ajustar a otra distribución como la dicotómica que nos da una binomial como se observa a continuación.

$$\ln\left(\frac{p_1}{1-p_1}\right) = \beta_0 + \beta_1x + \dots + \beta_ix$$

ln : logaritmo natural

p_1 : probabilidad de evento de interes

β_0, β_1 : parametros poblacionales.

En este caso no es necesario conocer el valor de “y” como una combinación lineal de los diferentes valores de “x” sino que buscamos la probabilidad del logaritmo de las diferentes razones que de 1.

2.2.5 Variables de respuesta y predictoras

La variable dependiente también conocida como variable de respuesta, estas variables son de interés dentro de un experimento y son aquellas que se pueden medir u observar. Por otra parte, las variables del experimento que perturban la respuesta medibles o establecidos por el investigador se conocen como variables independientes, predictoras, o explicativas (Minitab, 2022).

Ejemplo, si se desea conocer si los ingresos, el estado civil, el género, etc., afectan a las calificaciones escolares de los estudiantes en una determinada institución o a su vez conocer la frecuencia de riego necesario para el óptimo crecimiento de un cultivo de flores, debe quedar claro cuáles son las variables predictoras y de respuesta, como se muestra a continuación.

Asunto	Variables predictoras	Variable de respuesta
Calificaciones escolares	Género, estado civil, discapacidad, residencia, ocupación, ingresos.	Rendimiento
Crecimiento de cultivo de flores	Cantidad de luz, pH del suelo, frecuencia de riego	Calidad de la flor

Tabla 1

Fuente: Autor (2023)

Tabla que ejemplifica a las variables predictoras y de respuesta

A las variables predictoras continuas también se las conoce como covariables y a las variables predictoras categóricas se las conoce como factor.

En el caso de estudio de las calificaciones escolares, una covariable es: el estado civil y un factor en este caso vendría a ser los diferentes estados como soltero/a, casado/a, divorciado/a viudo/a o unión libre. En general, está en manos del investigador considerar variables predictoras “x” y variables de respuesta “y” acorde a su investigación.

2.2.6 P valor

Conocido también como p-value en inglés, hace referencia al valor probabilístico (valor p); este término hace énfasis al nivel de significancia no arbitrario mínimo (comúnmente 0.05) con el que se puede descartar la hipótesis nula “H₀” dado un estadístico de contraste y la función de distribución. Es decir, es el mínimo probabilístico dada por la distribución con el que se puede rechazar H₀, sin la necesidad de precisar el nivel de significancia del contraste. Para la regla de

rechazo: Se rechaza H_0 si, el p valor es menor que el nivel de significancia caso contrario; si el p valor es mayor que el nivel de significancia no se rechaza H_0 (Rodó, 2022).

2.2.7 Correlación

La asociación dos variables numéricas normalmente una de entrada con otra de salida es denominada como correlación, esta evalúa la tendencia ya sea creciente o a su vez decreciente de los datos a analizar. Cuando de entre dos variables una proporciona información de la otra se dice que están asociadas o correlacionadas; por otro lado, si al aumentar o disminuir una determinada variable no proporciona información acerca de la otra variable, se dice que no existe asociación o no están correlacionadas. Para determinar si las variables están correlacionadas o no, se sugiere realizar el diagrama de dispersión de estas para analizar y evitar ciertos problemas, que a simple inspección sería complicado conocerlas (Máxima Formación, 2022). Además, permite medir la magnitud y el signo de tendencia de las dos variables como se muestra en la figura 2:

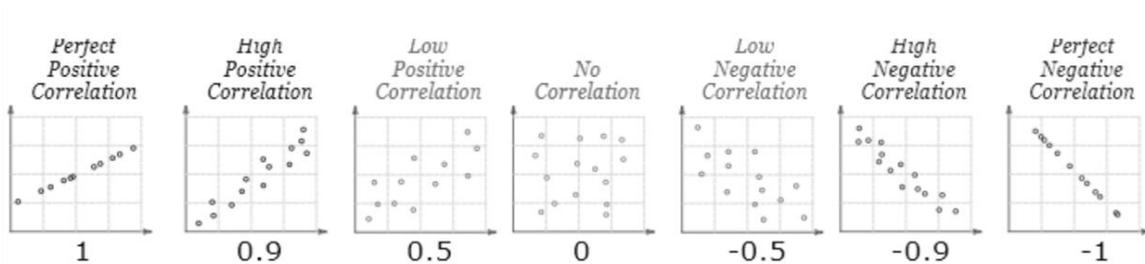


Figura 2
Fuente: Autor (2023)
Figura que muestra los Coeficientes de correlación y diagramas de dispersión

Estas correlaciones son medibles según las variables propuestas, por otra parte, en su mayoría tienden a formar una línea recta dentro del diagrama de dispersión. Se presenta, la correlación positiva perfecta, correlación positiva alta, correlación positiva baja, no correlación, baja correlación negativa, alta correlación

negativa, correlación negativa perfecta. Las mismas que son medibles mediante el coeficiente de correlación producto-momento en Pearson aplicables en variables numéricas, coeficiente de correlación de rango en Spearman en variables ordinales. (Universidad Tecnológica de Pereira Facultad de Ciencias de la Salud, 2017)

La dirección permite conocer el signo de la relación de la siguiente manera: signo positivo muestra una relación positiva o directa, signo negativo muestra una relación negativa, indirecta o inversa, un valor nulo muestra que no hay tendencia entre ambas variables o sea compleja ejemplo, que la relación sea en forma de U. La magnitud toma valores de -1 a 1 e indica con que fuerza se da la relación; entre más se acerque a los extremos ya sea -1 o 1 la tendencia será mayor entre las variables analizadas, caso contrario si se acerca a cero la tendencia será débil y la nube de puntos estará disperso; finalmente si la correlación es igual a 1 o -1 se dice que es perfecta, caso contrario si es igual a cero no tienen correlación.

2.2.8 Odd Ratio

El Odd Ratio (OR) también conocido como la razón de posibilidades, es una herramienta estadística de medida para la asociación de variables dicotómicas. El OR estima la relación de variables dicotómicas binarias con un cierto intervalo de confianza, y permite interpretar de una forma más adecuada en lo que a eficacia se refiere. El OR permite determinar la relación de variables mediante regresión logística. El concepto Odds hace referencia a la razón entre la probabilidad de que un evento ocurra sobre la probabilidad de no ocurrencia (Aedo, Pavlov, & Clavero, 2010). El valor de OR siempre va desde el cero pudiendo llegar hasta el infinito y su interpretación se lo puede realizar de la siguiente manera como se observa en la Tabla 2:

Odd Ratio	Interpretación
OR>1	Alta probabilidad de ocurrencia del evento
OR<1	Baja probabilidad de ocurrencia del evento
OR=1	Probabilidad nula, no hay correlación de variables

Tabla 2

Fuente: Autor (2023)

Tabla que muestra la Interpretación OR

Cuando OR es mayor que 1 significa que existe una correlación entre las variables analizadas, además, mientras más grande sea dicho valor mayor será la correlación y la probabilidad de ocurrencia del evento aumentará considerablemente.

CAPÍTULO III

MARCO METODOLÓGICO

La investigación realizada es de tipo descriptiva correlacional ya que trabaja con una muestra de la población y relaciona dos variables; descriptiva debido a que detalla las características específicas de los estudiantes, las mismas que fueron obtenidas de la ficha socioemocional aplicada a los estudiantes que ingresan al Instituto. Una vez conseguido los datos se hace uso del método cuantitativo para medir las propiedades, características o componentes obtenidas de la investigación exploratoria de la manera más precisa posible (Díaz & Calzadilla, 2016).

Por otra parte, se aplicó la investigación correlacional porque establece la relación que existe entre las variables de entrada o descriptoras con la variable de salida o independiente y mediante el uso de un modelo matemático se descarta las variables que no tengan relación entre sí, dejando a los que tengan un fuerte grado de correlación. Este tipo de investigación tiene un valor explicativo, aunque sea solo parcial, dado que al relacionar dos variables puede explicar un fenómeno en cierta manera, una vez obtenido la relación se debe analizar la correlación; se concluye que a mayor cantidad de variables relacionadas la explicación será mejor (Hernandez, Fernández, & Baptista, 1997).

3.1 Ubicación

La presente investigación se lo realizó a los estudiantes del primer ciclo del Instituto Superior Tecnológico “Tres de Marzo” sede Guaranda, pertenecientes a las Tres carreras Tecnológicas que se mencionan a continuación: Tecnología

Superior en Electricidad, Electrónica y Mecánica Automotriz; la cual está ubicada en la ciudad Guaranda de la provincia de Bolívar, Ecuador.

3.2 Equipos y materiales

Para la elaboración del presente proyecto de titulación los recursos empleados fueron los siguientes:

- Software (Office, Gretel, R Commander)
- World Wide Web
- Laptop
- Impresora
- Ficha socioemocional de la Institución (entrevista)

3.3. Tipo de investigación

3.3.1 Investigación bibliográfica

La investigación bibliográfica consiste en la recopilación de información secundaria, para la formulación del marco teórico y conceptual de la investigación, obteniendo a través de libros, artículos de revista, informes de investigación, artículos científicos publicados, entre otros, los mismos que permiten la obtención de información relacionada a los objetivos y necesidades de la investigación (Rivero, 2020).

3.3.2 Investigación seccional o transversal

La investigación transversal es aquella en el cual los datos de la población o muestra, motivo de estudio son obtenidos en un instante determinado. Desde esta perspectiva Briones (1985), menciona que este tipo de investigación se asemejan a

tomar una fotografía instantánea al fenómeno que se estudia. Finalmente, esta puede ser de carácter descriptiva o a su vez explicativa (Bernal, Metodología de la Investigación, 2010).

3.3.3 Investigación cuantitativa

La investigación cuantitativa usa las matemáticas mediante ciertas herramientas estadísticas informáticas y su objetivo es encontrar resultados a partir de estos. Esta permite cuantificar las variables del problema en estudio, de manera que se pueda generalizar y los resultados obtenidos puedan ser proyectados a otra población que tenga características similares a las ya estudiadas. Esta investigación por lo general es de carácter exploratoria y regularmente se suele aplicar después de la investigación cualitativa con el objetivo de medir la magnitud e interpretar los resultados de manera objetiva (Investigadores, 2020).

3.3.2 Investigación explicativa o causal

La investigación causal es ideal para realizar la investigación no experimental y tiene como base fundamental probar la hipótesis y mediante las conclusiones obtenidas formular ciertos principios científicos o leyes; para determinar si una investigación es de carácter explicativo se debe tomar en cuenta si el investigador estudia hechos, fenómenos o situaciones en los cuales se está buscando el porqué de las cosas. Esta investigación relaciona las variables de causa-efecto (Bernal, Metodología de la Investigación, 2010).

3.4. Prueba de hipótesis

En la presente investigación se procura verificar las hipótesis siguientes:

Hipótesis nula

El modelo matemático para establecer los factores asociados a las calificaciones escolares utilizando Regresión Logística NO es el apropiado.

Hipótesis alternativa

El modelo matemático para establecer los factores asociados a las calificaciones escolares utilizando Regresión Logística es el apropiado.

3.5. Población o muestra

La población estudiada en esta investigación son 489 estudiantes del primer ciclo pertenecientes a los años del 2019-2020 y 2021 del Instituto Superior Tecnológico “Tres de Marzo” de la provincia de Bolívar, siendo el total de toda la población, por lo tanto, no fue necesario obtener una muestra ya que se trabajó con la población completa.

3.6. Recolección de información

La recolección de datos es una técnica que el investigador utiliza para procesar los datos obtenidos en las actividades realizadas para la obtención de información, la cual es necesaria para responder la hipótesis de su investigación. Para alcanzar el objetivo de recolección de datos se debe plantear ciertos pasos como estrategia previa a la obtención de esta, esto conlleva que el investigador deba buscar el escenario adecuado para la recolección, análisis y posterior tabulación de datos. Además, debe contar con el personal adecuado, recursos económicos, físicos, recursos bibliográficos en temas referentes a investigaciones similares, muestra e instrumentos para la recolección de datos (Sanchez, Fernández, & Díaz, 2021).

Para la obtención de los datos en la presente investigación se hizo el uso de la base de datos del personal estudiantil perteneciente al Instituto Superior Tecnológico “Tres de Marzo”, que está desarrollada conforme a la guía de registro de matriculados de los Institutos públicos y privados proporcionados por la Senescyt y se aplica al instante que los estudiantes ingresan, como parte de los requisitos que deben cumplir para formar parte del Instituto, la información obtenida fue del primer semestre de Tecnología Superior en Electricidad, Electrónica y Mecánica Automotriz.

La guía para el registro de información de institutos públicos y particulares otorgado por la Secretaría de Educación Superior, Ciencia, Tecnología e Innovación del Ecuador es un manual que tiene como objetivo dar lineamientos para recolectar la información tanto de estudiantes como de personal administrativo pertenecientes a la educación superior y está dirigido a funcionarios que están encargados de la gestión de información (Senescyt, 2018).

3.7. Procesamiento de información y análisis estadístico

En la presente investigación para obtener la información de los estudiantes se hará uso de la base de datos existentes, para su posterior análisis estadístico. El procesamiento de dicha información se lo realizará en el Software de cálculo por hojas Microsoft Excel para lo cual se procederá a tabular los datos de los diferentes estudiantes matriculados en el primer semestre de Tecnología Superior de las carreras de Electricidad, Electrónica y Mecánica Automotriz.

Para el procesamiento de la información se debe realizar una serie de actividades con el objetivo de obtener datos inequívocos por lo tanto, primeramente se debe determinar las fuentes reales de información aplicando la comprensión e

interpretación correcta de éstas; consecutivamente, se debe sistematizar dicha información en una base de datos con la finalidad de poder resumir, jerarquizar, descomponer o integrar, y descubrir relaciones entre unos datos y otros para reproducir información de carácter científico. Otra actividad es relacionar los datos obtenidos con investigaciones previas para valorar, depurar y emitir juicios de valor en datos atípicos, para llegar a conclusiones. Finalmente, se puede hacer uso de software como SPSS, SVS, GRETL, SIG, ARGIS, STIMATE, R COMMANDER, etc., para organizar la información, conocer la correlación de datos, determinar tendencias, realizar gráficas, estadísticas, realizar informes y exponer los principales resultados, para dar posibles soluciones a los problemas encontrados en caso de haberlos (Figueredo, León, & Martínez, 2019).

En base a lo anteriormente mencionado como primera actividad se depurará los datos que no sean relevantes para establecer los factores asociados a las calificaciones escolares; de los 25 descriptores proporcionados por el Instituto se los analizará en el programa R Commander para medir el nivel de correlación con la variable de salida y se obtendrá la Tabla 3. Como segunda actividad será elaborar los diagramas estadísticos de los estos factores para una mejor interpretación.

Luego de haber definido los descriptores para el modelo matemático se convertirá a cada uno de ellos a cofactores, debido a que dichos descriptores son de carácter discreto más no continuo y finalmente se modelará mediante regresión logística estos cofactores para determinar el “p valor” de cada una de ellas, de esta manera se valorará cuáles son los factores asociados a las calificaciones escolares de los estudiantes del primer ciclo del Instituto Superior Tecnológico Tres de Marzo de la provincia Bolívar.

La tabla 3 que se presenta a continuación establece a las variables descriptoras que serán tomadas para el procesamiento, análisis y recolección de los datos, en el cual se observa las variables que serán tomadas en cuenta en base al manual de guía para el registro de información de institutos públicos y particulares otorgado por la Secretaría de Educación Superior, Ciencia, Tecnología e Innovación del Ecuador, para diseñar el modelo matemático utilizando regresión logística multivariante que permitirá predecir variables académicas, físicas o actitudinales asociadas al fracaso académico y al bajo rendimiento.

Calificación	estadocivilId	nivelFormacionPadre	nivelFormacionMadre	IngresoTotal Hogar	cantidadMiembros Hogar
--------------	---------------	---------------------	---------------------	--------------------	------------------------

Tabla 3

Fuente: Autor (2023)

Tabla que muestra los Factores asociados a las calificaciones escolares

3.8. Variables

Mediante la revisión literaria realizada en el Capítulo II de la presente Investigación con respecto a estudios previos relacionados al tema, el sector escolar es uno de los más grandes a nivel mundial y uno de los más susceptibles a ser condicionado por los factores influyentes, es decir se ven afectados por cualquier factor asociado. Según Capera (2015) la deserción escolar es una de las problemáticas más costosas a nivel económico puesto que la inversión realizada en cada inicio del ciclo escolar se ve interrumpida por que muchos de los estudiantes carecen de sostenibilidad debido al estatus económico, niveles académicos de los progenitores y el estado civil del estudiante (Guerra, 2015).

Inicialmente la base de datos original maneja 56 variables entre ellas discapacidad, etnia, tipo de sangre, cedula de identidad, código del estudiante, entre

otros; pero en base a lo anteriormente investigado se considera como variables independientes o descriptoras a estado civil, nivel Formación Padre, nivel Formación Madre, ingreso Total Hogar y cantidad Miembros Hogar una vez realizado la revisión literaria, como se observa en la siguiente tabla 4:

Variables según (Capera, 2015)	Variables según (Guerra, 2015)	Variables según Autor (2023)
Género	Variables socioeconómicas	Estado civil
Edad	Nivel de ingreso	Nivel Formación Padre
Estado Civil	Situación laboral	Nivel Formación Madre
Tiempo dedicado al trabajo	Lugar de residencia	Ingreso Total Hogar
Lugar de residencia	Variables académicas	Cantidad Miembros Hogar una vez
Ingreso individual mensual	Rendimiento académico	Calidad de la educación
Promedio académico de bachillerato con que ingresó a la universidad	Variables de adaptación	Integración social
Meta educativa (grado que espera(ba) obtener)	Apoyo de la familia	Variables de motivación
Carga académica en número de créditos tomados por semestre	Satisfacción con la carrera	Interés por los estudios
Número de cursos a distancia que toma actualmente (que tomó en total)		
Promedio académico actual (alcanzado en general)		
Tipo de colegio de donde obtuvo el grado de bachiller		

Tabla 4

Fuente: Autor (2023)

Tabla que muestra las Variables utilizadas por cada uno de los autores.

Variables independientes: estadoCivilId, nivelFormacionPadre, nivelFormacionMadre, ingresoTotalHogar y cantidadMiembrosHogar.

Por otra parte, en cambio con respecto a la variable dependientes o respuesta para esta investigación es la calificación que obtiene el estudiante:

Variable dependiente: Calificación

Como se puede observar las variables que se están trabajando son de tipo cualitativas por lo que es necesario convertir a cofactores para que el software que vaya a realizar la modelación matemática la tome como datos discretos y determine la correlación variable entrada-salida.

Los datos que se obtuvieron para cada una de las variables fueron de la base de datos de la secretaría del Instituto, el cual recaba la información mediante una entrevista al estudiante para llenar el formulario de matriculación, al momento que ingresa a una de las carreras. Para una mejor comprensión de los datos, las variables: cantidad de miembros en el hogar y total de ingreso del hogar, se trabaja con datos agrupados para clasificar la información de mejor manera como se observa en la Tabla 5.

Variables	Cofactores
Calificación	APRUEBA REPRUEBA
Estado Civil	SOLTERO/A CASADO/A DIVORCIADO/A UNION LIBRE VIUDO/A
Nivel de Formación del Padre	NINGUNO JARDIN ALFABETIZACION EDUCACION BASICA PRIMARIA EDUCACION MEDIA SECUNDARIA SUPERIOR NO UNIVERSITARIO SUPERIOR UNIVERSITARIO POSTGRADO
Nivel de Formación de la Madre	NINGUNO JARDIN ALFABETIZACION EDUCACION BASICA

	PRIMARIA EDUCACION MEDIA SECUNDARIA SUPERIOR NO UNIVERSITARIO SUPERIOR UNIVERSITARIO POSTGRADO
Ingreso Total del Hogar	MENOR A \$300 ENTRE \$300-\$599 ENTRE \$600-\$899 MAYOR A \$899
Cantidad de Miembros del Hogar.	ENTRE 1-3 ENTRE 4-6 ENTRE 7-9 MAS DE 9

Tabla 5

Fuente: Autor (2023)

Tabla que detalla a las Variables con sus respectivos cofactores

CAPÍTULO IV

RESULTADOS Y DISCUSIÓN

De la entrevista aplicada a los estudiantes del primer semestre, se obtiene la información para la presente investigación, en el cual se busca determinar los factores asociados a las calificaciones escolares, en los estudiantes del primer ciclo del Instituto Superior Tecnológico “Tres de Marzo”, dando como resultado lo siguientes datos que se detallan a continuación.

4.1 Análisis descriptivo de las variables

De acuerdo con las variables determinadas para esta investigación se tiene como variables independientes: el estado civil, al nivel de formación del padre, el nivel de formación de la madre, ingreso total del hogar y cantidad de miembros del hogar. Por otra parte, en cambio con respecto a la variable dependiente se tiene a la Calificación.

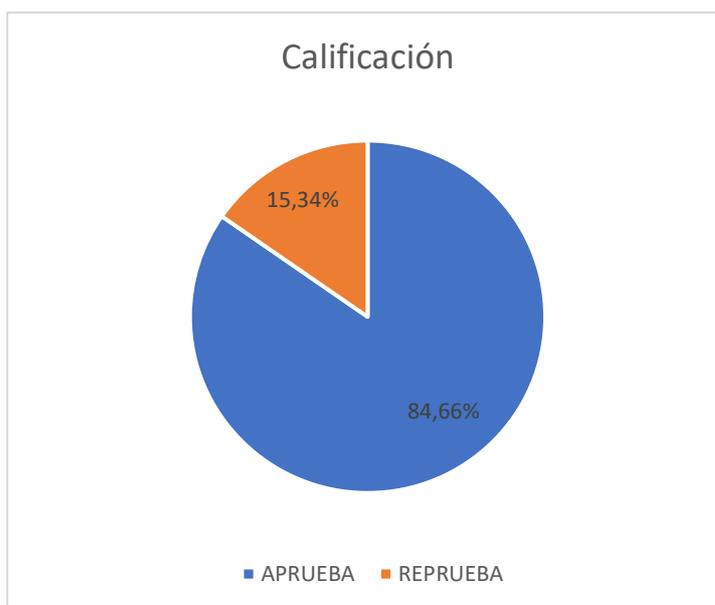


Figura 3
Fuente: Autor (2023)
Estadística que corresponde a las calificaciones de los estudiantes

Una vez tabulado los datos de los estudiantes del primer semestre con respecto a las calificaciones obtenidas se puede observar que el 84,66% obtuvieron notas en el rango de 7 a 10 puntos que son considerados como buenas calificaciones, por otra parte, el 15,34% de los estudiantes obtuvieron notas menores a 7 consideradas como malas, debido a que no cumple el mínimo para aprobar el semestre por lo que es necesario determinar los factores asociados a este problema dentro del instituto, para tomar las debidas medidas del caso.

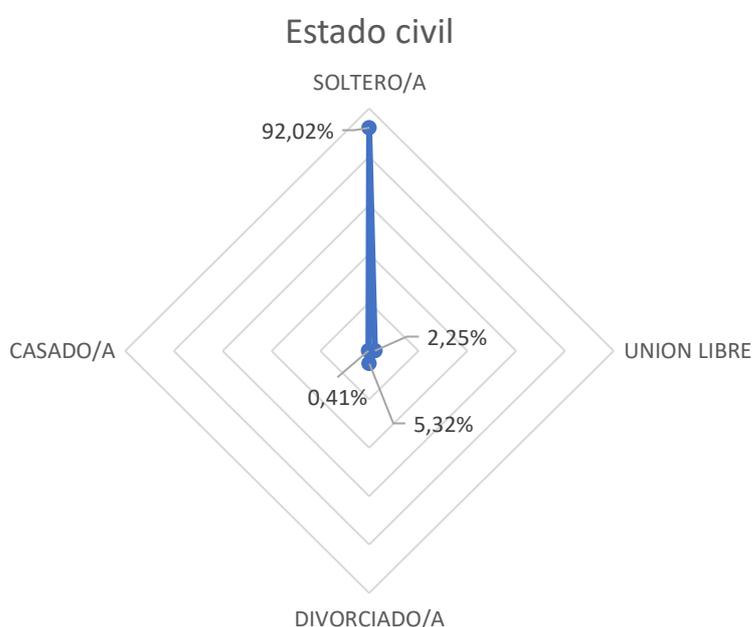


Figura 4
Fuente: Autor (2023)
Estadística que corresponde al Estado Civil de los estudiantes

La entrevista realizada a los estudiantes con respecto al estado civil dio como resultado que el 92,02% son solteros/as, el 2,25% están en unión libre, el 5,32% son divorciados/as y el 0,41% son casados/as, siendo este cofactor el de menor porcentaje registrado.

Nivel de Formación del Padre

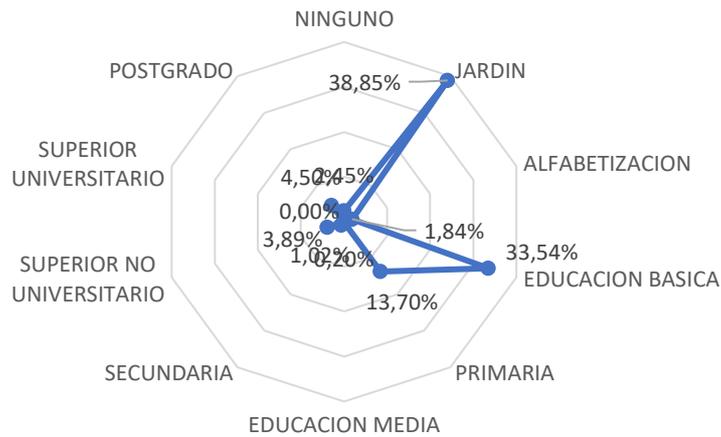


Figura 5
 Fuente: Autor (2023)
 Estadística que corresponde al Nivel de formación del padre

Con respecto al nivel de formación del padre la mayoría de los estudiantes contestan que el padre solo tiene formación de jardín lo que representa el 38,85%, el 2,45% responde que ninguno, el 1,84% alfabetización, el 33,54% educación básica, el 13,70% primaria, el 0,20% educación media, el 1,02% secundaria, el 3,89% superior no universitario, el 4,50% postgrado y 0% superior universitario.

Nivel de Formación de la Madre

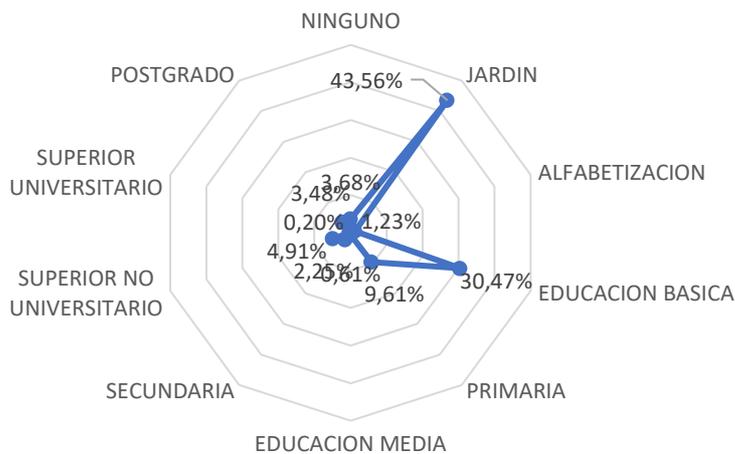


Figura 6
 Fuente: Autor (2023)
 Estadística que corresponde al Nivel de formación de la madre

Con respecto al nivel de formación de la madre, la mayoría de los estudiantes contestan que la madre alcanzó sus estudios hasta el jardín lo que representa el 43,56%, el 3,68% responde que ninguno, el 1,23% alfabetización, el 30,47% educación básica, el 9,61% primaria, el 0,61% educación media, el 2,25% secundaria, el 4,91% superior no universitario, el 0,20% superior universitario y finalmente el 3,48% postgrado.

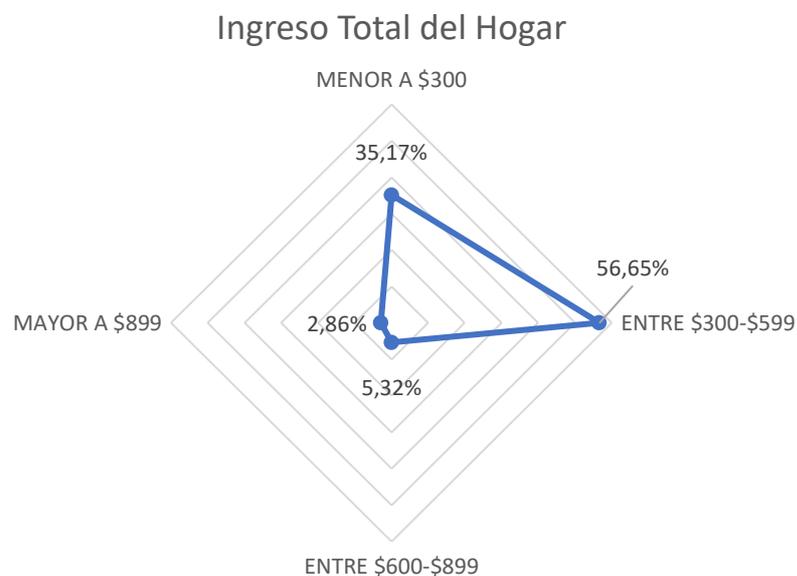


Figura 7
Fuente: Autor (2023)
Estadística que corresponde al Ingreso total del hogar

En lo que corresponde al ingreso económico que percibe el hogar de los estudiantes entrevistados la mayoría de los estudiantes responden que tienen un ingreso entre \$300-\$599 que representa el 56,65%, entre \$600-\$899 el 5,32%, menor a \$300 el 35,17% y mayor a \$899 solamente el 2,86% que representa el menor porcentaje, en el cual los entrevistados contestan al reactivo realizado. Esta información será útil al momento de realizar la correlación de variables, para determinar cuál de los cofactores formaran parte de los factores asociados a las calificaciones motivo de estudio.

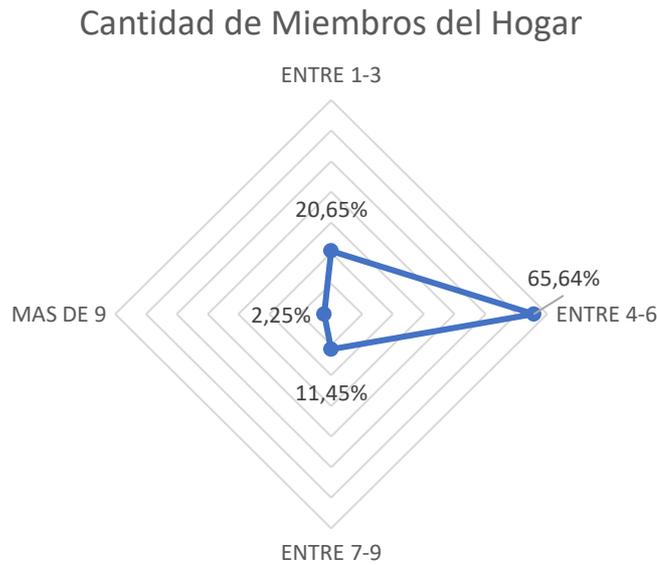


Figura 8
Fuente: Autor (2023)
Estadística que corresponde a Cantidad de miembros del hogar

Finalmente, en el reactivo de la entrevista concerniente a la cantidad de miembros del hogar la mayoría de las estudiantes contestaron que dentro de su hogar viven entre 4-6 que representa el 65,64%, entre 1-3 el 20,65%, entre 7-9 el 11,45% y más de 9 el 2,25% que representa la minoría de toda la población entrevistada. Abordar acerca del rendimiento académico es un tema complejo ya que las variables intervinientes varían dependiendo el contexto; por ejemplo, en esta investigación no existen estudiantes que vivan solos por lo cual no se puede tomar en cuenta en el modelo matemático así lo ratifica (Martínez, Ferrás, Bermúdez, Cabrera, & Pérez, 2020) en su trabajo titulado “Rendimiento académico en estudiantes Vs factores que influyen en sus resultados: una relación a considerar”, al mencionar que este tipo de estudio es difícil de abordar en un sola investigación por ser de carácter multifactorial.

4.2 Estadísticos descriptivos

```
> summary(Dataset)
  Calificación  estadocivilId  nivelFormacionPadre nivelFormacionMadre ingresoTotalHogar cantidadMiembros
Min.   :0.0000  Min.   :0.2500  Min.   :0.1000    Min.   :0.1000    Min.   :0.0500  Min.   :0.0000
1st Qu.:0.0000  1st Qu.:0.2500  1st Qu.:0.3000    1st Qu.:0.3000    1st Qu.:0.1300  1st Qu.:0.2500
Median :0.0000  Median :0.2500  Median :0.4000    Median :0.4000    Median :0.1500  Median :0.2500
Mean   :0.1534  Mean   :0.2822  Mean   :0.4288    Mean   :0.4213    Mean   :0.1765  Mean   :0.2725
3rd Qu.:0.0000  3rd Qu.:0.2500  3rd Qu.:0.5000    3rd Qu.:0.5000    3rd Qu.:0.2000  3rd Qu.:0.2500
Max.   :1.0000  Max.   :1.0000  Max.   :1.0000    Max.   :1.0000    Max.   :1.0000  Max.   :1.0000
```

Figura 9

Fuente: Autor (2023)

Estadística que corresponde a los Estadísticos descriptivos del Dataset

En base a la información obtenida según el instrumento de recolección de datos de carácter categórico, se establece el análisis descriptivo de variables; el cual se interpreta de forma discreta ya que las variables numéricas representan factores. Por esta razón, el rango va de 0 a 1, motivo por el cual los máximos y mínimos son “1” y “0” respectivamente, por lo tanto, las medidas de tendencia central como la media, la mediana y los terceros cuartiles se representan en la figura 9. Así se tiene como valores máximos de media, al nivel de formación del padre con un valor de 0.428 y el mínimo de 0.153 a la calificación obtenida por el estudiante, debido que se está trabajando en una escala binaria. Además, el tercer cuartil nos da información sobre la tendencia que tiene el 75% de los datos de la investigación realizada en este caso, por ejemplo, para la variable estado civil representa un 0.25 = “SOLTERO/A” lo que representa que la mayoría de los estudiantes tiene ese tipo de estado civil.

4.3 Modelación matemática – Regresión Logística

Modelo de Regresión Logística (Logit-binomial)

$$\ln\left(\frac{\pi_1}{1 - \pi_1}\right) = \beta_0 + \beta_1 x_1 + \dots + \beta_i x_i$$

Expresada explícitamente en función de probabilidad (π_1) el modelo Logit se tiene:

$$\pi_1 = \beta_0 + \beta_i x_i$$

Donde π_1 , es la probabilidad de que la variable x_i sea uno de los factores asociados a las calificaciones escolares el cual esté interviniendo al rendimiento académico del estudiante. Dado que nuestro modelo es de carácter binario o dicotómico, para el caso de dos variables explicativas finalmente nuestro modelo quedaría expresado de la siguiente manera:

$$\pi_{(y=1)} = \frac{1}{1 - e^{-(\beta_0 + \beta_1 x_1 + \dots + \beta_i x_i)}}$$

Donde la variable dependiente “y” hace referencia a Calificación, para determinar si aprueba o reprueba. Dicho esto, del modelo de regresión logística, se construye el modelo matemático para la presente investigación con las variables intervinientes basadas en la entrevista aplicada a los estudiantes del instituto educativo, dando como resultados los siguientes coeficientes.

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-16.2442	1457.5416	-0.011	0.99111
cantidadMiembrosHogar[T.ENTRE 4-6]	0.2356	0.3752	0.628	0.53015
cantidadMiembrosHogar[T.ENTRE 7-9]	0.1580	0.5120	0.309	0.75757
cantidadMiembrosHogar[T.MAS DE 9]	2.9760	0.9694	3.070	0.00214 **
estadocivilId[T.DIVORCIADO/A]	-17.0143	2739.4642	-0.006	0.99504
estadocivilId[T.SOLTERO/A]	-1.3454	0.5268	-2.554	0.01065 *
estadocivilId[T.UNION LIBRE]	0.7538	0.8189	0.920	0.35734
ingresoTotalHogar[T.ENTRE \$600-\$899]	-0.5342	0.8159	-0.655	0.51264
ingresoTotalHogar[T.MAYOR A \$899]	0.7566	0.9033	0.838	0.40225
ingresoTotalHogar[T.MENOR A \$300]	0.5907	0.2934	2.013	0.04411 *
nivelFormacionMadre[T.EDUCACION BASICA]	15.9411	1457.5418	0.011	0.99127
nivelFormacionMadre[T.EDUCACION MEDIA]	17.5169	1457.5424	0.012	0.99041
nivelFormacionMadre[T.JARDIN]	16.7099	1457.5418	0.011	0.99085
nivelFormacionMadre[T.NINGUNO]	-0.7615	1670.6708	0.000	0.99964
nivelFormacionMadre[T.POST-GRADO]	15.9170	1457.5420	0.011	0.99129
nivelFormacionMadre[T.PRIMARIA]	17.0699	1457.5418	0.012	0.99066
nivelFormacionMadre[T.SECUNDARIA]	16.3220	1457.5421	0.011	0.99107
nivelFormacionMadre[T.SUPERIOR NO UNIVERSITARIO]	16.7133	1457.5420	0.011	0.99085
nivelFormacionMadre[T.SUPERIOR UNIVERSITARIO]	1.1523	4216.1347	0.000	0.99978
nivelFormacionPadre[T.EDUCACION BASICA]	-0.8301	1.0361	-0.801	0.42304
nivelFormacionPadre[T.EDUCACION MEDIA]	-16.1522	3956.1805	-0.004	0.99674
nivelFormacionPadre[T.JARDIN]	-1.1701	0.9699	-1.206	0.22767
nivelFormacionPadre[T.NINGUNO]	-2.1990	1.5063	-1.460	0.14432
nivelFormacionPadre[T.POST-GRADO]	-0.4371	1.0819	-0.404	0.68620
nivelFormacionPadre[T.PRIMARIA]	-2.3451	1.0634	-2.205	0.02744 *
nivelFormacionPadre[T.SECUNDARIA]	-16.5149	1640.1675	-0.010	0.99197
nivelFormacionPadre[T.SUPERIOR NO UNIVERSITARIO]	-1.6104	1.4314	-1.125	0.26055

En los modelos de Regresión Logística, para determinar si existe correlación de variables, el p-valor debe ser menor a 0.05; por lo tanto, la hipótesis nula tiende a ser falsa. Esto significa que dichos resultados son importantes para determinar si los factores analizados afectan al estudiante para que este repruebe.

En consecuencia, se puede observar que ciertos factores tienen una correlación alta que incide en la reprobación del ciclo educativo de los estudiantes. De esta forma se tiene: cantidadMiembrosHogar [T.MAS DE 9] con un p-value = 0.00214 **, estadocivilId [T.SOLTERO/A] con su respectivo p-value = 0.01065 *, ingresoTotalHogar [T.MENOR A \$300] con un p-value = 0.04411 *, y finalmente nivelFormacionPadre [T.PRIMARIA] con un p-value = 0.02744 *. Donde los asteriscos representan que, a mayor cantidad, mayor será la correlación de las

variables. Por lo tanto, para el contexto aplicado, el modelo quedaría expresado de la siguiente manera:

$$\Pi_{(y=1)} = \frac{1}{1 - e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4)}}$$

β_0 : término de intercepto.

$\beta_{1,2,3,4}$: parámetro pendiente.

x_1 : cantidadMiembrosHogar [T.MAS DE 9]

x_2 : estadocivilId [T.SOLTERO/A]

x_3 : ingresoTotalHogar [T. MENOR A \$300]

x_4 : nivelFormacionPadre [T.PRIMARIA]

Reemplazando los coeficientes en base a los valores obtenidos al modelo matemático para determinar los factores asociados al rendimiento académico de los estudiantes del primer ciclo de las carreras Tecnología Superior en Electricidad, Electrónica y Mecánica Automotriz, del Instituto Superior Tecnológico “Tres de Marzo” sede Guaranda se tiene:

$$\Pi_{(y=1)} = \frac{1}{1 - e^{17.05245 - 3.20832x_1 + 1.33910x_2 - 0.58332x_3 + 2.39738x_4}}$$

Cabe recalcar que el tener una correlación grande entre variables, en este caso cofactores, no garantiza que dicha variable este haciendo que los estudiantes reprobren, por lo que es necesario realizar otro paso adicional para poder interpretar la probabilidad de ocurrencia del evento. Para lo cual, se debe sacar el exponencial a los coeficientes del modelo ya que los coeficientes miden el cambio del logaritmo conocida como odds ratio (OR) el cual es la razón de la probabilidad de éxito vs fracaso en el que ocurre el evento cuando “X” incrementa una unidad.

Como se pudo observar en el modelo matemático la función logit no permite que los coeficientes se puedan interpretar de forma directa, por lo que es necesario obtener su exponencial para obtener la razón par (razón de momios, odds ratios) de la variable x_i , como se demuestra a continuación.

$$e^{\ln\left(\frac{\pi_1}{1-\pi_1}\right)} = e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4}$$

$$\frac{\pi_1}{1 - \pi_1} = e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4}$$

$$\frac{\pi_1}{1 - \pi_1} = e^{\beta_0} e^{\beta_1 x_1} e^{\beta_2 x_2} e^{\beta_3 x_3} e^{\beta_4 x_4}$$

e^{β_i} → es la razón par, de la variable interviniente x_i

Como se demuestra, los coeficientes en si no se interpretan debido a que están relacionadas a la razón odd del logaritmo, por lo que es necesario interpretar su exponencial para conocer la razón odd; además, se debe tomar en cuenta que la razón odd del exponencial debe ser mayor a 0, porque si es igual a cero significa que dicha variable no influye dentro del modelo y si es negativa influye de forma inversa dentro del modelo desarrollado. Por lo tanto, solo se toma en cuenta los factores que influyan en las calificaciones escolares para un posible “reprobado”. Entre estas se tiene: cantidad miembros en el hogar [entre 4-6, entre 7-9, más de 9], estado civil [unión libre], ingreso total del hogar [mayor a \$899, menor a \$300] y nivel formación de la madre [educación básica, educación media, jardín, postgrado, primaria, secundaria, superior no universitario, superior universitario]; en la figura 10, se puede observar los valores Odd Ratios para cada una de las variables intervinientes antes detalladas.

	(Intercept)	8.815239e-08
cantidadMiembrosHogar	[T.ENTRE 4-6]	1.265626e+00
cantidadMiembrosHogar	[T.ENTRE 7-9]	1.171196e+00
cantidadMiembrosHogar	[T.MAS DE 9]	1.960964e+01
estadocivilId	[T.DIVORCIADO/A]	4.081112e-08
estadocivilId	[T.SOLTERO/A]	2.604237e-01
estadocivilId	[T.UNION LIBRE]	2.124981e+00
ingresoTotalHogar	[T.ENTRE \$600-\$899]	5.861464e-01
ingresoTotalHogar	[T.MAYOR A \$899]	2.131003e+00
ingresoTotalHogar	[T.MENOR A \$300]	1.805220e+00
nivelFormacionMadre	[T.EDUCACION BASICA]	8.377475e+06
nivelFormacionMadre	[T.EDUCACION MEDIA]	4.050349e+07
nivelFormacionMadre	[T.JARDIN]	1.807316e+07
nivelFormacionMadre	[T.NINGUNO]	4.669648e-01
nivelFormacionMadre	[T.POST-GRADO]	8.178335e+06
nivelFormacionMadre	[T.PRIMARIA]	2.590334e+07
nivelFormacionMadre	[T.SECUNDARIA]	1.226155e+07
nivelFormacionMadre	[T.SUPERIOR NO UNIVERSITARIO]	1.813390e+07
nivelFormacionMadre	[T.SUPERIOR UNIVERSITARIO]	3.165441e+00
nivelFormacionPadre	[T.EDUCACION BASICA]	4.360067e-01
nivelFormacionPadre	[T.EDUCACION MEDIA]	9.664929e-08
nivelFormacionPadre	[T.JARDIN]	3.103431e-01
nivelFormacionPadre	[T.NINGUNO]	1.109091e-01
nivelFormacionPadre	[T.POST-GRADO]	6.458931e-01
nivelFormacionPadre	[T.PRIMARIA]	9.583945e-02
nivelFormacionPadre	[T.SECUNDARIA]	6.724919e-08
nivelFormacionPadre	[T.SUPERIOR NO UNIVERSITARIO]	1.998076e-01

Figura 10
Fuente: Autor (2023)
Valores Odds Ratios de las variables intervinientes

Una vez hallado el valor odd ratio de cada coeficiente se debe tomar en cuenta solo a las variables que obtuvieron odds positivos por lo mencionado anteriormente como se muestra en la Tabla 6.

Variable	Coefficiente	Razón Odd	Interpretación
cantidadMiembrosHogar[T.ENTRE 4-6]	0,2356	1,265626	El estudiante que vive con una cantidad de miembros de hogar entre 4-6 tiene un 26,5% de reprobación
cantidadMiembrosHogar[T.ENTRE 7-9]	0,1580	1,171196	El estudiante que vive con una cantidad de miembros de hogar entre 7-9 tiene un 17,1% de reprobación
cantidadMiembrosHogar[T.MAS DE 9]	2,9760	19,60964	El estudiante que vive con una cantidad de miembros mayor a 9 tiene una probabilidad de 19000,61% de reprobación
EstadocivilId [T.UNION LIBRE]	0,7538	2,124981	Ser estudiante que tiene una relación de Unión libre tiene el 200,12% de probabilidad de reprobación
IngresoTotalHogar [T.MAYOR A \$899]	0,7566	2,131003	Si el ingreso en el hogar del estudiante es mayor a \$899 tiene el 200,13% de probabilidad de reprobación a que tenga un ingreso menor
IngresoTotalHogar [T.MENOR A \$300]	0,5907	1,805220	Si el ingreso en el hogar del estudiante es menor a \$300 la probabilidad de reprobación es menor al 80,6%
NivelFormacionMadre [T.EDUCACION BASICA]	15,9411	8377475	
NivelFormacionMadre	17,5169	40503490	Como se puede observar en la presente

[T.EDUCACION MEDIA]			tabla, en el análisis de los odds ratios en el nivel de formación de la madre la razón se dispara a valores muy altos lo que significa que, en sí, para esta investigación si afecta y en gran manera a excepción de no tenerlo.
NivelFormacionMadre [T.JARDIN]	16,7099	18073160	
NivelFormacionMadre [T.POSTGRADO]	15,9170	8178335	
NivelFormacionMadre [T.PRIMARIA]	17,0699	25903340	
NivelFormacionMadre [T.SECUNDARIA]	16,3220	12261550	Así, por ejemplo, tenemos en el nivel de formación superior universitario.
NivelFormacionMadre [T.SUPERIOR NO UNIVERSITARIO]	16,7133	18133900	
NivelFormacionMadre [T.SUPERIOR UNIVERSITARIO]	1,1523	3,165441	La probabilidad de que el estudiante repruebe aumenta a un 300,16%

Tabla 6

Fuente: Autor (2023)

Tabla que detalla el análisis de los Odds Ratios

Al interpretar los resultados obtenidos en el cálculo de los odd ratios se puede determinar que los cofactores de la variable nivel de formación de la madre son los que incrementan en gran manera a la reprobación del ciclo académico; por lo que se puede afirmar que, esta variable es uno de los factores asociados a las calificaciones escolares del estudiante y para corroborar se puede comprobar mediante de una revisión literaria.

Los resultados anteriormente mencionados se alinean a la investigación realizada en la Universidad Pontificia Comillas de España, bajo el tema "Influencia del Nivel Educativo de los Padres en el Rendimiento Académico de los Estudiantes de ADE. Un Enfoque de Género", en el cual concluye que el nivel de formación del padre no afecta en el rendimiento académico de los estudiantes más no así el nivel de formación de la madre donde sí existe una diferencia significativa y si afecta a

las calificaciones escolares y por ende a su rendimiento académico (*Martinez De Ibarreta, 2022*). Además, en otra investigación realizada en México a 21724 estudiantes se llega a la misma conclusión; en el cual, se evidencia que el nivel de formación de la madre si influye en el rendimiento académico, con un valor de media cuadrática del conglomerado de 4999.71, una media cuadrática del error de 0.65 y un valor F de ANOVA de 7685.82* siendo el dato que más contribuye a diferenciar el rendimiento académico (*Chaparro, González, & Caso, 2016*).

La variable cantidad de miembros en el hogar también influye en el rendimiento académico del estudiante para que este repruebe el curso, por lo tanto, esta variable también se considera como uno de los factores asociados a la calificación escolar dentro del sistema educativo del mismo.

Según la prensa El Mercurio de Cuenca-Ecuador en su artículo escrito por Lourdes Alvarado con el título “Problemas familiares y rendimiento académico”, menciona que: *“El bajo rendimiento académico se puede dar por diversas causas, tales como: desintegración familiar, estilos de crianza, padres trabajadores, desinterés de los padres, adicciones, hijos predilectos, hijos no deseados, etc.”* (Alvarado, Problemas familiares y rendimiento académico, 2000). Por lo tanto, que dentro de una familia exista más de 9 miembros viviendo dentro de un mismo hogar produce que esto afecte en el rendimiento académico; ya que no existe la capacidad salarial para cubrir los gastos Universitarios, ni los recursos necesarios para amoblar una sala de estudios acorde a la necesidad de un estudio superior para tantos miembros.

Finalmente, la variable estado civil Unión Libre dentro de esta investigación es uno de los factores asociados a las calificaciones escolares; esta información se

corroborar en el estudio de caso de (Ruiz, García, Luis, & Ximhai) 2014, donde expresa que los estudiantes que cursan la Universidad y están en una relación de Unión Libre tienden a desertar debido a que sus parejas no les dan el apoyo necesario para continuar, sino que exigen dedicarse al hogar. De tal manera que, los estudios previamente revisados acreditan a los resultados que se obtienen en esta investigación. Además, en otra investigación similar, Duarte y otros (2021) publican el artículo bajo el tema “Factores asociados al rendimiento académico en estudiantes de enfermería” en el cual se demuestra que la variable incidente en el rendimiento académico es la motivación; donde el estado civil también es influyente debido a las ocupaciones familiares que este conlleva, como en el caso de Unión Libre que con un $r = 0,7524$ y un $r^2 = 0,5661$ afecta en un 4,5 al nivel de desempeño (Duarte, y otros, 2021).

CONCLUSIONES Y RECOMENDACIONES

5.1 Conclusiones

- Se logró identificar las variables académicas, físicas y actitudinales mediante el estudio del arte en investigaciones previas realizadas por la comunidad científica mediante una revisión literaria y, permitió determinar las causas del bajo rendimiento académico; estos factores están asociados a las calificaciones, entre estos se tiene: la familia, al estado civil, nivel económico y académico de los progenitores.
- Se determinó las variables socioeconómicas mediante una investigación documental a los estudiantes del primer ciclo del Instituto Superior Tecnológico “Tres de Marzo de la Provincia Bolívar”, y se estableció las variables involucradas a la deserción escolar aplicando regresión logística multivariante, dando como resultado que el estado civil del estudiante, ingresos totales del hogar, cantidad de miembros en el hogar y nivel académico de los padres, influyen en la probabilidad de que los estudiantes tengan un bajo rendimiento académico
- Se desarrolló un modelo matemático predictivo de regresión logística multivariante utilizando el software estadístico R, para evaluar el desempeño de este; el mismo que predice el 84.9% de los casos posibles para lo cual, considera las variables más significativas mediante la razón de momios como el nivel de formación de la madre que afecta en gran manera ya que la razón odd es muy alta, vivir dentro de una familia con más de 9 miembros aumenta hasta 19 veces la probabilidad de reprobar y, hasta el doble la variable estado civil unión libre que representa el 200,12%.

5.2 Recomendaciones

- Una vez reconocido los factores que influyen en gran manera en la deserción escolar, se debe considerar tomar guías o rutas que prevenga la existencia de esto o evitar el gran impacto en los estudiantes, siendo así que se evite tener niveles exagerados de bajo rendimiento académico que provoque reprobar el ciclo escolar en el Instituto. Como los valores factores asociados a las calificaciones fueron la familia, al estado civil, nivel económico y académico de los progenitores; se recomienda trabar en propuestas psicológicas, charlas motivacionales, programas para padres para el buen vivir y, un plan micro curricular enfocado al trabajo entre pares y al emprendimiento para hacer frente al problema.
- Se debe considerar utilizar el modelo matemático para futuras investigaciones y también para evaluar constantemente el impacto de estos factores en la vida escolar e intentar erradicar estas situaciones bajo avance académico. Esta investigación maneja datos del contexto de los estudiantes del primer ciclo del Instituto Superior Tecnológico “Tres de Marzo de la Provincia Bolívar” por lo que no se debe usar los valores odds ratios para investigaciones similares ya que cada contexto maneja datos diferentes y por ende los coeficientes del modelo también.
- Los datos obtenidos se deben poner en consideración al Instituto Superior Tecnológico Tres de Marzo de la Provincia Bolívar”, para que sea aplicable en otras instituciones para predecir problemáticas académicas y sociales como la reprobación o deserción de estudios que estancan el avance escolar de las personas. A pesar de que el modelo Logit desarrollado predice el 84.9%, se recomienda trabajar en otros modelos usando Software para la

libre distribución de información para una mayor predicción y mejora de los procesos de corrección del problema.

Bibliografía

- Acuña, C. (2018). Deserción escolar en educación a distancia: datos de estudio. *Revista Vinculando*. Obtenido de <https://vinculando.org/educacion/desercion-escolar-en-educacion-a-distancia-datos-de-estudio.html>
- Aedo, S., Pavlov, S., & Clavero, F. (2010). Riesgo relativo y Odds ratio ¿Qué son y cómo se interpretan? *NOTA ESTADÍSTICA*, 51-54.
- Alvarado, L. (2 de 10 de 2000). Problemas familiares y rendimiento académico. *Problemas familiares*.
- Alvarado, L. (2 de 10 de 2020). Problemas familiares y rendimiento académico.
- Arismendy, C., & Morales, N. (2018). *Modelo de regresión logística como alternativa para medir la probabilidad de deserción temprana en la Universidad de los Llanos periodo 2015-2 – 2018-1*. Villavicencio: FACULTAD DE CIENCIAS ECONÓMICAS.
- Armas Sopalo, S. M. (2020). Análisis de los factores que inciden en la deserción estudiantil de los niños, niñas y adolescentes ecuatorianos en el periodo 2009-2019. *Bibdigital*, 64.
- Barrientos, Z. (2010). *Deserción estudiantil en posgrados semipresenciales de la Universidad . costa rica: Universidad Estatal a Distancia*.
- Bernal, C. (2010). *Metodología de la investigación* (Tercera ed.). Colombia: Pearson Educación. Obtenido de <https://ceduna.jimdofree.com/app/download/13872381378/Bernal+Cap%C>

%20DE%20INVESTIGACI%C3%93N%20S%C3%81NCHEZ%20ALLY
SON.pdf

Cerda, J., Vera, C., & Rada, G. (2013). Odds ratio: aspectos teóricos y prácticos. *Revista médica Chile, 141*, 1329 - 1335. Retrieved from <https://scielo.conicyt.cl/pdf/rmc/v141n10/art14.pdf>

Cerpa, W. F. (2015). Analisis multivariado para determinar los factores mas relevantes de desrcion estudiantil presentes en el programa de ingenieria industrial de una Universidad del Caribe Colombia. *Dialnet*, 86-98. Obtenido de ojs.uac.edu.co/index.php/prospectiva/article/download/363/328

Chaparro, A., González, C., & Caso, J. (2016). Familia y rendimiento académico: configuración de perfiles estudiantiles en secundaria. *Scielo*, 53-68.

Constitución del Ecuador. (2013). *Ministerio de educación*. Obtenido de Constitución de la República del Ecuador: https://educacion.gob.ec/wp-content/uploads/downloads/2014/01/TRANSP-NORMAS_CONSTITUCIONALES.pdf#:~:text=para%20sus%20habitantes.-,Art.,indispensable%20para%20el%20buen%20vivir.

Cuji, B., Gavilanes, W., & Sanchez, R. (2017). Modelo predictivo de deserción estudiantil basado en arboles de decisión. *Revista Espacios*, 17.

Díaz, C. J. (2009). *Factores de Deserción Estudiantil en Ingeniería: Una*. Chile: Universidad Católica de la Santísima Concepción.

Díaz, C. J. (2009). *Factores de Deserción Estudiantil en Ingeniería: Una . chile :* Universidad Católica de la Santísima Concepción.

Díaz, V., & Calzadilla, A. (2016). Artículos científicos, tipos de investigación y productividad científica en las Ciencias de la Salud. *Scielo*, 115–121.

dssd. (2014). *Redalyc*.

Duarte, C., Alvarado, I., Suarez, J., Madriz, D., Nieto, Z., Moreno, Y., . . . Bravo, A. (2021). Factores asociados al rendimiento académico en estudiantes de enfermería. *Redalyc*, 927-933.

El Mercurio. (2 de Octubre de 2020). *Problemas familiares y rendimiento académico*, págs. 1-2.

el, F. q. (2016). *Mary Nieves Cruz Zuluaga*. medellin : Institución Universitaria Esumer.

Elkin Castaño, S. (2008). *Análisis de los factores asociados a la deserción estudiantil*. Colombia: Centro de Investigaciones Económicas (.

Fernández, T., Solís, M., Hernández, M., & Moreira, T. (2019). Un análisis multinomial y predictivo de los factores asociados a la deserción universitaria. *Educare Electronic Journal*, 1-25.

Ferre, M. (4 de Abril de 2019). *Universidad de Murcia*. Obtenido de Modelos de Regresión:

https://gauss.inf.um.es/feir/40/#122_la_correlaci%C3%B3n_simple_en_r

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using r* . Sage Publications Ltd.

Fierros, R., Gómez, L., & Sánchez , M. (2014). Factores que determinan la Deserción Escolar en la Licenciatura de Trabajo Social de la Universidad de Guadalajara en los calendarios escolares A y B del 2012. *Universidad de Guadalajara*. Obtenido de <https://www.ecorfan.org/handbooks/Educacion%20T-IV/ARTICULO%2010.pdf>

Figueredo, A., León, R., & Martínez, M. (2019). Procedimiento para el procesamiento de información científica en la DPI de la carrera Ingeniería Forestal. *Journal of Librarianship and Information Science*, 46-61.

FUENTES, C. A. (2015). *MODELO DE REGRESIÓN LOGÍSTICA COMO ALTERNATIVA PARA MEDIR LA* . colombia: UNIVERSIDAD DE LOS LLANOS.

Goldenhersch, H., Coria, A., & Saino, M. (2011). DESERCIÓN ESTUDIANTIL: UNA FORMA DE ABORDAJE DESDE LA UNIVERSIDAD PÚBLICA. Obtenido de <https://core.ac.uk/download/pdf/234020259.pdf>

González , M., & Maranto, M. (2015). *Fuentes de Información*. Obtenido de Universidad autónoma del estado de Hidalgo: <https://repository.uaeh.edu.mx/bitstream/bitstream/handle/123456789/16700/LECT132.pdf>

González, E. (2017). *Factores que inciden en el rendimiento académico de los estudiantes de la Universidad Politécnica del Valle de Toluca*. Toluca, Mexico: Universidad Politécnica del Valle de Toluca.

Guerra, E. D. (2015). *Repositorio UNIVERSIDAD PEDAGÓGICA NACIONAL*. Obtenido de <http://repositorio.pedagogica.edu.co/bitstream/handle/20.500.12209/911/T0-17863.pdf?sequence=1&isAllowed=y>

Guevara, G., Verdesoto, A., & Castro, N. (2020). Metodologías de investigación educativa (descriptivas, experimentales, participativas, y de investigación-acción). *RECIMUNDO*, 4(3), 163 - 173. Retrieved from <https://dialnet.unirioja.es/descarga/articulo/7591592.pdf>

Hernandez, R., Fernández, C., & Baptista, P. (1997). *Metodología de la Investigación*. México: Mc Graw Hill. Retrieved from <https://www.studocu.com/cl/document/universidad-de-aconcagua/proyecto-de-titulo/1-4-hernandez-sampieri-metodologia-de-la-investigacion/8226374>

Hernández, S., & Avila, D. (2020). Técnicas e instrumentos de recolección de datos. *Boletín Científico de las Ciencias Económico Administrativas del ICEA*, 9(17), 51 - 53. Obtenido de <https://repository.uaeh.edu.mx/revistas/index.php/icea/issue/archive>

Instituto Nacional de Estadísticas y Geografía (México). (2012). *Procesamiento de la información*. INEGI. Obtenido de

https://www.snieg.mx/documentacionportal/normatividad/vigente/doctos_genbasica/procesamiento_informacion.pdf

Investigadores. (4 de Marzo de 2020). *Investigación cuantitativa*. Obtenido de Técnicas de Investigación: <https://tecnicasdeinvestigacion.com/investigacion-cuantitativa/>

Jano, D., & Ortiz, S. (2006). *Los factores que afectan al rendimiento académico en la educación superior*. Madrid: Universidad Autónoma.

Loaiza, W., & Rojas, A. (2015). *Rendimiento Académico en Matemáticas un estudio vía Regresión Logística*. Duitama: UNIVERSIDAD PEDAGÓGICA Y TECNOLÓGICA DE COLOMBIA.

López, C. N. (2021). *Diseño de un modelo matemático para estimar la deserción estudiantil*. ambato: ACULTAD DE INGENIERÍA EN SISTEMAS, ELECTRÓNICA .

Martinez De Ibarreta, C. R. (2022). Influencia del Nivel Educativo de los Padres en el Rendimiento Académico de los Estudiantes de ADE. *Un Enfoque de Género*.

Martínez, B. H. (2015). Causas del bajo rendimiento escolar origina un alto nivel de deserción escolar y habilidades para estudiar ayudan a mejorar el rendimiento escolar.

Martínez, J., Ferrás, Y., Bermúdez, L., Cabrera, Y., & Pérez, E. (2020). Rendimiento académico en estudiantes Vs factores que influyen en sus resultados: una relación a considerar. *Scielo*, 105-121.

- Máxima Formación. (21 de 09 de 2022). *¿QUÉ ES LA CORRELACIÓN ESTADÍSTICA Y CÓMO INTERPRETARLA?* Obtenido de Máxima formación: <https://www.maximaformacion.es/blog-dat/que-es-la-correlacion-estadistica-y-como-interpretarla/>
- Medina, E. (2003). *Modelos de elección discreta*. Obtenido de <https://www.cartagena99.com/recursos/alumnos/apuntes/logit.pdf>
- Meybelin, J. (2018). *Análisis del rendimiento académico*. colombia: REVISTA ELECTRÓNICA DE CONOCIMIENTOS, SABERES Y PRÁCTICAS.
- Minitab. (20 de 09 de 2022). *¿Qué son variables de respuesta y predictoras?* . Obtenido de Minitab.com: <https://support.minitab.com/es-mx/minitab/21/help-and-how-to/statistical-modeling/regression/supporting-topics/basics/what-are-response-and-predictor-variables/>
- Morocho Valarezo, K. M., & García Bustos, S. (2020). Modelo de regresión logística de la deserción estudiantil en un instituto tecnológico en el cantón daule (Doctoral dissertation, ESPOL. FCNM). Obtenido de https://scholar.google.es/scholar?hl=es&as_sdt=0%2C5&q=la+desercion+estudiantil+con+regresion+logistica+aplicado+en+r+studio&btnG=
- Obregón, J. M. (2018). *Análisis del rendimiento académico* . colombia: REVISTA ELECTRÓNICA DE CONOCIMIENTOS, SABERES Y PRÁCTICAS.
- Pachay, M., & Rodríguez, M. (2021). La deserción escolar: Una perspectiva compleja en tiempos de pandemia. *Polo del conocimiento*, 6(54), 130 - 155.

Obtenido de
<https://polodelconocimiento.com/ojs/index.php/es/article/view/2129/4239#>

Pachay-López, M., & Rodríguez-Gómez, M. (2021). Abandono escolar: Una perspectiva compleja en tiempos de pandemia. *Polo del Conocimiento*, 50-62.

Palacios, K. R. (2021). *Comparativa de modelos de clasificación para inferir la probabilidad de deserción estudiantil en la Facultad de Ciencias Químicas de la Universidad de Cuenca*. Cuenca: Facultad de Ciencias Químicas de la Universidad.

Paladino, M. (19 de 09 de 2022). *Modelos logit con R*. Obtenido de [Institutomora.edu.mx](https://www.institutomora.edu.mx):
https://www.institutomora.edu.mx/testU/SitePages/martinpaladino/modelos_logit_con_R.html

Paramo, G., & Correa, C. (1999). *Desercion Estudiantil Universitaria*. Madrid: Universidad Eafid.

Pérez, A., Escobar, C., Ruiz, M., Benites, L., & Reye, G. (2018). Modelo de predicción de la deserción estudiantil de primer año en la Universidad Bernardo O'Higgins. *Scielo*.

Pérez, C., Cerón, E., Suárez, R., Mera, M., Briones, N., Zambrano, L., & Barreto, M. (2017). Deserción y repitencia en estudiantes de la carrera de Enfermería matriculados en el período 2010-2015. *Elsevier*, 84-90.

- Prieto, C. (2015). *Uso de Regresión Logística para predecir deserción estudiantil temprana*. Bogotá: Departamento de Ingeniería de Sistemas y Computación.
- Ramírez, M., Casas, V., Téllez, L., & Arroyo, A. (2015). Deserción escolar y menor infractor. *Revista de Psicología y Ciencias del Comportamiento de la U.A.C.J.S.*, 6(1), 1 - 32. Retrieved from <http://www.scielo.org.mx/pdf/rpcc/v6n1/2007-1833-rpcc-6-01-1.pdf>
- Reyes, J., Escobar, C., & Duarte, J. (2017). UNA APLICACION DEL MODELO DE REGRESION LOGISTICA EN LA PREDICCION DEL RENDIMIENTO ESTUDIANTIL. *Revista Uach*, 101-120.
- Rivero, R. (19 de Abril de 2020). *Metodología de la Investigación*. Obtenido de Instituto de Hematología e Inmunología: <https://instituciones.sld.cu/ihi/metodologia-de-la-investigacion/>
- Rocabado, J. R. (2014). *UNA APLICACION DEL MODELO DE REGRESION LOGISTICA*. Chile : Universidad de Antofagasta, Departamento de Matemáticas, Avenida Angamos.
- Rochin , F. (2021). Deserción escolar en la educación superior en México: revisión de literatura. *RIDE*, 12(22). Obtenido de <http://www.scielo.org.mx/pdf/ride/v11n22/2007-7467-ride-11-22-e01.pdf>
- Rodó, P. (21 de 09 de 2022). *Valor-p*. Obtenido de Economipedia: <https://economipedia.com/definiciones/valor-p.html>
- Roldán, P. (19 de 09 de 2022). *Modelo matemático*. Obtenido de economipedia.com: <https://economipedia.com/definiciones/modelo->

matematico.html#:~:text=Un%20modelo%20matem%C3%A1tico%20es%
20una,%E2%80%9Cteor%C3%ADa%20de%20los%20modelos%E2%80
%9D.

Ruiz, M., García, R., Luis, J., & Ximhai, R. (2014). CAUSAS Y CONSECUENCIAS DE LA DESERCIÓN ESCOLAR EN EL BACHILLERATO: CASO UNIVERSIDAD AUTÓNOMA DE SINALOA. *Redalyc*, 51-57.

Ruíz, R., Ayala, M., & Zapata, E. (2014). Estereotipos de género en la deserción escolar, caso El Fuerte, Sinaloa. *Ra Ximha*, 10(7), 165 - 184. Obtenido de <https://www.redalyc.org/pdf/461/46132451012.pdf>

Ruiz, R., García, J., & Pérez, M. (2014). CAUSAS Y Consecuencias de la deserción escolar en el bachillerato: Caso universidad autónoma de Sinaloa. *Ra Ximha*, 10(5), 51 - 74. Retrieved from <https://www.redalyc.org/pdf/461/46132134004.pdf>

Sanchez, M., Fernández, M., & Díaz, J. (2021). Técnicas e instrumentos de recolección de información: análisis y procesamiento realizado por el investigador cualitativo. *Revista Científica UISRAEL*, 107-121.

Senescyt. (2018). *Guia de registro de institutos públicos y particulares*. Quito: Secretaría de Educación Superior, Ciencia, Tecnología e Innovación del Ecuador.

Senescyt. (20 de Marzo de 2020). *Secretaría de Educación Superior, Ciencia, Tecnología e Innovación*. Obtenido de

<https://www.educacionsuperior.gob.ec/informacion-estadistica-sobre-educacion-superior-ciencia-tecnologia-e-innovacion/>

Tapasco Alzate, O. A., Ruiz Ortega, F. J., Osorio García, D., & Ramírez Ramírez, D. (2019). Deserción estudiantil: incidencia de factores institucionales relacionados con los procesos de admisión. . *Educación y Educadores*.

Tapasco Alzate, O. A., Ruiz Ortega, F. J., Osorio García, D., & Ramírez Ramírez, D. (2019). Deserción estudiantil: incidencia de factores institucionales relacionados con los procesos de admisión. *Educación y Educadores*.

Tapasco, O., Ruiz, F., Osorio, D., & Ramírez, D. (2019). Deserción estudiantil: incidencia de factores institucionales relacionados con los procesos de admisión. *Educación y Educadores*, 81-100.

Tuñón, I., & Poy, S. (2016). Factores asociados a las calificaciones escolares como proxy del rendimiento educativo. *Revista Electrónica de Investigación Educativa*, 98-111.

Universidad Tecnológica de Pereira Facultad de Ciencias de la Salud. (Marzo de 2017). *Universidad Tecnológica de Pereira*. Obtenido de <https://academia.utp.edu.co/seminario-investigacion-II/files/2017/03/06a.An%C3%A1lisisDeCorrelaciones.pdf>

Urrego, A. I. (2015). Influencia de los factores personales y académicos sobre la deserción de estudiantes en la educación a distancia. *Revista de Investigación UNAD* , 1-19.

- Valera, J., Sinha, S., Varela, J., & Ponsot Balaguer, E. (2009). Una explicación del rendimiento estudiantil universitario mediante modelos de regresión. *Redalyc*, 415-427.
- Velez, A., Meerbeke, v., & Gonzales, N. (2005). *Factores asociados al rendimiento académico en estudiantes de medicina*. Bogotá, Colombia: Facultad de Medicina Universidad del Rosario.
- Vereau, E. V. (2012). *ANÁLISIS COMPARATIVO DE MODELOS DE CLASIFICACIÓN EN EL*. lima: Revista digital de la Facultad de Ingeniería de Sistemas.
- Vergaray, K. R. (2021). Modelo predictivo para la detección temprana de estudiantes con alto riesgo de deserción académica. Obtenido de https://scholar.google.es/scholar?hl=es&as_sdt=0%2C5&q=la+desrcion+e+estudiantil+aplicado+la+regresion+logistica+&btnG=
- Villa, M. S., & Omar, B. A. (2010). ADAPTACIÓN DEL MODELO DE DESERCIÓN UNIVERSITARIA DE TINTO A LA UNIVERSIDAD CATÓLICA DE LA SANTÍSIMA CONCEPCIÓN, CHILE (2010).
- Vinueza, C. (2021). *Diseño de un modelo matemático para estimar la deserción estudiantil mediante técnicas de análisis multivariado en una institución de educación superior tecnológica*. Ambato: FACULTAD DE INGENIERÍA EN SISTEMAS, ELECTRÓNICA E INDUSTRIAL.
- Viteri, D., & Uquillas, M. (2011). *Estudio sobre la deserción estudiantil en la Pontificia Universidad Católica del Ecuador - Matriz, en los niveles Iro,*

2do y 3ero de todas las Facultades y Escuelas del primer semestre del año académico 2007-2008. Quito: Facultad de Psicología.

Vizuite, R. (2017). *Deserción Escolar de los estudiantes del colegio técnico Vicente Rocafuerte de la ciudad de Quito de la especialidad de mecánica automotriz y la repercusión en el ámbito laboral (año 2017-2017)*. Quito: Universidad Central del Ecuador. Obtenido de <http://www.dspace.uce.edu.ec/bitstream/25000/14152/1/T-UCE-0010-ET002-2018-.pdf>

Vries, W., Leon, P., Romero, J., & Hernandez, I. (2011). *¿Desertores o decepcionados? Distintas causas para abandonar los estudios universitarios*. *Scielo*.

ANEXOS

Anexo 1

Modelación matemática en R Commander.

R Commander
Modelo lineal generalizado

Introducir un nombre para el modelo: GLM2

Variables (doble clic para enviar a la fórmula)

- Calificación [factor]
- estadoocivilid [factor]
- ingresoTotalHogar [factor]
- nivelFormacionMadre [factor]
- nivelFormacionPadre [factor]

Fórmula del modelo

Operadores (pulsar para formular): + * / %in% - ^ ()

Splines/Polinomiales (seleccionar variable y pulsar)

- B-spline
- spline natural
- polinomio ortogonal
- polinomio crudo
- g. de los splines: 5
- g. de polinomios: 2

Expresión de selección: < todos los casos válidos >

Weights: < ninguna variable sele >

Familia (doble clic para seleccionar)

- gaussian
- binomial
- poisson
- Gamma
- inverse gaussian
- quasibinomial
- quasipoisson

Función de enlace

- logit
- probit
- loglog

Ejecutar

R Commander

Conjunto de datos: Dataset | Editar conjunto de datos | Visualizar conjunto de datos | Modelo: GLM2

R Script | R Markdown

```
Dataset <- readXL("D:/TESIS GUANABUO/MATRII/base_de_datos4.xlsx",  
rownames=FALSE, header=TRUE, na="", sheet="deparados",  
stringsAsFactors=TRUE)  
GLM.1 <- glm(Calificación ~ cantidadMiembrosHogar + estadoocivilid + ingresoTotalHogar + nivelFormacionPadre, family="binomial", data=Dataset)  
summary(GLM.1)  
exp(coef(GLM.1)) # Exponentiated coefficients ("odds ratios")  
Dataset <- readXL("D:/TESIS GUANABUO/MATRII/base_datos3.xlsx", rownames=FALSE, header=TRUE, na="", sheet="deparados", stringsAsFactors=TRUE)  
GLM.2 <- glm(Calificación ~ cantidadMiembrosHogar + estadoocivilid + ingresoTotalHogar + nivelFormacionMadre + nivelFormacionPadre, family="binomial", data=Dataset)  
summary(GLM.2)  
exp(coef(GLM.2)) # Exponentiated coefficients ("odds ratios")
```

Salida

cantidadMiembrosHogar [T.MAS DE 9]	2.9760	0.9694	3.070	0.00214 **
estadoocivilid [T.DIVORCIADO/A]	-17.0149	2739.4664	-0.006	0.99804
estadoocivilid [T.SOLTERO/A]	-1.3454	0.5265	-2.854	0.01065 *
estadoocivilid [T.UNION LIBRE]	0.7538	0.8189	0.920	0.35734
ingresoTotalHogar [T.ENTRE \$400-\$899]	-0.5342	0.8159	-0.655	0.51264
ingresoTotalHogar [T.MAYOR A \$899]	0.7866	0.9033	0.838	0.40225
ingresoTotalHogar [T.MENOR A \$300]	0.8907	0.2934	2.013	0.04411 *
nivelFormacionMadre [T.EDUCACION BASICA]	15.9411	1457.5418	0.011	0.99127
nivelFormacionMadre [T.EDUCACION MEDIA]	17.1169	1457.5424	0.012	0.99061
nivelFormacionMadre [T.JARDIN]	16.7099	1457.5418	0.011	0.99085
nivelFormacionMadre [T.NINGUNO]	-0.7615	1670.6708	0.000	0.99964
nivelFormacionMadre [T.POST-GRADO]	15.9270	1457.5420	0.011	0.99129
nivelFormacionMadre [T.PRIMARIA]	17.5699	1457.5418	0.012	0.99066
nivelFormacionMadre [T.SECUNDARIA]	16.3220	1457.5421	0.011	0.99107
nivelFormacionMadre [T.SUPERIOR NO UNIVERSITARIO]	16.7133	1457.5420	0.011	0.99085
nivelFormacionPadre [T.SUPERIOR UNIVERSITARIO]	1.1523	4216.1957	0.009	0.99978
nivelFormacionPadre [T.EDUCACION BASICA]	-0.8301	1.0361	-0.801	0.42304
nivelFormacionPadre [T.EDUCACION MEDIA]	-16.1522	3956.1805	-0.004	0.99674
nivelFormacionPadre [T.JARDIN]	-1.1701	0.9699	-1.206	0.22767
nivelFormacionPadre [T.NINGUNO]	-2.1990	1.5063	-1.460	0.14432
nivelFormacionPadre [T.POST-GRADO]	-0.4371	1.0819	-0.404	0.68620

Mensajes

[5] NOTA: El conjunto de datos Dataset tiene 489 filas y 6 columnas.
[6] AVISO: Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

R Commander

Conjunto de datos: Dataset | Editar conjunto de datos | Visualizar conjunto de datos | Modelo: GLM2

R Script | R Markdown

```
Dataset <- readXL("D:/TESIS GUANABUO/MATRII/base_de_datos4.xlsx",  
rownames=FALSE, header=TRUE, na="", sheet="deparados",  
stringsAsFactors=TRUE)  
GLM.1 <- glm(Calificación ~ cantidadMiembrosHogar + estadoocivilid + ingresoTotalHogar + nivelFormacionPadre, family="binomial", data=Dataset)  
summary(GLM.1)  
exp(coef(GLM.1)) # Exponentiated coefficients ("odds ratios")  
Dataset <- readXL("D:/TESIS GUANABUO/MATRII/base_datos3.xlsx", rownames=FALSE, header=TRUE, na="", sheet="deparados", stringsAsFactors=TRUE)  
GLM.2 <- glm(Calificación ~ cantidadMiembrosHogar + estadoocivilid + ingresoTotalHogar + nivelFormacionMadre + nivelFormacionPadre, family="binomial", data=Dataset)  
summary(GLM.2)  
exp(coef(GLM.2)) # Exponentiated coefficients ("odds ratios")
```

Salida

```
> exp(coef(GLM.2)) # Exponentiated coefficients ("odds ratios")  
(Intercept)                cantidadMiembrosHogar [T.ENTRE 4-6]          cantidadMiembrosHogar [T.ENTRE 7-9]  
      8.815239e+08                1.268216e+00                1.171196e+00  
cantidadMiembrosHogar [T.MAS DE 9]          estadoocivilid [T.DIVORCIADO/A]          estadoocivilid [T.SOLTERO/A]  
      1.960964e+01                4.08112e-09                2.604237e-01  
estadoocivilid [T.UNION LIBRE]          ingresoTotalHogar [T.ENTRE $400-$899]          ingresoTotalHogar [T.MAYOR A $899]  
      2.121898e+00                1.861464e-01                2.131003e+00  
ingresoTotalHogar [T.MENOR A $300]          nivelFormacionMadre [T.EDUCACION BASICA]          nivelFormacionMadre [T.EDUCACION MEDIA]  
      1.805220e+00                3.377475e+06                4.050349e+07  
      1.807316e+07                nivelFormacionMadre [T.NINGUNO]          nivelFormacionMadre [T.POST-GRADO]  
      4.669648e-01                8.178335e+06  
nivelFormacionMadre [T.PRIMARIA]          nivelFormacionMadre [T.SECUNDARIA]          nivelFormacionMadre [T.SUPERIOR NO UNIVERSITARIO]  
      2.590334e+07                1.226155e+07                1.213390e+07  
nivelFormacionMadre [T.SUPERIOR UNIVERSITARIO]          nivelFormacionPadre [T.EDUCACION BASICA]          nivelFormacionPadre [T.EDUCACION MEDIA]  
      3.165441e+00                4.360067e-01                9.664929e-09  
      nivelFormacionPadre [T.JARDIN]          nivelFormacionPadre [T.NINGUNO]          nivelFormacionPadre [T.POST-GRADO]  
      3.103431e-01                1.1109091e-01                4.458931e-01  
      nivelFormacionPadre [T.PRIMARIA]          nivelFormacionPadre [T.SECUNDARIA]          nivelFormacionPadre [T.SUPERIOR NO UNIVERSITARIO]  
      9.583945e-02                6.724919e-08                1.998076e-01
```

Mensajes

[5] NOTA: El conjunto de datos Dataset tiene 489 filas y 6 columnas.
[6] AVISO: Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred