



ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO

**Diseño de un modelo matemático enfocado a la reducción del
abandono estudiantil a través de un análisis multivariante, para la
Universidad Regional Amazónica Ikiam**

ROBERTO CARLOS AYABACA LARA

**Trabajo de Titulación modalidad Proyectos de Investigación y Desarrollo, presentado ante
el Instituto de Posgrado y Educación Continua de la ESPOCH, como requisito parcial para
la obtención del grado de**

MAGÍSTER EN MATEMÁTICA MENCIÓN MODELACIÓN Y DOCENCIA

RIOBAMBA - ECUADOR

NOVIEMBRE 2022

©2022, Roberto Carlos Ayabaca Lara

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo a la cita bibliográfica del documento, siempre y cuando se reconozca el derecho de autor.



ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO

EL TRIBUNAL DE TRABAJO DE TITULACIÓN CERTIFICA QUE:

El Trabajo de Titulación modalidad Proyectos de Investigación y Desarrollo, titulado Diseño de un modelo matemático enfocado a la reducción del abandono estudiantil a través de un análisis multivariante, para la Universidad Regional Amazónica Ikiam, de responsabilidad del Sr. Roberto Carlos Ayabaca Lara ha sido prolijamente revisado y se autoriza su presentación.

Dra. Silvia Mariana Haro Rivera; Mag.
PRESIDENTE



Firmado electrónicamente por:
**SILVIA
MARIANA HARO**

Ing. Christian Fabian Alcocer Castillo; Mag.
DIRECTOR



Firmado electrónicamente por:
**CHRISTIAN FABIAN
ALCOECER CASTILLO**

Lic. Jorge Manuel Ríos Obregón; Ph.D.
MIEMBRO



Firmado electrónicamente por:
**JORGE MANUEL
RIOS OBREGON**

Lic. Willin Gabriel Álvarez Irausquin; Ph.D.
MIEMBRO



Firmado electrónicamente por:
**WILLIN GABRIEL
ALVAREZ
IRAUSQUIN**

Riobamba, noviembre 2022

DECLARACIÓN DE AUTENCIDAD Y CESIÓN DE DERECHOS DE AUTOR

Yo, ROBERTO CARLOS AYABACA LARA, declaro que el presente trabajo de titulación es de mi autoría y que los resultados del mismo son auténticos. Los textos en el documento que provienen de otras fuentes están debidamente citados y referenciados.

Como autor asumo la responsabilidad legal y académica de los contenidos de este trabajo de titulación. El patrimonio intelectual pertenece a la Escuela Superior Politécnica de Chimborazo.

ROBERTO CARLOS AYABACA LARA

1600461808

DEDICATORIA

A mi familia, compañeros y conocidos que me han impulsado y colaborado en que el presente trabajo sea técnico y ético son sentido lógico, convirtiéndose en el punto de partida para futuras investigaciones.

A mis padres, por haber confiado en mis capacidades, por su apoyo incondicional y permanente, convirtiéndose en un pilar fundamental para lograr cada objetivo que me he propuesto que en un inicio parecían inalcanzables.

AGRADECIMIENTO

Quiero expresar mi agradecimiento a mi familia que ha estado en esos momentos difíciles y supieron brindarme su apoyo de diferentes maneras. Un agradecimiento especial, a mis padres que fueron los que creyeron en mí cuando dudaba de mis capacidades y ser un apoyo incondicional durante todo este proceso.

A mi tutor y miembros, quienes me han guiado de manera impecable durante la realización de este trabajo de investigación. Agradecer, de la misma manera a las diferentes coordinadoras que me han apoyado desde el inicio hasta el término de este programa de maestría.

Al personal que conforma el Instituto de Posgrado y Educación Continua de la Escuela Superior Politécnica de Chimborazo por su buena predisposición en la enseñanza, por la paciencia y guía en los diferentes procesos y por ayudarnos a crecer académicamente en este hermoso campo de la Matemática.

TABLA DE CONTENIDO

| | |
|--|-----------|
| RESUMEN..... | xiv |
| ABSTRACT..... | xv |
| | |
| CAPÍTULO I | 1 |
| | |
| 1. INTRODUCCIÓN | 1 |
| 1.1. Planteamiento Del Problema..... | 1 |
| <i>1.1.1. Situación problemática</i> | <i>2</i> |
| <i>1.1.2. Formulación Del Problema</i> | <i>4</i> |
| <i>1.1.3. Preguntas Directrices.....</i> | <i>5</i> |
| 1.1. Justificación de la Investigación | 5 |
| <i>1.1.2. Justificación metodológica</i> | <i>6</i> |
| <i>1.2.1. Objetivo General.....</i> | <i>6</i> |
| <i>1.2.2. Objetivos específicos.....</i> | <i>7</i> |
| | |
| CAPÍTULO II | 8 |
| | |
| 2. MARCO TEÓRICO | 8 |
| 2.1. Estado del Arte | 8 |
| 2.2. Bases Teóricas | 11 |
| <i>2.2.1. La Educación Superior en el Ecuador</i> | <i>11</i> |
| <i>2.2.2. La Deserción Estudiantil en el Ecuador</i> | <i>13</i> |
| <i>2.2.3. La Deserción desde la perspectiva Institucional.....</i> | <i>14</i> |
| 2.2.6. Modelos de Análisis de la Deserción Estudiantil..... | 15 |
| <i>2.2.6.1. Modelos Psicológicos.....</i> | <i>15</i> |
| <i>2.2.6.2. Modelos Sociológicos.....</i> | <i>16</i> |
| <i>2.2.6.3. Modelos Económicos</i> | <i>16</i> |
| <i>2.2.6.4. Modelo Organizacional</i> | <i>16</i> |
| <i>2.2.6.5. Modelo de Interacción</i> | <i>16</i> |
| 2.2.7. Tipos de Variables..... | 17 |
| <i>2.2.7.1. Variables Cualitativas.....</i> | <i>17</i> |
| <i>2.2.7.2. Variables Cuantitativas.....</i> | <i>17</i> |
| 2.2.8. Técnicas de Análisis Multivariante de Datos | 17 |
| 2.2.9. Técnicas de Análisis de Dependencia..... | 19 |

| | | |
|-------------------------|--|-----------|
| 2.2.10. | Regresión Múltiple | 20 |
| 2.2.10.1. | <i>Consideraciones dentro del Análisis de Regresión Múltiple</i> | 21 |
| 2.2.11. | Técnicas de Análisis de Interdependencia | 22 |
| 2.2.12. | Análisis de Componentes Principales | 22 |
| 2.2.12.1. | <i>Supuestos del Análisis</i> | 22 |
| 2.2.13. | Modelo Matemático | 23 |
| CAPÍTULO III | | 24 |
| 3. | METODOLOGÍA DE LA INVESTIGACIÓN | 24 |
| 3.1. | Tipos de Investigación | 24 |
| 3.1.1. | <i>Tipos de investigación según el nivel</i> | 24 |
| 3.1.2. | <i>Tipos de Investigación según el diseño</i> | 25 |
| 3.1.3. | <i>Tipos de Investigación según el propósito</i> | 25 |
| 3.2.1. | Pasos de la Investigación Científica | 26 |
| 3.2.1.1. | <i>Recolección y análisis de datos</i> | 26 |
| 3.2.1.2. | <i>Propuesta del Modelo Matemático</i> | 26 |
| 3.2.1.3. | <i>Validar el Modelo Matemático</i> | 27 |
| 3.2.1.4. | <i>Optimizar el Modelo Matemático</i> | 27 |
| 3.5. | Selección del Software Estadístico | 28 |
| 3.6. | Recolección de Datos | 29 |
| 3.6.1. | <i>Población de estudio</i> | 29 |
| 3.6.2. | <i>Unidad de análisis</i> | 30 |
| 3.6.3. | <i>Selección de la muestra</i> | 30 |
| 3.7. | Medidas Estadísticas | 30 |
| 3.8. | Identificación de variables | 30 |
| 3.9. | Operacionalización De Variables | 31 |
| 3.10. | Adecuación de la información | 34 |
| CAPÍTULO IV | | 35 |
| 4.1. | RESULTADOS Y ANÁLISIS | 35 |
| 4.1. | Adecuación de la Información | 40 |
| 4.1.1. | Componentes factoriales | 40 |
| 4.1.1.1. | <i>Prueba de esfericidad de Bartlett</i> | 44 |
| 4.1.1.2. | <i>Índice KMO de Kaiser-Meyer- Olkin</i> | 45 |
| 4.1.1.3. | <i>Valores Propios</i> | 47 |

| | |
|--|-----------|
| 4.1.1.4. <i>Matriz de componentes</i> | 49 |
| 4.1.1.5. <i>La Comunalidad</i> | 52 |
| 4.1.1.6. <i>Rotación de los ejes</i> | 53 |
| 4.1.1.7. <i>Identidad de ejes</i> | 53 |
| 4.1.2. <i>Análisis de Regresión Lineal Múltiple</i> | 55 |
| 4.1.2.1. <i>Obtención de los coeficientes parciales de regresión</i> | 56 |
| 4.1.2.2. <i>Coefficiente de Determinación R^2</i> | 57 |
| 4.1.2.3. <i>Coefficiente de Determinación ajustado</i> | 58 |
| 4.1.2.4. <i>Prueba de significancia general de la ecuación de Regresión Múltiple</i> | 59 |
| 4.1.2.5. <i>Prueba de significancia de los coeficientes de regresión parcial</i> | 60 |
| 4.1.2.6. <i>Prueba de colinealidad</i> | 62 |
| 4.1.3. <i>Validación del Modelo Matemático</i> | 62 |
| | |
| CAPÍTULO V | 64 |
| | |
| 5. PROPUESTA | 64 |
| CONCLUSIONES | 66 |
| RECOMENDACIONES | 67 |
| GLOSARIO | |
| BIBLIOGRAFÍA | |
| ANEXOS | |

ÍNDICE DE TABLAS

| | |
|---|----|
| Tabla 1-3. Selección del Software..... | 29 |
| Tabla 2-3. Operacionalización De Variables | 31 |
| Tabla 3-3. Matriz De Consistencia..... | 32 |
| Tabla 4-4. Factores asociados agrupados por provincia de procedencia..... | 37 |
| Tabla 5-4. Tasa de incidencia de las variables respecto a la procedencia..... | 38 |
| Tabla 6-4. Matriz de correlaciones de todos los factores asociados..... | 41 |
| Tabla 7-4. Matriz de correlaciones reducida..... | 44 |
| Tabla 8-4. Valores para el Test de Esfericidad de Bartlett..... | 44 |
| Tabla 9-4. Matriz de correlaciones parciales..... | 46 |
| Tabla 10-4. Prueba de esfericidad de Bartlett e índice KMO | 46 |
| Tabla 11-4. Matriz de valores propios | 47 |
| Tabla 12-4. Raíces de la ecuación de séptimo grado | 48 |
| Tabla 13-4. Cálculo de la varianza total explicada | 48 |
| Tabla 14-4. Cálculo de Varianza total explicada con SPSS..... | 49 |
| Tabla 15-4. Sistema de ecuaciones en forma matricial..... | 49 |
| Tabla 16-4. Matriz de componentes extraídas..... | 51 |
| Tabla 17-4. Matriz de componentes hallada con software SPSS | 51 |
| Tabla 18-4. Tabla de Comunalidades..... | 52 |
| Tabla 19-4. Matriz de datos para Regresión | 56 |
| Tabla 20-4. Coeficientes de regresión..... | 57 |
| Tabla 21-4. Prueba de significancia de la ecuación | 60 |
| Tabla 22-4. Prueba de significancia de los coeficientes de regresión parcial | 61 |
| Tabla 23-4. Comparación de tasas de deserción | 62 |
| Tabla 24-5. Error entre las tasas de deserción..... | 65 |

ÍNDICE DE FIGURAS

| | |
|---|----|
| Figura 1-2. Clasificación de las Técnicas Multivariante..... | 19 |
| Figura 2-2. Técnicas de Análisis de Dependencia según la naturaleza de las variables. | 20 |
| Figura 3-3. Pasos del Método Científico. | 27 |

ÍNDICE DE GRÁFICOS

| | |
|---|----|
| Gráfico 1-2. Distribución de establecimientos de Educación Superior..... | 11 |
| Gráfico 2-2. Tasa Bruta de Matrícula en Educación Superior | 12 |
| Gráfico 3-2. Capacidad disponible para estudiantes en instituciones Superiores | 13 |
| Gráfico 4-2. Aceptación de cupo para Ingreso a la Educación Superior..... | 13 |
| Gráfico 5-4. Procedencia de los estudiantes..... | 35 |
| Gráfico 6-4. Autoidentificación étnica..... | 36 |
| Gráfico 7-4. Nivel de estudio de los estudiantes. | 42 |
| Gráfico 8-4. Género de los estudiantes. | 43 |
| Gráfico 9-4. Nivel socioeconómico de los estudiantes. | 43 |
| Gráfico 10-4. Representación gráfica de las componentes 1 y 2 | 54 |
| Gráfico 11-4. Representación gráfica de las componentes 2 y 3 | 55 |

ÍNDICE DE ANEXOS

ANEXO A. TABLA DE DISTRIBUCIÓN F

ANEXO B. TABLA DE DISTRIBUCIÓN T

RESUMEN

El presente estudio tiene como finalidad determinar una ecuación matemática que prediga la Tasa de Deserción Estudiantil en la Universidad Regional Amazónica Ikiam con base en la información disponible de los estudiantes matriculados en los años 2017, 2018 y 2019, para lo cual se utilizó el software estadístico SPSS y Microsoft Excel para la adecuación de los datos. Las variables analizadas fueron: nivel de estudio, género, etnia, provincia de procedencia y nivel socioeconómico, las mismas que son del tipo cualitativa. Previo al análisis se completaron los campos faltantes de los registros de los estudiantes utilizando la información disponible por el Registro Civil según el número de cédula. Se convirtió las variables cualitativas en cuantitativas utilizando el concepto de variable ficticia. Luego, con el software Microsoft Excel se efectuó el ordenamiento y adecuación de la información según la provincia de procedencia. Posteriormente, se ejecutó el análisis de componentes principales para establecer los factores más relevantes que explican el fenómeno, siendo estos los parámetros mestizo e indígena. Se aplicó el Análisis de Regresión Lineal variables mencionadas para conseguir una ecuación que prediga la Tasa de Deserción Estudiantil. El modelo resultante se verificó para comprobar la exactitud y precisión en el cálculo. En los años 2018 y 2019 el error en el cálculo es de 8.67% y 9.2% respectivamente respecto de los valores registrados. En el año de 2017 el error es de 28.08%, siendo sustancialmente significativo con respecto de los otros años de estudio. Se concluye que el análisis realizado cumple con el objetivo trazado

Palabras claves: MODELO MATEMÁTICO, DESERCIÓN ESTUDIANTIL, REGRESIÓN LINEAL, ANÁLISIS FACTORIAL, ANÁLISIS CUANTITATIVO



07-10-2022

0141-DBRA-UPT-IPEC-2022

ABSTRACT

The present study aims to determine a mathematical equation that predicts the Student Dropout Rate at the Universidad Regional Amazónica Ikiam based on the available information of the students enrolled in the years 2017, 2018, and 2019, whereby the statistical software SPSS and Microsoft Excel were used to determine whether the data were adequate. Level of study, gender, ethnicity, province of origin, and socioeconomic status were the qualitative characteristics evaluated. Before the analysis, the missing fields of the students' records were completed using the information available from the Civil Registry according to the ID card number. The qualitative variables were converted into quantitative variables using the concept of fictitious variable. Then, using Microsoft Excel software, the information was sorted and adjusted according to the province of origin. Subsequently, Principal Component Analysis was then used to determine the mestizo and indigenous parameters as the most important aspects to consider in explaining the phenomena. Linear Regression Analysis was applied to the variables mentioned to obtain an equation that predicts the Student Dropout Rate. The model that resulted from the calculation was examined for precision and correctness. In the years 2018 and 2019, the error in the calculation is 8.67% and 9.2% respectively concerning the recorded values. In the year 2017, the error is 28.08%, being substantially significant to the other years of study. In conclusion, the analysis performed meets the objective set.

Keywords:

MATHEMATICAL MODEL, STUDENT DROPOUT, LINEAR REGRESSION, FACTORIAL ANALYSIS, QUANTITATIVE ANALYSIS

CAPÍTULO I

1. INTRODUCCIÓN

1.1. Planteamiento Del Problema

A nivel mundial existe una marcada diferencia entre personas que logran obtener un título profesional frente a otras que por diversos motivos no lo consiguen. Este aspecto es muy importante y relevante, por ende, es necesario realizar una valoración detallada para determinar las causas que conllevan a que una persona abandone de manera voluntaria o involuntaria el Sistema de Educación Superior. Los alumnos que no logran culminar sus estudios universitarios tienen como consecuencia un costo social, económico, académico y familiar, es por ello, que las Universidades se ven en la necesidad de planificar e implementar estrategias y políticas que promuevan la retención de los estudiantes en los primeros años de estudios universitarios.

Es así que para (García Fernández et al., 2016) según el pensamiento de Przybylski, la aplicación exitosa de las políticas educativas conlleva a reducir el abandono escolar. En nuestro país la tasa de deserción estudiantil en los centros universitarios alcanza el 51.65%, según los datos proporcionados por el repositorio digital PUCE del 2011. Por ello, la calidad educativa que ofrecen los centros de Educación Superior es un factor muy importante y en ningún momento debe dejarse de lado, sino más bien cuestionarlo.

Es esencial proponer medidas que disminuyan la deserción, por ende, es preciso desarrollar modelos que expliquen los orígenes, causas, y así proponer soluciones. Por ello, los centros de Educación Superior deberían implementar sistemas para monitorear, analizar y prevenir la deserción estudiantil desde el momento en que ingresa hasta el momento que el estudiante interrumpe o abandona sus estudios.

El interés que un estudiante pueda tener para iniciar sus estudios en una cierta carrera, depende el puntaje obtenido en la prueba de ingreso al Sistema de Educación Superior. Este condicionamiento, de no ser cumplido causa en el estudiante incertidumbre, desinterés y apatía hacia la carrera designada y a la universidad que lo acoge, terminando quizá en un futuro abandono.

La escasa información que disponen los estudiantes de los últimos años de la Educación Media, sobre las Universidades y de las carreras que tienen disponibles, sumada con la deficiente

orientación vocacional impartida hacen que el estudiante no contemple todo el abanico de oportunidades que tiene y provocan que el futuro alumno opte por una carrera que quizá no es de su agrado.

La falta de dinero es un importante condicionamiento que debe superar un estudiante desfavorecido para cubrir con las exigencias de la carrera, sumado a la carencia de financiamiento de instituciones gubernamentales. De igual forma, la baja expectativa de encontrar un trabajo y una remuneración que le permita suplir todas sus necesidades.

El bajo nivel de estudios agravado por problemas familiares provoca que repruebe las asignaturas de la malla escolar desencadenando afectaciones en la salud física y mental del estudiante, terminando en un potencial abandono.

En la región Amazónica, en la provincia de Napo, hace aproximadamente 8 años atrás se funda la Universidad Regional Amazónica Ikiam, con cuatro facultades y carreras destinadas principalmente al estudio y protección del medio Ambiente. Muchos de los estudiantes que pertenecen a la mencionada Universidad fueron designados bajo una prueba de Ingreso al Sistema de Educación Superior. Según (Montánchez & Martínez, 2017), en el Plan de Acción Tutorial manifiestan que el 66% de los estudiantes matriculados en la Universidad Regional Amazónica Ikiam reprueban o desertan. Por lo tanto, se requiere de un estudio que permita identificar las causas del abandono en este centro de Educación Superior. Por ello, es imprescindible proponer un modelo matemático para cuantificar la deserción estudiantil y así, las autoridades del establecimiento opten por decisiones que ataquen y disminuyan el alto porcentaje de estudiantes reprobados o desertados.

1.1.1. Situación problemática

Según la Constitución de la República del Ecuador en su Artículo 26 establece a la Educación como un derecho y un deber ineludible e inexcusable del estado. Esto indica que se debe permitir que todas las personas sin impedimento de edad puedan acceder a un establecimiento educativo hasta culminar sus estudios. Sin embargo, esto no se cumple, pues en el informe de rendición de cuentas de 2019 del Ministerio de Educación sostiene que de 4 998 169 personas entre 3 y 17 años, solamente 4 385 168 estudiantes pertenecen a una institución educativa. De igual manera, según (SENESCYT, 2021), para el año 2021 están habilitados 218 centros de Educación Superior para cubrir una población de 94 349, empero, a nivel del país cada año se incorporan cerca de trescientos mil nuevos bachilleres, dejando sin opción de estudiar a dos terceras partes de ellos,

quienes deberán esperar a un nuevo periodo para intentar conseguir un cupo, teniendo en cuenta que en cada periodo se incorporan más bachilleres.

En el artículo 29 de la Constitución de la República del Ecuador, determina que la familia tiene la libertad de seleccionar una educación para sus representados según sus principios, creencias y opciones pedagógicas. Esto dista mucho de la realidad, pues como se ha mencionado en el párrafo anterior solamente la tercera parte de los estudiantes accede a un cupo en un establecimiento de Educación Superior. Sin embargo, aunque disponga del cupo, la carrera universitaria no siempre está acorde a sus capacidades y habilidades adquiridas, lo cual, si bien la mayoría opta por iniciar sus estudios superiores, esto por lo general termina con un potencial abandono, pues la carrera asignada no cubre las expectativas que tienen los estudiantes para su vida futura.

Lo mencionado, se ha detectado que los establecimientos de Educación Superior, no disponen de los cupos suficientes para que un estudiante pueda optar por una carrera teniendo en cuenta sus capacidades y habilidades. Es la frustración de tener que adaptarse a una carrera que no es de su interés, que incrementa la posibilidad de que una persona pueda abandonar o no sus estudios

La constitución del Ecuador, en el artículo 348 y 356, determinan la gratuidad de la Educación en los establecimientos públicos hasta el tercer nivel. A su vez, en el artículo 357 sostiene que el estado garantizará el financiamiento. Para el año 2020, el gobierno presupuestó cerca de 1 586 millones de dólares para el sistema Universitario, sin embargo, por la pandemia del COVID-19, realizó un recorte resultando 1 425 millones como el nuevo presupuesto (Torres & Guerra, 2020). Esto ha conllevado, a que las Universidades reduzcan gastos y reasignen partidas, profundizando aún más la delicada situación que ha venido atravesando la Educación. Por lo tanto, para ampliar los cupos del Sistema de Educación Superior, se debe destinar una cifra muy superior a la citada para inversión en infraestructura y partidas presupuestarias.

Como se ha expuesto, en nuestro país, el presupuesto destinado a Educación no cubre con las expectativas en todos sus niveles. Los estudiantes que logran incorporarse de Bachilleres, que luego optarán por un cupo en las instituciones de Educación Superior provienen de diferentes establecimientos y con realidades particulares. Según (Benalcázar, 2017) no existe diferencia en que un estudiante universitario proceda de un establecimiento público o privado. Por lo tanto, la diferencia entre el éxito o no en la culminación de una carrera profesional, no solo depende de las capacidades del estudiante sino también, del lugar de procedencia, de las condiciones económicas y personales.

La deserción en nuestro país ha disminuido en los últimos años. Según (Torres M. , 2019) con base en lo publicado por la SENESCYT, la deserción promedio a nivel Nacional ha pasado del 50% en el 2012, a un 26% en el año 2016. En la Universidad de Guayaquil, una de las instituciones de Educación Superior más grandes del país, la deserción promedio pasó del 21% en 2017 hasta el 28% en el 2019. Cabe notar que en otros establecimientos de Educación Superior los promedios de deserción estudiantil son bajos, sin embargo, preocupan a sus autoridades; por ejemplo, la ESPOL pasó del 4.49% en el 2016 al 2.99% en el 2018, con una población estudiantil que bordea los quince mil estudiantes.

En la Universidad Regional Amazónica Ikiam, según cifras de la SENESCYT, la tasa de deserción estudiantil pasó del 12,5% en el 2015 al 13.4% en el 2017. Según los datos registrados por la institución en el 2017 se matricularon 647 estudiantes, y en el 2020 se matricularon 1606 estudiantes. Por lo tanto, la tasa de deserción no ha sufrido un cambio importante entre los años citados. Sin embargo, en cada periodo crece el número de alumnos matriculados, por ende, la cantidad de alumnos que abandonan sus estudios se ha incrementado. Según (Torres & Guerra, 2020), el promedio nacional por estudiante destinado a educación es 5 924 dólares por año. Esto es un indicativo de que es necesario realizar un análisis detallado de los factores asociados a la deserción estudiantil, así como de un modelo que permita predecir, en diferentes escenarios, la tasa de deserción para que las autoridades puedan optar por mejores estrategias y políticas educativas a fin de reducir la deserción.

Como se ha expuesto, la deserción estudiantil es un tema de análisis. Sin embargo, a nivel del país no se ha analizado a profundidad, ni se cuenta con modelos que predigan tasas de deserción ante escenarios definidos a pesar que vivimos en una era digital. Por ello, es imprescindible disponer en la Universidad Ikiam del primer modelo matemático para el análisis de la deserción estudiantil que predigan este fenómeno, y la institución con el uso de esta nueva herramienta en tecnología se podrá analizar y evaluar varios escenarios y optar por la mejor solución a nuestra problemática.

1.1.2. Formulación Del Problema

¿Puede medirse la tasa de deserción estudiantil, mediante la aplicación de un modelo Matemático, basado en el método regresión lineal múltiple de variables ficticias y/o dicotómicas con la información existente de los estudiantes matriculados en la Universidad Regional Amazónica Ikiam?

1.1.3. Preguntas Directrices

- a. ¿Cuáles son los factores más importantes asociados a la deserción estudiantil, con base en la información existente de los estudiantes matriculados en la Universidad Regional Amazónica Ikiam?
- b. ¿Qué técnica de análisis multivariante es la más idónea para predecir la tasa de deserción en la Universidad Regional Amazónica Ikiam?
- c. ¿Se puede determinar la tasa de deserción estudiantil mediante la aplicación de un modelo matemático obtenido a partir de una técnica de análisis multivariante, con la información existente de los estudiantes matriculados en la Universidad Regional Amazónica Ikiam?
- d. ¿Cuál es el error entre la tasa de deserción calculado aplicando el modelo matemático y el valor real?

1.1. Justificación de la Investigación

1.1.1. Justificación teórica

En la república del Ecuador, la deserción estudiantil en los establecimientos de Educación Superior del país es relativamente alta. Según (SENESCYT, 2018) en el año 2015 la deserción estudiantil Universitaria en promedio era del 29.6%, en el año 2016 era de 26.2%, mientras que para el 2017, fue del 23.2%, evidenciándose que las mayores tasas de deserción se producen en la Universidades Autofinanciadas y Cofinanciadas.

La deserción en las instituciones es producto de varios factores, los mismos que son conocidos por las Universidades, a través de estudios de investigaciones de carácter cualitativo, es decir, identificar las causas que producen que un estudiante abandone sus estudios, mas no establecen que aspectos del estudiante se debería abordar para que el establecimiento de Educativo establezca políticas educativas que reduzcan el abandono estudiantil.

Por lo tanto, es necesario que las instituciones de Educación Superior, opten entre sus tantos campos de investigación, por analizar puertas a dentro más al detalle la deserción estudiantil. También se requiere de modelos matemáticos capaces de predecir la tasa de deserción estudiantil utilizando recursos tecnológicos disponibles que permitan procesar la información de manera más eficiente.

1.1.2. Justificación metodológica

Con este estudio se procurará obtener un modelo matemático que pronostique la tasa de deserción estudiantil en la Universidad Regional Amazónica Ikiam a partir de la información existente en el Centro Educativo. Dicho modelo es el resultado de un análisis estadístico detallado de los datos basado en técnicas de análisis estadístico conocidas y altamente fiables. Para ello, se recurrirá al uso de software especializados en el tratamiento y procesamiento de datos. En este sentido, el presente estudio servirá como base para futuras investigaciones ya que se brindará información y sugerencias, las cuales pueden ser acogidas por futuras investigaciones sobre la problemática para mejorar o implementar acciones frente a este escenario.

1.1.3. Justificación práctica

El presente trabajo de investigación se basa en la necesidad de optimizar recursos para la educación estudiantil en la Universidad Regional Amazónica Ikiam, ya que, a la presente fecha, según publicaciones de la SENESCYT, en el año 2020 el promedio nacional anual por educación en establecimientos de Educación Superior es de 5 924 dólares por estudiante. Por lo tanto, si un estudiante abandona sus estudios los costos económicos perdidos por la Universidad son elevados.

El resultado de la presente investigación permitirá, disponer de un modelo matemático que prediga la tasa de deserción bajo ciertos parámetros, así como identificar las causas más relevantes que lo predicen, y de esta manera permitir que la institución disponga de datos que respalden la implementación de políticas y estrategias que minimicen la deserción estudiantil.

El trabajo de investigación evaluará la información disponible actualmente por la institución de cada estudiante legalmente matriculado en el establecimiento sobre aspectos, tales como: social, académico, familiar y económico; caso contrario recomendar que la institución implemente mecanismos para para obtener dicha información.

1.2. Objetivos de la Investigación

1.2.1. Objetivo General

Diseñar un modelo matemático enfocado a la reducción del abandono estudiantil a través de un análisis multivariante para la Universidad Regional Amazónica Ikiam

1.2.2. Objetivos específicos

- a. Identificar los factores sociales y económicos que motivan la deserción con base en la información existente de los estudiantes de la Universidad Regional Amazónica Ikiam.
- b. Establecer un modelo matemático que prediga la tasa de deserción estudiantil en la Universidad Regional Amazónica Ikiam mediante la aplicación de las Técnicas de Análisis multivariante más idóneas.
- c. Validar la exactitud en la predicción del modelo matemático con los registros históricos del establecimiento educativo.

CAPÍTULO II

2. MARCO TEÓRICO

2.1. Estado del Arte

En Ecuador, las instituciones de nivel educativo superior han realizado investigaciones para identificar las causas que motivan a que un estudiante abandone sus estudios. Un estudio desarrollado por Zambrano, Rodríguez y Guevara en el año 2018, titulado “Análisis de la Deserción estudiantil en las Universidades del Ecuador y América Latina”, tienen como objetivo identificar los factores que interaccionan e inciden en el aumento de la deserción estudiantil. Por ello, se analiza las causas y factores en los aspectos socioeconómico, educativo y psicológico, logrando identificar los más importantes que motivan la deserción estudiantil en el nivel superior de nuestro país y Latinoamérica, mediante el análisis estadístico descriptivo. Concluyen que entre los factores más destacados están: los estudiantes fueron asignados a una carrera no acorde a sus intereses, la carencia del dominio de contenidos esenciales y la aplicación de técnicas tradicionales en la enseñanza (Zambrano et al., 2018).

De igual manera, los investigadores Bravo, Illescas, Larriva y Peña de la Universidad de Cuenca, realizaron un estudio en el año 2017 titulado “Causas de Deserción en el Ingreso a la Universidad; un Estudio de Caso”, cuyo objetivo es identificar las causas que inciden en la deserción de los estudiantes del curso de nivelación de la Universidad de Cuenca. Para ello, analizaron las fichas socioeconómicas de los estudiantes, pruebas antes y después del curso y encuestas dirigidas a los estudiantes que desertaron. Algunas respuestas obtenidas fueron: carencia de bases, carrera no acorde a sus intereses, alta exigencia académica, sistema de evaluación complejo, escasez de recursos económicos, entre otras. Esto permite tomar e implementar acciones para disminuir los índices de abandono (Bravo et al., 2017).

Se presenta la investigación con tema “Adaptación de un modelo matemático para la predicción del abandono estudiantil en la educación superior” cuya finalidad es diseñar un modelo matemático para el cálculo del porcentaje de abandono estudiantil en el primer nivel de educación superior a través del agrupamiento de estudiantes según factores adecuados. Se establece un modelo matemático (ecuación) con base en la información existente de la Escuela Politécnica Nacional (EPN) a través del Análisis de Regresión Lineal Múltiple que pronostica el porcentaje de abandono estudiantil, considerando como variables predictoras: el promedio del examen “Ser

Bachiller”, sostenimiento del colegio del que procede el estudiante (público o particular) con gran precisión y exactitud respecto del registro histórico (Alcocer, 2020).

El estudio titulado “Aplicación de un modelo de duración en programas de prevención de deserción universitaria” efectuado por Herrero, Merlino Ayllón y Escanés en el 2012 para la Universidad Empresarial Siglo 21 propone aplicar un modelo predictivo de riesgo de deserción en los estudiantes universitarios para proponer medidas que mitiguen este suceso. Se basa en los modelos de duración que permiten determinar la probabilidad de que el abandono se suscite en cierto periodo de riesgo. La aplicación de este tipo de modelo se basa en la premisa que la deserción no es fácilmente detectable, sino que depende de diferentes etapas del desarrollo estudiantil. El modelo abarca dos funciones relacionadas que dan respuesta a la probabilidad que un estudiante permanezca en sus estudios. Una función estima la probabilidad de abandono, mientras que otra indica la probabilidad de permanencia de un estudiante con el paso de los periodos. El modelo se aplica en cuatro fases que son: al realizar la inscripción, durante el curso de nivelación, en el primer cuatrimestre y al finalizar el cuatrimestre. La precisión y exactitud mejoran cuando los resultados de la o las fases previas retroalimentan a la o las siguientes, ya que se incluyen nuevos factores que inicialmente no eran considerados. A su vez esto, a criterio de los investigadores es positivo, pues a medida que existe más información de los estudiantes analizados se puede reducir el tamaño de la muestra concentrándose únicamente en aquellos que presentan más riesgo de abandono, lo cual permite focalizar los esfuerzos para evitar la deserción (Herrero et al., 2013).

Otro estudio titulado “Modelación de la deserción Universitaria mediante Cadenas de Markov” propuesto por González, Carvajal y Aspeé, cuya meta es definir un índice como estimador del riesgo a la Deserción individual para retener a un estudiante. Para determinar dicho índice utilizan las Cadenas de Markov. Se basa en el concepto de que la deserción es un fenómeno de ocurrencia aleatoria. En tal sentido, el modelo determina la probabilidad de deserción de un estudiante en el transcurso de una carrera universitaria. Los investigadores concluyen que las probabilidades son altas cuando el alumno cursa los dos primeros niveles y va disminuyendo conforme el estudiante avanza en sus estudios. Esta tendencia se replica entre facultades y carreras sin diferencias significativas en una misma Universidad. Sin embargo, este comportamiento no se imitan entre centros educativos, lo que incurre en que se debería realizar una investigación más profunda para generalizar (J. González et al., 2020).

Se ha desarrollado un estudio desde el punto de vista económico titulado “Deserción Estudiantil y el Costo Económico en Universidades Chilenas” por Ramírez en el año 2020. La investigación tiene como finalidad estimar y analizar los costos que genera el abandono estudiantil desde la

perspectiva institucional, proyectando los recursos que se dejaría de percibir a consecuencia de la deserción. El análisis se centra en la deserción del primer año de carrera universitaria, estimando los costos provocados por la deserción en cuatro aspectos: cantidad de matriculados, número de desertores, docentes y administración de las Universidades. El autor concluye que es de vital importancia para los establecimientos educativos estimar los recursos económicos perdidos a causa de la deserción para formular estrategias que conlleven a minimizar los costos. Propone que es necesario profundizar la investigación para detallar al máximo los costos directos e indirectos de la educación superior para especificar estrategias que permitan disminuir los costos (Ramírez, 2021).

En la investigación titulada “Modelo Predictivo de deserción estudiantil basado en árboles de decisión” elaborado por Cuji, Gavilanes y Sanchez en el 2017, tiene como finalidad pronosticar la probabilidad que un estudiante abandone el programa académico mediante técnicas de clasificación basadas en árboles de decisión. Se considera la información académica y personal del alumno. El árbol de decisión contiene un nodo raíz de cual se parte al resto de nodos división y posteriormente a nodos hoja, que contienen un conjunto de observaciones. Cada regla asigna una observación a un segmento de acuerdo al valor de entrada. Una nueva regla se aplica finalizada la anterior, obteniendo una jerarquía de los segmentos dentro del conjunto de segmentos formado. La jerarquía resultante se denomina árbol y cada segmento formado se llama nodo. La complejidad del análisis requiere potentes dispositivos tecnológicos y de modificaciones continuas de ciertos parámetros. Por ello, los autores concluyen que el modelo propuesto motiva a que se creen nuevos modelos de predicción, usando técnicas de clasificación entre ellas redes neuronales y regresión logística para realizar análisis comparativos de los factores que inciden en la deserción estudiantil. (Cuji et al., 2017).

El trabajo realizado por Argote, Iván; Jiménez, Robinson y Gómez, Jair titulado “Detección de Patrones de Deserción en los Programas de Pregrado de la Universidad Mariana de San Juan de Pasto, aplicando el proceso de descubrimiento de conocimiento sobre base de datos (KDD) y su implementación en Modelos Matemáticos de Predicción”, tiene como objetivo aportar al proceso de toma de decisiones para disminuir los niveles de deserción estudiantil en los programas de Pregrado de la Universidad Mariana, mediante el descubrimiento de conocimiento sobre bases de datos (KDD) e implementación en el diseño de unos modelos de predicción en relación a cuatros factores establecidos por el Ministerio de Educación de Colombia de los aspectos socioeconómico, personal, académico e institucional. Las expresiones se han obtenido mediante la aplicación de modelos basados en la minería de datos y el índice de deserción institucional es el promedio de los cuatro. Los autores concluyen que para tener modelos confiables es necesario que la institución disponga de información a detalle de los estudiantes en la mayor cantidad de

aspectos para que la predicción sea más precisa y exacta, facilitando la toma de decisiones a fin de establecer estrategias para disminuir la deserción estudiantil (Argote, I., Jimenez, R. y Gómez, 2014).

2.2. Bases Teóricas

2.2.1. La Educación Superior en el Ecuador

El Sistema de Educación Superior nacional de nuestro país está compuesto en el año 2021, de 218 centros educativos, siendo; 55 (25%) Universidades y Escuelas Politécnicas, de las cuales 29 (13%) son públicas, mientras que 26 (12%) corresponden a Universidades privadas y 163 (75%) son institutos técnicos y tecnológicos, de los cuales 83 (38%) son públicos y 80 (37%) son particulares (SENESCYT, 2021). En el Gráfico 1 se muestra la distribución de forma esquemática

Distribución de Establecimientos de Educación Superior

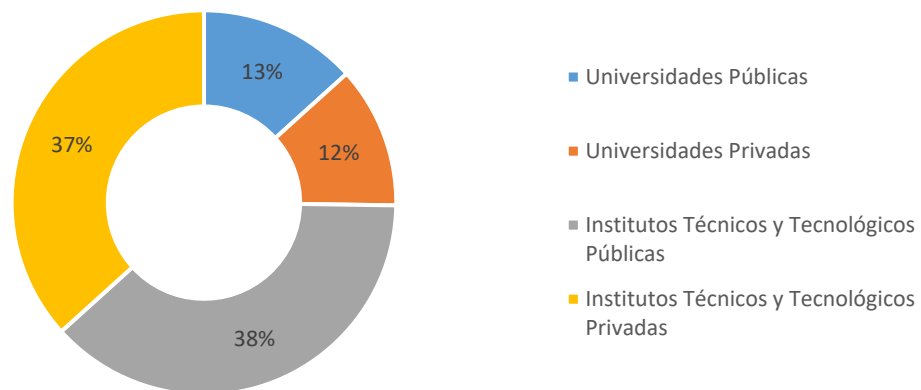


Gráfico 1-2. Distribución de establecimientos de Educación Superior

Realizado por: Ayabaca, Roberto. 2022

De acuerdo al acceso a la Educación Superior para (Ponce & Carrasco, 2017) la tasa bruta de matrícula universitaria en 1999 era de 20.2%, para el año 2006 se situó en 26.5% y en el 2014 alcanzó la cifra de 29.6%. En contraste, según (SENESCYT, 2018) en el 2014 era del 27.17%, para el 2015, fue de 27.76% y en el 2016 se incrementó hasta 27.81%. Hasta el 2016 se publican cifras oficiales, y para los años subsiguientes solo se dispone de proyecciones realizadas a partir de los datos históricos antes mencionados, por ende en la proyección para el año 2021 se alcanzaría el 31.21% (SENESCYT, 2020). En la Gráfico 2 se sintetiza lo manifestado.

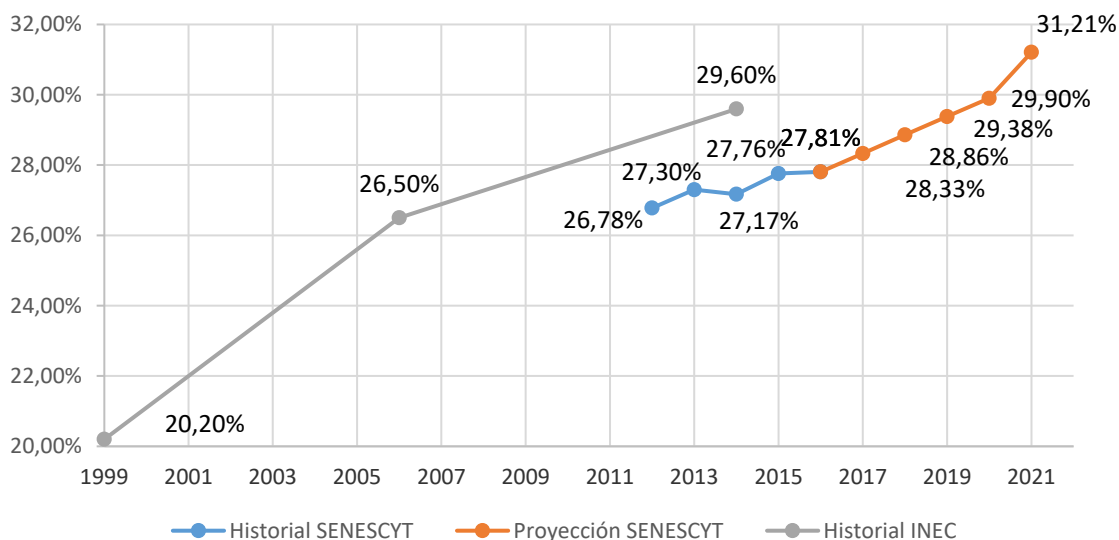


Gráfico 2-2. Tasa Bruta de Matrícula en Educación Superior

Realizado por: Ayabaca, Roberto. 2022

En 2008 entró en vigencia la nueva Constitución, la cual modificó las políticas de acceso a la Educación Superior. La tasa bruta de asistencia universitaria pasó del 18.5% en el 2000 al 27% en 2008. Posterior a ello, se registra un aumento considerable en la asistencia universitaria llegando al 31% en 2011. En parte este incremento está ligado a la aplicación de las nuevas políticas, entre ellas la gratuidad de la Educación Superior. A partir del 2011, empero, la asistencia cae considerablemente hasta el 21% en el 2015. Esta reducción importante puede ser asociada por la implementación del Examen ENES y por el cierre definitivo de las Universidades que hasta esa fecha no alcanzaron los estándares mínimos de calidad educativa. Hechos que se suscitaron en 2012 (Ponce & Carrasco, 2017).

En lo referente a la capacidad de los establecimientos de Educación Superior en la Admisión de nuevos estudiantes. Según (SENESCYT, 2020) los cupos ofertados para el ingreso a una carrera universitaria en los establecimientos de Educación Superior a nivel nacional cubrían el 49% para el 2015, 45% en el 2016, 31% en el 2017 y 57% en el 2018 de la demanda de estudiantes postulados, aceptando el cupo el 97%, 95%, 75% y 89% respectivamente de los cupos disponibles de los años citados. En el Gráfico 3 y Gráfico 4, se muestra para una mejor comprensión.

Cobertura estudiantil en la Educación Superior

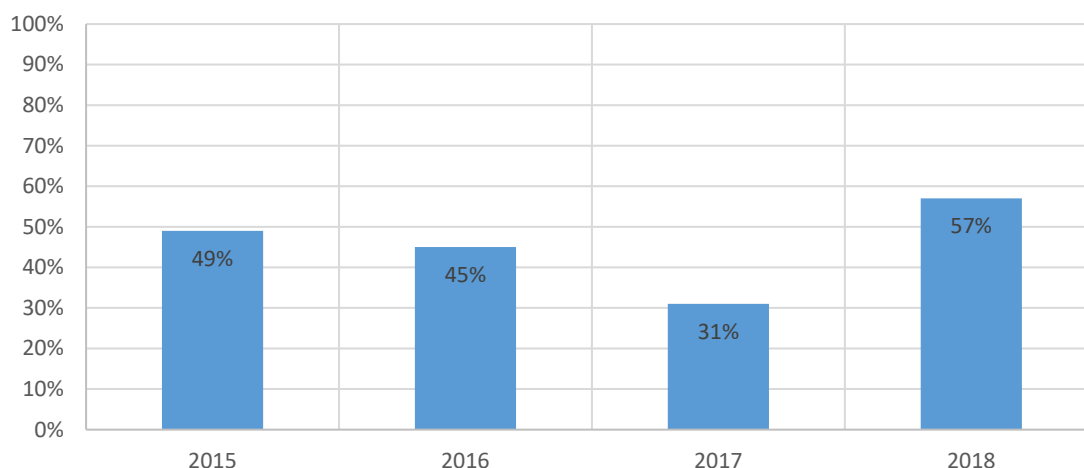


Gráfico 3-2. Capacidad disponible para estudiantes en instituciones Superiores

Realizado por: Ayabaca, Roberto. 2022

Aceptación del cupo para la Educación Superior

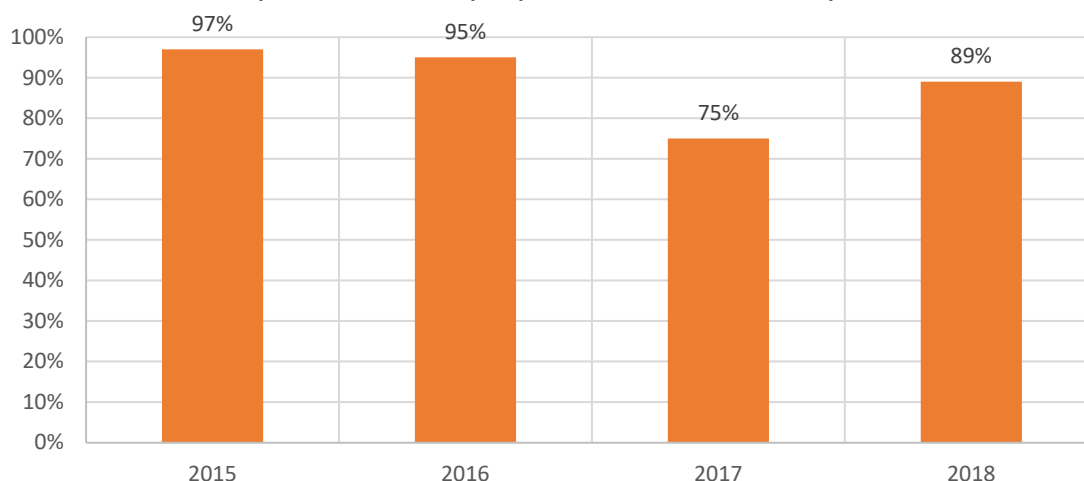


Gráfico 4-2. Aceptación de cupo para Ingreso a la Educación Superior

Realizado por: Ayabaca, Roberto. 2022

2.2.2. *La Deserción Estudiantil en el Ecuador*

En la República del Ecuador no se dispone de estudios detallados sobre la deserción estudiantil en los establecimientos de Educación Superior. Se dispone de estudios particulares desarrollados en ciertas Universidades que indagan sobre los factores que inciden en la deserción estudiantil en dicho centro Educativo, siendo estos aspectos personales, académicos, sociales y económicos. Sin embargo, la SENESCYT ha mostrado preocupación por los efectos de este fenómeno. Dispone de tasas de retención y deserción estudiantil en el periodo comprendido entre 2015 y 2017 de las

Universidades tanto públicas y privadas, descartando los institutos técnicos y tecnológicos. En el 2017, la tasa de deserción estudiantil a nivel nacional fue del 23.2%.

La deserción estudiantil tiene consecuencias económicas en términos de expectativa de un empleo a tiempo completo. En la zona urbana, aquellas personas que carecen de instrucción o disponen de estudios primarios solamente el 20% de ellos tiene un empleo pleno, mientras que tres de cada cinco (60%) personas que han finalizado sus estudios, es decir, que cuentan con un título profesional tienen un empleo pleno. En contraste, en la zona rural una persona sin instrucción tiene una mayor expectativa de poseer un empleo con un ingreso estable superior a una persona con nivel de instrucción superior que tiene limitadas opciones de empleo a tiempo completo (INEC, 2015). Por lo tanto, una persona que haya abandonado sus estudios tendrá restringidas opciones para optar por un empleo pleno.

2.2.3. La Deserción desde la perspectiva Institucional

Desde la perspectiva institucional, el alumno que abandona sus estudios es catalogado como desertor. En consecuencia, se genera un puesto vacante que pudo haber sido ocupado por un estudiante que permaneciera en sus estudios, ocasionando dificultades en el financiamiento del establecimiento (Tinto, 1982). En este sentido, no todos los abandonos obedecen a determinadas causas que se puedan agrupar de manera sencilla, sino responden a particularidades en ciertos casos dificultando su detección y análisis (Tapasco Alzate et al., 2019). Por lo tanto, el análisis que se realice debe orientarse a un determinado tipo de abandono para detallar sus causas y obtener resultados satisfactorios.

2.2.4. Definición de Deserción Universitaria

Según (Chow & Flores, 2021) bajo el pensamiento de Quintero, afirma que la deserción es un fenómeno determinado por varias causas que agrupados resultan en factores sociales, familiares e individuales. Para (M. González et al., 2018) la deserción es la interrupción de una persona en sus estudios. (Zambrano et al., 2018) según el criterio de Rodríguez y Hernández, sostiene que la deserción estudiantil es la finalización del vínculo entre el estudiante legalmente matriculado y la universidad.

La deserción desde el punto de vista del individuo es el fracaso por no completar una carrera universitaria, producto de la escasez de carácter o interés para responder a las exigencias del centro Educativo en la consecución de un título universitario (Tinto, 1982).

La deserción estudiantil puede comprenderse como el abandono definitivo, sin oportunidad de retorno para culminar una carrera universitaria. La cantidad de desertores cambia dependiendo de la Universidad (Páramo & Correa, 2019)

2.2.5. Tipos de Deserción Universitaria

Existen dos tipos de deserción claramente identificables siendo éstas respecto del tiempo y del espacio. Refiriéndose al tiempo como la adaptación del alumno con la carrera al inicio, en los primeros semestres y cuando cursa los últimos; mientras al espacio como la asociación del individuo con la institución y el sistema académico (Aparicio & Seminara, 2018).

La deserción respecto del tiempo más frecuente es aquella que se producen en los primeros semestres, conocida como deserción temprana provocada generalmente por una expectativa errónea que el estudiante pueda interpretar en base a la información camuflada que presenta la Universidad sobre su oferta académica (Tinto, 1982).

Para (Páramo & Correa, 2019) mencionan los siguientes tipos de deserción: discriminada por causas (debido a factores que influyen en su decisión final), por facultad (cambio de facultad), por programa (cambio de carrera), por primer semestre (por desinterés o desmotivación en la adaptación a la carrera universitaria) y acumulada.

2.2.6. Modelos de Análisis de la Deserción Estudiantil

La deserción estudiantil es producto de varios factores que inciden para que el alumno opte por abandonar sus estudios de manera definitiva (Chimba, 2020). Por ello, es necesario analizar al detalle las causas involucradas para que las instituciones tomen decisiones que disminuyan al máximo el abandono (Lázaro, 2020). Cabe resaltar que este fenómeno no se podrá eliminar por más grandes esfuerzos que se realicen, pues la gran cantidad de factores de los que depende imposibilita tal acción (Poveda et al., 2020). Con ello, no se pretende dejar de lado un análisis, sino más bien estudiarlo para reducirlo. A continuación, se describen algunos modelos los cuales permiten abordar un estudio desde diferentes enfoques.

2.2.6.1. Modelos Psicológicos

Este tipo de modelos abarcan patrones relacionados con la personalidad de los estudiantes que culminan su carrera de quienes no lo logran. La decisión de continuar o no sus estudios es producto de conductas previas, del concepto que disponga sobre la deserción y del debilitamiento

de sus intenciones iniciales. Con este tipo de análisis se ha determinado que el rendimiento académico previo incide de manera directa en el actuar futuro del estudiante, afectando su percepción en sus estudios, sus metas y expectativas futuras (Guglielmi et al., 2018). Este criterio es compartido por (Bernal-Cifuentes, 2018).

2.2.6.2. Modelos Sociológicos

Los modelos sociológicos a más de estudiar la parte psicológica del estudiante abarcan factores externos al individuo (Lázaro Alvarez et al., 2017). Para (Albarrán Peña, 2019) los factores socioeconómicos e institucionales conllevan al abandono.

2.2.6.3. Modelos Económicos

Este tipo de modelos estudian el razonamiento que realiza un estudiante sobre el beneficio que obtendría si culmina una carrera universitaria a pesar de que para conseguirlo deba optar por algunas decisiones trascendentales como trabajar para costear sus estudios. De igual manera, se estudia como las ayudas proporcionadas a los alumnos como becas, financiamiento y subsidios favorece la permanencia de ellos en los establecimientos educativos (Díaz et al., 2021), (Smulders Chaparro, 2018).

2.2.6.4. Modelo Organizacional

Con este tipo de modelos según (Mejía & Vanegas, 2018), (Bautista, 2019) se analiza la calidad y experiencia de los docentes, la administración institucional que afectan directamente en el vínculo social del alumno.

2.2.6.5. Modelo de Interacción

Este tipo de modelos estudia la relación entre el estudiante y el establecimiento educativo, considerando la procedencia y cultura del y entre estudiantes con la cultura universitaria, incidiendo de manera directa o indirecta en el sentido de pertinencia de la persona con el establecimiento (Poveda et al., 2020), (Galindo & Ramírez, 2018).

2.2.7. Tipos de Variables

2.2.7.1. Variables Cualitativas

Son de naturaleza no métrica, es decir, representan cualidades, atributos y categorías no numéricas. Pueden ser nominales que se agrupan por categorías sin un orden específico y, las ordinales que se las categorías están ordenadas bajo un criterio, de menor a mayor o viceversa (Ramos et al., 2020).

2.2.7.2. Variables Cuantitativas

Las variables cuantitativas son de naturaleza métrica, es decir, que se expresan mediante cantidades numéricas. Se subdividen en dos grupos: continuas y discretas. Las variables continuas pueden tomar cualquier valor numérico a partir de una medición. Las variables discretas provienen del conteo asociado a los números enteros (Matos et al., 2020). No obstante, también son discretas las variables ficticias que proceden de la transformación de las variables categóricas, cuando a cada clase se le asigna un valor numérico (Pérez, 2004).

2.2.8. Técnicas de Análisis Multivariante de Datos

En un estudio, los investigadores disponen de una gran variedad de variables, las mismas que son recogidas mediante observaciones y/o mediciones. Para el estudio, por lo general se suele recurrir al análisis estadístico de la información. Por la diversidad de datos obtenidos se disponen de una gran variedad de técnicas, por ende, se debe detectar la más idónea y que mejor resulte para el estudio de los datos (Pérez, 2004).

Por otra parte, una clasificación de las técnicas de análisis puede ser según la autonomía entre las variables utilizadas (Meneses, 2019). Por un lado, los métodos de interdependencia que analizan el grado de relación entre las variables. Estos métodos sirven para disminuir el número de variables a estudiar y/o clasificar los factores según alguna característica común, que describen un fenómeno (Mercedes & Gutiérrez, 2019). Por otro lado, los métodos de dependencia, en los cuales un parámetro es explicado y/o pronosticado a partir de varios otros utilizando una ecuación o modelo para tal efecto (Sagaró del Campo & Zamora, 2020). En la Figura 1, se observa la clasificación de los métodos multivariantes con base en la naturaleza de los factores existentes y del Análisis más idóneo a aplicarse.

Las técnicas también se pueden clasificar dependiendo el tipo de variable. Cuando las variables son interdependientes entre sí y también son cuantitativas se puede aplicar un Análisis de Componentes Principales (ACP). A su vez, si los factores son cualitativos se opta por un Análisis de Correspondencias o un Análisis de Escalamiento Multidimensional. Sin embargo, si los parámetros son cuantitativos y cualitativos, se estima un Análisis Tipología. Caso contrario, cuando existe dependencia, es decir, una variable está condicionada por otras. En este sentido, si el factor vinculado es cuantitativo, al igual que los parámetros explicativos se elige Regresión Lineal. No obstante, si las variables independientes son cualitativas se aplica Modelos de Análisis de la Varianza que puede extenderse a un Modelo Loglineal. En cambio, si la variable dependiente es cualitativa y los factores son cuantitativos se somete a un Análisis Discriminante. Finalmente, si la variable dependiente es cuantitativa o cualitativa y los parámetros explicativos son cualitativos se rige por un Análisis de Segmentación (Pérez, 2004), (Hidalgo, 2019).

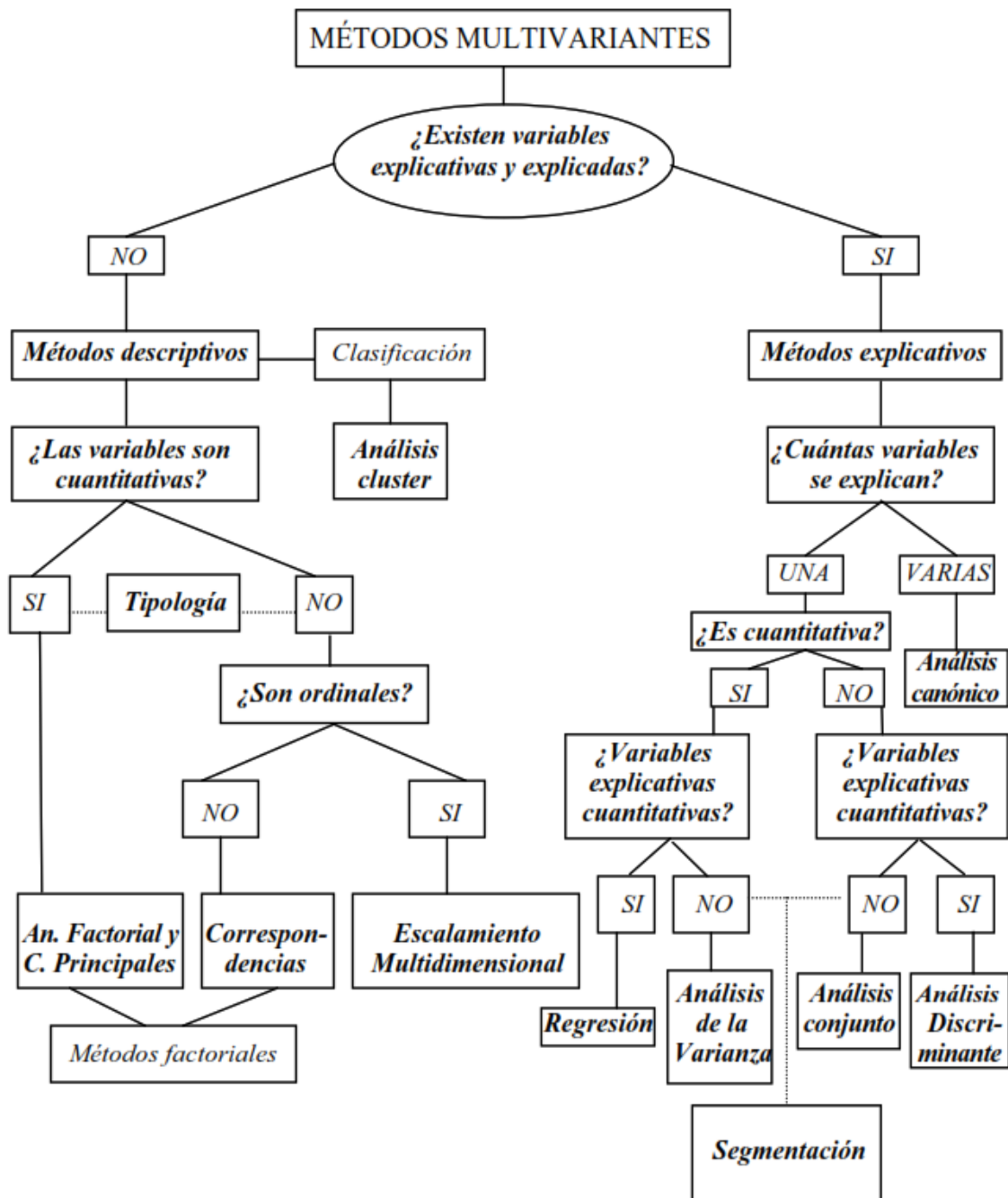


Figura 1-2. Clasificación de las Técnicas Multivariante

Fuente: Pérez, C. 2004

2.2.9. Técnicas de Análisis de Dependencia

Las Técnicas de Análisis de Dependencia, como se mencionó en la sección anterior se basan en la existencia de un factor explicado a partir de varios parámetros explicativos. Asimismo, se estableció que tanto las variables independientes como la dependiente pueden ser cuantitativas (métricas) o cualitativas (no métrica).

En la Figura 2 se muestra el tipo de análisis recomendado en función de la naturaleza de los factores explicativos y explicados.

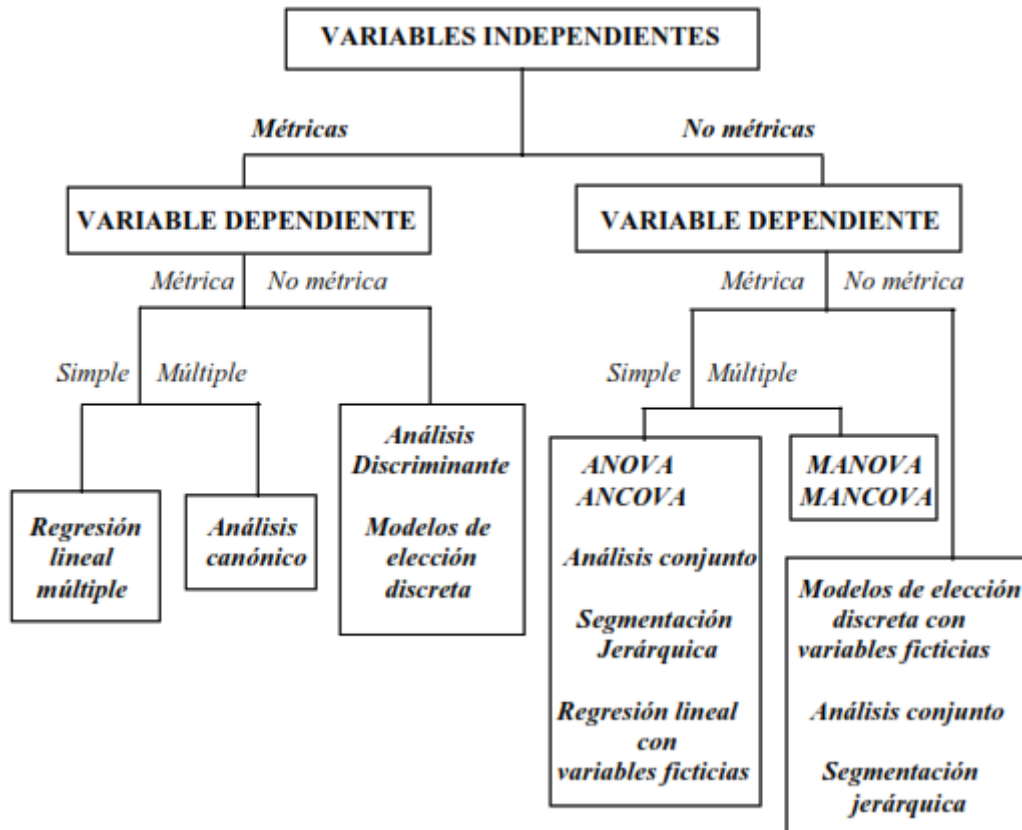


Figura 2-2. Técnicas de Análisis de Dependencia según la naturaleza de las variables.

Realizado por: Pérez, Cesar. 2004

2.2.10. Regresión Múltiple

El análisis de Regresión múltiple es una técnica estadística cuyo objetivo es explicar una variable endógena a través de uno o varios factores explicativos. Las variables independientes son de naturaleza métrica. Si existen variables exógenas cualitativas y éstas son relevantes dentro del análisis tienen que ser transformadas en parámetros métricos utilizando variables ficticias (Pérez, 2004), (Vilá et al., 2019).

En el análisis, si se dispone de una variable independiente que explica la variable dependiente la técnica utilizada es la regresión lineal simple. Ahora bien, cuando existen dos o más factores que describen la variable dependiente se recurre a la regresión lineal múltiple (Alcocer, 2020).

2.2.10.1. Consideraciones dentro del Análisis de Regresión Múltiple

Para determinar que el modelo esté correctamente especificado debe satisfacer las consideraciones de linealidad, aditividad, representatividad, multicolinealidad y principio de parsimonia.

Para determinar que el modelo esté correctamente especificado deben cumplirse los siguientes supuestos: independencia de residuos, homogeneidad, normalidad, homocedasticidad y multicolinealidad (Sierra, 2021), (Valle, 2018). Además, un modelo debe ser sencillo, es decir, cumplir con el principio de parsimonia (Gutiérrez-Cánovas & Escribano-Ávila, 2019), (Ruiz-Benito et al., 2018).

Los factores exógenos que explican la variable dependiente deben basarse en el principio de parsimonia, es decir, describir con la menor cantidad de parámetros el fenómeno analizado (Yunkor & Ochoa, 2021). Para (Camarero, 2004) la representatividad de las variables que detallan una situación no resulta ser un inconveniente desde la perspectiva matemática formal, sino mas bien la dificultad radica en la imposibilidad de conocer todos los aspectos que influyen de manera directa o indirecta en la explicación de la problemática. Por ello, se pretende delimitar el análisis y obtener variables que expliquen en un alto porcentaje el problema.

Para que el modelo explique de manera adecuada, es preciso que las variables explicativas sean linealmente independientes, esto es, no mantengan relación alguna entre sí. Cuando existe dependencia entre estos factores se denomina multicolinealidad (Coa, 2019). No obstante, se puede determinar un modelo sin considerar la multicolinealidad, sin embargo, el modelo obtenido tendrá estimadores imprecisos y carentes de significancia ante la imposibilidad de detectar por separado la influencia de los parámetros explicativos (Minchón, 2020). Para identificarla se puede calcular el Factor de Inflación de la Varianza (VIF). Si dicho valor es mayor o igual a diez se verifica la existencia de multicolinealidad (Soria, 2021).

El coeficiente de correlación lineal diagnostica el vínculo existente entre dos variables, siendo el mismo de carácter lineal, pero no determina ninguna relación si no es lineal (J. Hernández et al., 2018). Cabe resaltar que la relación no lineal de no ser detectada provoca efectos nocivos en el modelo y para eliminar tales consecuencias es necesario recurrir a un análisis de regresión no lineal (Pérez, 2004).

Para eliminar el problema causado por la multicolinealidad se emplea el Análisis de Componentes Principales (ACP). De igual manera, se puede optar por eliminar ciertas variables, justificando su

baja significancia e importancia estadística dentro del estudio o recurriendo a métodos específicos (Guerrero, 2018).

2.2.11. Técnicas de Análisis de Interdependencia

En el desarrollo de un estudio, en las etapas iniciales de la investigación, por lo general, se dispone de escasa información sobre el tema a abordar. Por ello, las personas involucradas extienden al máximo la recolección de datos a través de variables, que en un inicio según el criterio del investigador considera importantes. En consecuencia, se recogen datos de una gran cantidad de factores que al momento de analizarlos resulta poco práctico, invirtiendo una mayor cantidad de recursos en el análisis y a su vez dificultando la detección de parámetros que inciden de manera significativa (Meneses, 2019).

Con la finalidad de comprender la situación en el análisis por medio de variables representativas que expliquen de manera concisa y precisa, se han desarrollado técnicas permiten resumir las variables facilitando el análisis. Por ende, cuando los factores son cuantitativos, se emplean las técnicas de: Análisis Factorial, Análisis de Componentes Principales, Análisis de Conglomerados o Análisis de Escalamiento Multidimensional; siendo el más utilizado el primero que se ha mencionado. En cambio, cuando los parámetros son cualitativos se aplican las técnicas de: Análisis de Correspondencias, Análisis de Conglomerados o Análisis de Escalamiento Multidimensional (Chávez, 2018).

2.2.12. Análisis de Componentes Principales

El Análisis de Componentes Principales reduce la cantidad de variables, las cuales sintetizan la información de todos los factores iniciales (Pérez, 2004). Este criterio es compartido por (Hair et al., 1999). El ACP determina factores que expliquen la mayor parte de la varianza total (Pérez, 2004).

2.2.12.1. Supuestos del Análisis

Para (Hair et al., 1999) desde la perspectiva estadística, por lo general, no se considera la normalidad, homocedasticidad y linealidad entre variables, considerando que la omisión produce una reducción en la correlaciones analizadas. En este tipo de técnicas incluso se omite la multicolinealidad presente, ya que la finalidad es detectar la interrelación entre los factores. El análisis es inapropiado cuando el valor de las correlaciones es inferior a 0.3. Asimismo, si el

cálculo de las correlaciones parciales, considerando el efecto de las otras variables es alto, se concluye que el análisis es inadecuado.

2.2.13. Modelo Matemático

Un modelo matemático es un objeto que relaciona una parte del universo a través de una expresión matemática que defina alguna característica específica estudiada (Huinchahue et al., 2018). Para (Cardona et al., 2020) es una representación matemática de la realidad mediante relaciones de variable medibles. Según (R. V. Hernández, 2021), define a un modelo matemático como una ecuación o un conjunto de ellas que representan una realidad de manera aproximada, bajo ciertas consideraciones.

La modelación bajo ningún concepto es la aplicación de algoritmos, tampoco solucionar situaciones en la que se dispone de la totalidad de datos, sino más bien debe representar un reto para la persona que lo propone (Guerrero-Ortiz, y otros, 2021).

CAPÍTULO III

3. METODOLOGÍA DE LA INVESTIGACIÓN

3.1. Tipos de Investigación

Se comprende como investigación a un estudio basado en el cumplimiento de un procedimiento lógico y meditado destinado a responder un problema o inquietudes de carácter científico, a través de la generación de nuevo conocimiento o la ampliación del mismo, obteniendo una solución o respuesta a las interrogantes planteadas (F. Arias, 2012). Para (Carhuacho et al., 2019), con base en el pensamiento de Ander Egg define a la investigación como el propósito de descubrir, describir, explicar y/o interpretar los fenómenos que se muestran en un determinado aspecto de la realidad. (Espinoza Freire, 2019) cataloga a la investigación como un aspecto imprescindible que encamina al investigador a la respuesta de interrogantes en el ámbito social, económico, político y cultural constituyéndose en un factor esencial en el desarrollo profesional. En este sentido, existen diferentes modelos y por ende distintas clasificaciones.

Según (F. Arias, 2012), entre los tipos de investigación se destacan los siguientes:

- a. Tipos de investigación según el nivel
- b. Tipos de Investigación según el diseño
- c. Tipos de Investigación según el propósito

3.1.1. Tipos de investigación según el nivel

Se refiere al grado de profundidad con el cual se aborda un problema (F. Arias, 2012). Este tipo de estudio, según el autor se subdivide en:

- a. Investigación exploratoria: Analiza un tema poco estudiado obteniendo resultados superficiales aproximados (F. Arias, 2012). Se realiza con la finalidad de comprender las características principales del fenómeno (Ramos-Galarza, 2020).
- b. Investigación descriptiva: Consiste en la descripción de un hecho a fin describir su comportamiento, siendo de mayor profundidad que el exploratorio. Abarcan estudios de medición de variables independientes e investigaciones correlacionales (F. Arias, 2012). No se propone una hipótesis al igual que en la investigación exploratoria (Ortiz, 2015).

- c. Investigación explicativa: Involucra tanto una investigación post facto, buscando relaciones de causa – efecto y una investigación experimental determinando los efectos, mediante la comprobación de hipótesis. Es en definitiva es el nivel más profundo en búsqueda el conocimiento (F. Arias, 2012). Además, permite al investigador manipular las variables para analizar el efecto provocado (Guevara et al., 2020).

3.1.2. Tipos de Investigación según el diseño

Se denomina al planteamiento que decide el investigador para responder ante una situación. Se dividen en (F. Arias, 2012):

- a. Investigación documental: Basado en el análisis de datos de otros investigadores con el propósito de aportar un nuevo conocimiento. Los estudios que aportan a la investigación pueden proceder de fuentes impresas, audiovisuales y electrónicas (F. Arias, 2012).
- b. Investigación de campo: También llamada investigación no experimental. Se centra en la recolección datos de la mayor cantidad de factores sobre un suceso dado. Dicha información no es controlada ni manipulada. Se denomina extensiva cuando incluye toda la población en estudio e intensiva cuando se centra en casos específicos, imposibilitando la generalización de los resultados (F. Arias, 2012).
- c. Investigación experimental: Consiste en la observación de las reacciones o consecuencias que experimenta un objeto o conjunto de individuos a determinadas estímulos y/o condiciones (F. Arias, 2012). Su resultado es únicamente explicativo, estableciendo una relación de causa – efecto, probando de esta manera la hipótesis planteada (J. Arias & Covinos, 2021).

3.1.3. Tipos de Investigación según el propósito

Este tipo de investigaciones pretenden profundizar el conocimiento sobre un determinado tema o a su vez proponer modificaciones en el campo de análisis o teorías. Se clasifican en (J. Arias & Covinos, 2021):

- a. Investigación Básica: Se caracteriza por ser netamente teórica. Se fundamenta en el marco teórico, permaneciendo siempre en el campo conceptual. Su finalidad es profundizar e incrementar el conocimiento científico sin refutar con aspectos prácticos o empíricos (F. Arias, 2012).
- b. Investigación Aplicada: Conocida también como Investigación práctica o empírica. Depende de las conclusiones establecidas por la investigación básica, por ende, requiere de un marco

teórico. Por lo tanto, aplica el conocimiento teórico existente evaluando en la práctica las consecuencias (F. Arias, 2012).

En el presente estudio se aplicará una combinación de los diferentes tipos de investigación para dar respuesta a las diferentes interrogantes planteadas.

3.2. Método de la Investigación

Según (De Franco & Solorzano, 2020), puntualiza que el método es el cumplimiento estricto de un proceso lógico secuencial para generar conocimiento. En un aspecto muy general, un método constituye los pasos seguidos que conllevan al cumplimiento de un objetivo (F. Arias, 2012).

3.2.1. Pasos de la Investigación Científica

La investigación científica involucra conocimiento científico, de ahí la denominación de método científico. En este sentido, el método científico se constituye en un conglomerado de pasos, técnicas y procesos destinados a formular y responder a los problemas de investigación a través de la verificación de hipótesis (F. Arias, 2012).

Los pasos establecidos para llegar un resultado preciso, siguiendo la estructura del método científico se detallan a continuación y se muestran en la Figura 3.

3.2.1.1. Recolección y análisis de datos

El establecimiento de Educación Superior analizado, dispone del 37% de fichas socioeconómicas de los estudiantes legalmente matriculados. Para el análisis se limitará a analizar la información común para todos los alumnos.

3.2.1.2. Propuesta del Modelo Matemático

Según la Figura 1, se clasifican los diferentes tipos de análisis según la presencia de variables explicativas y/o explicadas. El tipo de análisis seleccionado será el que mejor se ajuste a las condiciones planteadas. Por lo tanto, dado que el objetivo de este estudio es obtener la tasa de deserción estudiantil, siendo ésta una variable explicada de carácter cuantitativa, así como los factores explicativos también lo son, entonces el método explicativo idóneo es regresión múltiple.

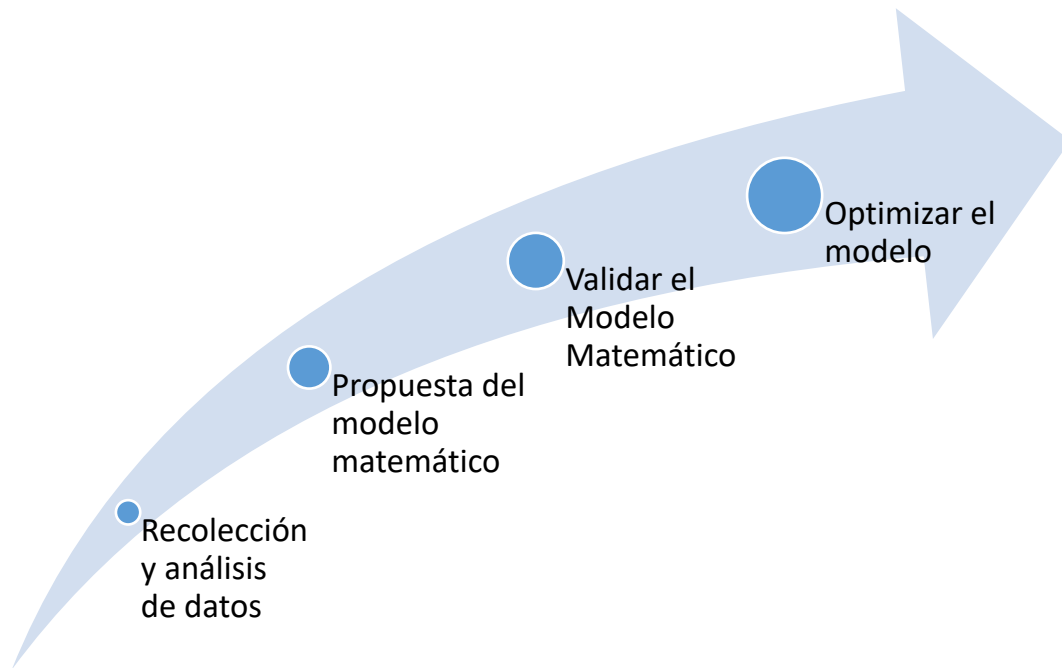


Figura 3-3. Pasos del Método Científico.

Realizado por: Ayabaca, Roberto. 2022

Para determinar la ecuación, que es el resultado de la aplicación del método que mejor se ajuste a las condiciones, se emplearán softwares especializados en el tratamiento y procesamiento de datos que permitan obtener una ecuación que prediga de la manera más precisa y exacta la deserción, aplicando los criterios que el investigador considere pertinentes.

3.2.1.3. Validar el Modelo Matemático

Una vez encontrada la ecuación deberá ser validada, es decir, el valor calculado tendrá que ser lo más preciso y exacto posible con relación a la tasa de deserción registrada. Por ello, se calculará el error entre el valor registrado y el valor calculado mediante la ecuación del modelo matemático. El investigador definirá si el error de cálculo es aceptable.

3.2.1.4. Optimizar el Modelo Matemático

Luego de establecer el modelo matemático, el mismo que ha sido validado se procederá a estimar los factores que deben manipularse para obtener una menor deserción, lo cual permitirá optimizar a un más el modelo existente, reduciendo variables que a priori se las pueda considerar significativas.

3.3. Enfoque de la investigación

El enfoque de la investigación será de carácter cuantitativo, es decir, cuantificar las variables que intervienen en el análisis para controlar y predecir. En tal sentido, identificar variables y comprobar hipótesis (Cortez Torrez, 2018).

3.4. Alcance investigativo

La presente investigación tiene como finalidad predecir la tasa de deserción de los estudiantes matriculados en un año calendario en la Universidad Regional Amazónica Ikiam empleando un modelo matemático basado en el análisis de Regresión lineal múltiple. De igual manera, detectar y jerarquizar los factores que inciden en el abandono utilizando el Análisis de Componentes Principales (ACP). Esto permitirá a las autoridades a establecer políticas y estrategias que conlleven a la reducción de la deserción estudiantil.

3.5. Selección del Software Estadístico

En la actualidad, existen a disposición una amplia variedad de softwares especializados en el tratamiento y procesamiento de datos lo cual imposibilita al investigador conocer al detalle cada uno de ellos. Por ende, se establecerán parámetros que le permitan al investigador seleccionar de manera adecuada un programa acorde a su conocimiento y habilidades. Dichos parámetros se valorarán en una escala de uno al cinco. La valoración de cada característica se detalla a continuación:

- a. Económico: Relacionado con el coste de adquisición del producto, así como licencias de vigencia. Siendo uno el software más caro y cinco el menos costoso.
- b. Reputación del software: Referido a la calidad del producto o al éxito de la empresa que lo provee. Siendo uno la menor calidad hasta cinco la máxima calidad.
- c. Especialización: el grado de experiencia del software para un uso determinado, es decir, cuenta con un menú detallado para el análisis. Siendo uno la menor especialización y cinco la máxima.
- d. Tipo de software: Referido si el software es cerrado (no es posible personalizar y se está supeditado a los cambios realizados por el proveedor) o abierto (el usuario puede realizar cambios). Siendo uno un software cerrado hasta cinco un software con código abierto.
- e. Conectividad: El software es compatible o vincula con otras herramientas. Siendo uno la mínima vinculación con otros programas hasta cinco la máxima compatibilidad.

Se ha seleccionado tres potenciales programas, en base al dominio del investigador. Estos son: Excel, Statistical Package for the Social Sciences (SPSS) y R. Los resultados mostrados en la Tabla 1 evidencian que SPSS, es un programa especializado en el análisis de datos enfocado a investigaciones en ciencias sociales y aplicadas, de alta reputación y confiabilidad, disponible para su adquisición y conocido por el investigador. Excel es un software desarrollado por Microsoft incorporado en el paquete de Office destinado a la manipulación de datos organizados en filas y columnas, altamente disponible y la información tratada puede ser manipulada por otro software. R es un software libre destinado al análisis estadístico. Utilizado principalmente en minería de datos, investigación biomédica, bioinformática y matemática financiera. Por lo tanto, se elige a Excel como el programa para “adecuar” los datos y SPSS para el análisis y obtención del Modelo Matemático.

Tabla 1-3. Selección del Software

| Programa Parámetro | Excel | SPSS | R |
|-----------------------|-----------|-----------|-----------|
| Económico | 4 | 4 | 5 |
| Reputación | 5 | 5 | 4 |
| Especialización | 1 | 5 | 3 |
| Tipo | 3 | 1 | 3 |
| Conectividad | 4 | 4 | 3 |
| TOTAL | 17 | 19 | 18 |

Realizado por: Ayabaca, Roberto. 2022

3.6. Recolección de Datos

Las técnicas de recolección de datos que se emplearán en la presente investigación serán: la encuesta socioeconómica de los estudiantes matriculados para un nivel en un año calendario y la entrevista dirigida a los estudiantes que abandonan sus estudios en el transcurso del mismo año.

3.6.1. Población de estudio

La población de estudio abarcará 2797 registros de estudiantes matriculados en todas las carreras que ofertó la Universidad Regional Amazónica Ikiam en los años 2017, 2018 y 2019. Del total de registros, 646 corresponden al 2017, 929 al año 2018 y 1222 al 2019.

3.6.2. Unidad de análisis

El estudio engloba a todos los estudiantes matriculados en todos los niveles de la totalidad de las carreras universitarias que ofertó la Universidad Regional Amazónica Ikiam en los años 2017, 2018 y 2019.

3.6.3. Selección de la muestra

Para la selección de la muestra se aplicará un muestreo no probabilístico ya que se analizarán factores cualitativos. En este sentido, la selección de los elementos de la muestra no depende de la probabilidad sino más bien de quien realiza la investigación. Por lo tanto, a criterio del investigador la muestra estará constituida por todos los registros que pertenecen a la población de estudio.

3.7. Medidas Estadísticas

En la búsqueda del modelo matemático que calcule de manera precisa la tasa de deserción estudiantil a partir de la información existente se analizará la dependencia de esta variable calculada a través de varios factores explicativos que a su vez deberán ser independientes entre sí, para garantizar un resultado confiable. Por lo tanto, se utilizarán técnicas comprobadas y ampliamente utilizadas en este tipo de estudios.

3.8. Identificación de variables

Según un análisis previo de la información existente en la Universidad con respecto a los estudiantes matriculados durante un año calendario. Se han detectado variables cualitativas, las cuales son: nivel, sexo, etnia, quintil, procedencia. De igual forma, se ha identificado una variable cuantitativa, siendo ésta la tasa de deserción estudiantil.

- a. Nivel. Se considera dos categorías: Nivelación (asocia a los estudiantes que cursa la Nivelación en la Universidad), y carrera (agrupa a los estudiantes que están cursando algún semestre dentro de una carrera universitaria).
- b. Género. Conformada por dos grupos diferenciados por la capacidad de producir gametos masculinos (hombre) o femeninos (mujer).
- c. Etnia. Agrupa a las personas según el grupo étnico, es decir, que provienen del mismo origen. Se establecen siete categorías: Afroecuatoriano, Blanco, Indígena, Mestizo, Montubio, Mulato y otro (aquellos que provienen de lugares muy distantes a nuestro país)

- d. Nivel socioeconómico. Según (Agualongo & Garcés, 2020) el término se utiliza para indicar la posición social de una persona, familia o comunidad. Determinado por dos componentes. Por un lado, el aspecto social que involucra características referentes a la educación, ocupación, procedencia étnica e historial familiar y comunitario; por otro lado, el factor económico del entorno donde se desenvuelve el individuo, familia o comunidad.
- e. Procedencia. Se refiere a la provincia de la cual proviene el estudiante. Por ende, se establecen veinticuatro categorías
- f. Tasa de deserción estudiantil. Es el porcentaje de estudiantes que han abandonado respecto de los estudiantes que permanecen en sus estudios.

3.9. Operacionalización De Variables

En la Tabla 2, se describe la operacionalización de las variables que se utilizarán en el presente trabajo investigativo:

Tabla 2-3. Operacionalización De Variables

| VARIABLE | TIPO DE VARIABLE | CONCEPTO | INDICADOR | DESCRIPCION | INSTRUMENTOS |
|-------------------------------|------------------|--|--|---|---|
| Nivel | Independiente | El nivel indica la antigüedad que tiene un estudiante en una carrera universitaria | Carrera Nivelación | Carrera agrupa los estudiantes que se hallan cursando una carrera profesional, mientras que nivelación agrupa a estudiantes nuevos que recién ingresan o repiten los primeros niveles. | Directa Observación Indirecta Entrevista |
| Género | Independiente | El sexo determina la condición sexual del individuo. Si produce gametos masculinos es hombre, y si produce gametos femeninos es mujer. | Masculino Femenino | Hombre agrupa a los individuos cuyos órganos reproductores producen gametos masculinos (espermatozoides) Mujer agrupa a los individuos cuyos órganos reproductores producen gametos femeninos (óvulos) | Directa Observación Indirecta Entrevista |
| Etnia | Independiente | Identifica al estudiante a la etnia o pueblo al que pertenece. | Afroecuatoriano Blanco Mestizo Indígena Montubio Mulato Otro | Se enlista las “principales” etnias presentes en la República del Ecuador. | Directa Observación Indirecta Entrevista |
| Nivel socioeconómico | Independiente | Indica el nivel económico del cual procede o es parte el estudiante. | N1 N2 N3 N4 N5 | Representa la posición social específica de un estudiante respecto al aspecto social y económico; siendo N1 el nivel más bajo y N5 nivel más alto | Directa Observación Indirecta Entrevista |
| Tasa de deserción estudiantil | Dependiente | El porcentaje de estudiantes que desertan de sus estudios en un año. | Tasa de deserción estudiantil | Valor obtenido aplicando un modelo matemático basado en el método de regresión Lineal Múltiple. | Excel SPSS |

Fuente: IPEC, 2021

Realizado por: Ayabaca, Roberto. 2022

En la Tabla 3, se describe la matriz de consistencia que se utilizara en el presente trabajo investigativo

Tabla 3-3. Matriz De Consistencia

| FORMULACIÓN DEL PROBLEMA | OBJETIVOS | HIPÓTESIS | Variables | Indicadores | Técnicas | Instrumentos |
|--|--|---|---|-------------------------------|---|--|
| GENERAL | | | | | | |
| ¿Cuál es la tasa de deserción estudiantil, mediante la aplicación de un modelo Matemático, basado en el método regresión lineal múltiple de variables ficticias y/o dicotómicas con la información existente de los estudiantes matriculados en la Universidad Regional Amazónica Ikiam? | Desarrollar una propuesta de modelo matemático a través del procesamiento de información existente en la Universidad para el análisis de deserción estudiantil en la Universidad Regional Amazónica - Ikiam. | El modelamiento de la tasa de deserción estudiantil, a partir de los datos disponibles predice con precisión el abandono estudiantil en la Universidad Regional Amazónica Ikiam | VARIABLE DEPENDIENTE Tasa de deserción estudiantil VARIABLE INDEPENDIENTE Nivel, Sexo, Etnia, Nivel socioeconómico | Tasa de deserción estudiantil | Análisis de Componentes Principales. Regresión Múltiple Lineal | Observación Directa Observación Indirecta Entrevista |
| ESPECIFICOS | | | | | | |
| ¿Cuáles son los factores más importantes asociados a la deserción estudiantil, en base a la información existente de los estudiantes matriculados en la Universidad Regional Amazónica Ikiam? | Recopilar, analizar y procesar la información existente desde la creación de la Universidad Regional Amazónica Ikiam | El modelo matemático propuesto identifica las principales causas que conllevan a la deserción estudiantil dentro del establecimiento Educativo | VARIABLE INDEPENDIENTE Nivel, Sexo, Etnia, Nivel socioeconómico | Correlación entre variables | Análisis de Componentes Principales | Observación Directa Observación Indirecta Entrevista |
| ¿Se puede determinar la tasa de deserción estudiantil mediante la aplicación de un modelo matemático basado en el método de regresión lineal múltiple, con la información existente de | Proponer el modelo matemático para cuantificar la deserción estudiantil en la Universidad Regional Amazónica Ikiam | La información existente desde la creación de la Universidad es suficiente para establecer un modelo matemático que pronostique con precisión la deserción estudiantil | VARIABLE INDEPENDIENTE Nivel, Sexo, Etnia, Nivel socioeconómico VARIABLE DEPENDIENTE | Tasa de deserción estudiantil | Regresión Múltiple Lineal | Observación Directa Observación Indirecta Entrevista |

| | | | | | | |
|--|--|--|---|---|--------------------------|--|
| los estudiantes matriculados en la Universidad Regional Amazónica Ikiam? | | | Tasa de deserción estudiantil | | | |
| ¿Cuál es el error entre la tasa de deserción calculado aplicando el modelo matemático y el valor real? | Realizar pruebas de validación del modelo matemático | El modelo matemático cuantifica con precisión la deserción estudiantil en la Universidad Regional Amazónica Ikiam | VARIABLE INDEPENDIENTE Tasa de deserción calculada, Tasa de deserción real. VARIABLE DEPENDIENTE Error | Error de la tasa de deserción calculada respecto de la tasa de deserción real | Error relativo | Observación Directa Observación Indirecta Entrevista |
| ¿Qué factores explicativos se deberían reducir para iniciar una disminución en la tasa de deserción calculada mediante el modelo matemático? | Optimizar del Modelo Matemático | La manipulación de los factores intervinientes en el modelo matemático permite la disminución de la tasa de deserción. | VARIABLE INDEPENDIENTE Nivel, Sexo, Etnia, Nivel socioeconómico VARIABLE DEPENDIENTE Tasa de deserción estudiantil | Tasa de deserción Mínima | Análisis de Optimización | Observación Directa Observación Indirecta Entrevista |

Fuente: IPEC, 2021

Realizado por: Ayabaca, Roberto. 2022

3.10. Adecuación de la información

La información procesada con base en los datos proporcionados por la Universidad Regional Ikiám sobre los estudiantes matriculados y de aquellos que abandonaron el establecimiento. El proceso seguido es el siguiente:

- a. Seleccionar los años de análisis en que se disponga la mayor cantidad de información y que la misma sea común.
- b. Crear un nuevo libro en Microsoft Excel con la información seleccionada de los años 2017, 2018 y 2019.
- c. Ordenar la información de los diferentes años bajo parámetros comunes.
- d. Completar información faltante según el número de cédula del estudiante.
- e. Transformar las variables cualitativas en cuantitativas aplicando el concepto de variable ficticia.
- f. Consolidar la información y adecuarla considerando la provincia de procedencia.
- g. Aplicar la técnica de Análisis de Componentes Principales para determinar los factores más influyentes.
- h. Aplicar la técnica de Regresión Lineal Múltiple para determinar una ecuación que permita calcular la tasa de deserción estudiantil.
- i. Verificar la precisión en la predicción del modelo

CAPÍTULO IV

4.1. RESULTADOS Y ANÁLISIS

En la sección anterior se estableció parámetros comunes en todos los registros disponibles de la institución a lo largo de los años. La información se ha adecuado según la procedencia de los estudiantes matriculados. En el Gráfico 5, se evidencia que los alumnos provienen de las 24 provincias del Ecuador, siendo Napo la provincia que más aporta, seguida de Pichincha.

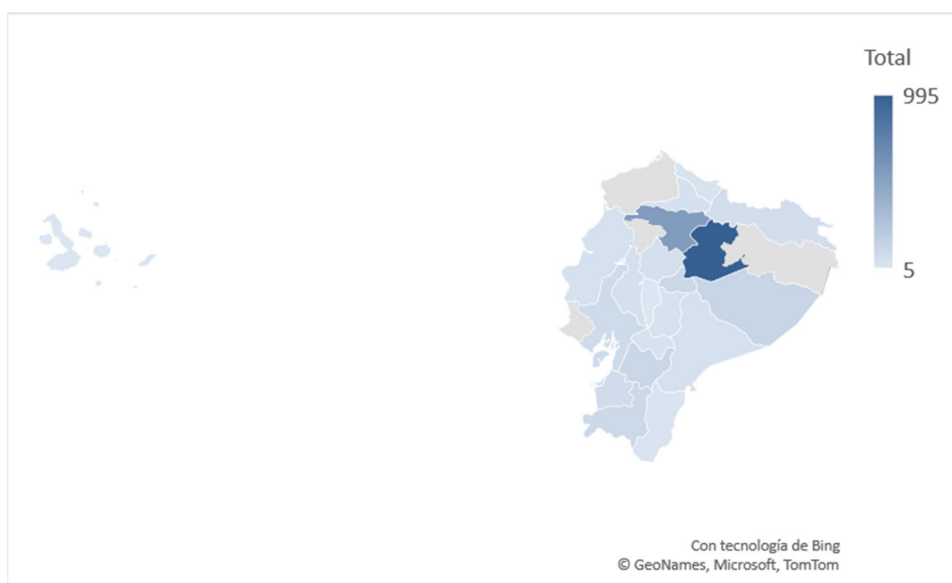


Gráfico 5-4. Procedencia de los estudiantes

Realizado por: Ayabaca, Roberto. 2022

Este filtro realizado indica que la distancia no necesariamente incide de manera directa en el abandono escolar, sino más bien son otros factores que afectan en mayor medida en la decisión de desertar. Por lo tanto, a criterio del investigador la información se ha adecuado según este parámetro.

En base a la información proporcionada de los años en estudio, se determina que una mayoría importante de los estudiantes se autoidentifican como mestizo, seguido por la etnia indígena y en menor escala, muy escasa, de otras etnias (Afroecuatoriano, blanco, mulato, montubio, otro). En el Gráfico 6, se muestra la distribución según la autoidentificación étnica. Este factor incide en un potencial abandono, puesto que son los estudiantes indígenas quienes registran la mayor cantidad de abandono. La mayoría de los alumnos indígenas pertenece, en su mayoría a los tres primeros niveles socioeconómicos establecidos por la institución. En el primer nivel socioeconómico se concentra cerca del 40% de estudiantes autoidentificados como indígenas.

Adicional, el 60% de alumnos indígenas matriculados cursan la nivelación y de ellos el 85% deserta, por falta de conocimiento, por desinterés o falta de recursos económicos. Los mestizos que representan el mayor sector de los educandos del establecimiento educativo, desertan en mayor proporción aquellos que cursan los niveles iniciales, debido fundamentalmente porque la carrera no es acorde a sus intereses y la falta de recursos económicos para solventar los gastos que incurre una carrera universitaria fuera de su lugar de residencia.

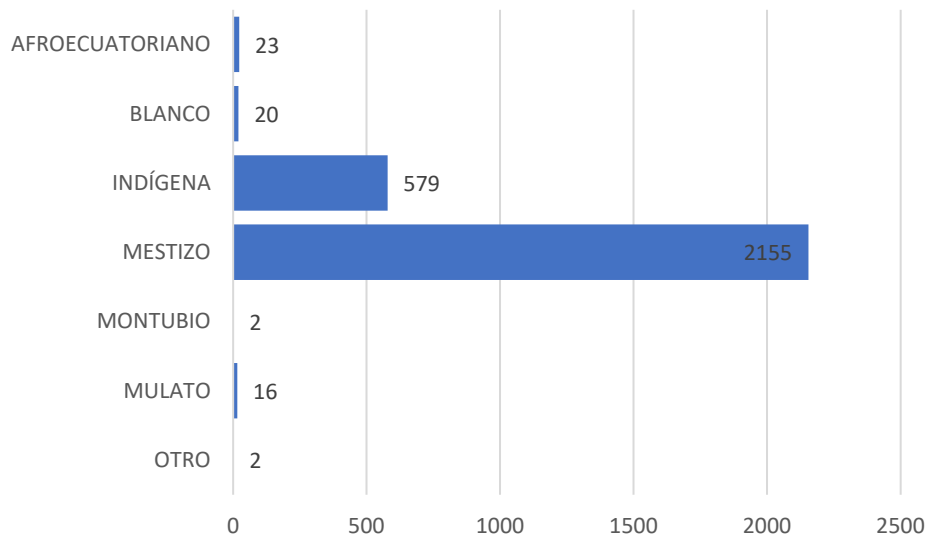


Gráfico 6-4. Autoidentificación étnica

Realizado por: Ayabaca, Roberto. 2022

En la Tabla 4, se observa la agrupación realizada tomando como referencia el lugar de procedencia, es decir, la provincia de la cual proviene. Los factores conocidos son: nivel, sexo, etnia, nivel socioeconómico y condición actual (el estudiante estudia o ha abandonado sus estudios). Se han establecido nuevas variables cuantitativas, producto de la transformación de las diferentes categorías contenidas en cada variable principal, utilizando el concepto de variable ficticia.

En la Tabla 5, se muestra la relación entre el valor de cada variable respecto del total de estudiantes pertenecientes a cada provincia. El resultado se muestra en porcentaje. Se evidencia que las provincias con mayor índice de abandono son: Zamora Chinchipe, Orellana y Santa Elena; seguidas de: Napo, Cotopaxi, Los Ríos, Pastaza y Esmeraldas que bordea el 8%. En contraste, las provincias en las cuales el abandono es nulo son: Bolívar, Carchi, Chimborazo, Imbabura y Galápagos.

Tabla 4-4. Factores asociados agrupados por provincia de procedencia

| PROVINCIA | NIVEL | | GÉNERO | | ETNIA | | | | | | | NIVEL SOCIOECONÓMICO | | | | | Abandona |
|-----------------|---------|------------|-----------|----------|-----------------|--------|---------|----------|----------|--------|------|----------------------|-----|-----|----|----|----------|
| | Carrera | Nivelación | Masculino | Femenino | Afroecuatoriano | Blanco | Mestizo | Indígena | Montubio | Mulato | Otro | N1 | N2 | N3 | N4 | N5 | |
| Azuay | 97 | 10 | 58 | 58 | 1 | 0 | 113 | 2 | 0 | 0 | 0 | 12 | 21 | 67 | 14 | 2 | 4 |
| Bolívar | 3 | 3 | 2 | 5 | 0 | 0 | 7 | 0 | 0 | 0 | 0 | 1 | 1 | 5 | 0 | 0 | 0 |
| Cañar | 28 | 4 | 18 | 17 | 0 | 0 | 33 | 2 | 0 | 0 | 0 | 7 | 3 | 21 | 0 | 4 | 1 |
| Carchi | 19 | 5 | 10 | 18 | 0 | 0 | 28 | 0 | 0 | 0 | 0 | 1 | 5 | 14 | 6 | 2 | 0 |
| Cotopaxi | 31 | 10 | 19 | 28 | 0 | 0 | 46 | 1 | 0 | 0 | 0 | 12 | 14 | 19 | 1 | 1 | 4 |
| Chimborazo | 27 | 5 | 22 | 15 | 0 | 0 | 35 | 2 | 0 | 0 | 0 | 12 | 6 | 13 | 6 | 0 | 0 |
| El Oro | 57 | 11 | 47 | 29 | 0 | 0 | 75 | 1 | 0 | 0 | 0 | 11 | 20 | 37 | 8 | 0 | 2 |
| Esmeraldas | 28 | 4 | 16 | 21 | 5 | 0 | 31 | 0 | 0 | 1 | 0 | 9 | 15 | 10 | 3 | 0 | 3 |
| Guayas | 60 | 10 | 36 | 51 | 0 | 0 | 86 | 0 | 0 | 1 | 0 | 10 | 25 | 27 | 19 | 6 | 5 |
| Imbabura | 31 | 10 | 21 | 20 | 0 | 1 | 39 | 1 | 0 | 0 | 0 | 9 | 12 | 15 | 2 | 3 | 0 |
| Loja | 69 | 29 | 52 | 52 | 0 | 2 | 100 | 2 | 0 | 0 | 0 | 10 | 20 | 51 | 16 | 7 | 5 |
| Los Ríos | 39 | 8 | 22 | 26 | 0 | 0 | 48 | 0 | 0 | 0 | 0 | 6 | 8 | 23 | 9 | 2 | 4 |
| Manabí | 43 | 1 | 28 | 17 | 0 | 0 | 45 | 0 | 0 | 0 | 0 | 7 | 8 | 21 | 5 | 4 | 1 |
| Morona Santiago | 16 | 15 | 11 | 29 | 0 | 0 | 36 | 4 | 0 | 0 | 0 | 20 | 6 | 12 | 2 | 0 | 1 |
| Napo | 371 | 495 | 464 | 531 | 8 | 9 | 482 | 491 | 0 | 4 | 1 | 391 | 201 | 306 | 91 | 6 | 97 |
| Pastaza | 51 | 72 | 71 | 60 | 0 | 1 | 107 | 22 | 0 | 1 | 0 | 33 | 17 | 68 | 7 | 6 | 11 |
| Pichincha | 381 | 139 | 270 | 283 | 4 | 4 | 519 | 22 | 0 | 3 | 1 | 41 | 132 | 289 | 78 | 13 | 29 |
| Tungurahua | 71 | 32 | 53 | 57 | 1 | 2 | 105 | 2 | 0 | 0 | 0 | 9 | 22 | 54 | 14 | 11 | 8 |

| | | | | | | | | | | | | | | | | | |
|--------------------------------|----|----|----|----|---|---|----|----|---|---|---|----|----|----|----|---|----|
| Zamora Chinchipe | 19 | 4 | 19 | 4 | 0 | 0 | 18 | 5 | 0 | 0 | 0 | 4 | 8 | 9 | 2 | 0 | 3 |
| Galápagos | 3 | 2 | 3 | 2 | 1 | 0 | 4 | 0 | 0 | 0 | 0 | 1 | 1 | 2 | 1 | 0 | 0 |
| Sucumbíos | 32 | 27 | 36 | 30 | 1 | 0 | 58 | 5 | 0 | 2 | 0 | 26 | 15 | 20 | 4 | 1 | 4 |
| Orellana | 50 | 47 | 49 | 60 | 2 | 0 | 88 | 17 | 0 | 2 | 0 | 38 | 15 | 40 | 11 | 5 | 12 |
| Santo Domingo de los Tsáchilas | 25 | 9 | 21 | 17 | 0 | 1 | 35 | 0 | 2 | 0 | 0 | 5 | 7 | 20 | 6 | 0 | 2 |
| Santa Elena | 15 | 3 | 6 | 13 | 0 | 0 | 17 | 0 | 0 | 2 | 0 | 2 | 5 | 8 | 1 | 3 | 2 |

Fuente: Universidad Regional Amazónica IKIAM

Realizado por: AYABACA, Roberto.2021

Tabla 5-4. Tasa de incidencia de las variables respecto a la procedencia

| Provincia | NIVEL | | SEXO | | ETNIA | | | | | | | NIVEL SOCIOECONÓMICO | | | | | Abandona |
|------------|---------|------------|--------|-------|-----------------|--------|---------|----------|----------|--------|------|----------------------|-------|-------|-------|-------|----------|
| | Carrera | Nivelación | Hombre | Mujer | Afroecuatoriano | Blanco | Mestizo | Indígena | Montubio | Mulato | Otro | N1 | N2 | N3 | N4 | N5 | |
| Azuay | 83,62 | 8,62 | 50,00 | 50,00 | 0,86 | 0,00 | 97,41 | 1,72 | 0,00 | 0,00 | 0,00 | 10,34 | 18,10 | 57,76 | 12,07 | 1,72 | 3,45 |
| Bolívar | 42,86 | 42,86 | 28,57 | 71,43 | 0,00 | 0,00 | 100,0 | 0,00 | 0,00 | 0,00 | 0,00 | 14,29 | 14,29 | 71,43 | 0,00 | 0,00 | 0,00 |
| Cañar | 80,00 | 11,43 | 51,43 | 48,57 | 0,00 | 0,00 | 94,29 | 5,71 | 0,00 | 0,00 | 0,00 | 20,00 | 8,57 | 60,00 | 0,00 | 11,43 | 2,86 |
| Carchi | 67,86 | 17,86 | 35,71 | 64,29 | 0,00 | 0,00 | 100,0 | 0,00 | 0,00 | 0,00 | 0,00 | 3,57 | 17,86 | 50,00 | 21,43 | 7,14 | 0,00 |
| Cotopaxi | 65,96 | 21,28 | 40,43 | 59,57 | 0,00 | 0,00 | 97,87 | 2,13 | 0,00 | 0,00 | 0,00 | 25,53 | 29,79 | 40,43 | 2,13 | 2,13 | 8,51 |
| Chimborazo | 72,97 | 13,51 | 59,46 | 40,54 | 0,00 | 0,00 | 94,59 | 5,41 | 0,00 | 0,00 | 0,00 | 32,43 | 16,22 | 35,14 | 16,22 | 0,00 | 0,00 |
| El Oro | 75,00 | 14,47 | 61,84 | 38,16 | 0,00 | 0,00 | 98,68 | 1,32 | 0,00 | 0,00 | 0,00 | 14,47 | 26,32 | 48,68 | 10,53 | 0,00 | 2,63 |
| Esmeraldas | 75,68 | 10,81 | 43,24 | 56,76 | 13,51 | 0,00 | 83,78 | 0,00 | 0,00 | 2,70 | 0,00 | 24,32 | 40,54 | 27,03 | 8,11 | 0,00 | 8,11 |
| Guayas | 68,97 | 11,49 | 41,38 | 58,62 | 0,00 | 0,00 | 98,85 | 0,00 | 0,00 | 1,15 | 0,00 | 11,49 | 28,74 | 31,03 | 21,84 | 6,90 | 5,75 |

| | | | | | | | | | | | | | | | | | |
|--------------------------------|-------|-------|-------|-------|-------|------|-------|-------|------|-------|------|-------|-------|-------|-------|-------|-------|
| Imbabura | 75,61 | 24,39 | 51,22 | 48,78 | 0,00 | 2,44 | 95,12 | 2,44 | 0,00 | 0,00 | 0,00 | 21,95 | 29,27 | 36,59 | 4,88 | 7,32 | 0,00 |
| Loja | 66,35 | 27,88 | 50,00 | 50,00 | 0,00 | 1,9 | 96,15 | 1,92 | 0,00 | 0,00 | 0,00 | 9,62 | 19,23 | 49,04 | 15,38 | 6,73 | 4,81 |
| Los Ríos | 81,25 | 16,67 | 45,83 | 54,17 | 0,00 | 0,00 | 100,0 | 0,00 | 0,00 | 0,00 | 0,00 | 12,50 | 16,67 | 47,92 | 18,75 | 4,17 | 8,33 |
| Manabí | 95,56 | 2,22 | 62,22 | 37,78 | 0,00 | 0,00 | 100,0 | 0,00 | 0,00 | 0,00 | 0,00 | 15,56 | 17,78 | 46,67 | 11,11 | 8,89 | 2,22 |
| Morona Santiago | 40,00 | 37,50 | 27,50 | 72,50 | 0,00 | 0,00 | 90,00 | 10,00 | 0,00 | 0,00 | 0,00 | 50,00 | 15,00 | 30,00 | 5,00 | 0,00 | 2,50 |
| Napo | 37,29 | 49,75 | 46,63 | 53,37 | 0,80 | 0,90 | 48,44 | 49,35 | 0,00 | 0,40 | 0,10 | 39,30 | 20,20 | 30,75 | 9,15 | 0,60 | 9,75 |
| Pastaza | 38,93 | 54,96 | 54,20 | 45,80 | 0,00 | 0,76 | 81,68 | 16,79 | 0,00 | 0,76 | 0,00 | 25,19 | 12,98 | 51,91 | 5,34 | 4,58 | 8,40 |
| Pichincha | 68,90 | 25,14 | 48,82 | 51,18 | 0,72 | 0,72 | 93,85 | 3,98 | 0,00 | 0,54 | 0,18 | 7,41 | 23,87 | 52,26 | 14,10 | 2,35 | 5,24 |
| Tungurahua | 64,55 | 29,09 | 48,18 | 51,82 | 0,91 | 1,82 | 95,45 | 1,82 | 0,00 | 0,00 | 0,00 | 8,18 | 20,00 | 49,09 | 12,73 | 10,00 | 7,27 |
| Zamora Chinchipe | 82,61 | 17,39 | 82,61 | 17,39 | 0,00 | 0,00 | 78,26 | 21,74 | 0,00 | 0,00 | 0,00 | 17,39 | 34,78 | 39,13 | 8,70 | 0,00 | 13,04 |
| Galápagos | 60,00 | 40,00 | 60,00 | 40,00 | 20,00 | 0,00 | 80,00 | 0,00 | 0,00 | 0,00 | 0,00 | 20,00 | 20,00 | 40,00 | 20,00 | 0,00 | 0,00 |
| Sucumbíos | 48,48 | 40,91 | 54,55 | 45,45 | 1,52 | 0,00 | 87,88 | 7,58 | 0,00 | 3,03 | 0,00 | 39,39 | 22,73 | 30,30 | 6,06 | 1,52 | 6,06 |
| Orellana | 45,87 | 43,12 | 44,95 | 55,05 | 1,83 | 0,00 | 80,73 | 15,60 | 0,00 | 1,83 | 0,00 | 34,86 | 13,76 | 36,70 | 10,09 | 4,59 | 11,01 |
| Santo Domingo de los Tsáchilas | 65,79 | 23,68 | 55,26 | 44,74 | 0,00 | 2,63 | 92,11 | 0,00 | 5,26 | 0,00 | 0,00 | 13,16 | 18,42 | 52,63 | 15,79 | 0,00 | 5,26 |
| Santa Elena | 78,95 | 15,79 | 31,58 | 68,42 | 0,00 | 0,00 | 89,47 | 0,00 | 0,00 | 10,53 | 0,00 | 10,53 | 26,32 | 42,11 | 5,26 | 15,79 | 10,53 |

Fuente: Universidad Regional Amazónica IKIAM

Realizado por: AYABACA, Roberto.2021

4.1. Adecuación de la Información

El objetivo radica en determinar una ecuación matemática que prediga y/o pronostique la deserción estudiantil. La variable explicada es cuantitativa al igual que las variables explicativas como se observa en la Tabla 5. El método idóneo para hallar la expresión matemática es la regresión lineal múltiple, como ya se ha explicado en secciones anteriores. De igual manera, se mencionó que se deben tomar ciertas consideraciones para que el producto estime de manera exacta y precisa. A continuación, se analizará dichos parámetros:

4.1.1. Componentes factoriales

Un modelo matemático debe ser explicado por la menor cantidad de factores, constituyéndose una expresión sencilla de fácil manipulación. Esto se conoce como Principio de Parsimonia. Para lograrlo, es necesario que las variables explicativas sean lo más independientes posible entre ellas, es decir, que la multicolinealidad sea evitada o mínima. Para conseguirlo se aplica el Análisis de Componentes Principales (ACP).

Para comprobar el grado de asociación de las variables, se calcula el determinante de la matriz de correlaciones. Para calcular cada coeficiente de correlación entre dos factores explicativos se utiliza la ecuación (1)

$$r_{xy} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}} \quad (1)$$

Dónde, n , es el tamaño de la muestra y x_i e y_i son los puntos muestrales individuales vinculados con i .

En la Tabla 6, se muestra los coeficientes de correlación entre dos variables, para una muestra de 24 registros.

Tabla 6-4. Matriz de correlaciones de todos los factores asociados

| | Carrera | Nivelación | Masculino | Femenino | Afroecuatoriano | Blanco | Mestizo | Indígena | Montubio | Mulato | Otro | N1 | N2 | N3 | N4 | N5 |
|-----------------|---------|------------|-----------|----------|-----------------|--------|---------|----------|----------|--------|-------|-------|-------|-------|-------|-------|
| Carrera | 1,00 | -0,93 | 0,38 | -0,38 | -0,03 | -0,07 | 0,49 | -0,50 | 0,00 | 0,06 | -0,15 | -0,58 | 0,33 | 0,13 | 0,23 | 0,35 |
| Nivelación | -0,93 | 1,00 | -0,18 | 0,18 | 0,11 | 0,19 | -0,56 | 0,53 | -0,02 | -0,05 | 0,18 | 0,49 | -0,33 | -0,06 | -0,23 | -0,28 |
| Hombre | 0,38 | -0,18 | 1,00 | -1,00 | 0,11 | 0,10 | -0,17 | 0,18 | 0,11 | -0,31 | -0,02 | -0,07 | 0,18 | -0,06 | 0,18 | -0,23 |
| Mujer | -0,38 | 0,18 | -1,00 | 1,00 | -0,11 | -0,10 | 0,17 | -0,18 | -0,11 | 0,31 | 0,02 | 0,07 | -0,18 | 0,06 | -0,18 | 0,23 |
| Afroecuatoriano | -0,03 | 0,11 | 0,11 | -0,11 | 1,00 | -0,17 | -0,28 | -0,13 | -0,07 | 0,04 | -0,06 | 0,08 | 0,27 | -0,28 | 0,21 | -0,28 |
| Blanco | -0,07 | 0,19 | 0,10 | -0,10 | -0,17 | 1,00 | -0,04 | 0,02 | 0,54 | -0,19 | 0,11 | -0,19 | 0,00 | 0,11 | 0,08 | 0,12 |
| Mestizo | 0,49 | -0,56 | -0,17 | 0,17 | -0,28 | -0,04 | 1,00 | -0,89 | 0,03 | -0,12 | -0,34 | -0,53 | -0,09 | 0,44 | 0,13 | 0,27 |
| Indígena | -0,50 | 0,53 | 0,18 | -0,18 | -0,13 | 0,02 | -0,89 | 1,00 | -0,12 | -0,07 | 0,38 | 0,54 | -0,07 | -0,32 | -0,21 | -0,24 |
| Montubio | 0,00 | -0,02 | 0,11 | -0,11 | -0,07 | 0,54 | 0,03 | -0,12 | 1,00 | -0,08 | -0,06 | -0,12 | -0,08 | 0,17 | 0,17 | -0,19 |
| Mulato | 0,06 | -0,05 | -0,31 | 0,31 | 0,04 | -0,19 | -0,12 | -0,07 | -0,08 | 1,00 | -0,05 | -0,01 | 0,26 | -0,23 | -0,20 | 0,48 |
| Otro | -0,15 | 0,18 | -0,02 | 0,02 | -0,06 | 0,11 | -0,34 | 0,38 | -0,06 | -0,05 | 1,00 | -0,03 | 0,05 | 0,01 | 0,08 | -0,15 |
| N1 | -0,58 | 0,49 | -0,07 | 0,07 | 0,08 | -0,19 | -0,53 | 0,54 | -0,12 | -0,01 | -0,03 | 1,00 | -0,11 | -0,60 | -0,40 | -0,40 |
| N2 | 0,33 | -0,33 | 0,18 | -0,18 | 0,27 | 0,00 | -0,09 | -0,07 | -0,08 | 0,26 | 0,05 | -0,11 | 1,00 | -0,50 | 0,01 | -0,15 |
| N3 | 0,13 | -0,06 | -0,06 | 0,06 | -0,28 | 0,11 | 0,44 | -0,32 | 0,17 | -0,23 | 0,01 | -0,60 | -0,50 | 1,00 | -0,13 | 0,16 |
| N4 | 0,23 | -0,23 | 0,18 | -0,18 | 0,21 | 0,08 | 0,13 | -0,21 | 0,17 | -0,20 | 0,08 | -0,40 | 0,01 | -0,13 | 1,00 | -0,07 |
| N5 | 0,35 | -0,28 | -0,23 | 0,23 | -0,28 | 0,12 | 0,27 | -0,24 | -0,19 | 0,48 | -0,15 | -0,40 | -0,15 | 0,16 | -0,07 | 1,00 |

Realizado por: Ayabaca, Roberto. 2022.

El determinante de la matriz es cero, por ende, es una matriz singular, lo cual indica que existen variables que son linealmente dependientes entre sí.

Ante esta situación, el investigador deduce que ciertas variables pueden descartarse, previo a un análisis de la información.

La variable “Nivel” que se subdivide en dos nuevas variables ficticias (Nivelación y Carrera) son dicotómicas, es decir, que la presencia de una de ellas automáticamente ausenta a la otra. Por ende, al observar el Gráfico 7, se considera la variable “Nivelación” ya que es precisamente en la nivelación donde se registra la mayor deserción, descartando el parámetro “Carrera”.

Nivel de estudio de los estudiantes

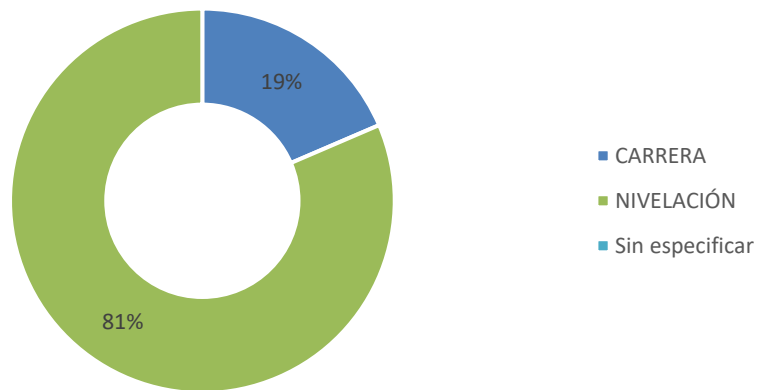


Gráfico 7-4. Nivel de estudio de los estudiantes.

Fuente: Universidad Regional Amazónica IKIAM

Realizado por: Ayabaca, Roberto. 2022

El factor “Género” es una variable dicotómica, ya que si un alumno es masculino anula la otra categoría (femenino). Las mujeres desertan en mayor cantidad que los hombres, aunque la diferencia no es tan significativa, tal y como se evidencia en el Gráfico 8. Por lo tanto, se optará por la variable “Femenino”, descartando el parámetro “Masculino”.

Género de los estudiantes

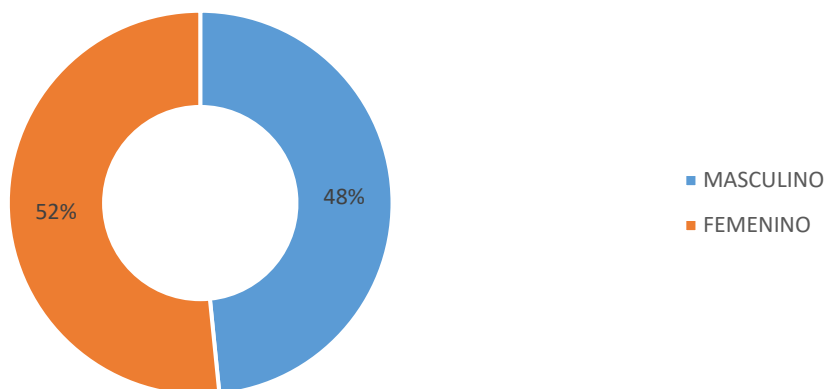


Gráfico 8-4. Género de los estudiantes.

Fuente: Universidad Regional Amazónica IKIAM

Realizado por: Ayabaca, Roberto. 2022

Nivel socioeconómico de los estudiantes

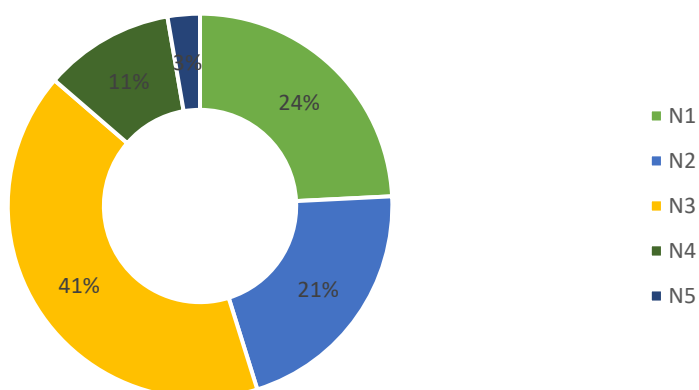


Gráfico 9-4. Nivel socioeconómico de los estudiantes.

Fuente: Universidad Regional Amazónica IKIAM

Realizado por: Ayabaca, Roberto. 2022

Los parámetros ficticios derivados del factor “Etnia” más relevantes son “Indígena” y “Mestizo”, como se observa en el Gráfico 6. De igual forma, las variables ficticias obtenidas del “Nivel socioeconómico”, que mayor afectación tienen sobre la deserción son los tres primeros niveles socioeconómicos más bajos llegando a abarcar el 85% del alumnado. Esto se pone en relieve en el Gráfico 9.

Por lo tanto, en la Tabla 7, se muestra que, de las dieciséis variables adecuadas inicialmente, se ha reducido a siete. El determinante obtenido es 0,0147 distinto de cero, y al ser cercano a cero existen correlaciones altas entre las variables en estudio. De estas se procurará reducir a un más, aplicando el Análisis de Componentes Principales.

Tabla 7-4. Matriz de correlaciones reducida

| | Nivelación | Femenino | Mestizo | Indígena | N1 | N2 | N3 |
|------------|------------|----------|---------|----------|-------|-------|-------|
| Nivelación | 1,00 | 0,18 | -0,56 | 0,53 | 0,49 | -0,33 | -0,06 |
| Femenino | 0,18 | 1,00 | 0,17 | -0,18 | 0,07 | -0,18 | 0,06 |
| Mestizo | -0,56 | 0,17 | 1,00 | -0,89 | -0,53 | -0,09 | 0,44 |
| Indígena | 0,53 | -0,18 | -0,89 | 1,00 | 0,54 | -0,07 | -0,32 |
| N1 | 0,49 | 0,07 | -0,53 | 0,54 | 1,00 | -0,11 | -0,60 |
| N2 | -0,33 | -0,18 | -0,09 | -0,07 | -0,11 | 1,00 | -0,50 |
| N3 | -0,06 | 0,06 | 0,44 | -0,32 | -0,60 | -0,50 | 1,00 |

Realizado por: Ayabaca, Roberto. 2022.

4.1.1.1. Prueba de esfericidad de Bartlett

La prueba de esfericidad de Bartlett χ^2 se utiliza para determinar el grado de ausencia o no de correlación entre las variables. Cuando $\chi^2 > \chi^2_{(\alpha, t-1)}$ se rechaza la hipótesis nula, es decir, existe correlación entre las variables. Siendo $\chi^2_{(\alpha, t-1)}$ la prueba de Chi-cuadrado para una muestra de cierto número de variables. La ecuación (2) se utiliza para cuantificar la esfericidad.

$$\chi^2 = - \left[n - 1 - \frac{1}{6}(2v + 5) \right] \ln(|R|) \quad (2)$$

Dónde; n , es el tamaño de la muestra, v es el número de variables y, $|R|$ es el determinante de la matriz de correlaciones simple.

Tabla 8-4. Valores para el Test de Esfericidad de Bartlett

| Parámetro | Símbolo | Valor |
|--|----------------------------|----------------|
| Tamaño de la muestra | n | 24 |
| Número de variables | v | 7 |
| Determinante | $ R $ | 0,0147 |
| Test de esfericidad de Bartlett | χ^2 | 83,6654 |

Realizado por: Ayabaca, Roberto. 2022

En la Tabla 8, se observa que la esfericidad es 83.6654 y la prueba de Chi-cuadrado es 46.7963 para el tamaño de la muestra (24 registros) y número de variables (7 variables). En este caso, se rechaza la hipótesis nula, por ende, se concluye que existe correlación entre las variables, lo cual indica que se puede aplicar el mencionado análisis.

4.1.1.2. Índice KMO de Kaiser-Meyer- Olkin

El índice KMO se utiliza para estimar la existencia de coeficientes de correlación parcial diferentes de cero. Mientras más bajo es el coeficiente de correlación parcial las variables comparten factores comunes viabilizando la aplicación del Análisis Factorial. El coeficiente de correlación parcial indica el grado de relación entre dos variables descartando la influencia de las otras variables. Para determinar el índice se aplica la ecuación (3):

$$KMO = \frac{\sum_{i \neq j} \sum r_{ij}^2}{\sum_{i \neq j} \sum r_{ij}^2 + \sum_{i \neq j} \sum a_{ij}^2} \quad (3)$$

Dónde, r_{ij}^2 es la correlación simple entre variables y a_{ij}^2 es la correlación parcial

Para determinar los coeficientes de correlación parciales, existen varias formas. El procedimiento clásico, el mismo que en décadas anteriores se lo realizaba a mano, se obtienen a partir de la correlación de Pearson. La ecuación (4), muestra la expresión de primer orden para calcular el coeficiente de correlación parcial entre dos variables, eliminando el efecto de una tercera.

$$r_{12,3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{1 - r_{13}^2} \sqrt{1 - r_{23}^2}} \quad (4)$$

Dónde, r_{12} es el coeficiente de correlación entre la variable 1 y 2, r_{13} es el coeficiente de correlación entre la variable 1 y 3, r_{23} es el coeficiente de correlación entre la variable 2 y 3.

Cuando se tienen más de tres variables, la expresión se traduce en incorporar términos de orden inferior. Por ejemplo, la expresión matemática de segundo orden se muestra en la ecuación (5):

$$r_{12,34} = \frac{r_{12,3} - r_{14,3}r_{24,3}}{\sqrt{1-r_{14,3}^2}\sqrt{1-r_{24,3}^2}} \quad (5)$$

Dónde, $r_{12,3}$, $r_{14,3}$ Y $r_{24,3}$ son coeficientes parciales de primer orden.

En este caso, como se disponen de siete variables se llegará a obtener los coeficientes de correlación parcial de quinto orden. En la Tabla 9, se exhiben los valores obtenidos.

Tabla 9-4. Matriz de correlaciones parciales

| | Nivelación | Femenino | Mestizo | Indígena | Q1 | Q2 | Q3 |
|------------|------------|----------|---------|----------|--------|--------|--------|
| Nivelación | | 0,275 | -0,428 | -0,130 | 0,339 | -0,142 | 0,294 |
| Femenino | 0,275 | | 0,114 | -0,098 | 0,030 | -0,080 | -0,076 |
| Mestizo | -0,428 | 0,114 | | -0,822 | 0,164 | -0,165 | 0,288 |
| Indígena | -0,130 | -0,098 | -0,822 | | 0,256 | -0,134 | 0,194 |
| Q1 | 0,339 | 0,030 | 0,164 | 0,256 | | -0,441 | -0,729 |
| Q2 | -0,142 | -0,080 | -0,165 | -0,134 | -0,441 | | -0,598 |
| Q3 | 0,294 | -0,076 | 0,288 | 0,194 | -0,729 | -0,598 | |

Realizado por: Ayabaca, Roberto. 2022

Para determinar el índice KMO se debe obtener la sumatoria de los cuadrados de los coeficientes de correlación simple y parcial de la Tabla 7 y Tabla 9 respectivamente. Una vez conseguidos estos valores, se reemplazan en la ecuación (3) como se muestra a continuación.

$$KMO = \frac{6.7584}{6.7584 + 5.1006}$$

$$KMO = 0.570$$

Tabla 10-4. Prueba de esfericidad de Bartlett e índice KMO

| KMO y la prueba de Bartlett | | |
|--|---------------------|--------|
| Kaiser-Meyer-Olkin Medida de adecuación de muestreo. | | 0,570 |
| | Aprox. Chi-Cuadrado | 83,665 |
| La prueba de esfericidad de Bartlett | df | 21 |
| | Sig. | 0,000 |

Fuente: Análisis de Componentes Principales con SPSS

Realizado por: AYABACA. Roberto. 2021

En la Tabla 10, se muestran los resultados obtenidos mediante el software SPSS. Los valores obtenidos mediante la ecuación (2) respecto a la prueba de esfericidad coinciden con los del software utilizado. Asimismo, el índice KMO es el mismo tanto el calculado como el

conseguido por el programa SPSS. Sin embargo, es muy bajo, lo cual indica que no es tan recomendable iniciar un Análisis Factorial. Se debe principalmente a que existen correlaciones parciales muy elevadas que provocan una disminución importante en el índice KMO. No obstante, se prosigue con el análisis a fin de desestimar alguna o algunas variables que todavía manifiestan una moderada correlación parcial entre ellas.

4.1.1.3. Valores Propios

Tabla 11-4. Matriz de valores propios

| | | | | | | |
|----------|----------|----------|----------|----------|----------|----------|
| 1-λ | 0,18359 | -0,55890 | 0,52686 | 0,48974 | -0,32899 | -0,05500 |
| 0,18359 | 1-λ | 0,17279 | -0,17746 | 0,07099 | -0,18174 | 0,05888 |
| -0,55890 | 0,17279 | 1-λ | -0,88791 | -0,52656 | -0,08850 | 0,44499 |
| 0,52686 | -0,17746 | -0,88791 | 1-λ | 0,54060 | -0,06879 | -0,31863 |
| 0,48974 | 0,07099 | -0,52656 | 0,54060 | 1-λ | -0,11169 | -0,60499 |
| -0,32899 | -0,18174 | -0,08850 | -0,06879 | -0,11169 | 1-λ | -0,50304 |
| -0,05500 | 0,05888 | 0,44499 | -0,31863 | -0,60499 | -0,50304 | 1-λ |

Realizado por: Ayabaca, Roberto. 2022

Con los valores de la Tabla 11 se calculan los valores propios, considerando que el determinante de la matriz debe ser cero. Para hallar el determinante se aplica la regla de Laplace (método de cofactores). El resultado es una ecuación polinómica de grado siete (al tener siete variables explicativas), cuya forma viene dada por la ecuación (6)

$$a\lambda^7 + b\lambda^6 + c\lambda^5 + d\lambda^4 + e\lambda^2 + f\lambda + g = 0 \quad (6)$$

Dónde, λ son los valores propios y a, b, \dots, g son los coeficientes de cada término. El resultado obtenido es la ecuación (7)

$$\begin{aligned} &5 \times 10^{62} \lambda^7 - 3,5 \times 10^{63} \lambda^6 + 8,81040988475571 \times 10^{63} \lambda^5 \\ &- 1,02423610259958 \times 10^{64} \lambda^4 + 5,83223644453659 \times 10^{63} \lambda^3 \\ &- 1,59131477293971 \times 10^{63} \lambda^2 + 1,87116392403417 \times 10^{62} \lambda \\ &- 7,36093460438694 \times 10^{60} = 0 \end{aligned} \quad (7)$$

La solución a la ecuación de séptimo grado se ha obtenido mediante el método numérico de Newton-Raphson. Los valores obtenidos son los siguientes:

Tabla 12-4. Raíces de la ecuación de séptimo grado

| λ_1 | λ_2 | λ_3 | λ_4 | λ_5 | λ_6 | λ_7 |
|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| 0,079 | 0,152 | 0,366 | 0,605 | 1,085 | 1,668 | 3,045 |

Realizado por: Ayabaca, Roberto. 2022

De los valores obtenidos y expuestos en la Tabla 12, se observa que λ_7 representa a la componente principal, ya que absorbe el 43.5% ($3.045/7 = 0.435$) de toda la información. Análogamente, si se toma los dos valores más altos, esto es 3.045 y 1.668, se alcanza un 67.33% de explicación de la información. Siguiendo el análisis, para tres componentes con mayor incidencia se esperaría explicar el 82.83% de la información. En consecuencia, a mayor cantidad de componentes más exacta y precisa será la explicación. No obstante, esto dificulta el análisis. En la Tabla 13, se muestran los datos obtenidos para una mejor comprensión.

Tabla 13-4. Cálculo de la varianza total explicada

| Componente | Valor propio | Varianza | |
|------------|--------------|----------|-----------|
| | | Parcial | Acumulada |
| 1 | 3,045 | 43,50% | 43,50% |
| 2 | 1,668 | 23,83% | 67,33% |
| 3 | 1,085 | 15,50% | 82,83% |
| 4 | 0,605 | 8,64% | 91,47% |
| 5 | 0,366 | 5,23% | 96,70% |
| 6 | 0,152 | 2,17% | 98,87% |
| 7 | 0,079 | 1,13% | 100,00% |

Realizado por: Ayabaca, Roberto. 2022

Se consideran las tres primeras componentes debido a los siguientes criterios:

Requiriente

- Para una adecuada explicación, la varianza acumulada de las componentes debe explicar al menos el 70% de la información.
- El valor propio de cada componente debe ser superior a la unidad

Esto es corroborado con los resultados obtenidos mediante el software SPSS, los mismos que se muestran en la Tabla 14.

Tabla 14-4. Cálculo de Varianza total explicada con SPSS

| Componente | Valores propios iniciales | | | Sumas de extracción de cargas cuadradas | | | Sumas de rotación de cargas cuadradas | | |
|------------|---------------------------|---------------|-------------|---|---------------|-------------|---------------------------------------|---------------|-------------|
| | Total | % de varianza | % acumulado | Total | % de varianza | % acumulado | Total | % de varianza | % acumulado |
| 1 | 3,045 | 43,497 | 43,497 | 3,045 | 43,497 | 43,497 | 2,967 | 42,387 | 42,387 |
| 2 | 1,668 | 23,824 | 67,322 | 1,668 | 23,824 | 67,322 | 1,648 | 23,537 | 65,924 |
| 3 | 1,085 | 15,494 | 82,816 | 1,085 | 15,494 | 82,816 | 1,182 | 16,892 | 82,816 |
| 4 | ,605 | 8,649 | 91,464 | | | | | | |
| 5 | ,366 | 5,230 | 96,694 | | | | | | |
| 6 | ,152 | 2,173 | 98,867 | | | | | | |
| 7 | ,079 | 1,133 | 100,000 | | | | | | |

Método de extracción: Análisis de componentes principales.

Fuente: Análisis de Componentes Principales con SPSS

Realizado por: Ayabaca, Roberto. 2022

4.1.1.4. Matriz de componentes

Una vez que se ha determinado los valores propios, se requiere determinar el vector propio o vector característico, bajo la siguiente relación que se expresa en la ecuación (8)

$$(R - \lambda I)u_k = 0 \quad (8)$$

Dónde, R es la matriz de los coeficientes de correlación, λ son los valores propios, I es la matriz identidad y u_k es el vector propio.

En la Tabla 15, se muestra el sistema de ecuaciones en forma matricial para $\lambda_7 = 3.045$.

Tabla 15-4. Sistema de ecuaciones en forma matricial

$$\begin{pmatrix} -2,045 & 0,1836 & -0,5589 & 0,5269 & 0,4897 & -0,3290 & -0,0550 & | & 0 \\ 0,1836 & -2,045 & 0,1728 & -0,1775 & 0,0710 & -0,1817 & 0,0589 & | & 0 \\ -0,5589 & 0,1728 & -2,045 & -0,8879 & -0,5266 & -0,0885 & 0,4450 & | & 0 \\ 0,5269 & -0,1775 & -0,8879 & -2,045 & 0,5406 & -0,0688 & -0,3186 & | & 0 \\ 0,4897 & 0,0710 & -0,5266 & 0,5406 & -2,045 & -0,1117 & -0,6050 & | & 0 \\ -0,3290 & -0,1817 & -0,0885 & -0,0688 & -0,1117 & -2,045 & -0,5030 & | & 0 \\ -0,0550 & 0,0589 & 0,4450 & -0,3186 & -0,6050 & -0,5030 & -2,045 & | & 0 \end{pmatrix}$$

Realizado por: Ayabaca, Roberto. 2022

Aplicando eliminación gaussiana, se tiene el siguiente resultado mostrado en la ecuación (9):

$$u_1 = \begin{pmatrix} u_{11} \\ u_{12} \\ u_{13} \\ u_{14} \\ u_{15} \\ u_{16} \\ u_{17} \end{pmatrix} = \begin{pmatrix} -1.132u_{17} \\ 0.140u_{17} \\ 1.528u_{17} \\ -1.478u_{17} \\ -1.345u_{17} \\ -0.019u_{17} \\ u_{17} \end{pmatrix} \quad (9)$$

El vector obtenido debe satisfacer la condición enunciada en la ecuación (10). Por ello se debe hallar un valor para u_7 que satisfaga tal requerimiento:

$$u_{11}^2 + u_{12}^2 + u_{13}^2 + u_{14}^2 + u_{15}^2 + u_{16}^2 + u_{17}^2 = \lambda \quad (10)$$

El valor de u_7 que se aproxima, para satisfacer la ecuación (10) para $\lambda_I = 3.045$ es $0,5940136$. Por lo tanto, el vector propio (redondeado a cuatro decimales) para la componente uno es:

$$u_1 = \begin{pmatrix} u_{11} \\ u_{12} \\ u_{13} \\ u_{14} \\ u_{15} \\ u_{16} \\ u_{17} \end{pmatrix} = \begin{pmatrix} -0.6724 \\ 0.0832 \\ 0.9076 \\ 0.8779 \\ 0.7989 \\ -0.0113 \\ 0.5940 \end{pmatrix}$$

De igual manera, se obtienen los vectores propios (redondeado a cuatro decimales) para la componente dos y tres, cuyos valores propios son 1.668 y 1.085 respectivamente.

$$u_2 = \begin{pmatrix} u_{22} \\ u_{22} \\ u_{23} \\ u_{24} \\ u_{25} \\ u_{26} \\ u_{27} \end{pmatrix} = \begin{pmatrix} 0.5473 \\ 0.4456 \\ 0.0570 \\ 0.0700 \\ 0.0465 \\ -0.8807 \\ 0.6198 \end{pmatrix} \quad u_3 = \begin{pmatrix} u_{31} \\ u_{32} \\ u_{33} \\ u_{34} \\ u_{35} \\ u_{36} \\ u_{37} \end{pmatrix} = \begin{pmatrix} -0.0319 \\ -0.8082 \\ -0.2193 \\ 0.2965 \\ -0.3263 \\ -0.1495 \\ 0.4084 \end{pmatrix}$$

En la Tabla 16, se muestran las tres componentes analizadas. Estos resultados son idénticos a los calculados por el software SPSS, como se evidencia en la Tabla 17.

Tabla 16-4. Matriz de componentes extraídas

| Variable | Componente | | |
|------------|------------|---------|---------|
| | 1 | 2 | 3 |
| Nivelación | -0,6724 | 0,5473 | -0,0319 |
| Femenino | 0,0832 | 0,4456 | -0,8082 |
| Mestizo | 0,9077 | 0,0570 | -0,2193 |
| Indígena | 0,8780 | 0,0700 | 0,2965 |
| N1 | 0,7989 | 0,0465 | -0,3263 |
| N2 | -0,0113 | -0,8807 | -0,1495 |
| N3 | 0,5940 | 0,6198 | 0,4084 |

Realizado por: Ayabaca, Roberto. 2022

Tabla 17-4. Matriz de componentes hallada con software SPSS

Matriz de componentes^a

| | Componente | | |
|------------|------------|-------|-------|
| | 1 | 2 | 3 |
| Nivelación | ,672 | ,547 | ,032 |
| Femenino | -,083 | ,446 | ,808 |
| Mestizo | -,908 | ,057 | ,219 |
| Indígena | ,878 | ,070 | -,296 |
| N1 | ,799 | ,047 | ,326 |
| N2 | ,011 | -,880 | ,149 |
| N3 | -,594 | ,620 | -,408 |

Método de extracción: Análisis de componentes principales.

a. 3 componentes extraídos.

Fuente: Análisis de Componentes Principales con SPSS.

Realizado por: Ayabaca, Roberto. 2022

4.1.1.5. La Comunalidad

La comunalidad representa la varianza compartida de una variable con otras a través de los factores comunes. Al iniciar el estudio la comunalidad de cada variable es la unidad, pues comparte el 100% de la varianza. No obstante, cuando se han establecido las componentes la comunalidad de cada parámetro es inferior a la unidad. Para calcular la comunalidad de una variable se utiliza la ecuación (11):

$$h_j^2 = \sum_{k=1}^n a_{jk}^2 \quad (11)$$

Dónde, n es la cantidad de componentes obtenidas, a_{jk} es el factor de la variable en cada componente.

Los valores altos son relevantes ya que identifican la importancia de la variable en la creación de una componente, mientras que si el valor es bajo (cerca a cero) no contribuyen para la descripción de la componente, ya que comparten muy poca de la varianza común.

Los valores de la comunalidad se obtienen luego de que se haya hallado la matriz de componentes o de saturación.

Utilizando la ecuación (11) para los valores de los componentes de las variables mostrados en la Tabla 16, se deriva las siguientes comunalidades mostradas en la Tabla 18:

Tabla 18-4. Tabla de Comunalidades

| Variable | Comunalidad | |
|------------|-------------|----------|
| | Inicial | Obtenida |
| Nivelación | 1 | 0,7527 |
| Mujer | 1 | 0,8587 |
| Mestizo | 1 | 0,8752 |
| Indígena | 1 | 0,8636 |
| N1 | 1 | 0,7470 |
| N2 | 1 | 0,7982 |
| N3 | 1 | 0,9038 |

Realizado por: Ayabaca, Roberto. 2022

Los valores de la comunalidad mostrados en la Tabla 18, determinan que las variables “nivelación”, “N1” y “N2” tienen valores de comunalidad inferiores respecto de las otras, sin embargo, las diferencias no son decisivas, por lo tanto, todas ellas contribuyen fuertemente al establecimiento de las componentes mencionadas.

4.1.1.6. Rotación de los ejes

Los factores obtenidos y mostrados en la Tabla 16, indican el peso o factor que cada variable tiene dentro de cada componente. El análisis se puede continuar desde este punto, sin embargo, no siempre se puede establecer de manera clara la relación entre los factores y las variables.

La rotación de los ejes consiste en la transformación de la matriz de componentes en otra de mayor simplicidad que adecúe los valores de las variables en un análisis de la manera más aproximada posible con los ejes, facilitando de esta forma la interpretación de la estructura de los datos sin alterar la solución factorial, comunalidades y porcentaje de explicación de la varianza entre los parámetros en estudio.

En este caso de estudio, a criterio del investigador la diferencia entre los factores el peso de cada una de las variables dentro de una misma componente es definitiva. Por ende, no se considera una rotación de variables.

4.1.1.7. Identidad de ejes

La identidad del eje consiste en otorgar de un nombre al conjunto de variables que tienen elevados factores de peso dentro de una componente. Por ende, partimos analizando la Tabla 16, en la cual se observa que:

- a. En la componente 1, las variables que tienen mayor peso son: mestizo, indígena y N1,
- b. En la componente 2, la variable cuyo peso sobresale del resto es “N2”,
- c. En la componente 3, la variable que tiene el mayor peso es “mujer”.

Diagrama de dispersión de las componentes 1 y 2

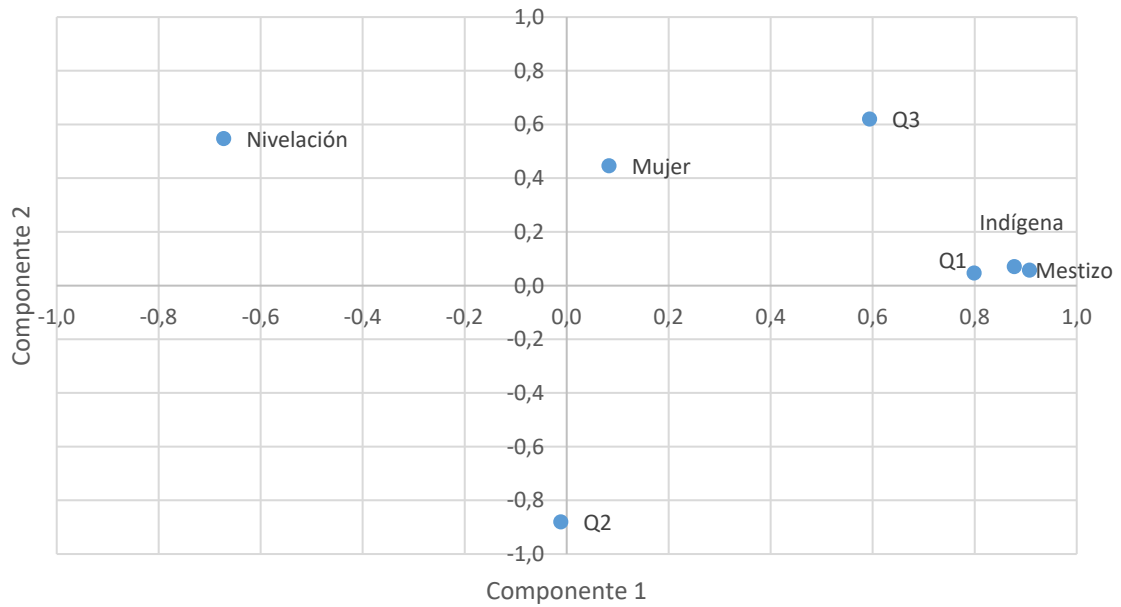


Gráfico 10-4. Representación gráfica de las componentes 1 y 2

Realizado por: Ayabaca, Roberto. 2022

En el Gráfico 10, se muestra que en el eje horizontal correspondiente a la componente 1 las variables cuyo factor de peso se acerca a la unidad son “mestizo” e “indígena”. A esta componente se le podría denominar “Autoidentificación étnica”. De manera similar, en la componente 2 se evidencia que la variable “N2” tiene el factor más cercano a la unidad, pudiendo denominar a esta componente como “nivel económico”. Mientras que el resto de variables quedan excluidas de las componentes analizadas.

De igual manera, en el Gráfico 11 se representa gráficamente las componentes 2 y 3. Se nota que en la componente 3, la variable que sobresale es “mujer”. Por ende, podría denominarse como “identidad de sexo” o “identidad de género”.

Diagrama de dispersión de las componentes 2 y 3

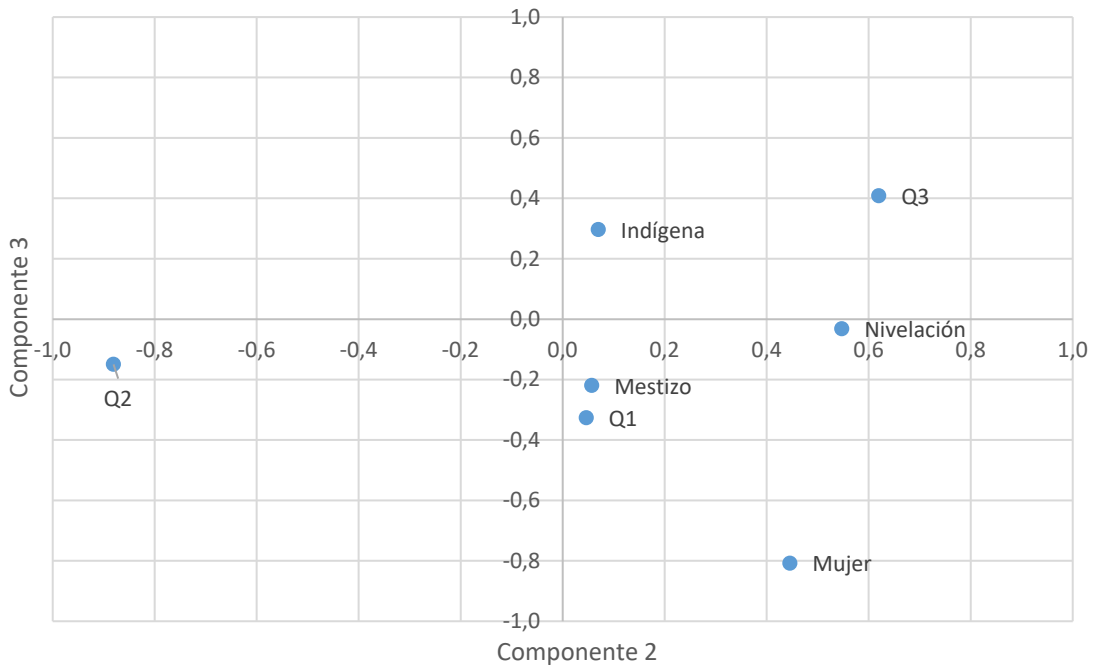


Gráfico 11-4. Representación gráfica de las componentes 2 y 3

Realizado por: Ayabaca, Roberto. 2022

4.1.2. *Análisis de Regresión Lineal Múltiple*

En la sección anterior se determinaron tres componentes que acumulan la mayor parte de la varianza, siendo la componente uno la que más varianza abarca. Las variables que más destacaron son “mestizo” e “indígena” perteneciendo a la clasificación denominada “Autoidentificación Étnica”. En este sentido, son dos las variables que explican el índice de deserción estudiantil, según los datos analizados. Para determinar una ecuación que relacione los factores exógenos con la variable endógena se aplica el Análisis de Regresión Múltiple. La forma general viene dada por la ecuación (12).

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n + \varepsilon \tag{12}$$

Dónde, $b_0 + b_1 + b_2 + \dots + b_n$ son los coeficientes parciales de regresión, $x_1 + x_2 + \dots + x_n$ son las variables predictoras que se miden sin error y ε es el error de cada variable explicativa.

El modelo de regresión múltiple se determina a partir de la ecuación (13):

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n \tag{13}$$

Dónde, \hat{y} es el valor promedio pronosticado

4.1.2.1. Obtención de los coeficientes parciales de regresión

Los coeficientes de regresión se obtienen aplicando el método de mínimos cuadrados que permite determinar el vector de coeficientes mediante un estimador, como se denota en la ecuación (16).

$$\hat{b} = (X^T X)^{-1} X^T Y \quad (14)$$

Dónde, \hat{b} es el vector de coeficientes de regresión, X^T es la matriz transpuesta de la matriz X , X es la matriz de las variables explicativas e Y es el vector de la variable explicada.

Tabla 19-4. Matriz de datos para Regresión

| Provincia | Abandona | Mestizo | Indígena |
|--------------------------------|----------|---------|----------|
| | y_1 | x_1 | x_2 |
| Azuay | 3,45 | 97,41 | 1,72 |
| Bolívar | - | 100,00 | - |
| Cañar | 2,86 | 94,29 | 5,71 |
| Carchi | - | 100,00 | - |
| Cotopaxi | 8,51 | 97,87 | 2,13 |
| Chimborazo | - | 94,59 | 5,41 |
| El Oro | 2,63 | 98,68 | 1,32 |
| Esmeraldas | 8,11 | 83,78 | - |
| Guayas | 5,75 | 98,85 | - |
| Imbabura | - | 95,12 | 2,44 |
| Loja | 4,81 | 96,15 | 1,92 |
| Los Ríos | 8,33 | 100,00 | - |
| Manabí | 2,22 | 100,00 | - |
| Morona Santiago | 2,50 | 90,00 | 10,00 |
| Napo | 9,75 | 48,44 | 49,35 |
| Pastaza | 8,40 | 81,68 | 16,79 |
| Pichincha | 5,24 | 93,85 | 3,98 |
| Tungurahua | 7,27 | 95,45 | 1,82 |
| Zamora Chinchipe | 13,04 | 78,26 | 21,74 |
| Galápagos | - | 80,00 | - |
| Sucumbíos | 6,06 | 87,88 | 7,58 |
| Orellana | 11,01 | 80,73 | 15,60 |
| Santo Domingo de los Tsáchilas | 5,26 | 92,11 | - |
| Santa Elena | 10,53 | 89,47 | - |

Fuente: Universidad Regional Amazónica Ikiam

Realizado por: Ayabaca, Roberto. 2022.

En la Tabla 19, se muestran las variables que intervienen en la explicación de la variable dependiente. Los valores expuestos se representan en tanto por ciento. La matriz de variables explicativas X está conformada de dos variables (mestizo e indígena) con veinticuatro registros. De igual manera, el vector de la variable explicada corresponde al abandono escolar con la misma cantidad de mediciones.

Se aplica la ecuación (16), para obtener los coeficientes de regresión. Se considera el valor de cero a la constante del modelo dado que la tasa de deserción es cero si no hay estudiantes matriculados. En consecuencia, la media aritmética de las variables independientes (\bar{x}_1, \bar{x}_2) toma el valor de cero. Las cifras de los coeficientes se muestran en la Tabla 20.

Tabla 20-4. Coeficientes de regresión

| Coeficiente | Valor |
|-------------|--------|
| B1 | 0,0433 |
| B2 | 0,2081 |

Realizado por: Ayabaca, Roberto. 2022

Reemplazando los valores en la ecuación (13), se tiene la ecuación (17) que modela el fenómeno de deserción estudiantil.

$$\hat{y} = 0.0433x_1 + 0.2081x_2 \quad (15)$$

Para afirmar que la ecuación (17), predice de forma adecuada la deserción estudiantil es necesario analizar los siguientes aspectos:

- El coeficiente de determinación (R^2),
- El coeficiente de determinación ajustado,
- Significancia general de la ecuación de regresión múltiple y,
- Significancia de los coeficientes de regresión parcial.

4.1.2.2. Coeficiente de Determinación R^2

La ecuación (16), permite analizar la variación total en la variable de salida y

$$V_{total} = V_{regresión} + V_{error} \quad (16)$$

Dónde, $V_{regresión}$ es la variación explicada usando la ecuación de regresión y V_{error} es la variación residual en los datos no explicados por las variables independientes.

La ecuación (17) y (18) cuantifican la variación explicada por la regresión y la variación total de la variable dependiente en análisis respectivamente.

$$V_{regresión} = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = \sum_{i=1}^n \hat{y}_i^2 \quad (17)$$

$$V_{Total} = \sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n y_i^2 \quad (18)$$

Dónde, $\sum_{i=1}^n \hat{y}_i^2$ es la suma de los cuadrados para la regresión y $\sum_{i=1}^n y_i^2$ es la suma de los cuadrados para el error.

Los valores obtenidos aplicando las ecuaciones (17) y (18) para $V_{regresión}$ y V_{Total} son 728.647 y 1014.437 respectivamente. La ecuación (16) permite el cálculo de V_{error} cuyo resultado es 285.790

La ecuación (19) permite cuantificar el coeficiente de determinación (R^2).

$$R^2 = \frac{V_{regresión}}{V_{Total}} \quad (19)$$

El valor conseguido es 0.718, es decir, que el modelo se ajusta de manera aceptable.

4.1.2.3. Coeficiente de Determinación ajustado

El coeficiente de determinación a medida que se incluyen variables predictoras al modelo tiende a incrementarse dando la idea de que a mayor cantidad de factores explicativos mejora el ajuste. No obstante, es no es cierto del todo. Para restringir esta apreciación errónea se ha establecido como medida alternativa el uso de una expresión con base en los grados de libertad. La ecuación (20) posibilita el cálculo.

$$R^2_{ajustado} = 1 - \frac{MCE}{V_{total}}(n-1) \quad (20)$$

Dónde, MCE es la varianza no explicada por las variables predictoras, n es la cantidad de registros disponibles.

La varianza no explicada por las variables predictoras se halla mediante la ecuación (21)

$$MCE = \frac{V_{error}}{n-k-1} \quad (21)$$

Dónde, k representa los grados de libertad de la regresión y $n-k-1$ representa los grados de libertad total.

En este caso se tienen veinticuatro registros y dos grados de libertad. Dentro de la ecuación (21) no se considera el término “-1”, ya que este número estima el uso de una constante, que en el presente estudio no se contempla. Reemplazando los valores, MCE adopta la cantidad de 12.9905 (redondeado a cuatro decimales)

Por lo tanto, reemplazando los valores en la ecuación (20) se tiene que $R^2_{ajustado}$ es 0.6927 (aproximado a cuatro decimales).

El valor obtenido indica que los dos parámetros predictores contemplan el 69.27% de la varianza explicada.

4.1.2.4. Prueba de significancia general de la ecuación de Regresión Múltiple

Para determinar si el modelo es significativo en la explicación de la variable dependiente se establece la prueba F bajo las siguientes consideraciones:

- a. Hipótesis nula: $\beta_i = 0$, se rechaza cuando $F > F_{\alpha}(k, n-k-1)$, por ende, se considera que el modelo es significativo.
- b. Hipótesis alternativa: $\beta_i \neq 0$, se acepta cuando $F < F_{\alpha}(k, n-k-1)$, es decir, el modelo no es significativo.

La ecuación (22) que permite cuantificar el estadístico F.

$$F = \frac{R^2}{1-R^2} \cdot \frac{n-k-1}{k} \quad (22)$$

Se tienen dos grados de libertad de la regresión y veintidós grados de libertad totales. Reemplazando los valores en la ecuación (22) se tiene que F equivale a 28.0454. De tablas sobre la distribución F de Fisher se tiene un valor de 19.450. Este valor rechaza la hipótesis nula. Por lo tanto, se afirma que la regresión es significativa. En la Tabla 21 se muestra lo obtenido con el programa SPSS verificando lo calculado.

Tabla 21-4. Prueba de significancia de la ecuación

| | | ANOVA ^{a,b} | | | | |
|---|-----------|-----------------------|----|--------------------|--------|-------------------|
| | Modelo | Suma de cuadrados | df | Media de cuadrados | F | Sig. |
| 1 | Regresión | 728,647 | 2 | 364,323 | 28,045 | ,000 ^c |
| | Residuos | 285,790 | 22 | 12,990 | | |
| | Total | 1014,437 ^d | 24 | | | |

a. Variable dependiente: Abandona

b. Regresión lineal a través del origen

c. Predictoras: Indígena, Mestizo

d. La suma de los cuadrados no es correcta por la constante porque la constante es cero para la regresión a través del origen.

Fuente: Análisis de regresión con SPSS

Realizado por: AYABACA, Roberto. 2021

4.1.2.5. Prueba de significancia de los coeficientes de regresión parcial

Una vez que se ha establecido que el modelo es útil, se procede a determinar si las variables son significativas dentro de la ecuación, es decir, si aportan información relevante para la explicación. Para el estudio de este aspecto se procede a realizar pruebas a cada una de las variables intervinientes bajo las siguientes hipótesis:

- Hipótesis nula: $\beta_i = 0$, $t > t_\alpha (n-k-1)$ se rechaza la hipótesis y se establece que la variable es significativa dentro del modelo.
- Hipótesis alternativa: $\beta_i \neq 0$, $t < t_\alpha (n-k-1)$ se acepta la hipótesis y se determina que la variable no es significativa dentro del modelo.

Esta prueba se basa en el estadístico t de Student. Para calcular se emplea la ecuación (23)

$$t = \frac{b_i - \beta_i}{\sqrt{\frac{MCE^*}{S_{xx}}}} \quad (23)$$

Dónde, S_{xx} es la varianza de la variable independiente, MCE^* es la media cuadrática de la varianza de los errores, b_i es el coeficiente de regresión.

Para el coeficiente de regresión de la variable “mestizo”, el valor calculado de t es 4.925. La cantidad obtenida de tablas de distribución t de Student para una confianza del 95% y 23 grados de libertad es de 1.7139. Dado que el valor hallado es mayor que el tabulado, se rechaza la hipótesis nula, considerándose el factor analizado como significativo.

El valor de t para el coeficiente de regresión de la variable “indígena” calculado es de 3.218. El valor tabulado para una distribución t de Student con una cola para una confianza del 95% y 23 grados de libertad es de 1.7139. Por lo tanto, el valor conseguido es superior al tabulado, por ende, el parámetro “indígena” es significativo dentro del modelo.

En la Tabla 22 se muestran los valores obtenidos por el programa SPSS, que contrasta los valores calculados a partir de la ecuación (23) para cada factor interviniente de la ecuación (15).

Tabla 22-4. Prueba de significancia de los coeficientes de regresión parcial

| | | Coeficientes ^{a,b} | | | | | Estadísticas de colinealidad | |
|--------|----------|--------------------------------|------|-----------------------------|-------|------|------------------------------|-------|
| Modelo | | Coeficientes no estandarizados | | Coeficientes estandarizados | t | Se. | Tolerancia | VIF |
| | B | Std. Error | Beta | | | | | |
| 1 | Mestizo | ,043 | ,009 | ,607 | 4,925 | ,000 | ,842 | 1,188 |
| | Indígena | ,208 | ,065 | ,397 | 3,218 | ,004 | ,842 | 1,188 |

a. Dependiente Variable: Abandona

b. Regresión lineal a través del origen

Fuente: Análisis de regresión lineal con SPSS

Realizado por: AYABACA, Roberto. 2022

4.1.2.6. Prueba de colinealidad

El Factor de Inflación de la Varianza (VIF) mide la colinealidad de una variable con respecto de las otras. Cuando el valor es superior a 10 se concluye que existe colinealidad fuerte. Si este supera el valor de 30 la colinealidad es severa. Caso contrario, si el valor es inferior a 10 se acepta. En la Tabla 22 se observa que los valores obtenidos son inferiores a diez, por lo cual, la colinealidad entre las variables intervinientes es bajo.

4.1.3. Validación del Modelo Matemático

En esta sección se analiza la precisión en la predicción de la Tasa de Deserción Estudiantil (TDE) a través de la ecuación producto del análisis de regresión lineal. La expresión obtenida con base en la información registrada por la Universidad está en función de dos variables. En la ecuación (24) se reescribe a partir de la ecuación (15), en la cual X_1 el parámetro “mestizo” y X_2 el factor “indígena”. Por consiguiente:

$$TDE = 0.0433 \cdot \%mestizo + 0.2081 \cdot \%indígena \quad (24)$$

Para verificar la ecuación encontrada, se compara los porcentajes de deserción registrada con los valores calculados para los años en estudio. Esto indicará la precisión del modelo. En la Tabla 23 se evidencia que en los años 2018 y 2019 el abandono calculado, así como el registrado son aproximados. En el año 2017 los valores del abandono difieren sustancialmente entre el calculado y el registrado, esto debido muy probablemente a que en el citado año solo se clasificaron a los estudiantes en dos categorías “mestizo” e “indígena” de acuerdo a su autoidentificación étnica, mientras que en los años posteriores las categorías del parámetro “etnia” se diversificaron.

Tabla 23-4. Comparación de tasas de deserción

| Año | Registros | | | Abandono Calculado | Error |
|------|-----------|---------|----------|--------------------|---------|
| | Abandono | Mestizo | Indígena | | |
| 2017 | 5,57% | 79,26% | 20,74% | 7,75% | -28,08% |
| 2018 | 8,72% | 74,60% | 23,04% | 8,02% | 8,67% |
| 2019 | 6,63% | 77,74% | 18,90% | 7,30% | -9,20% |

Realizado por: AYABACA, Roberto. 2022

La estimación desarrollada toma en consideración la procedencia étnica. Dentro de ella, son relevantes la categoría “mestizo” e “indígena”, siendo esta última la que mayor impacto tiene dentro de la ecuación (en una relación de cinco a uno con respecto de la primera). Puede deberse

a factores implícitos, tales como: nivel socioeconómico y el nivel de educación recibida en el nivel predecesor.

CAPITULO V

5. PROPUESTA

El presente estudio tiene como finalidad calcular la tasa de deserción estudiantil de la Universidad Regional Amazónica Ikiam con base en la información facilitada por la institución de los años 2017, 2018 y 2019. Las variables analizadas son: nivel, provincia de procedencia, etnia y quintil socioeconómico. Las variables son de naturaleza no métrica. La información se adecúa según la provincia de procedencia ya que se concluyó que la distancia es un factor irrelevante dentro del análisis. Las variables al ser cualitativas involucran categorías, y estas categorías se transforman en variables cuantitativas aplicando el concepto de *variable ficticia*. Se obtuvieron dieciséis factores.

Un modelo matemático debe contener la menor cantidad de variables que expliquen la mayor parte de la varianza. Bajo esta premisa, es necesario reducir las variables. Dado que el determinante de la matriz de datos es cero, se deduce que existen factores que presentan una elevada correlación entre sí. Por ello, mediante un análisis de la información, conforme el criterio del investigador, se descartan nueve variables, quedando siete variables (nivelación, femenino, indígena, blanco, N1, N2 y N3) para un análisis más profundo.

El Análisis de Componentes Principales es una técnica que permite obtener los factores más significativos que expliquen el fenómeno. Se aplica la técnica a las siete variables seleccionadas. Los factores se agrupan en tres componentes, explicando el 82.83% de la varianza. En la componente uno, los factores más relevantes e icorrelados son: “mestizo” e “indígena”.

Una vez que se han establecido las variables más representativas que aglutinan la parte mayoritaria de la explicación de la varianza se aplica un Análisis de Regresión Lineal Múltiple, con la finalidad de conseguir una expresión matemática que permita el cálculo de la tasa de deserción. El resultado es la siguiente expresión:

$$TDE = 0.0433 \cdot \%mestizo + 0.2081 \cdot \%indigena$$

Para calcular la tasa de deserción, es necesario conocer el porcentaje de estudiantes autoidentificados como mestizo e indígena. La ecuación alcanza el 69.27% de la explicación de la deserción.

Conocida la expresión que predice la deserción se verifica la precisión de cálculo. Para ello, se compara los valores registrados por la institución con las cantidades calculadas para los años de estudio 2017, 2018 y 2019. En la Tabla 24, se muestra las cifras y el error de cálculo presente.

Tabla 24-5. Error entre las tasas de deserción

| Año | Abandono | | Error |
|------|------------|-----------|---------|
| | Registrado | Calculado | |
| 2017 | 5,57% | 7,75% | -28,08% |
| 2018 | 8,72% | 8,02% | 8,67% |
| 2019 | 6,63% | 7,30% | -9,20% |

Realizado por: AYABACA, Roberto. 2022

La ecuación obtenida predice de manera aproximada la tasa de deserción del establecimiento educativo, constituyéndose en una herramienta útil y práctica. Además, estima que los estudiantes indígenas tienen una mayor probabilidad de desertar en relación a aquellos que se autoidentifican como mestizo.

CONCLUSIONES

En esta investigación se analizó la información registrada por la Universidad en los años 2017, 2018 y 2019 con la finalidad de determinar la tasa de deserción estudiantil. Las variables analizadas fueron cualitativas y se aplicó el concepto de variable ficticia para transformar las categorías en nuevas variables de naturaleza métrica. Posterior a ello, se clasificó la información según la provincia de procedencia en dieciséis variables. Algunos factores fueron descartados por provenir de variables dicotómicas (género, nivel de estudio) y variables irrelevantes como: Q4, Q5, Afroecuatoriano, Blanco, Montubio, Mulato y Otro. Mediante un ACP se definen las variables mestizo e indígena como predictoras del modelo.

A priori, se tenía la premisa que: “*A mayor distancia, mayor dificultad para el estudio, por ende, mayor deserción*”. Sin embargo, al analizar la información se determina que la distancia no es un factor significativo dentro del estudio de la deserción estudiantil.

Se aplicó un Análisis de Regresión Lineal Múltiple con las variables explicativas mestizo e indígena y abandono como variable explicada. La ecuación es la siguiente:

$$TDE = 0.0433 \cdot \%mestizo + 0.2081 \cdot \%indígena$$

La razón entre los coeficientes de regresión de la variable indígena y mestizo es de 4.8, por ende, un estudiante autoidentificado como indígena es más propenso de desertar que un alumno mestizo.

La ecuación obtenida explica el 69.27% de la varianza de la deserción. Por ende, el 31.73% queda sin explicación, lo cual, incrementa el error del cálculo. Para el año 2017, el error calculado respecto del registrado es del 28.08%. Para los años posteriores 2018 y 2019 el error es de 8.67% y 9.2% respectivamente. La diferencia significativa del error en el 2017 se produce, debido a que, los estudiantes estaban identificados como mestizos o indígenas. Para los años posteriores se amplía las categorías de autoidentificación étnica.

La tasa de deserción calculada es de 7.75%, 8.02% y 7.30% para los años 2017, 2018, y 2019 respectivamente. No se revela un incremento significativo en el porcentaje de deserción. No obstante, se matricularon 646 en 2017, 929 en 2018 y 1222 en 2019. Esto es un indicativo que cada año la cantidad de estudiantes que abandonan sus estudios se incrementa.

RECOMENDACIONES

Se recomienda a la Universidad que en sus registros de la información de los estudiantes conste a más de los campos vigentes los siguientes: Calificación de la prueba de ingreso al Sistema de Educación Superior, estado civil, cantidad de hijos, edad, fuente de ingresos económicos y tipo de sostenimiento de la institución predecesora (unidad educativa).

Realizar la aplicación de la encuesta y/o ficha socioeconómica de los estudiantes previo o en el momento mismo de la matrícula, pues en los registros existentes, un porcentaje de estudiantes no disponían de información necesaria para ser considerados en el actual estudio.

Cabe indicar que los porcentajes de deserción para los años 2017, 2018 y 2019 son 5.57%, 8.72% y 6.63% respectivamente. Se recomienda a la Universidad que opte por estudiar la deserción de forma más profunda y detallada a fin de obtener herramientas que permitan reducirla.

Sintetizar la base de datos existente de estudiantes que abandonan el establecimiento, incorporando campos que faciliten el análisis de la información.

La tasa de deserción estudiantil, como se explica, debería ser tenido en cuenta en el sistema de Administración de la Universidad para la detección temprana de los posibles estudiantes que pertenecen al grupo propenso a deserción y darle la orientación y seguimiento respectivo a fin de minimizar el abandono.

GLOSARIO

A

Abandono

Acción y efecto de abandonar o abandonarse..... 1

Académico

Referido al nivel de instrucción alcanzado por un individuo. 1

Análisis de Componentes Principales

Transforma un conjunto de variables correlacionadas en un nuevo conjunto de variables no correlacionadas. 21

C

Coefficiente de correlación lineal

Relación lineal existente entre dos variables. 21

D

Deserción

Acción de desertar. Separarse o abandonar los estudios. 1

Determinante

Valor numérico de una matriz cuadrada. 44

E

Económico

Gasto monetario vinculado a la obtención de un determinado fin 1

Ecuación

Igualdad que contiene una o más incógnitas. 23

Educación Superior

Tercer nivel de instrucción de educación en el cual una persona al finalizarla obtiene un título profesional en

diferentes ámbitos en Universidades e Institutos Tecnológicos..... 1

Estrategias

Arte, traza para dirigir un asunto 1

Etnia

Agrupación natural de individuos de igual cultura. 43

F

Femenino

Propio de las mujeres. Dotado de órganos para ser fecundado. 42

G

Género

Modo o manera. El de los nombres de personas de una sola terminación para el masculino y el femenino. 42

I

Indígena

Originario del país de que se trata, autóctono. En América latina, indio o mestizo no asimilado. 43

M

Masculino

Propio del varón. Que está dotado de órganos para fecundar..... 42

Mestizo

Aplíquese a la persona nacida de padre y madre de diferente raza..... 43

Modelo Matemático

Es un objeto que relaciona una parte del universo a través de una expresión matemática que defina alguna característica específica estudiada. ... 23

| | | |
|--|--|----|
| Multicolinealidad | Disminuir, estrechar o ceñir..... | 1 |
| La multicolinealidad es la relación de dependencia lineal fuerte entre más de dos variables explicativas en una regresión múltiple que incumple el supuesto de Gauss-Markov cuando es exacta. | | 21 |
| N | | |
| Nivel socioeconómico | | |
| Representa la posición social específica de un estudiante respecto al aspecto social y económico..... | | 31 |
| P | | |
| Políticas | | |
| Arte o traza con que se conduce un asunto o se emplean los medio para alcanzar un fin determinado. | | 1 |
| Procedencia | | |
| Referido a la provincia ecuatoriana de procedencia de la cual proviene el estudiante | | 31 |
| R | | |
| Recolección de datos | | |
| La recolección de datos se refiere al enfoque sistemático de reunir y medir información de diversas fuentes a fin de obtener un panorama completo y preciso de una zona de interés..... | | 22 |
| Reducir | | |
| | Regresión múltiple | |
| | Es una técnica estadística, en la cual, una variable es explicada a través de uno o varios factores..... | 20 |
| | Retención | |
| | Proceso de aprendizaje que da lugar a la consolidación de hábitos adquiridos.. | 1 |
| | S | |
| | SENESCYT | |
| | Secretaría de Educación Superior, Ciencia, Tecnología e Innovación. Organismo gubernamental encargado en la dirección de la Educación Superior | 4 |
| | T | |
| | Tasa de deserción estudiantil | |
| | Es la razón porcentual entre los estudiantes que han abandonado sus estudios respecto de los estudiantes matriculados para un periodo académico. | 31 |
| | V | |
| | Variable endógena | |
| | Es un factor que es explicado a partir de uno o varios factores, es decir, las causas que originan el fenómeno. Que se origina por causas internas | 20 |

BIBLIOGRAFÍA

- Agualongo, D., & Garcés, A. (2020). El nivel socioeconómico como factor de influencia en temas de salud y educación. *Revista Vínculos*, 5(2), 19–27. Recuperado el 12 de diciembre de 2021 de <https://doi.org/10.24133/vinculospe.v5i2.1639>
- Albarrán Peña, J. (2019). La deserción estudiantil en la Universidad de Los Andes (Venezuela). *Educación y Humanismo*, 21(36), 60–92. Recuperado el 3 de marzo de 2021 de <https://doi.org/10.17081/eduhum.21.36.2806>
- Alcocer, C. (2020). Adaptación de un modelo matemático para la predicción del abandono estudiantil en la educación superior. Versión preliminar para la Universidad Internacional de La Rioja. Manuscrito no publicado.
- Aparicio, M. T., & Seminara, M. P. (2018). La deserción universitaria ¿Un concepto equívoco? Revisión de estudios latinoamericanos sobre conceptos alternativos. *Universidad Nacional de San Juan*, 44–72
- Argote, I., Jimenez, R. y Gómez, J. (2014). Detección de patrones de deserción en los programas de pregrado de la Universidad Mariana de San Juan de Pasto, aplicando el proceso de descubrimiento de conocimiento sobre base de datos (KDD) y su implementación en modelos matemáticos de predicción (pp. 1–7). Colombia: Universidad Mariana. *Cuarta Conferencia Latinoamericana Sobre El Abandono En La Educación Superior*.
- Arias, F. (2012). *El Proyecto de Investigación: Introducción a la Metodología Científica* (6ª ed.). Caracas. Episteme.
- Arias, J., & Covinos, M. (2021). *Diseño y metodología de la investigación*. Recuperado el 12 de diciembre de 2021, de <http://hdl.handle.net/20.500.12390/2260>
- Bautista, E. (2019). El comportamiento organizacional y la deserción estudiantil en la institución educativa Micaela Bastidas de San Juan de Lurigancho – 2018. Lima: Universidad César Vallejo. Tesis de posgrado para optar al título de Maestro en Administración de la Educación, Escuela de Posgrado. Obtenido de [http://repositorio.uncp.edu.pe/bitstream/handle/UNCP/3000/SilvaAcosta.pdf?sequence=1 &isAllowed=y%0Ahttps://repositorio.comillas.edu/xmlui/handle/11531/1046](http://repositorio.uncp.edu.pe/bitstream/handle/UNCP/3000/SilvaAcosta.pdf?sequence=1&isAllowed=y%0Ahttps://repositorio.comillas.edu/xmlui/handle/11531/1046)
- Benalcázar, M. (2017). Educación privada versus educación pública en el Ecuador. *Revista Publicando*, 4(11), 484–498. Obtenido de <https://dialnet.unirioja.es/servlet/articulo?codigo=6115199>
- Bernal-Cifuentes, O. O. (2018). Evaluación financiera ocasionada por la deserción estudiantil en la Universidad Libre Seccional Cúcuta. *Interfaces*, 1(1), 67–81.
- Bravo, F., Illescas, L., Larriva, S., & Peña, M. (2017). Causas de Deserción en el Ingreso a la Universidad. *Revista de La Facultad de Ciencias Químicas Universidad de Cuenca*, 18, 48–59

- Cardona, J. P., Leal, J. J., & Ustariz, J. E. (2020). Modelado matemático de caja blanca y negra en educación en ingeniería. *Formación Universitaria*, 13(6), 105–118. Obtenido de <https://doi.org/10.4067/S0718-50062020000600105>
- Carhuancho, I., Nolazco, F., Sicheri, L., Guerrero, M., & Casana, K. (2019). *Metodología para la investigación holística* (1ª ed.). Guayaquil. Universidad Internacional del Ecuador.
- Chávez, G. (2018). Estructura de las comunidades de Lepidopteras en función de variables climáticas y hábitat natural y antrópico mediante análisis de interdependencia, altos mirandinos Venezuela, 2013. Riobamba: Escuela Superior Politécnica de Chimborazo. Escuela de Física y Matemática.
- Chimba, E. (2020). Modelo para la determinación de factores de deserción estudiantil en la universidad técnica de Cotopaxi utilizando minería de datos. Latacunga: Universidad Técnica de Cotopaxi, Dirección de Posgrado. Obtenido de: <http://repositorio.utc.edu.ec/bitstream/27000/4501/1/PI-000727.pdf>
- Chow, V., & Flores, W. (2021). Deserción estudiantil: Un estudio en una Universidad Comunitaria Intercultural. *Revista Electrónica de Conocimientos, Saberes y Prácticas*, 4(2).
- Coa, C. (2019). Consecuencias de alta multicolinealidad en un modelo de regresión lineal. *Revista Varianza*, 16, 22–27.
- Cortez Torrez, J. A. (2018). El marco teórico referencial y los enfoques de investigación. *Apthapi* 4(1), 1036–1062. Obtenido de: <http://www.ojs.agro.umsa.bo/index.php/ATP/article/view/198>
- Cuji, B., Gavilanes, W., & Sanchez, R. (2017). Modelo predictivo de deserción estudiantil basado en arboles de decisión. *Espacios*, 38(55), 17. Obtenido de: <http://www.revistaespacios.com/a17v38n55/a17v38n55p17.pdf>
- De Franco, M. F., & Solorzano, J. L. V. (2020). Paradigmas, enfoques y métodos de investigación: análisis teórico. *Mundo Recursivo*, 3(1), 1–24. Obtenido de: <https://www.atlantic.edu.ec/ojs/index.php/mundor/article/view/38>
- Díaz, L., & Morales, M. (2012). *Estadística Multivariada: Inferencia y Métodos*. Bogotá: Universidad Nacional de Colombia.
- Díaz, P., De León, A. T., & Saavedra, L. M. (2021). Modelos de análisis y prevención de deserción estudiantil universitaria dirigidos al contexto panameño. *South Florida Journal of Development*, 2(2), 1349–1357. Obtenido de: <https://doi.org/10.46932/sfjdv2n2-018>
- Espinoza Freire, E. (2019). Las variables y su operacionalización en la Investigación Educativa. *Revista Conrado*, 15(69), 171–180. Obtenido de: <http://conrado.ucf.edu.cu/index.php/conrado>.
- Galindo, C., & Ramírez, T. (2018). Factores psicológicos, sociológicos e interaccionistas que inciden en la permanencia y/o deserción de estudiantes universitarios. Bogotá: Universidad Santo Tomás, Facultad de Cultura Física Deporte y Recreación. Obtenido de:

<https://repository.usta.edu.co/handle/11634/15450>

- García Fernández, B., Jiménez, A. M., & Bejarano Franco, M. T. (2016). Training Teachers with a Virtual Learning Community: Connecting Peers with an International Dimension. *Pedagogika*, 122(2), 124–140. Obtenido de: <https://doi.org/10.15823/p.2016.25>
- González, J., Carvajal, C., & Aspeé, J. (2020). Modelación de la deserción universitaria mediante cadenas de Markov. *Uniciencia*, 34, 129-146. Obtenido de: <https://www.redalyc.org/jatsRepo/4759/475962995008/html/index.html>
- González, M., Miquilena, L., Rebolledo, A., & Vásquez, H. (2018). Deserción Estudiantil en la Universidad del Zulia ¿Problema Coyuntural? *Revista Venezolana De Gerencia*, 23(1), 403–431.
- Guerrero, J. (2018). Validación de constructo a un cuestionario relacionado con el diagnóstico estratégico de las TIC en la educación superior. Caso de estudio. *Acción Pedagógica*, 27, 22–33.
- Guerrero-Ortiz, C., Morales-Soto, A., Ramos-Rodriguez, E., Alarcón Relmucao, N. A., Aravena Díaz, M., Bonilla Barraza, D., & Cabrera Baqueadano, A. (2021). Aportes desde la didáctica de la matemática, innovar y mejorar en y sobre la práctica docente: Modelación matemática. Barcelona: Graó.
- Guevara, G., Verdesoto, A., & Castro, N. (2020). Metodologías de investigación educativa (descriptivas, experimentales, participativas, y de investigación-acción). *Revista Científica Mundo de La Investigación y El Conocimiento*, 3, 163–173. Obtenido de: [https://doi.org/10.26820/recimundo/4.\(3\).julio.2020.163-173](https://doi.org/10.26820/recimundo/4.(3).julio.2020.163-173)
- Gutiérrez-Cánovas, C., & Escribano-Ávila, G. (2019). Inferencia estadística a partir de varios modelos y su utilidad en ecología. *Ecosistemas*, 28(1), 118–120. Obtenido de: <https://doi.org/10.7818/ecos.1699>
- Hair, J., Rolph, A., Ronald, T., & Black, W. (1999). *Análisis Multivariante* (5ª ed.). Madrid. Prentice Hall Iberia.
- Hernández, J., Espinosa, F., Rodríguez, J. E., Chacón, J., Toloza, C., Arenas, M., Carrillo, S., & Bermúdez, V. (2018). Sobre el uso adecuado del coeficiente de correlación de Pearson. *Archivos Venezolanos de Farmacología y Terapéutica*, 37(5). Obtenido de: <https://www.redalyc.org/articulo.oa?id=55963207025>
- Hernández, R. V. (2021). Procesos de modelación matemática y modelación estadística en tiempos contemporáneos: similitudes y diferencias. *Revista Boletín Redipe*, 10(7), 334–356. Obtenido de: <https://doi.org/10.36260/rbr.v10i7.1372>
- Herrero, V., Merlino, A., Ayllón, S., Escanés, G., & P, L. B. C. (2013). Aplicación de un modelo de duración en programas de prevención de deserción universitaria. *Revista Electrónica de Investigación Educativa*, 15(3), 38–52. Obtenido de: <http://redie.uabc.mx/index.php/redie/article/view/554/812>

- Hidalgo, A. (2019). Técnicas Estadísticas en el Análisis cuantitativo de datos. *Revista Sigma*, 15(3), 44–48.
- Huincahue, J., Borromeo-Ferri, R., & Mena-Lorca, J. (2018). El conocimiento de la modelación matemática desde la reflexión en la formación inicial de profesores de matemática. *Enseñanza de Las Ciencias*, 36(1), 99–115. Obtenido de: <https://doi.org/10.5565/rev/ensciencias.2277>
- Instituto Nacional de Estadísticas y Censos. (2015). Análisis y Proyección de la Población Económicamente Activa del Ecuador. *Estudios Demográficos En Profundidad*, 4–5. Quito. Obtenido de: [http://www.ecuadorencifras.gob.ec/documentos/web-inec/Bibliotecas/Estudios/Estudios_Socio-demograficos/AnalisisyProyeccion_de_la_Poblacion_Economicamente_Activa_\(PEA\)_del_Ecuador.pdf](http://www.ecuadorencifras.gob.ec/documentos/web-inec/Bibliotecas/Estudios/Estudios_Socio-demograficos/AnalisisyProyeccion_de_la_Poblacion_Economicamente_Activa_(PEA)_del_Ecuador.pdf).
- Lázaro Alvarez, N., Callejas, Z., Griol, D., & Durán Benejam, M. (2017). La deserción estudiantil en educación superior: S.O.S. en carreras de ingeniería informática. En Conferencia: *Séptima conferencia latinoamericana sobre el Abandono en la Educación Superior CLABES*. Córdoba, Argentina: Universidad Nacional de Córdoba. Obtenido de: <http://www.revistas.utp.ac.pa/index.php/clabes/article/view/1674/2410>
- Lázaro, N. (2020). Acciones tutoriales con tic atendiendo a factores predictivos de la deserción estudiantil en carreras de ingeniería informática. Granada: Universidad de Granada, Programa de Doctorado en ciencias de la Educación. Obtenido de: <https://digibug.ugr.es/handle/10481/64571>
- Matos, F., Contreras, F., & Olaya, C. J. (2020). *Estadística descriptiva y probabilidad para las ciencias de la información con el uso del SPSS* (1ª ed.). Lima. Asociación de Bibliotecólogos del Perú. Obtenido de: http://eprints.rclis.org/40470/1/ESTADISTICA_DESCRIPTIVA.pdf
- Mejía, J., & Vanegas, J. (2018). La deserción de aprendices SENA. Versión preliminar para la Corporación Universitaria Minuto de Dios. Manuscrito no publicado.
- Meneses, J. (2019). *Introducción al análisis multivariante*. Barcelona. FUOC.
- Mercedes, L., & Gutiérrez, L. (2019). Cómo realizar e interpretar un análisis factorial exploratorio utilizando SPSS. *REIRE Revista d Innovación i Recerca En Educació*, 12 (2), 1–14. Obtenido de: <https://doi.org/10.1344/reire2019.12.227057>
- Minchón, C. (2020). Multilinealidad y falta de ajuste en superficies de respuesta estimadas en investigaciones publicadas de scientia agropecuaria. *Revista Electrónica Universidad Nacional de Trujillo*, 3(1), 43–57.
- Montánchez, M. L., & Martínez, P. C. (2017). Plan de Acción Tutorial como actividad pedagógico formativa en la Universidad Regional Amazónica, IKIAM. *CienciAmérica*, 6(2), 76–81.
- Moreno Tapia, J. (2014). Tipos de Investigación. *DIVULGARE Boletín Científico De La Escuela*

Superior De Actopan.

- Ortiz, A. (2015). *Enfoques y métodos de investigación*. Bogotá. Buena Semilla.
- Páramo, G., & Correa, C. (2019). Deserción Estudiantil Universitaria. Contextualización. *Revista Universidad EAFIT*, 14
- Pérez, C. (2004). *Técnicas de análisis multivariante de datos*. Madrid. Pearson Prentice Hall.
- Ponce, J., & Carrasco, F. (2017). Acceso y equidad a la educación superior y posgrado en el Ecuador, un enfoque descriptivo. *Mundos Plurales - Revista Latinoamericana de Políticas y Acción Pública*, 3(2), 9. Obtenido de: <https://doi.org/10.17141/mundosplurales.2.2016.2841>
- Poveda, J. C., Poveda, I. M., & España, I. A. (2020). Análisis de la deserción estudiantil en una universidad pública de Bolivia. *Revista Iberoamericana de Educación*, 82(2), 151–172. Obtenido de: <https://doi.org/10.35362/rie8223572>
- Ramírez, V. (2021). Deserción Estudiantil y el Costo Económico en Universidades Chilenas. Concepción: Universidad del Bio-Bio, Programa Intensivo.
- Ramos-Galarza, C. A. (2020). Los Alcances de una investigación. *CienciAmérica*, 9(3), 5. Obtenido de: <https://doi.org/10.33210/ca.v9i3.336>
- Ramos, J. R., Del Águila, V., & Bazalar, A. B. (2020). *Estadística básica para los negocios*. Lima. Fondo Editorial.
- Ruiz-Benito, P., Andivia, E., Archambeau, J., Astigarraga, J., Barrientos, R., Cruz-Alonso, V., Florencio, M., Gómez, D., Martínez-Baroja, L., Quiles, P., Rohrer, Z., Santos, A. M. C., Velado, E., Villén-Pérez, S., & Morales-Castilla, I. (2018). Ventajas de la estadística bayesiana frente a la frecuentista: ¿por qué nos resistimos a usarla? *Ecosistemas*, 27(2), 136–139. Obtenido de: <https://doi.org/10.7818/ecos.1591>
- Sagaró del Campo, N. M., & Zamora, L. (2020). Técnicas estadísticas para identificar posibles relaciones bivariadas. *Revista Cubana de Anestesiología y Reanimación*, 19(2), 1–23. Obtenido de: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1726-67182020000200008
- Secretaría de Educación Superior, Ciencia, Tecnología e Innovación. (2018). Educación Superior, Ciencia, Tecnología, Innovación y Saberes Ancestrales en Cifras. Quito: Castañeda, J.
- Secretaría de Educación Superior, Ciencia, Tecnología e Innovación. (2020). Educación Superior, Ciencia, Tecnología, Innovación y Saberes Ancestrales en Cifras. Versión 2.0, 9. Quito.
- Secretaría de Educación Superior, Ciencia, Tecnología e Innovación. (1 de abril de 2021). Secretaría de Educación Superior, Ciencia, Tecnología e Innovación. Obtenido de: <https://www.educacionsuperior.gob.ec/94-349-cupos-estaran-disponibles-en-218-instituciones-de-educacion-superior/>
- Sierra, L. (2021). Cálculo de Modelos de Regresión Robusta para la Valoración Masiva de Construcciones Residenciales No PH en la zona UPZ 65 - Arborizadora. Bogotá:

- Universidad Distrital Francisco José Caldas, Facultad de Ingeniería especialización en avalúos.
- Smulders Chaparro, M. E. (2018). Factores que influyen en la deserción de los Estudiantes Universitarios. *Academo Revista de Investigación En Ciencias Sociales y Humanidades*, 5(2), 127–132. Obtenido de: <https://doi.org/10.30545/academo.2018.jul-dic.5>
- Soria, E. (2021). Modelos de ecuaciones estructurales en innovación de empresas de manufactura peruanas con información del Instituto Nacional de Estadística e Informática. Lima: Universidad Nacional Agraria La Molina, Facultad de economía y planificación. Obtenido de: <http://repositorio.lamolina.edu.pe/bitstream/handle/UNALM/2131/L02-C389-T.pdf?sequence=1&isAllowed=y>
- 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*, 22(1), 81–100. Obtenido de: <https://doi.org/10.5294/edu.2019.22.1.5>
- Tinto, V. (1982). Definir la deserción: una cuestión de perspectiva. 1–9. Obtenido de: <http://publicaciones.anuies.mx/>
- Torres, M. (15 de diciembre de 2019). EXPRESO. Obtenido de www.expreso.ec: <https://www.expreso.ec/guayaquil/desercion-universitaria-pais-alcanza-26-1456.html>
- Torres, W., & Guerra, S. (8 de mayo de 2020). PRIMICIAS. Obtenido de PRIMICIAS: <https://www.primicias.ec/noticias/economia/sistema-universitario-cifras-ecuador/>
- Valle, D. (2018). Modelo De Regresión Lineal Multivariado Aplicado A La Relación Entre Variables Químicas Del Clinker Y Variables Físicas Del Cemento. *In Angewandte Chemie International Edition*, 6(11), 951–952.
- Vilá, R., Torrado, M., & Reguant, M. (2019). Análisis de regresión lineal múltiple con SPSS: un ejemplo práctico. *REIRE Revista d Innovació i Recerca En Educació*, 12 (2), 1–10. Obtenido de: <https://doi.org/10.1344/reire2019.12.222704>
- Yunkor, Y., & Ochoa, J. (2021). Validación Interna de Instrumentos de Investigación en las Ciencias Sociales. *Angewandte Chemie International Edition*, 6(11), 951–952., 10–27
- Zambrano, G. J., Rodríguez, K. G., & Guevara, L. H. (2018). Análisis de la deserción estudiantil en las universidades del Ecuador y América Latina. *Revista Pertinencia Académica*, 91(5), 1689–1699. Obtenido de: <http://revista-academica.utb.edu.ec/index.php/pertacade/article/view/127/91>

ANEXOS

ANEXO A. DISTRIBUCIÓN F

$$1 - \alpha = 0.9$$

$$1 - \alpha = P(F \leq f_{\alpha, v_1, v_2})$$

| $v_2 \backslash v_1$ | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 | 200 | 500 | 1000 |
|----------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 1 | 61.815 | 61.883 | 61.945 | 62.002 | 62.055 | 62.103 | 62.148 | 62.189 | 62.229 | 62.265 | 62.529 | 62.688 | 62.794 | 62.871 | 62.927 | 62.972 | 63.007 | 63.167 | 63.264 | 63.296 |
| 2 | 9.444 | 9.446 | 9.448 | 9.450 | 9.451 | 9.453 | 9.454 | 9.456 | 9.457 | 9.458 | 9.466 | 9.471 | 9.475 | 9.477 | 9.479 | 9.480 | 9.481 | 9.486 | 9.489 | 9.490 |
| 3 | 5.182 | 5.180 | 5.178 | 5.176 | 5.175 | 5.173 | 5.172 | 5.170 | 5.169 | 5.168 | 5.160 | 5.155 | 5.151 | 5.149 | 5.147 | 5.145 | 5.144 | 5.139 | 5.136 | 5.135 |
| 4 | 3.841 | 3.837 | 3.834 | 3.831 | 3.828 | 3.826 | 3.823 | 3.821 | 3.819 | 3.817 | 3.804 | 3.795 | 3.790 | 3.788 | 3.782 | 3.780 | 3.778 | 3.769 | 3.764 | 3.762 |
| 5 | 3.202 | 3.198 | 3.194 | 3.191 | 3.187 | 3.184 | 3.181 | 3.179 | 3.178 | 3.174 | 3.157 | 3.147 | 3.140 | 3.135 | 3.132 | 3.129 | 3.126 | 3.116 | 3.109 | 3.107 |
| 6 | 2.831 | 2.827 | 2.822 | 2.818 | 2.815 | 2.811 | 2.808 | 2.805 | 2.803 | 2.800 | 2.781 | 2.770 | 2.762 | 2.756 | 2.752 | 2.749 | 2.746 | 2.734 | 2.727 | 2.725 |
| 7 | 2.589 | 2.584 | 2.580 | 2.575 | 2.571 | 2.568 | 2.564 | 2.561 | 2.558 | 2.555 | 2.535 | 2.523 | 2.514 | 2.508 | 2.504 | 2.500 | 2.497 | 2.484 | 2.476 | 2.473 |
| 8 | 2.419 | 2.414 | 2.409 | 2.404 | 2.400 | 2.396 | 2.392 | 2.389 | 2.386 | 2.383 | 2.361 | 2.348 | 2.339 | 2.333 | 2.328 | 2.324 | 2.321 | 2.307 | 2.298 | 2.295 |
| 9 | 2.292 | 2.287 | 2.282 | 2.277 | 2.272 | 2.268 | 2.265 | 2.261 | 2.258 | 2.255 | 2.232 | 2.218 | 2.208 | 2.202 | 2.196 | 2.192 | 2.189 | 2.174 | 2.165 | 2.162 |
| 10 | 2.194 | 2.189 | 2.183 | 2.178 | 2.174 | 2.170 | 2.166 | 2.162 | 2.159 | 2.155 | 2.132 | 2.117 | 2.107 | 2.100 | 2.095 | 2.090 | 2.087 | 2.071 | 2.062 | 2.059 |
| 11 | 2.117 | 2.111 | 2.105 | 2.100 | 2.095 | 2.091 | 2.087 | 2.083 | 2.080 | 2.076 | 2.052 | 2.036 | 2.026 | 2.019 | 2.013 | 2.009 | 2.005 | 1.989 | 1.979 | 1.975 |
| 12 | 2.053 | 2.047 | 2.041 | 2.036 | 2.031 | 2.027 | 2.022 | 2.019 | 2.015 | 2.011 | 1.986 | 1.970 | 1.960 | 1.952 | 1.946 | 1.942 | 1.938 | 1.921 | 1.911 | 1.907 |
| 13 | 2.000 | 1.994 | 1.988 | 1.983 | 1.978 | 1.973 | 1.969 | 1.965 | 1.961 | 1.958 | 1.931 | 1.915 | 1.904 | 1.896 | 1.890 | 1.886 | 1.882 | 1.864 | 1.853 | 1.850 |
| 14 | 1.955 | 1.949 | 1.943 | 1.938 | 1.933 | 1.928 | 1.923 | 1.919 | 1.916 | 1.912 | 1.885 | 1.869 | 1.857 | 1.849 | 1.843 | 1.838 | 1.834 | 1.816 | 1.805 | 1.801 |
| 15 | 1.917 | 1.911 | 1.905 | 1.899 | 1.894 | 1.889 | 1.885 | 1.880 | 1.876 | 1.873 | 1.845 | 1.828 | 1.817 | 1.808 | 1.802 | 1.797 | 1.793 | 1.774 | 1.763 | 1.759 |
| 16 | 1.884 | 1.877 | 1.871 | 1.866 | 1.860 | 1.855 | 1.851 | 1.847 | 1.843 | 1.839 | 1.811 | 1.793 | 1.782 | 1.773 | 1.766 | 1.761 | 1.757 | 1.738 | 1.726 | 1.722 |
| 17 | 1.855 | 1.848 | 1.842 | 1.836 | 1.831 | 1.826 | 1.821 | 1.817 | 1.813 | 1.809 | 1.781 | 1.763 | 1.751 | 1.742 | 1.735 | 1.730 | 1.726 | 1.706 | 1.694 | 1.690 |
| 18 | 1.829 | 1.823 | 1.816 | 1.810 | 1.805 | 1.800 | 1.795 | 1.791 | 1.787 | 1.783 | 1.754 | 1.736 | 1.723 | 1.714 | 1.707 | 1.702 | 1.698 | 1.678 | 1.665 | 1.661 |
| 19 | 1.807 | 1.800 | 1.793 | 1.787 | 1.782 | 1.777 | 1.772 | 1.767 | 1.763 | 1.759 | 1.730 | 1.711 | 1.699 | 1.690 | 1.683 | 1.677 | 1.673 | 1.652 | 1.639 | 1.635 |
| 20 | 1.786 | 1.779 | 1.773 | 1.767 | 1.761 | 1.756 | 1.751 | 1.746 | 1.742 | 1.738 | 1.708 | 1.690 | 1.677 | 1.667 | 1.660 | 1.655 | 1.650 | 1.629 | 1.616 | 1.612 |
| 21 | 1.768 | 1.761 | 1.754 | 1.748 | 1.742 | 1.737 | 1.732 | 1.728 | 1.723 | 1.719 | 1.689 | 1.670 | 1.657 | 1.647 | 1.640 | 1.634 | 1.630 | 1.608 | 1.595 | 1.591 |
| 22 | 1.751 | 1.744 | 1.737 | 1.731 | 1.726 | 1.720 | 1.715 | 1.711 | 1.706 | 1.702 | 1.671 | 1.652 | 1.639 | 1.629 | 1.622 | 1.616 | 1.611 | 1.590 | 1.576 | 1.571 |
| 23 | 1.736 | 1.729 | 1.722 | 1.716 | 1.710 | 1.705 | 1.700 | 1.695 | 1.691 | 1.686 | 1.655 | 1.636 | 1.622 | 1.613 | 1.605 | 1.599 | 1.594 | 1.572 | 1.558 | 1.554 |
| 24 | 1.722 | 1.715 | 1.708 | 1.702 | 1.696 | 1.691 | 1.686 | 1.681 | 1.676 | 1.672 | 1.641 | 1.621 | 1.607 | 1.597 | 1.590 | 1.584 | 1.579 | 1.556 | 1.542 | 1.538 |
| 25 | 1.710 | 1.702 | 1.695 | 1.689 | 1.683 | 1.678 | 1.672 | 1.668 | 1.663 | 1.659 | 1.627 | 1.607 | 1.593 | 1.583 | 1.576 | 1.569 | 1.565 | 1.542 | 1.527 | 1.523 |
| 26 | 1.698 | 1.690 | 1.683 | 1.677 | 1.671 | 1.666 | 1.660 | 1.656 | 1.651 | 1.647 | 1.615 | 1.594 | 1.581 | 1.570 | 1.562 | 1.556 | 1.551 | 1.528 | 1.514 | 1.509 |
| 27 | 1.687 | 1.680 | 1.673 | 1.666 | 1.660 | 1.655 | 1.649 | 1.645 | 1.640 | 1.636 | 1.603 | 1.583 | 1.569 | 1.558 | 1.550 | 1.544 | 1.539 | 1.515 | 1.501 | 1.496 |
| 28 | 1.677 | 1.669 | 1.662 | 1.656 | 1.650 | 1.644 | 1.639 | 1.634 | 1.630 | 1.625 | 1.592 | 1.572 | 1.558 | 1.547 | 1.539 | 1.533 | 1.528 | 1.504 | 1.489 | 1.484 |
| 29 | 1.668 | 1.660 | 1.653 | 1.647 | 1.640 | 1.635 | 1.630 | 1.625 | 1.620 | 1.616 | 1.583 | 1.562 | 1.547 | 1.537 | 1.529 | 1.522 | 1.517 | 1.493 | 1.478 | 1.472 |
| 30 | 1.659 | 1.651 | 1.644 | 1.638 | 1.632 | 1.626 | 1.621 | 1.616 | 1.611 | 1.606 | 1.573 | 1.552 | 1.538 | 1.527 | 1.519 | 1.512 | 1.507 | 1.482 | 1.467 | 1.462 |
| 40 | 1.596 | 1.588 | 1.581 | 1.574 | 1.568 | 1.562 | 1.556 | 1.551 | 1.546 | 1.541 | 1.508 | 1.483 | 1.467 | 1.455 | 1.447 | 1.439 | 1.434 | 1.406 | 1.389 | 1.383 |
| 50 | 1.559 | 1.551 | 1.543 | 1.536 | 1.529 | 1.523 | 1.517 | 1.512 | 1.507 | 1.502 | 1.465 | 1.441 | 1.424 | 1.412 | 1.402 | 1.395 | 1.388 | 1.359 | 1.340 | 1.333 |
| 60 | 1.534 | 1.526 | 1.518 | 1.511 | 1.504 | 1.498 | 1.492 | 1.486 | 1.481 | 1.476 | 1.437 | 1.413 | 1.395 | 1.382 | 1.372 | 1.364 | 1.358 | 1.326 | 1.306 | 1.299 |
| 70 | 1.517 | 1.508 | 1.500 | 1.493 | 1.486 | 1.479 | 1.473 | 1.467 | 1.462 | 1.457 | 1.418 | 1.392 | 1.374 | 1.361 | 1.350 | 1.342 | 1.335 | 1.302 | 1.281 | 1.273 |
| 80 | 1.503 | 1.495 | 1.487 | 1.479 | 1.472 | 1.465 | 1.459 | 1.453 | 1.448 | 1.443 | 1.403 | 1.377 | 1.358 | 1.344 | 1.334 | 1.325 | 1.318 | 1.284 | 1.261 | 1.253 |
| 90 | 1.493 | 1.484 | 1.476 | 1.468 | 1.461 | 1.455 | 1.448 | 1.442 | 1.437 | 1.432 | 1.391 | 1.365 | 1.346 | 1.332 | 1.321 | 1.312 | 1.304 | 1.269 | 1.245 | 1.237 |
| 100 | 1.485 | 1.476 | 1.468 | 1.460 | 1.453 | 1.446 | 1.440 | 1.434 | 1.428 | 1.423 | 1.382 | 1.355 | 1.336 | 1.321 | 1.310 | 1.301 | 1.293 | 1.257 | 1.232 | 1.223 |
| 200 | 1.448 | 1.438 | 1.430 | 1.422 | 1.414 | 1.407 | 1.400 | 1.394 | 1.388 | 1.383 | 1.339 | 1.310 | 1.289 | 1.273 | 1.261 | 1.250 | 1.242 | 1.199 | 1.168 | 1.157 |
| 500 | 1.425 | 1.416 | 1.407 | 1.399 | 1.391 | 1.384 | 1.377 | 1.370 | 1.364 | 1.358 | 1.313 | 1.282 | 1.260 | 1.243 | 1.229 | 1.218 | 1.209 | 1.160 | 1.122 | 1.106 |
| 1000 | 1.418 | 1.408 | 1.399 | 1.391 | 1.383 | 1.376 | 1.369 | 1.362 | 1.356 | 1.350 | 1.304 | 1.273 | 1.250 | 1.232 | 1.218 | 1.207 | 1.197 | 1.145 | 1.103 | 1.084 |

ANEXO B. DISTRIBUCIÓN T

Tabla t-Student



| Grados de libertad | 0.25 | 0.1 | 0.05 | 0.025 | 0.01 | 0.005 |
|--------------------|--------|--------|--------|---------|---------|---------|
| 1 | 1.0000 | 3.0777 | 6.3137 | 12.7062 | 31.8210 | 63.6559 |
| 2 | 0.8165 | 1.8856 | 2.9200 | 4.3027 | 6.9645 | 9.9250 |
| 3 | 0.7649 | 1.6377 | 2.3534 | 3.1824 | 4.5407 | 5.8408 |
| 4 | 0.7407 | 1.5332 | 2.1318 | 2.7765 | 3.7469 | 4.6041 |
| 5 | 0.7267 | 1.4759 | 2.0150 | 2.5706 | 3.3649 | 4.0321 |
| 6 | 0.7176 | 1.4398 | 1.9432 | 2.4469 | 3.1427 | 3.7074 |
| 7 | 0.7111 | 1.4149 | 1.8946 | 2.3646 | 2.9979 | 3.4995 |
| 8 | 0.7064 | 1.3968 | 1.8595 | 2.3060 | 2.8965 | 3.3554 |
| 9 | 0.7027 | 1.3830 | 1.8331 | 2.2622 | 2.8214 | 3.2498 |
| 10 | 0.6998 | 1.3722 | 1.8125 | 2.2281 | 2.7638 | 3.1693 |
| 11 | 0.6974 | 1.3634 | 1.7959 | 2.2010 | 2.7181 | 3.1058 |
| 12 | 0.6955 | 1.3562 | 1.7823 | 2.1788 | 2.6810 | 3.0545 |
| 13 | 0.6938 | 1.3502 | 1.7709 | 2.1604 | 2.6503 | 3.0123 |
| 14 | 0.6924 | 1.3450 | 1.7613 | 2.1448 | 2.6245 | 2.9768 |
| 15 | 0.6912 | 1.3406 | 1.7531 | 2.1315 | 2.6025 | 2.9467 |
| 16 | 0.6901 | 1.3368 | 1.7459 | 2.1199 | 2.5835 | 2.9208 |
| 17 | 0.6892 | 1.3334 | 1.7396 | 2.1098 | 2.5669 | 2.8982 |
| 18 | 0.6884 | 1.3304 | 1.7341 | 2.1009 | 2.5524 | 2.8784 |
| 19 | 0.6876 | 1.3277 | 1.7291 | 2.0930 | 2.5395 | 2.8609 |
| 20 | 0.6870 | 1.3253 | 1.7247 | 2.0860 | 2.5280 | 2.8453 |
| 21 | 0.6864 | 1.3232 | 1.7207 | 2.0796 | 2.5176 | 2.8314 |
| 22 | 0.6858 | 1.3212 | 1.7171 | 2.0739 | 2.5083 | 2.8188 |
| 23 | 0.6853 | 1.3195 | 1.7139 | 2.0687 | 2.4999 | 2.8073 |
| 24 | 0.6848 | 1.3178 | 1.7109 | 2.0639 | 2.4922 | 2.7970 |
| 25 | 0.6844 | 1.3163 | 1.7081 | 2.0595 | 2.4851 | 2.7874 |
| 26 | 0.6840 | 1.3150 | 1.7056 | 2.0555 | 2.4786 | 2.7787 |
| 27 | 0.6837 | 1.3137 | 1.7033 | 2.0518 | 2.4727 | 2.7707 |
| 28 | 0.6834 | 1.3125 | 1.7011 | 2.0484 | 2.4671 | 2.7633 |
| 29 | 0.6830 | 1.3114 | 1.6991 | 2.0452 | 2.4620 | 2.7564 |
| 30 | 0.6828 | 1.3104 | 1.6973 | 2.0423 | 2.4573 | 2.7500 |
| 31 | 0.6825 | 1.3095 | 1.6955 | 2.0395 | 2.4528 | 2.7440 |
| 32 | 0.6822 | 1.3086 | 1.6939 | 2.0369 | 2.4487 | 2.7385 |
| 33 | 0.6820 | 1.3077 | 1.6924 | 2.0345 | 2.4448 | 2.7333 |
| 34 | 0.6818 | 1.3070 | 1.6909 | 2.0322 | 2.4411 | 2.7284 |
| 35 | 0.6816 | 1.3062 | 1.6896 | 2.0301 | 2.4377 | 2.7238 |
| 36 | 0.6814 | 1.3055 | 1.6883 | 2.0281 | 2.4345 | 2.7195 |
| 37 | 0.6812 | 1.3049 | 1.6871 | 2.0262 | 2.4314 | 2.7154 |
| 38 | 0.6810 | 1.3042 | 1.6860 | 2.0244 | 2.4286 | 2.7116 |
| 39 | 0.6808 | 1.3036 | 1.6849 | 2.0227 | 2.4258 | 2.7079 |
| 40 | 0.6807 | 1.3031 | 1.6839 | 2.0211 | 2.4233 | 2.7045 |
| 41 | 0.6805 | 1.3025 | 1.6829 | 2.0195 | 2.4208 | 2.7012 |
| 42 | 0.6804 | 1.3020 | 1.6820 | 2.0181 | 2.4185 | 2.6981 |
| 43 | 0.6802 | 1.3016 | 1.6811 | 2.0167 | 2.4163 | 2.6951 |
| 44 | 0.6801 | 1.3011 | 1.6802 | 2.0154 | 2.4141 | 2.6923 |
| 45 | 0.6800 | 1.3007 | 1.6794 | 2.0141 | 2.4121 | 2.6896 |
| 46 | 0.6799 | 1.3002 | 1.6787 | 2.0129 | 2.4102 | 2.6870 |
| 47 | 0.6797 | 1.2998 | 1.6779 | 2.0117 | 2.4083 | 2.6846 |
| 48 | 0.6796 | 1.2994 | 1.6772 | 2.0106 | 2.4066 | 2.6822 |
| 49 | 0.6795 | 1.2991 | 1.6766 | 2.0096 | 2.4049 | 2.6800 |