



Factores asociados a la deserción estudiantil para los estudiantes pertenecientes al Programa ASES en el año 2017

Nathalia Gonzalez Ortiz
Jhonn Wilman Martinez Diafara

Universidad del Valle
Facultad de Ingeniería, Escuela de Estadística
Santiago de Cali, Colombia
2022

Factores asociados a la deserción estudiantil para los estudiantes pertenecientes al programa ASES en el año 2017

Nathalia Gonzalez Ortiz
Jhonn Wilman Martinez Diafara

Trabajo de grado presentado como requisito parcial para optar al título de:
Estadístico(a)

Director(a):
Ph.D. Alvaro José Flórez Poveda
Codirector(a):
Mg.Ed. Fabio Andrés Barbosa Gomez

Universidad del Valle
Facultad de Ingeniería, Escuela de Estadística
Santiago de Cali, Colombia
2022

Agradecimientos

Primeramente queremos darle gracias a Dios por darnos la vida, habernos guiado a lo largo de ella, por ser nuestro apoyo, nuestra luz y nuestro camino. Gracias a nuestros padres y hermanos por su apoyo y compañía a lo largo de nuestra carrera universitaria. Gracias a nuestras parejas por su amor incondicional y apoyarnos cuando nuestros ánimos decaían.

También le damos las gracias a nuestro tutor Álvaro Florez, quien con su conocimiento y apoyo nos guió a lo largo del desarrollo de este trabajo de grado. Gracias al Programa ASES y al señor Fabio Barbosa por brindarnos todos los recursos y herramientas que fueron necesarios para llevar a cabo el proceso de investigación. No hubiésemos podido arribar a estos resultados de no haber sido por su incondicional ayuda.

Por último, gracias a todos nuestros compañeros, amigos, profesores y en especial a nuestras monitoras socieducativas de ASES Alejandra Oleas, Daniela Sarria y Catalina Medina.

Muchas gracias a todos.

Resumen

El siguiente trabajo tiene como objetivo estudiar los factores que influyen en la deserción de los estudiantes admitidos por Condición de Excepción (CE2017A y CE2017B) y Ser Pilo Paga III (SPP3) en el año 2017 durante su ciclo básico de formación, para ello se analizarán los riesgos que tienen 472 estudiantes en las diferentes dimensiones con las que trabaja el programa ASES, mediante un modelo logístico en el que se tiene como efecto aleatorio la Facultad a la que pertenece el estudiante, pues se tiene como hipótesis que es un factor de correlación. En el análisis se encontró que los factores que inciden en la deserción son: los riesgos en las dimensiones académica, familiar y vida universitaria, también el número de encuentros, el número de bajos, inasistencias a los encuentros y el sexo del estudiante; destaca que la condición de admisión a la que pertenece el estudiante no es significativa, es decir que el comportamiento de la deserción es similar en los estudiantes.

Palabras clave: Deserción, Condición de Excepción, Ser Pilo Paga, ASES, Modelo logístico con intercepto aleatorio

Abstract

The following work aims to study the factors that influence the dropout of students admitted by exception condition (CE2017A and CE2017B) and Ser Pilo Paga III (SPP3) in 2017 during their basic professional training cycle, for this it will be analyzed the risks that 472 students have in the different dimensions with which the ASES program works, through a logistic model in which the student's faculty is taken as a random effect, since it is hypothesized that it is a correlation factor. In the analysis, it was found that the factors that affect desertion are: the risks in the academic, family and university life dimensions, as well as the number of meetings, the number of absences, absences from meetings and the student's sex; highlights that the admission condition to which the student belongs is not significant, that is, the behavior of desertion is similar in students.

Keywords: Desertion, Exception condition, Ser Pilo Paga, ASES, logistic model with random intercept

Contenido

Resumen	IV
Lista de Figuras	VII
Lista de Tablas	VIII
1. Introducción	1
1.1. Planteamiento del Problema	2
1.2. Justificación	4
1.3. Pregunta de Investigación	5
2. Objetivos	6
2.1. Objetivo General	6
2.2. Objetivos Específicos	6
3. Antecedentes	7
4. Marco Teórico	12
4.1. Marco Teórico Conceptual	12
4.1.1. Deserción Estudiantil	12
4.1.2. Programa de Acompañamiento y Seguimiento Estudiantil (ASES) . .	12
4.1.3. Semáforo de Alertas	12
4.1.4. Dimensiones	13
4.1.5. Temáticas y Niveles de Riesgo	14
4.1.6. Ser Pilo Paga	14
4.1.7. Condición de Excepción	14
4.2. Marco Teórico Estadístico	17
4.2.1. Modelos Lineales Generalizados	17
4.2.2. Familia Exponencial	17
4.2.3. Modelo Logístico	17
4.2.4. Datos por Cluster	18
4.2.5. Modelo Lineal Mixto	18
4.2.6. Modelo Lineal Mixto Generalizado	19
4.2.7. Modelo Logístico con Intercepto Aleatorio	19
4.2.8. Estimación por Máxima Verosimilitud	20

4.2.9. Método de Cuadratura Gaussiana	20
4.2.10. Estimación Bayesiana Empírica	21
4.2.11. Selección de Variables	21
4.2.12. Criterios de Información	22
4.2.13. Pruebas de Hipotesis	22
4.2.14. Matriz de Confusión	23
5. Metodología	25
5.1. Población Objeto de Estudio	26
5.2. Consolidación de la Información	26
5.2.1. Primera Base: Información Académica	26
5.2.2. Segunda Base: Puntuaciones Encuentros ASES	26
5.2.3. Base de Datos Final	27
5.3. Unidad Experimental	29
5.4. Datos de Entrenamiento y Datos de Prueba	30
5.5. Modelo Logístico	30
5.6. Modelo Logístico con Intercepto Aleatorio	30
5.7. Método de Cuadratura Gaussiana	31
5.8. Selección de Variables	31
5.9. Pruebas de Hipótesis	31
5.10. Diagnóstico del Modelo	32
5.11. Efectos Aleatorios	32
6. Resultados	33
6.1. Análisis Exploratorio	33
6.2. Modelo Logístico con Intercepto Aleatorio	38
6.3. Selección de Variables	39
6.4. Test de Razón de Verosimilitud	39
6.5. Modelo Logístico con Intercepto Aleatorio Reducido	40
6.6. Diagnóstico del Modelo	41
6.6.1. Modelo Completo	41
6.6.2. Modelo Reducido	42
6.7. Efectos Aleatorios	43
7. Conclusiones y Recomendaciones	44
7.1. Conclusiones	44
7.2. Recomendaciones	45
A. Anexo: Temáticas y Niveles de Riesgo	49

Lista de Figuras

4-1. Dimensiones semáforo de alertas. Tomado de: Estrategia ASES.	13
4-2. Matriz de confusión. Tomado de: Barrios.	24
5-1. Fases de la metodología	25
6-1. Inasistencias	35
6-2. Riesgos según cada dimensión	36
6-3. Índice de deserción por tipo de admisión en cada semestre	37
6-4. Curva ROC: Modelo completo	41
6-5. Curva ROC: Modelo reducido	42
A-1. Autoconocimiento. Tomado de: Estrategia ASES.	49
A-2. Proyecto de vida. Tomado de: Estrategia ASES.	50
A-3. Historia de vida. Tomado de: Estrategia ASES.	50
A-4. Rasgos de personalidad. Tomado de: Estrategia ASES.	51
A-5. Salud. Tomado de: Estrategia ASES.	51
A-6. Relación erótica-afectiva. Tomado de: Estrategia ASES.	52
A-7. Identificación. Tomado de: Estrategia ASES.	52
A-8. Aspectos motivacionales. Tomado de: Estrategia ASES.	53
A-9. Diversidad sexual. Tomado de: Estrategia ASES.	53
A-10Desempeño académico. Tomado de: Estrategia ASES.	54
A-11Elección vocacional. Tomado de: Estrategia ASES.	55
A-12Manejo del tiempo. Tomado de: Estrategia ASES.	55
A-13Apoyos económicos institucionales. Tomado de: Estrategia ASES.	56
A-14Apoyo económico familiar. Tomado de: Estrategia ASES.	56
A-15Situación laboral y ocupacional. Tomado de: Estrategia ASES.	57
A-16Motivaciones para el acompañamiento. Tomado de:Estrategia ASES.	57
A-17Referenciación geográfica. Tomado de:Estrategia ASES.	58
A-18Adaptación a la ciudad y universidad. Tomado de:Estrategia ASES.	58
A-19Oferta de servicios. Tomado de: Estrategia ASES.	59
A-20Vinculación a grupos y realización de actividades extracurriculares. Tomado de: Estrategia ASES.	59
A-21Dinámica familiar. Tomado de: Estrategia ASES.	60

Lista de Tablas

5-1. Información académica	26
5-2. Puntuaciones encuentros ASES	27
5-3. Base de datos final	28
5-4. Descripción de las variables	29
5-5. Método de Cuadratura Gaussiana	31
5-6. Matriz de confusión	32
6-1. Índice de deserción temprana por Facultad	33
6-2. Índice de deserción temprana por tipo de admisión	34
6-3. Índice de deserción temprana por Facultad y tipo de admisión	34
6-4. Índice de deserción temprana por Condición de Excepción	34
6-5. Promedio de encuentros por semestre	35
6-6. Matricula académica por semestre	37
6-7. Resumen del modelo completo	38
6-8. Selección de variables: Método backward	39
6-9. Test de razón de verosimilitud	39
6-10. Resumen del modelo reducido	40
6-11. Matriz de confusión: Modelo completo	42
6-12. Matriz de confusión: Modelo reducido	43
6-13. Estimación efectos aleatorios	43

1. Introducción

Uno de los principales problemas que enfrenta el sistema de educación superior colombiano concierne a los altos niveles de deserción académica en el pregrado según el Ministerio de Educación Nacional (Ruiz et al., 2009). Pese a que los últimos años se han caracterizado por aumentos de cobertura e ingreso de estudiantes nuevos, el número de alumnos que logra culminar sus estudios superiores no es alto, dejando entrever que una gran parte de éstos abandonan sus estudios, principalmente en los primeros semestres. Según estadísticas del Ministerio de Educación Nacional (MEN), de cada cien estudiantes que ingresan a una institución de educación superior cerca de la mitad no logra culminar su ciclo académico y obtener la graduación (Ruiz et al., 2009).

La deserción a nivel nacional para la carrera de formación universitaria para el año 2017 fue de 9,15 % y para la formación tecnológica de 10,52 % según estadísticas del Sistema para la Prevención y Análisis de la Deserción en las Instituciones de Educación Superior (SPADIES, s.f.).

El programa de Acompañamiento y Seguimiento Estudiantil (ASES) de la Universidad del Valle, diseña e implementa estrategias de acompañamiento a los estudiantes de las diferentes cohortes de Ser Pilo Paga, Condición de Excepción y Generación E en su proceso de adaptación, autoconocimiento y transformación frente a las nuevas exigencias que les plantea su vida como universitarios (Estrategia ASES, 2022a).

La estrategia ASES produce información ágil y oportuna sobre la experiencia que los estudiantes están viviendo en la Universidad: su desempeño académico, dificultades, necesidades, retos, etc. Esta información resulta valiosa para realizar actividades de apoyo a los estudiantes y ajustar el diseño de estrategias pertinentes que garanticen la permanencia de los estudiantes (Estrategia ASES, 2022a).

La principal motivación del proyecto se debe a nuestra participación dentro del programa ASES, bajo el programa de Condición de Excepción y Ser Pilo Paga II, esto hizo que nos planteáramos la pregunta sobre qué factores influyen en la deserción de los estudiantes durante su admisión y acompañamiento del programa ASES; para ello se utilizará un modelo de regresión logística donde se incluye a la Facultad como efecto aleatorio, ya que esta metodología estadística permite analizar tanto datos longitudinales como datos

agrupados (cluster).

La estructura del trabajo se desarrolla de la siguiente manera: En el Capítulo 1 se discute el planteamiento del problema y su justificación. En el Capítulo 2 se presentan los objetivos propuestos. En el Capítulo 3 se hace la revisión de la literatura, donde se presentan antecedentes que abordan el problema de la deserción dentro de la Universidad del Valle y el Programa ASES, con su respectiva metodología estadística. En el Capítulo 4 se presenta el marco teórico de la investigación dividida en dos partes, en la primera se hacen las definiciones más importantes sobre la deserción estudiantil y el Programa ASES, luego se presenta el marco teórico estadístico el cual contiene toda la información de la metodología estadística a implementar. En el Capítulo 5 se plantea la metodología, donde se describen las características de la población de estudio y una serie de pasos del respectivo análisis estadístico. En el Capítulo 6 se presentan los resultados obtenidos en el análisis exploratorio y de la modelación estadística. Finalmente, en el Capítulo 7 se hace la discusión de las conclusiones y recomendaciones del trabajo de grado.

1.1. Planteamiento del Problema

En el estudio realizado por Escobar et al. en la Universidad del Valle (2006), se encontró que entre los estudiantes clasificados como afrocolombianos la tasa de supervivencia disminuía de 90 % a casi 60 % a lo largo de la carrera, mientras que la tasa de supervivencia de los indígenas se encontraba en 90 % y disminuye progresivamente hasta llegar a 50 % a lo largo de la carrera. Por esta razón, el hecho de ser indígena incrementa la tasa de deserción de un estudiante.

Lozano (2018) analizó a los estudiantes de la Universidad del Valle admitidos por Condición de Excepción en el año 2017-I, donde las facultades con más alto número de estudiantes inactivos después del primer año de estudio fueron de Ingeniería, Ciencias Naturales y Exactas y Humanidades respectivamente, y la Condición de Excepción con mayor cantidad de estudiantes inactivos fue la de Comunidades Afrocolombianas.

Para el año 2017, la deserción de los estudiantes pertenecientes al programa Ser Pilo Paga III fue de 6,1 % para el primer semestre y de 8,5 % para el segundo semestre en la Universidad del Valle (Alvarado, 2019). Aunque el porcentaje de deserción es bajo, llama la atención que estudiantes beneficiarios en los dos primeros semestres hayan decidido desertar, esto se puede deber a que cambiaron de carrera, porque se dieron cuenta que no les gustó o que hayan decidido abandonar la universidad durante el primer año por otros factores (Alvarado, 2019).

Según la resolución No. 081 del 21 de mayo de 1998, se destina el 4 % de cupos para los indígenas en cada programa ofertado (Universidad del Valle, 1998), de igual manera, la resolución No. 097 del 16 de octubre de 2003, que crea la Condición de Excepción “Comunidad Afrocolombiana” donde se destina el 4 % de cupos adicionales ofertados por un programa académico (Universidad del Valle, 2003). Además, se define como comunidad negra a las personas que pertenecen a los siguientes grupos: Cuenca del Pacífico, ríos de la Cuenca del Pacífico (de la vertiente del Pacífico y las cuencas de los ríos Atrato, Acandí, y Tolo de la vertiente del Caribe), zonas rurales ribereñas, raizales de San Andrés, Providencia y Santa Catalina y demás regiones con asentamiento afrodescendiente del resto del país (Universidad del Valle, 2003).

En el Plan de Desarrollo 2005-2015 de la Universidad del Valle, se creó un proyecto llamado Universidad y Culturas, que busca tanto el ingreso como la permanencia de los estudiantes pertenecientes a las comunidades indígenas y afrodescendientes, mediante el fortalecimiento de conocimientos académicos y acompañamiento social a los estudiantes, con el fin de lograr una mejor adaptación al ambiente universitario, sin que esto implique la pérdida de identidad de las personas de estas comunidades (Universidad del Valle, 2005).

La deserción para un estudiante que ingresó a la Universidad por Condición de Excepción está asociado con la falta de adaptación académica y social (Meneses, 2009). Para evitar que esto siga sucediendo, se creó el programa ASES de la Universidad del Valle en el año 2016, que alberga a los estudiantes de Ser Pilo Paga, Generación E y Condición de Excepción, con el fin de brindar acompañamiento personal, académico y psicológico.

Se pueden diferenciar tres tipos de abandono en estudiantes universitarios: respecto al tiempo, espacio y período. La deserción con respecto al tiempo se clasifica a su vez en: deserción precoz donde el individuo que habiendo sido admitido por la Institución de Educación Superior no se matricula; deserción temprana donde el individuo abandona sus estudios en los primeros semestres del programa; y deserción tardía donde el individuo abandona los estudios en los últimos semestres de la carrera. La deserción con respecto al espacio, por su parte, se divide en: deserción institucional, caso en el cual el estudiante abandona la Institución; deserción interna o del programa académico, se refiere al alumno que decide cambiarse a otro programa que ofrece la misma Institución de Educación Superior; y la deserción del sistema, manifiesta el abandono de todo el sistema de Educación Superior. Por último, respecto al periodo tenemos: deserción por cohorte que contabiliza la deserción acumulada en cada semestre para un grupo de estudiantes que ingresaron a primer curso en un mismo período académico (cohorte); y la deserción anual, corresponde a la proporción de estudiantes que estando matriculados dos semestres atrás son clasificados como desertores un año después (Ruiz et al., 2009).

En este estudio se considera la deserción institucional de los estudiantes que ingresaron a la Universidad del Valle en el año 2017, por Condición de Excepción y Ser Pilo Paga III, a los cuales se hizo seguimiento durante los primeros 5 semestres (ciclo básico), para ver si en algún periodo desertaron de la Universidad. Por ende, se trabajará con la deserción temprana y por cohorte; dicho esto, el estudio se enfocará en analizar las 5 dimensiones con las que trabaja el programa ASES: individual (información que brinda el estudiante sobre aspectos personales), familiar (descripción de la relación del estudiante con su familia y si es un obstáculo o una ayuda en la elección y permanencia del estudiante en la carrera escogida), académica (información que el estudiante le proporciona al monitor que lo acompaña durante el semestre académico, acerca de su desempeño en las diferentes asignaturas, teniendo en cuenta sus fortalezas y debilidades), económica (información acerca de la situación económica del estudiante y el uso del dinero que tiene a su disposición) y, vida universitaria (corresponde a la opinión que tiene el estudiante sobre su vida en la universidad y la ciudad; además de la participación del estudiante en los diferentes espacios que brinda la universidad como monitorias, actividades extracurriculares, grupos universitarios, etc. y la ciudad como el transporte, el barrio, la vivienda, etc.) (Estrategia ASES, 2022a).

El estudio de los factores que influyen en la deserción es de gran importancia, puesto que afecta a los estudiantes y a la Universidad tanto socialmente como económicamente. Una persona que abandona el estudio pierde los recursos invertidos e incluso sus relaciones interpersonales pueden verse afectadas, como también su estado psicológico. Para la Universidad, un aumento en la tasa de deserción estudiantil puede llevar a una pérdida de apoyo económico por parte del Estado; por lo que la solución a esta problemática es una ayuda para los estudiantes admitidos a la Universidad, como a la Universidad misma.

1.2. Justificación

Como se expuso anteriormente, uno de los principales problemas que enfrenta el Sistema de Educación Superior colombiano está relacionado con los altos niveles de deserción académica en el pregrado, a pesar de los diferentes esfuerzos que han venido realizando en los últimos años en la Universidad del Valle, esta situación parece afectar más a los estudiantes que son admitidos bajo Condición de Excepción. Los estudiantes afrodescendientes tenían una alta tasa de deserción en los primeros semestres de la carrera que estudiaban (Lozano, 2018), la tasa de supervivencia disminuye desde 90 % a casi 60 % (Escobar et al., 2006). Además, la mitad de los estudiantes indígenas admitidos no lograron graduarse (Escobar et al., 2006).

Por otra parte, se tiene que los estudiantes pertenecientes al programa Ser Pilo Paga III, provienen de hogares con bajos recursos económicos y aunque logran buenos resultados en las pruebas de Estado ICFES, no hubieran podido estudiar en una Universidad acreditada sin la ayuda dada por el gobierno de Juan Manuel Santos.

Desde el programa ASES se realiza un acompañamiento a los estudiantes admitidos bajo Condición de Excepción y Ser Pilo Paga III con el fin de que logren una adaptación adecuada a la universidad tanto académicamente como socialmente, estableciendo si en algún momento tienen un riesgo alto de desertar según una o varias dimensiones (académica, personal, familiar, vida universitaria y económica), para tomar medidas y ver si es posible evitar esa deserción.

Este estudio trabajará con el nivel de riesgo de deserción que tiene un estudiante admitido a la Universidad del Valle, sede Cali en el año 2017, bajo Ser Pilo Paga III y Condición de Excepción, en las diferentes dimensiones con las que trabaja el programa ASES durante el ciclo básico de formación académica. Esta información está dada por el encuentro que hay entre el estudiante y el monitor de acompañamiento semana a semana, donde el monitor evalúa el nivel de riesgo en una o varias dimensiones dependiendo de la información que brinda el estudiante. La importancia de trabajar con esta información, se debe a que las situaciones de un estudiante cambian constantemente por lo que el nivel de riesgo de deserción no corresponde únicamente a su condición por la cual fue admitido, sino también a circunstancias de su entorno que pueden influir, incluso cuando un estudiante está por pasar a su ciclo profesional de la carrera.

1.3. Pregunta de Investigación

¿Cuáles son los factores que influyen en la deserción de los estudiantes admitidos por Condición de Excepción y el programa Ser Pilo Paga III a la Universidad del Valle en el año 2017 y durante su proceso de acompañamiento en el programa ASES?

2. Objetivos

2.1. Objetivo General

Determinar los factores que influyen de forma positiva y negativa en la deserción a nivel institucional en el ciclo básico de la formación académica de los estudiantes admitidos a la Universidad del Valle sede Cali, en el año 2017 pertenecientes a Condición de Excepción y Ser Pilo Paga III.

2.2. Objetivos Específicos

- Caracterizar a los estudiantes admitidos en el año 2017 a la Universidad del Valle sede Cali bajo alguna Condición de Excepción y Ser Pilo Paga III.
- Identificar las dimensiones y otros factores que influyen en la deserción de los estudiantes admitidos en el año 2017 a la Universidad del Valle sede Cali, bajo alguna Condición de Excepción y Ser Pilo Paga III.
- Ajustar un modelo predictivo para la probabilidad de deserción de los estudiantes admitidos en el año 2017 por Condición de Excepción y Ser Pilo Paga III a la Universidad del Valle sede Cali.

3. Antecedentes

La deserción académica universitaria ha sido un gran problema, donde varios investigadores desde diferentes campos de conocimiento han aportado a dicha problemática. En este capítulo se presentan algunos estudios e investigaciones relacionados con la deserción estudiantil dentro de la Universidad del Valle. Se comienza con dos investigaciones de Escobar et al. (2006 y 2008) realizadas a toda la comunidad universitaria, luego se presentan antecedentes enfocados a nuestra población de interés que son los estudiantes de Condición de Excepción y Ser Pilo Paga (Lozano, 2018; Alvarado, 2019 y Salamanca, 2019), finalmente dos antecedentes con relación a la metodología propuesta (Arismendy & Morales, 2018; García & Díaz, 2011).

Escobar et al. (2006)

Realizaron un estudio para conocer los factores asociados a la deserción y permanencia estudiantil en la Universidad del Valle durante el período 1994-2006, donde muestran secuencialmente el comportamiento de la deserción, permanencia y graduación en la universidad para las cohortes que ingresaron a la Universidad desde el 1994 hasta el 2001, esto con el fin de poder analizar a los estudiantes desde el inicio de su carrera universitaria hasta el final. En la estructuración de los datos, se contó con información de las Oficinas de Registro Académico correspondiente a información personal como: código del estudiante, número de identificación, nombres y apellidos, sexo, código y nombre del programa académico, para un total de 39.022 estudiantes. Matrícula Financiera aportó información sobre las características socioeconómicas de los estudiantes que ingresaron a la Universidad en 2001 en los dos períodos académicos con un total de 3.588 estudiantes. Bienestar Universitario entregó información sobre los estudiantes beneficiarios de subsidios para beneficiarios en el período comprendido entre 2000 y 2005 con un total de 932 estudiantes. La metodología utilizada para estimar la supervivencia, fueron el modelo no paramétrico de Kaplan-Meier y modelos paramétricos de duración, donde los datos siguieron una distribución Weibull.

Una de las dificultades que se tuvo en la realización de esta investigación fue la imprecisión de los datos con respecto al tiempo de supervivencia, por otra parte, durante el cruce de bases de datos se encontró con información repetida de estudiantes, también con estudiantes que estaban matriculados en más de un programa académico en el mismo semestre, como última dificultad fue realizar un seguimiento a los estudiantes que desertaron ya que la no respuesta es alta, especialmente esos estudiantes que llevan más de una década en la Universidad.

A nivel de resultados se encontró que características como la edad, el rendimiento en el colegio y durante el primer año de estudios en la Universidad, no convivir con los padres, ser hombre, y ser indígena son factores que incrementan el riesgo de deserción, cabe resaltar que no se contaron con estudiantes afrodescendientes admitidos por lo que esta característica no aparece; también invita a realizar una investigación a nivel de Facultad o Programa que tenga en cuenta aspectos propios del entorno y/o variables cualitativas más específicas, además de estudios que permitan reflexionar sobre los criterios de admisión a la Universidad y un seguimiento semestral de los factores que influyen en la deserción a nivel de Programa Académico y/o Facultad.

Escobar et al. (2008)

Otra investigación realizada por Jaime Escobar y su equipo de trabajo se enfocó en el rendimiento académico en la Universidad del Valle: determinantes y su relación con la deserción estudiantil que está justificado por un ejercicio previo en el que se estableció que uno de los factores más importantes para explicar el fenómeno de la deserción estudiantil en la Universidad es el rendimiento académico de los estudiantes. Para la construcción de la base de datos se tuvieron en cuenta los principales aspectos socioeconómicos, personales y académicos de los estudiantes de las cohortes de 2001, así como la de los profesores que estuvieron a cargo de los cursos matriculados por los estudiantes en el período comprendido entre los años 2001-2006. La metodología planteada fue un modelo compuesto por cuatro vectores que contienen información de los diferentes aspectos evaluados, empleó el modelo log-log que arroja como resultados las elasticidades, o la sensibilidad del rendimiento académico, ante cambios en las variables explicativas y el modelo logit en el que la variable explicativa toma el valor de 1 cuando el estudiante pierde una asignatura y 0 en otro caso. En este caso se estiman los efectos marginales de las variables explicativas incluidas en términos de la probabilidad de perder una materia.

Los resultados obtenidos muestran que características personales y socioeconómicas de los docentes pueden influir en el rendimiento académico de los estudiantes, sin embargo el investigador invita a que se realice un estudio que permita examinar las características cualitativas de los docentes, ya que pueden haber factores de confusión que produzcan un sesgo en las estimaciones; también a medida que el tamaño del grupo es grande el rendimiento académico de un estudiante disminuye, otros factores como las características socioeconómicas, provenir de un colegio público, ser menor de edad, residir en la zona de la ladera o en el distrito especial de Aguablanca, ser hombre y trabajar al momento de ingresar a la Universidad influyen negativamente en el rendimiento académico; por otra parte el rendimiento académico influye en el riesgo de desertar. Los hallazgos encontrados no se pueden generalizar hasta que se realice un estudio más enfocado a nivel de facultades o programas académicos.

Lozano (2018)

Propuso un tipo de estudio exploratorio y descriptivo. Exploratorio, porque abordará un problema que, aunque ha sido estudiado en la Universidad del Valle, no se ha realizado tomando en cuenta las condiciones al momento exacto de ocurrir el fenómeno de la deserción. Lozano propuso una aproximación a los determinantes de la deserción académica en la Universidad del Valle sede Cali, para el caso de los estudiantes admitidos por Condición de Excepción en la cohorte 2017-1 en su primer año de estudio, mediante el análisis documental de estudios y la aplicación de una encuesta a los estudiantes que desertaron. Los resultados obtenidos mostraron que el análisis cualitativo ayuda a precisar información que puede contribuir a un mejor conocimiento de lo ocurrido en el momento de desertar y permite identificar patrones de información que deberían ser tenidos en cuenta en los análisis cuantitativos. Dado que estos últimos pueden arrojar variables que ocultan información, y los resultados podrían ser explicados por diferencias particulares que no son percibidas por este tipo de análisis.

Alvarado (2019)

Una investigación enfocada en el programa ASES fue realizado por Guido Alvarado, su interés estaba enfocado en el impacto del programa en el rendimiento académico durante el ciclo básico de los estudiantes beneficiarios de Ser Pilo Paga admitidos en el primer semestre del año 2016. La metodología utilizada por Alvarado fue realizar una estimación por Propensity Score Matching usando el modelo de Roy-Rubin. Su planteamiento no experimental fue comparar el rendimiento académico de los estudiantes que siendo beneficiarios de Ser Pilo Paga pertenecen al programa ASES, con el rendimiento académico de los beneficiarios que pudieron haber pertenecido al programa ASES y ser beneficiarios de Ser Pilo Paga, para mirar una aproximación de los resultados en el rendimiento de la población de interés en caso de que no existiera el programa ASES. Una de las dificultades es la subestimación o sobreestimación del impacto que tiene el programa, ya que los estudiantes beneficiarios de ASES y no beneficiarios usualmente tienen diferencias que influyen en el rendimiento académico, por lo que el autor plantea una serie de supuestos con el fin de solucionar este problema. El principal es que la participación dentro del programa ASES y en el rendimiento está dada por variables que se pueden observar y se incluirán dentro del modelo. Se contó con la información de registro académico y socioeconómica de 98 estudiantes beneficiarios de Ser Pilo Paga II dada por ASES y 1102 estudiantes donde los registros académicos fueron dados por Registro Académico y la información socioeconómica fue dada por la Oficina de Planeación de la Universidad. Los principales hallazgos fueron una mejora en el rendimiento académico de los estudiantes pertenecientes a Ser Pilo Paga transcurrido un año de acompañamiento en el programa ASES, y que puede existir un sesgo en el rendimiento académico debido a la variable socioeconómica ya que los estudiantes que pertenecen a Ser

Pilo Paga reciben un subsidio, mientras que los estudiantes admitidos bajo otro criterio no, sin embargo, las diferencias no son significativas durante el primer año de estudios, por lo que el autor considera que el efecto de este subsidio no es significativo.

Salamanca (2019)

Con su investigación enfocada en el impacto de la estrategia ASES sobre el rendimiento académico de los estudiantes admitidos por Condición de Excepción a la Universidad del Valle en el año 2017 considera que los estudiantes pertenecientes a minorías étnicas tienen mayor riesgo en su permanencia en las Instituciones de Educación Superior, por lo que está interesado en conocer si la estrategia de acompañamiento implementada por la Universidad del Valle ha sido útil. La metodología que usa consiste en tener dos grupos, uno de estudiantes admitidos por Condición de Excepción en el año 2016 (424 estudiantes) que será un grupo control, y los estudiantes admitidos por Condición de Excepción en el año 2017 (502 estudiantes) que serán el grupo de tratamiento, dado que a partir de este año empezó el acompañamiento de la estrategia ASES a los estudiantes admitidos por Condición de Excepción; se construyó un modelo econométrico de estimación en diferencias, el cual tiene como variable de respuesta un indicador de esfuerzo que tiene en cuenta la ponderación de las asignaturas obligatorias del programa que cursa el estudiante, dando un menor peso a las electivas complementarias, y la información socioeconómica de los estudiantes que resulte significativa como factor explicativo del rendimiento estudiantil. Los resultados obtenidos mostraron que no se encontró una diferencia significativa entre el grupo control y de tratamiento, pero se da la posibilidad de que el programa ASES tenga un efecto positivo sobre el rendimiento académico, de los estudiantes admitidos por Condición de Excepción a medida que avanzan en la carrera.

Arismendy & Morales (2018)

Desarrollaron un estudio en la Universidad de los Llanos titulado Modelo de regresión logística como alternativa para medir la probabilidad de deserción temprana en la Universidad de los Llanos periodo 2015-2 - 2018-1, quienes utilizaron 13 variables sociodemográficas para estudiar la influencia que tienen en la probabilidad de deserción por medio de un modelo logístico. Para la construcción del modelo se tomó como muestra la información de 574 estudiantes que ingresaron en la cohorte 2015-2 y que para el período 2018-1 eran registrados como vigentes o reportaban como último período matriculado cualquiera de los cuatro primeros semestres. Primero calcularon los coeficientes estimados para las variables independientes, así como el estadístico Wald. Para la selección de las variables a incluir en el modelo trabajaron con un nivel de significancia ($P \leq 0,05$). Después presentaron las relaciones de probabilidad de las variables independientes. Luego

presentan los resultados de las pruebas de bondad de ajuste para validar el modelo: Estadístico Razón de Verosimilitud (Jennings, 1986), y la prueba de Hosmer & Lemesbow (1980). Finalmente evaluaron el modelo mediante los criterios de sensibilidad y especificidad (Castaño & Ramírez, 2005). Los resultados que arrojó esta investigación dan cuenta de que un buen puntaje en la prueba Saber 11, el ser mujer, no haber reprobado años durante el bachillerato, el haber cursado estudios antes y si los padres conviven disminuyen la probabilidad de deserción temprana; asimismo, el haber egresado de un colegio privado aumenta la probabilidad de deserción.

García & Díaz (2011)

Escribieron un artículo en la Revista de Economía y Estadística titulado Modelos mixtos generalizados para el estudio del desempleo en los grandes aglomerados urbanos de Argentina en el que querían identificar los factores de riesgo socioeconómicos y demográficos que influyen en la condición de empleo en los principales aglomerados urbanos de Argentina, para ello utilizaron los datos provenientes de la Encuesta Permanente de Hogares período 2004-2005. La metodología utilizada fue un modelo logístico que incorpora efectos fijos y aleatorios y evaluaron la evolución promedio marginal inducida por el modelo a través de la marginalización de sus resultados. Los principales resultados fueron que la pertenencia del individuo a un hogar con necesidades básicas insatisfechas triplica la probabilidad de desempleo, hay un mayor riesgo de desempleo para los individuos sin instrucción, no jefes de hogar y sin pareja, en la edad se advierte una disminución en la probabilidad de desempleo hasta alrededor de los 45 años, momento a partir del cual comienza nuevamente a aumentar.

4. Marco Teórico

En este capítulo se definirán los aspectos más importantes que deben conocerse para entender bien el problema, además se dividirá en dos secciones: marco teórico conceptual, el cual está relacionado con el contexto del problema de estudio y el marco teórico estadístico, el cual se enfoca en la metodología estadística utilizada.

4.1. Marco Teórico Conceptual

En esta sección se definirán los principales conceptos fundamentales relacionados con el problema de estudio, en el cual se orientan las búsquedas del investigador e identifica la metodología necesaria. Además, permite establecer un consenso mínimo entre el investigador y el lector al respecto del lenguaje y los conceptos manejados.

4.1.1. Deserción Estudiantil

Estado de un estudiante que de manera voluntaria o forzosa no registra matrícula por dos o más períodos académicos consecutivos del programa en el que se matriculó; y no se encuentra como graduado, o retirado por motivos disciplinarios (Ruiz et al., 2009).

4.1.2. Programa de Acompañamiento y Seguimiento Estudiantil (ASES)

El programa ASES fue creado en el año 2016 con el fin de acompañar a los estudiantes de Ser Pilo Paga en sus diferentes versiones en el ciclo básico de la carrera que estudian en la Universidad del Valle, con el fin de disminuir la deserción académica, y en el año 2017 empezó el acompañamiento a los estudiantes admitidos a la Universidad del Valle por Condición de Excepción (Estrategia ASES, 2022a).

4.1.3. Semáforo de Alertas

El Semáforo de Alertas es un desarrollo metodológico de la Estrategia ASES que integra el conocimiento de diferentes disciplinas, alrededor de los fenómenos de la deserción y fracaso académico, permitiendo realizar una aproximación a la comprensión de las diferentes situaciones que experimentan los estudiantes y los riesgos que pueden estar asociados a las

mismas (Estrategia ASES, 2022b).

La herramienta del Semáforo de Alertas considera un amplio abanico de situaciones que se encuentran categorizadas en 5 dimensiones: individual, familiar, académica, económica y de vida universitaria y ciudad. A su vez, cada dimensión incluye una serie de temáticas que con frecuencia son abordadas durante los acompañamientos. Finalmente, estas temáticas se desarrollan mediante una serie de descriptores que sirven de guía al monitor socioeducativo para categorizar los niveles de riesgo en cada caso (Estrategia ASES, 2022b).

4.1.4. Dimensiones

Según Estrategia ASES, las cinco dimensiones, y su descripción, se pueden observar en la Figura 4-1.

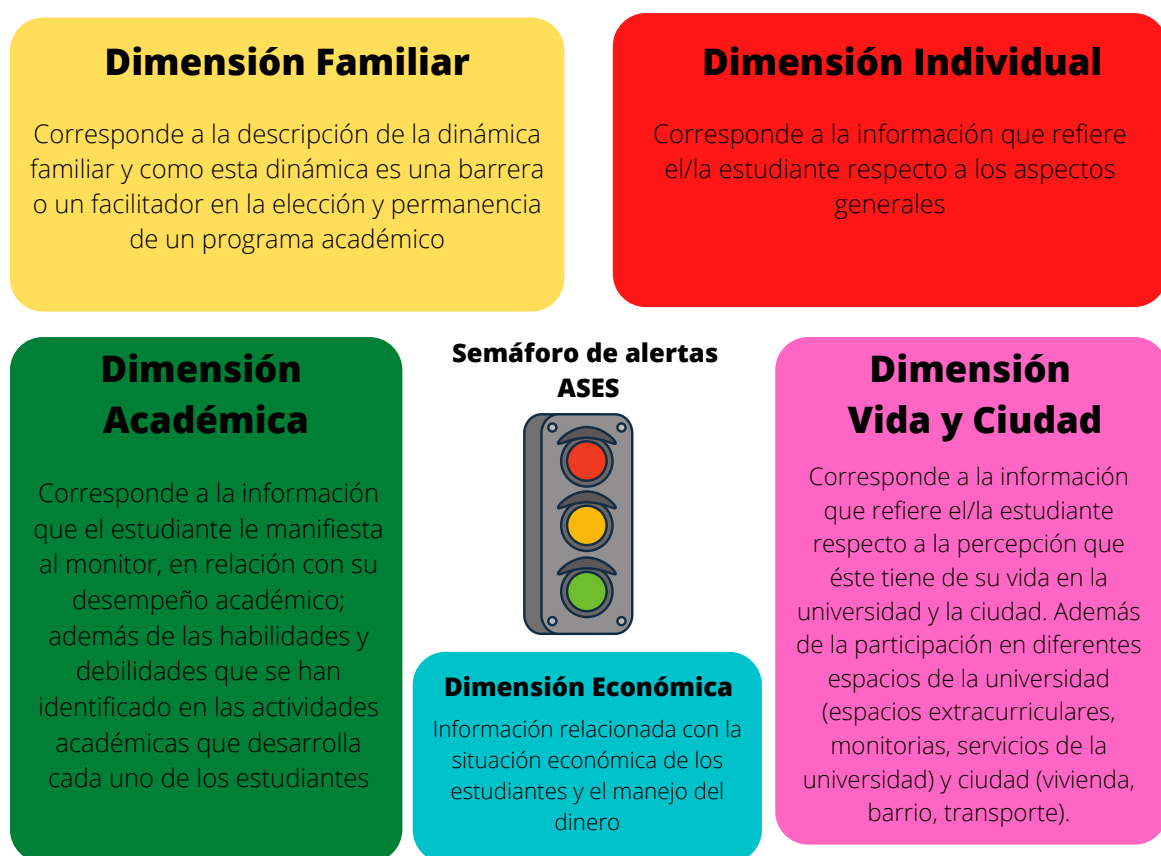


Figura 4-1.: Dimensiones semáforo de alertas. Tomado de: Estrategia ASES.

4.1.5. Temáticas y Niveles de Riesgo

Las cinco dimensiones están conformadas por diferentes temáticas, en las que se miden los niveles de riesgo que indican la probabilidad de que un estudiante pueda desertar, y se clasifican en bajo, medio y alto; cuando el nivel de riesgo es alto desde el programa ASES se realiza una intervención con el propósito de disminuir este riesgo en la temática evaluada (para una descripción de las temáticas y niveles de riesgo más detallada, ver Anexo A).

4.1.6. Ser Pilo Paga

Ser Pilo Paga es un programa que fue creado en Colombia en el año 2016 durante el gobierno del expresidente Juan Manuel Santos con la finalidad de que los estudiantes de bajos recursos (registrados en la base de datos de SISBEN hasta el 22 de septiembre de 2016, en el caso de ser indígena estar registrado en la base de datos del Ministerio del Interior hasta el 30 de septiembre de 2016, priorizando a los estudiantes con un bajo puntaje) que recién se graduaron de bachillerato y obtuvieron un puntaje en la prueba saber ICFES, igual o superior al punto de corte para cada versión (en este caso 342), pudieran acceder a la formación académica en una Institución de Educación Superior acreditada, otorgándoles un crédito condonable en el que incluye el pago de la matrícula de cada semestre de la carrera que escogieran, y un subsidio económico que sirviera para el sostenimiento del estudiante en la Universidad, el cual se otorgaba de la siguiente manera, un salario mínimo mensual legal vigente (SMMLV) semestral para el estudiante que viviera con su familia en la misma ciudad donde estudiaba, 1.5 SMMLV semestrales para el estudiante que viviera sin su familia pero en la misma ciudad donde estudiaba y 4 SMMLV semestrales para el estudiante que tuviera que cambiar de ciudad de residencia durante la duración del semestre, ya que la Universidad se ubicaba en otra ciudad diferente a la que residía (ICETEX, s.f.).

4.1.7. Condición de Excepción

La Universidad del Valle con el fin de lograr mayor participación de las minorías y en cumplimiento de la ley tiene reglamentadas las siguientes CONDICIONES DE EXCEPCIÓN:

- **Indígenas:** La resolución creada el 21 de mayo de 1998 que destina el 4 % de cupos adicionales a los establecidos por el Consejo Académico en cada carrera ofertada por la Universidad del Valle en la que la admisión no dependa un 100 % de una prueba de aptitud, a los estudiantes pertenecientes a una comunidad o resguardo indígena, que se encuentre registrada ante la Dirección de Asuntos Indígenas, Rom y Minorías del Ministerio del Interior y que los estudiantes se encuentren registrados en el censo que el Cabildo envía anualmente al Ministerio del Interior - Dirección de Etnias (Área de Admisiones, s.f.).

- **Los más altos puntajes en el Examen de Estado (Nivel nacional y departamental M.A.P):** quienes obtengan los 50 puntajes más altos en el ICFES a nivel nacional y los 2 puntajes más altos a nivel departamental tienen destinados cupos adicionales a los establecidos por el Consejo Académico de la Universidad del Valle, en la carrera que elijan, siempre y cuando la admisión no dependa un 100 % de una prueba de aptitud (Área de Admisiones, s.f.).
- **Afrodescendientes:** Los estudiantes pertenecientes a una comunidad negra descrita en el artículo 2 y en el artículo 45 de la Ley 70 del 27 de agosto de 1993, que se enmarcan en los siguientes grupos y que se encuentren debidamente registradas en el censo que elabora el Ministerio del Interior: Cuenca del Pacífico, ríos de la Cuenca de Pacífico (Área de Admisiones, s.f.):

a) De la Vertiente del Pacífico.

b) De las Cuencas de los Ríos Atrato, Acandí y Tolo de la vertiente del Caribe, zonas rurales ribereñas, raizales de San Andrés, Providencia y Santa Catalina, demás regiones con asentamientos afrodescendientes del resto del país

Para los cuales se crea otra resolución en el año 2003 que destina el 4 % de cupos adicionales a los establecidos por el Consejo Académico en cada carrera ofertada por la Universidad del Valle, siempre y cuando la admisión no dependa un 100 % de una prueba de aptitud (Universidad del Valle, 2003).

- **Reserva de cupos:** (estos cupos están dentro del total de cupos establecidos por el Consejo Académico de la Universidad del Valle) para los estudiantes admitidos en los dos períodos académicos anteriores que reservaron cupo y cumplen con uno de estos causales: la Universidad del Valle reservará hasta por un (1) año el cupo de los estudiantes que hayan sido admitidos y sean llamados a prestar el servicio militar obligatorio o que participen en un programa de intercambio estudiantil en otros países o que presenten quebrantos de salud o por razones de calamidad doméstica, caso fortuito y fuerza mayor que les impidan iniciar sus estudios. Cualquiera de estas situaciones deberá ser certificada por las instancias pertinentes o autoridades competentes (Área de Admisiones, s.f.).
- **Reinserción** destina el 2 % de los cupos establecidos por el Consejo Académico en cada carrera ofertada por la Universidad del Valle, siempre y cuando la admisión no dependa un 100 % de una prueba de aptitud a las personas que se hayan acogido al programa de reinserción o plan de paz del gobierno (Área de Admisiones, s.f.).
- **Los más altos puntajes en el examen de Estado de los colegios oficiales en los municipios del departamento del Valle del Cauca (M.P.M.):** se destina el 7 % de los cupos establecidos por el Consejo Académico en cada carrera ofertada por la

Universidad del Valle siempre y cuando la admisión a la carrera no dependa un 100 % de una prueba de aptitud, a los estudiantes que obtengan los dos más altos puntajes en el ICFES en cada colegio oficial del departamento del Valle del Cauca, si cumplen con los requisitos establecidos en la resolución No. 058 de julio 09 de 2010 expedida por el Consejo Superior de la Universidad del Valle (Área de Admisiones, s.f.).

- **Aspirantes que estén realizando actualmente su último año de bachillerato, provenientes de departamentos donde no existen sedes ni seccionales de Instituciones de Educación Superior(D.N.I.):** La condición D.N.I. otorga un cupo adicional a los cupos establecidos por el Consejo Académico de la Universidad del Valle para cada carrera ofertada, siempre y cuando la admisión a la carrera no dependa un 100 % de una prueba de aptitud, a los bachilleres que deseen estudiar una carrera ofertada durante ese año en la Universidad del Valle y provengan de estos departamentos (Área de Admisiones, s.f.).
- **Aspirantes que estén realizando actualmente su último año de bachillerato, que provengan de municipios de difícil acceso o con problemas de orden público (M.D.P.):** La condición M.D.P. otorga un cupo adicional a los cupos establecidos por el Consejo Académico de la Universidad del Valle para cada carrera ofertada, siempre y cuando la admisión a la carrera no dependa un 100 % de una prueba de aptitud, a los bachilleres que deseen estudiar una carrera ofertada durante ese año en la Universidad del Valle y provengan de estos municipios (Área de Admisiones, s.f.).
- **Población desplazada:** otorga el 2 % de cupos adicionales a los establecidos por el Consejo Académico de la Universidad del Valle en cada carrera ofertada, siempre y cuando la admisión a la carrera no dependa el 100 % de una prueba de aptitud, a las personas que se encuentren debidamente registradas en el Registro Único de Atención a la Población Desplazada de la Unidad para la Atención y Reparación Integral a las Víctimas (Área de Admisiones, s.f.).

4.2. Marco Teórico Estadístico

En esta sección se definen los conceptos estadísticos fundamentales que se deben tener en cuenta para entender la metodología estadística a utilizar. Para este proyecto se va a modelar la probabilidad de deserción de los estudiantes del Programa ASES a través de un Modelo Logístico con Intercepto Aleatorio.

4.2.1. Modelos Lineales Generalizados

En muchas situaciones los modelos de regresión lineal clásicos no se pueden aplicar directamente. Por ejemplo, en casos donde se pretende modelar respuestas no normales como los conteos o proporciones. Esta situación motivó el desarrollo del modelo lineal generalizado (GLM), que es una unificación de los modelos de regresión lineal y no lineal. Con los modelos de regresión generalizados se pueden modelar respuestas categóricas, binarias de proporciones y conteo, entre otras (Montgomery et al., 2021).

4.2.2. Familia Exponencial

Montgomery et al. (2021) define una variable aleatoria Y sigue una distribución que pertenece a la familia exponencial si la densidad es de la forma:

$$f(y) \equiv f(y|\theta, \phi) = \exp\{\phi^{-1}[y\theta - \psi(\theta)] + c(y, \phi)\} \quad (4-1)$$

Para un conjunto específico de parámetros desconocidos θ y ϕ , y para funciones conocidas $\psi(\cdot)$ y $c(\cdot, \cdot)$. A menudo, θ y ϕ se denominan “parámetro natural” (o “canónico”) y “parámetro de escala”, respectivamente.

4.2.3. Modelo Logístico

El modelo logístico es útil para modelar variables de tipo binario a partir de un conjunto de variables explicativas. Una variable binaria adopta solo dos posibles valores, 0 si se presenta un fracaso en el evento de interés y 1 si se presenta un éxito (Montgomery et al., 2021).

Sea Y una distribución de Bernoulli con probabilidad de éxito $P(Y = 1) = \pi$. La función de densidad se puede escribir como:

$$f(y) = \exp \left\{ y \ln \left(\frac{\pi}{1 - \pi} \right) + \ln(1 - \pi) \right\} \quad (4-2)$$

Lo que implica que la distribución de Bernoulli pertenece a la familia exponencial, con parámetro natural $\theta = \ln\left(\frac{\pi}{1-\pi}\right)$ y el parametro de escala $\phi = 1$. La función de enlace natural es el enlace logit, que conduce al modelo de regresión logístico clásico. Este se puede definir como $Y_i \sim \text{Bernoulli}(\pi_i)$, donde

$$\pi_i = \frac{\exp(x_i' \beta)}{[1 + \exp(x_i' \beta)]} \quad (4-3)$$

Donde x_i es un vector de covariables asociadas al individuo i , y β son los coeficientes de regresión. Adicionalmente cada observación es independiente para cada individuo i .

4.2.4. Datos por Cluster

Los datos agrupados son aquellos que están clasificados en función a un criterio, mostrando una frecuencia para cada clase o grupo formado. Es decir, los datos agrupados están separados por categorías, y cada dato u observación solo puede pertenecer a una única categoría (Westreicher, 2021).

También se puede expresar como Y_{ij} es el j -ésimo resultado medido para el grupo i , $i = 1, \dots, N$, $j = 1, \dots, n_i$ y Y_i es el vector n_i -dimensional de todas las medidas disponibles para el grupo i .

4.2.5. Modelo Lineal Mixto

El enfoque de efectos aleatorios para extender el modelo de regresión lineal univariado a estructuras de datos correlacionados o agrupados se basa en el supuesto de que, para cada sujeto, la respuesta se puede modelar mediante un modelo de regresión lineal pero con coeficientes de regresión específicos del cluster (Montgomery et al., 2021).

El modelo lineal mixto general supone que el vector Y_i de respuesta asociado para el i -ésimo cluster satisface

$$Y_i | b_i \sim N(X_i \beta + Z_i b_i, \Sigma_i) \quad (4-4)$$

donde $b_i \sim N(0, D)$

Para $n_i * p$ y $n_i * q$ matrices de diseño conocidas X_i y Z_i , para un vector β p -dimensional de coeficientes de regresión desconocidos, para un vector b_i q -dimensional de coeficientes de regresión específicos del cluster que supone que siguen una distribución normal multivariada q -dimensional con media cero y covarianza D , y Σ_i una matriz de covarianza desconocida.

Los componentes en β se denominan ‘efectos fijos’, los componentes en b_i se denominan ‘efectos aleatorios’. El hecho de que el modelo contenga así como los efectos aleatorios motivan el término ‘modelos mixtos’.

Marginalmente el modelo lineal mixto sigue la distribución

$$Y_i|b_i \sim N(X_i\beta, Z_i'DZ_i + \Sigma_i) \quad (4-5)$$

4.2.6. Modelo Lineal Mixto Generalizado

El modelo lineal mixto generalizado (GLMM) es el modelo de efectos aleatorios más utilizado en el contexto de mediciones discretas repetidas (Montgomery et al., 2021).

Se supone que, condicionalmente en q-dimensiones los efectos aleatorios b_i se distribuyen independientemente como $N(0, D)$, los resultados Y_{ij} son independientes con densidades de la forma

$$f_i(y_{ij}|b_i, \beta, \phi) = \exp\{\phi^{-1}[y_{ij}\theta_{ij} - \psi(\theta_{ij})] + c(y_{ij}, \phi)\} \quad (4-6)$$

Con $\eta(\mu_{ij}) = \eta[E(Y_{ij}|b_i)] = x'_{ij}\beta + z'_{ij}b_i$ conocida como la función de enlace $\eta(\cdot)$, con x_{ij} y z_{ij} vectores p-dimensional y q-dimensional de covariables conocidas, con β un vector p-dimensional con coeficientes de regresión fijos desconocidos y con ϕ un parametro de escala. Finalmente, sea $f(b_i|D)$ la densidad con distribución $N(0, D)$ para los efectos aleatorios b_i .

4.2.7. Modelo Logístico con Intercepto Aleatorio

$Y_{ij}|b_i \sim Bernoulli(\pi_{ij})$, con el uso de la función logit, el modelo logístico de intercepto aleatorio está dado por (Verbeke et al., 2008):

$$\pi_{ij} = \frac{\exp(b_i + x'_{ij}\beta)}{[1 + \exp(b_i + x'_{ij}\beta)]} \quad (4-7)$$

Para la interpretación de los parametros β se realizará a traves de los odds ratio (OR = $\exp(\beta)$).

4.2.8. Estimación por Máxima Verosimilitud

Los modelos de efectos aleatorios pueden ser ajustados por maximización de la verosimilitud marginal obtenida integrando los efectos aleatorios. Para Verbeke et al. (2008), la función de densidad marginal del cluster i entonces se convierte en

$$f_i(y_i|\beta, D, \phi) = \int \prod_{j=1}^{n_i} f_{ij}(y_{ij}|b_i, \beta, \phi) f(b_i|D) db_i \quad (4-8)$$

Donde se deriva la función de verosimilitud marginal:

$$L(\beta, D, \phi) = \prod_{i=1}^N f_i(y_i|\beta, D, \phi) = \prod_{i=1}^N \int f_{ij}(y_{ij}|b_i, \beta, \phi) f(b_i|D) db_i \quad (4-9)$$

El problema clave para maximizar la ecuación 4-9 es la presencia de N integrales sobre los efectos aleatorios q-dimensionales b_i . En algunos casos (como en el caso de la ecuación 4-5) estas integrales se pueden resolver analíticamente. En general, no hay expresiones analíticas disponibles para las integrales en 4-9 y se necesitan aproximaciones numéricas. Hay varios métodos para hacerlo como la cuadratura gaussiana, la cuasi-verosimilitud penalizada (PQL) y la cuasi-verosimilitud marginal (MQL) (Verbeke et al., 2008).

4.2.9. Método de Cuadratura Gaussiana

La cuadratura puede ser útil en la aproximación de la verosimilitud marginal por integración numérica. Específicamente, si la distribución asumida para los efectos aleatorios es normal, la cuadratura Gauss-Hermite puede aproximar la integral a un grado alto de exactitud. La integral es aproximada por una sumatoria sobre un número especificado de puntos de cuadratura Q para cada dimensión de la integración. Si se asume que los efectos aleatorios se rigen por una distribución distinta de la normal, se escogen otros puntos de cuadratura pero la exactitud es más débil (García et al., 2012). Según Verbeke et al. (2008) Para la aproximación de integrales de la forma

$$\int f(z)\phi(z)dz \quad (4-10)$$

Para una función conocida $f(z)$ y para $\phi(z)$ la densidad de la distribución normal estándar (multivariante). Esta integral (4-10) se aproxima por la suma ponderada

$$\int f(z)\phi(z)dz \approx \sum_{q=1}^Q w_q f(z_q) \quad (4-11)$$

Donde Q es el orden de la aproximación. Cuanto mayor sea Q , más precisa será la aproximación. Además, los llamados nodos (o puntos de cuadratura) z_q son soluciones del polinomio de Hermite de orden Q , mientras que los w_q son pesos elegidos apropiadamente. (Verbeke et al., 2008).

La cuadratura Gaussiana puede funcionar bien si las respuestas son dicotómicas y el tamaño de los cluster o grupos de datos, es pequeño. Sin embargo, frecuentemente se necesita un gran número de puntos de cuadratura para obtener una buena aproximación de la verosimilitud. Como consecuencia, el método puede ser computacionalmente intensivo, particularmente si hay muchos efectos aleatorios (García et al., 2012).

4.2.10. Estimación Bayesiana Empírica

Los efectos aleatorios describen la variabilidad entre clusters, lo que los hace útiles para detectar perfiles especiales (es decir, grupos periféricos). La estimación de los efectos aleatorios se basará en su distribución posterior con densidad dada por

$$f_i(b_i|y_i\beta, D, \phi) = \frac{f_i(y_i|b_i\beta, \phi)f(b_i|D)}{\int f_i(y_i|b_i\beta, \phi)f(b_i|D)db_i} \quad (4-12)$$

A diferencia del caso lineal, esta densidad posterior no es, en general, normal. Por lo tanto, se usa la moda posterior, en lugar de la media posterior, como estimador puntual para b_i . Más específicamente, el estimador \hat{b}_i es el valor para b_i que maximiza la ecuación 4-12, reemplazando los parámetros desconocidos por sus estimaciones por máxima verosimilitud. Estas estimaciones se denominan estimaciones empíricas de Bayes (EB) (Verbeke et al., 2008).

4.2.11. Selección de Variables

Los métodos de selección de variables se encargan de abordar el problema de construcción o selección del modelo. En general, si se incluyen cada vez más variables en un modelo de regresión, el ajuste a los datos mejora, aumenta la cantidad de parámetros a estimar pero disminuye su precisión individual (mayor varianza) y por tanto la de la función de regresión estimada, se produce un sobreajuste. Por el contrario, si se incluyen menos variables de las necesarias en el modelo, las varianzas se reducen pero los sesgos aumentarán obteniéndose una mala descripción de los datos. Por otra parte, algunas variables predictoras pueden perjudicar la confiabilidad del modelo, especialmente si están correlacionadas con otras. De esta manera, el objetivo de los métodos de selección de variables es buscar un modelo que se ajuste bien a los datos y que a la vez sea posible buscar un equilibrio entre bondad de ajuste y sencillez (González et al., 2015).

La idea de los algoritmos de selección de variables es elegir el mejor modelo en forma secuencial pero incluyendo o excluyendo una sola variable predictora en cada paso de acuerdo a ciertos criterios. El proceso secuencial termina cuando se satisface una regla de parada establecida. Los algoritmos más usados son el forward (selección hacia adelante), el backward (selección hacia atrás) y el stepwise (combinación de forward y backward) (González et al., 2015).

Método Backward

El método backward utiliza como punto de partida un modelo muy complejo, el cual incorpora todos los efectos que pueden influir en la respuesta y en cada etapa se elimina la variable menos influyente hasta que no proceda eliminar otra variable (González et al., 2015).

4.2.12. Criterios de Información

Criterio de Información de Akaike

El Criterio de Información de Akaike (AIC) es un criterio paramétrico comparativo entre modelos y representa por esto una herramienta para selección de modelos. Dado un conjunto de datos, es posible encontrar varios modelos que se ajustan a ellos. La idea es ranquearlos de acuerdo al AIC (Correa & Salazar, 2016).

Cuando se aplica este tipo de procedimientos para seleccionar las variables se debe tener en cuenta cual será la condición para suprimir o incluir un término. Según González et al. (2015) el cálculo del AIC se realiza mediante la siguiente ecuación:

$$AIC = 2k - 2\log(L(\theta)) \quad (4-13)$$

En donde k es el número de parámetros en el modelo y $L(\theta)$ es el valor de la verosimilitud obtenida por el modelo estimado. Esta medida tiene en cuenta el ajuste y a la vez compensa el exceso de parámetros. El objetivo es buscar un modelo cuyo AIC sea pequeño, pues en ese caso habría una verosimilitud grande y pocos parámetros (González et al., 2015).

4.2.13. Pruebas de Hipotesis

Estadístico de Wald

Es utilizada para poner a prueba el verdadero valor del parámetro basado en la estimación de la muestra. Este contraste está basado en la normalidad asintótica de los estimadores. Se quiere contrastar si un parámetro $\beta_i = 0$, con $i = 1, \dots, r$ frente a que no lo sea (Reche, 2013).

$$H_0 : \beta_i = 0 \text{ vs } H_1 : \beta_i \neq 0$$

Wald, demostró que bajo la hipótesis nula, el estadístico de contraste es:

$$Z_i = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)} \quad (4-14)$$

donde $\hat{\beta}_i$ es la estimación del parámetro β_i y $SE(\hat{\beta}_i)$ es el error estándar asociado para $i = 1, \dots, r$

Test de Razón de Verosimilitud

En las pruebas de hipótesis de modelos mixtos se busca comparar si el modelo completo y el modelo reducido no presentan diferencias estadísticamente significativas, esta prueba se puede realizar con el estadístico de wald o con el test de verosimilitud. Supongamos que queremos estudiar $H_0 : \theta \in \Theta_0 \text{ vs } H_1 : \theta \in \Theta$, por lo que el estadístico en el test de verosimilitud (LR) se calcularía de la siguiente forma según Hernández & López (2022):

$$LR = -2 \log \left(\frac{\sup_{\theta \in \Theta_0} L(\theta)}{\sup_{\theta \in \Theta} L(\theta)} \right) \quad (4-15)$$

Usualmente la prueba de razón de verosimilitud se expresa en función de los valores de log-verosimilitud del modelo así:

$$LR = -2(\log L(\hat{\theta}_0) - \log L(\hat{\theta})) \quad (4-16)$$

En donde el estadístico de prueba $LR \sim \chi^2_{k-k_0}$, k es el número de parámetros del modelo completo y k_0 es el número de parámetros asumiendo que H_0 es verdadera, $\hat{\theta}_0$ es el vector de parámetros estimados asumiendo H_0 como verdadera y $\hat{\theta}$ es el vector de estimaciones de los parámetros del modelo completo (Hernández & López, 2022).

4.2.14. Matriz de Confusión

Barrios (2019) dice que la matriz de confusión permite ver la cantidad de aciertos y errores que tiene un modelo a la hora de predecir datos binarios comparándola con el comportamiento real de los datos. La prueba clasifica un valor como positivo o negativo utilizando la predicción de la probabilidad a partir del modelo y un punto de corte definido. Si la probabilidad estimada es mayor que el punto de corte entonces la predicción es positiva, en caso contrario la predicción es negativa.

VALORES PREDICCIÓN	Verdaderos positivos	Falsos Positivos
	Falsos Negativos	Verdaderos Negativos
	VALORES REALES	

Figura 4-2.: Matriz de confusión. Tomado de: Barrios.

En la matriz de confusión (Ver figura 4-2) se muestra la siguiente información:

- Verdaderos positivos (VP): El valor real es positivo y la prueba predijo también que era positivo
- Verdaderos negativos (VN): El valor real es negativo y la prueba predijo también que era negativo
- Falsos positivos (FP): El valor real es negativo, y la prueba predijo que el resultado es positivo.
- Falsos negativos (FN): El valor real es positivo, y la prueba predijo que el resultado es negativo.

A partir de la matriz de confusión se pueden obtener métricas que permitan evaluar el modelo, entre ellas la sensibilidad y especificidad.

La sensibilidad o tasa de verdaderos positivos (TP) es la proporción de casos positivos que fueron correctamente identificados por el modelo, se calcula así:

$$TP = \frac{VP}{VP + FN} \quad (4-17)$$

La especificidad o tasa de verdaderos negativos (TN) es la proporción de casos negativos que fueron correctamente identificados por el modelo, se calcula así:

$$TN = \frac{VN}{VN + FP} \quad (4-18)$$

Estos indicadores dependen del punto de corte que se defina. Por esta razón, se construye las curvas ROC. Una curva ROC es una representación gráfica de la sensibilidad y la especificidad para diferentes puntos de corte.

5. Metodología

En este capítulo se presenta la metodología implementada para dar cumplimiento a los objetivos planteados. En primera instancia se caracterizará a la población de interés, luego se consolidará la base de datos identificando los tipos de variables, a continuación se aplicará un modelo logístico con efectos aleatorios, que permita identificar los factores asociados a la deserción estudiantil.

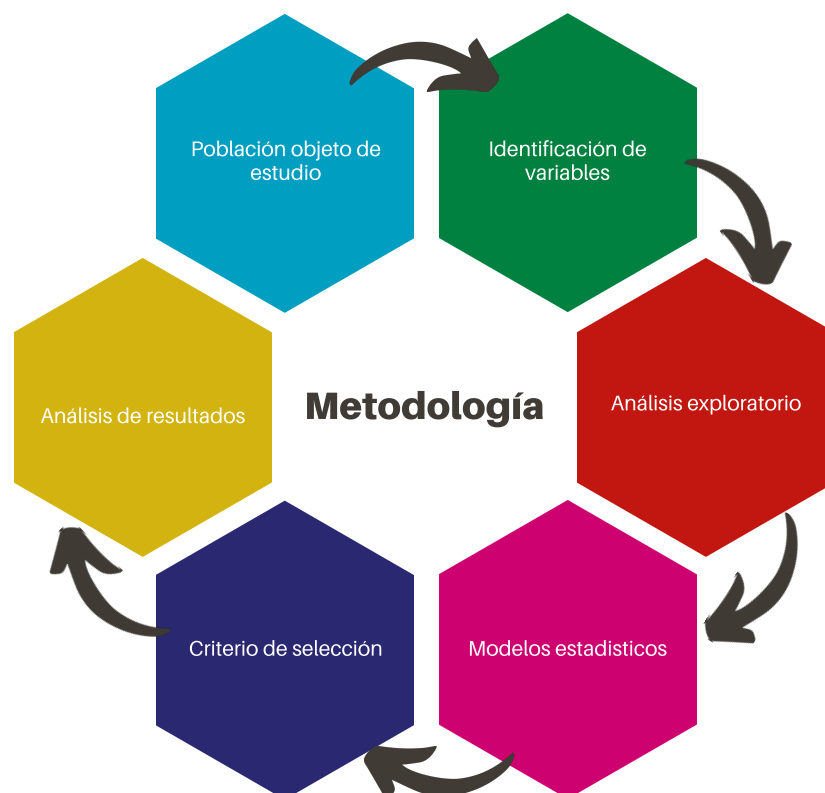


Figura 5-1.: Fases de la metodología

5.1. Población Objeto de Estudio

Estudiantes admitidos por alguna Condición de Excepción o el programa del Gobierno Nacional Ser Pilo Paga en el año 2017 a la Universidad del Valle sede Cali, pertenecientes al programa ASES los cuales recibieron acompañamiento en su ciclo básico universitario.

5.2. Consolidación de la Información

Las bases de datos utilizadas fueron otorgadas por parte del programa ASES, la cual cuenta con una muestra de 472 estudiantes, donde 73 estudiantes pertenecen a Ser Pilo Paga III y 399 estudiantes pertenecen a alguna Condición de Excepción. Además de 12 variables asociadas al acompañamiento socio-académico que tuvieron los estudiantes durante su ciclo básico.

5.2.1. Primera Base: Información Académica

La primera base de datos tiene registrada toda la información académica de los estudiantes como: código del estudiante, programa al que se encuentra matriculado, Facultad a la que pertenece, tipo de admisión con la que ingreso a la Universidad del Valle, Condición de Excepción (CE), si estudiante deserto durante su ciclo básico de formación (primeros 5 semestres), cantidad de bajos académicos, cantidad de semestres cursados y su sexo. En la tabla 5-1 se puede observar la estructuración de la base de datos tomando algunos estudiantes de ejemplo, para mantener la confidencialidad de los estudiantes se suprimen los últimos dos dígitos de cada código.

Tabla 5-1.: Información académica

Código	Programa	Facultad	Admisión	CE	Desertó	Bajos	Semestres	Sexo
17222xy	3744	Ingeniería	CE2017-1	Indígena	No	0	5	M
17222xy	3461	Psicología	SPP3	-	No	0	5	F
17223xy	3740	Ingeniería	CE2017-1	M.D.P.	No	0	5	F
17223xy	3340	C. Sociales	CE2017-1	C.A.	No	0	5	F
17714xy	3841	C. Admin.	SPP3	-	No	0	5	F
17714xy	3754	Ingeniería	SPP3	-	Si	1	1	M
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

5.2.2. Segunda Base: Puntuaciones Encuentros ASES

La segunda base de datos tiene registrada la información de los encuentros socio-educativos entre los estudiantes y sus correspondientes monitores. La base de datos recopila en

cada encuentro la siguiente información: código del estudiante, fecha del encuentro, una puntuación correspondiente al nivel de riesgo que se percibe del estudiante en cada dimensión. En la tabla **5-2** se puede observar la estructuración de la base de datos tomando algunos estudiantes de ejemplo, para mantener la confidencialidad de los estudiantes se suprimen los últimos dos dígitos de cada código.

Tabla 5-2.: Puntuaciones encuentros ASES

Código	fecha	Individual	Familiar	Académico	Económico	Vida Universitaria
17222xy	2017-02-17	bajo	alto	-	-	bajo
17222xy	2017-03-01	alto	-	-	-	-
17222xy	2017-03-03	bajo	bajo	bajo	-	bajo
17222xy	2017-03-14	bajo	-	bajo	-	-
17222xy	2017-03-19	-	-	-	-	-
17222xy	2017-03-30	-	-	bajo	-	bajo
:	:	:	:	:	:	:

5.2.3. Base de Datos Final

Para la consolidación, depuración y limpieza de la base de datos final utilizada en este proyecto se hizo uso del software KNIME, donde se realizaron los siguientes pasos:

- De la segunda base de datos, se contabiliza la cantidad de encuentros efectivos realizados por cada estudiante mediante el nodo *GroupBy*.
- De la segunda base de datos, se contabiliza la cantidad de inasistencias a los encuentros pactados por cada estudiante mediante el nodo *GroupBy*.
- De la segunda base de datos, se contabiliza la cantidad de veces que un estudiante fue calificado con un nivel de riesgo medio o alto en cada una de las dimensiones con las que trabaja el programa ASES mediante el nodo *GroupBy*.
- Se realiza el emparejamiento de la primera y segunda base de datos, utilizando el nodo *Joiner* el cual permite combinar dos tablas a partir de una serie de condiciones, para este caso se utilizó el código de los estudiantes.

La estructura final de la base de datos utilizada para este proyecto se muestra en la tabla **5-3**, tomando algunos estudiantes de ejemplo, para mantener la confidencialidad de los estudiantes se suprimen los últimos dos dígitos de cada código:

Tabla 5-3.: Base de datos final

Código (S)	Facultad	Desertó	Admisión	Bajos	Encuentros	Indi.	Fami.	Acad.	Econ.	VidaUni	Inasistencia
17222xy (M)	Ingeniería	No	CE2017	0	18	5	1	4	1	0	2
17222xy (F)	Psicología	No	SPP3	0	44	2	0	0	1	0	3
17223xy (F)	Ingeniería	No	CE2017	0	77	6	0	16	0	12	24
17223xy (F)	C. Sociales	No	CE2017	0	25	1	1	4	1	0	3
17714xy (F)	C. Admin.	No	SPP3	0	23	2	0	5	2	1	2
17714xy (M)	Ingeniería	Si	SPP3	1	55	0	0	35	0	10	10
17800xy (F)	Ingeniería	No	CE2017	0	10	0	2	3	0	0	0
17800xy (F)	Artes	No	CE2017	0	7	1	0	1	0	0	1
17272xy (M)	Ingeniería	SI	CE2017	2	35	5	5	17	10	5	0
17278xy (F)	IEP	SI	CE2017	2	12	5	3	7	0	0	4
17300xy (M)	C. Admin.	NO	CE2017	1	18	3	2	7	10	5	5
17312xy (F)	Salud	NO	CE2017	0	20	12	7	12	2	3	5
17318xy (M)	Ingeniería	NO	CE2017	0	26	2	0	4	0	0	2
17406xy (F)	C. Admin.	NO	CE2017	0	30	5	16	1	5	0	2
17437xy (F)	C. Admin.	NO	CE2017	1	10	0	0	3	0	0	5
17713xy (F)	Ingeniería	NO	SPP3	0	44	5	2	10	0	2	8
:	:	:	:	:	:	:	:	:	:	:	:

Debido al gran tamaño de la tabla **5-3**, la variable sexo (S) se incluyo junto a la variable código del estudiante para poder incorporar todas las variables.

La descripción detallada de cada una de las variables se encuentra en la tabla 5-4:

Tabla 5-4.: Descripción de las variables

Variable	Descripción
Desertó	El estudiante deserto durante su ciclo básico de formación (Si o No)
Admitido	Tipo de admisión (SPP3 o CE2017)
Facultad	Facultad a la que pertenece el estudiante
Bajos	Cantidad de bajos académicos registrados
Encuentros	Cantidad de encuentros registrados por semestre
Inasistencias	Cantidad de veces que el estudiante faltó a un encuentro
Sexo	Sexo del estudiante (Masculino o Femenino)
DAcademica	Cantidad de veces que tuvo un riesgo medio o alto en la dimensión académica
DEconomica	Cantidad de veces que tuvo un riesgo medio o alto en la dimensión económica
DFamiliar	Cantidad de veces que tuvo un riesgo medio o alto en la dimensión familiar
DIndividual	Cantidad de veces que tuvo un riesgo medio o alto en la dimensión individual
DVidaUni	Cantidad de veces que tuvo un riesgo medio o alto en la dimensión universitaria

5.3. Unidad Experimental

Interés: Identificar cuales son los factores con los que trabaja el programa ASES e influyen en la deserción de los estudiantes admitidos por Condición de Excepción o Ser Pilo Paga III durante su ciclo básico de formación académica en la Universidad del Valle.

Experimento: Observar si los estudiantes admitidos por Condición de Excepción o Ser Pilo Paga III en el año 2017 desertaron durante todo su proceso de acompañamiento en el programa ASES

Evento A: El estudiante j de la Facultad i desertó durante el acompañamiento recibido por el programa ASES

Variable Aleatoria: Sea Y_{ij} la variable aleatoria asociada al evento A.

$$Y_{ij} = \begin{cases} 1 & \text{si ocurre el evento A} \\ 0 & \text{si lo contrario, para } i = 1, \dots, j = 1, \dots, n_i. \end{cases} \quad (5-1)$$

Dado que la variable es binaria, se le asocia una distribución Bernoulli y se utilizará un modelo logístico.

5.4. Datos de Entrenamiento y Datos de Prueba

Para poder calibrar si un modelo funciona, se necesita probarlo con un conjunto de datos diferente, por lo tanto, los datos del trabajo se dividen en dos partes: datos de entrenamiento y datos de prueba. Se utilizó la función *sample* de librería *dplyr* en R-Studio para dividir la base de datos y probar el poder predictivo del modelo, se utilizó el 80 % de los datos para entrenar el modelo y el 20 % restante para evaluar la especificidad y sensibilidad del modelo con el fin de determinar si el modelo predice adecuadamente si un estudiante tiene mayor riesgo de desertar o si por el contrario no tiene este riesgo.

5.5. Modelo Logístico

El modelo logístico suele utilizarse cuando la variable de respuesta tiene una distribución Bernoulli, en este caso la variable de respuesta que está asociada con si un estudiante desertó o no tiene el valor de 0 cuando un estudiante no desertó en un período académico y tiene el valor de 1 si el estudiante desertó. Sin embargo, se espera que las observaciones de estudiantes que pertenecen a la misma Facultad estén correlacionados, se propone trabajar con el modelo logístico con efectos aleatorios.

5.6. Modelo Logístico con Intercepto Aleatorio

En este caso se propone modelar la probabilidad de que un estudiante deserte en función de la cantidad de veces que un estudiante tuvo una calificación de riesgo medio o alto en los niveles de riesgo en cada dimensión de las que trabaja el programa ASES, el numero de bajos académicos, el número de acompañamientos que tuvo el estudiante durante un período académico, el número de encuentros a los que faltó el estudiante, el tipo de admisión que realizo a la Universidad del Valle y su sexo; además la Facultad se tomará como efecto aleatorio, por ende la variable aleatoria toma la siguiente forma

$$Y_{ij}|b_i \sim \text{Bernoulli}(\pi_{ij})$$

$$\text{donde } b_i \sim N(0, d)$$

Finalmente, se ajusta el modelo de la siguiente manera:

$$\begin{aligned} \text{logit}(\pi_{ij}) = & \beta_0 + b_i + \beta_1 \text{Admitido}_{ij} + \beta_2 \text{Bajos}_{ij} + \beta_3 \text{Encuentros}_{ij} \\ & + \beta_4 \text{DAcademica}_{ij} + \beta_5 \text{DEconomica}_{ij} + \beta_6 \text{DFamiliar}_{ij} + \beta_7 \text{DIndividual}_{ij} \\ & + \beta_8 \text{DVidaUni}_{ij} + \beta_9 \text{Inasistencias}_{ij} + \beta_{10} \text{Sexo}_{ij} \end{aligned} \quad (5-2)$$

Las estimaciones del modelo 5-2 se hace por máxima verosimilitud mediante el método de cuadratura gaussiana utilizando la función *mixed_model* de la librería *GLMMadaptive* en R-Studio.

5.7. Método de Cuadratura Gaussiana

En la tabla 5-5 se muestran las estimaciones para cada coeficiente β con diferentes puntos de cuadratura. Se elige trabajar con 30 puntos de cuadratura, ya que a partir de ese punto se estabiliza la estimación de los coeficientes generando una mayor exactitud de los valores.

Tabla 5-5.: Método de Cuadratura Gaussiana

	Q = 1	Q = 5	Q = 10	Q = 30	Q = 50
β_0	-2.7598	-2.5608	-2.5593	-2.5581	-2.5581
β_1	0.0203	-0.0520	-0.0529	-0.0534	-0.0533
β_2	1.4908	1.4884	1.4884	1.4883	1.4883
β_3	-0.2741	-0.2674	-0.2672	-0.2671	-0.2671
β_4	0.6689	0.6528	0.6526	0.6525	0.6525
β_5	0.2621	0.2640	0.2636	0.2633	0.2633
β_6	0.7251	0.6617	0.6610	0.6605	0.6605
β_7	-0.0261	-0.0431	-0.0433	-0.0432	-0.0432
β_8	-1.5487	-1.4580	-1.4563	-1.4554	-1.4555
β_9	0.3891	0.4022	0.4024	0.4025	0.4025
β_{10}	1.0208	0.9872	0.9868	0.9870	0.9870

5.8. Selección de Variables

Para la selección de variables del modelo se utilizará el método backward en el que se eliminará la variable que menos aporte en la explicación de la deserción, además se utilizará como métrica de evaluación la disminución en el valor del AIC, donde se eliminan covariables hasta que el AIC deje de disminuir.

5.9. Pruebas de Hipótesis

Estadístico de Wald

Con el fin de evaluar si los coeficientes estimados son significativos en el modelo. Es decir, determinar si una variable explicativa tiene un efecto significativo para la estimación de la deserción de los estudiantes. Se utiliza la ecuación 4-14 para realizar el cálculo del estadístico de Wald y se contrasta con un nivel de significancia del 5 %.

Test de Razón de Verosimilitud

Para probar que el modelo dado por el método de selección de variables proporciona el mismo ajuste que el modelo completo. Se utilizó la función *anova* de la librería *GLMMadaptive* en R-Studio y el resultado de esta prueba se comparará con un valor de significancia del 5 % para decidir si se rechaza la hipótesis nula o no.

5.10. Diagnóstico del Modelo

Para la construcción de la matriz de confusión, se utilizó la función *ConfusionMatrix* de la librería *caret*, la cual permite ver la cantidad de aciertos y errores que tiene el modelo completo y el modelo reducido a la hora de predecir (Para las predicciones, se hizo uso de la base de prueba y la función *predict*). La tabla 5-6 muestra la estructuración de la matriz de confusión, donde a partir de un punto de corte definido por la curva ROC se evaluará la correcta clasificación de las predicciones.

Tabla 5-6.: Matriz de confusión

		Valores reales	
		No desertor	Desertor
Predicción	No desertor	VN	FN
	Desertor	FP	VP

Mediante la matriz de confusión se obtienen las métricas de sensibilidad (ecuación 4-17) y especificidad (ecuación 4-18). Finalmente, se podrá evaluar la capacidad del modelo completo y el modelo reducido para predecir observaciones nuevas.

5.11. Efectos Aleatorios

La estimación de los efectos aleatorios asociados a la Facultad se hace por medio del estimador EB (Ecuación 4-12) utilizando la función *ranef* de la librería *GLMMadaptive* en R-studio. Estas estimaciones permiten visualizar las facultades en las que se requiere hacer un mayor esfuerzo por parte de la universidad por tener un indicio de deserción más alto.

6. Resultados

En este Capítulo se presentan los resultados de la investigación. Se empieza con un análisis descriptivo con el fin de ordenar la información y extraer características importantes de los datos a partir de gráficos y tablas. Luego, se realizan inferencias sobre la deserción estudiantil de los estudiantes pertenecientes al Programa ASES en el año 2017 mediante el uso de un Modelo Logístico con Intercepto Aleatorio.

6.1. Análisis Exploratorio

El análisis exploratorio de los datos se realiza antes de construir un modelo estadístico, generalmente con el fin de generarse una idea acerca de sus características.

Tabla 6-1.: Índice de deserción temprana por Facultad

Facultad	Desertores	Matriculados	Índice de deserción temprana
Artes Integradas	5	24	20.83 %
Ciencias de la Administración	3	51	5.58 %
Ciencias Naturales y Exactas	10	46	21.73 %
Ciencias Sociales y Económicas	1	20	5 %
Humanidades	15	66	22.72 %
Ingeniería	43	174	24.71 %
Instituto de Educación y Pedagogía	7	33	21.21 %
Psicología	0	7	0 %
Salud	7	51	13.72 %

En la tabla **6-1** se muestra el índice de deserción temprana, por Facultad de los estudiantes que fueron admitidos por Condición de Excepción (CE2017-1, admitidos en el primer semestre del año 2017 y CE2017-2, admitidos en el segundo semestre del año 2017) y Ser Pilo Paga III (SPP3) en la Universidad del Valle; aquí se observa una alta participación de los estudiantes en la Facultad de Ingeniería, los cuales presentan el mayor índice de deserción, seguido por la Facultad de Humanidades; también destaca el bajo índice de deserción de las Facultades de las Ciencias de la Administración, Ciencias Sociales y Económicas y Psicología manteniendo un índice de deserción inferior al 6 %, donde en esta última Facultad su participación es baja y no presenta desertores.

Tabla 6-2.: Índice de deserción temprana por tipo de admisión

Tipo de admisión	Desertores	Matriculados	Índice de deserción temprana
CE2017-1	58	252	23.01 %
CE2017-2	24	147	16.32 %
SPP3	9	73	12.33 %

En la tabla **6-2** se realiza el mismo análisis anterior, pero distinguiendo según el tipo de admisión que tuvo en la universidad. Se observa que los estudiantes SPP3 tienen un índice de deserción menor a los estudiantes admitidos por Condición de Excepción.

Tabla 6-3.: Índice de deserción temprana por Facultad y tipo de admisión

Facultad	CE2017-1	CE2017-2	SPP3
Artes Integradas	20.8 %	0 %	0 %
Ciencias de la Administración	1.96 %	1.96 %	1.96 %
Ciencias Naturales y Exactas	10.9 %	6.5 %	4.3 %
Ciencias Sociales y Económicas	0 %	5 %	0 %
Humanidades	19.7 %	1.5 %	1.5 %
Ingeniería	14.4 %	7.5 %	2.9 %
Instituto de Educación y Pedagogía	12.1 %	9.1 %	0 %
Psicología	0 %	-	0 %
Salud	9.8 %	3.9 %	0 %

En la tabla **6-3** se realiza un análisis bivariado de los índices de deserción temprana de los estudiantes admitidos al programa ASES en el año 2017 agrupando por Facultad y el tipo de admisión. Los estudiantes admitidos por CE2017-1 muestran unos altos índices en la mayoría de las Facultades, mientras los estudiantes admitidos por SPP3 en 5 de las 9 Facultades no presenta estudiantes desertores.

Tabla 6-4.: Índice de deserción temprana por Condición de Excepción

Condición de Excepción	Desertores	Matriculados	Índice de deserción temprana
Comunidades afrocolombianas	26	156	16.67 %
Indígena	29	116	25 %
Población desplazada	14	71	19.72 %
M.D.P.	10	37	27.03 %
M.P.M.	3	19	15.78 %

En la tabla **6-4** se muestran los índices de deserción temprana de los estudiantes según la Condición de Excepción a la que ingreso a la Universidad del Valle. El interés se radica, ya que los estudiantes admitidos por Condición de Excepción son los que mayor índice de

deserción registran. Donde se observa que los estudiantes indígenas y los estudiantes que están realizando su último año de bachillerato y provienen de municipios de difícil acceso o con problemas de orden público (M.D.P.) son los que mayor índice de deserción presentan.

Tabla 6-5.: Promedio de encuentros por semestre

Facultad	CE2017-1	CE2017-2	SPP3
Artes Integradas	14.9	11.2	36.4
Ciencias de la Administración	16.2	14.2	18.5
Ciencias Naturales y Exactas	18.6	15.6	30.8
Ciencias Sociales y Económicas	15	11.7	15.5
Humanidades	14.5	11.2	15
Ingeniería	22.2	18	29.1
Instituto de Educación y Pedagogía	11.12	8.94	14.8
Psicología	14.2	-	24.3
Salud	14.3	14.7	28.2

En la tabla **6-5** se realizó el análisis bivariado por Facultad y tipo de admisión del estudiante teniendo en cuenta el número promedio de encuentros por semestre, ya que por revisión de antecedentes se sabe que la relación entre el índice de deserción y el número de acompañamientos que tuvo el estudiante en el semestre es inversa (Alvarado, 2019); los estudiantes admitidos por SPP3 son los que registran los mayores encuentros en las diferentes Facultades y esto podría explicar el por qué tienen el menor índice de deserción; los encuentros por semestre que tuvieron los estudiantes de las Facultades de Ciencias de la Administración, Humanidades e Ingeniería no son tan variables entre los diferentes tipos de admisión.

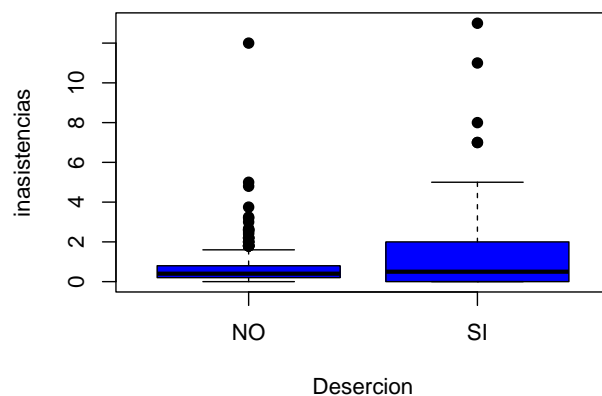


Figura 6-1.: Inasistencias

En la figura **6-1** se muestra el comportamiento de las inasistencias entre los estudiantes que desertaron y los que no, donde se observa que los estudiantes que decidieron no continuar con sus estudios académicos eran los que mayor tendencia tenían a no asistir a los encuentros pactados con sus monitores.

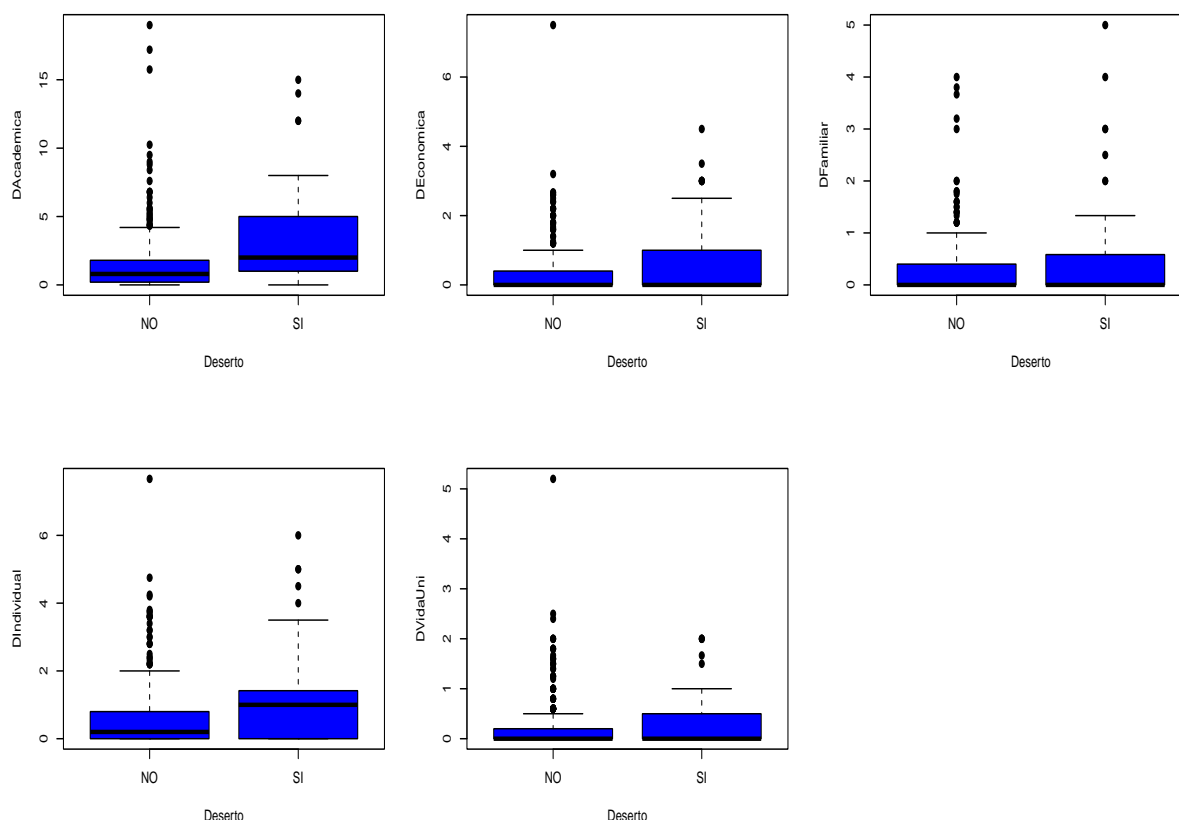


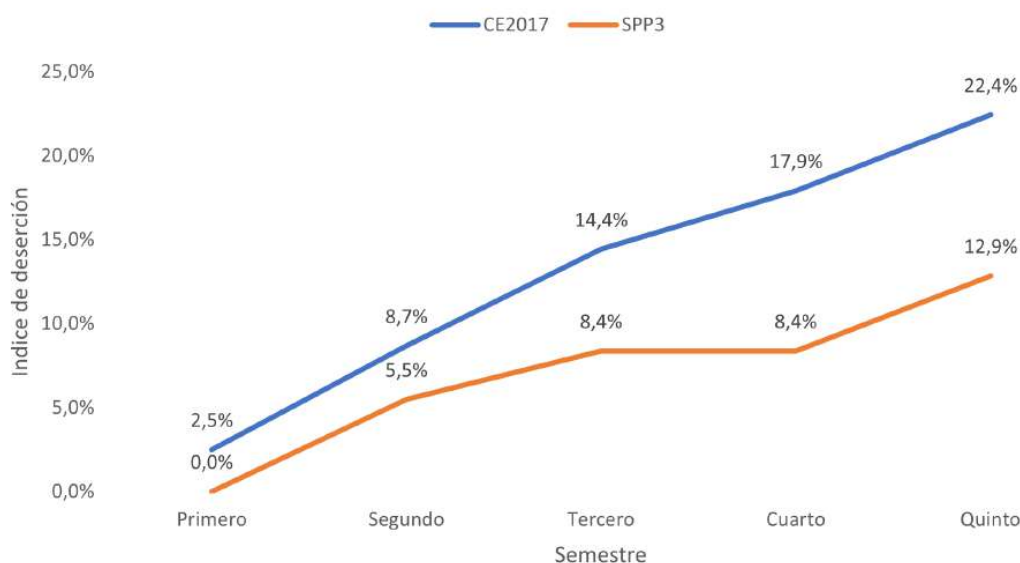
Figura 6-2.: Riesgos según cada dimensión

La figura **6-2** ilustra el número de riesgos (medio o alto) que tuvieron los estudiantes en el ciclo básico de formación en cada una de las dimensiones con las que trabaja el programa ASES diferenciando a los estudiantes que desertaron y los que no lo hicieron, recordemos que el riesgo es una calificación que da el monitor al estudiante después de cada encuentro basado en ciertos criterios dependiendo de la dimensión trabajada. Comparando cada dimensión, se observa que los estudiantes que desertaron presentan mayores puntuaciones en todas las dimensiones, pero en la dimensión académica se observa una notable diferencia entre los estudiantes que desertaron y los que no desertaron, además de ser la dimensión con las puntuaciones más altas, indicándonos que es la dimensión que posiblemente más se trabaja en cada encuentro.

Tabla 6-6.: Matricula académica por semestre

	Primero	Segundo	Tercero	Cuarto	Quinto
Deserción por semestre	10	28	23	12	18
Deserción por semestre acumulada	10	38	61	73	91
Índice por semestre	2.11 %	6.06 %	5.29 %	2.91 %	4.51 %
Índice por semestre acumulada	2.11 %	8.17 %	13.46 %	16.37 %	20.88 %
Cantidad estudiantes	472	462	434	411	399

En la tabla **6-6** se presenta la cantidad de estudiantes que desertaron, el índice de deserción y la cantidad que se matriculó en cada semestre de los estudiantes admitidos en el año 2017. En el segundo y tercero semestre del ciclo básico de formación hay una mayor tendencia a desertar en los estudiantes pertenecientes al programa ASES dado que su índice de deserción supera el 5 %. En el quinto semestre el índice de deserción acumulado es de 20.88 % y la cantidad de estudiantes que finalizaron su ciclo básico fue de 381 de 472 estudiantes.

**Figura 6-3.:** Índice de deserción por tipo de admisión en cada semestre

La figura **6-3** muestra la variación del índice de deserción acumulativo en cada semestre de los estudiantes pertenecientes al programa ASES según su tipo de admisión. Se observa que en el caso de los estudiantes admitidos por el programa Ser Pilo Paga (SPP3) no presentaron desertores en su primer y cuarto semestre, además que en comparación a las Condiciones de Excepción, logro obtener satisfactoriamente un menor índice de deserción temprana.

6.2. Modelo Logístico con Intercepto Aleatorio

Se ajusta el modelo 5-2 para describir la probabilidad de deserción a partir de las variables escogidas, teniendo una diferencia para los estudiantes que están en la misma Facultad de la universidad, ya que se espera que la deserción este correlacionada entre los estudiantes que pertenecen a una misma Facultad y sea independiente si los estudiantes pertenecen a diferentes Facultades.

Tabla 6-7.: Resumen del modelo completo

Coeficientes	β	Estimación	Std error	Z-Value	P-Value	OR
Intercepto	β_0	-2.5581	0.3813	-6.7094	0.00001	
AdmitidoSPP3	β_1	-0.0534	0.5065	-0.1054	0.91607	0.819
Bajos	β_2	1.4883	0.2939	5.0633	0.00001	4.429
Encuentros	β_3	-0.2671	0.0724	-3.6893	0.00022	0.765
DAcademica	β_4	0.6525	0.1395	4.6773	0.00001	1.995
DEconomica	β_5	0.2633	0.2623	1.0037	0.31550	1.491
DFamiliar	β_6	0.6605	0.2923	2.2596	0.02384	1.354
DIndividual	β_7	0.0432	0.2280	-0.1897	0.84956	1.04
DVidaUni	β_8	-1.4554	0.4806	-3.0283	0.00245	0.233
Inasistencia	β_9	0.4025	0.1668	2.4136	0.01579	1.495
SexoMasculino	β_{10}	0.9870	0.3535	2.7919	0.00524	2.683
	d	0.0854				

De la tabla **6-7** se puede inferir que los riesgos que se presentan en las dimensiones académica, económica, individual y familiar aumentan la probabilidad de desertar de un estudiante, mientras ocurre lo contrario con la dimensión vida universitaria, esto puede deberse a que no se observa una diferencia en las puntuaciones del riesgo de los estudiantes desertores y no desertores, o también puede ser por la presencia de casos atípicos.

A partir de los odds ratio (OR) se realizan las interpretaciones de las variables. Por cada bajo rendimiento que presente un estudiante su probabilidad de desertar es 4 veces mayor; también por cada inasistencia a los encuentros pactados con su monitor incrementa en 1.5 su probabilidad de desertar; en caso contrario, por cada encuentro efectivo disminuye en 0.76 la probabilidad de desertar del estudiante. En cuanto a los estudiantes admitidos por SPP3 tienen menor probabilidad de desertar que los estudiantes admitidos por CE2017.

También se observa que las variables dimensión económica, dimensión individual y el tipo de admisión no son significativas en el modelo. Es decir, que el coeficiente podría ser igual a cero y no presentar ninguna relevancia. Por lo tanto, se procede a realizar el método de selección de variables backward, el cual nos permita reducir la complejidad del modelo.

6.3. Selección de Variables

La tabla 6-8 muestra secuencialmente el funcionamiento del método backward, iniciando con un modelo utilizando todas las variables y cada vez excluyendo a la variable que menos aporte en la explicación de la deserción. Se observa que el modelo sin incluir el tipo de admisión, la dimensión económica y la dimensión individual se obtiene la máxima disminución del AIC.

Tabla 6-8.: Selección de variables: Método backward

	Modelo	AIC
1	Admitido + Bajos + Encuentros + DAcademica + DEconomica + DFamiliar + DIndividual + DVidaUni + Inasistencias + Sexo	274.06
2	Bajos + Encuentros + DAcademica + DEconomica + DFamiliar + DIndividual + DVidaUni + Inasistencias + Sexo	272.05
3	Bajos + Encuentros + DAcademica + DEconomica + DFamiliar + DVidaUni + Inasistencias + Sexo	270.16
4	Bajos + Encuentros + DAcademica + DFamiliar + DVidaUni + Inasistencias + Sexo	269.15
5	Bajos + Encuentros + DAcademico + DFamiliar + DVidaUni + Inasistencias	275.1

6.4. Test de Razón de Verosimilitud

A continuación se plantea una comparación entre el modelo completo y el modelo reducido que no incluye las variables AdmitidoSPP3, DIndividual y DEconomica, con el fin de determinar si estas covariables que se están excluyendo del modelo aportan significativamente a la explicación de la deserción estudiantil mediante el análisis de varianzas

$$H_0 : \beta_1 = \beta_5 = \beta_7 = 0 \text{ vs } H_1 : \beta_i \neq 0, i = 1, 5, 7$$

Tabla 6-9.: Test de razón de verosimilitud

Modelo	AIC	RELE	LR	valor-p
Reducido	269.15	-125.58		
Completo	274.06	-125.03	1.09	0.67966

Comparando el modelo completo con el modelo reducido se observa que las variables excluidas no son variables significativas cuando las otras variables ya se encuentran en

el modelo. Es decir, que el modelo reducido proporciona el mismo ajuste que el modelo completo, por lo tanto se decide trabajar con el modelo reducido.

6.5. Modelo Logístico con Intercepto Aleatorio Reducido

Nuevamente se ajusta un modelo logístico con intercepto aleatorio, pero esta sin tener en cuenta las variables excluidas por el método backward. El modelo reducido 6-1 se describe en la tabla 6-10:

$$\text{logit}(\pi_{ij}) = \beta_0 + b_i + \beta_1 \text{Bajos}_{ij} + \beta_2 \text{Encuentros}_{ij} + \beta_3 \text{DAcademica}_{ij} + \beta_4 \text{DFamiliar}_{ij} + \beta_5 \text{DVidaUni}_{ij} + \beta_6 \text{Inasistencias}_{ij} + \beta_7 \text{Sexo}_{ij} \quad (6-1)$$

Tabla 6-10.: Resumen del modelo reducido

Coficientes	β	Estimación	Std error	Z-Value	P-Value	OR
Intercepto	β_0	-2.5351	0.3793	-6.6834	0.00001	
Bajos	β_1	1.5412	0.2897	5.1690	0.00001	4.660
Encuentros	β_2	-0.0531	0.0720	-3.6972	0.00021	0.948
DAcademica	β_3	0.7021	0.1359	4.8333	0.00001	2.017
DFamiliar	β_4	0.7193	0.2401	2.9962	0.00273	2.052
DVidaUni	β_5	-1.3369	0.4604	-2.9035	0.00369	0.262
Inasistencias	β_6	0.4004	0.1673	2.3934	0.01669	1.491
SexoMasculino	β_7	0.9557	0.3489	2.7392	0.00615	2.600
	d	0.1934				

Los coeficientes del modelo se interpretan a partir de los odds ratio (OR), es decir, que tan probable es que un individuo deserte según cierta característica. Del análisis se observa que por cada bajo rendimiento que presente un estudiante tendrá 4.6 veces más chances de desertar, por cada encuentro efectivo con el monitor la probabilidad de desertar es 0.94 veces menor, también, por cada inasistencia registrada por el estudiante su probabilidad de desertar aumenta en 1.5 veces. Con relación a las dimensiones, la dimensión académica y familiar son los factores más determinantes para predecir la deserción de una estudiante, por cada vez que un estudiante sea calificado como alto o medio su probabilidad de desertar se duplica. Además, un estudiante de sexo masculino tiene 2.6 probabilidades de desertar más que una estudiante de sexo femenino.

6.6. Diagnóstico del Modelo

Con el fin de realizar un diagnóstico de los modelos propuestos, se determina la sensibilidad y especificidad del modelo reducido y se compara con la del modelo completo, se presenta la curva ROC con la que se elegirá el punto de corte óptimo que se comparará con los score generados por ambos modelos y así clasificar a los estudiantes como desertores o no, después se presenta la matriz de confusión para ambos modelos.

6.6.1. Modelo Completo

Curva ROC

La curva ROC (Ver figura 6-4) permitirá establecer el punto de corte óptimo para maximizar la sensibilidad y especificidad del modelo, para ello se calcularán las probabilidades de deserción y se clasificará como desertor aquel estudiante que tenga una probabilidad mayor a 0.076.

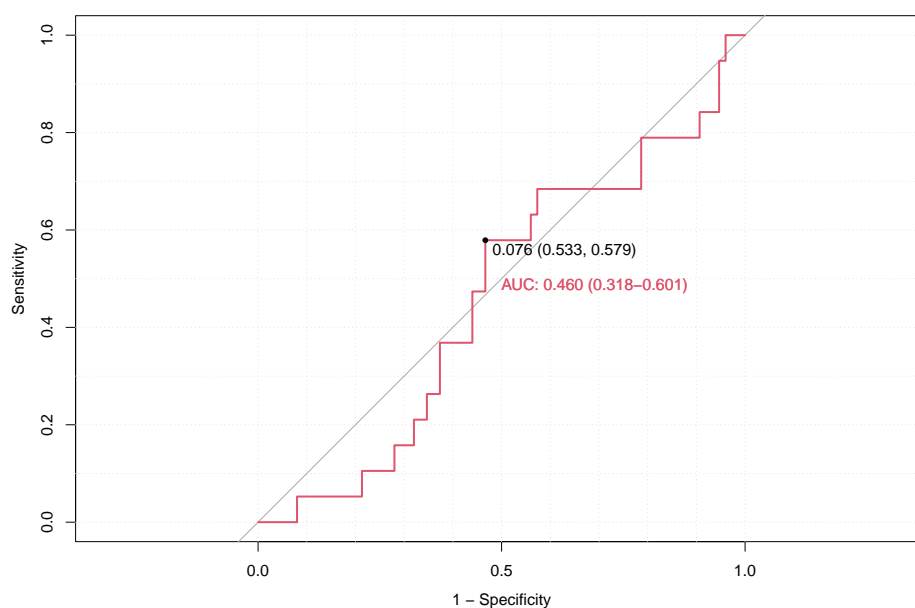


Figura 6-4.: Curva ROC: Modelo completo

Matriz de Confusión Modelo Completo

En la tabla **6-11** se presenta la clasificación de los estudiantes según los score obtenidos. El modelo completo (5-2) clasifico al 57.8 % de los desertores adecuadamente, este resultado se conoce como la sensibilidad del modelo, es decir la proporción de desertores que fueron correctamente identificados por el modelo, en el caso de los estudiantes no desertores el modelo clasificó correctamente al 53.3 % de estos. Globalmente, el modelo clasificó al 54.6 % de las observaciones correctamente.

Tabla 6-11.: Matriz de confusión: Modelo completo

		Valores reales	
		No desertor	Desertor
Predicción	No desertor	40	8
	Desertor	35	11

6.6.2. Modelo Reducido

Curva ROC

La curva ROC (Ver figura **6-5**) permitirá establecer el punto de corte óptimo para maximizar la sensibilidad y especificidad del modelo, para ello se calcularán las probabilidades de deserción y se clasificará como desertor aquel estudiante que tenga una probabilidad mayor a 0.102.

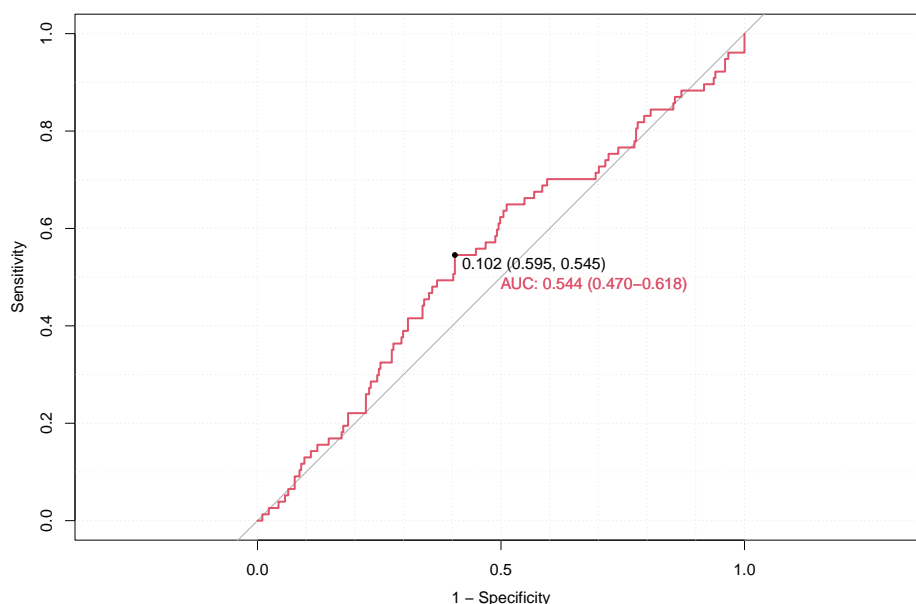


Figura 6-5.: Curva ROC: Modelo reducido

Matriz de Confusión Modelo Reducido

En la tabla **6-12** se presenta la clasificación de los estudiantes según los score obtenidos. El modelo reducido (6-1) clasifico al 64.3 % de los desertores adecuadamente, este resultado se conoce como la sensibilidad del modelo, en el caso de los estudiantes no desertores el modelo clasificó correctamente al 72.5 % de estos. Globalmente, el modelo clasificó al 71.2 % de las observaciones correctamente.

Tabla 6-12.: Matriz de confusión: Modelo reducido

		Valores reales	
		No desertor	Desertor
Predicción	No desertor	58	5
	Desertor	22	9

6.7. Efectos Aleatorios

Por último, la tabla **6-13** muestra la estimación del efecto aleatorio que se incluyó en el modelo con el fin de determinar cuales son las Facultades donde los estudiantes tienen una probabilidad de desertar más alta.

Tabla 6-13.: Estimación efectos aleatorios

Facultad	intercepto
Artes Integradas	0.0135726601
Ciencias de la Administración	-0.0954945853
Ciencias Naturales y Exactas	-0.0766348633
Ciencias Sociales y Económicas	-0.0071866759
Humanidades	0.0951908563
Ingeniería	0.0738621867
Instituto de Educación y Pedagogía	0.0223721373
Psicología	-0.0213462258
Salud	-0.0001821549

Las estimaciones de los efectos aleatorios indican que la Facultad de Ingenierías y la Facultad de Humanidades son las dos Facultades en la que los estudiantes tienen mayor probabilidad de desertar dentro del programa ASES, tal cual como se evidenció en el análisis exploratorio. Ahora bien, resulta interesante que las estimaciones de los efectos aleatorios de la Facultad de Ciencias Naturales y Exactas es bastante baja, siendo esta la cuarta Facultad con mayor índice de deserción, esto se debe a la alta variabilidad dentro del cluster en las variables estudiadas.

7. Conclusiones y Recomendaciones

7.1. Conclusiones

El anterior trabajo aborda la deserción de los estudiantes admitidos por Condición de Excepción y Ser Pilo Paga 3 en el año 2017, mediante el uso de un modelo logístico con un efecto aleatorio asociado a la Facultad, ya que el comportamiento de los factores que influyen en la deserción puede ser similar en dos estudiantes que pertenecen a la misma Facultad y diferente en aquellos que no pertenecen a una misma Facultad. El estudio de la deserción estudiantil dentro del programa ASES es de gran importancia porque permite identificar aspectos en los que deben enfatizar mediante la creación de estrategias que permitan disminuir la deserción estudiantil y así continuar siendo un programa de gran ayuda para los estudiantes a los que les realizan el acompañamiento.

En el análisis descriptivo se observó que los estudiantes admitidos por Ser Pilo Paga presentaban menor índice de deserción a comparación a los estudiantes admitidos por Condición de Excepción. En el modelo se identificó que las dimensiones académica, económica y vida universitaria, el número de bajos que tiene el estudiante, el número de encuentros e inasistencias que tuvo un estudiante y el sexo son variables que influyen a la hora de que un estudiante deserte ya que fueron variables importantes en el momento de predecir adecuadamente si un estudiante desertará; además el riesgo relacionado con la vida universitaria es negativo que puede deberse a la influencia de un dato atípico; también por cada bajo académico la probabilidad de desertar se incrementa casi 5 veces; el hecho de ser un hombre duplica la probabilidad de desertar a comparación de las mujeres y la dimensión más riesgosa es la académica ya que por cada vez que un estudiante tenga riesgo medio o alto en la dimensión académica la probabilidad de desertar se duplica.

Incluir la Facultad como efecto aleatorio en el modelo nos permitió tener un análisis de la ocurrencia de deserción a nivel de facultades en donde la Facultad de Ingeniería y la Facultad de Humanidades son quienes presentan mayor probabilidad de desertar en función de las variables observadas, ya que poseen las estimaciones con valores más altos. También, a pesar de que la Facultad de Ciencias Naturales y Exactas es la cuarta Facultad que presenta mayor índice de deserción, las estimaciones de los efectos aleatorios permitió identificar que existe una gran variabilidad en la deserción de los estudiantes por lo que la explicación de la deserción se deba a otras variables que no fueron incluidas en el estudio.

Este trabajo aporta una visión diferente de la deserción ya que utiliza información del estudiante desde que entra a estudiar la carrera hasta que finaliza el ciclo básico de formación, la razón es que la situación de un estudiante cambia a lo largo del tiempo y por ende los factores que llevan a un estudiante a desertar cambian constantemente. El modelo permitió identificar las dimensiones en las que debe enfatizar estrategias para disminuir las probabilidades de deserción con una adecuada clasificación de los estudiantes que desertaron o no en un 70 % de las veces; el incluir otras variables en el estudio podría aumentar la cantidad de predicciones correctas.

7.2. Recomendaciones

Con el fin de aprovechar al máximo toda la información que recopila el programa ASES de sus estudiantes semanalmente a partir de los encuentros socioeducativos, se recomienda implementar un aplicativo donde se pueda recrear el modelo construido en este trabajo y se pueda llevar un control de alerta temprana de los estudiantes con el fin de realizar una intervención con el propósito de disminuir el riesgo de deserción que presenta cada estudiante.

Una de las limitaciones con las que nos encontramos fue poca consolidación de información sociodemográfica de los estudiantes, pues al momento de solicitar la información, lo máximo que se pudo acceder fue a una encuesta realizada por los estudiantes que aunque contaba con información interesante como la edad, lugar de nacimiento, lugar de residencia actual, servicio de salud, número de personas a cargo, entre otras variables; no nos resultó útil agregarlas al modelo por el alto porcentaje de datos faltantes, generando inconsistencias en los resultados. Por lo tanto, se recomienda para futuras investigaciones el incluir este tipo de variables ya que puede aumentar la precisión del modelo a la hora de identificar si un estudiante va a desertar.

Bibliografía

Acevedo, J., Vélez, L. & González, L. (2012), 'Estrategias de acompañamiento familiar para el restablecimiento de derechos de la niñez en el contexto de hogares sustitutos', *Katharsis* pp. 151–176.

Alvarado, G. (2019), Impacto de la estrategia de acompañamiento y seguimiento estudiantil ASES, en el rendimiento académico de los estudiantes de (Ser Pilo paga) en la Universidad del Valle: una estimación por propensity score matching, Cali, Colombia: Universidad del Valle., PhD thesis.

Área de Admisiones (s.f.), 'Condiciones de excepción - Sede Cali', *Cali, Colombia: Universidad del Valle*. Disponible en: <http://admisiones.univalle.edu.co/new/condiciones/index.php> .

Arismendy, C. & Morales, N. (2018), 'Modelo de regresión logística como alternativa para medir la probabilidad de deserción temprana en la Universidad de los Llanos en el periodo 2015-2 - 2018-1', *Villavicencio, Colombia: Universidad de los Llanos* .

Barrios, J. (2019), 'La matriz de confusión y sus métricas', *Health Big Data* .

Castañó, H. F. & Ramírez, F. O. P. (2005), 'El modelo logístico: una herramienta estadística para evaluar el riesgo de crédito', *Revista Ingenierías Universidad de Medellín* 4(6), 55–75.

Correa, J. & Salazar, J. (2016), 'Introducción a los modelos mixtos', *Medellín, Colombia: Universidad Nacional de Colombia-Sede Medellín* .

Escobar, J., Largo, E. & Perez, C. (2006), 'Factores asociados a la admisión, deserción y permanencia estudiantil en la Universidad del Valle', *Cali, Colombia: Universidad del Valle* .

Escobar, J., Largo, E. & Perez, C. (2008), 'Rendimiento académico en la Universidad del Valle: determinantes y su relación con la deserción estudiantil', *Cali, Colombia: Universidad del Valle* .

Estrategia ASES (2022a), 'Estrategia ASES - Universidad del Valle', *Cali, Colombia: Universidad del Valle*. Disponible en: <https://ases.univalle.edu.co/> .

- Estrategia ASES (2022b), ‘Semaforo de alertas’, *Cali, Colombia: Universidad del Valle*.
Disponible en: <https://asesinteractiva.univalle.edu.co/semaforoalertas/> .
- García, C. P. et al. (2012), Estimación del modelo logístico mixto: revisión y nueva propuesta, PhD thesis, Medellin, Colombia: Universidad EAFIT.
- García, F. & Díaz, M. (2011), ‘Modelos mixtos generalizados para el estudio del desempleo en los grandes aglomerados urbanos de Argentina’, *Revista de Economía y Estadística* **49**(1), 79–98.
- González, A. et al. (2015), ‘Selección de variables: Una revisión de métodos existentes’, *La Coruña, España: Universidad da Coruña* .
- Hernández, F. & López, J. (2022), ‘Modelos mixtos con r’.
- Hosmer, D. W. & Lemeshow, S. (1980), ‘Goodness of fit tests for the multiple logistic regression model’, *Communications in statistics-Theory and Methods* **9**(10), 1043–1069.
- ICETEX (s.f.), ‘Ser pilo paga 3’, *Bogotá, Colombia: Disponible en: https://web.icetex.gov.co/es/-/ser-pilo-paga-3* .
- Jennings, D. E. (1986), ‘Judging inference adequacy in logistic regression’, *Journal of the American Statistical Association* **81**(394), 471–476.
- Lozano, S. (2018), ‘Una aproximación a los determinantes de la deserción academica en la Universidad del Valle sede Cali, para el caso de los estudiantes admitidos por condición de excepción en su primer año de estudio’, *Cali, Colombia: Universidad del Valle* .
- Meneses, A. (2009), ‘La deserción de los estudiantes que ingresaron por condición de excepción indígena a la Universidad del Valle en 2001-ii-2006’, *Cali, Colombia: Universidad del Valle* .
- Montgomery, D. C., Peck, E. A. & Vining, G. G. (2021), *Introduction to linear regression analysis*, John Wiley & Sons.
- Reche, J. L. C. (2013), ‘Regresión logística. tratamiento computacional con r.’.
- Ruiz, C. G., Muriel, D. M. D., Gallego, J. F., Velez, E. C., Gomez, S. G. & Portilla, K. G. (2009), *Deserción estudiantil en la educación superior colombiana: Metodología de seguimiento, diagnóstico y elementos para su prevención*, Ministerio de educación nacional.
- Salamanca, D. (2019), Efecto de la estrategia ASES sobre el desempeño académico de los estudiantes de Condición de Excepción en la Universidad del Valle durante el año 2017, Cali, Colombia: Universidad del Valle., PhD thesis.

SPADIES (s.f.), ‘Estadísticas de deserción’, *Sistema para la Prevención y Análisis de la Deserción en las Instituciones de Educación Superior*, Ministerio de educación nacional .

Universidad del Valle (1998), ‘Por la cual se modifican, parcialmente, las resoluciones no. 039 y 067 de 1998 que establecen los criterios de admisión y la oferta educativa para los programas de pregrado en la Sede Cali y del sistema de regionalización en lo que concierne a las condiciones de excepción para los indígenas’, *Cali, Colombia* .

Universidad del Valle (2003), ‘“por la cual se crea la Condición de Excepción comunidades afrocolombianas para el ingreso a los Programas Académicos de Pregrado’, *Cali, Colombia* .

Universidad del Valle (2005), ‘Plan estratégico de desarrollo 2005 - 2015’, *Cali, Colombia* .

Verbeke, G., Fitzmaurice, G., Davidian, M. & Molenberghs, G. (2008), *Longitudinal data analysis*, CRC press.

Westreicher, G. (2021), ‘Datos agrupados’.

A. Anexo: Temáticas y Niveles de Riesgo

Las 5 dimensiones están conformadas por diferentes temáticas, en las que se miden los niveles de riesgo que indican la probabilidad de que un estudiante pueda desertar, y se clasifican en bajo, medio y alto; cuando el nivel de riesgo es alto desde el programa ASES se realiza una intervención con el propósito de disminuir este riesgo en la temática evaluada.

Dimensión Individual

Autoconocimiento

Tipo de saber que tiene cada sujeto de sí mismo, ya sea de sus representaciones, estados mentales, percepciones, acciones, de su cuerpo, entre otros. En esta temática se incluyen todos los aspectos de conocimiento de sí mismos que los estudiantes expresan sobre cómo se sienten, lo que desean, lo que piensan, lo que los impulsa a actuar, sus valores, todo aquello que constituye su ser desde tres pilares: identidad, autoestima y autoconstrucción.

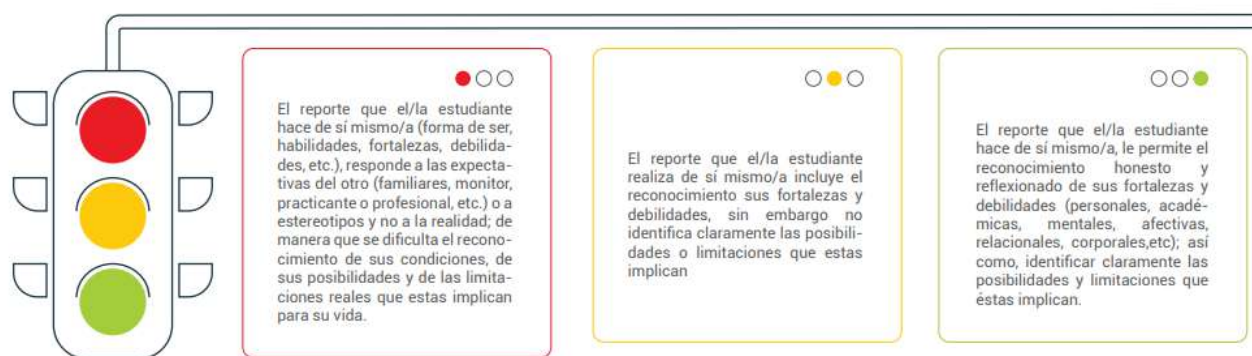


Figura A-1.: Autoconocimiento. Tomado de: Estrategia ASES.

Proyecto de vida

Esta temática engloba todos los proyectos y metas a mediano y largo plazo que los estudiantes manifiestan en los acompañamientos entre pares. El proyecto de vida articula la identidad personal-social en las perspectivas de su dinámica temporal y posibilidades de desarrollo futuro.

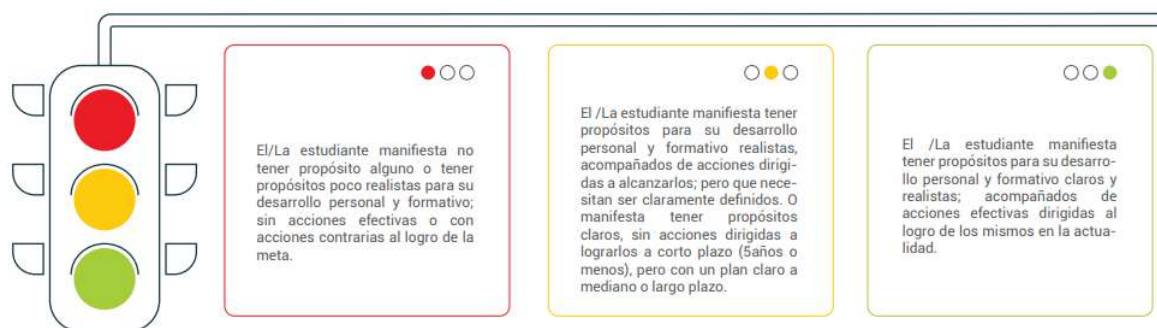


Figura A-2.: Proyecto de vida. Tomado de: Estrategia ASES.

Historia de vida

Esta temática hace referencia a todas las narraciones de vida en tiempo pasado que realizan los estudiantes desde los diferentes contextos socioculturales que vivieron; da cuenta de las prácticas, creencias y valores familiares y culturales que influyen en sus decisiones y en sus formas de ver el mundo.

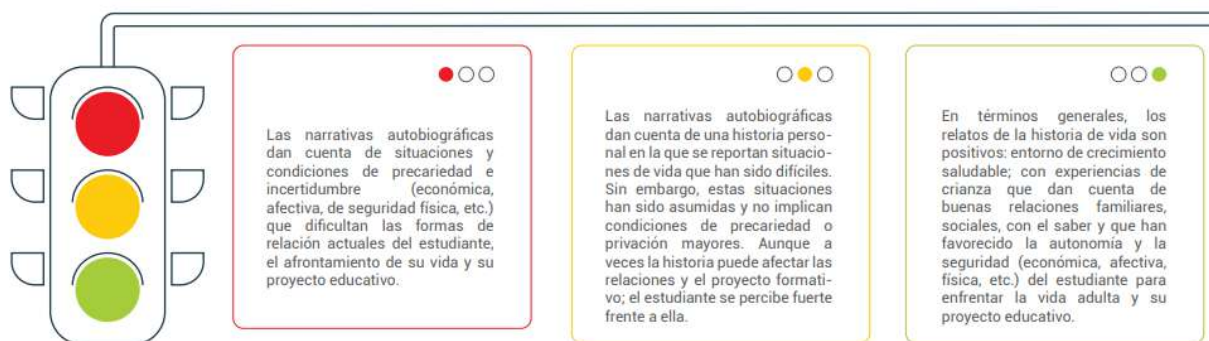


Figura A-3.: Historia de vida. Tomado de: Estrategia ASES.

Rasgos de personalidad

Esta temática está compuesta por las descripciones y apreciaciones que los monitores realizan sobre sus estudiantes asignados en cuanto a características cognitivas, comportamentales, anímicas, y temperamentales.

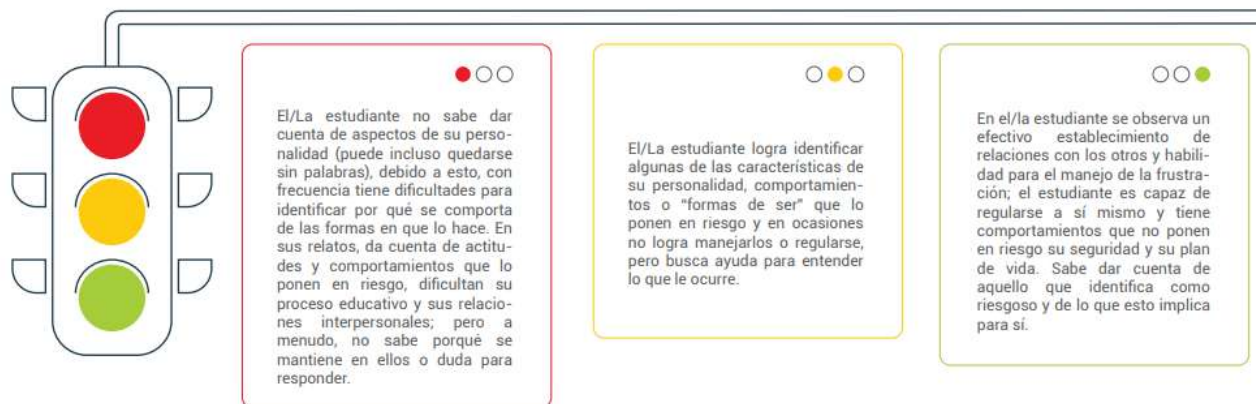


Figura A-4.: Rasgos de personalidad. Tomado de: Estrategia ASES.

Salud

Esta temática hace referencia aquellas acciones que permitan identificar y conocer el estado de salud del estudiante, así como su evolución en caso de enfermar o sufrir algún accidente. También se refiere a los reportes que se tengan acerca de la salud mental del estudiante.

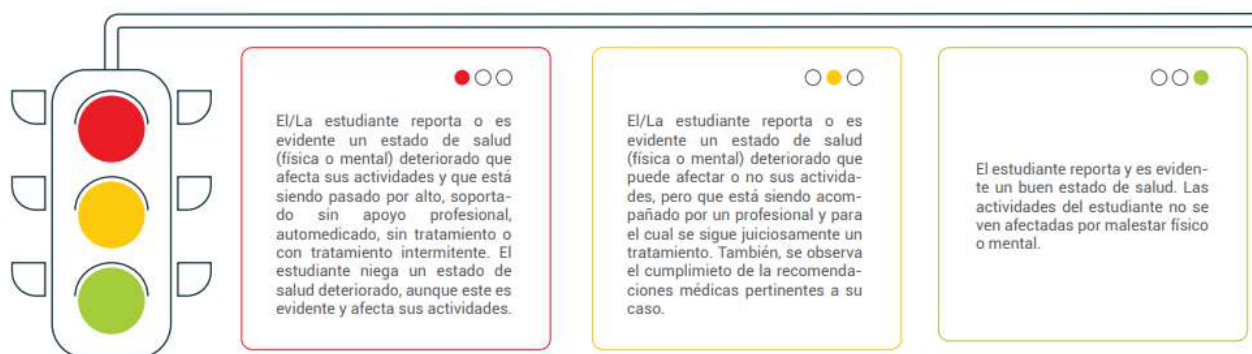


Figura A-5.: Salud. Tomado de: Estrategia ASES.

Relación erótica-afectiva

Esta temática contiene todo lo referido por los estudiantes narrativamente en relación con sus parejas afectivas y sentimentales.

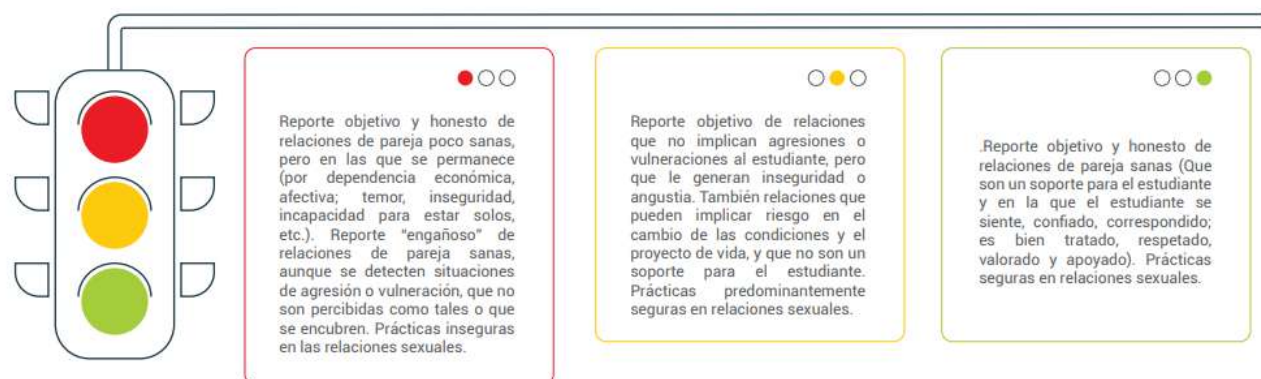


Figura A-6.: Relación erótica-afectiva. Tomado de: Estrategia ASES.

Identificación

Esta temática comprende todo lo referente a la obtención, cambio o pérdida de documentos esenciales que se encuentran inscritos en registros oficiales y que le permiten al estudiante el acceso a los servicios sociales del Estado.

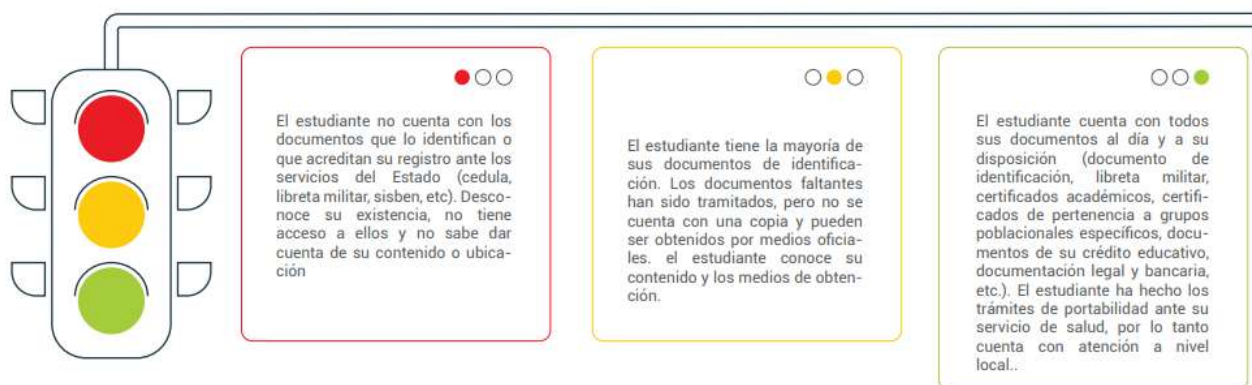


Figura A-7.: Identificación. Tomado de: Estrategia ASES.

Aspectos motivacionales

El proceso psicológico básico de la motivación contiene dos componentes principales: “los direccionales (que dan cuenta de la elección) y los energizadores (que dan cuenta de la iniciación, la persistencia y el vigor) de la conducta dirigida a meta”.



Figura A-8.: Aspectos motivacionales. Tomado de: Estrategia ASES.

Diversidad sexual

Todos aquellos reconocimientos, prácticas y relaciones establecidas por el estudiante; frente a su género, identidad sexual, preferencia sexual y, ante la diversidad sexual de las personas en su contexto cotidiano.



Figura A-9.: Diversidad sexual. Tomado de: Estrategia ASES.

Dimensión Académica

Desempeño académico

Esta temática incluye las narraciones realizadas por los estudiantes en torno a las metodologías utilizadas en las clases, información de las diversas asignaturas y seguimientos académicos; ejemplo: reporte del rendimiento académico (notas), reporte de la carga académica, informe de las necesidades académicas identificadas con el estudiante (bases conceptuales para las diversas asignaturas) y refuerzo académico (asesorías académicas). Estos seguimientos tienen como finalidad identificar y fortalecer las habilidades y destrezas de los estudiantes.



Figura A-10.: Desempeño académico. Tomado de: Estrategia ASES.

Elección vocacional

Esta temática hace referencia a las preferencias de áreas disciplinares y carreras de los estudiantes, la búsqueda de información y conocimiento de sus carreras de elección y las estrategias implementadas por los monitores para contribuir en sus procesos de orientación vocacional concebida ésta como un vínculo conversacional en el que el estudiante recibe apoyo en el marco de encontrar alternativas y tomar decisiones, de manera consciente voluntaria y comprometida (De Mori Santiviago, s.f.).

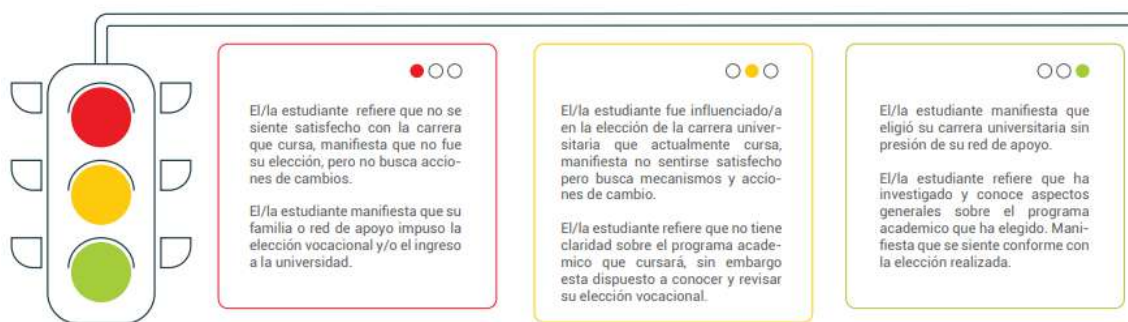


Figura A-11.: Elección vocacional. Tomado de: Estrategia ASES.

Manejo del tiempo

Esta temática contiene lo referente a la forma como los estudiantes manejaban el tiempo de acuerdo al establecimiento de sus rutinas diarias, comprendiendo tres aspectos claves: 1) el establecimiento de metas, 2) las herramientas para la gerencia del tiempo y 3) la percepción de control o verificación del uso del tiempo personal.



Figura A-12.: Manejo del tiempo. Tomado de: Estrategia ASES.

Dimensión Económica

Apoyos económicos institucionales

Apoyos económicos institucionales: incluye todos los procesos que los estudiantes realizan para recibir apoyos económicos. (ICETEX, Jóvenes en Acción, Bienestar Universitario, monitorias, etc.)

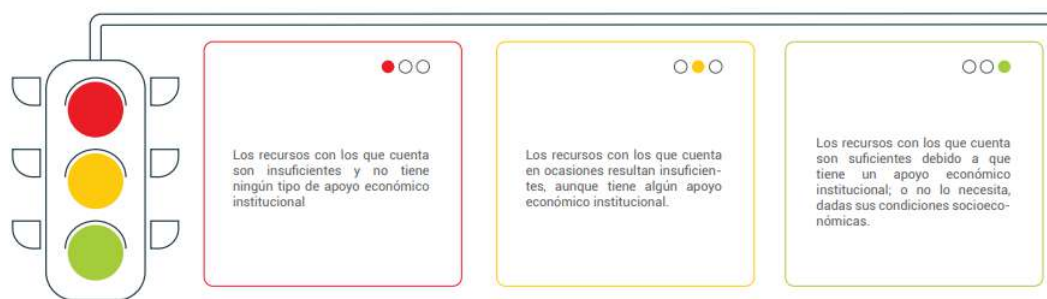


Figura A-13.: Apoyos económicos institucionales. Tomado de: Estrategia ASES.

Apoyo económico familiar

Lo socioeconómico, entendido como un “enfoque teórico y metodológico necesariamente transdisciplinar, que pretende entender integralmente la complejidad social a partir de la observación, descripción y análisis orientada a la acción en y desde la realidad” (Coraggio Arancibia, 2006). Esta temática involucra aspectos económicos y sociológicos como la preparación laboral, ubicación social y familiar en la sociedad.

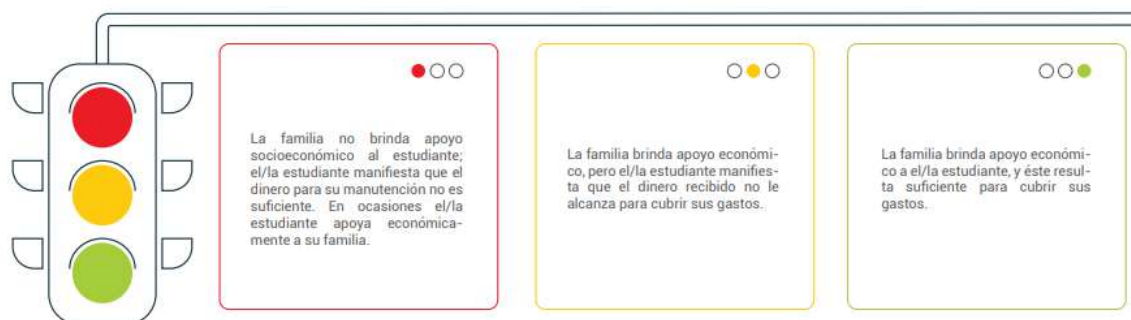


Figura A-14.: Apoyo económico familiar. Tomado de: Estrategia ASES.

Situación laboral y ocupacional

Se refiere a las diversas actividades que realiza el estudiante para cubrir sus gastos personales y académicos; y la relación de estas ocupaciones con su actividad académica en la Universidad.



Figura A-15.: Situación laboral y ocupacional. Tomado de: Estrategia ASES.

Dimensión Vida universitaria y Ciudad

Motivaciones para el acompañamiento

Esta temática aborda los momentos de presentación entre el monitor y el estudiante, incluyendo además la explicación de la estrategia ASES y expectativas de ingreso a la universidad de parte de los estudiantes y del acompañamiento que van a tener.



Figura A-16.: Motivaciones para el acompañamiento. Tomado de: Estrategia ASES.

Referenciación geográfica

En este apartado nos encontramos con el conocimiento de nuevos lugares en sus territorios cotidianos de parte de los estudiantes (tanto en la universidad como en la ciudad), muchos de ellos guiados por sus monitores.



Figura A-17.: Referenciación geográfica. Tomado de:Estrategia ASES.

Adaptación a la ciudad y universidad

Esta temática involucra todo lo expresado por los estudiantes en relación a la adaptación que se encuentran realizando al nuevo contexto de ciudad en el que se encuentra, para el caso de estudiantes que proceden de otros municipios y regiones del país. Además de las diferentes experiencias que expresan los estudiantes con relación a su adaptación a la Universidad.

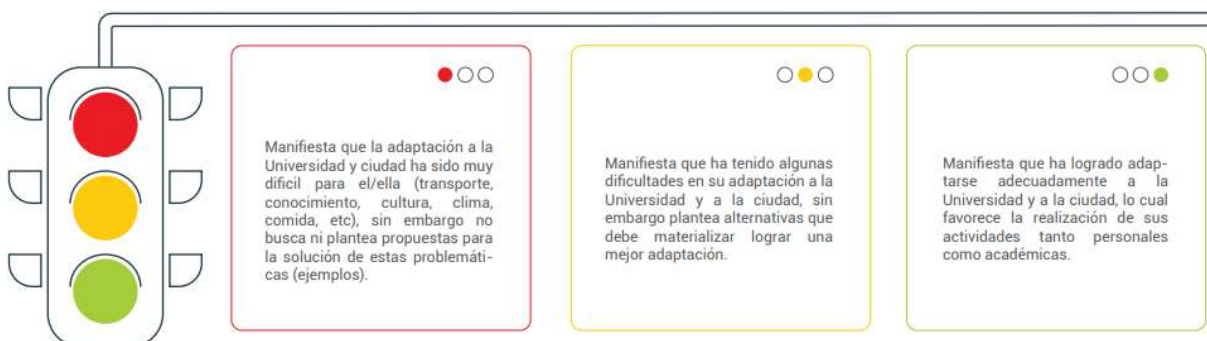


Figura A-18.: Adaptación a la ciudad y universidad. Tomado de:Estrategia ASES.

Oferta de servicios

Este apartado contiene la oferta de servicios institucionales y no formales que los monitores realizaban a sus estudiantes en relación al ámbito universitario, de ciudad y nacional.



Figura A-19.: Oferta de servicios. Tomado de: Estrategia ASES.

Vinculación a grupos y realización de actividades extracurriculares

Este apartado contiene lo relacionado con el interés, exploración y vinculación de los estudiantes a grupos estudiantiles, académicos, investigativos, culturales, y deportivos de la Universidad del Valle o externos.

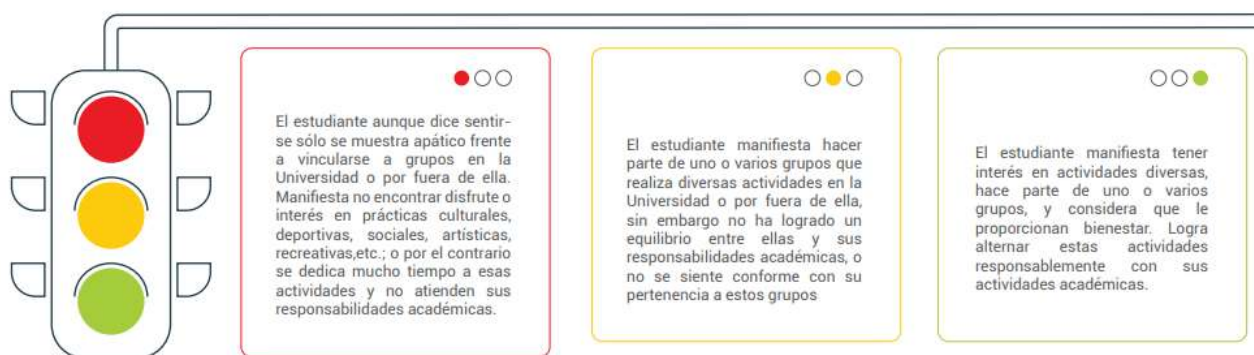


Figura A-20.: Vinculación a grupos y realización de actividades extracurriculares. Tomado de: Estrategia ASES.

Dimensión Familiar

Dinámica familiar: Se define: “como los encuentros entre las subjetividades, encuentros mediados por una serie de normas, reglas, límites, jerarquías y roles, entre otros, que regulan la convivencia y permite que el funcionamiento de la vida familiar se desarrolle armónicamente”. El seguimiento familiar implica reconocer la existencia de lazos de parentesco, afecto, comunicación, límites, jerarquías, roles, toma de decisiones, los cuales se van construyendo entre sus miembros de la familia y definen la realización del ser y la búsqueda del bienestar colectivo (Acevedo et al., 2012).



Figura A-21.: Dinámica familiar. Tomado de: Estrategia ASES.