

**UNIVERSIDAD NACIONAL DE SAN ANTONIO ABAD DEL CUSCO
FACULTAD DE CIENCIAS QUÍMICAS, FÍSICAS Y MATEMÁTICAS
ESCUELA PROFESIONAL DE MATEMÁTICA**



TESIS

**ANÁLISIS DE DESERCIÓN ESTUDIANTIL EN LA ESCUELA
PROFESIONAL DE MATEMÁTICA DE LA UNSAAC EN EL AÑO
2015 A 2020 MEDIANTE MODELOS DE SUPERVIVENCIA**

PRESENTADO POR:

Br. DEMETRIO PAUCCAR QUISPE
Br. LUIS ALBERTO HUAMAN HUARCO

**PARA OPTAR AL TÍTULO PROFESIONAL
DE LICENCIADO EN MATEMÁTICA
MENCION ESTADÍSTICA**

ASESOR:

Dr. CLETO DE LA TORRE DUEÑAS

CUSCO – PERÚ

2026



Universidad Nacional de San Antonio Abad del Cusco

INFORME DE SIMILITUD

(Aprobado por Resolución Nro.CU-321-2025-UNSAAC)

El que suscribe, el Asesor CLETO DE LA TORRE DUEÑAS.....
..... quien aplica el software de detección de similitud al
trabajo de investigación/tesistitulada: ANÁLISIS DE DESERCIÓN ESTUDIANTIL
EN LA ESCUELA PROFESIONAL DE MATEMÁTICA DE LA UNSAAC EN
EL AÑO 2015 A 2020 MEDIANTE MODELOS DE SUPERVIENCIA
.....

Presentado por: DEMETRIO PAUCCAR QUISPE..... DNI N° 76353999.....;
presentado por: LUIS ALBERTO HUAMAN HUARCO..... DNI N°: 71140971
Para optar el título Profesional/Grado Académico de LICENCIADO EN MATEMÁTICA
MENCION ESTADÍSTICA.....

Informo que el trabajo de investigación ha sido sometido a revisión por 03... veces, mediante el Software de Similitud, conforme al Art. 6° del **Reglamento para Uso del Sistema Detección de Similitud en la UNSAAC** y de la evaluación de originalidad se tiene un porcentaje de 09.....%.

Evaluación y acciones del reporte de coincidencia para trabajos de investigación conducentes a grado académico o título profesional, tesis

Porcentaje	Evaluación y Acciones	Marque con una (X)
Del 1 al 10%	No sobrepasa el porcentaje aceptado de similitud.	X
Del 11 al 30 %	Devolver al usuario para las subsanaciones.	
Mayor a 31%	El responsable de la revisión del documento emite un informe al inmediato jerárquico, conforme al reglamento, quien a su vez eleva el informe al Vicerrectorado de Investigación para que tome las acciones correspondientes; Sin perjuicio de las sanciones administrativas que correspondan de acuerdo a Ley.	

Por tanto, en mi condición de Asesor, firmo el presente informe en señal de conformidad y **adjunto** las primeras páginas del reporte del Sistema de Detección de Similitud.

Cusco, 18 de FEBRERO..... de 2026.....


.....
Firma

Post firma CLETO DE LA TORRE DUEÑAS.....

Nro. de DNI 23988416.....

ORCID del Asesor 0000 - 0003 - 0921 - 7217.....

Se adjunta:

- Reporte generado por el Sistema Antiplagio.
- Enlace del Reporte Generado por el Sistema de Detección de Similitud: **oid:** 27259:558158158283.....

DEMETRIO y LUIS ALBERTO PAUCCAR QUISPE y H...

ANÁLISIS DE DESERCIÓN ESTUDIANTIL EN LA ESCUELA PROFESIONAL DE MATEMÁTICA DE LA UNSAAC EN EL AÑO 2...

 Universidad Nacional San Antonio Abad del Cusco

Detalles del documento

Identificador de la entrega

trn:oid:::27259:558158283

137 páginas

Fecha de entrega

18 feb 2026, 7:01 a.m. GMT-5

24.025 palabras

Fecha de descarga

18 feb 2026, 7:28 a.m. GMT-5

136.732 caracteres

Nombre del archivo

FALTA TURNITING TESIS 2026.docx

Tamaño del archivo

8.4 MB

9% Similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para.ca...




Filtrado desde el informe

- Bibliografía
- Texto citado
- Texto mencionado
- Coincidencias menores (menos de 8 palabras)
- Trabajos entregados

Exclusiones


- N.º de coincidencias excluidas

Fuentes principales

- 9%  Fuentes de Internet
- 3%  Publicaciones
- 0%  Trabajos entregados (trabajos del estudiante)

Marcas de integridad

N.º de alerta de integridad para revisión

-  **Texto oculto**
319 caracteres sospechosos en N.º de páginas
El texto es alterado para mezclarse con el fondo blanco del documento.

Los algoritmos de nuestro sistema analizan un documento en profundidad para buscar inconsistencias que permitirían distinguirlo de una entrega normal. Si advertimos algo extraño, lo marcamos como una alerta para que pueda revisarlo.

Una marca de alerta no es necesariamente un indicador de problemas. Sin embargo, recomendamos que preste atención y la revise.

DEDICATORIA

A nuestro señor Dios, por darme la vida, por estar siempre conmigo cuidándome bendiciéndome en los momentos difíciles, en mis logros, en mis tristezas, en mis alegrías que siempre me escuchó mi oración que hoy se cumple mi meta trazada, uno de mis deseos de ser profesional.

¡Doy gracias por siempre! A mi madre, GREGORIA QUISPE YAPU, por siempre darme las fuerzas necesarias para seguir adelante, por su apoyo, por su amor incondicional que me da, por enseñarme los valores de la vida para ser una persona buena en la sociedad.

A mi padre, MATEO PAUCCAR CHAUCA, por darme las fuerzas necesarias desde el cielo, a mis hermanos, ALCIDES, MILAGROS Y JORGE por estar siempre en los momentos buenas y malas, por su apoyo moral, por sus motivaciones y alientos para concluir mi meta.

A mis tíos, JUAN PAUCCAR, CECILIA, MARIO, JULIA, por su apoyo incondicional para poder culminar mis estudios estaré agradecido eternamente. A mis primos, JAVIER, BERTHA, WILSON, WASHINGTON Y FERNANDO.

A mis cuñados, GRIMALDO Y SANDRA y a mis DOCENTES, de la Escuela Profesional de MATEMATICA, por sus enseñanzas, por los conocimientos adquiridos y ser un ejemplo a seguir.

Demetrio

A ti, mi Dios, todo poderoso por darme sabiduría, salud, fuerza y guiarme para alcanzar mis metas como persona y como profesional.

Con mucho amor a mis queridos padres Jose Huaman Huarco y Benedicta Huarco Tapia, por haberme brindado su amor y apoyo incondicional ya que son el principal cimiento y sustento para culminar mi carrera profesional.

A mis queridas hermanas Maritsa Huaman Huarco, Erica, Jiulisa, Santusa, Rosa y Cerila, gracias por su apoyo y ser los faros en mi vida; de igual manera, dedico este trabajo a mis familiares por su apoyo forjando mi camino profesional.

A Carmen M. mi mayor fortaleza durante este camino académico, gracias por tu apoyo y la motivación que me impulsó a seguir adelante, por tu paciencia infinita ante mis ausencias, por las palabras de aliento que me calmaban mi cansancio.

Luis Alberto

AGRADECIMIENTO

A Dios todo poderoso, por darnos el milagro de vida, guiarnos por el camino de bien con sabiduría y por permitirnos lograr nuestro sueño anhelado.

A nuestro asesor Dr. Cleto De La Torre Dueñas, por su visión crítica sobre muchos aspectos cotidianos de la vida, por la rectitud en su profesión como docente, quien, con sus consejos, guía, y motivación ha logrado que podamos culminar nuestra tesis de pregrado.

A todos los docentes del Departamento Académico de Matemática y Estadística de la Universidad Nacional de San Antonio Abad del Cusco, por sus apoyos incondicionales y colaboraciones durante los años de formación en nuestra vida profesional.

También queremos expresar nuestro especial agradecimiento a nuestra amiga Elizabeth Ponce Pumacayo, por el apoyo y la compañía brindada durante la realización de esta tesis. Su ayuda, paciencia y cariño fueron fundamentales para culminar con éxito este trabajo, y siempre recordaremos con gratitud su valiosa contribución.

Son muchas las personas que han sido parte fundamental de nuestra formación profesional, a quienes agradecemos profundamente por su amistad, su consejo, su apoyo, su aliento y compañía en los momentos más difíciles de nuestra vida.

Los tesistas

ÍNDICE GENERAL

DEDICATORIA	ii
AGRADECIMIENTO	iv
ÍNDICE GENERAL	v
ÍNDICE DE TABLA	x
ÍNDICE DE FIGURAS.....	xi
PRESENTACIÓN.....	xii
RESUMEN	xiii
ABSTRACT.....	xiv
INTRODUCCIÓN	xv
CAPÍTULO I	1
PLANTEAMIENTO DEL PROBLEMA	1
1.1 Situación Problemática	1
1.2 Formulación del problema	2
1.2.1 Problema general	2
1.2.2 Problemas específicos	3
1.3 Justificación	3
1.3.1 Justificación Social	3
1.3.2 Justificación Económica	4

1.3.3 Justificación Práctica	5
1.3.4 Justificación teórica	5
1.3.5 Justificación metodológica.....	6
1.4 Objetivos de la investigación.....	6
1.4.1 Objetivo General.....	6
1.4.2 Objetivos Específicos.....	7
1.5 Delimitación de la investigación.....	7
1.5.1 Delimitación espacial.....	7
1.5.2 Delimitación Temporal	7
1.5.3 Delimitación conceptual	8
CAPÍTULO II.....	9
MARCO TEÓRICO.....	9
2.2 Bases Teóricas	9
2.2.1 Análisis de Supervivencia.....	9
2.2.1.1 Definiciones de términos utilizados en el análisis de supervivencia	10
2.2.1.2 Elementos Básicos	11
2.2.1.3 Función de supervivencia	11
2.2.1.4 Función de riesgo.....	12
2.2.1.5 Censura	13
2.2.1.5.1 Censura por la derecha.....	14

2.2.1.5.2 Censura por izquierda	14
2.2.1.5.3 Censura por intervalos	14
2.2.2 Método paramétrico	14
2.2.2.1 Modelo Exponencial.	14
2.2.2.2 Modelo Weibull	16
2.2.2.3 Modelo log-Normal	17
2.2.3 Método no Paramétrico.....	17
2.2.3.1 Estimador de Kaplan-Meier.....	17
2.2.3.2 Varianza del Estimador Kaplan-Meier	18
2.2.3.3 Comparación funciones de supervivencia	23
2.2.3.3.1 Prueba de hipótesis para la función de supervivencia	25
2.2.3.3.2 Prueba de Log- Rank	26
2.2.4 Método Semiparamétrico.....	27
2.2.4.1 Modelo de Regresión de Cox.....	27
2.2.4.2 Estimación de los parámetros en el modelo de Cox	29
2.2.4.2.1 Estimación de β	30
2.2.4.3 Estimación de la función de riesgo base $h_0(t)$	32
2.2.4.4 Prueba de Wald.....	33
2.2.5 DESERCIÓN ESTUDIANTIL	34
2.2.5.1 Definición de Deserción	34

2.2.5.2 Clases y actores de Deserción.....	34
2.2.5.3 Factores asociados	35
2.2.5.3.1 Factores individuales	35
2.2.5.3.2 Factores socioeconómicos	35
2.2.5.3.3 Factores académicos	36
2.2.5.3.4 Factores institucionales.....	36
2.3 Marco conceptual.....	36
2.4 Antecedentes de la investigación	38
2.4.1 Antecedentes internacionales.....	38
2.4.2 Antecedentes nacionales	43
2.4.3 Antecedentes locales.....	48
CAPÍTULO III.....	49
HIPÓTESIS Y VARIABLES	49
3.1 Hipótesis	49
3.1.1 Hipótesis general.....	49
3.1.2 Hipótesis específicas.....	49
3.2 Identificación de variables	50
3.2.1 Variable dependiente	50
3.2.2 Variable independiente	50
3.3 Operacionalización de variable.....	51

CAPÍTULO IV.....	52
METODOLOGÍA.....	52
4.1 Ámbito de estudio, localización y geografía.....	52
4.2 Tipo y nivel de investigación.....	53
4.3 Diseño de Investigación.....	53
4.4 Unidad de análisis.....	53
4.5 Población de estudio.....	54
4.6 Tamaño de muestra.....	55
4.7 Técnicas de selección de muestra.....	56
4.8 Técnicas de recolección de información.....	56
4.9 Técnicas de análisis e interpretación de la información.....	56
RESULTADOS.....	58
DISCUSIÓN.....	84
CONCLUSIONES.....	89
RECOMENDACIONES.....	91
REFERENCIAS BIBLIOGRÁFICAS.....	93
ANEXOS.....	98
Matriz de consistencia.....	99

ÍNDICE DE TABLA

Tabla 1 Prueba de igualdad de la función de supervivencia de dos grupos en el tiempo de observación t_i	23
Tabla 2 Funciones básicas de supervivencia cálculo de estadísticas de prueba de igualdad funcional	24
Tabla 3 Población estudiantil de la facultad de Escuela profesional de Matemática.	54
Tabla 4 Muestra del estudio	55
Tabla 5 Análisis de Supervivencia.....	58
Tabla 6 Supervivencia - Modalidad de Ingreso	59
Tabla 7 Supervivencia - Sexo	62
Tabla 8 función de riesgo acumulado	70
Tabla 9 Función de riesgo acumulado según modalidad de ingreso.....	71
Tabla 10 Estimación paramétrica de la función de supervivencia.....	73
Tabla 11 Comparación de modelos.....	78
Tabla 12 Regresión de Cox con variables cuantitativas	80
Tabla 13 Regresión de cox con variables cualitativas	82

ÍNDICE DE FIGURAS

Figura 1 Análisis de Supervivencia	59
Figura 2 Supervivencia - Modalidad de Ingreso.....	61
Figura 3 Supervivencia - Género	63
Figura 4 Supervivencia - Procedencia	64
Figura 5 Supervivencia según dependencia económica.	65
Figura 6 Supervivencia según tipo de colegio.....	67
Figura 7 Supervivencia según sexo y procedencia.	68
Figura 8 Riesgo acumulado meses.....	73
Figura 9 Probabilidad de supervivencia tiempo	74
Figura 10 Comparación entre curvas de supervivencia.....	75
Figura 11 Comparación entre riesgo acumulado	77
Figura 12 Ajuste paramétrico 1	79
Figura 13 Ajuste paramétrico 2	80

PRESENTACIÓN

SEÑOR DECANO DE LA FACULTAD DE CIENCIAS QUÍMICAS, FÍSICAS Y MATEMÁTICAS DE LA UNIVERSIDAD NACIONAL DE SAN ANTONIO ABAD DEL CUSCO.

SEÑOR DIRECTOR DE LA ESCUELA PROFESIONAL DE MATEMÁTICA.

SEÑORES DOCENTES MIEMBROS DEL JURADO.

En cumplimiento con las normas y reglamentos de Grados y Títulos establecidos por la Escuela Profesional de Matemática, tenemos el honor de presentar a su consideración el presente trabajo de tesis, titulado: **ANÁLISIS DE DESERCIÓN ESTUDIANTIL EN LA ESCUELA PROFESIONAL DE MATEMÁTICA DE LA UNSAAC EN EL AÑO 2015 A 2020 MEDIANTE MODELOS DE SUPERVIVENCIA**, con el fin de Optar el Título profesional de Licenciado en Matemática Mención Estadística.

El presente trabajo de investigación nace de la preocupación por el incremento en las tasas de abandono universitario que afecta a distintas universidades del país, y que también se evidencia en la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco. En este contexto, se propuso utilizar herramientas estadísticas como los modelos de supervivencia paramétrica, no paramétricas y semiparamétricos, para analizar y entender los factores que afectan en la permanencia o deserción de los estudiantes.

Esta tesis espera ser una herramienta de referencia útil para las autoridades universitarias y los investigadores interesados en comprender y afrontar el problema de la deserción, aportando herramientas estadísticas relevantes para la formulación de estrategias que favorezcan la permanencia estudiantil con el objetivo de:

MODELAR MEDIANTE MÉTODOS PARAMÉTRICOS Y NO PARAMÉTRICOS LA SUPERVIVENCIA DE LOS ESTUDIANTES ANTES DE SU DESERCIÓN DE LA ESCUELA PROFESIONAL DE MATEMÁTICA DE LA UNIVERSIDAD NACIONAL DE SAN ANTONIO ABAD DEL CUSCO DEL AÑO 2015 A 2020.

RESUMEN

Este estudio titulado como “Análisis de deserción estudiantil en la escuela profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco en el año 2015 a 2020 mediante modelos de supervivencia” tiene como objetivo modelar mediante métodos paramétricos y no paramétricos la supervivencia de los estudiantes antes de su deserción de la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco del año 2015 a 2020. La investigación se desarrolla bajo un enfoque cuantitativo, con un alcance explicativo. Se asume un diseño no experimental. La población es de 273 estudiantes con una muestra de 163 estudiantes. Se aplica la técnica de documental para registro de información. Los resultados muestran que la probabilidad de supervivencia disminuye significativamente, la probabilidad de supervivencia es alta (98.2% a los 2 meses), pero esta disminuye significativamente con el tiempo, alcanzando un 66.6% a los 14 meses y un 54.6% a los 34 meses y después se estabiliza. Los factores asociados que influyen en el tiempo de deserción de los estudiantes, obtenidos de la regresión de Cox, son la edad, el rendimiento académico, el ingreso familiar y la modalidad de ingreso. Se concluye que, en las primeras semanas, el modelo paramétrico exponencial y el no paramétrico Kaplan-Meier presentan similar ajuste de supervivencia. Después de 5 meses, la curva de Kaplan-Meier presenta un mejor ajuste del tiempo de estudios de los estudiantes de la Escuela Profesional de Matemática de la UNSAAC en el año 2015 a 2020.

Palabras claves: Supervivencia, Estudiantes de matemática, Métodos paramétricos, Métodos no paramétricos.

ABSTRACT

This study, entitled "Analysis of student dropout in the Professional School of Mathematics of the National University of San Antonio Abad del Cusco from 2015 to 2020 using survival models," aims to model, using parametric and non-parametric methods, the survival of students in the Professional School of Mathematics before their dropout from the National University of San Antonio Abad del Cusco from 2015 to 2020. The research is developed under a quantitative approach, with an explanatory scope. A non-experimental design is assumed. The population is 273 students with a sample of 163 students. The documentary technique is applied to record information. The results show that the probability of survival decreases significantly; the probability of survival is high (98.2% at 2 months), but this decreases significantly over time, reaching 66.6% at 14 months and 54.6% at 34 months, and then stabilizes. The associated factors influencing student dropout time, obtained from Cox regression, are age, academic performance, family income, and admission mode. It is concluded that, in the first few weeks, the exponential parametric model and the nonparametric Kaplan-Meier model present similar survival fit. After 5 months, the Kaplan-Meier curve presents a better fit to the study time of students at the Professional School of Mathematics of the National University of San Antonio Abad del Cusco from 2015 to 2020.

Keywords: Survival, Mathematics Students, Parametric and nonparametric methods.

INTRODUCCIÓN

La deserción estudiantil es uno de los problemas que aborda en la mayoría de las Instituciones de Educación Superior de toda Latinoamérica. A través de distintas investigaciones, se da cuenta de un número importante de estudiantes que no logran culminar sus estudios Universitarios, con el consecuente costo social asociado a este fenómeno. Por lo tanto, algunos países han comenzado a diseñar profundos procesos de mejoramiento para aumentar la retención en los primeros años de estudios Universitarios (Díaz, 2008).

Según Vaira, Avila, Ricardi y Bergesio (2010) citado por Paucar y Quispe (2021) la deserción estudiantil Universitaria en nuestros tiempos se ha convertido en un paradigma que conlleva a consecuencias negativas en las Instituciones de Educación Superior (IES) y en general en toda población que ingresa a realizar sus estudios superiores. La deserción estudiantil es de gran complejidad difícil de explicar y describir, debido al gran número de factores asociados que intervienen, razón por la cual se hace difícil la implementación de políticas efectivas que permitan disminuir su impacto en la sociedad. Su existencia origina problemas sociales, institucionales y personales. Desde el punto de vista social se incrementa el desempleo por lo que crece el ciclo de la pobreza, disminuye el aporte intelectual y genera e incrementa el costo de la educación debido a la suboptimización de los recursos. En el contexto institucional genera problemas económicos que afectan a la Universidad, debido a que se invierte recursos en jóvenes que no terminan sus estudios. En lo personal y familiar se percibe la sensación del fracaso, pérdida de oportunidades laborales implicando menor posibilidad de conseguir empleos satisfactorios.

La información sobre las características de los estudiantes de la Escuela Profesional de Matemática de los semestres 2015-I, a 2020-II permitirán realizar este trabajo de investigación, proporcionada por la Oficina de Bienestar Universitario y centro de cómputo UNSAAC.

Este presente trabajo de investigación está estructurado de la siguiente forma:

1) PLANTEAMIENTO DE PROBLEMA: Se presenta la caracterización del problema, formulación del problema (problema general y problemas específicos), objetivos (objetivo general y objetivos específicos), justificación, limitación del trabajo, hipótesis (hipótesis general e hipótesis específica).

2) MARCO TEÓRICO: Se presentan los antecedentes de la investigación, los conceptos básicos de análisis de supervivencia, descripción del método no paramétrico Kaplan Meier para estimar la función de supervivencia, el método semi-paramétrico del modelo de regresión de Cox y la definición de deserción estudiantil.

3) DISEÑO METODOLÓGICO: Se presenta el tipo de investigación, diseño de investigación, población, variables de estudio, técnicas e instrumentos de recolección de datos, procedimiento de recolección de datos, técnicas de procedimiento y análisis de datos.

4) RESULTADOS: Se obtuvieron mediante la aplicación del modelo de Kaplan-Meier, prueba de Log-Rank, el modelo de regresión de Cox y prueba de Wald, donde se realizó su respectivo análisis e interpretación de datos y discusión con los antecedentes. Conclusiones, recomendaciones, referencias y anexos.

CAPÍTULO I

PLANTEAMIENTO DEL PROBLEMA

1.1 Situación Problemática

A nivel internacional, la deserción estudiantil en la educación superior representa uno de los principales retos de los sistemas educativos. Aproximadamente un 30% de los estudiantes universitarios abandonan sus estudios antes de finalizar el primer año. Este fenómeno afecta tanto a países desarrollados como en vías de desarrollo, limitando la formación de capital humano calificado y generando pérdidas económicas para las instituciones educativas y las familias. La deserción universitaria no solo refleja problemáticas estructurales en los modelos educativos, sino también factores sociales, económicos y psicológicos que influyen en la decisión de los estudiantes de abandonar sus estudios (UNESCO, 2021).

En el contexto nacional, este problema no es ajeno. Informes del Ministerio de Educación y la SUNEDU evidencian que un porcentaje considerable de estudiantes de universidades públicas y privadas no concluye su carrera universitaria, siendo las razones más frecuentes: dificultades económicas, bajo rendimiento académico, falta de orientación vocacional, y problemas familiares. Además, se observa que las carreras de ciencias básicas, como Matemática, Física o Química, presentan mayores tasas de deserción en comparación con otras carreras profesionales. Esta situación preocupa a las autoridades educativas, ya que impacta directamente en la formación de profesionales para el desarrollo científico y tecnológico del país (Educacion, 2021, págs. 18 -25).

En el ámbito local, específicamente en la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco (UNSAAC), la deserción estudiantil ha

sido constante durante los últimos años. Entre los años 2015 a 2020, se ha observado una disminución considerable en el número de estudiantes que culminan la carrera, a pesar de los esfuerzos por mejorar los planes de estudio y brindar apoyo académico. No obstante, no se cuenta con un estudio estadístico riguroso que permita identificar los factores asociados a la deserción ni los patrones temporales de abandono. En este sentido, el uso de modelos de supervivencia puede ofrecer una herramienta poderosa para analizar el tiempo hasta la deserción y sus variables asociadas, permitiendo a las autoridades universitarias tomar decisiones informadas y diseñar estrategias efectivas de retención estudiantil.

De acuerdo a los datos proporcionados a través de la dirección de sistemas de información (DSI) y de la Unidad de Estadística – UNSAAC (2016 y 2017), del total de estudiantes que ingresan a la Universidad Nacional de San Antonio Abad del Cusco, el 28.2 % y el 29.69%, respectivamente, optaron por abandonar sus estudios universitarios. Asimismo, se observó que, dentro del número global de estudiantes que ingresaron a la Escuela Profesional de Matemática de la UNSAAC, el 64% y el 68.62%, respectivamente, tomaron la decisión de abandonar sus estudios superiores.

1.2 Formulación del problema

1.2.1 Problema general

¿Cuál es el modelo de supervivencia paramétrico y no paramétrico, que mejor ajuste presenta al tiempo de estudios de los estudiantes antes de su deserción de la Escuela Profesional de Matemática, de la UNSAAC del año 2015 a 2020?

1.2.2 Problemas específicos

- PE.1** ¿Cuál es la función de riesgo del tiempo de estudios de los estudiantes antes de su deserción de la Escuela Profesional de Matemática de la UNSAAC del año 2015 a 2020?
- PE.2** ¿Cuál es la función de supervivencia del tiempo de estudios de los estudiantes antes de su deserción de la Escuela Profesional de Matemática de la UNSAAC del año 2015 a 2020?
- PE.3** ¿Cuáles son los factores asociados del tiempo de estudios de los estudiantes antes de su deserción de la Escuela Profesional de Matemática de la UNSAAC del año 2015 a 2020?

1.3 Justificación

La deserción estudiantil en la educación superior es un problema crítico que afecta la calidad y eficiencia de las universidades, especialmente en carreras como Matemática. En la Escuela Profesional de Matemática de la UNSAAC, se ha evidenciado una disminución en la permanencia estudiantil entre 2015 a 2020.

Este estudio es necesario porque permite analizar mediante modelos de supervivencia, identificando los factores que influyen en la deserción y el tiempo que los estudiantes permanecen en la carrera. Los resultados obtenidos son útiles para proponer estrategias que mejoren la retención estudiantil, fortaleciendo así la calidad educativa y la toma de decisiones a nivel institucional.

1.3.1 Justificación Social

La deserción estudiantil en la escuela profesional estudiada de la UNSAAC tiene implicaciones importantes para la sociedad en general. Por una parte, la deserción estudiantil crea

una disminución económica importante para los estudiantes y sus familias, que han invertido una cantidad alta de dinero y tiempo en educar a sus hijos.

Desde otras perspectivas, este problema afecta de manera negativa a la sociedad en términos de falta de profesionales en matemática y demás áreas relacionadas. La escasez de profesionales influye en la capacidad de la sociedad para solucionar problemas y de la misma manera generar soluciones creativas e innovadoras.

Por ello, es fundamental realizar la investigación de las razones que influyen en la decisión de los estudiantes de la Escuela Profesional de Matemática de la UNSAAC y buscar alternativas para la reducción de la tasa de deserción estudiantil. Esta investigación ayuda al desarrollo de estrategias efectivas para la continuidad de los estudiantes y finalicen con éxito sus estudios superiores.

El estudio es de gran utilidad para otras instituciones de educación superior que enfrentan desafíos parecidos con la deserción de los estudiantes de las diversas carreras técnicas y universitarias, es por ello que esta investigación tiene mayor valor social y ayuda a la creación de políticas, así como estrategias efectivas para contribuir al mejoramiento integral de la educación superior.

1.3.2 Justificación Económica

La deserción estudiantil que se da en la Escuela Profesional de Matemática en la UNSAAC conlleva a repercusiones económicas de importancia para la universidad y para la sociedad en general.

Como punto de inicio, el índice de deserción estudiantil tiene un impacto negativo en los ingresos de la universidad, ya que los estudiantes, al momento de abandonar la carrera profesional, no pagan las matrículas semestrales, lo que estaría ocasionando pérdidas de ingreso para la universidad y, además, limitar su capacidad para ofrecer programas y servicios de calidad.

Por otra parte, la deserción por parte de los estudiantes tiene un impacto económico a largo plazo en la sociedad, ya que esto implica la falta de profesionales en Matemática, así como áreas relacionadas que tendrá efectos perjudiciales para el crecimiento económico y competitividad.

Entonces, es crucial estudiar los factores que afectan en la deserción estudiantil y encontrar estrategias para reducir la deserción estudiantil, de esta manera, se puede crear políticas y estrategias más efectivas para mejorar la retención estudiantil y les ayuden a llegar hasta la graduación.

1.3.3 Justificación Práctica

Tiene un propósito muy claro: ofrecer información útil para que la universidad pueda diseñar mejores políticas y estrategias para que los estudiantes permanezcan en la carrera, superen diferentes dificultades y logren terminar sus estudios.

La universidad puede usar los resultados de esta investigación para comprender cuáles son los factores que conlleva a que los estudiantes quieran desertar de su carrera; con esa información, será posible tomar medidas para enfrentar el problema. Por ejemplo, si se llega a identificar que los estudiantes tienen dificultades para equilibrar los estudios académicos con otras responsabilidades, la universidad puede brindar apoyo en la flexibilidad de horarios.

1.3.4 Justificación teórica

Tiene una justificación teórica importante, ya que de los hallazgos obtenidos en esta investigación podrían ser aplicables a otras profesiones u otras universidades, que podrían ayudar al desarrollo y mejora sobre la retención estudiantil en general.

El uso de modelos para realizar el análisis de la deserción estudiantil contribuye en la educación superior. Esta aplicación de modelos, que podríamos denominar como “modelos de

supervivencia”, ayudará a desarrollar y mejorar la teoría de cómo aplicar estos modelos en la educación superior.

Esta investigación contribuye al desarrollo teórico en el campo de la educación superior y la retención estudiantil, lo que puede tener consecuencias para la práctica y la política educativa.

1.3.5 Justificación metodológica

Tiene una justificación metodológica más sólida, porque se trata de un análisis complejo que exige un enfoque metódico y riguroso.

El uso de los modelos de supervivencia para estudiar la deserción estudiantil en la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco es una metodología adecuada y rigurosa. Esta técnica permite analizar los datos a través de periodos de tiempo conocidos como “datos longitudinales” y toma en cuenta el tiempo que los estudiantes permanecen en el centro educativo antes de abandonarlo.

Estos modelos de supervivencia permiten analizar los factores de la deserción estudiantil, lo que ayuda a la identificación de patrones y tendencias que podrían pasar desapercibidos con otros métodos.

Es fundamental señalar, que la recopilación de datos constituye un proceso importante para la validación de los resultados obtenidos en el estudio sobre la deserción estudiantil, es por ello, se garantiza la precisión y la confiabilidad de los datos recolectados.

1.4 Objetivos de la investigación

1.4.1 Objetivo General

Modelar mediante métodos paramétricos y no paramétricos la supervivencia de los estudiantes antes de su deserción de la Escuela Profesional de Matemática de la UNSAAC del año 2015 a 2020.

1.4.2 Objetivos Específicos

- OE.1** Determinar la función de riesgo del tiempo de estudios de los estudiantes antes de su deserción de la Escuela Profesional de Matemática de la UNSAAC del año 2015 a 2020.
- OE.2** Determinar la función de supervivencia del tiempo de estudios de los estudiantes antes de su deserción de la Escuela Profesional de Matemática de la UNSAAC del año 2015 a 2020.
- OE.3** Determinar los factores asociados del tiempo de estudios de los estudiantes antes de su deserción de la Escuela Profesional de Matemática de la UNSAAC del año 2015 a 2020.

1.5 Delimitación de la investigación

1.5.1 Delimitación espacial

La delimitación espacial del análisis de deserción estudiantil en la Escuela Profesional estudiada de la Universidad Nacional de San Antonio Abad del Cusco mediante modelos de supervivencia se circunscribe a esta institución académica ubicada en la ciudad de Cusco, Perú.

El estudio se enfocó en los estudiantes matriculados en la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco en el período comprendido entre los años 2015 a 2020. No se consideraron otros programas académicos de la Universidad Nacional de San Antonio Abad del Cusco ni otras instituciones educativas en la zona geográfica en cuestión.

1.5.2 Delimitación Temporal

La delimitación temporal del análisis de deserción estudiantil en la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco, mediante modelos de supervivencia, se enfoca en el periodo comprendido entre los años 2015 a 2020. Para Mondragón (2013), esto significa que la investigación considera únicamente los datos correspondientes a los estudiantes matriculados en la Escuela Profesional de Matemática de la Universidad Nacional de

San Antonio Abad del Cusco durante este lapso de tiempo (Mondragón, 2013), no se consideran datos anteriores al año 2015 ni posteriores al año 2020. Esta permite tener información importante acerca de los factores asociados con la deserción por parte de los estudiantes en la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco en un tiempo de cinco años.

1.5.3 Delimitación conceptual

Esta delimitación del análisis de deserción por parte de los estudiantes en la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco a través del modelo de supervivencia se enfoca en el estudio de las variables que pueden afectar en la permanencia o deserción de los estudiantes matriculados en la presente escuela profesional.

Se comprende como deserción estudiantil a la decisión de un estudiante de abandonar sus estudios antes de haber completado su estudio superior. En este trabajo, se consideran las variables asociadas a la permanencia de los estudiantes de la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco, tal y cual como el rendimiento académico, la situación socioeconómica, la motivación el apoyo familiar y las características personales de los estudiantes tales como el rendimiento académico, la situación socioeconómica, la motivación, el apoyo familiar y las características personales de los estudiantes. Se hace uso de modelos de supervivencia para determinar la probabilidad de que un estudiante deje la carrera en un momento en específico en función de estas variables (Mondragón, 2013).

Este tipo de delimitación facilita tener una comprensión más profunda de los factores que influyen en esta decisión en la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco y otorgar información importante para diseñar estrategias efectivas de prevención y retención del estudiante.

CAPÍTULO II

MARCO TEÓRICO

2.2 Bases Teóricas

2.2.1 Análisis de Supervivencia

La disciplina de análisis de supervivencia se enfoca en el estudio de los procesos estocásticos relacionados a eventos como la finalización de la vida de seres vivos, el fallo de sistemas biológicos y sistemas físicos como lo son las máquinas mecánicas o electrónicas.

El análisis de supervivencia es parte de la estadística que integra varias técnicas para estudiar la sincronización de eventos de interés, tiempo para aprender habilidades, detección de enfermedades, fallas de equipos, clientes disminuidos o perdidos, roturas de línea, apagones, etc. (Mondragón, 2013).

El análisis de supervivencia constituye un grupo de métodos estadísticos en los que la variable respuesta es el tiempo transcurrido desde el inicio del seguimiento hasta la ocurrencia del evento de interés (muerte) en los individuos del estudio. A menudo, sucede que las personas se retiran de la encuesta antes del evento, por lo que solo información parcial (censura) sobre variables de interés (Mondragón, 2013). Por ello, la finalidad del análisis de supervivencia es incluir parte de la información provista por la persona que está siendo revisada de las siguientes maneras método desarrollado para este propósito (San José et al., 2009).

El análisis de supervivencia engloba una serie de métodos, técnicas y procedimientos estadísticos que se emplean para analizar información relacionada con la ocurrencia de eventos particulares conocidos como tiempos de supervivencia o duración de vida (Mondragón, 2013).

El objetivo principal del análisis de supervivencia es anticipar la probabilidad de que ocurra el evento de interés para un individuo determinado, como por ejemplo la probabilidad de

supervivencia o de no experimentar el evento (como la muerte) hasta un tiempo específico (Avendaño, 2013).

2.2.1.1 Definiciones de términos utilizados en el análisis de supervivencia

Fecha de inicio: Momento de diagnóstico, comienzo de tratamiento o remisión completa (Mondragón, 2013).

Fecha de último reporte: Día de la información más reciente sobre el caso.

Monitoreo: Implica la supervisión continua de los miembros de un grupo desde el punto de partida, con el fin de determinar su condición vital (vivo o fallecido) (Mondragón, 2013).

Intervalo de observación: El lapso temporal que abarca desde el punto de partida hasta la fecha de finalización del análisis (Mondragón, 2013).

Fecha de culminación del estudio: Fecha establecida por el investigador para concluir el monitoreo de los individuos (Mondragón, 2013).

Período de sobrevida: Intervalo temporal entre las fechas de inicio y la última notificación (Mondragón, 2013).

Registro truncado: Ocurre cuando hay interrupción del seguimiento debido a fallecimiento por causas ajenas, retiro del estudio sin haber experimentado el evento, o ausencia del suceso de interés durante el período de observación (Mondragón, 2013).

Probabilidades de sobrevida: A partir de una serie de lapsos de sobrevida de un conjunto de sujetos (incluyendo los truncados), podemos calcular la fracción de la población que continuará existiendo a lo largo de un periodo determinado, bajo iguales circunstancias (Mondragón, 2013).

En el análisis de supervivencia, el análisis de los datos puede ser realizado utilizando técnicas paramétricas y no paramétricas (Mondragón, 2013).

- a) Métodos Paramétricos: Entre los más comunes, se encuentran las distribuciones exponenciales, Weibull y log-normal, junto con otras distribuciones de probabilidad usadas para modelar el análisis de supervivencia (Mondragón, 2013).
- b) Método no Paramétrico: Entre los métodos no paramétricos, se encuentran las tablas de vida, métodos de Kaplan-Meier (Mondragón, 2013).
- c) Método Semiparamétrico: El método semiparamétrico se refiere a la regresión de Cox, la cual combina elementos paramétricos y no paramétricos, este enfoque se explicará con mayor detalle en secciones posteriores (Tineo, 2005).

2.2.1.2 Elementos Básicos

La variable aleatoria T representa el tiempo hasta que ocurre el evento de interés (muerte, abandono, falla del sistema, etc.). La probabilidad de que el evento ocurra en un tiempo menor o igual a t se denota como $(T \leq t)$ (Tineo, 2005). Además, la función de supervivencia $S(t) = 1 - F(t) = P(T > t)$, también es esencial en el análisis de supervivencia, ya que representa la probabilidad de que el evento ocurra después del tiempo t (Barrera, 2008).

2.2.1.3 Función de supervivencia

La función de supervivencia, representada por $S(t)$, indica la probabilidad de que un individuo sobreviva (no experimente el evento de interés, como muerte o abandono) hasta el tiempo t o más

Definición: Si consideramos una variable aleatoria no negativa T con una función de distribución acumulativa $F(t)$ y una función de densidad de probabilidad $f(t)$, la función de supervivencia $S(t)$ se formula de la siguiente manera :

$$(1)$$

$$\int_t^{\infty} f(u)du$$

dónde: $S(t)$ es una función no creciente, tal que

$$S(t) = 1 \text{ cuando } t \rightarrow 0$$

$$S(t) = 0 \text{ cuando } t \rightarrow \infty$$

En otras palabras, la probabilidad de sobrevivir en el momento inicial es de uno, mientras que la probabilidad de sobrevivir en un tiempo infinitamente largo es cero (García, 2014).

2.2.1.4 Función de riesgo

De acuerdo con Barrera (2008), la función de riesgo $h(t)$ denota el riesgo instantáneo de que un evento suceda en un intervalo extremadamente reducido de tiempo $(t, t + \Delta t)$, considerando que dicho evento aún no ha tenido lugar hasta el instante t . La expresión de esta función es la siguiente:

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t \mid T > t)}{\Delta t} \quad (2)$$

Al aplicar la probabilidad condicional a la expresión (2), se llega a:

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \frac{P(t \leq T < t + \Delta t)}{P(T > t)} \quad (3)$$

Al expandir el cálculo de las probabilidades y considerando que $P(T > t) = S(t)$ en la expresión (3), se llega a:

$$h(t) = \frac{1}{S(t)} \lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} \quad (4)$$

Utilizando la definición de la derivada en la ecuación (4), se llega a:

$$h(t) = \frac{f(t)}{S(t)} \quad (5)$$

Se tiene entonces una expresión que relaciona la función de riesgo y la función de supervivencia.

Aplicando integrales a la ecuación (5), se llega a la función acumulada de riesgo.

$$H(t) = \int_0^t h(s) ds = \int_0^t \frac{f(s)}{S(s)} ds \quad (6)$$

También se puede escribir como:

$H(t) = \int_0^t \frac{f(s)}{1-F(s)} ds$ Sustituyendo $u = 1 - F(s)$ y $du = -f(s) ds$, la expresión (5) se convierte en:

$$H(t) = \int \frac{-du}{u}$$

de donde resulta que:

$$H(t) = -\ln S(t) \quad 6.1$$

$$S(t) = e^{-H(t)} \quad 6.2$$

$$h(t) = -\frac{d}{dt}(\ln S(t)) \quad 6.3$$

$$f(t) = -\frac{d}{dt}S(t) \quad 6.4$$

2.2.1.5 Censura

Los tiempos de supervivencia individuales se censuraron si el evento de interés (deserción o muerte) no se observó durante el periodo de seguimiento.

Definición: Sea T la variable aleatoria que denota el tiempo hasta que un estudiante experimenta el evento de interés, donde:

$$T = \begin{cases} 0, & \text{si } T > t_k \\ 1, & \text{si } T \leq t_k \end{cases}$$

El tiempo de supervivencia se considera como censurado cuando $T > t_k$, donde t_k es el último tiempo de seguimiento, y como no censurado cuando $T \leq t_k$.

Por lo tanto, otra característica importante del análisis de supervivencia es que utiliza parte de la información obtenida de la censura (Puma Laura, 2018).

2.2.1.5.1 Censura por la derecha

Esta forma de censura es el escenario más frecuente de datos incompletos, identificable por eventos que no se manifiestan durante la observación de un individuo. La ausencia de datos puede estar vinculada a diversas causas, entre ellas, que el evento de interés no se materialice antes del término del estudio (mientras el período de seguimiento está acotado), la salida de personas del estudio u otros eventos que impiden la observación del acontecimiento (García, 2014).

2.2.1.5.2 Censura por izquierda

Esta clase de censura ocurre cuando un individuo es observado por primera vez y el evento observable ya ha ocurrido (García, 2014).

2.2.1.5.3 Censura por intervalos

Se refiere a la situación en la que solo se tiene conocimiento de que el individuo experimenta el evento de interés en un rango de tiempo, es decir, entre los momentos t_i y t_k

Los datos de supervivencia se organizan de forma (t_i, δ_i) , donde t_i representa el tiempo de observación, y δ_i toma el valor 0 si la observación está censurada y δ_i toma el valor 1 si ocurre el evento de interés, como la deserción (García, 2014).

2.2.2 Método paramétrico

2.2.2.1 Modelo Exponencial.

Definición. Sea X una variable aleatoria se dice que sigue una distribución exponencial de parámetro λ , con $\lambda > 0$, denotado $X \sim Exp(\lambda)$, una función de densidad viene dada por

$$f(x) = \lambda e^{-\lambda x} \cdot I(x \in [0, \infty >))$$

Esta distribución exponencial describe el tiempo (o la distancia, por ejemplo, en otros contextos)

entre dos sucesos consecutivos que tienen lugar a una tasa constante. Es decir, si el número de sucesos por unidad de tiempo se modela mediante una $Poisson(\lambda)$, entonces el tiempo que pasa entre dos de esos sucesos consecutivos sigue una $Exp(\lambda)$ (Silva Torres, A. 2020).

Observación 1. Se caracteriza también la distribución exponencial a través de su parámetro de escala θ , que es el inverso del parámetro λ de la Definición 1, $\theta = 1/\lambda$. De hecho, el parámetro θ se puede considerar un “parámetro de supervivencia”, pues indica el tiempo medio entre dos sucesos. Este parámetro a veces es llamado parámetro o tasa de decaimiento, o bien, parámetro de intensidad del proceso (Silva Torres, A. 2020).

Para la variable aleatoria tiempo de vida T con $T \sim Exp(1/\theta)$, la función de distribución es:

$$F(t) = \int_0^t \frac{1}{\theta} \exp(-t/\theta) \cdot I(x \in [0, \infty >) dt$$

$$F(t) = \int_0^t \frac{1}{\theta} \exp\left(-\frac{t}{\theta}\right) dt = 1 - \exp(-t/\theta), \text{ si } t \geq 0$$

A partir de esta se obtiene la función de supervivencia, la función de riesgo y la función acumulada.

Para $t \geq 0$ (Silva Torres, A. 2020).

Función de supervivencia:

$$S(t) = 1 - F(t) = \exp(-t/\theta)$$

Función de riesgo:

$$h(t) = \frac{f(t)}{S(t)} = \frac{1}{\theta}$$

Función acumulada:

$$H(t) = \int_0^t h(x) dx = t/\theta$$

Como se observa que la distribución exponencial no es muy flexible, pues la función de riesgo es constante con el tiempo. En términos estadísticos, supone que el riesgo instantáneo de muerte (“justo después de t ”) no depende del factor del individuo (Silva Torres, A. 2020).

2.2.2.2 Modelo Weibull

El modelo Weibull es una generalización de modelo exponencial, se incorpora un parámetro de forma α se mantiene el parámetro de escala $\theta = 1/\lambda$ este modelo es uno de los mas para los tiempos de supervivencia (Silva Torres, A. 2020).

Definición 2. Una variable aleatoria continua X se dice sigue distribución Weibull con parámetros de forma y escala, respectivamente, $\alpha, \theta > 0$, denotado $X \sim Weibull(\alpha, \theta)$, si su función de densidad viene dada por:

$$f_X(x) = \frac{\alpha}{\theta} \left(\frac{x}{\theta}\right)^{\alpha-1} e^{-\left(\frac{x}{\theta}\right)^\alpha} \cdot I(x \in [0, \infty))$$

Para $\alpha = 1$, la expresión (2.1) se reduce a la distribución exponencial.

Para la variable aleatoria tiempo de vida T con $T \sim Weibull(\alpha, \theta)$ la función de supervivencia y de riesgo con, para $t \geq 0$ (Silva Torres, A. 2020).

Función de supervivencia:

$$S(t) = \exp(-(t/\theta)^\alpha)$$

función de riesgo:

$$h(t) = \frac{f(t)}{S(t)} = (\alpha/\theta)(\alpha/\theta)^{\alpha-1}$$

El tiempo de vida esperado y su varianza viene dado por:

$$E(t) = \theta \Gamma(1 + \theta) ; \text{Var}(T) = \theta^2 [\Gamma(1 + 2\theta) - (\Gamma(1 + \theta))^2]$$

Donde Γ es la función Gamma definido por:

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$$

2.2.2.3 Modelo log-Normal

Para el modelo Log-Normal, las funciones de densidad, riesgo y supervivencia vienen dadas por:

Función de supervivencia:

$$f(t) = \frac{1/t}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log t - \mu)^2}{2\sigma^2}\right)$$

Función de riesgo:

$$h(t) = \frac{\frac{1/t}{\sigma\sqrt{2\pi}}}{1 - \Phi\left(\frac{(\log t - \mu)^2}{\sigma}\right)} \exp\left(-\frac{1}{2} \left(\frac{(\log t - \mu)^2}{\sigma}\right)\right)$$

Función acumulada:

$$S(t) = 1 - \Phi\left(\frac{(\log t - \mu)^2}{\sigma}\right)$$

La distribución normal tiene dos parámetros que hereda el modelo Log-Normal: la media, $\mu \in R$, y la desviación estándar, $\sigma > 0$. Estas dos distribuciones se corresponden unívocamente a través de las transformaciones exponencial y logaritmo (Silva Torres, A. 2020).

2.2.3 Método no Paramétrico

2.2.3.1 Estimador de Kaplan-Meier

Es un método no paramétrico para estimar la función de supervivencia $S(t)$, basado en muestras aleatorias independientes con tiempos de falla censurados basado en los principios de supervivencia condicional. El método de construcción de estimaciones no es paramétrico porque no asume ninguna estructura para la función de distribución de probabilidad de por vida.

El propósito del estimador es analizar la función de supervivencia Kaplan-Meier que utilizará toda la información disponible de casos revisados y no revisados en un momento dado (Barrera, 2008).

Esta estimación se deriva de múltiples productos obtenidos de una estimación de la probabilidad

condicional de supervivencia basado en el número observado de casos en riesgo n_i y la "muerte" d_i se observa en un instante de tiempo dado y se calcula como $\left(\frac{n_i - d_i}{n_i}\right)$, con n_i que representa el número de casos en riesgo y d_i representa el número de los que han desertado.

Definición. Sea $t_1 < t_2 < \dots < t_k$ los k diferentes tiempos de falla observados en la muestra el tamaño es aleatorio n , la posibilidad de cometer más de un error t_k está tolerado, entonces, denotamos el número de fallas en este período de tiempo por $d_i (d_i \geq 1)$ (Barrera, 2008). Para t fijo, el estimador de supervivencia de Kaplan-Meier $\hat{S}(t)$ se calcula de la siguiente manera (Barrera, 2008):

$$\hat{S}(t) = \prod_{t:t_i < t} \left(1 - \frac{d_i}{n_i}\right) = \prod_{t:t_i < t} \frac{n_i - d_i}{n_i} \quad (7)$$

Donde:

d_i : Número de desertores observados en el instante $t_i, i = 1, 2, \dots, k$

n_i : Número de estudiantes observados en riesgo en el instante t_k , que corresponde a individuos que no fallaron y no fueron censurados hasta el instante anterior a t_k .

Considerando la ecuación 7, se puede notar que cuando $t = 0$, $\hat{S}(t) = 1$; esto indica que todos los individuos comienzan vivos en el estudio (Puma Laura, 2018).

2.2.3.2 Varianza del Estimador Kaplan-Meier

La varianza del estimador de la función de supervivencia es muy importante, por lo tanto, este estudio utiliza el método delta basado en la extensión de términos en series de Taylor del primer orden (Barrera, 2008).

Método Delta

El propósito del método delta es utilizar la aproximación de la serie de Taylor para obtener funciones lineales que se aproximen a funciones más complejas, como los estimadores de funciones de supervivencia (Barrera, 2008).

Supongamos que $f(x)$ es una función de densidad de probabilidad de una variable aleatoria X . La expansión en series de Taylor de primer orden alrededor de la media es (Barrera, 2008).

$$f(x) \cong f(u) + (x - u)f'(u)$$

Aplicando propiedades de varianza a la función (x) .

$$Var(f(x)) \cong [f'(u)]^2 Var(x - u)$$

$$Var(f(x)) \cong [f'(u)]^2 \sigma^2$$

Donde σ^2 es la varianza de la variable aleatoria X . A modo de ejemplo, se ilustra la aplicación de las propiedades de varianza para las siguientes funciones

Función logaritmo natural:

$$\ln(X) \cong \ln(u) + (X - u) \frac{1}{u}$$

$$\widehat{Var}(\ln(x)) \cong \frac{1}{\hat{u}^2} \hat{\sigma}^2$$

Función exponencial:

$$\exp(X) \cong \exp(u) + (X - u)\exp(u)$$

$$\widehat{Var}(\exp(X)) \cong \hat{\sigma}^2 (\exp(\hat{u}))^2$$

Varianza del estimador Kaplan - Meier

Usaremos el método Delta para calcular la varianza del estimador de Kaplan-Meier. Tomando el logaritmo de ambos lados de la ecuación (7) (Puma Laura, 2018).

$$\log (\hat{S}(t)) = \sum_{t:t_i < t} \log \frac{n_i - d_i}{n_i} \quad (8)$$

Realizando $\hat{p}_i = \left(\frac{n_i - d_i}{n_i}\right)$, se tiene:

$$\log (\hat{S}(t)) = \sum_{t:t_i < t} \log (\hat{p}_i) \quad (9)$$

Aplicando en una expansión de series de Taylor el $\log (\hat{p}_i)$ donde $\hat{p}_i = p_i$, por tanto:

$$\log (\hat{p}_i) = \log (p_i) + \frac{d}{dp_i}(\log (p_i)) (\hat{p}_i - p_i) + \frac{1}{2} \cdot \frac{d^2}{dp_i^2}(\log (p_i)) (\hat{p}_i - p_i)^2 + R(\hat{p}_i, p_i)$$

$$\log (\hat{p}_i) = \frac{d}{dp_i}(\log (p_i)) (\hat{p}_i - p_i), \text{ donde: } i = 1, 2, 3, 4, \dots, n$$

Aplicando las propiedades de varianzas a ambos miembros

$$\text{Var}(\log (\hat{p}_i)) \cong \left[\frac{d}{dp_i}(\log (p_i)) \right]^2 \text{Var}(\hat{p}_i - p_i)$$

$$\text{Var}(\log (\hat{p}_i)) = \left[\frac{d}{dp_i}(\log (p_i)) \right]^2 [\text{Var}(\hat{p}_i) + \text{Var}(p_i)]$$

Aplicando la varianzas de un constante es cero, es decir:

$$\text{Var}(\log (\hat{p}_i)) \cong \text{Var}(\hat{p}_i) \left[\frac{d}{dp_i}(\log (p_i)) \right]^2, \text{ donde: } i = 1, 2, 3, 4, \dots, n \quad (10)$$

Asumiendo que $n_i \hat{p}_i$ tiene una distribución binomial con parámetros n_i y p_i es decir:

$$n_i \hat{p}_i \sim \text{Binomial}(n_i, p_i), i = 1, 2, 3, 4, \dots, n$$

Por consiguiente;

$$\text{Var}(n_i \hat{p}_i) = n_i p_i (1 - p_i) \rightarrow n_i^2 \text{Var}(\hat{p}_i) = n_i p_i (1 - p_i)$$

Simplificando n_i

$$\text{Var}(\hat{p}_i) = \frac{p_i(1 - p_i)}{n_i}, \text{ donde: } i = 1, 2, 3, 4, \dots, n \quad (11)$$

Reemplazando la ecuación (11) en (10)

$$\text{Var}(\log(\hat{p}_i)) \cong \frac{p_i(1-p_i)}{n_i} \times \frac{1}{p_i^2}, \text{ con: } i = 1, 2, 3, 4, \dots, n$$

Reduciendo p_i

$$\text{Var}(\log(\hat{p}_i)) = \frac{(1-p_i)}{n_i p_i}, \text{ con: } i = 1, 2, 3, 4, \dots, n$$

Suponiendo que $\log \hat{p}_i$ es independiente para $i = 1, 2, 3, \dots, n$. En la ecuación (9), aplicamos la varianza.

$$\text{Var}(\log(\hat{S}(t))) = \sum_{i:t_i < t} \text{Var}(\log(\hat{p}_i))$$

$$\text{Var}(\log(\hat{S}(t))) = \sum_{i:t_i < t} \left[\frac{(1-\hat{p}_i)}{n_i \hat{p}_i} \right]$$

Se reemplaza $\hat{p}_i = \left(\frac{n_i - d_i}{n_i} \right)$

$$\text{Var}(\log(\hat{S}(t))) = \sum_{i:t_i < t} \left[\frac{\left(1 - \left(\frac{n_i - d_i}{n_i} \right) \right)}{n_i \left(\frac{n_i - d_i}{n_i} \right)} \right] = \sum_{i:t_i < t} \left[\frac{\frac{d_i}{n_i}}{n_i - d_i} \right]$$

$$\text{Var}(\log(\hat{S}(t))) = \sum_{i:t_i < t} \left[\frac{d_i}{n_i(n_i - d_i)} \right] \quad (12)$$

Empleando la serie de Taylor a $\log \hat{S}(t)$ alrededor de $\hat{S}(t) = S(t)$ se tiene:

$$\log \hat{S}(t) = \log S(t) + \frac{d}{dS(t)} (\log S(t)) (\hat{S}(t) - S(t)) + R(\hat{S}(t), S(t)).$$

De donde:

$$\log \hat{S}(t) \cong \frac{d}{dS(t)} (\log S(t)) [\hat{S}(t) - S(t)],$$

Aplicando derivada respecto a $S(t)$

$$\log \hat{S}(t) = \frac{1}{S(t)} (\hat{S}(t) - S(t))$$

Empleando las propiedades de varianza

$$\text{Var}(\log \hat{S}(t)) \cong \text{Var}\left(\frac{1}{S(t)} (\hat{S}(t) - S(t))\right) = \frac{1}{S^2(t)} \text{var}(\hat{S}(t) - S(t))$$

Aplicando varianza de un constante, es igual a cero.

$$\text{Var}(\log \hat{S}(t)) = \frac{1}{S^2(t)} \text{var}(\hat{S}(t))$$

Reemplazando este resultado en la ecuación (12) se tiene:

$$\frac{1}{S^2(t)} \text{Var}(\hat{S}(t)) \cong \sum_{i:t_i < t} \left[\frac{d_i}{n_i(n_i - d_i)} \right]$$

Finalmente hallamos la varianza del estimador de Kaplan-Meier

$$\text{Var}(\hat{S}(t)) \cong \text{Var}(\hat{S}^2(t)) \sum_{i:t_i < t} \left[\frac{d_i}{n_i(n_i - d_i)} \right] \quad (13)$$

De ello el error estándar de $S(t)$ es la raíz cuadrada de la varianza.

$$e.e(\hat{S}(t)) \cong \hat{S}(t) \left(\sum_{i:t_i < t} \left[\frac{d_i}{n_i(n_i - d_i)} \right] \right)^{\frac{1}{2}} \quad (14)$$

Se calcula el intervalo de confianza $S(t)$ al 95% definido de la siguiente manera:

$$S(t) \pm 1.96 e.e(\hat{S}(t))$$

ello nos permite observar los intervalos de confianza para $S(t)$.

Para hallar la mediana para el tiempo de supervivencia $S(t)$

$$\bar{S}(t) = \frac{\sum_{i=1}^{10} t_i}{n_i} \cong S(t_{\text{mediana}}) = 0.50 \quad (15)$$

2.2.3.3 Comparación funciones de supervivencia

Las comparaciones gráficas de las funciones de supervivencia son importantes e ilustrativas, pero a veces tales comparaciones pueden volverse muy subjetivas y, en algunos casos, difíciles de realizar. Para abordar esta deficiencia, se utilizan pruebas estadísticas para determinar si las diferencias observadas entre las funciones de supervivencia son significativas, como la prueba de Log-Rank no paramétrica (Barrera, 2008).

Tabla 1

Prueba de igualdad de la función de supervivencia de dos grupos en el tiempo de observación t_i .

<i>Evento</i>	<i>Grupo 1</i>	<i>Grupo 0</i>	<i>Total</i>
<i>Muerte</i>	d_{1i}	d_{0i}	d_i
<i>No muerte</i>	$n_{1i} - d_{1i}$	$n_{0i} - d_{0i}$	$n_i - d_i$
<i>En riesgo</i>	n_{1i}	n_{0i}	n_i

Una estadística de prueba es la suma de las diferencias observadas entre eventos y las cantidades esperadas de muertes en todos los períodos de supervivencia observados, dividida por las diferencias en el número de muertes en uno de los dos grupos, como, por ejemplo, el grupo 1, el número estimado de muertes es aproximadamente :

$$\hat{e}_{1i} = \frac{n_{1i}d_i}{n_i} \quad (16)$$

Y para la varianza suponiendo una distribución hipergeométrica:

$$\hat{v}_{1i} = \frac{n_{1i}n_{0i}d_i(n_i - d_i)}{n_i^2(n_i - 1)} \quad (17)$$

La estadística del test de Log-Rank para verificar la hipótesis de igualdad entre las dos funciones de supervivencia se calcula de la siguiente manera (Barrera, 2008):

$$Q = \frac{[\sum_{i=1}^K (d_{1i} - \hat{e}_{1i})]^2}{\sum_{i=1}^K \hat{v}_{1i}} \quad (18)$$

donde la distribución de Q sigue una distribución chi-cuadrado con un grado de libertad ($p = P(\chi^2(1)) \geq Q$). Si hay más de dos grupos, se utiliza de manera similar a lo que se ha explicado previamente para dos grupos. La siguiente tabla es una tabla relacionada con las diversas funciones básicas de supervivencia cálculo de estadísticas de prueba de igualdad funcional.

Tabla 2

Funciones básicas de supervivencia cálculo de estadísticas de prueba de igualdad funcional

Evento	1	2	...	M	...	M	Total
Muerte	d_{1i}	d_{2i}	...	d_{mi}	...	d_{2i}	d_{2i}
No muerte	$n_{1i} - d_{1i}$	$n_{2i} - d_{2i}$...	$n_{mi} - d_{mi}$...	$n_{Mi} - M_{2i}$	$n_i - d_i$
En riesgo	n_{1i}	n_{2i}	...	n_{mi}	...	n_{Mi}	n_i

En primer lugar, el número esperado de muertes en cada grupo debe estimarse de acuerdo con los siguientes supuestos de igualdad de funciones de supervivencia, a saber:

$$\hat{e}_{ki} = \frac{n_{ki}d_i}{n_i}, \text{ donde: } i = 1, \dots, m \quad (19)$$

Hay más grupos, el cálculo ya no es la varianza, sino la matriz de covarianza vector d_i . Para obtener la matriz de covarianza, se debe suponer que las cantidades observados los eventos siguen una distribución hipergeométrica multivariante central y enlaces diagonales la matriz representada por \hat{V}_i se calcula de la siguiente manera:

$$\hat{v}_{mmi} = \frac{n_{mi}(n_i - n_{mi})d_i(n_i - d_i)}{n_i^2(n_i - 1)}, m = 1, 2, \dots, m - 1 \quad (20)$$

Los elementos fuera de la diagonal se calculan como:

$$\hat{v}_{lmi} = \frac{n_{mi}(n_i - n_{mi})d_i(n_i - d_i)}{n_i^2(n_i - 1)}, l, m = 1, 2, \dots, M - 1; l \neq m \quad (21)$$

Con los estimadores del número de muertes y la varianza se construye el siguiente estadístico:

$$Q = \left[\sum_{i=1}^K W_i(d_i - \hat{e}_i) \right] \left[\sum_{i=1}^K W_i \hat{V}_i W_i \right]^{-1} \left[\sum_{i=1}^K W_i(d_i - \hat{e}_i) \right], \quad (22)$$

La ecuación involucra W_i como matriz diagonal con pesos w_i , d_i como vector con cantidades de muertes por grupo t_i , y \hat{e}_i como vector con muertes esperadas en el tiempo t_i . Similar al caso de dos grupos, en múltiples grupos, Q sigue distribución chi-cuadrado con $M - 1$ grados de libertad ($p = P(\chi^2(1)) \geq Q$), bajo hipótesis nula de igualdad de supervivencia (Barrera, 2008).

2.2.3.3.1 Prueba de hipótesis para la función de supervivencia

La prueba de hipótesis es un procedimiento que nos permite evaluar hasta que grado el conjunto de los datos observados es compatible con una hipótesis (Puma Laura, 2018).

Formulación de hipótesis.

Hipótesis nula: $H_0: S_1(t) = S_2(t) = \dots = S_M(t)$

Hipótesis alterna: $H_1: S_1(t) \neq S_2(t) \neq \dots \neq S_M(t)$

Interpretación:

Si $p - value$ ($p \leq 0.05$) se rechaza H_0 , se concluye que existe diferencia significativa

Si $p - value$ ($p > 0.05$) no se rechaza H_0 , se concluye que no existe diferencia significativa.

2.2.3.3.2 Prueba de Log- Rank

El test de log-rank (Mantel-Haenszel) es un método que posibilita la comparación de las curvas de supervivencia entre grupos compuestos por dos o más individuos.

Supongamos que se quiere comparar la supervivencia en dos grupos, muerte y no muerte que se dispone de una muestra de cada grupo de tamaño n_{1i} (*Varon*) y n_{0i} (*Mujer*) para cada uno (Puma Laura, 2018).

$$n_{1i} + n_{0i}$$

El tamaño de la muestra combinada y, $t_{(1)} < t_{(2)} < t_{(3)} < \dots < t_{(k)}$ los k tiempos de muertes diferentes observados en la muestra conjunta ordenada en forma creciente; además, por d_{1i} y d_{0i} el número de fallos ocurridos en un instante de tiempo t_i en cada grupo y por d_i el número total de fallos en ese instante, es decir (Barrera, 2008):

$$d_i = d_{1i} + d_{0i}$$

Del mismo modo, se denota por $n_{1i} + n_{0i}$, número de personas en riesgo en cada grupo antes del tiempo t_i y por n_i la suma de:

$$n_{1i} + n_{0i}$$

La mayoría de la información de las situaciones, de pruebas estadísticas se fundamentan en tablas de contingencia agrupadas por estado para cada período de supervivencia (Puma Laura, 2018).

GRUPO	Nº. de fallos	Nº. de individuos vivos en $t_{(i)}$	Nº. de individuos en riesgo en $t_{(i)}$
<i>1</i>	d_{1i}	$n_{1i} - d_{1i}$	n_{1i}
<i>0</i>	d_{0i}	$n_{0i} - d_{0i}$	n_{0i}
<i>total</i>	d_i	$n_i - d_i$	n_i

Para probar la función de supervivencia entre grupos, se hacen las siguientes suposiciones:

H_0 : la supervivencia de los grupos es la misma.

H_1 : la supervivencia de los grupos es diferente.

2.2.4 Método Semiparamétrico

2.2.4.1 Modelo de Regresión de Cox

El modelo de regresión de Cox (1972), también conocido como modelo de riesgos proporcionales de Cox, se emplea para analizar la relación entre un conjunto fijo de variables explicativas llamadas covariables X_1, X_2, \dots, X_n y el tiempo de respuesta o supervivencia (Puma Laura, 2018).

En realidad, se utiliza la función de riesgo $h(t; X)$, que representa la tasa instantánea del evento de interés

En los modelos de regresión de Cox (1972), la función de la tasa de riesgo de tiempo de ocurrencia del evento, para un modelo con un vector de covariables representado por x , se define como:

$$h(t; X) = h_0(t) \exp(\beta^T X) \quad (23)$$

Donde t es el tiempo de vida, X es el vector de covariables, $h_0(t)$ es la función de riesgo base; y $\beta^T = (\beta_1, \beta_2, \dots, \beta_n)$ es el vector de parámetros de la regresión.

El modelo de regresión de Cox $h(t, X)$, mostrado en la expresión (23), es un modelo semi-paramétrico porque tiene una parte paramétrica y una parte no paramétrica (Puma Laura, 2018).

- a) *La parte paramétrica del modelo es $\exp(\beta^T X)$, llamada función de riesgo relativo, donde indicar claramente la influencia relativa de los parámetros de regresión estimados sobre el riesgo (Puma Laura, 2018).*
- b) *La parte no paramétrica del modelo es $h_0 t$, la función riesgo base, que es una función arbitraria y no especificada (Puma Laura, 2018).*

Además, como se mencionó anteriormente, el modelo de regresión de Cox, también conocido como el modelo de riesgos proporcionales de Cox, evidencia que la relación es constante

entre el riesgo de dos individuos con el mismo vector de covarianza, e independiente del tiempo, es decir (Barrera, 2008):

$$\frac{h(t; X_i)}{h(t; X_j)} = \frac{h_o(t) \exp(\beta^t X_i)}{h_o(t) \exp(\beta^t X_j)} = \exp[\beta^t (X_i - X_j)] = Cte. \quad (24)$$

Para $i, j = 1, 2, \dots, n$ y $i \neq j$, no dependen del tiempo de fallo t , razón dada en la ecuación (23) es constante para cada valor de t .

El cociente expresado en la ecuación (24), es conocida como razón de riesgo relativo instantáneo en el tiempo t (Puma Laura, 2018).

Observación: La componente no paramétrica absorbe la constante β_0 , por lo que no aparece en la componente paramétrica $\exp(X\beta_j^T)$ (Puma Laura, 2018).

Si una persona tiene 4 veces el riesgo de otra persona al comienzo del estudio, entonces es el mismo en cualquier momento del estudio, lo que implica que la tasa de error es proporcional, una suposición subyacente al modelo de riesgos proporcionales de Cox.

En la regresión lineal, la media de la variable de respuesta está relacionada linealmente con el factor pronóstico, mientras que, en el modelo de regresión de Cox, el logaritmo de la razón de la función de riesgo está relacionado linealmente con el factor pronóstico (covariable).

Reemplazando la expresión (6) en la expresión (6.2), se tiene que

$$S(t) = \exp \left\{ - \int_0^t h(t) dt \right\}, \forall t \geq 0.$$

Luego, sabemos que la función de supervivencia

$$S(t/X) = \exp \left(- \int_0^t h(t, X) dt \right).$$

Entonces, la función de supervivencia base se expresa de la siguiente manera

$$S_o(t) = \exp\left(-\int_0^t h_o(t)dt\right).$$

Donde:

$$\int_0^t h_o(t)dt = H_o(t) \rightarrow S_o(t) = \exp(-H(t)) \quad (25)$$

$H(t)$ Es conocido como el riesgo base acumulado, de modo que

$$S(t/X) = \exp\left(-\int_0^t h_o(t)dt \exp(X\beta^T)\right) = \exp(-H_o(t))\exp(X\beta^T)$$

$$S(t/X) = (S_o(t))^{\exp(X\beta^T)} \quad (26)$$

Este resultado nos permite encontrar la función de densidad $f(t/X)$, puesto que:

$$h(t/X) = \frac{f(t/X)}{S(t/X)} \rightarrow f(t/X) = h(t/X) \cdot S(t/X) = h_o(t/X) \exp(X\beta^T) \cdot S(t/X)$$

$$= h_o(t/X) \exp(X\beta^T) \cdot (S_o(t))^{\exp(X\beta^T)}$$

Donde $S_o(t)$ es la función de supervivencia base, y al comparar su gráfica con la gráfica de $(S_o(t))^{\exp(X\beta^T)}$, estas curvas son aproximadamente paralelas. Por lo tanto, se cumple la suposición de proporcionalidad de riesgos en el modelo de Cox (Puma Laura, 2018).

2.2.4.2 Estimación de los parámetros en el modelo de Cox

El método más prevalente para estimar los parámetros de un modelo de regresión es la máxima verosimilitud, que implica estimar los parámetros que maximizan el logaritmo natural de la verosimilitud (Barrera, 2008). Sin embargo, en el caso de los modelos de Cox, esta técnica no es viable debido a la falta de conocimiento de las funciones de riesgo subyacentes que están presentes en estos modelos (Barrera, 2008).

En 1972, Cox desarrolló un método que permite estimar los parámetros de regresión sin asumir una forma particular de la función de riesgo subyacente (Barrera, 2008), proponiendo un

término llamado "*función de verosimilitud parcial*" que depende solo del parámetro. interés La función de probabilidad parcial se escribe de la siguiente manera.

$$L(\beta) = \prod_{i=1}^K \frac{e^{x_{(i)}\beta}}{\sum_{j \in R(t_i)} e^{x_j \beta}} \quad (27)$$

2.2.4.2.1 Estimación de β

Una característica de los modelos de Cox es que los coeficientes β deben estimarse a partir de las observaciones. La presencia de componentes no paramétricos invalida el uso de la máxima verosimilitud. Para estimar el vector de parámetros β , Cox propone máxima verosimilitud parcial, donde la función de verosimilitud viene dada por $L(\beta)$ en la expresión (27) porque tiene en cuenta la probabilidad de individuos censurados. La función $L(\beta)$ consiste en una muestra de n individuos y existen k ($k \leq n$).

A veces, errores obvios t_1, t_2, \dots, t_k . Una manera fácil de entender la funcionalidad la probabilidad parcial es tener en cuenta los siguientes parámetros condicionales:

P [estudiante falla en t_i del grupo de falla en t_i]

$$\begin{aligned} &= \frac{P(\text{estudiante falla en } t_i / \text{estaba vivo antes})}{P(\text{un fallo en } t_i \text{ del grupo en riesgo/estaba vivo antes})} \\ &= \frac{h(t_i/X_j)}{\sum_{j \in R(t_i)} h^{(t_i/X_j)}} \\ &= \frac{h_o(t_i) \exp(X_i \beta^T)}{\sum_{j \in R(t_i)} h_o(t_i) \exp(X_j \beta^T)} \\ &= \frac{\exp(X_i \beta^T)}{\sum_{j \in R(t_i)} \exp(X_j \beta^T)} \quad (28) \end{aligned}$$

Donde $t_i = \{j: t_j > t_i\}$ se refiere al conjunto de individuos en riesgo antes de t_i , es decir, aquellos que están siendo observados y aún no han experimentado el evento (Barrera, 2008)

Ahora, la función de verosimilitud para β consiste en el producto de varias probabilidades (Puma Laura, 2018), abarcando todas las instancias de fallo $t_i, i = 1, 2, \dots, n$

Esto es;

$$L(\beta) = \prod_{i=1}^n \left[\frac{\exp(X_i \beta^T)}{\sum_{j \in R(t_i)} \exp(X_j \beta^T)} \right]^{\delta_i} \quad (29)$$

Donde

$$\delta = \begin{cases} 1, & \text{si } t_i \text{ no es censurado} & i = 1, 2, 3, \dots, n \\ 0, & \text{si } t_i \text{ es censurado} & i = 1, 2, 3, \dots, n \end{cases}$$

Seguidamente, aplicamos logaritmo en ambos lados de la ecuación (29) tendremos:

$$\begin{aligned} \ln[L(\beta)] &= \sum_{i=1}^n \delta_i \ln \left[\frac{\exp(X_i \beta^T)}{\sum_{j \in R(t_i)} \exp(X_j \beta^T)} \right] \\ &= \sum_{i=1}^n \left[\delta_i \ln \left(\exp(X_i \beta^T) - \sum_{j \in R(t_i)} \exp(X_j \beta^T) \right) \right] \\ l(\beta) = \ln[L(\beta)] &= \sum_{i=1}^n \delta_i \left[\left((X_i \beta^T) - \ln \left[\sum_{j \in R(t_i)} \exp(X_j \beta^T) \right] \right) \right] \end{aligned} \quad (30)$$

Derivando la ecuación (30) respecto a los parámetros β^T tenemos:

$$\frac{\partial l(\beta)}{\partial \beta_k} = \sum_{i=1}^n \delta_i \left[X_{ik} - \frac{\sum_{j \in R(t_i)} X_{jk} \exp(X_j \beta^T)}{\sum_{j \in R(t_i)} \exp(X_j \beta^T)} \right], k = 1, 2, \dots, p.$$

Haciendo:

$$\frac{\partial l(\beta)}{\partial \beta_k} = 0$$

Obtenemos el siguiente sistema de ecuaciones por (ponderaciones).

$$\sum_{i=1}^n \delta_1 \left[X_{ik} - \frac{\sum_{j \in R(t_i)} X_{ik} \exp(X_j \beta^T)}{\sum_{j \in R(t_i)} \exp(X_j \beta^T)} \right], = 0, k = 1, 2, 3, \dots, p. \quad (31)$$

A partir de este sistema de ecuaciones, podemos obtener los valores estimados de $\hat{\beta}$. (Puma Laura, 2018).

2.2.4.3 Estimación de la función de riesgo base $h_0(t)$

Para estimar la función de riesgo, suponga que $\hat{h}_0(t)$ es una constante entre los tiempos de supervivencia no hay revisión (Puma Laura, 2018). Las vidas sin censura primero deben ponerse en forma ascendente (Puma Laura, 2018):

$$y_{(1)} \leq y_{(2)} \leq \dots \leq y_{(k)}$$

y, el conjunto de riesgo $R_{(j)}$ asociado a estos tiempos de supervivencia ordenados es (Puma Laura, 2018):

$$R_{(j)} = \{T_i: T_i \geq y_j\}, \quad j = 1, 2, \dots, k.$$

Sea $\hat{h}_{(0)}, \hat{h}_{(1)}, \dots$ constantes, es decir:

$$\hat{h}_{(0)}(t) = \begin{cases} \hat{h}_{(0)}; & 0 \leq t < y_{(1)} \\ \hat{h}_{(1)}; & y_{(1)} \leq t < y_{(2)} \\ \dots & \dots \end{cases}$$

Queremos estimar $\hat{h}_{(0)}$, el individuo del conjunto de riesgos en $y_{(2)}$ que está en el grupo $R_{(2)}$. Si un individuo falla en $y_{(2)}$ dado $(y_{(2)}, R_{(2)})$ entonces:

$$\begin{aligned} 1 &= \sum_{j \in R_{(2)}} P[\text{el } j - \text{ésimo estudiante falla en } y_{(2)}/y_{(2)}, R_{(2)}] \\ &= \sum_{j \in R_{(2)}} (y_{(3)} - y_{(2)}) \hat{h}_{(2)} \exp(X_j \beta^T), \text{ puesto que } h(t) \Delta(t) \end{aligned}$$

$$= p[t \leq t < t + \Delta t / T \geq t] = (y_{(3)} - y_{(2)}) \hat{h}_{(2)} \sum_{j \in R_{(2)}} \exp(X_j \beta^T)$$

Así, el riesgo de probabilidad entre $y_{(2)}$ y $y_{(3)}$ es:

$$(y_{(3)} - y_{(2)}) \hat{h}_{(2)} = \frac{1}{\sum_{j \in R_{(2)}} \exp(X_j \beta^T)} \quad (32)$$

Ahora usamos $\hat{\beta}$ (obtenido mediante el método de máxima verosimilitud parcial) para obtener:

$$\hat{h}_{(2)} = \frac{1}{(y_{(3)} - y_{(2)}) \sum_{j \in R_{(2)}} \exp(X_j \hat{\beta}^T)} = \hat{h}_{(0)}(t) \quad (33)$$

Con lo cual queda estimado el riesgo base de $h_0(t)$

Por otro lado, este resultado nos permite obtener, también, la estimación de la función de riesgo acumulado $\hat{h}_0(t)$ (Barturén & Niño, 2018).

2.2.4.4 Prueba de Wald

La prueba de Wald se calcula como la relación entre el coeficiente calculado y la desviación estándar de este coeficiente de la siguiente manera:

$$z = \frac{\hat{\beta}_j}{\sqrt{\text{Var} \hat{\beta}_j}} \quad (34)$$

Bajo la hipótesis nula de que el coeficiente es cero, el estadístico de Wald sigue una distribución normal estándar. Para obtener un intervalo de confianza para el coeficiente $\hat{\beta}_j$, tenemos:

$$\hat{\beta}_j \pm z_{\alpha/2} \sqrt{\text{Var} \hat{\beta}_j} \quad (35)$$

Si necesitamos probar $H_0: \beta = \beta_0$ contra $H_1: \beta \neq \beta_0$, usamos la prueba de Wald, que bajo la hipótesis nula sigue una distribución chi-cuadrado con p grados de libertad. (Puma Laura, 2018)

2.2.5 DESERCIÓN ESTUDIANTIL

2.2.5.1 Definición de Deserción

De acuerdo con Páramo y Correa (1999), los estudiantes que abandonan los estudios se deben a una variedad de razones o razones y, finalmente, abandonan los estudios, una ruptura con el molde académico de que la sociedad quiere que se gradúen felices de la universidad.

Al respecto, Cruz y Ortega (2008), si un estudiante se da de baja de un programa académico por dos semestres consecutivos, se matricula en otro programa, pero no continúa afiliado al primer programa, o completa su programa de estudios, pero no recibe un título por más de dos semestres después de completar el estudio.

La deserción estudiantil se refiere a la salida permanente de una institución de educación superior dentro de un período de tiempo específico (Barrera, 2008). El desarrollo del abandono es todo un proceso, a veces es lento, crece y se consolida en los alumnos, y el alumno demuestra en el último abandono si es bueno o malo para el alumno y su entorno (Barrera, 2008).

2.2.5.2 Clases y actores de Deserción

De acuerdo con Páramo y Correa (1999), se identifican diversas categorías de deserción en educación, que pueden coexistir:

- Deserción total: abandono definitivo de la educación individual (Barrera, 2008).
- Deserción discriminada por causas: clasificada según la razón de la decisión
- Deserción por facultad (escuela o departamento): cambio entre diferentes facultades (Barrera, 2008).
- Deserción por programa: cambio de programa dentro de una misma facultad (Barrera, 2008)..

- Deserción en el primer semestre de la carrera: debido a una adaptación inadecuada a la vida universitaria (Barrera, 2008).
- Deserción acumulada: suma de abandonos en una institución. Los involucrados en el fenómeno de la deserción son (Barrera, 2008).
- Desertores: estudiantes que abandonan su educación universitaria (Barrera, 2008)
- Padres de familia de desertores.
- Antiguos compañeros de clase: estudiantes que estuvieron en el mismo semestre que el desertor antes de su retiro (Barrera, 2008).
- Docentes: aquellos que estuvieron involucrados en el proceso educativo del estudiante hasta el momento de su retiro (Barrera, 2008).
- Directivas y administradores académicos

2.2.5.3 Factores asociados

Cruz y Ortega (2008) definen los siguientes factores asociados a la deserción.

2.2.5.3.1 Factores individuales

Algunos factores son edad, género, ocurrencia de desastres o problema hogar, fácil integración social de los individuos, expectativas personales insatisfechas (Barrera, 2008), en:

2.2.5.3.2 Factores socioeconómicos

Los aspectos socioeconómicos abarcan el nivel socioeconómico del estudiante, su situación ocupacional, la situación laboral de los padres, la dependencia económica, la existencia de apoyo familiar, la educación de los padres, el contexto familiar (tamaño familiar, estado civil de los padres, etc.).

2.2.5.3.3 Factores académicos

Los aspectos académicos comprenden: acceso a orientación profesional previa a la admisión, procedencia de la escuela secundaria (pública o privada), rendimiento académico individual, enfoques de estudio empleados, acceso a exámenes de admisión, carga académica (créditos por semestre), descontento con el plan de estudios, entre otros (Barrera, 2008).

2.2.5.3.4 Factores institucionales

Los factores institucionales (una categoría más relevante para las instituciones) son: reglamentos académicos, recursos disponibles para las instituciones educativas, relaciones personales con el personal y otros estudiantes, la calidad de los cursos en los que se matriculan, anti-bullying o acción afirmativa en las actividades educativas, instituciones, etc.

2.3 Marco conceptual

Deserción: se contextualiza como la acción de renunciar o abandonar algo, especialmente una obligación o un compromiso. En términos militares, este se contextualiza como abandonar el servicio o no cumplir el tiempo específico establecido. En lo académico, esta se entiende como una falta de compromiso o una renuncia a las responsabilidades que se ha asumido.

Parámetro: este viene a ser un valor o una variable que se utiliza en una función, en una ecuación matemática o en una declaración de programa, el cual puede cambiar y ajustar para cambiar el comportamiento o el resultado de la función, ecuación o programas estos dan más facilidad a los programadores o usuarios personalizar el comportamiento de una función o programa sin tener que cambiar directamente el código de fuente. En el ámbito estadístico, los parámetros se refieren a los valores desconocidos que se usan para describir una población, como la medida, la desviación estándar o la proporción. De manera resumida, un parámetro es una

variable que se utiliza para definir o controlar el comportamiento la salida de una función o programa.

Tiempo: este viene a ser una medida física y abstracta que se usa para calcular cuánto es la duración o la secuencia de eventos o procesos. En términos físicos, el tiempo se define como una dimensión que ayuda a ordenar los eventos en forma secuencial es una de eventos o procesos ordenar los eventos en una secuencia lineal, desde el pasado, a través del presente y hacia el futuro. En el contexto de la física, el tiempo se mide en unidades como segundos, minutos, horas, días, semanas, meses y años. En el lenguaje rutinario, el tiempo se usa para describir la duración de las actividades a la hora en la ocurren, y se pueden calcular con relojes y otros dispositivos de medición de tiempo. Es más, el tiempo puede ser percibido a la duración de eventos. En concreto, el tiempo es una medida abstracta y física que se usa para describir la duración y la secuencia de eventos, así como procesos.

Supervivencia: esta se entiende como la capacidad de mantenerse con vida funcionando en condiciones complicadas y con obstáculos. En lo biológico, la supervivencia se entiende como la capacidad de un organismo para mantenerse con vida y reproducirse en ambiente, aguantando factores estresantes como por ejemplo los climatológicos, la comida insuficiente o a los que lo intentan depredar. Desde el punto de vista humano, esta se puede entender como la capacidad de una persona para continuar con vida frente a enfermedades, un accidente o una situación de emergencia. En lo empresarial, este se considera como la capacidad que tiene la empresa para seguir en el mercado empresarial frente a sus competidores. De manera general, la supervivencia se refiere a la capacidad de un organismo, sistema o entidad para resistir y superar condiciones desafiantes, así como seguir funcionando.

Retención: esta se comprende como la acción de mantener algo o alguien en su lugar actual, en lo fiscal, este viene a ser el impuesto que se paga por los ingresos de una persona o empresa el cual es pagado directamente al estado. Desde el punto de vista empresarial, este es la capacidad que tiene la entidad para retener clientes proveedores y colaboradores clave; en lo legal, se puede determinar como la acción de mantener bajo custodia o detención; en lo académico, consiste en la capacidad de la institución para retener a sus estudiantes ayudarles a cumplir con el objetivo final que es graduarse; para finalizar, la retención se entiende como la capacidad de mantener a algo o a alguien en su lugar o posición actual, ya sea en el ámbito fiscal, empresarial, legal o académico.

2.4 Antecedentes de la investigación

2.4.1 Antecedentes internacionales

García et al. (2022), en su artículo de investigación titulado, *Deserción estudiantil en Colombia: un análisis de los factores predictores y sus indicadores*, llegaron a las siguientes conclusiones:

II-2006. Un análisis de la cohorte de 2018 que da seguimiento a las tasas específicas de deserción mostró que el 51,7% de los estudiantes de la Universidad de la Costa CUC desertaron y el 9,6% se graduaron, mientras que el 38,7% restante de la población estudiantil total permaneció activo en la institución.

En cuanto al género de los estudiantes, se puede observar que el sexo femenino presenta la mayor tasa de deserción con un 52% frente al 48% de los varones, lo que indica que el 80% de los estudiantes de la institución pertenecen a los niveles 2 y 3, con la mayor tasa de deserción. en el semestre II - III-IV, el 39% (70,9% del total de abandonos del año).

Los programas académicos asociados con las tasas de deserción más altas son ingeniería de sistemas y psicología con un 63 %, seguidos de derecho con un 59 % y administración de empresas con un 57 %. Le sigue la construcción civil con un 34% y la ingeniería ambiental con un 37%.

En este sentido, se proponen cuatro factores que inciden en el fenómeno de la deserción: socioeconómico, personal, institucional y académico; entre ellos, los factores socioeconómicos representan el 53%, seguido de los factores personales el 25%, seguido de los factores académicos el 17% y, finalmente, los factores institucionales el 5%.

En cuanto al análisis factorial y multivariado, se puede observar que las variables que inciden positivamente en el fenómeno de la deserción son bajo nivel educativo de los padres, dependencia económica de padres de escasos recursos, junto con el desarrollo del empleo activo, métodos de los docentes y finanzas pasivas (crédito a corto plazo).

En cuanto al análisis de sobrevivencia, se puede observar que entre las carreras que ofrece la institución que presentan una mayor tasa de sobrevivencia al llegar al 10° semestre se encuentra la carrera de Ingeniería Electrónica, la cual tiene un nivel de gestión acumulativo. Aprobar un 99% de contabilidad seguido de un 80% y, finalmente, ingeniería eléctrica un 90%.

Sin embargo, la carrera de derecho fue la menos satisfactoria en términos de tasa de supervivencia, ya que, a partir del segundo semestre, la tasa de supervivencia de esta carrera fue disminuyendo en cada semestre subsiguiente. En el sexto semestre, las mujeres no pueden participar. Por lo tanto, en el décimo período, la tasa de supervivencia acumulada era solo del 45%, lo que demuestra que el sexo femenino tenía menos probabilidades de permanecer, mientras que el sexo masculino se comportaba mejor. Juntos, a través de un análisis global de los datos examinados, ilustran la escala del problema y resaltan los enormes desafíos que la agencia ha

enfrentado para mantener los programas que se han entregado a las comunidades durante más de 43 años. Vale la pena señalar que, si bien las políticas persistentes han logrado buenos resultados, es de suma importancia profundizar en el estudio del fenómeno de la deserción y evaluar las estrategias implementadas para aumentar la duración de los diversos programas académicos.

Medina y Romero (2020), en su artículo de investigación titulado *Modelo de supervivencia para analizar el abandono escolar de la Uraccan-Nueva Guinea*, llegaron a las siguientes conclusiones:

Las principales características de los participantes de 2013 son el 61,7% desertaron, el 71,76% eran de Nueva Guinea urbana, el origen no tiene nada que ver con la deserción, la profesión con mayor tasa de deserción es la administración de empresas con 26,17% y la ocupación con menor tasa de retención es Zootecnia con 20%.

Factores que influyeron en la interrupción: 58,4% de las deserciones fueron en la primera evaluación, 60,6% mujeres, 77,37% solteras, 69,35% tenían entre 15 y 22 años, 56,20% estaban ocupadas, 75,18% había profesiones urbanas, zootécnicas y contables. de la zona urbana el 43,80% pertenece al régimen sabatino. El índice de error promedio está entre 2,33 y 3,27, el índice de error promedio de la URACCAN es de 44,41%, el 84,67% se encuentra en colegios públicos, y, por último, el cuarto y quinto años. 74.93% a 76.98%, respectivamente, es un grado aprobado

Las conclusiones del método de supervivencia utilizado son que un alumno tiene un 37% de probabilidad de abandonar el primer año, un 38% de probabilidad de finalizar el programa y un 5% de probabilidad de abandonar el último año. tasa más baja de abandono de negocios. Las profesiones son Administración, Contabilidad, Deportes, Construcción e inglés y las ocupaciones escogidas que no afectaron la deserción fueron género, antecedentes y estilo de vida, estado civil,

edad, promedio de cuarto y quinto grado, trabajo, tipo de escuela (pública o privada), y cómo cursó el bachillerato.

Sobre el Plan de Retención, URACCAN no cuenta con un plan de retención de estudiantes, los motivos de la deserción son: falta de formación profesional, falta de trabajo y falta de motivación, la universidad adopta un programa orientado a pasar la supervisión, la supervisión racial, cultural y deportiva.

Paz (2022), en su artículo de investigación titulado *Aplicación De Técnicas De Análisis De Supervivencia Al Estudio De La Deserción Estudiantil En Una Carrera De Ingeniería Civil*, llegó a las siguientes conclusiones:

Aplicación de tres métodos de análisis de supervivencia a la deserción en investigación carreras de ingeniería FACET-UNT Enfoque en los resultados del curso primer año del juego. Como confirma Panaia (2019), la rendición es un proceso los momentos pasados se pueden identificar como punteros. Si la institución para monitorear completamente el futuro de los estudiantes puede predecir este comportamiento, tome precauciones para evitar o minimizar el proceso de abandono, es el propósito de este trabajo.

Hubo una diferencia significativa en el tiempo de supervivencia entre hombres (5,03 años) y mujeres (2,63 años). En un estudio Paz y Abdala (2022), se observó que las mujeres experimentan violencia, aunque leve, en mayor medida que los hombres. Si bien esta no es obviamente la única causa, puede indicar síntomas que pueden explicar parte de esa diferencia. Para esclarecer todos los factores que influyen en este debate, se deberían realizar estudios más específicos sobre este tema. Este es un factor que necesita ser explorado más profundamente para determinar la razón de esta discrepancia.

Analizando las variables que crean el riesgo de deserción, aplicando el modelo semiparamétrico de Cox a las asignaturas de primer semestre, se puede establecer que el hecho de que las asignaturas no se formalicen en el primer año de carrera aumenta significativamente el riesgo de deserción.

Cabe destacar que este aumento del riesgo no fue el mismo para todos los sujetos. Álgebra y Geometría Analítica (23,3%), y Cálculo I (18,3%) fueron los más altos mayor valor del riesgo de abandono.

Si se incluyen en el análisis todas las materias de primer año, el mayor riesgo se desplaza a Cálculo II y Física II. Esto se debe a que, en este caso, las asignaturas de segundo semestre se incluyeron como nueva covariable en el cálculo de la función de riesgo propuesta por el modelo. Calcular el valor para el modelo de Cox maximiza el logaritmo de la llamada "función de verosimilitud parcial", que solo considera la probabilidad del número de errores, no la probabilidad de errores. Datos censurados. Sin embargo, de acuerdo con Cox (1972), al calcular la probabilidad de los tiempos de falla, se tienen en cuenta todos los sujetos de riesgo (censurados o no a posteriori) al inicio de los diferentes tiempos de falla. En este caso, hasta el final del segundo tema para asignaturas de 1er semestre que suspendan (bajen), el valor de riesgo se trasladará a las asignaturas de 2do semestre, ya que la probabilidad temporal de los datos censurados no se incluye en el cálculo.

Estos resultados son consistentes con Harvey et al. (2006). En cierto sentido, los resultados preliminares del primer año (Q1) predicen mucho mejores notas finales de primer año y persistencia. El resultado también se adjunta a él. El primer año indica una mayor conservación; el mayor riesgo de abandono de materias del primer cuatrimestre (álgebra y geometría analítica, física I y Cálculo I). En tales casos, las agencias deben actuar.

2.4.2 Antecedentes nacionales

Lopez (2018), en su trabajo de investigación titulado *Estrategia Para La Prevención Del Riesgo De Deserción En Los Estudiantes En Condiciones De Pobreza De La Universidad Técnica De Manabí – Ecuador*, de la Universidad Nacional Mayor de San Marcos, Lima- Perú, para optar por el grado académico de Doctora en Educación, tiene como objetivo principal: prevenir, a través de una estrategia, el riesgo de deserción en los estudiantes bajo condiciones de pobreza en la Universidad Técnica de Manabí, y llegó a las siguientes conclusiones (Barrera, 2008):

Las estrategias de prevención de riesgos probadas experimentalmente para la deserción escolar de los estudiantes pobres son efectivas esencialmente en un intento de reducir significativamente el número de abandonos entre los estudiantes desfavorecidos de la Universidad Técnica de Manabí, para lograr convertir una debilidad en una fortaleza y cambiar las amenazas en oportunidades. Esto afecta en gran medida la efectividad de los títulos y la creación de condiciones óptimas para la actividad académica de los estudiantes, donde la prevención de riesgos es la principal herramienta para evitar la deserción cuando se trata de estudiantes con riesgo de abandono (Barrera, 2008), académicamente, cuando la situación todavía tiene una solución. Si bien la estrategia se aplica a la Universidad Técnica de Manabí, los resultados, también, se pueden aplicar a otras instituciones de educación superior del país (Barrera, 2008). Por lo tanto, tiene un valor teórico, metodológico y práctico importante.

En los fundamentos teóricos, acerca de la problemática en contextos de pobreza, es evidente que las teorías nuevas muestran limitaciones importantes y no consideran la prevención de la deserción y su intrigante relación con el riesgo. Por estas razones, este estudio se basó en la teoría de la deserción y subrayó la necesidad de adoptar un enfoque cauteloso hacia la deserción y

el riesgo, introduciendo conceptos clave como el riesgo de deserción y los factores de riesgo de deserción, los cuales aportan un valioso valor metodológico.

Utilizando la herramienta, el análisis estadístico de la interpretación de los datos obtenidos, se pueden probar las variables explicativas en el comportamiento en torno a la profesión, 224 estudiantes en situación de pobreza, en un ambiente de frustración, donde existen diversas debilidades y amenazas. Se deben mostrar los tipos: Formas de deserción Factores de riesgo, Factores académicos, familiares y económicos. Se puede demostrar que el factor más importante en el proceso de deserción es el tipo académico. Estos factores de riesgo permiten estrategias que convierten las debilidades en fortalezas y las amenazas en oportunidades, al tiempo que minimizan las oportunidades de riesgo al abordar el riesgo de abandono escolar, cuando se combina con los impactos socioeconómicos, las esperanzas y los sueños se desvanecen un aprendiz rápido

Las investigaciones muestran que los factores académicos, especialmente los relacionados con el rendimiento académico, son los más influyentes el mayor riesgo de abandono está estrechamente relacionado con otras variables: el entorno familiar y las condiciones económicas, que las refuerzan y vulneran a los estudiantes, lo que provoca el abandono.

La estrategia para prevenir el riesgo de deserción de los estudiantes pobres fue aprobada inicialmente con criterio de expertos o método Delphi, donde 32 expertos consideraron que tenía un impacto significativo en la comunidad estudiantil, y los expertos y su criterio ayudaron a refinar, afinar y presentar un juicio sobre la eficacia. El próximo momento de vigencia de esta estrategia es enviarla a un experimento piloto donde su correlación e impacto en la sociedad al reducir significativamente el riesgo de abandono escolar para los estudiantes desfavorecidos que tienen más probabilidades de abandonar los estudios.

Saldivar (2020), en su trabajo de investigación titulado *Factores Protectores Y De Riesgo De La Deserción De Estudiantes Del Segundo Semestre De La Escuela De Administración Industrial De Una Institución De Educación Superior*, de la Universidad Cayetano Heredia para optar por el título profesional Maestro En Docencia Profesional Tecnológica, tiene como objetivo principal (Barrera, 2008); identificar los factores de protección y riesgo de la deserción de estudiantes del segundo semestre de la Escuela de Administración Industrial de una Institución de Educación Superior en el periodo 2018-II, y llego a las siguientes conclusiones:

La motivación y la comunicación familiar son dos aspectos sociales que pueden llegar a ser transgresores en el ámbito familiar afecta significativamente el riesgo de deserción del estudiante, porque a los padres no les importa la orientación de sus hijos la educación superior, porque no basta con que el entorno del hogar proporcione recursos para el desarrollo de la personalidad del estudiante o de la institución educativa como aspecto protector. También, es necesario el seguimiento y la participación a largo plazo del entorno inmediato del hogar y la mejora de la comunicación, basado en los objetivos del estudiante acosador no me dejes rendirme.

En el ámbito económico, el desempleo por la falta de oportunidades laborales y el bajo nivel de ingresos en el ámbito familiar en nuestro país ha provocado que los estudiantes creen que si completan el sistema de educación superior tendrán las mismas oportunidades laborales que si no lo hubieran completado, lo que crea un riesgo de deserción. Por ello, frente a esta situación, como muestran nuestros resultados, el apoyo a las familias y el trabajo temporal no relacionado con su educación parece ser moderado, apoyar al alumno para que continúe con su proyecto de convertirse en un buen profesional.

La visión y la motivación externa son conductas de aspecto personal que, como muestran nuestros resultados, se convierten en factores de riesgo debido a que los estudiantes no tienen una

visión de futuro cercano, lo que, combinado con un proyecto de vida, los desalienta a culminar sus estudios. En el segundo comportamiento, el proceso la razón de la motivación es que algunos estudiantes tienen un bajo nivel de educación de los padres y ningún apoyo familiar, como resultado insatisfacción juvenil y deserción estudiantil. Sin embargo, esto puede cambiar si factores protectores como la motivación intrínseca permiten que los estudiantes conozcan su programa de formación profesional y se sientan capaces de completarlo para lograr sus metas y deseos. Otro factor protector es la misión de que los estudiantes desarrollen las habilidades para encajar en el mundo del trabajo si el ambiente es favorable para la realización de sus metas y deseos.

Si se cuenta con el apoyo de las familias, los docentes y las instituciones educativas, los estudiantes ausentes pueden ser detectados tempranamente e incluidos en el proceso educativo, motivarlos en su camino educativo porque lo que hagamos juntos afectará de alguna manera la satisfacción profesional y personal de la próxima generación de jóvenes se espera que la investigación proporciona pruebas a la dirección modo y le permite comparar los resultados desde que el fenómeno la basura afecta a cada estudiante de manera diferente.

Sifuentes (2018), en su trabajo de investigación titulado, *Modelo Predictivos De La Deserción Estudiantil En Una Universidad Privada Del Perú*, de la Universidad Nacional de San Marcos, para optar por el título profesional de Doctor en Ingeniería Industrial, tiene como objetivo principal determinar cómo el uso de modelos predictivos en asignaturas críticas contribuye a identificar, al inicio del curso, a los estudiantes en riesgo de deserción.

Si bien el ausentismo estudiantil es efectivamente un problema potencial en las universidades públicas y privadas de todo el Perú, existen mecanismos o estrategias que pueden ayudar a prevenirlo, como lo demuestra este estudio al elaborar y desarrollar modelos predictivos

basados en las características de los estudiantes que permitan la detección temprana de la juventud, personas que tienen mayor riesgo de reprobación de cursos, lo que más tarde puede llevarlos a retirarse de la vida universitaria.

Armado con este pronóstico, una vez que se cumplan los siete modelos de pronóstico al inicio del semestre, la institución de educación superior puede tomar las medidas que considere adecuadas para ayudar a los estudiantes. En este caso, la universidad privada en estudio implementó un ciclo de formación académica de 16 semanas durante el semestre, enfocándose en los estudiantes con menor probabilidad de aprobar las materias.

Fue una misión, y el esfuerzo rindió frutos, coordinando con diversas áreas de apoyo a los nuevos estudiantes, como oficinas de consejería y orientación, así como orientación vocacional y educativa en carreras. Al final del semestre, se logró reducir entre un 40% y un 50% el número de rechazos de las asignaturas involucradas en el estudio, en comparación con años anteriores, sin utilizar modelos predictivos.

Los siete modelos de predicción son efectivos porque cada modelo está diseñado concretamente para cada área, ya que las particularidades de las disciplinas y las necesidades de los estudiantes difieren, y los factores analizados en el historial académico fueron proporcionados por los registradores de las universidades privadas participantes. en los estudiantes del estudio.

En cuanto a las variables consideradas en el desarrollo del modelo predictivo, tal como se describe en la sección de Métodos, la variable que mejor predijo la impopularidad del curso fue la ocupación estudiada por los jóvenes, es decir, empleo, número de admitidos en la materia y calificaciones en Matemáticas o comunicación al finalizar el quinto año de secundaria, ya que los estudios obtenidos en estas materias son la base para iniciar estudios en una universidad.

2.4.3 Antecedentes locales

Paucar y Quispe (2021), en su trabajo de investigación titulado *Análisis De Supervivencia: Deserción Estudiantil En La Escuela Profesional De Matemática De La Universidad Nacional De San Antonio Abad Del Cusco*, de la Universidad San Antonio Abad del Cusco, con el propósito fundamental de obtener el título académico de Licenciado en Matemática con mención en Estadística, tienen como objetivo central el desarrollo del análisis de supervivencia para abordar la problemática de la deserción estudiantil (Barrera, 2008). Tras llevar a cabo el análisis y el estudio de los datos, se arribó a las siguientes conclusiones:

Del conjunto de estudiantes matriculados en los semestres 2010-I, 2010-II, 2011-I, 2011-II, 2012-I y 2012-II en la Facultad de Matemática de la Universidad Nacional San Antonio Abad de Cusco, el 61,8% logró completar exitosamente el período de estudios de 10 semestres, mientras que el 38,2% no lo lograron.

Los estudiantes de la Escuela Profesional de Matemática optan por utilizar la censura por la derecha, ya que se tiene información sobre el inicio de sus estudios (semestre de matrícula y admisión), pero el final de su período de estudio es desconocido (ya que no todos completan exactamente 10 semestres o más).

Las variables género, edad, ingreso del hogar y ocupación del jefe de hogar (con un valor de $p < 0,05$) mostraron una asociación significativa con la desnutrición crónica.

Aquellos con mayor riesgo de abandonar u optar por no participar en la educación superior, como hombres menores de 22 años, se refieren a estudiantes que son cabeza de familia profesional, ingreso económico familiar, salario base y salario medio.

CAPÍTULO III

HIPÓTESIS Y VARIABLES

3.1 Hipótesis

3.1.1 Hipótesis general

El modelo no paramétrico Kaplan-Meier es el que mejor ajuste presenta al tiempo de estudios de los estudiantes antes de su deserción de la Escuela Profesional de Matemática de la UNSAAC del año 2015 a 2020.

3.1.2 Hipótesis específicas

HE1: La función de riesgo del tiempo de estudios de los estudiantes antes de su deserción de la Escuela Profesional de Matemática de la UNSAAC del año 2015 a 2020 es cambiante de manera ascendente con respecto a cada año.

HE2: La función de supervivencia del período de estudio de los estudiantes antes de su deserción de la Escuela Profesional de Matemática de la UNSAAC entre 2015 a 2020 se modela mediante la función de Kaplan-Meier.

HE3: Los factores asociados vinculados con la duración de los estudios de los estudiantes antes de su deserción de la Escuela Profesional de Matemática de la UNSAAC entre 2015 a 2020 incluyen la carga académica, limitaciones económicas y una orientación vocacional deficiente.

3.2 Identificación de variables

3.2.1 Variable dependiente

- Deserción estudiantil

3.2.2 Variable independiente

Modelos de supervivencia

Factor

- Edad
- Sexo
- Rendimiento académico
- Pensión mensual
- Ingreso familiar
- Modalidad de ingreso
- Tipo de colegio
- Lugar de procedencia
- Pensión mensual
- Depende del mismo
- Ocupación del jefe del hogar
- Casa que ocupa

3.3 Operacionalización de variable

VARIABLE	DEFINICIÓN OPERACIONAL	DIMENSIONES	INDICADORES
Variable dependiente Deserción estudiantil	Deserción estudiantil La deserción estudiantil se define como el escenario en el cual un estudiante abandona su proceso educativo antes de concluir su formación académica prevista, La deserción puede manifestarse en todos los niveles educativos, desde la educación primaria hasta la educación superior, La deserción estudiantil puede ser causada por diferentes factores, como problemas financieros, falta de motivación, dificultades académicas, problemas personales, según (MINEDU, 2019).	<ul style="list-style-type: none"> • Socioeconómica • Académica • Personal 	<ul style="list-style-type: none"> • Tasa de Deserción • Tiempo de Permanencia • Rendimiento Académico • Ingreso Familiar • Modalidad de ingreso

VARIABLE	DEFINICIÓN OPERACIONAL	DIMENSIONES	INDICADORES
Variable independiente Modelos de supervivencia	Modelos de supervivencia Los modelos de supervivencia, también conocidos como modelos de análisis de supervivencia o modelos de tiempo hasta el evento, son una clase de modelos estadísticos utilizados para analizar datos de eventos que tienen una duración o un tiempo de ocurrencia Estos modelos se utilizan comúnmente en ciencias de la salud, ciencias sociales, economía, ingeniería y otros campos donde es importante analizar el tiempo hasta la ocurrencia de un evento según, (Barrera, 2008).	<ul style="list-style-type: none"> • Función de riesgo • Función de supervivencia • Factores asociados 	<ul style="list-style-type: none"> • Orientación vocacional • Método paramétrico y no paramétrico • Estimador de Kaplan-Meier

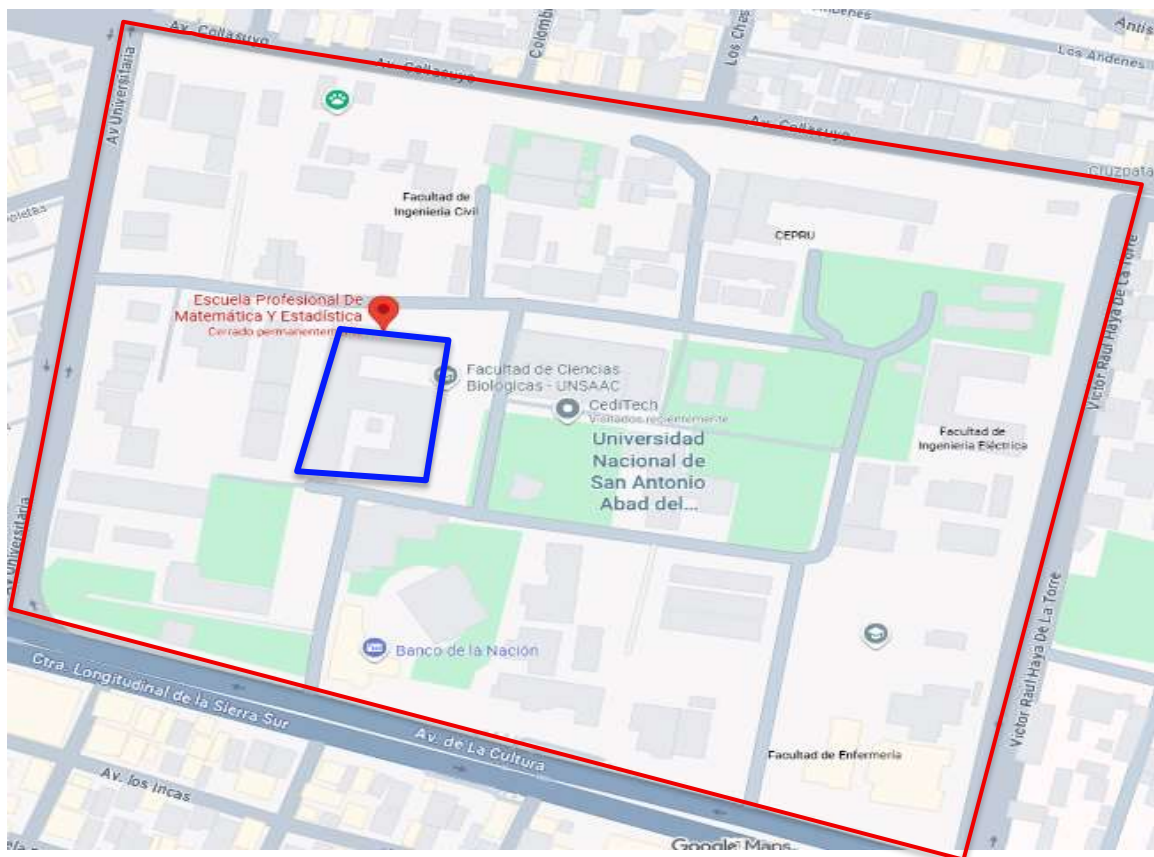
CAPÍTULO IV

METODOLOGÍA

4.1 Ámbito de estudio, localización y geografía

La investigación se desarrolla en el ámbito geográfico correspondiente al Distrito de Cusco, Provincia de Cusco y Departamento de Cusco.

Universidad Nacional San Antonio Abad del cusco se encuentra ubicada en la avenida la cultura N° 773 según las coordenadas $13^{\circ}31'18''S$ $71^{\circ}57'31''O$ / -13.521666666667, -71.958611111111



4.2 Tipo y nivel de investigación

Esta investigación emplea un enfoque cuantitativo, ya que implica la recopilación y el análisis de datos numéricos con el propósito de responder a las preguntas de investigación planteadas y poner a prueba las hipótesis formuladas (Barrera, 2008).

El tipo de la investigación es básica, en razón de que se basa en los conocimientos sobre el tema y las teorías relacionadas a las variables de estudio.

Carrasco (2009) indica que: Este estudio se destaca por contar con objetivos prácticos, ya que no busca una aplicación inmediata. Su propósito fundamental es mejorar las bases teóricas científicas de un análisis exhaustivo y minucioso.

El estudio posee un nivel explicativo, ya que va más allá de la mera descripción de la deserción estudiantil. Su propósito es identificar los factores que están vinculados a la manifestación de dicho evento.

4.3 Diseño de Investigación

El diseño de la investigación es no experimental, longitudinal, retrospectivo, ya que los individuos que pertenecen al estudio son observados en el tiempo hasta que ocurra el evento de interés (deserción) tal como se da en su contexto natural sin manipularlos deliberadamente de los factores académicos y socioeconómicos que asocian a este evento, para analizar y explicar los cambios a través del tiempo en los estudiantes que ingresaron en el periodo 2015 a 2020.

4.4 Unidad de análisis

La unidad de análisis está constituida por los estudiantes matriculados en la escuela profesional de Matemática de la UNSAAC, entre los años 2015 a 2020. Cada caso representa las características y atributos específicos de un estudiante individual.

4.5 Población de estudio

Para el propósito de la tesis la población está comprendida por los 273 estudiantes de la Escuela profesional de Matemática UNSAAC, 2015 al 2020.

Distribuidos en los semestres 2015-I, 2015-II, 2016-I, 2016-II, 2017-I, 2017-II, 2018-I, 2018-II, 2019-I, 2019-II, 2020-I, 2020-II, de estudio. En la siguiente tabla se especifica la población estudiantil:

Tabla 3

Población estudiantil de la facultad de Escuela profesional de Matemática.

Semestre de Ingreso	Número de estudiantes
2015-I	37
2015-II	10
2016-I	36
2016-II	15
2017-I	34
2017-II	23
2018-I	45
2018-II	12
2019-I	26
2019-II	07
2020-I	22
2020-II	06
TOTAL	273

Fuente: centro de cómputo UNSAAC 2015 a 2020.

Criterios de inclusión:

- En el presente trabajo de investigación se incluye o se considera en la población a estudiantes que se han matriculado desde el primer semestre inmediatamente después de haber ingresado a la Escuela Profesional de Matemática.
- Estudiantes con información completa requerida para el estudio.

Criterios de exclusión

- Se excluye de la población a estudiantes que ingresaron y no se matricularon inmediatamente, y no se tiene la información necesaria para el estudio por eso son excluidos del estudio.

4.6 Tamaño de muestra

La muestra es un subgrupo de una población o universo, sobre la cual se recolectaron los datos pertinentes, y fue representativa de dicha población (Hernández & Mendoza, 2018).

Para la determinación de la muestra de estudio se aplicó la técnica de muestreo no probabilístico por conveniencia, en este caso la muestra se determinó por conveniencia de los investigadores, por lo tanto, se eligió a 163 estudiantes de los semestres 2015-I al 2020-II de la Escuela profesional de Matemática, UNSAAC.

Tabla 4

Muestra del estudio

Semestre de Ingreso	Número de estudiantes
2015-I	19
2015-II	10
2016-I	18
2016-II	12
2017-I	21
2017-II	11
2018-I	23
2018-II	06
2019-I	15
2019-II	07
2020-I	15
2020-II	06
TOTAL	163

Fuente: centro de cómputo UNSAAC 2015 a 2020.

4.7 Técnicas de selección de muestra

Para determinar la muestra, se utilizó un método de muestreo no probabilístico intencional / juicio, considerando el procedimiento de selección orientado por las características y contexto de la investigación más que por criterio estadístico de generalización. (Hernández & Mendoza, 2018).

4.8 Técnicas de recolección de información

Técnica: en esta investigación, se empleó un enfoque de análisis documental que se basará en el análisis de las fichas socioeconómicas completadas por los estudiantes al momento de su ingreso a la UNSAAC y el rendimiento académico.

El instrumento: se emplearon fichas documentales para la recolección de datos; se registró información relevante sobre los estudiantes matriculados, el rendimiento académico y las fichas socioeconómicas; para obtener dichos datos, se realizó el acceso al Centro de Cómputo y a la Unidad de Asistencia Social.

4.9 Técnicas de análisis e interpretación de la información

Los datos se organizan en una tabla compuesta de Excel relacionada con las variables explicativas utilizadas en el estudio una vez que se hayan recopilado para este proyecto de investigación.

A medida que avanzamos, continuamos nuestro trabajo en el paquete estadístico R Studio. En primer lugar, estimamos la función de supervivencia mediante el estimador no paramétrico de Kaplan-Meier ajustado a cada variable independiente en función de la variable temporal, y luego comparamos las curvas de supervivencia (Barrera, 2008).

Luego comparamos los modelos de supervivencia paramétrico exponencial y log-logística, para después compara los modelos paramétricos y no paramétricos

Por último, procedemos a estimar los parámetros utilizando el modelo paramétrico y semiparamétrico (modelo de regresión de Cox) con el fin de determinar las variables independientes que presentan una contribución significativa en la toma de decisión de abandonar, desertar o retirarse de la Universidad (Barrera, 2008).

RESULTADOS

ESTIMACIÓN NO PARAMÉTRICA DE LA FUNCIÓN DE SUPERVIVENCIA

Tabla 5

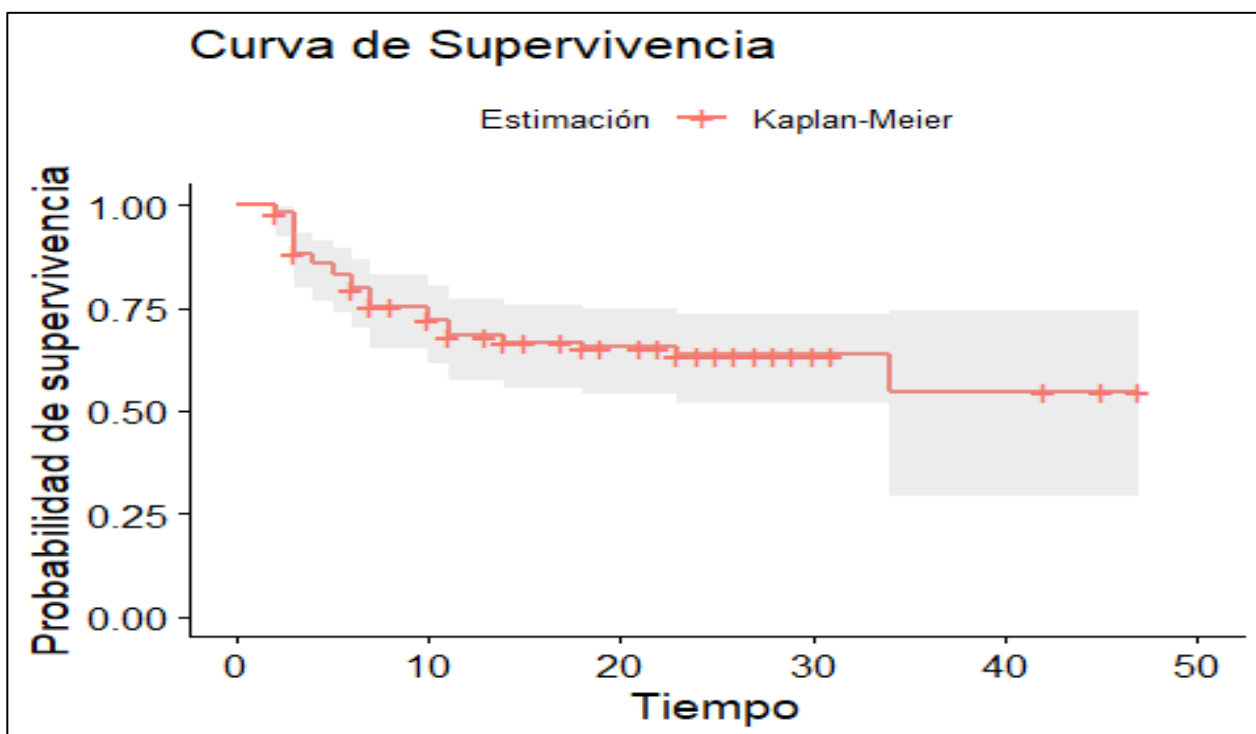
Análisis de Supervivencia

Tiempo (Meses)	N° en riesgo (n.risk)	N° de eventos (n.event)	Supervivencia estimada	Error estándar (std.err)	Límite inferior IC 95%	Límite superior IC 95%
2	163	3	0.982	0.0105	0.961	1.000
3	158	16	0.882	0.0254	0.834	0.933
4	137	4	0.856	0.0277	0.804	0.913
5	133	4	0.831	0.0297	0.774	0.891
6	129	5	0.798	0.0319	0.738	0.863
7	122	7	0.753	0.0344	0.688	0.823
10	97	4	0.722	0.0363	0.654	0.796
11	92	5	0.682	0.0384	0.611	0.762
14	83	2	0.666	0.0392	0.593	0.747
18	61	1	0.655	0.0400	0.581	0.738
23	37	1	0.637	0.0427	0.559	0.727
34	7	1	0.546	0.0919	0.393	0.760

Fuente: elaborado en base a los resultados de R.Studios

En el análisis de supervivencia realizado para los estudiantes de la Escuela Profesional de Matemáticas de la Universidad Nacional de San Antonio Abad del Cusco, se utilizó el estimador de Kaplan-Meier para evaluar la probabilidad de que los estudiantes permanezcan matriculados en la escuela profesional durante diferentes períodos de tiempo. A partir de los datos analizados, se observó una disminución progresiva en la función de supervivencia, reflejando el porcentaje acumulado de estudiantes que no desertaron.

Inicialmente, la probabilidad de supervivencia fue alta (98.2% a los 2 meses), pero esta disminuyó significativamente con el tiempo, alcanzando un 66.6% a los 14 meses y un 54.6% a los 34 meses. Los intervalos de confianza del 95% sugieren una disminución constante en la proporción de estudiantes que permanecen matriculados, con mayor incertidumbre en los períodos más largos debido al reducido número de estudiantes aún en riesgo.

Figura 1*Análisis de Supervivencia*

Fuente: elaborado en base a los resultados de R-Studios

ESTIMACIÓN NO PARAMÉTRICA DE LA FUNCIÓN DE SUPERVIVENCIA SEGÚN MODALIDAD DE INGRESO.

Tabla 6

Supervivencia - Modalidad de Ingreso

Modalidad de Ingreso (Primera oportunidad)							
Tiempo (meses)	Nº en riesgo (n.risk)	Nº de eventos (n.event)	Supervivencia estimada	Error estándar (std.err)	Límite inferior IC 95%	Límite superior IC 95%	
2	28	1	0.964	0.0351	0.898	1.000	
3	27	2	0.893	0.0585	0.785	1.000	
4	24	3	0.781	0.0790	0.641	0.953	
6	21	1	0.744	0.0836	0.597	0.927	
7	20	2	0.670	0.0903	0.514	0.872	
11	11	1	0.609	0.1005	0.440	0.841	
23	4	1	0.457	0.1518	0.238	0.876	

Modalidad de ingreso (Dirimencia)						
2	38	1	0.974	0.0260	0.924	1.000
3	37	1	0.947	0.0362	0.879	1.000
5	35	1	0.920	0.0442	0.838	1.000
6	34	2	0.866	0.0557	0.764	0.983
7	30	2	0.808	0.0653	0.690	0.947
11	23	1	0.773	0.0713	0.645	0.926
Modalidad de ingreso (CEPRU)						
7	29	1	0.966	0.0339	0.901	1.000
10	23	1	0.924	0.0523	0.827	1.000
11	22	1	0.882	0.0646	0.764	1.000
14	19	1	0.835	0.0761	0.699	0.998
34	2	1	0.418	0.2977	0.103	1.000
Modalidad de ingreso (Ordinario)						
2	65	1	0.985	0.0153	0.955	1.000
3	63	13	0.781	0.0516	0.687	0.890
4	49	1	0.765	0.0530	0.668	0.877
5	48	3	0.718	0.0564	0.615	0.837
6	45	2	0.686	0.0582	0.581	0.810
7	43	2	0.654	0.0597	0.547	0.782
10	40	3	0.605	0.0616	0.495	0.738
11	36	2	0.571	0.0626	0.461	0.708
14	33	1	0.554	0.0631	0.443	0.692
18	28	1	0.534	0.0638	0.423	0.675

Fuente: elaborado en base a los resultados de R-Studios

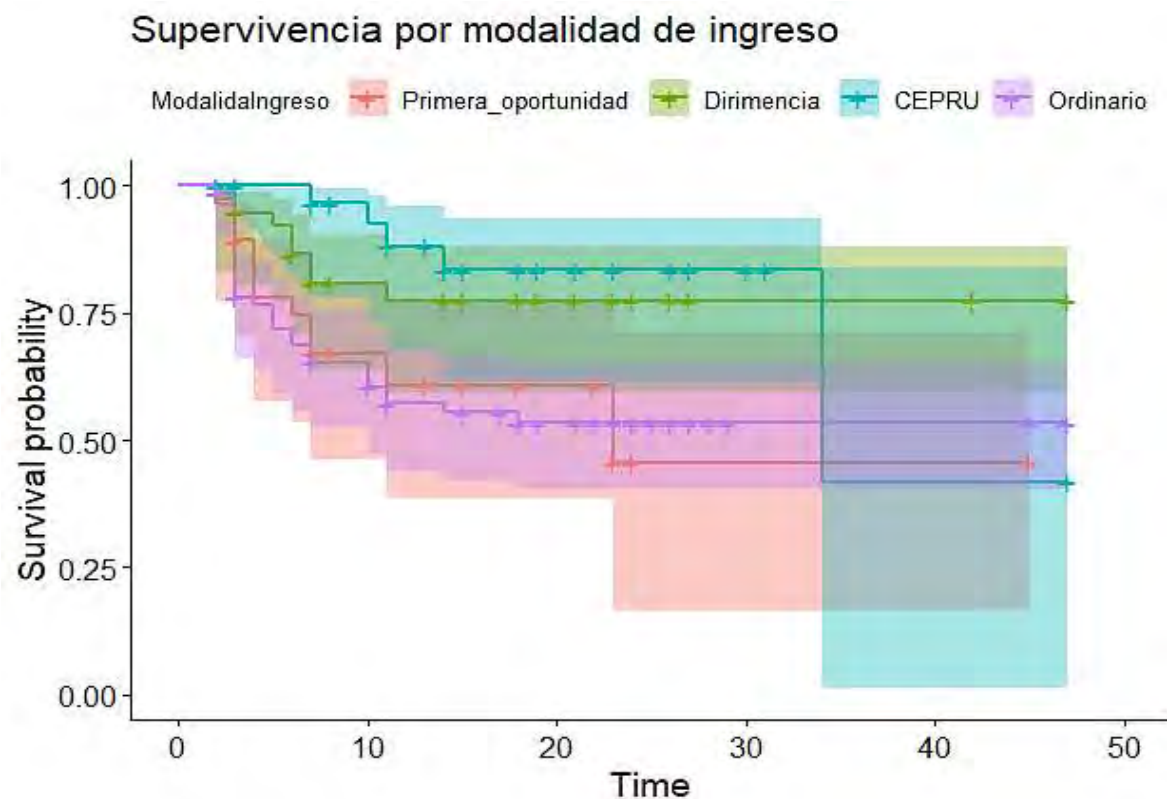
En este trabajo de supervivencia, se estudió la influencia de la modalidad de ingresos sobre la probabilidad de permanencia de los estudiantes en la Escuela estudiada de la UNSAAC, las modalidades tomadas en cuenta fueron la Primera Oportunidad, Dirimencia, CEPRU y Ordinario. Las curvas de Kaplan-Meier enseña diferencias resaltantes en las tasas de supervivencia entre estas modalidades.

Los estudiantes que ingresan mediante Primera Oportunidad y Dirimencia mostraron más probabilidades de permanencia a lo largo del tiempo, con tasas de supervivencia de 60.9% y 77.3%, en ese mismo orden, a los 11 meses. Desde otro punto de vista, los estudiantes que ingresaron por la modalidad de CEPRU y examen ordinario mostraron tasas más altas de deserción llegando a tasas de supervivencia de 41.8%, 53.4% a los 34 y 18 meses, de manera

correspondiente. Estos resultados sugieren que las modalidades por la ingresan están relacionadas con diferencias en el compromiso y la retención estudiantil, resaltando la necesidad de políticas específicas de apoyo según la vía de admisión.

Figura 2

Supervivencia - Modalidad de Ingreso



Fuente: elaborado en base a los resultados de R-Studios

ESTIMACIÓN NO PARAMÉTRICA DE LA FUNCIÓN DE SUPERVIVENCIA SEGÚN SEXO.

Tabla 7

Supervivencia - Sexo

Sexo 1 (Varones)						
-Tiempo (meses)	N° en riesgo (n.risk)	N° de eventos (n.event)	Supervivencia estimada	Error estándar (std.err)	Límite inferior IC 95%	Límite superior IC 95%
2	122	2	0.984	0.0115	0.961	1.000
3	119	13	0.876	0.0299	0.819	0.937
4	101	3	0.850	0.0326	0.789	0.916
5	98	4	0.815	0.0356	0.749	0.888
6	94	5	0.772	0.0386	0.700	0.852
7	88	7	0.711	0.0419	0.633	0.798
10	71	4	0.671	0.0441	0.590	0.763
11	66	3	0.640	0.0455	0.557	0.736
14	60	2	0.619	0.0464	0.534	0.717
18	45	1	0.605	0.0474	0.519	0.705
23	25	1	0.581	0.0513	0.489	0.691
34	3	1	0.387	0.1617	0.171	0.878
Sexo 2 (Mujeres)						
2	41	1	0.976	0.0241	0.930	1.000
3	39	3	0.901	0.0472	0.813	0.998
4	36	1	0.876	0.0521	0.779	0.984
11	26	2	0.808	0.0664	0.688	0.949

Fuente: elaborado en base a los resultados de R-Studios

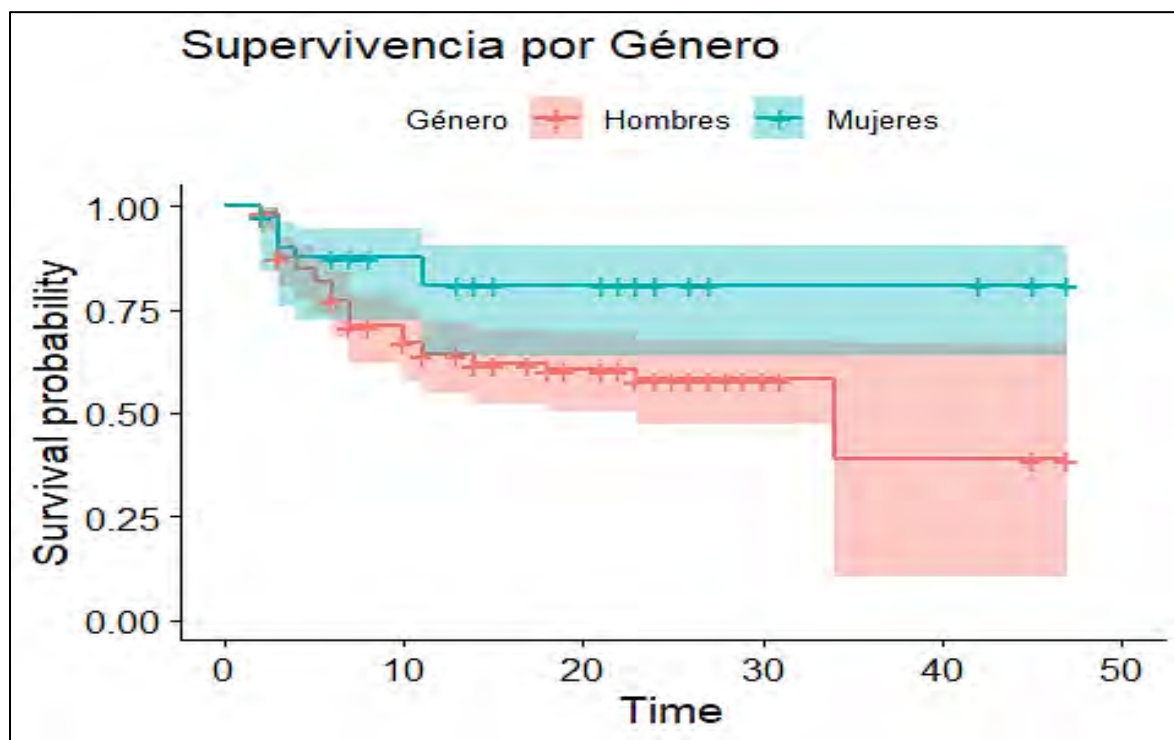
Se evaluó la influencia del sexo en la probabilidad de permanencia de los estudiantes de la Escuela Profesional de Matemáticas de la Universidad Nacional de San Antonio Abad del Cusco. Las curvas de Kaplan-Meier se calcularon para los grupos de varones (1) y mujeres (2), mostrando diferencias en las tasas de supervivencia.

Los resultados indican que las mujeres presentaron mayores probabilidades de permanencia en comparación con los varones. A los 11 meses, la supervivencia acumulada para las mujeres fue del 80.8%, mientras que para los varones fue del 64.0%. Además, los varones

mostraron una disminución más pronunciada en la supervivencia acumulada, alcanzando un 38.7% a los 34 meses, mientras que los datos para las mujeres no registraron eventos de deserción después del mes 11. Estos hallazgos resaltan posibles disparidades de género en los factores asociados con la retención estudiantil, subrayando la importancia de implementar estrategias diferenciadas para abordar las necesidades de cada grupo.

Figura 3

Supervivencia - Género

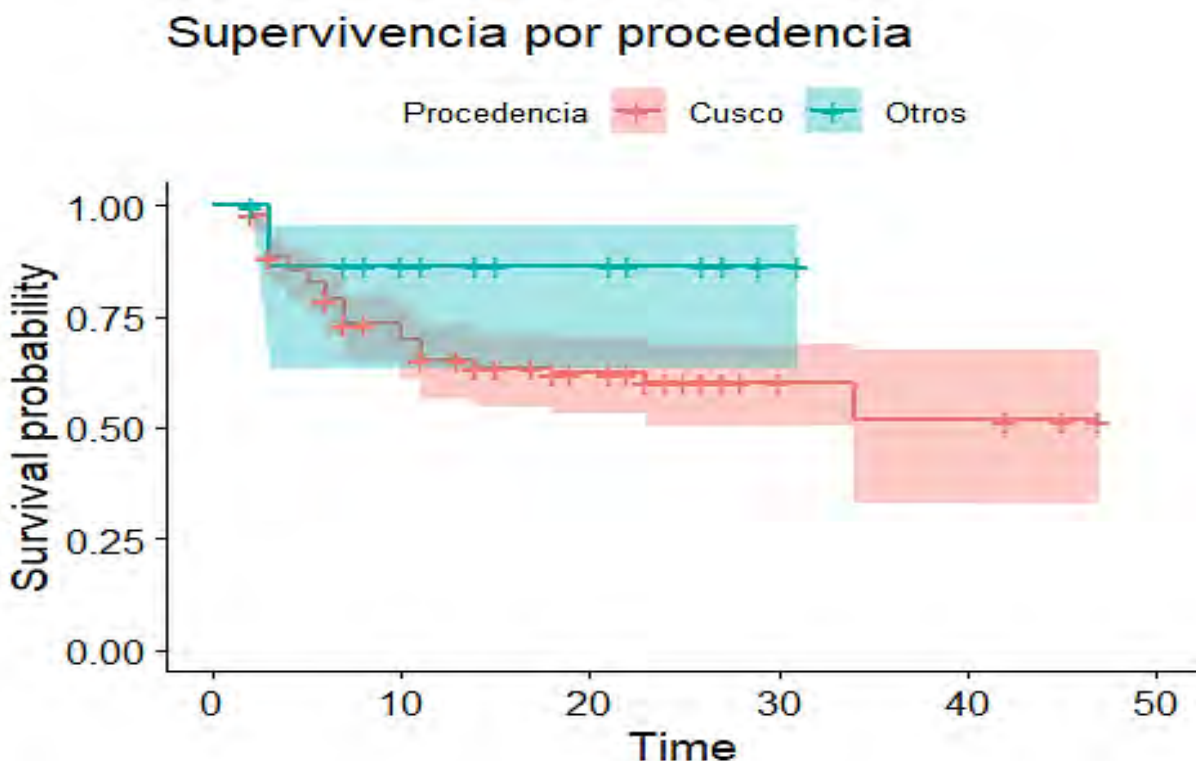


Fuente: elaborado en base a los resultados de R-Studios

ESTIMACIÓN NO PARAMÉTRICA DE LA FUNCIÓN DE SUPERVIVENCIA SEGÚN PROCEDENCIA.

Figura 4

Supervivencia - Procedencia



Fuente: elaborado en base a los resultados de R-Studios

El gráfico de supervivencia por procedencia muestra las probabilidades acumuladas de supervivencia (o permanencia) a lo largo del tiempo, diferenciadas entre estudiantes cuya procedencia es Cusco y aquellos provenientes de otros lugares.

Los estudiantes provenientes de Cusco (línea roja) presentan una tasa de supervivencia notablemente más baja en comparación con los de otras procedencias (línea azul). La caída en la supervivencia acumulada de los estudiantes de Cusco es más pronunciada en los primeros meses, estabilizándose posteriormente, pero en niveles más bajos. A medida que avanza el tiempo, los estudiantes de Cusco muestran una probabilidad acumulada cercana al 50% hacia el final del

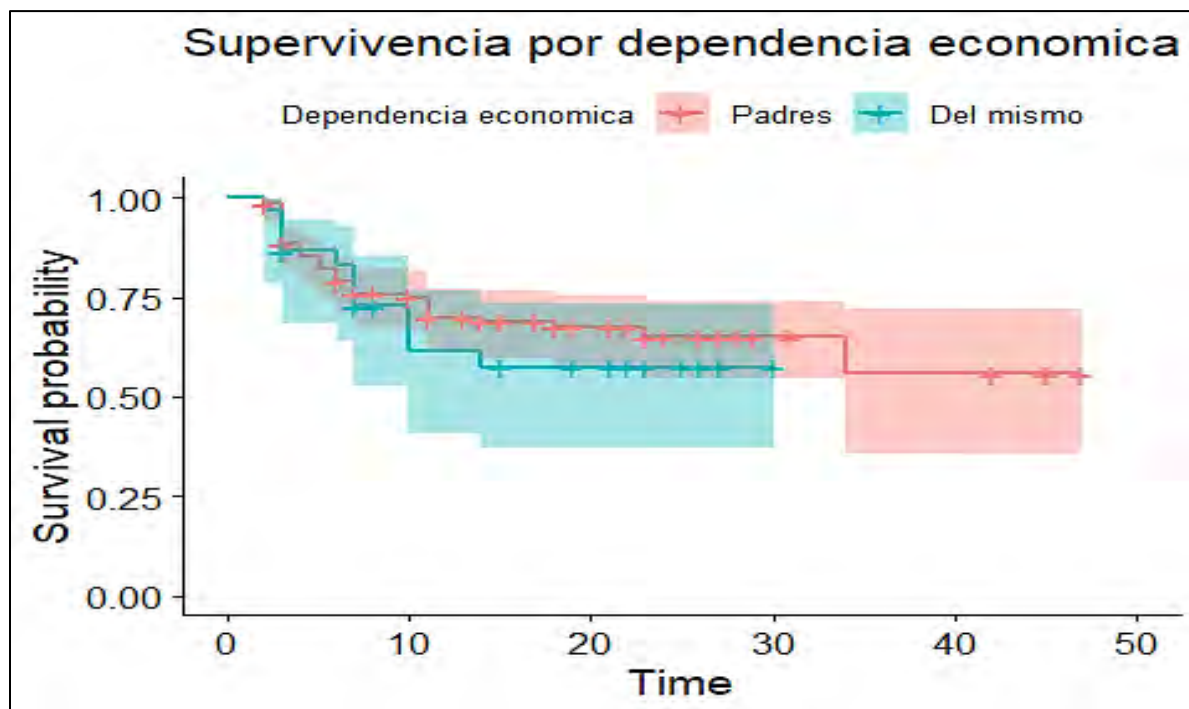
periodo de seguimiento. En contraste, los estudiantes de otras procedencias mantienen una supervivencia acumulada significativamente mayor, con menor reducción a lo largo del tiempo.

En conclusión, la procedencia parece ser un factor relevante en la permanencia estudiantil. Los estudiantes de Cusco enfrentan mayores riesgos de deserción en comparación con los de otras regiones. Esto podría estar relacionado con factores socioeconómicos, familiares, o diferencias en la adaptación al ambiente académico. Es fundamental profundizar en estas posibles causas para diseñar estrategias focalizadas que ayuden en la mejora de la retención de los estudiantes en el grupo más afectado.

ESTIMACIÓN NO PARAMÉTRICA DE LA FUNCIÓN DE SUPERVIVENCIA SEGÚN DEPENDENCIA ECONÓMICA.

Figura 5

Supervivencia según dependencia económica.



Fuente: elaborado en base a los resultados de R-Studios

El gráfico de supervivencia por dependencia económica muestra las probabilidades acumuladas de supervivencia en función de si los estudiantes son económicamente dependientes de sus progenitores o de ellos mismos. Por eso, los estudiantes que dependen de sus padres (línea roja) tienen más posibilidad de supervivencia acumulada a lo largo del tiempo, en comparación con aquellos que son económicamente independientes (línea azul). Los estudiantes que no dependen de nadie experimentan una disminución más avanzada en la probabilidad de supervivencia en las primeras etapas de seguimiento.

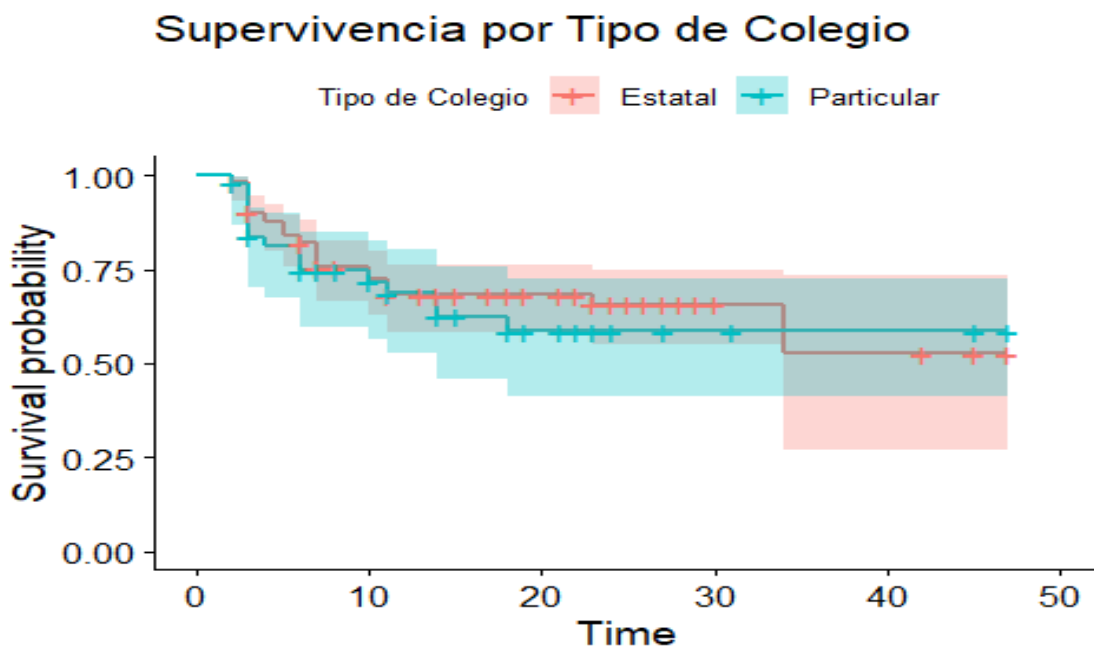
Para los estudiantes que dependen de sus progenitores, la probabilidad de supervivencia se mantiene relativamente más estable, mostrando una mayor retención a largo plazo sin embargo el lado de los estudiantes que dependen de ellos mismos presenta niveles más bajos de supervivencia.

Para concluir, la dependencia económica parece ser un factor determinante, En conclusión, la dependencia económica parece ser un factor determinante en la supervivencia estudiantil. Los estudiantes que dependen de sus padres muestran una ventaja en términos de retención académica, posiblemente porque no enfrentan tantas presiones económicas como aquellos que son responsables de su propio sustento. Esto subraya la importancia del apoyo financiero familiar en la continuidad educativa.

ESTIMACIÓN NO PARAMÉTRICA DE LA FUNCIÓN DE SUPERVIVENCIA SEGÚN TIPO DE COLEGIO.

Figura 6

Supervivencia según tipo de colegio.



Fuente: elaborado en base a los resultados de R-Studios

El gráfico de supervivencia por tipo de colegio presenta las probabilidades de supervivencia por tipo de colegio presenta las probabilidades de supervivencia (o retención) en función del tiempo, diferenciando entre estudiantes en función del tiempo, diferenciando entre estudiantes provenientes de colegios estatales (línea roja) y particulares (línea azul). Los estudiantes provenientes de colegios particulares tienen, en general, una probabilidad acumulada de supervivencia ligeramente mayor que aquellos de colegios estatales durante los primeros períodos de seguimiento.

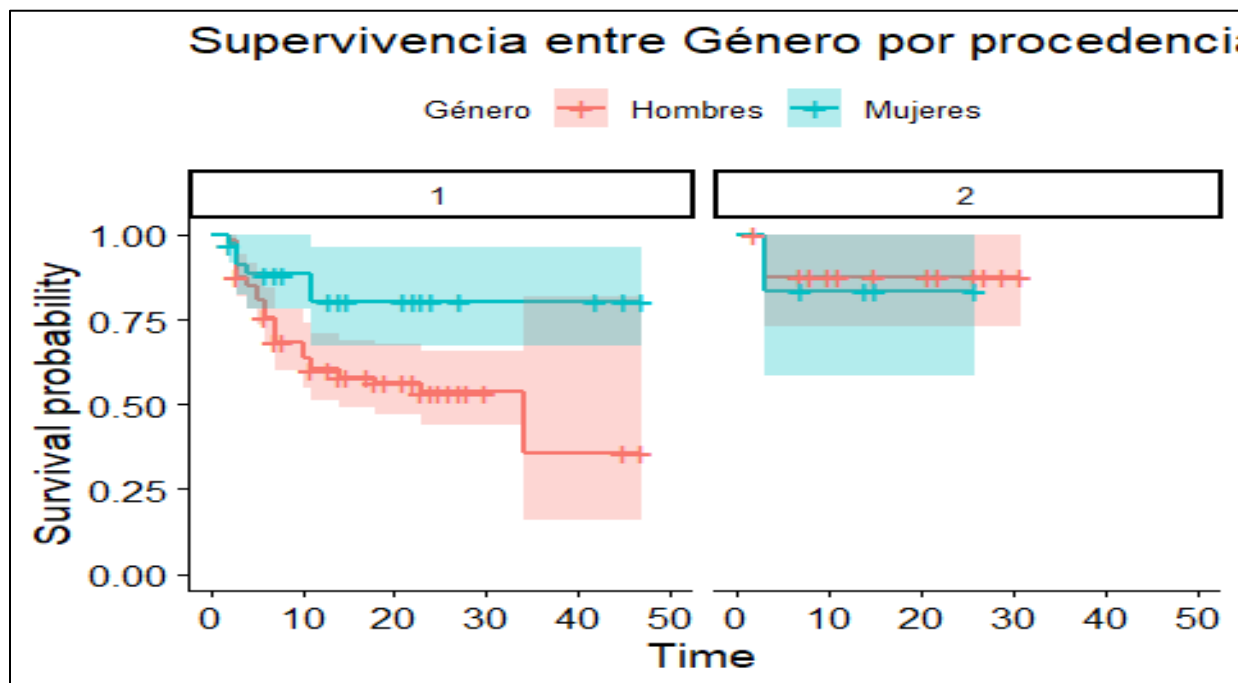
Sin embargo, las curvas se estabilizan hacia el final, presentando diferencias menores entre los grupos.

En los primeros periodos, se ve una disminución más pronunciada para ambo grupos lo que sugiere una disminución inicial significativa en la retención. Los estudiantes de colegios privados parecen tener cierta ventaja con inicial en términos de probabilidad de supervivencia. A medida que avanza el tiempo, las curvas tienden a converger. Esto muestra las diferencias iniciales se reducen y las probabilidades de supervivencia son parecidos a los periodos finales.

Los estudiantes de colegios particulares muestran inicialmente mejores probabilidades de retención que los de colegios estatales, posiblemente debido a diferencias en la preparación académica o recursos. Sin embargo, esta ventaja se reduce con el tiempo. Es importante considerar otros factores contextuales, como el apoyo académico, económico o personal, que podrían influir en las tasas de retención independientemente del tipo de colegio.

ESTIMACIÓN NO PARAMÉTRICA DE LA FUNCIÓN DE SUPERVIVENCIA SEGÚN SEXO Y PROCEDENCIA.

Figura 7
Supervivencia según sexo y procedencia.



Fuente: elaborado en base a los resultados de R-Studios

Este gráfico de supervivencia estratificada por género y procedencia presenta las probabilidades de supervivencia a lo largo del tiempo para hombres y mujeres (indicados con colores rojo y azul, respectivamente) en dos grupos de procedencia (1 y 2).

Grupo de Procedencia (1) Cusco; hombres (rojo): la probabilidad de supervivencia disminuye más rápidamente en comparación con las mujeres, especialmente en los primeros 20-30 intervalos de tiempo. Esto sugiere que los hombres tienen un mayor riesgo de experimentar el evento (deserción) en este grupo. Mientras que las mujeres (azul) tienen una curva de supervivencia más estable, con una probabilidad de supervivencia mayor en todos los tiempos observados.

Grupo (2) de Procedencia otros lugares; ambas curvas, para hombres y mujeres, son más cercanas y muestran menores diferencias en comparación con el grupo Cusco. Esto podría indicar que la procedencia en este caso tiene un efecto moderador, reduciendo las diferencias de riesgo entre géneros. El grupo de procedencia Cusco presenta mayor dispersión en las probabilidades de supervivencia a lo largo del tiempo, sugiriendo una posible influencia de factores contextuales o características asociadas a este grupo.

ESTIMACIÓN DE LA FUNCIÓN DE RIESGO ACUMULADO

Tabla 8

función de riesgo acumulado

Tiempo(meses)	N° en riesgo (n.risk)	N° de eventos (n.event)	N° de censura (n.censor)	Supervivencia estimada (surv)	Error estándar (std.err)	Límite superior IC 95%	Límite inferior IC 95%	Hazard Acumulado (CumHaz)
2	163	3	2	0.9815951	0.01072524	10.000.000	0.9611763	0.01840491
3	158	16	5	0.8821931	0.02877795	0.9333821	0.8338113	0.11967073
4	137	4	0	0.8564356	0.03236815	0.9125287	0.8037905	0.14886781
5	133	4	0	0.8306781	0.03578880	0.8910381	0.7744070	0.17894300
6	129	5	2	0.7984813	0.03991762	0.8634609	0.7383917	0.21770269
7	122	7	12	0.7526668	0.04574218	0.8232630	0.6881244	0.27507974
8	103	0	6	0.7526668	0.04574218	0.8232630	0.6881244	0.27507974
10	97	4	1	0.7216290	0.05035630	0.7964845	0.6538086	0.31631685
11	92	5	2	0.6824100	0.05621783	0.7619002	0.6112132	0.37066468
13	85	0	2	0.6824100	0.05621783	0.7619002	0.6112132	0.37066468
14	83	2	4	0.6659664	0.05880417	0.7473199	0.5934691	0.39476106
15	77	0	14	0.6659664	0.05880417	0.7473199	0.5934691	0.39476106
17	63	0	2	0.6659664	0.05880417	0.7473199	0.5934691	0.39476106
18	61	1	5	0.6550489	0.06108318	0.7383595	0.5811385	0.41115451
19	55	0	3	0.6550489	0.06108318	0.7383595	0.5811385	0.41115451
21	52	0	6	0.6550489	0.06108318	0.7383595	0.5811385	0.41115451
22	46	0	9	0.6550489	0.06108318	0.7383595	0.5811385	0.41115451
23	37	1	8	0.6373449	0.06694704	0.7267080	0.5589708	0.43818153
24	28	0	3	0.6373449	0.06694704	0.7267080	0.5589708	0.43818153
25	25	0	1	0.6373449	0.06694704	0.7267080	0.5589708	0.43818153
26	24	0	7	0.6373449	0.06694704	0.7267080	0.5589708	0.43818153
27	17	0	6	0.6373449	0.06694704	0.7267080	0.5589708	0.43818153
28	11	0	1	0.6373449	0.06694704	0.7267080	0.5589708	0.43818153
29	10	0	1	0.6373449	0.06694704	0.7267080	0.5589708	0.43818153
30	9	0	1	0.6373449	0.06694704	0.7267080	0.5589708	0.43818153
31	8	0	1	0.6373449	0.06694704	0.7267080	0.5589708	0.43818153
34	7	1	0	0.5462956	0.16820056	0.7596269	0.3928757	0.58103868
42	6	0	1	0.5462956	0.16820056	0.7596269	0.3928757	0.58103868
45	5	0	2	0.5462956	0.16820056	0.7596269	0.3928757	0.58103868
47	3	0	3	0.5462956	0.16820056	0.7596269	0.3928757	0.58103868

Fuente: elaborado en base a los resultados de R-Studios

La estimación de la función de riesgo acumulado proporciona una medida del riesgo total acumulado de que ocurra el evento de interés (la deserción) en un periodo de tiempo dado. A continuación, se realiza la interpretación basada en los resultados otorgados. El riesgo acumulado inicia en un valor bajo (0.0184 en el tiempo 2) e incrementa progresivamente hasta llegar a 0.5810 en el tiempo 47. Esto muestra que el riesgo de experiencia del evento se incrementa con el tiempo, como era de esperarlo y ese tipo de análisis.

Algunos cambios notables en el riesgo acumulado, se puede observar que entre el tiempo 2 y el tiempo 6, se observa un aumento progresivo en el riesgo acumulado, pasando de 0.0184 a 0.2177. Este crecimiento se desacelera hacia el tiempo 18, con un incremento más moderado en el riesgo acumulado (de 0.2750 a 0.4111).

Un incremento significativo se observa entre los tiempos 27 y 34 (de 0.4381 a 0.5810). Este salto refleja una concentración de eventos durante este intervalo.

Con conclusión, el riesgo acumulado en este caso, indica cómo la exposición al riesgo se incrementa con el tiempo, especialmente en ciertos intervalos críticos (tiempo 27-34). La relación inversa entre el riesgo acumulado y la supervivencia refuerzan la importancia de abordar factores que podrían estar contribuyendo al incremento del riesgo en los períodos clave.

ESTIMACIÓN DE LA FUNCIÓN DE RIESGO ACUMULADO SEGÚN MODALIDAD DE INGRESO

Tabla 9

Función de riesgo acumulado según modalidad de ingreso

Tiempo (MESES)	N° en riesgo (n.risk)	N° de eventos (n.event)	N° de censura (n.censor)	de Supervivencia estimada (surv)	Error estándar (std.err)	Límite superior IC 95%	Límite inferior IC 95%
2	28	1	0	0.964	0.0364	1.000	0.898
3	27	2	1	0.893	0.0655	1.000	0.785
4	24	3	0	0.781	0.101	0.953	0.641
6	21	1	0	0.744	0.112	0.927	0.597
7	20	2	5	0.670	0.135	0.872	0.514
8	13	0	2	0.670	0.135	0.872	0.514
11	11	1	0	0.609	0.165	0.841	0.440
13	10	0	1	0.609	0.165	0.841	0.440
15	9	0	3	0.609	0.165	0.841	0.440
18	6	0	1	0.609	0.165	0.841	0.440

Fuente: Elaborado en base a los resultados de R-Studios

El riesgo acumulado en función del tiempo en meses, estratificado por las modalidades de ingreso:

- A (Primera oportunidad)

- B (Dirimencia)
- C (CEPRU)
- D (Examen ordinario)

En las curvas de riesgo acumuladas, la pendiente de cada curva indica la velocidad con la que se acumulan eventos (como abandonos, fracasos o el evento de interés) a lo largo del tiempo. Una pendiente más pronunciada indica un mayor riesgo acumulado en menos tiempo.

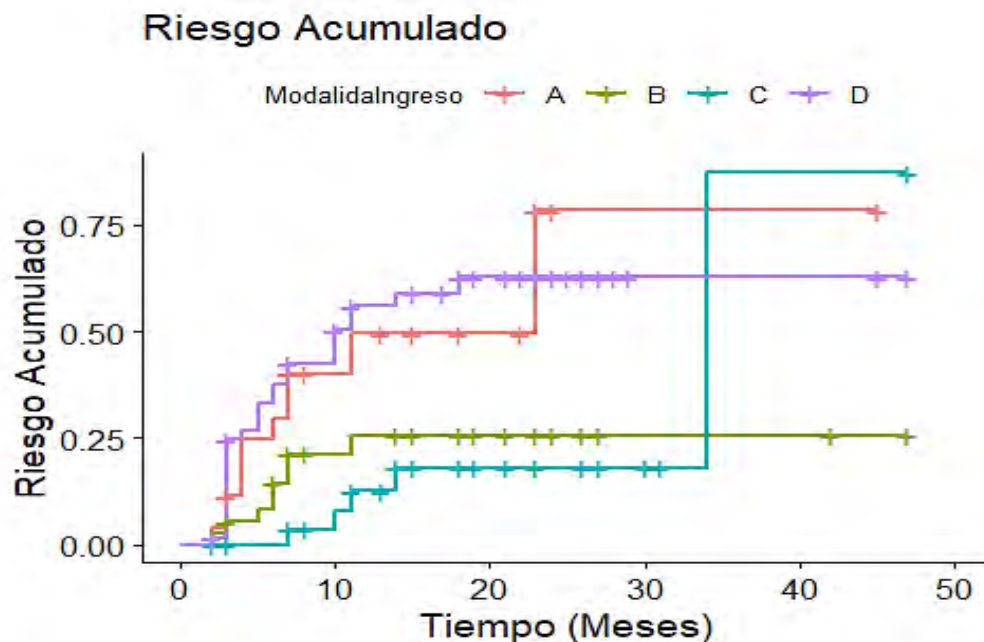
Comparaciones entre modalidades

- *Modalidad A (Primera oportunidad)*: tiene un incremento temprano y relativamente pronunciado en el riesgo acumulado. Esto sugiere que los estudiantes de esta modalidad presentan un mayor riesgo de experimentar el evento al inicio, estabilizándose después de cierto tiempo.
- *Modalidad B (Dirimencia)*: su curva es más baja y menos pronunciada, indicando un menor riesgo acumulado y una mayor estabilidad en el tiempo.
- *Modalidad C (CEPRU)*: la curva muestra un comportamiento intermedio, con un riesgo acumulado que se incrementa más lentamente que la Modalidad A, pero más rápido que la B.
- *Modalidad D (Examen ordinario)*: Presenta el mayor riesgo acumulado hacia el final del período, con incrementos importantes, lo que sugiere que los estudiantes en esta modalidad son más vulnerables al evento en etapas tardías.

Los estudiantes que ingresan por Primera Oportunidad (A) tienen un riesgo inicial más alto, probablemente debido a una falta de preparación o adaptación al sistema educativo. Los estudiantes que ingresan por Dirimencia (B) parecen tener un perfil más estable, con el menor riesgo acumulado a lo largo del tiempo. Los de Examen Ordinario (D) experimentan un riesgo creciente hacia el final del período, lo que podría estar asociado con factores estructurales o personales que se manifiestan más tarde.

Este gráfico refleja diferencias significativas en la evolución del riesgo acumulado entre las modalidades de ingreso, destacando que el tipo de ingreso podría estar vinculado con las probabilidades de ocurrencia del evento. Esto sugiere que sería relevante analizar factores asociados (como el soporte académico, económico o emocional) para diseñar intervenciones que disminuyan el riesgo en las modalidades más vulnerables.

Figura 8
Riesgo acumulado meses



Fuente: elaborado en base a los resultados de R-Studios.

ESTIMACIÓN PARAMÉTRICA DE LA FUNCIÓN DE SUPERVIVENCIA

Tabla 10

Estimación paramétrica de la función de supervivencia

Parámetro	Estimación (est)	Límite inferior IC 95% (L95%)	Límite superior IC 95% (U95%)	Error estándar (se)
Tasa de eventos (rate)	0.02232	0.01705	0.02921	0.00307
Métrica	Valor			
Número total de sujetos (N)	163			
Número de eventos	53			
Número de censurados	110			
Tiempo total en riesgo	2375			
Log-verosimilitud	-2.545.304			
Grados de libertad (df)	1			
Criterio de información de Akaike (AIC)	5.110608			

Fuente: elaborado en base a los resultados de R-Studios.

La estimación paramétrica de la función de supervivencia con un modelo exponencial permite caracterizar la supervivencia y el riesgo bajo la suposición de que el tiempo hasta el evento

sigue una distribución exponencial, donde el riesgo es constante en el tiempo. A continuación, se interpretan los resultados proporcionados.

La distribución exponencial asume que la tasa de riesgo es constante en el tiempo, lo que significa que la probabilidad de que ocurra un evento no depende del tiempo transcurrido.

Tasa de riesgo $\hat{\lambda} = 0.02232$ viene a ser la tasa de eventos esperados por unidad de tiempo es 0.02232. Una tasa baja como esta indica que el evento (fallo) es poco frecuente en cada unidad de tiempo. Con respecto al Intervalo de confianza, al 95% [0.01705, 0.02921], este rango sugiere la posible variación de la estimación de la tasa de riesgo en la población.

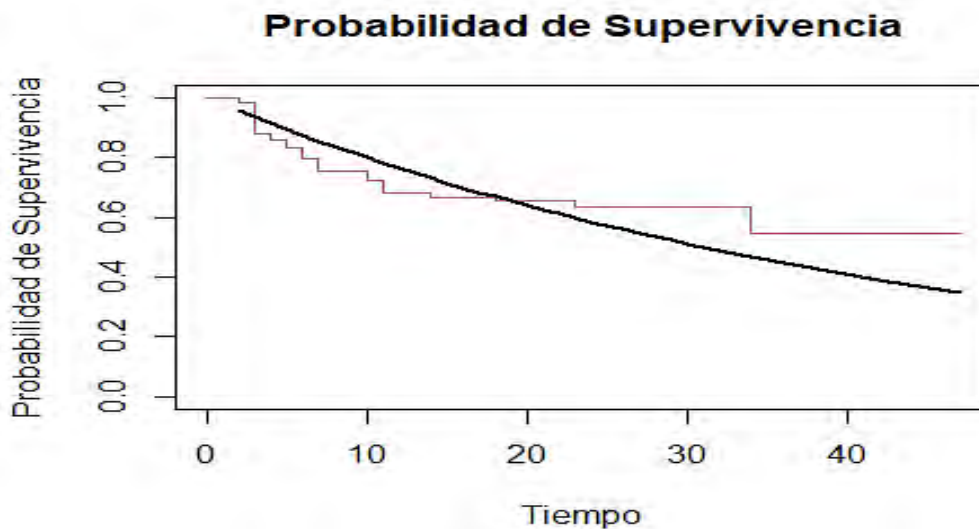
1. La media del tiempo de supervivencia para un modelo exponencial es el inverso de la tasa. Esto sugiere que, en promedio, los individuos permanecerán en "supervivencia" por 44.8 unidades de tiempo antes de que ocurra el evento.

2. Para un modelo exponencial, la función de supervivencia es: $S(t) = e^{-\lambda t}$, con

$\hat{\lambda} = 0.02232$, la probabilidad de que un individuo "sobreviva" más allá del tiempo t disminuye exponencialmente a medida que t aumenta.

Figura 9

Probabilidad de supervivencia tiempo

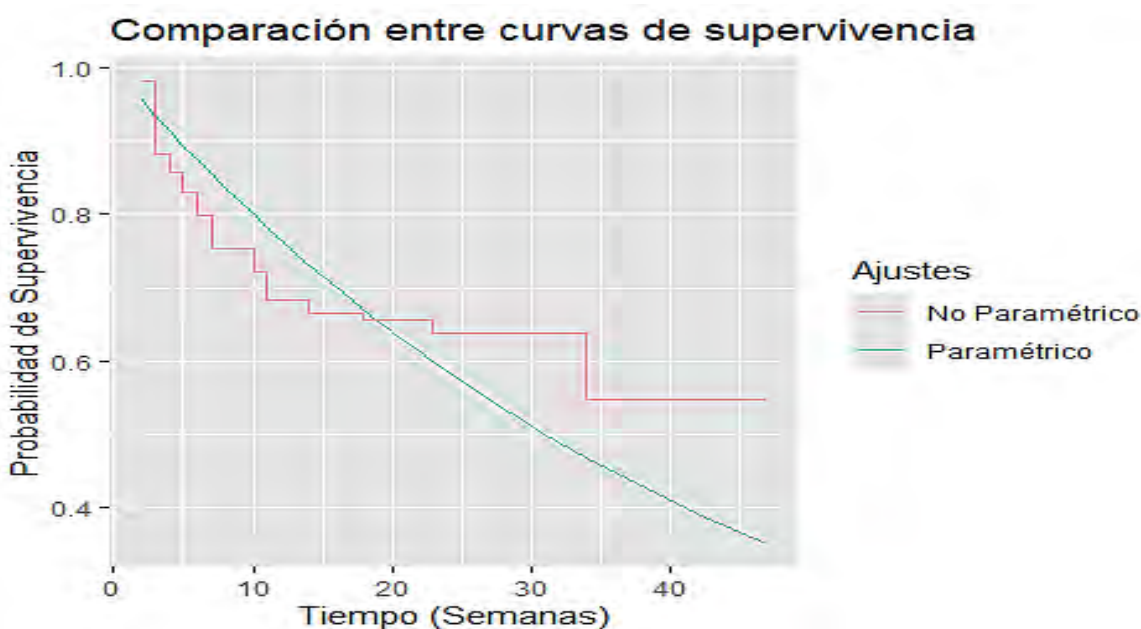


Fuente: elaborado en base de resultados R-Studios

COMPARACIÓN PARAMÉTRICA Y NO PARAMÉTRICA DE CURVAS DE SUPERVIVENCIA

Figura 10

Comparación entre curvas de supervivencia



Fuente: elaborado en base a los resultados de R-Studios

Modelo no paramétrico (Kaplan-Meier) está representada por la curva escalonada indica estimaciones puntuales de la probabilidad de supervivencia basadas en los datos observados. Cada "escalón" corresponde a un evento (como muerte, fallo o abandono), mientras que los tramos planos reflejan períodos en los que no ocurren eventos. Es un modelo flexible que no asume una distribución específica para los tiempos de supervivencia, lo que lo hace útil para descripciones empíricas de los datos.

La línea continua azul representa un modelo que asume que los tiempos de supervivencia siguen una distribución exponencial. Este modelo implica que el riesgo de ocurrencia del evento es constante a lo largo del tiempo (una suposición fuerte que puede no cumplirse en muchos casos).

En las primeras semanas, ambas curvas están relativamente cercanas, lo que indica que el modelo exponencial describe razonablemente bien los datos en este período. A medida que aumenta el tiempo (después de aproximadamente 20 semanas), el modelo exponencial subestima la probabilidad de supervivencia en comparación con Kaplan-Meier. Esto sugiere que la suposición de un riesgo constante no se ajusta bien a los datos a largo plazo. La curva de Kaplan-Meier, al ser más detallada, muestra períodos específicos donde el riesgo de evento podría estar disminuyendo o siendo irregular.

El modelo paramétrico exponencial es más simplificado y hace suposiciones fuertes, que no se ajustan perfectamente a estos datos, como lo muestra la divergencia con Kaplan-Meier en el tiempo. La elección del modelo debe depender del objetivo del análisis y la validez de las suposiciones paramétricas.

COMPARACIÓN PARAMÉTRICA Y NO PARAMÉTRICA DEL RIESGO ACUMULADO.

Figura 11

Comparación entre riesgo acumulado



Fuente: elaborado en base a los resultados de R-Studios.

En las primeras semanas, las curvas son relativamente similares, lo que sugiere que el modelo exponencial describe razonablemente el riesgo acumulado inicial. A medida que pasa el tiempo, el modelo exponencial sobreestima el riesgo acumulado en comparación con Kaplan-Meier, especialmente después de la semana 20. Esto indica que la suposición de riesgo constante del modelo exponencial no es completamente válida para estos datos.

El modelo no paramétrico muestra que el riesgo no se incrementa de manera constante, sino en "saltos" relacionados con los eventos observados, con períodos en los que el riesgo acumulado se estabiliza (tramos planos). Kaplan-Meier es útil para describir los datos observados en detalle, especialmente en situaciones donde el riesgo no es constante y varía a lo largo del tiempo. Modelo

exponencial: Aunque más simple, podría no ser adecuado si los datos no cumplen la suposición de tasa de riesgo constante, como se observa aquí.

El modelo paramétrico exponencial, al asumir un aumento constante del riesgo acumulado, puede ser útil en análisis simplificados o contextos donde esta suposición sea razonable. Sin embargo, el modelo Kaplan-Meier proporciona una descripción más precisa de los datos cuando el riesgo acumulado tiene una evolución irregular, como en este caso.

COMPARACIÓN DE MODELOS

Tabla 11

Comparación de modelos

Modelo	Parámetro	Estimación (est)	Límite inferior IC 95% (L95%)	Límite superior IC 95% (U95%)	Error estándar (se)	Log-verosimilitud	AIC
Exponencial (exp)	Rate	0.02232	0.01705	0.02921	0.00307	-2.545304	5.110608
Weibull	Shape	0.860	0.681	1.088	0.103	-2.536863	5.113726
	Scale	51.922	34.691	77.712	10.683		
Log-logística (llogis)	Shape	1.013	0.806	1.275	0.119	-2.510294	5.060588
	Scale	33.120	22.414	48.942	6.599		
Gompertz	Shape	-0.06959	-0.11242	-0.02677	0.02185	-2.480504	5.001009
	Rate	0.04083	0.02736	0.06094	0.00834		
Generalizado F (genf)	Mu	1.20e+00	7.09e-01	1.69e+00	2.49e-01	-2.321704	4.723409
	Sigma	8.15e-01	4.74e-01	1.40e+00	2.25e-01		
	Q	-6.05e+00	-	-	1.96e+00		
	P	9.76e-03	6.11e-27	1.56e+22	2.78e-01		

Fuente: elaborado en base a los resultados de R-Studios

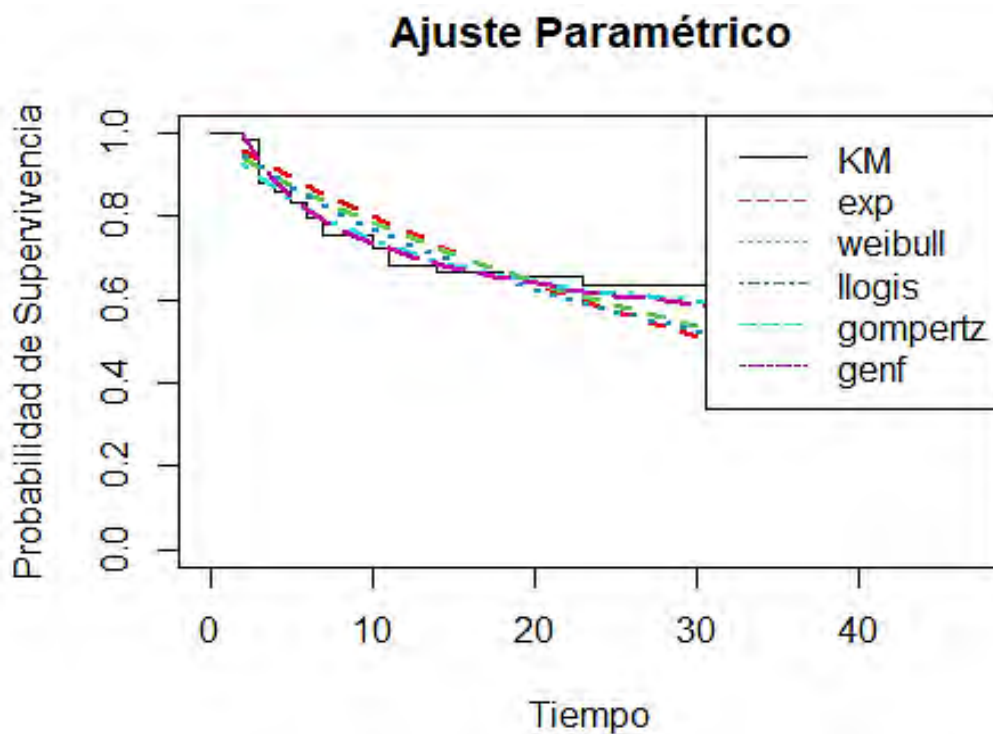
Comparación de modelos (AIC) el modelo GenF (AIC = 472.34) ajusta mejor a los datos. Gompertz (AIC = 500.10), así como es competitivo, pero menos flexible que GenF. Log-logística (AIC = 506.06); es una opción intermedia, con buen ajuste en casos de riesgos no constantes.

Exponencial y Weibull tienen un AIC similar (cercano a 511), indicando un ajuste más pobre comparado con los modelos más flexibles.

El modelo Generalizado de F (GenF) es el más adecuado para estos datos, según el criterio de AIC. Esto se debe a su capacidad para capturar patrones complejos en los datos de supervivencia. Sin embargo, si se prioriza la simplicidad, la Gompertz o la Log-logística podrían ser opciones razonables dependiendo de los objetivos del análisis.

Figura 12

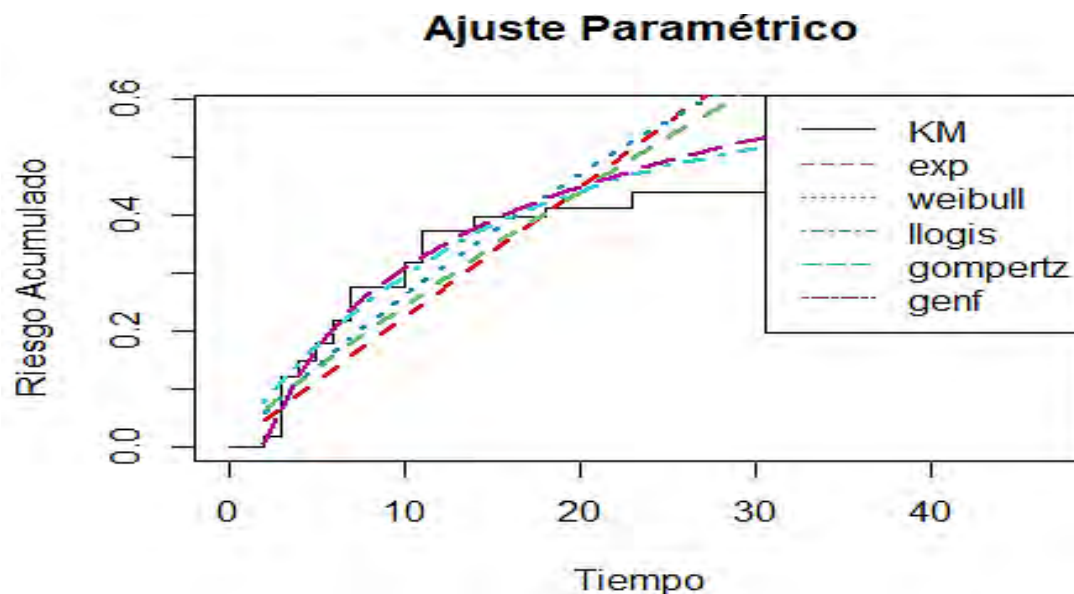
Ajuste paramétrico 1



Fuente: elaborado en base a los resultados de R-Studios.

Figura 13

Ajuste paramétrico 2



Fuente: elaborado en base a los resultados de R-Studio.

REGRESIÓN DE COX CON VARIABLES CUANTITATIVAS

Tabla 12

Regresión de Cox con variables cuantitativas

Variable	Coefficiente (coef)	Exp(coef)	Error estándar (se(coef))	Z	p-valor (p)
Edad	0.1269227	1.1353292	0.0436408	2.908	0.00363
Rend. Académico	-0.1070493	0.8984813	0.0431495	-2.481	0.01311
Pensión Mensual	-0.0002996	0.9997005	0.0008557	-0.350	0.72627
Ingreso Familiar	0.0003473	1.0003474	0.0002456	1.414	0.15736
Métrica	Valor				
Test de razón de verosimilitud	16.06				
Grados de libertad (df)	4				
p-valor del modelo	0.002942				
Número de sujetos (n)	163				
Número de eventos	53				

Fuente: elaborado en base a los resultados de R-Studios

La regresión de Cox es un modelo de riesgos proporcionales utilizado para analizar la relación entre varias variables y el tiempo hasta que ocurre un evento (en este caso, un evento de interés como la condición de un paciente o individuo). A continuación, se interpretan los resultados de la regresión de Cox con variables numéricas:

Edad: un aumento de una unidad en la variable Edad incrementa el riesgo de experimentar el evento en un 13.5% ($\exp(\text{coef}) - 1 = 0.135$), lo cual es estadísticamente significativo ($p < 0.05$). Esto indica que la edad tiene un efecto positivo en el riesgo del evento.

Rendimiento Académico (Rend. Académico): un aumento de una unidad en el rendimiento académico disminuye el riesgo de experimentar el evento en un 10.15% ($1 - 0.898$). Este efecto es estadísticamente significativo ($p < 0.05$), lo que sugiere que mejores resultados académicos están asociados con un menor riesgo de que ocurra el evento.

Pensión Mensual (PensionMensual): el efecto de la pensión mensual es muy pequeño ($\exp(\text{coef})$ cercano a 1), y el valor $p > 0.05$, lo que indica que no tiene un impacto significativo sobre el riesgo de experimentar el evento en este modelo.

Ingreso Familiar (IngrFamiliar): el efecto del ingreso familiar, también, es muy pequeño ($\exp(\text{coef})$ cercano a 1) y no es estadísticamente significativo ($p > 0.05$), lo que sugiere que no hay una relación fuerte entre el ingreso familiar y el riesgo del evento en este modelo.

Por tanto, la Edad es un factor de riesgo significativo, ya que incrementa el riesgo de que ocurra el evento, como se refleja en el coeficiente positivo. A medida que la edad aumenta, también, lo hace la probabilidad de que ocurra el evento. Asimismo, el rendimiento académico tiene un efecto protector, dado que una mayor puntuación en esta variable disminuye el riesgo del evento. Esto podría reflejar un mayor bienestar o un entorno que disminuye la probabilidad de

experimentar el evento. Por otro lado, la pensión mensual e ingreso familiar no presentan un efecto significativo en el modelo, lo que sugiere que estas variables no tienen un impacto claro sobre el riesgo del evento en este caso específico.

REGRESIÓN DE COX CON VARIABLES CUALITATIVAS

Tabla 13

Regresión de cox con variables cualitativas

Variable	Coficiente (coef)	Exp(coef)	Error estándar (se(coef))	Z	p-valor (p)
Sexo Mujer	-0.91016	0.40246	0.42281	-2.153	0.03135
Procedencia Otros	-1.23131	0.29191	0.61407	-2.005	0.04495
Tipo de Colegio Particular	-0.10893	0.89679	0.31499	-0.346	0.72947
Depende del mismo	0.53889	1.71410	0.36234	1.487	0.13695
Ocupación del jefe: No profesional	-0.77883	0.45894	0.34977	-2.227	0.02597
Casa Ocupa: Alquilado	-0.00496	0.99505	0.32550	-0.015	0.98784
Modalidad de Ingreso: Dirimencia	-0.78512	0.45607	0.47797	-1.643	0.10047
Modalidad de Ingreso: CEPRU	-1.53897	0.21460	0.57392	-2.682	0.00733
Modalidad de Ingreso: Ordinario	-0.18668	0.82971	0.37459	-0.498	0.61824

Métrica	Valor
Test de razón de verosimilitud	26.39
Grados de libertad (df)	9
p-valor del modelo	0.001765
Número de sujetos (n)	163
Número de eventos	53

Fuente: elaborado en base a los resultados de R-Studios

De los resultados de este modelo, nos indica que ser mujer reduce el riesgo del evento en un **59.75%** ($1 - 0.40246$), lo cual es estadísticamente significativo ($p < 0.05$). Este efecto negativo sugiere que las mujeres tienen menos riesgo de experimentar el evento (deserción) en comparación con los hombres. Proceder de un lugar distinto (Otros) reduce el riesgo del evento en un **70.81%** ($1 - 0.29191$), lo cual es significativo ($p < 0.05$). Por otro lado, el tipo de colegio no tiene un impacto significativo sobre el riesgo del evento ($p > 0.05$), lo que indica que no hay

una diferencia clara en el riesgo entre los estudiantes de colegios particulares y otros tipos de colegios.

La ocupación de jefe en un puesto no profesional reduce el riesgo del evento en un 54.11% ($1 - 0.45894$), lo cual es estadísticamente significativo ($p < 0.05$). Asimismo, la modalidad de ingreso "CEPRU" reduce el riesgo del evento en un 78.54% ($1 - 0.21460$), y este efecto es estadísticamente significativo ($p < 0.05$).

A modo de conclusiones, se puede sugerir intervenciones dirigidas a mujeres, dado que las mujeres tienen un menor riesgo de experimentar el evento. Podría considerarse un enfoque específico de intervención para otros grupos con mayores riesgos., enfoque en la modalidad de ingreso "CEPRU": la modalidad de ingreso "CEPRU" tiene un impacto significativo, lo que podría indicar un área importante de enfoque para mejorar los resultados.

DISCUSIÓN

El objetivo general de modelar la supervivencia de los estudiantes en la Escuela Profesional de Matemática de la UNSAAC mediante métodos paramétricos y no paramétricos se concretó con éxito, permitiendo un análisis integral de los factores que influyen en la deserción estudiantil. En esta investigación, se aplicaron técnicas como el método de Kaplan-Meier (no paramétrico), regresión de Cox (semiparamétrico) y diversos modelos paramétricos (Exponencial, Weibull, Gompertz, Log-logístico y Generalizado F), lo que permitió identificar los periodos críticos de deserción y las variables asociadas a la misma.

En comparación con los antecedentes, como el estudio de Medina y Romero (2020), quienes utilizaron modelos de supervivencia para analizar el abandono escolar en la URACCAN-Nueva Guinea, se observan similitudes metodológicas al emplear análisis de Kaplan-Meier y regresión de Cox para evaluar las probabilidades de deserción. Sin embargo, a diferencia de su enfoque limitado a métodos semiparamétricos, en esta investigación, se incorporaron modelos paramétricos más complejos, como el Generalizado F, que proporcionaron un mejor ajuste y una mayor capacidad para capturar la heterogeneidad en los datos.

Por otro lado, los resultados de esta investigación, también, coinciden con lo reportado por **García et al. (2022)** en Colombia, quienes destacaron la relevancia de los factores socioeconómicos y académicos en la deserción estudiantil. Sin embargo, mientras que, en el caso colombiano, se identificó un impacto significativo del bajo nivel educativo de los padres, en la UNSAAC, se observó que la modalidad de ingreso y el rendimiento académico fueron factores más determinantes.

Finalmente, esta investigación contribuye a la literatura local al confirmar, como lo sugieren Paucar y Quispe (2021), que las metodologías de análisis de supervivencia son

herramientas efectivas para estudiar la deserción estudiantil en contextos peruanos. A diferencia de sus hallazgos, que se centraron en un enfoque no paramétrico, el uso combinado de modelos paramétricos y semiparamétricos permitió una evaluación más robusta y detallada, lo que evidencia la aplicabilidad y efectividad de estos métodos para modelar la supervivencia estudiantil y abordar el fenómeno de la deserción en la UNSAAC.

En esta investigación, se logró determinar la función de riesgo asociada al tiempo de estudios, mostrando que el riesgo de deserción aumenta significativamente a partir del tercer año, con un incremento notable en los periodos críticos, especialmente entre los estudiantes mayores de 22 años y aquellos con bajo rendimiento académico. Este hallazgo es consistente con los resultados de Medina y Romero (2020), quienes identificaron que la probabilidad de deserción es mayor en los primeros años de estudio, especialmente durante el primer año con un riesgo del 37%. Sin embargo, en el caso de la UNSAAC, el incremento más pronunciado ocurre después del tercer año, lo que podría atribuirse a diferencias estructurales en los planes de estudio y las condiciones sociales de los estudiantes.

Asimismo, en concordancia con lo señalado por García et al. (2022), en Colombia, los factores socioeconómicos, como la dependencia económica y el ingreso familiar bajo, fueron identificados como determinantes clave del riesgo de deserción. En esta investigación, estas variables también influyeron en el riesgo acumulado, especialmente para estudiantes provenientes de entornos rurales y admitidos mediante modalidades como CEPRU.

Por otro lado, en el ámbito local, los hallazgos de Paucar y Quispe (2021) destacaron que variables como la ocupación del jefe de hogar y los ingresos familiares tienen una asociación significativa con el riesgo de abandono. Nuestros resultados corroboran estas observaciones, pero también muestran que la modalidad de ingreso y el rendimiento académico tienen un impacto

particularmente relevante, lo que añade una dimensión más específica al análisis del riesgo en la UNSAAC.

En contraste con lo señalado por Saldívar (2020), quien destacó que la comunicación familiar y la motivación son factores protectores clave, nuestro análisis no incluyó directamente estas variables, aunque los resultados sugieren que podrían jugar un papel importante en la mitigación del riesgo. Esto resalta la necesidad de incorporar factores sociales y emocionales en futuros estudios para obtener una visión más completa del fenómeno.

La función de supervivencia calculada mediante el método Kaplan-Meier mostró que la probabilidad de que los estudiantes continúen sus estudios disminuye significativamente a medida que avanzan los años, llegando a un 28% al finalizar el sexto año. Estos resultados coinciden con los hallazgos de Medina y Romero (2020), quienes identificaron que la probabilidad de supervivencia en la URACCAN, también, disminuye drásticamente durante los primeros años, especialmente el primer año, que representó un punto crítico para la deserción estudiantil. Sin embargo, en el caso de la UNSAAC, el riesgo de abandono se incrementa progresivamente a lo largo de los años, con un punto crítico después del tercer año.

Asimismo, los resultados son consistentes con lo reportado por García et al. (2022), donde se destacó que las tasas de supervivencia son más bajas entre las mujeres y los estudiantes provenientes de entornos socioeconómicos desfavorables. En este estudio, las mujeres presentaron una menor probabilidad de continuar sus estudios ($HR = 0.40$; $p = 0.031$), y los estudiantes de áreas rurales, también, mostraron mayor vulnerabilidad, lo que refleja patrones similares entre ambos contextos.

En contraste, Paucar y Quispe (2021) encontraron que los hombres jóvenes menores de 22 años enfrentaban un mayor riesgo de abandono en la UNSAAC. Esta discrepancia puede deberse

a diferencias en los periodos analizados y las variables incluidas, ya que, en nuestro estudio, se consideraron factores adicionales como la modalidad de ingreso, que mostró un impacto significativo en la supervivencia.

El análisis mediante la regresión de Cox permitió identificar factores significativos asociados al tiempo de estudios, destacando variables como la edad ($HR = 1.13$; $p = 0.0036$), el rendimiento académico ($HR = 0.89$; $p = 0.0131$) y la modalidad de ingreso (CEPRU: $HR = 0.21$; $p = 0.0073$). Estos hallazgos son consistentes con los reportados por Garcia et al. (2022), quienes encontraron que factores como el bajo rendimiento académico y el nivel socioeconómico son determinantes clave en la deserción estudiantil. Sin embargo, en nuestro estudio, la modalidad de ingreso emergió como un factor distintivo, lo que resalta características particulares del sistema de admisión de la UNSAAC.

Por otro lado, los resultados concuerdan con los hallazgos de Sifuentes (2018), quien demostró que los modelos predictivos basados en características individuales, como las calificaciones previas, son efectivos para identificar estudiantes en riesgo de abandono. De manera similar, nuestro estudio muestra que el rendimiento académico es uno de los principales predictores de la deserción en la UNSAAC.

En el ámbito local, Paucar y Quispe (2021) destacaron que el ingreso del hogar y la ocupación del jefe de familia tienen un impacto significativo en la deserción. Aunque estas variables no resultaron significativas en nuestro análisis, los resultados indican que los factores socioeconómicos, como la modalidad de ingreso y la dependencia económica, son claves para entender el fenómeno de la deserción en este contexto.

Finalmente, se observaron discrepancias con lo reportado por Saldivar (2020), quien destacó la importancia de factores sociales como la comunicación familiar y la motivación

personal como protectores contra la deserción. Si bien estas variables no fueron directamente incluidas en nuestro análisis, los resultados sugieren que podrían complementar los factores identificados en esta investigación, abriendo la puerta a estudios futuros que integren aspectos emocionales y sociales en el análisis de la deserción estudiantil.

CONCLUSIONES

- Primero:** En las primeras semanas, el modelo paramétrico exponencial y el no paramétrico descrito Kaplan-Meier presentan similar ajuste de supervivencia. Después de 20 semanas, la curva de Kaplan-Meier presenta mejor ajuste del tiempo de estudios de los estudiantes de la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco 2015-2020.
- Segundo:** El riesgo acumulado de deserción de los estudiantes en la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco al segundo mes es solo de 1.84% y aumenta progresivamente hasta alcanzar 58.10% en el mes 47. Esto refleja que el riesgo de experimentar la deserción se incrementa con el tiempo.
- Tercero:** La función de supervivencia del período de estudio de los estudiantes en la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco entre 2015 y 2020 se modela adecuadamente mediante la función de Kaplan-Meier. Los resultados muestran que la probabilidad de supervivencia disminuye significativamente, la probabilidad de supervivencia fue alta (98.2% a los 2 meses), pero esta disminuyó significativamente con el tiempo, alcanzando un 66.6% a los 14 meses y un 54.6% a los 34 meses y después se estabiliza.
- Cuarto:** Los factores asociados que influyen en el tiempo de deserción de los estudiantes, obtenidos de la regresión de Cox, son la edad, el rendimiento académico, el ingreso familiar y la modalidad de ingreso. Los estudiantes que ingresaron por Primera Oportunidad y Dirimencia presentaron mayores probabilidades de permanencia a lo largo del tiempo, por otro lado, los estudiantes que ingresaron por CEPRU y Examen

Ordinario mostraron mayores tasas de deserción. Las mujeres presentaron mayores probabilidades de permanencia en comparación con los varones. Los estudiantes provenientes de Cusco presentan una tasa de supervivencia notablemente más baja en comparación con los de otras procedencias. Los estudiantes dependientes de sus padres tienen una mayor probabilidad de supervivencia acumulada a lo largo del tiempo, en comparación con aquellos que son económicamente independientes.

RECOMENDACIONES

Primero: A la Dirección de la Escuela Profesional de Matemática y al Departamento de Bienestar Universitario, se recomienda implementar y fortalecer programas de tutoría académica, así como asesoramiento psicológico para los estudiantes en riesgo de deserción. Estos programas deben enfocarse especialmente en aquellos con bajo rendimiento académico, estudiantes de sexo femenino y aquellos provenientes de áreas rurales o con antecedentes socioeconómicos desfavorables. Las tutorías personalizadas pueden ayudar a mejorar las habilidades académicas, mientras que el asesoramiento psicológico puede abordar problemas emocionales y de motivación que afectan la continuidad en los estudios. Además, talleres sobre gestión del tiempo y técnicas de estudio podrían ser una herramienta eficaz para mejorar la retención estudiantil.

Segundo: A la Oficina de Orientación Vocacional y al Departamento de Admisión Es necesario desarrollar programas de orientación vocacional desde el inicio del ciclo académico, para que los estudiantes comprendan mejor las expectativas del programa y alineen sus intereses con sus metas profesionales. Charlas informativas, seminarios con egresados y profesionales del campo, así como visitas a empresas, pueden reducir la deserción causada por la falta de conexión entre las expectativas del estudiante y las realidades del programa. Estos esfuerzos son especialmente relevantes para los estudiantes ingresados mediante modalidades como CEPRU y Dirimencia, quienes demostraron ser más vulnerables al abandono académico.

Tercero: Al Consejo Universitario y a la Facultad de Ciencias se recomienda introducir políticas de flexibilidad académica que permitan a los estudiantes adaptar su horario de clases a sus responsabilidades personales y laborales. Opciones como horarios de clases más

flexibles, modalidades de estudio a tiempo parcial y programas de educación continua pueden beneficiar especialmente a estudiantes mayores de 22 años, quienes enfrentan mayores cargas familiares o laborales. La implementación de estas medidas puede ser clave para reducir la deserción en este grupo demográfico.

Cuarto. Al Rectorado y a la Unidad de Bienestar Estudiantil es fundamental establecer más becas y ayudas económicas destinadas a estudiantes provenientes de familias de bajos ingresos. Estas ayudas deben incluir apoyo financiero directo, acceso a materiales de estudio, servicios de transporte y alimentación. Además, se debe garantizar que los estudiantes tengan acceso a recursos básicos que les permitan enfocarse en sus estudios. Estas medidas ayudarán a disminuir significativamente la deserción por motivos económicos y a garantizar que todos los estudiantes tengan la oportunidad de completar sus programas académicos con éxito.

REFERENCIAS BIBLIOGRÁFICAS

- APN. (2025). *Educación universitaria: hay 30 % de deserción por falta de orientación y de recursos*. Cusco: Andina. Obtenido de <https://andina.pe/agencia/noticia-educacion-universitaria-hay-30-desercion-falta-orientacion-y-recursos-670189.aspx>
- Avendaño, M. (2013). *Modelos Generalizados de Riesgos Proporcionales para el Analisis de Supervivencia*. Tesis doctoral, Universidad Complutense de Madrid, Madrid. Madrid-Ma....rid: <http://eprints.ucm.es/22026/1/T34592.pdf>.
- Barrera, M. (2008). *Análisis de Supervivencia Aplicado al Problema de la Deserción Estudiantil en la Universidad Tecnológica de Pereira*. Tesis de maestría, Universidad Tecnológica de Pereira, Pereira. Pereira- Colombia: <http://repositorio.utp.edu.co/dspace/bitstream/handle/11059/2043/3712913B272.pdf;sequence=1>.
- Barturén, D., & Niño, A. (2018). *Modelos de sobrevivencia para la deserción estudiantil en la Escuela Profesional de Estadística de la Universidad Nacional Pedro Ruiz Gallo, Lambayeque 2007-2008*. 8. Universidad Nacional Pedro Ruiz Gallo, Lambayeque, Lambayeque.: <http://repositorio.unprg.edu.pe/bitstream/handle/UNPRG/3004/BC-TES-TMP-1823.pdf?sequence=1&isAllowed=y>.
- Carrasco, S. (2009). *Metodología de la investigación científica*. San Marcos.
- Cox, D. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2), 187-220. doi:<https://doi.org/10.1111/j.2517-6161.1972.tb00899.x>

Cruz, D. O., & Ortega, J. C. (2008). *Análisis de la deserción estudiantil en la facultad de Ciencias Exactas y Naturales de la Universidad de Nariño desde la cohorte 2001-2 hasta la cohorte 2006-2 utilizando e*. Nariño: Universidad de Nariño. Obtenido de <https://www.semanticscholar.org/paper/An%C3%A1lisis-de-la-deserci%C3%B3n-estudiantil-en-la-facultad-DiegoCruz-Jos%C3%A9Ortega/27edacd473d16c30e2c64948aa9a7231b0041b1c>

Educacion, M. d. (2021). *REPORTE SOBRE LA*. Obtenido de <file:///C:/Users/Usuario/Downloads/Reporte%20sobre%20la%20interrupci%C3%B3n%20de%20estudios%20universitarios%20en%20el%20Per%C3%BA,%20en%20el%20contexto%20del%20COVID-19.pdf>: <file:///C:/Users/Usuario/Downloads/Reporte%20sobre%20la%20interrupci%C3%B3n%20de%20estudios%20universitarios%20en%20el%20Per%C3%BA,%20en%20el%20contexto%20del%20COVID-19.pdf>

García, C. (2014). *Estudio de Análisis de Supervivencia. Doble Grado en Ingeniería Informática y Matemáticas, Universidad Autónoma de Madrid, Madrid*. Madrid: https://repositorio.uam.es/bitstream/handle/10486/661556/pruena_garcia_hinojosa_cristina_tfg.pdf?sequence=1.

García, K., Rodríguez, J., Ferrer, Y., & Fragozo, Z. (2022). *DESERCIÓN ESTUDIANTIL EN COLOMBIA: UN ANÁLISIS DE LOS FACTORES PREDICTORES Y SUS INDICADORES*. Barranquilla- Colombia: Revista de Ciencias Humanas, Teoría Social y Pensamiento Crítico.

- Hernández, R., & Mendoza, C. (2018). *Metodología de la Investigación: Las rutas cuantitativa, cualitativa y Mixta* (Primera edición ed.). Mc Graw Hill Interamericana Editores. Obtenido de https://d1wqtxts1xzle7.cloudfront.net/64312353/Investigaci%C3%B3n_Rutas_cualitativa_y_cuantitativa-libre.pdf?1598813734=&response-content-disposition=inline%3B+filename%3DMETODOLOGIA_DE_LA_INVESTIGACION_LAS_RUTA.pdf&Expires=1748541910&Signature=U~SQmfP9yB
- Lopez, M. (2018). *Estrategia para la prevención del riesgo de deserción en los estudiantes en condiciones de pobreza de la Universidad Técnica de Manabí - Ecuador*. Lima Peru : https://cybertesis.unmsm.edu.pe/bitstream/handle/20.500.12672/6825/Lopez_pm.pdf?sequence=3&isAllowed=y.
- Medina, M., & Romero, T. (2020). *Modelo de supervivencia para analizar el abandono escolar de la URACCAN-Nueva Guinea*. Nueva Guinea: REVISTA CIENCIA E INTERCULTURALIDAD ISSN 1997-9231.
- Mondragón, J. (2013). *Una Aplicación de Intervalos de Confianza para la Mediana de Supervivencia en el Modelo de Regresión de Cox*. Tesis de maestría, Pontificia Universidad Católica del Perú , Lima. Lima-Peru: file:///H:/Libros%20para%20imprimir/MONDRAGON_JORGE_INTERVALOS_DE_CONFIANZA.pdf.
- Páramo, G., & Correa, C. (1999). Deserción Estudiantil Universitaria Conceptualización. *Revista Universidad Eafit* 66-78., <file:///C:/Users/USUARIO/Downloads/1075-1-3396-1-10-20120703.pdf>.

Paucar, J., & Quispe, M. (2021). *Análisis De Supervivencia: Deserción Estudiantil En La Escuela Profesional De Matemática De La Universidad Nacional De San Antonio Abad Del Cusco*.

Cusco-Peru:

file:///C:/Users/Administrador/Downloads/TESIS%20DE%20ESTADISTICA.pdf.

Paz, R. (2022). *Aplicación De Técnicas De Análisis De Supervivencia Al Estudio De La Deserción Estudiantil En Una Carrera De Ingeniería Civil*. Tucuman- Mexico:

file:///C:/Users/Administrador/Downloads/Aplicaci%C3%B3n+de+T%C3%A9cnicas+de+An%C3%A1lisis+de+Supervivencia+al+Estudio+de+la+Deserci%C3%B3n+Estudiantil+en+una+Carrera+de+Ingenier%C3%ADa+Civil.pdf.

Puma Laura, C. V. (2018). *Análisis de supervivencia de la deserción en los estudiantes de la Facultad de Ciencias Naturales y Formales a través del modelo de COX*. Arequipa:

Universidad Nacional de San Agustín de Arequipa. Obtenido de <https://repositorio.unsa.edu.pe/server/api/core/bitstreams/68c147f2-8c47-4f5e-9c50-062bd646627e/content>

UNESCO. (2021). file:///C:/Users/Usuario/Downloads/360-Texto%20del%20art%C3%ADculo-1846-1-10-20211203.pdf. Obtenido de file:///C:/Users/Usuario/Downloads/360-Texto%20del%20art%C3%ADculo-1846-1-10-20211203.pdf:

<file:///C:/Users/Usuario/Downloads/360-Texto%20del%20art%C3%ADculo-1846-1-10-20211203.pdf>

Silva Torres, A. (2020). Modelos paramétricos de regresión en análisis de supervivencia.

Universidad Nacional Autónoma de Mexico. Obtenido de:

<https://repositorio.unam.mx/contenidos/ficha/modelos-parametricos-de-regresion-en-analisis-de-supervivencia-317145>.

- Saldivar, R. (2020). *Factores Protectores Y De Riesgo De La Deserción De Estudiantes Del Segundo Semestre De La Escuela De Administración Industrial De Una Institución De Educación Superior*. Lima- Peru: https://repositorio.upch.edu.pe/bitstream/handle/20.500.12866/8981/Factores_SaldivarA%C3%B1azco_Ramon.pdf?sequence=1&isAllowed=y.
- San José, B., Pérez, E., & Madero, R. (2009). *Hablemos de... Métodos estadísticos en estudios de supervivencia*, . Obtenido de www.apcontinuada.com/es/pdf/80000481/: 55-59. .
- Sifuentes, O. (2018). *Modelo predictivo de la deserción estudiantil en una universidad privada del Perú*. Lima- Peru : <https://core.ac.uk/reader/323348780>.
- SUNEDU. (2021). *III Informe bienal sobre la realidad universitaria en el Perú*. Lima: Superintendencia Nacional de Educación Superior Universitaria. Obtenido de <https://cdn.www.gob.pe/uploads/document/file/3018068/III%20Informe%20Bienal.pdf>
- Tineo, F. (2005). *Estimación de Kaplan Meier Bootstrap de la Curva de Supervivencia*. Monografía para optar el Título de Licenciado en Estadística, Universidad Nacional Mayor de San Marcos, Lima.: http://sisbib.unmsm.edu.pe/bibvirtualdata/monografias/basic/tineo_gf/intro.pdf.
- Vivanco, M. (2005). *Muestreo estadístico, Diseño y aplicaciones*. Chile: Editorial Universitaria. Obtenido de https://books.google.com.co/books?id=-_gr5l3LbpIC&printsec=frontcover#v=onepage&q&f=false

ANEXOS

Anexo 1

Matriz de consistencia

ANÁLISIS DE DESERCIÓN ESTUDIANTIL EN LA ESCUELA PROFESIONAL DE MATEMÁTICA DE LA UNSAAC EN EL AÑO 2015 A 2020 MEDIANTE MODELOS DE SUPERVIVENCIA.

Problemas	Objetivos	Hipótesis	Variable	Dimensiones	Metodología
<p>Problema General ¿Cuál es el modelo de supervivencia paramétrico y no paramétrico, que mejor ajuste presenta al tiempo de estudios de los estudiantes antes de su deserción de la Escuela Profesional de Matemática, de la UNSAAC del año 2015 a 2020?</p> <p>Problemas específicos: PE1: ¿Cuáles es la función de riesgo del tiempo de estudios de los estudiantes antes de su deserción en la Escuela Profesional de Matemática de la UNSAAC del año 2015 a 2020?</p> <p>PE 2: ¿Cuál es la función de supervivencia del tiempo de estudios de los estudiantes antes de su deserción de la escuela Profesional de Matemática de la UNSAAC del año 2015 a 2020?</p> <p>PE 3: ¿Cuáles son los factores asociados del tiempo de estudios de los estudiantes antes de su deserción de la escuela profesional de matemática de la UNSAAC del año 2015 a 2020?</p>	<p>Objetivo general Modelar mediante métodos paramétricos y no paramétricos la supervivencia de los estudiantes antes de su deserción de la Escuela Profesional de Matemática de la UNSAAC del año 2015 a 2020.</p> <p>Objetivos específicos OE1: Determinar la función de riesgo del periodo de estudios de los estudiantes antes de su deserción de la Escuela Profesional de Matemática de la UNSAAC del año 2015 a 2020.</p> <p>OE2: Determinar la función de supervivencia del tiempo de estudios de los estudiantes antes de su deserción de la Escuela Profesional de Matemática de la UNSAAC del año 2015 a 2020.</p> <p>OE3. Determinar los factores asociados del tiempo de estudios de los estudiantes antes de su deserción de la Escuela Profesional de Matemática de la UNSAAC del año 2015 a 2020.</p>	<p>Hipótesis general El modelo no paramétrico Kaplan-Meier es el que mejor ajuste presenta al tiempo de estudios de los estudiantes antes de su deserción de la Escuela Profesional de Matemática de la UNSAAC del año 2015 a 2020.</p> <p>Hipótesis específicas HE1: La función de riesgo del tiempo de estudios de los estudiantes antes de su deserción de la Escuela Profesional de Matemática de la UNSAAC del año 2015 a 2020 es cambiante de manera ascendente con respecto a cada año.</p> <p>HE2: La función de supervivencia del período de estudio de los estudiantes antes de su deserción de la Escuela Profesional de Matemática de la UNSAAC entre 2015 a 2020 se modela mediante la función de Kaplan-Meier.</p> <p>HE3: Los factores asociados vinculados con la duración de los estudios de los estudiantes antes de su deserción de la Escuela Profesional de Matemática de la UNSAAC entre 2015 a 2020 incluyen la carga académica, limitaciones económicas y una orientación vocacional deficiente.</p>	<p>VARIABLE DEPENDIENTE: Deserción estudiantil</p>	<ul style="list-style-type: none"> • Personal • Académica • Socioeconómica • Institucional 	<p>Enfoque: Cuantitativa</p> <p>Tipo: Básica</p> <p>Diseño: No experimental</p> <p>Alcance: Explicativo</p> <p>Población: 273 estudiantes de la Escuela profesional de Matemática</p> <p>Muestra: 163 estudiantes Escuela profesional de Matemática.</p> <p>Técnica: Análisis documental</p> <p>Instrumento: Formulario de captura de datos</p>
			<p>VARIABLE INDEPENDIENTE: Modelos de supervivencia</p>	<ul style="list-style-type: none"> • Función de riesgo • Función de supervivencia • Factores asociados 	

Fuente: Elaboración propia

Anexo 3

Relación de estudiantes matriculados según el lugar de nacimiento.

N°	CODIGO		SEXO	EST.CIVIL	FECHA	LUGAR DE NACIMIENTO			DOMC. CUSCO
						DEPRT.	PROV.	DISTR.	
1		MATRICULADO	0	S	30/01/1998	CUSCO	CHUMBIVICAS	LLUSCO	SAN SEBASTIAN
2		MATRICULADO	1	S	31/10/1993	CUSCO	CANAS	LAYO	SANTIAGO
3		MATRICULADO	0	S	18/04/1998	CUSCO	CUSCO	CUSCO	SAN JERONIMO
4		MATRICULADO	1	S	17/09/1994	CUSCO	URUBAMBA	YUCAY	WANCHAQ
5		MATRICULADO	1	S	4/09/1998	CUSCO	LA CONVENCION	ECHARATE	WANCHAQ
6		MATRICULADO	0	S	16/10/1997	CUSCO	CALCA	LAMAY	SAN SEBASTIAN
7		MATRICULADO	1	S	24/02/1998	CUSCO	CUSCO	SANTIAGO	SANTIAGO
8		MATRICULADO	0	S	15/10/1997	CUSCO	CUSCO	SAN JERONIMO	SAN JERONIMO
9		MATRICULADO	1	S	4/11/1998	CUSCO	ACOMAYO	POMACANCHI	WANCHAQ
10		MATRICULADO	1	S	2/07/1998	CUSCO	CUSCO	SAN SEBASTIAN	SAN SEBASTIAN
11		MATRICULADO	1	S	1/03/1998	CUSCO	CANAS	YANAQCA	SAN JERONIMO
12		NO MATRICULADO	1	S	12/09/1998	CUSCO	CUSCO	CUSCO	CUSCO
13		MATRICULADO	1	S	11/09/1997	CUSCO	CHUMBIVICAS	SANTO TOMAS	SAN SEBASTIAN
14		MATRICULADO	1	S	16/12/1997	CUSCO	CUSCO	SAN JERONIMO	SAN JERONIMO
15		NO MATRICULADO	1	S	16/06/1997	CUSCO	LA CONVENCION	HUAYOPATA	SAN SEBASTIAN
16		NO MATRICULADO	0	S	13/03/1996	CALLAO	CALLAO	CALLAO	WANCHAQ
17		NO MATRICULADO	0	S	2/01/1999	CUSCO	CUSCO	WANCHAQ	WANCHAQ
18		MATRICULADO	0	S	29/03/1998	CUSCO	CHUMBIVICAS	LLUSCO	SAN SEBASTIAN
19		NO MATRICULADO	1	S	15/04/1998	CUSCO	CUSCO	SAN JERONIMO	SAN JERONIMO
20		MATRICULADO	1	S	12/01/1997	CUSCO	ACOMAYO	SANGARARA	SAN JERONIMO
21		MATRICULADO	0	S	29/09/1998	CUSCO	CUSCO	CUSCO	CUSCO
22		MATRICULADO	0	S	20/05/1997	CUSCO	CANAS	YANAQCA	WANCHAQ
23		MATRICULADO	1	S	13/07/1998	CUSCO	CUSCO	SAN SEBASTIAN	SAN SEBASTIAN
24		MATRICULADO	1	S	23/12/1994	CUSCO	CUSCO	WANCHAQ	WANCHAQ
25		MATRICULADO	1	S	13/11/1996	CUSCO	CHUMBIVICAS	LLUSCO	SAN SEBASTIAN

Anexo 5

Cuadro de estudiantes según el tipo de colegio.

COLEGIO CONCLUIDO				TIPO DE COLEGIO					
AÑO	DEPRT.	PROV.	DISTR.	ESTATAL	PARTIC.	MILITAR	MIXTO	VARONES	MUJERES
2014	CUSCO	CHUMBIVILCAS	LLUSCO	1			1		
2009	CUSCO	CANAS	LAYO	1			1		
2014	CUSCO	CUSCO	CUSCO	1					
2010	CUSCO	URUBAMBA	YUCAY	1			1		
2014	CUSCO	LA CONVENCION	ECHARATE	1			1		
2013	CUSCO	CALCA	LAMAY	1			1		
2014	CUSCO	CUSCO	SANTIAGO		1				
2013	CUSCO	CUSCO	SAN JERONIMO		1				
2014	CUSCO	ACOMAYO	POMACANCHI		1				
2014	CUSCO	CUSCO	SAN SEBASTIAN	1			1		
2014	CUSCO	CANAS	YANAOCA	1			1		
2014	CUSCO	CUSCO	CUSCO						
2013	CUSCO	CHUMBIVILCAS	SANTO TOMAS	1					
2013	CUSCO	CUSCO	SAN JERONIMO	1					
2013	CUSCO	LA CONVENCION	HUAYOPATA	1					
2012	CALLAO	CALLAO	CALLAO	1			1		
2014	CUSCO	CUSCO	WANCHAQ		1				
2014	CUSCO	CHUMBIVILCAS	LLUSCO	1					
2014	CUSCO	CUSCO	SAN JERONIMO		1				
2013	CUSCO	ACOMAYO	SANGARARA	1					
2014	CUSCO	CUSCO	CUSCO		1				
2013	CUSCO	CANAS	YANAOCA	1					
2014	CUSCO	CUSCO	SAN SEBASTIAN	1					
2010	CUSCO	CUSCO	WANCHAQ	1					
2012	CUSCO	CHUMBIVILCAS	LLUSCO	1					

Anexo 7

Cuadro de situación económica de los estudiantes.

SITUACION ECONOMICA						
PEN. MENSUAL	P. SOSTIENE ECON. PADRE		P. SOSTIENE ECON. MADRE		P. SOSTIENE ECON. HERM	
	OCUPACION	INGRESO	OCUPACION	INGRESO	OCUPACION	INGRESO
200	AGRICULTOR	500				
150	ALBAÑIL	1000				
400	PROFESOR	1050				
150	CONDUCTOR	850				
200	COMERCIANTE	1200	COMERCIANTE	800		
180	OBRERO	1500				
320	TAXISTA	1300				
450	CONSTRUCCION	1500				
200	CONSTRUCCION	1100				
350			AMA DE CASA	850		
220	CONDUCTOR	950				
350	ALBAÑIL	900				
120	AGRICULTOR	600				
200	TAXISTA	1200				
320	CONSTRUCCION	1050				
600	PROFESOR	1500				
500	INGENIERO	1900				
300	AGRICULTOR	700				
400	CONSTRUCCION	1200				
300	AGRICULTOR	900				
450	ALBAÑIL	1500				
350	CONDUCTOR	1100				
500	OBRERO	1400				
300	CONSTRUCCION	1200				ALUMNO WV
400	OBRERO	1000				Va a Confian

Anexo 7

Cuadro según la vivienda que ocupan los estudiantes.

VIVIENDA													
LA CASA QUE OCUPA ES:					MAT. DE CONSTRUC.		SERVI. CON QUE CUENTA			DATOS DEL LUGAR DE PROC.		MAT. DE CONSTRUC.	
PROPIA	ALQUILADO	CUIDANTE	ALOJADO	OTROS	NOBLE	ADOBE	AGUA	LUZ	DESAGUE	VIVI. PROPIA	VIV. ALQUILAD	NOBLE	ADOBE
	1				1		1	1	1	1			1
	1				1		1	1	1	1		1	
1					1		1	1	1	1		1	
	1				1		1	1	1	1			1
	1				1		1	1	1	1			1
	1				1		1	1	1	1		1	
1						1	1	1	1	1			1
1					1		1	1	1	1		1	
	1				1		1	1	1	1			1
1						1	1	1	1	1			1
	1				1		1	1	1	1			1
1						1	1	1	1	1			1
	1				1		1	1	1	1			1
	1				1		1	1	1	1			1
1						1	1	1	1	1		1	
	1				1		1	1	1	1			1
1						1	1	1	1	1			1
	1				1		1	1	1	1			1
1						1	1	1	1	1			1
1					1		1	1	1	1			1
	1				1		1	1	1	1			1

Anexo 9

Ficha de SOCIO-ECONOMICA de los estudiantes.

UNIVERSIDAD NACIONAL DE SAN ANTONIO ABAD DEL CUSCO
 UNIDAD DE BIENESTAR UNIVERSITARIO AREA DE ASISTENCIA SOCIAL

FICHA SOCIO-ECONÓMICA

Escuela Profesional: Matemática Código matrícula: 192001

1. DATOS PERSONALES

Apellido paterno: <u>Ramos</u>	Apellido materno: <u>Huaycha</u>	Nombre(s): <u>Andrés Michel</u>
Lugar de Nacimiento: <u>Cuzco</u>	<u>Puno</u>	<u>Jaurisque</u> <u>Monique - Banos</u>
Departamento:	Provincia:	Distrito: Comun. Sector Otros:
Fecha de Nacimiento: <u>04</u> <u>Mayo</u> <u>2001</u>	Sexo: M <input checked="" type="checkbox"/> F ()	
Día:	Mes:	Año:
Estado Civil: soltero () casado <input checked="" type="checkbox"/> otro especificar:		DNI: <u>75761503</u>
Domicilio en la ciudad del cusco: <u>Santiago</u>		<u>camino de Dios</u>
Nº CEL: <u>982819095</u>	Distrito:	Urban. Av. Calle, Jr. Nº

2. DATOS ACADÉMICOS:

Año que concluyó secundaria: <u>2017</u>	<u>Cuzco</u>	<u>Puno</u>	<u>Paucartambo</u>
Departamento:	Provincia:	Distrito:	
Nombre del colegio: <u>CRFA Ayurkaway Huaywasin</u>			
Tipo del colegio: Estatal <input checked="" type="checkbox"/> Particular () Militar ()			
Mixto <input checked="" type="checkbox"/> Varones () Mujeres ()			
Año que ingreso a la universidad: semestre I <input checked="" type="checkbox"/> Semestre II ()			
Modalidad de ingreso:			
Primera opción () dirimencia () CEPRU <input checked="" type="checkbox"/> examen ordinario () seguida profesión ()			
Deportista () otros especifique:			

3. COMPOSICIÓN FAMILIAR:

PARENTESCO	APELLIDOS Y NOMBRES	EDAD	ESTADO CIVIL	GRADO DE INSTRUCCIÓN	OCCUPACIÓN	INGRESO MENSUAL
PADRES	Ramos Zavate Alejandro	51	Casado	2º de secundaria	Agricultor	3/50
MADRE	Huaycha Ramos Zoraida	47	Casada	3º de secundaria	Agricultora	3/50
HERMANOS	Ramos Huaycha Douglas Paul	27	Soltero	Univer.	Eng. Zootecnia	—
	Ramos Huaycha Isay Dora	17	Soltera	5º de secundaria		
	Ramos Huaycha Ayta Dora	11	Soltera	5º primari.		
CONYUGUE						
HIJOS						

En caso de fallecimiento de los padres, indique:

Anexo 10

Códigos de R-estudios

```

survfit(Surv(Tiempo, Condicion) ~ ModalidaIngreso, datos,
        conf.type = "log-log") %>% ggsurvplot(title = "Supervivencia por moda
        lidad de ingreso",
        conf.int = T, legend.title = "ModalidaIngreso", legend.labs = c("Primera_opor
        tunidad", "Dirimencia", "CEPRU", "Ordinario"))

#Procedemos a usar La Sexo como covariable
mie_km_prot <- survfit(Surv(Tiempo, Condicion)~Sexo,data = datos)
summary(mie_km_prot)

#Graficas de supervivencia
# supervivencia por sexo

survfit(Surv(Tiempo, Condicion)~Sexo, datos,
        conf.type = "log-log") %>% ggsurvplot(title = "Supervivencia por Géne
        ro",
        conf.int = T, legend.title = "Género", legend.labs = c("Hombres", "Mu
        jeres"))

# supervivencia por modalidad de procedencia

survfit(Surv(Tiempo, Condicion)~Procedencia, datos,
        conf.type = "log-log") %>% ggsurvplot(title = "Supervivencia por proc
        edencia",
        conf.int = T, legend.title = "P
        rocedencia", legend.labs = c("Cusco", "Otros"))

# supervivencia por modalidad de tipo de dependencia economica

survfit(Surv(Tiempo, Condicion)~Depende, datos,
        conf.type = "log-log") %>% ggsurvplot(title = "Supervivencia por depe
        ndencia economica",
        conf.int = T, legend.title = "D
        ependencia economica", legend.labs = c("Padres", "Del mismo"))

# supervivencia por tipo de colegio

survfit(Surv(Tiempo, Condicion)~TipoColegio, datos,
        conf.type = "log-log") %>% ggsurvplot(title = "Supervivencia por Tipo
        de Colegio",
        conf.int = T, legend.title = "T
        ipo de Colegio", legend.labs = c("Estatal", "Particular"))

```

```

# supervivencia por covariables sexo y procedencia

survfit(Surv(Tiempo, Condicion) ~ Sexo + Procedencia, datos, conf.type = "log-log") %>%
  ggsurvplot(title = "Supervivencia entre Género por procedencia", conf.int = T,
    facet.by = "Procedencia", legend.title = "Género", legend.labs = c("Hombres", "Mujeres"),
    panel.labs = list(Grupo_procedencia = c("Cusco", "Otros")), short.panel.labs = T)

# Guardando el Objeto Surv  time=tiempo, delta=Condicion, type=ModalidaIngreso, datas=datos

tongue.surv <- Surv(Tiempo, Condicion) #Creando objeto tipo Surv
tongue.km <- survfit(tongue.surv ~ 1, data = datos, type = "kaplan-meier")
#Estimación Kaplan Meier

# Sin Guardar el Objeto Surv
tongue.km <- survfit(Surv(Tiempo, Condicion) ~ 1, data = datos, type = "kaplan-meier")

summary(tongue.km)

# FORMA 2
tongue.km <- survfit(Surv(Tiempo, Condicion) ~ ModalidaIngreso, data = datos, type = "kaplan-meier") #Estimación Kaplan Meier por grupos
summary(tongue.km) #Resumen estadístico

# Intervalos de confianza de la estimación de la función de supervivencia
tongue.km <- survfit(Surv(Tiempo, Condicion) ~ 1, data = datos, type = "kaplan-meier",
  error = "tsiatis", conf.type = "log-log", conf.int = 0.99)

summary(tongue.km)

#Graficación de la curva de supervivencia
ggsurvplot(fit = tongue.km, data = datos, conf.int = T, title = "Curva de Supervivencia",
  xlab = "Tiempo", ylab = "Probabilidad de supervivencia", legend.title = "Estimación",
  legend.labs = "Kaplan-Meier")

```

#Estimación de la función de riesgo acumulado

```
tongue.km <- survfit(Surv(Tiempo, Condicion) ~ 1, datos)
R <- tongue.km %>% fortify %>% mutate(CumHaz = cumsum(n.event/n.risk))
R

#Estimación de La función de riesgo acumulado CON COVARIABLES
tongue.km <- survfit(Surv(Tiempo, Condicion) ~ ModalidaIngreso, datos)
R <- tongue.km %>% fortify %>% group_by(strata) %>% mutate(CumHaz = cumsum(n
.event/n.risk))
print(R, 60)

plot(tongue.km, fun = "cumhaz", conf.int = F, main = "Riesgo Acumulado", col
= 1:5,
      xlab = "Tiempo (Meses)", ylab = "Riesgo Acumulado")
```

```
ggsurvplot(tongue.km, fun = "cumhaz", xlab = "Tiempo (Meses)",
           censor = T,
           ylab = "Riesgo Acumulado",
           title = "Riesgo Acumulado",
           legend.title = "ModalidaIngreso",
           legend.labs = c("A", "B", "C", "D"))
```

```
table(ModalidaIngreso)
```

```
## ModalidaIngreso
##  1  2  3  4
## 28 38 32 65
```

#Estimación de La media, mediana y percentiles de Los tiempos de supervivencia.

```
print(tongue.km, print.rmean = TRUE)
```

#Estimación Paramétrica de La función de supervivencia

```
flex <- flexsurvreg(Surv(Tiempo, Condicion) ~ 1, data = datos, dist = "exp")
#Ajuste exponencial
flex
```

#riesgo acumulado

```
plot(flex, type = "cumhaz", ci = F, main = "Riesgo Acumulado", conf.int = F,
     col = 3, col.obs = 4, xlab = "Tiempo", ylab = "Riesgo Acumulado")
```

```
#prob supervivencia
plot(flex, type = "survival", ci = F, main = "Probabilidad de Supervivencia"
,
      conf.int = F, col = 1, col.obs = 2, xlab = "Tiempo", ylab = "Probabilidad de Supervivencia")
```

```
# comparacion parametrica y no parametrica

flexgg <- flexsurvreg(Surv(Tiempo, Condicion) ~ 1, data = datos, dist = "exp") %>%
  summary(type = "survival") %>% data.frame
kmgg <- survfit(Surv(Tiempo, Condicion) ~ 1, data = datos) %>% fortify
ggplot() + geom_line(aes(time, est, col = "Paramétrico"),
                     data = flexgg) + geom_step(aes(time,
                                                    surv, col = "No Paramétrico"), data = kmgg) + labs(x = "Tiempo (Semanas)",
y = "Probabilidad de Supervivencia", col = "Ajustes",
title = "Comparación entre curvas de supervivencia")
```

```
#riesgo acumulado

flexch <- flexsurvreg(Surv(Tiempo, Condicion) ~ 1, data = datos, dist = "exp") %>%
  summary() %>% data.frame

kmch <- survfit(Surv(Tiempo, Condicion) ~ 1, data = datos) %>% fortify

ggplot() + geom_line(aes(time, -log(est), col = "Paramétrico"), data = flexch) +
  geom_step(aes(time, -log(surv), col = "No Paramétrico"), data = kmch) +
  labs(x = "Tiempo (Semanas)", y = "Riesgo Acumulado", col = "Ajustes", title = "Comparación entre Riesgo acumulado")
```

```
# COMPARAC DE MODELOS
datosflex <- datos
names(datosflex)
```

```
Dist <- c("exp", "weibull", "llogis", "gompertz", "genf")
```

```
data.Surv <- Surv(datosflex$Tiempo, datosflex$Condicion)
```

```
# Inicia SCRIPT#
```

```

model <- sapply(Dist, function(x) flexsurvreg(data.Surv ~ 1, dist = x), USE.N
AMES = T,
              simplify = F)

# curva supervivencia parametrica

plot(model[[1]], ci = F, conf.int = F, lty = 2, main = "Ajuste Paramétrico",
      xlab = "Tiempo", ylab = "Probabilidad de Supervivencia")
for (i in 2:length(Dist)) plot(model[[i]], ci = F, conf.int = F, add = T, col
l = i +
                                1, lty = i)

legend("topright", c("KM", Dist), lty = 1:(length(Dist) + 1), col = 1:(length
h(Dist) +
1))

```

```

# curva supervivencia parametrica y no parametrica

KM <- survfit(data.Surv ~ 1, data = datosflex)
KMdist <- KM %>% fortify()

Flexdist <- names(model) %>% sapply(function(x) cbind(Dist = x, data.frame(s
ummary(model[[x]]))),
                                simplify = F) %>% do.call("rbind", .)

x <- (1:length(Dist)) + 1 #Colores
names(x) <- Dist #Cambiar si se Desea

ggplot() + geom_line(aes(time, est, col = factor(Dist)), Flexdist, size = 0.
9,
                    alpha = 0.5) + geom_step(aes(time, surv, col = "No Para
metrico", group = 1),
                                             KMdist, size = 1, alpha = 0.9)
+ scale_color_manual(values = c(`No Parametrico` = "black",
x)) + labs(col = "Distribuciones", linetype = "Ajustes", x = "Tiempo(Semanas)
",
y = "Probabilidad de Supervivencia", title = "Ajustes Paramétrico vs No Param
étrico")

```

```
#Riesgo acumulado parametrico

plot(model[[1]], ci = F, conf.int = F, lty = 2, main = "Ajuste Paramétrico",
      xlab = "Tiempo", ylab = "Riesgo Acumulado", type = "cumhaz")
for (i in 2:length(Dist)) plot(model[[i]], ci = F, conf.int = F, add = T, col = i +
                               1, lty = i, type = "cumhaz")

legend("topright", c("KM", Dist), lty = 1:(length(Dist) + 1), col = 1:(length(Dist) +
                                                                              1)
)
```

```
# riesgo acumulado parametrico y no parametrico

KM <- survfit(data.Surv ~ 1, data = datosflex)
KMdist <- KM %>% fortify()

Flexdist <- names(model) %>% sapply(function(x) cbind(Dist = x, data.frame(summary(model[[x]]))),
                                   simplify = F) %>% do.call("rbind", .)

x <- (1:length(Dist)) + 1 #Colores
names(x) <- Dist #Cambiar si se Desea

ggplot() + geom_line(aes(time, -log(est), col = factor(Dist)), Flexdist, size = 0.9,
                      alpha = 0.5) + geom_step(aes(time, -log(surv), col = "No Parametrico", group = 1),
                                                KMdist, size = 1, alpha = 0.9)
+ scale_color_manual(values = c(`No Parametrico` = "black",
                                x)) + labs(col = "Distribuciones", linetype = "Ajustes", x = "Tiempo(Semanas)",
y = "Riesgo Acumulado", title = "Ajustes Paramétrico vs No Paramétrico")
```

```
#####Riesgo proporcionales de Cox#####

# regresion con variables numericas
coxph(Surv(Tiempo, Condicion) ~ Edad+RendAcademico + PensionMensual +IngrFamiliar , data = datos)
```

```
# regresion con variables cualitativas
datos$Sexo <- factor(datos$Sexo , labels = c("Varon", "Mujer"))
datos$Procedencia <- factor(datos$Procedencia , labels = c("Cusco",
"Otros"))
datos$TipoColegio <- factor(datos$TipoColegio , labels = c("Estatal"
, "Particular"))
datos$Depende <- factor(datos$Depende , labels = c("Padres", "Del
mismo"))
datos$OcupaJefe <- factor(datos$OcupaJefe , labels = c("Profesiona
l", "No profesional"))
datos$CasaOcupa <- factor(datos$CasaOcupa , labels = c("Propio",
"Alquilado"))
datos$ModalidaIngreso <- factor(datos$ModalidaIngreso , labels = c("
Primera oportunidad", "Dirimencia", "CEPRU", "Ordinario" ))
attach(datos)
```

```
# regresion con variables cualitativas y cuantitativas
```

```
coxph(Surv(Tiempo, Condicion) ~ Edad * Sexo, data = datos)
```

Anexo 11

Respuesta de conformidad ante la solicitud de datos socioeconómicos

	UNIVERSIDAD NACIONAL DE SAN ANTONIO ABAD DEL CUSCO DIRECCION DE BIENESTAR UNIVERSITARIO UNIDAD DE ASISTENCIA SOCIAL	
<p>"AÑO DEL BICENTENARIO, DE LA CONSOLIDACIÓN DE NUESTRA INDEPENDENCIA, Y DE LA CONMEMORACIÓN DE LAS HEROICAS BATALLAS DE JUNÍN Y AYACUCHO"</p>		
<p>INFORME N° 002 - 2024 – UAS – DBU – UNSAAC</p>		
A	: Dra. ROXANA JULIA ABARCA ARRAMBIDE Directora de la Oficina de Bienestar Universitario.	
DE	: Asist. Soc. FELICIA FORTUNATA PUENTE DE LA VEGA TINTAYA Coordinadora de la Unidad de Asistencia Social.	
REFERENCIA	: MEMORANDO N° 150 -DBU-2023-UAS	
ASUNTO	: COMUNICA FACILIDADES OTORGADAS A ESTUDIANTES DE ESCUELA PROFESIONAL DE MATEMÁTICA	
FECHA	: 04 de enero del 2024	

Previo cordial saludo me dirijo a su jefatura y en atención al documento de la referencia, informo que se brindo facilidades a los estudiantes tesisistas de la Escuela Profesional de Matemática, quienes solicitaron en forma insistente a esta Unidad información de las fichas socioeconómicas de los estudiantes, información de carácter privado y confidencial que corresponde a cada estudiante. Mas, sin embargo, se les brindo la información de acuerdo a los solicitado a los bachilleres: Demetrio Paucar Quispe y Luis Alberto Huaman Huarco, quienes quedaron satisfechos y podrán continuar con su trabajo de investigación titulado "ANÁLISIS DE DESERCIÓN ESTUDIANTIL EN LA ESCUELA PROFESIONAL DE MATEMÁTICAS DE LA UNSAAC EN EL AÑO 2015 HASTA EL 2020"

es cuanto informo a Ud. para su conocimiento y fines consiguientes, es ope se deba comunicar al director de la Escuela Profesional de Matemática.

Atentamente,


 Lic. Felicia Fortunata Puente de la Vega Tintaya
 ASISTENTE SOCIAL
 COORDINADORA

Anexo 12

Respuesta de conformidad ante la solicitud de datos socio-económicos

589023

UNIVERSIDAD NACIONAL DE SAN ANTONIO ABAD DEL CUSCO
DIRECCIÓN DE BIENESTAR UNIVERSITARIO
"AÑO DE LA UNIDAD, LA PAZ Y EL DESARROLLO"

Cusco 05 de enero de 2024

Oficio N° 02-2024-DBU/UNSAAC

Señor
Dr. Epifanio Puma Huatuc
Director De La Escuela Profesional De Matematica

Presente

ASUNTO : EN RESPUESTA A SU SOLICITUD DE BRINDAR FACILIDADES A ALUMNOS DE LA ESCUELA PROFESIONAL DE MATEMÁTICA.

REF : Of. Circ. N° FCQFM-EPM-148-2023 y Exp. 589023.

Previa un cordial saludo, me dirijo a usted, en atención al expediente de la referencia, documento mediante el cual, solicita brindar facilidades a alumnos de la escuela profesional de matemática.

Con Of. Circ. N° FCQFM-EPM-148-2023, de fecha 15 de noviembre de 2023.

Se remite en respuesta a su solicitud el INFORME N° 002 - 2024 - UAS - DBU - UNSAAC emitido por la Unidad de Asistencia Social el 04 de enero del presente año.

En el cual informan que "se brindó facilidades a los estudiantes tesis de la Escuela Profesional de Matemática, quienes solicitaron en forma insistente a esta Unidad información de las fichas socioeconómicas de los estudiantes, información de carácter privado y confidencial que corresponde a cada estudiante. Mas, sin embargo, se les brindo la información de acuerdo a los solicitado a los bachilleres Demetrio Paucar Quispe y Luis Alberto Huaman Huarco, quienes quedaron satisfechos y podrán continuar con su trabajo de investigación titulado "ANÁLISIS DE DESERCIÓN ESTUDIANTIL LA ESCUELA PROFESIONAL DE MATEMÁTICAS DE LA UNSAAC EN EL AÑO 2015 HASTA EL 2020"

Sin otro particular, uso de la oportunidad para expresar las muestras de mi distinción personal.

Atentamente,



AGENCIACIÓN DE SERVICIOS
Dr. ROMANA SEARLA ABRAMIDE
 DIRECTORA

Cc
 Archivo
 DBU
 RJAA/ldtb

Anexo13

Solicitando datos socio-económicos, dirección de bienestar universitario.



*Anexo 14**Dirección de bienestar universitario – información facilitada.*

Anexo 15:

Dirección de bienestar universitario – información facilitada.



*Anexo 16**Procesamiento de data socio- económicos en R-Studio.*