

**UNIVERSIDAD NACIONAL DE SAN ANTONIO ABAD DEL
CUSCO
FACULTAD DE CIENCIAS
ESCUELA PROFESIONAL DE MATEMÁTICA**



TESIS:

**“ANÁLISIS DE SUPERVIVENCIA: DESERCIÓN ESTUDIANTIL EN LA
ESCUELA PROFESIONAL DE MATEMÁTICA DE LA UNIVERSIDAD NACIONAL
DE SAN ANTONIO ABAD DEL CUSCO”**

PRESENTADO POR:

BR.: JENNIFER PAUCAR AMAR

BR.: MARISEL GISELA QUISPE YMA

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

ASESORA:

MGT. RINA MARICELA ZAMALLOA CORNEJO

CUSCO – PERU

2021

DEDICATORIA

A Dios mi Padre Celestial, creador de todas las cosas y fuente de sabiduría, por permitirme haber llegado hasta este momento tan importante de mi formación profesional. Con mucho cariño, a mis padres Hilario Quispe Challco y Juana Yma Mamani que son de mucha bendición para mí, quienes con mucho esfuerzo y de manera incondicional me brindaron su apoyo para poder concluir con satisfacción este presente trabajo. Con gratitud a mi hermano Ronald, por su apoyo incondicional en los momentos difíciles, amistad en todo tiempo, animarme a perseverar y nunca rendirme en el camino de mi formación profesional.

Marisel Gisela

Con mucho amor y cariño dedico este trabajo a Dios padre todo poderoso y a mi mamita Virgen Purificada de Saylla, por darme mucha sabiduría, fuerza y guiarme siempre para lograr mis metas trazadas, con mucho respeto y amor a mis padres que me dieron la vida Sandra Amar Mormontoy y Ronald Paucar Cana, por haberme brindado todo su apoyo, a mis hermanos, Arnold, Ronald y Linda por apoyarme y tenerme paciencia, a mis abuelos y a toda mi familia que siempre me han apoyado y aconsejado en mi formación profesional y en la realización del presente trabajo.

Jennifer

AGRADECIMIENTO

A Dios, nuestro eterno amigo que nos dio la vida, sabiduría, salud y siempre nos ha dado la fortaleza para seguir adelante en momentos difíciles y valor para culminar el presente trabajo de investigación.

A todos los docentes del Departamento Académico de Matemática y Estadística de la Universidad Nacional San Antonio Abad del Cusco, que nos han compartido sus conocimientos teóricos- prácticos, dedicación y motivación durante nuestra formación profesional.

A nuestra asesora, Mgt. Rina Maricela Zamalloa Cornejo por habernos guiado durante la ejecución del presente trabajo de investigación, nuestra admiración hacia ella por ser una gran profesional y extraordinaria persona así mismo a nuestra docente Dra. Natalie Veronika Rondinel Mendoza por los conocimientos compartidos incondicionalmente para la culminación de la tesis.

Al Centro de Cómputo y la Oficina de Bienestar Universitario de la Universidad Nacional San Antonio Abad del Cusco agradecemos por la confianza, en brindarnos la base de datos de la cohorte 2010-I al 2012-II de la escuela profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco, con su apoyo se logró la captura de datos que han sido útiles en la aplicación de nuestro trabajo de investigación.

Las tesis

INDICE

INTRODUCCIÓN	1
1. PLANTEAMIENTO DEL PROBLEMA	3
1.1. Caracterización del problema.....	3
1.2. Formulación del problema	5
1.2.1. Problema general	5
1.2.2. Problemas específicos.....	5
1.3. Objetivo.....	5
1.3.1. Objetivo general.....	5
1.3.2. Objetivos específico.....	6
1.4. Justificación:.....	6
1.5. Limitación del Trabajo	6
1.6. Hipótesis.....	7
1.6.1. Hipótesis general.....	7
1.6.2. Hipótesis específica	7
2. MARCO TEÓRICO	8
2.1. Antecedentes de la investigación	8
2.1.1. Antecedentes internacionales.....	8
2.1.2. Antecedentes nacionales	10
2.1.3. Antecedentes locales.....	12

2.2.	Análisis de Supervivencia	15
2.2.1.	Definiciones de términos utilizados en el análisis de supervivencia	15
2.2.2.	Elementos Básicos	17
2.2.3.	Función de supervivencia	17
2.2.4.	Función de riesgo.....	18
2.2.5.	Censura	19
2.2.5.1.	Censura por la derecha	19
2.2.5.2.	Censura por izquierda.....	20
2.2.5.3.	Censura por intervalo	20
2.3.	Método no Paramétrico	20
2.3.1.	Estimador de Kaplan-Meier.....	20
2.3.2.	Varianza del Estimador Kaplan-Meier	21
2.3.3.	Comparación funciones de supervivencia	26
2.3.3.1.	Prueba de hipótesis para la función de supervivencia.....	26
2.3.3.2.	Prueba de Log- Rank.....	26
2.4.	Método Semi Paramétrico.....	29
2.4.1.	Modelo de Regresión de Cox.....	29
2.4.2.	Estimación de los parámetros en el modelo de Cox	32
2.4.2.1.	Estimación de β	32
2.4.3.	Estimación de la función de riesgo base $h_0(t)$	34

2.4.4.	Prueba de Wald.....	36
2.5.	DESERCIÓN ESTUDIANTIL	36
2.5.1.	Definición de Deserción	36
2.5.2.	Clases y actores de Deserción.....	37
2.5.3.	Factores asociados.	37
2.5.3.1.	<i>Factores individuales</i>	38
2.5.3.2.	Factores socioeconómicos.....	38
2.5.3.3.	Factores académicos.....	38
2.5.3.4.	Factores institucionales	38
3.	DISEÑO METODOLÓGICO.....	39
3.1.	Tipo de Investigación.....	39
3.2.	Diseño de Investigación	39
3.3.	Población.....	40
3.4.	Variables de estudio	41
3.4.1.	Variable dependiente	41
3.4.2.	Variable independiente	41
3.5.	Técnicas e instrumentos de recolección de datos.....	41
3.6.	Procedimiento de recolección de datos	42
3.7.	Técnicas de Procesamiento y Análisis de datos	43
3.8.	Equipos y Software	43

4.	RESULTADOS	44
4.1.	Resultado del Modelo de Kaplan Meier.....	44
4.1.1.	Deserción Estudiantil en la Escuela Profesional de Matemática (EPM)	44
4.1.1.1.	Comparación de Curvas de Supervivencia.....	47
4.2.	Resultado del Modelo de Regresión de Cox	89
	CONCLUSIONES	97
	RECOMENDACIONES.....	98
	REFERENCIA.....	99
	ANEXO	104

INDICE DE TABLAS

Tabla 1 Deserción del estudiante de pre-grado año 2015	14
Tabla 2 Deserción estimada del estudiante de pregrado año 2016	14
Tabla 3 Prueba de igualdad de la función de supervivencia de dos grupos en el tiempo de observación t_i	27
Tabla 4 Prueba de igualdad de la función de supervivencia de M grupos de observación t_i	28
Tabla 5 Número de estudiantes que ingresaron a la Escuela Profesional de Matemática según semestre de ingreso	40
Tabla 6 Resumen de Kaplan Meier (EPM)	44
Tabla 7 Resumen de Kaplan Meier (EPM)	45
Tabla 8 Resumen de Sexo	47
Tabla 9 Los tiempos de supervivencia, la proporción de estudiantes sobrevivientes en cada punto del tiempo en semestres para la variable sexo (mujer-varón).....	48
Tabla 10 Comparaciones Globales.....	49
Tabla 11 Resumen Tipo de Colegio	51
Tabla 12 Los tiempos de supervivencia, la proporción de estudiantes sobrevivientes en cada punto del tiempo en semestres para la variable tipo de colegio (estatal y particular).....	52
Tabla 13 Comparaciones Globales.....	53
Tabla 14 Resumen Tipo de Modalidad de Ingreso.....	55
Tabla 15 Los tiempos de supervivencia, la proporción de estudiantes sobrevivientes en cada punto del tiempo en semestres para tipo de modalidad de ingreso (CEPRU, dirimencia, ordinario y primera oportunidad).....	56
Tabla 16 Comparaciones Globales.....	57
Tabla 17 Resumen de Ocupación del Jefe de Hogar	59
Tabla 18 Los tiempos de supervivencia, la proporción de estudiantes sobrevivientes en cada punto del tiempo en semestres para la variable ocupación del jefe de hogar (no profesional y profesional).....	60
Tabla 19 Comparaciones Globales.....	61
Tabla 20 Resumen de Ingreso Económico familiar	63

Tabla 21 los tiempos de supervivencia, la proporción de estudiantes sobrevivientes en cada punto del tiempo en semestres para la variable ingreso económico familiar (sueldo alto, sueldo básico y sueldo medio).....	64
Tabla 22 Comparaciones Globales.....	65
Tabla 23 Resumen del Número de Hermanos	67
Tabla 24 Los tiempos de supervivencia, la proporción de estudiantes sobrevivientes en cada punto del tiempo en semestres para la variable número de hermanos (0-1, 2-3 y 3 a más).....	68
Tabla 25 Comparaciones Globales.....	69
Tabla 26 Resumen de Situación Familiar.....	71
Tabla 27 Los tiempos de supervivencia, la proporción de estudiantes sobrevivientes en cada punto del tiempo en semestres para la variable situación familiar (buena, regular, mala y separados).....	72
Tabla 28 Comparaciones Globales.....	72
Tabla 29 Resumen del Lugar de Procedencia	74
Tabla 30 Los tiempos de supervivencia, la proporción de estudiantes sobrevivientes en cada punto del tiempo en semestres para la variable lugar de procedencia (Cusco y otros).....	75
Tabla 31 Comparaciones Globales.....	76
Tabla 32 Resumen de Estudiante que Trabaja	78
Tabla 33 Los tiempos de supervivencia, la proporción de estudiantes sobrevivientes en cada punto del tiempo en semestres para la variable el estudiante trabaja (si, no)	79
Tabla 34 Comparaciones Globales.....	79
Tabla 35 Resumen de Edad.....	81
Tabla 36 Los tiempos de supervivencia, la proporción de estudiantes sobrevivientes en cada punto del tiempo en semestres para la variable edad (menor que 22 años y mayor o igual que 22 años).....	82
Tabla 37 Comparaciones Globales.....	83
Tabla 38 Resumen de los resultados de Kaplan-Meier	86
Tabla 39 Coeficiente de la regresión de Cox para las variables en la ecuación	90
Tabla 40 Comparación Global.....	92
Tabla 41 Resumen de los resultados de la regresión de Cox.....	95

INDICE DE FIGURAS

Figura 1 Función de supervivencia de la EPM.....	46
Figura 2 Función de riesgo de la EPM.....	47
Figura 3 Función de supervivencia Sexo	50
Figura 4 Función de riesgo de Sexo	51
Figura 5 Función de supervivencia tipo de Colegio	54
Figura 6 Función de Riesgo Tipo de Colegio.....	55
Figura 7 Función de supervivencia Tipo de Modalidad de Ingreso	58
Figura 8 Función de Riesgo Tipo de Modalidad de Ingreso	59
Figura 9 Función de Supervivencia Ocupación del jefe de hogar.....	62
Figura 10 Función de Riesgo Ocupación del jefe de hogar.....	63
Figura 11 Función de Supervivencia Ingreso Económico familiar.....	66
Figura 12 Función de Riesgo Ingreso Económico Familiar	67
Figura 13 Función de Supervivencia Número de Hermanos.....	69
Figura 14 Función de Riesgo Número de Hermanos	70
Figura 15 Función de Supervivencia Situación Familiar	73
Figura 16 Función de Riesgo Situación Familiar	74
Figura 17 Función de Supervivencia Lugar de Procedencia.....	77
Figura 18 Función de Riesgo Lugar de Procedencia	78
Figura 19 Función de Supervivencia el Estudiante Trabaja (no y sí).....	80
Figura 20 Función de Riesgo el Estudiante Trabaja (no y sí)	81
Figura 21 Función de Supervivencia Edad	84
Figura 22 Función de Riesgo Edad.....	85
Figura 23 Grafico de las Variables de la Regresión de Cox.....	93

PRESENTACIÓN

Sr. Decano de la Facultad de Ciencias.

Dr. Olintho Aguilar Condemayta

Presentamos el trabajo de investigación de tesis intitulado “ANÁLISIS DE SUPERVIVENCIA: DESERCIÓN ESTUDIANTIL EN LA ESCUELA PROFESIONAL DE MATEMÁTICA DE LA UNIVERSIDAD NACIONAL DE SAN ANTONIO ABAD DEL CUSCO” con el fin de optar al Título Profesional de Licenciadas en Matemática mención Estadística.

El objetivo de este trabajo de investigación es utilizar los modelos estadísticos no paramétrico y semi paramétrico relacionados con la deserción estudiantil de la Escuela Profesional de Matemática, lo que permitirá conocer el porcentaje de estudiantes que abandonan sus estudios Universitarios y los factores que influyen a tomar dicha decisión.

Sugiere el trabajo utilizar el análisis de supervivencia, el estimador de Kaplan-Meier y el modelo de Cox para hacer el análisis de deserción en los estudiantes de la Escuela Profesional de Matemática.

Estamos convencidas que este trabajo de investigación será de suma utilidad para el director de la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco, ayudará a tomar medidas pertinentes y necesarias para disminuir la cantidad de estudiantes desertores, al conocer los factores asociados a la deserción. Así mismo expresamos nuestra gratitud a los docentes del Departamento Académico de Matemática y Estadística por los conocimientos transmitidos y brindados durante nuestra formación profesional.

Atentamente

Las tesistas

RESUMEN

En este trabajo de investigación se presenta la metodología de análisis de supervivencia aplicada a un problema real de deserción estudiantil en los estudiantes de la Escuela Profesional de Matemática. El método no paramétrico de Kaplan-Meier se utiliza para estimar la función de supervivencia y la función de riesgo de una muestra en un intervalo de tiempo definido; para comparar las curvas de supervivencia se utiliza la prueba de Log-Rank y para saber que variables influyen más sobre el riesgo de desertar se utiliza el método semi paramétrico del modelo de regresión de Cox. Los objetivos son utilizar el análisis de supervivencia para explicar la deserción estudiantil, determinar el tipo de censura y determinar los factores asociados con la deserción estudiantil de los estudiantes de la Escuela Profesional de Matemática, se trabajó con una población de 131 estudiantes que ingresaron en los semestres 2010-I, 2010-II, 2011-I, 2011-II, 2012-I, 2012-II, y se les hizo un seguimiento de 10 semestres consecutivos. Los datos para nuestra aplicación han sido obtenidos de los registros de ingresantes y matriculados de los estudiantes de la Escuela Profesional de Matemática, datos proporcionados por la oficina del Centro de Cómputo y la oficina de la Unidad de Bienestar Universitario de la Universidad Nacional de San Antonio Abad del Cusco; las variables registradas son: sexo, colegio, modalidad de ingreso, situación familiar, edad, número de hermanos, ingreso económico familiar, lugar de procedencia, ocupación del jefe de hogar y el estudiante trabaja . Se concluye que hubo una deserción de 61.8% del total de estudiantes; los factores asociados a la deserción estudiantil de los estudiantes de la Escuela Profesional de Matemática son: sexo, edad, ingreso económico familiar y ocupación del jefe de hogar.

Palabras claves: Análisis de supervivencia, censura, estimador de Kaplan-Meier, modelo de regresión de Cox y deserción estudiantil.

ABSTRACT

This research work presents the survival analysis methodology applied to a real problem of student dropout in students of the Professional School of Mathematics. The non-parametric Kaplan-Meier method is used to estimate the survival function and the hazard function of a sample in a defined time interval; To compare the survival curves, the Log-Rank test is used and to find out which variables influence the risk of dropping out the most, the semi-parametric method of the Cox regression model is used. The objectives are to use the survival analysis to explain the student dropout, determine the type of censorship and determine the factors associated with the student dropout of the students of the Professional School of Mathematics, we worked with a population of 131 students who entered the 2010-I, 2010-II, 2011-I, 2011-II, 2012-I, 2012-II semesters, and they were followed up for 10 consecutive semesters. The data for our application has been obtained from the registration and enrollment records of the students of the Professional School of Mathematics, data provided by the office of the Computer Center and the office of the University Welfare Unit of the National University of San Antonio Abbot of Cusco; The variables recorded are: sex, school, type of income, family situation, age, number of siblings, family income, place of origin, occupation of the head of the household and the student works. It is concluded that there was a dropout of 61.8% of the total of students; The factors associated with the student dropout of the students of the Professional School of Mathematics are: sex, age, family income and occupation of the head of the household.

Key words: survival analysis, censorship, Kaplan-Meier estimator, Cox regression model and student dropout.

INTRODUCCIÓN

La deserción estudiantil es uno de los problemas que aborda 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 anterior, 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) 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 se percibe la sensación del fracaso, pérdida de oportunidades laborales implicando menor posibilidad de conseguir empleos satisfactorios y el impacto que esto produce en términos individuales y familiares.

La información sobre las características de los estudiantes de la Escuela Profesional de Matemática de los semestres 2010-I, 2010-II, 2011-I, 2011-II, 2012-I, 2012-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.

PLANTEAMIENTO DEL PROBLEMA

1.1. Caracterización del problema

Según Vélez y López (2004), la deserción estudiantil Universitaria se manifiesta en todos los países del mundo, tanto como en las Universidades públicas y privadas, en algunas naciones como España, Estados Unidos, Francia y Austria la tasa de deserción Universitaria estudiantil oscila entre el 30% y 50%, y en otros países europeos como Alemania (20%-25%), Suiza (7%-30%), Finlandia (10%), los países bajos (20%-30%) las cifras son menores. Por su parte, en América Latina el promedio de la deserción oscila cerca al 55% (Ministerio de Educación Nacional de Colombia, 2009).

Según Gonzales (2005) citado por (Díaz, 2008) en Chile, de acuerdo a estimaciones basadas en estadísticas nacionales, la tasa de deserción global de pregrado, evaluada para el año 2004, es cercana al 53.7%, siendo mayor en las Universidades privadas que en las públicas. Por su parte, las áreas del conocimiento más críticas son humanidades y derecho con cifras de orden de 80% y las más eficientes son las áreas de educación y salud con un 37% y un 27% respectivamente. Por género las mujeres poseen una tasa de deserción promedio más baja que los varones de 43% y 50%, respectivamente.

La deserción estudiantil Universitaria a través de estos primeros años del siglo XXI en nuestro país al igual que en el resto del mundo, es un problema que preocupa y alarma a toda la comunidad educativa, se observa que los estudiantes abandonan sus estudios en los primeros semestres, el abandono es principalmente por la falta de orientación vocacional, o a la carrera que ingresaron no es lo que ellos esperaban y no cumple su perspectiva.

Un estudio realizado por INSAN Consultores revela que la deserción Universitaria en el Perú alcanza el 30%, motivada especialmente, por la falta de una buena orientación vocacional y

razones económicas. Para el 2017, la proyección de ingresantes a diferentes Universidades supera los 300 mil, y de este grupo entre 40 y 50 mil jóvenes abandonarán sus estudios Universitarios cada año. Afirmó que el 70% de los que deciden no continuar pertenece a Universidades privadas y el 30% restante a estatales. “En términos económicos, tal decisión representa para los padres de familia una pérdida de al menos 100 millones de dólares”, según Justo Zaragoza, director de Expo Universidad 2017. (Educación al Futuro.com, 2017).

Según la Dirección de sistemas de información, unidad de estadística-UNSAAC (2016 y 2017), del total de estudiantes que ingresan a la Universidad Nacional de San Antonio Abad del Cusco el 27% y 29.69% respectivamente abandonan la Universidad; así mismo se observó que del total de estudiantes que ingresan a la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco el 64% y 68.62% respectivamente abandonan la Universidad.

Actualmente se ha notado que la cantidad de estudiantes que ingresan a la Escuela Profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco difiere mucho al de estudiantes que egresan de la misma y más aún, del número que finalizan satisfactoriamente sus estudios superiores en la mencionada Escuela.

Al abandonar los estudiantes la Escuela Profesional de Matemática, generan muchos problemas en lo personal, porque no pueden lograr sus objetivos profesionales, lo que genera que no pueda acceder a buenos puestos de trabajo, por lo que su situación económica no será sostenible y la Universidad Nacional de San Antonio Abad del Cusco habrá invertido recursos en jóvenes que abandonan sus proyectos educacionales, perdiendo así, el aporte intelectual que debería beneficiar a Universidad.

A través de este trabajo de investigación, se estudia la deserción de los estudiantes de la Escuela Profesional de Matemática mediante la utilización de la función de supervivencia estimada por el modelo de Kaplan-Meier y se establece los factores asociados a la deserción a partir del modelo de regresión de Cox que es utilizada para datos de sobrevivencia de los estudiantes de la Escuela Profesional de Matemática de la cohorte 2010-I al 2012-II.

1.2. Formulación del problema

1.2.1. Problema general

¿Cómo el análisis de supervivencia permite explicar la deserción estudiantil en la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco en la cohorte de estudiantes ingresantes del 2010-I al 2012-II?

1.2.2. Problemas específicos

- ¿Qué tipo de censura presentan los estudiantes en la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco en la cohorte de estudiantes ingresantes del 2010-I al 2012-II?
- ¿Cuáles son los factores asociados a la deserción estudiantil en la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco en la cohorte de estudiantes ingresantes del 2010-I al 2012-II?

1.3. Objetivo

1.3.1. Objetivo general

Desarrollar el análisis de supervivencia para explicar la deserción estudiantil en la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco en la cohorte de estudiantes ingresantes del 2010-I al 2012-II.

1.3.2. *Objetivos específico*

- Explicar el tipo de censura que presentan los estudiantes de la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco en la cohorte de estudiantes ingresantes del 2010-I al 2012-II.
- Determinar los factores asociados a la deserción estudiantil de la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco en la cohorte de estudiantes ingresantes del 2010-I al 2012-II.

1.4. *Justificación:*

La deserción estudiantil en estudios superiores es un problema que afecta de forma silenciosa a los estudiantes de las universidades, a sus familias y a la sociedad por las pérdidas económicas que se generan, especialmente la deserción que se presenta en la Escuela Profesional de Matemática y de igual manera en las diferentes Escuelas Profesionales de la UNSAAC.

El presente 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”, surge con el propósito de analizar en qué semestre un estudiante toma la decisión de abandonar sus estudios Universitarios y determinar el factor que influye a tomar dicha decisión de desertar.

1.5. *Limitación del Trabajo*

Una limitación importante para realizar este trabajo de investigación fue, que la oficina responsable de proporcionar la información obtenida de la ficha socioeconómica de los estudiantes ingresantes a la Universidad Nacional de San Antonio Abad del Cusco lo tenían:

- Desorganizado
- No digitalizado

Es decir, no cuenta con una base de dato disponible para los investigadores y no permiten el acceso a la información, por ello se tuvo que acudir a otras instancias u oficinas de la Universidad Nacional de San Antonio Abad del Cusco para completar la información necesaria para el desarrollo del presente trabajo de investigación.

1.6. Hipótesis

1.6.1. Hipótesis general

Según el análisis de supervivencia existe deserción estudiantil en la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco en la cohorte de estudiantes ingresantes del 2010-I al 2012-II.

1.6.2. Hipótesis específica

Existen factores asociados a la deserción estudiantil en la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco en la cohorte de estudiantes ingresantes del 2010-I al 2012-II.

MARCO TEÓRICO

2.1. Antecedentes de la investigación

2.1.1. Antecedentes internacionales

Vergel, Martínez y Ibarguen (2016) en el artículo titulado *Modelos estimados de análisis de supervivencia para el tiempo de permanencia de los estudiantes de la Universidad Francisco de Paula Santander*, Cúcuta, Colombia. Indicaron que la deserción es un problema presente en las instituciones educativas en el ámbito nacional e internacional, que implica el hecho de que un número de estudiantes matriculados, no sigan la trayectoria normal de los programas académicos, bien sea por retirarse de ella o demorar más tiempo del previsto para finalizar sus estudios; objetivo: Encontrar la función que describe el riesgo de cambio de estado en diferentes periodos de tiempo, determinar cuáles son las características que más se relacionan con la duración y riesgo de cada evento, donde llegaron a las siguientes conclusiones.

- a) Para el evento Graduarse, los estudiantes que obtuvieron buen desempeño en educación media, así mismo el conocimiento en áreas básicas aumentan el riesgo de Graduarse. De otra parte, no existe diferencia significativa en los estudiantes de género masculino o femenino. Estudiantes que reciben beneficio beca trabajo permanecen 12 semestres o más.
- b) Se evidencia como causa de abandono de los programas académicos, las sanciones, iniciar trabajos debido a situaciones socioeconómicas, así mismo la pérdida de empleo por parte de los padres de familia o el tener personas a cargo. El inicio de consumo de sustancias psicoactivas, o problemas psicológicos (en menor proporción) aumenta riesgo de abandono. Estudiantes que inician un programa que no es el de su preferencia o si la

universidad no es de su preferencia tienen alto riesgo de retiro. Los estudiantes que costean la matrícula con ingresos propios presentan mayor riesgo de retiro. Dos determinantes para estudiantes de postgrado relacionados con factor socio-económicos o terminación de trabajo de grado.

- c) Tres determinantes del tiempo de permanencia de los estudiantes, fueron identificados como razones de riesgo a través del modelo de regresión de Cox y análisis cualitativo, a saber, Factores Individuales y Socioeconómicos, Factores Académicos y Factores Institucionales.

Vergara, Boj del Val, Barriga y Díaz (2017) en el artículo *Factores explicativos de la deserción de estudiantes de pedagogía*, de la Universidad de Concepción, Bio Bio, Chile. La deserción Universitaria representa un problema que provoca múltiples efectos para la sociedad. A nivel social, influye en el incremento de las tasas de desempleo; a nivel institucional actúa en detrimento de los indicadores de eficiencia y calidad; y en el ámbito personal genera sentimientos de frustración. Propuso como objetivo: Identificar los factores que influyen en la deserción de estudiantes de pedagogía, considerando sus características individuales y académicas, donde llegaron a las siguientes conclusiones.

- a) Las variables académicas que explican la deserción Universitaria corresponden al promedio de notas de enseñanza media, el lugar en la lista de seleccionados, provenir de un establecimiento secundario científico-humanista, el total de asignaturas inscritas, el último promedio curricular y la suposición de estudios.
- b) Las variables individuales que explican la deserción de los estudiantes son: sexo y procedencia de la región del Bio Bio.

2.1.2. Antecedentes nacionales

Puma (2018) en su tesis *Análisis de supervivencia de la deserción en los estudiantes de la facultad de Ciencias Naturales y Formales a través del Modelo de Cox*, (tesis de pregrado).

Universidad Nacional de San Agustín, Arequipa, Perú. Propuso como objetivo: Encontrar los factores de riesgo frente a la deserción estudiantil de la Facultad de Ciencias Naturales y Formales de la Universidad Nacional de San Agustín, donde llegaron a las siguientes conclusiones:

- a) En los estudiantes de la escuela profesional de matemáticas tienen un nivel de supervivencia del 48.4%, que es relativamente superior al nivel de supervivencia de los estudiantes de las escuelas profesionales de física y química que tienen 31.4% y 43.4% respectivamente.
- b) Se observa que en los estudiantes varones, su nivel de supervivencia 39% es menor en comparación de las estudiantes mujeres que tienen un nivel de supervivencia del 48.4%.
- c) Los estudiantes que provienen de colegios particulares tienen un nivel de supervivencia del 61.7%, que es mayor en comparación de los estudiantes que provienen de colegios nacionales con un nivel de supervivencia del 36.3%.
- d) Los estudiantes que ingresan a la universidad hasta los 24 años tienen un nivel de supervivencia del 42.9%, que es mayor en comparación a aquellos estudiantes que ingresaron a la universidad de los 25 años a más con un nivel de supervivencia del 12.5 %
- e) Los estudiantes que provienen de provincia tienen un nivel de supervivencia del 69.2% mayor respecto a los estudiantes que provienen de ciudad con un nivel de supervivencia del 38%.

- f) Los factores en riesgo que determinan la deserción son: el colegio, edad y lugar de estudios.

Barturén y Niño (2018) en su tesis *Modelos de sobrevida para la deserción estudiantil en la Escuela Profesional de Estadística de la Universidad Nacional Pedro Ruiz Gallo, Lambayeque 2007-2008*, (tesis de pregrado) Lambayeque, Perú. En América latina la tasa bruta de matrícula aumentó en la última década, pero no todos llegan a culminar sus estudios satisfactoriamente y Perú no es la excepción, cada vez hay más jóvenes que ingresan a la Universidad para realizar sus estudios, donde la cantidad de estudiantes que ingresan es mayor a la cantidad de estudiantes que obtienen su título profesional. Propuso como objetivo: Determinar el modelo de sobrevida que explica la deserción de los estudiantes que ingresaron en el periodo 2007 al 2008 a la Escuela Profesional de Estadística de la Universidad Nacional Pedro Ruiz Gallo de Lambayeque, donde llegaron a las siguientes conclusiones.

- a) Las características más resaltantes de los estudiantes que ingresaron a la escuela profesional de Estadística en el periodo 2007 – 2008 son: el 45.5% de los estudiantes abandonaron sus estudios, 47.3% de los estudiantes fueron mujeres y 52.7% hombres, el 84.8% tenían menos de 20 años de edad, el 62.5% provienen de colegios públicos, el 60.7% provienen de la provincia de Chiclayo y solo el 17% provienen de la provincia de Lambayeque.
- b) Se propone para la validación de un modelo probabilístico del pronóstico de deserción estudiantil de los estudiantes que ingresaron a la escuela profesional de estadística (2007-2008), fundamentalmente en una variable: Total de créditos aprobados la cual incide en 0.896 el riesgo de desertar.

$$y = h(t; X) = h_0(t). e^{-0.110X_7}$$

- c) Se observó que la variable total de créditos aprobados tiene una influencia altamente significativa en el modelo probabilístico del pronóstico de deserción estudiantil de los estudiantes que ingresaron a la escuela profesional de estadística en el año 2007 y 2008, la cual incide en 0.899 y 0.869 el riesgo de desertar respectivamente, concluyendo que a más cursos desaprobados la probabilidad de desertar aumenta considerablemente.

2.1.3. Antecedentes locales

Huamán y Martínez (2018) en su investigación de tesis *factores relacionados con la deserción de estudiantes de enfermería, Universidad Nacional de San Antonio Abad del Cusco, semestres 2010- I al 2017-I*, Universidad Nacional de San Antonio Abad del Cusco, Cusco, Perú. La deserción trae consigo serias consecuencias como: sentimientos de frustración en los estudiantes, baja eficiencia del sistema de educación superior y dificultades en el cumplimiento de la función social de la educación, y pérdida de recursos estatales, institucionales y familiares. Cuyo objetivo es: Determinar los factores personales, económicos y académicos que se relacionan con la deserción de estudiantes de Enfermería - UNSAAC, semestres 2010-I al 2017-I, donde llegaron a las siguientes conclusiones.

- a) Dentro de las características generales de los estudiantes de enfermería del código 2010-I y 2010-II durante los semestres 2010-I al 2017-I en la escuela profesional de enfermería, en su mayoría son de 19 a 21 años, procedentes de zona urbana, con estado civil soltero, viven con sus padres y hermanos dedicándose solo a estudiar.
- b) En cuanto a los factores personales frente a la deserción universitaria, se encuentra en mayor proporción el indicador motivo de abandono de la carrera siendo problemas

familiares el de mayor porcentaje, en mayor proporción el desertor se encuentra entre el primer y segundo lugar en cuanto a posición de hermanos.

- c) En lo que respecta a los factores económicos que ocasionan la deserción apreciamos que en mayor proporción se obtuvo que los estudiantes desertores tienen un ingreso familiar menor de 800 soles, recibiendo en mayoría el apoyo económico por parte de familiares.
- d) En los factores académicos la mayoría de estudiantes desertores tienen expectativas algo satisfechas con la carrera, una regular percepción de las clases impartidas además que no reciben el servicio de tutoría en el desarrollo de su formación, afirmando que si tuvieron interés por otra carrera en el transcurso, presentando dificultad para adaptarse y prueba de ello es que desaprobaron de uno a dos cursos, mostrando poca vocación por la carrera, siendo esta elegida por presión de sus padres y que proceden de colegios públicos en mayoría. Asimismo, el semestre en el que decidieron abandonar la carrera difiere numéricamente en cada semestre siendo el 2do semestre en que cuenta con mayor porcentaje de desertores seguidamente el 3er semestre, ocurre lo mismo en 4to y 5to respectivamente, 1er semestre, 8vo y finalmente en 6to y 7mo semestre.
- e) Según la prueba de Chi-cuadrado se concluye; que el factor personal: motivo para el abandono, el factor económico: ingreso familiar y los factores académicos: dificultad para adaptarse al cambio, vocación por la carrera escogida, percepción sobre los docentes que imparten las clases, tienen una relación significativa con la deserción de los estudiantes de enfermería del código 2010 durante los semestres 2010-I al 2017-I en la Escuela Profesional de Enfermería por que los valores hallados fueron menores de 0.05 (sig = resultado < 0.05).

- f) El principal factor de deserción predominante es el factor académico con el indicador vocación por la carrera escogida podemos afirmar también que un estudiante si no cuenta con una vocación de servicio a la carrera es posible que abandone la carrera.

En la dirección de sistemas de información, unidad de estadística (2016) en el *compendio estadístico N°31* de la Universidad Nacional de San Antonio Abad del Cusco nos informa los siguientes resultados:

Tabla 1

Deserción del estudiante de pre-grado año 2015

Escuela profesional	Ingresantes 2009-2015	Matriculados Código 2009-2015	Graduados Titulados a partir del código 2009	Deserción N°	Estimación
Total	23777	16155	1183	6439	27%
Matemática	412	140	8	264	64%

Fuente: UNSAAC Unidad de estadística, compendio estadístico N° 31, 2016. Pág. 104

En la dirección de sistemas de información, unidad de estadística (2017) en el *compendio estadístico N°32* de la Universidad Nacional de San Antonio Abad del Cusco nos informa los siguientes resultados:

Tabla 2

Deserción estimada del estudiante de pregrado año 2016

Escuela profesional	Ingresantes 2009-2016	Matriculados Código 2009-2016	Graduados Titulados a partir del código 2009	Deserción N°	Estimación
Total	26542	17261	1400	7881	29.69%
Matemática	498	152	4	342	68.62%

Fuente: UNSAAC Unidad de estadística, compendio estadístico N° 32, 2017. Pág. 125

2.2. Análisis de Supervivencia

El análisis de supervivencia estudia los procesos aleatorios relacionados con la muerte de organismos vivos y la falla de sistemas físicos (mecánicos o electrónicos).

El análisis de supervivencia es una rama de la estadística que comprende una variedad de técnicas para estudiar los tiempos de ocurrencia de un evento de interés, como por ejemplo el tiempo de aprendizaje de una habilidad, la detección de una enfermedad, la falla de un equipo, la baja o pérdida de un cliente, la desactivación de una línea, el abandono de un estudiante etc. (Mondragón, 2013)

El análisis de supervivencia es un conjunto de técnicas estadísticas en la que la variable respuesta es el tiempo que transcurre entre el comienzo de seguimiento del individuo en el estudio y la aparición del evento de interés (muerte). Con frecuencia suele ocurrir que los individuos abandonen el estudio antes de que presente el evento, con lo que solo se obtiene información parcial (censura) de la variable de interés. El objetivo del análisis de supervivencia es incorporar esta información parcial que proporcionan los individuos censurados mediante métodos desarrollados para este fin. (San Jose et ál., 2009)

El análisis de supervivencia es un conjunto de métodos, técnicas y/o procedimientos estadísticos para analizar datos que corresponden al tiempo de ocurrencia de algún evento en particular, llamado tiempo de supervivencia o tiempo de vida. El objetivo del análisis de supervivencia es poder predecir la probabilidad de que ocurra el evento de interés (muerte) para algún individuo, esto es, la probabilidad de sobrevivir. (Avendaño, 2013)

2.2.1. *Definiciones de términos utilizados en el análisis de supervivencia*

Fecha inicial: Fecha de diagnóstico, de inicio del tratamiento o de remisión completa.

Fecha de última noticia: Fecha correspondiente a la última información que se tiene del caso.

Seguimiento: Es la observación de los individuos de un grupo a partir de la fecha inicial, para conocer su estado vital (vivo o muerto).

Periodo de seguimiento: El tiempo transcurrido entre la fecha de inicio y la fecha de corte del estudio.

Fecha de finalización del estudio: Fecha fijada por el investigador para el término del seguimiento de los individuos.

Tiempo de supervivencia: Es el intervalo de tiempo transcurrido entre las fechas de inicio y de última noticia.

Observación censurada: Ocurre cuando existe pérdida del seguimiento, muerte por otras causas, expulsión del estudio sin haber ocurrido el evento, o no ocurre el evento de interés durante el periodo de observación.

Probabilidades de supervivencia: De un conjunto de tiempos de supervivencia de una muestra de individuos (incluyendo los censurados), podemos estimar la proporción de la población que sobrevivirá un periodo de tiempo bajo las mismas condiciones.

En el análisis de supervivencia, el análisis de los datos puede ser realizado utilizando técnicas paramétricas y no paramétricas:

- a) **Métodos Paramétricos:** Son las más frecuentemente utilizadas, las distribuciones que se utilizan para modelar la supervivencia son: la Exponencial, Weibull y Log normal entre otras distribuciones de probabilidad.
- b) **Método no Paramétrico:** En la técnica no paramétrica se tienen las tablas de vida, el método de Kaplan-Meier.

- c) **Método Semi paramétrico:** El método semi paramétrico se refiere a la regresión de Cox y tiene una parte paramétrica y lo otra parte no paramétrica, se detallará más adelante en la definición de Cox. (Tineo, 2005)

2.2.2. *Elementos Básicos*

La variable aleatoria T representa el tiempo de duración hasta la ocurrencia de un evento de interés (muerte, deserción, abandono, falla de un sistema, entre otros) es de suma importancia en el análisis de supervivencia, cuya función de densidad de probabilidad es $f(t)$ y su función de distribución de probabilidad acumulada $F(t) = P(T \leq t)$ que es la probabilidad de que el evento de interés ocurra en un tiempo menor o igual que t . También la función de supervivencia $S(t) = 1 - F(t) = P(T > t)$ es importante para el análisis de supervivencia porque expresa la probabilidad de que el evento de interés ocurra en un tiempo mayor a t . (Barrera, 2008)

2.2.3. *Función de supervivencia*

La función de supervivencia denotada por $S(t)$, es la probabilidad de que un individuo sobreviva a la observación (no experimente el evento de interés (muerte o deserción)) al menos hasta el tiempo t .

Definición: Sea T una variable aleatoria (no negativa) con una función de distribución acumulada $F(t)$ y una función de densidad de probabilidad $f(t)$. La función de supervivencia $S(t)$ se expresa de la siguiente manera:

$$S(t) = 1 - F(t) = P(T > t) = \int_t^{\infty} f(t)dt \quad (1)$$

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

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

Es decir, la probabilidad de sobrevivir al tiempo cero es uno y la de sobrevivir al tiempo infinito es nula. (García, 2014).

2.2.4. Función de riesgo

Según Barrera (2008), la función de riesgo $h(t)$, representa el riesgo instantáneo de que un evento ocurra en un intervalo infinitamente pequeño de tiempo $(t, t + \Delta t)$, dado que no ha ocurrido hasta el tiempo t . Se escribe de la siguiente manera:

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

Aplicando la probabilidad condicional a la expresión (2) se obtiene:

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

Desarrollando el cálculo de probabilidades y del hecho de que $P(T > t) = S(t)$ en la expresión (3), se tiene que:

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

Aplicando la definición de la derivada a la ecuación (4) se obtiene:

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

Aplicando la integral a la ecuación (5) se obtiene la función de riesgo acumulada.

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

También se puede escribir como: $H(t) = \int_0^t \frac{f(t)}{1-F(t)} dt$ sustituyendo $u = 1 - F(t)$, y

$du = -f(t)dt$, la expresión (6) se convierte en $H(t) = \int_0^t \frac{-du}{u}$, de donde resulta que:

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

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

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

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

2.2.5. *Censura*

El tiempo de supervivencia de un individuo esta censurado cuando el evento de interés (deserción o muerte) no ha sido observado durante el período de seguimiento o periodo de observación.

Definición: Sea T la variable aleatoria que representa el tiempo hasta el evento de interés de un estudiante, 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 se considera como no-censurado cuando $T \leq t_k$.

Entonces, otra característica distintiva del Análisis de Supervivencia es que aprovecha esa información parcial proporcionada por la censura. (Puma, 2018).

2.2.5.1. *Censura por la derecha*

Este mecanismo de censura es el caso más común de datos incompletos y se caracteriza por el hecho de que, durante el tiempo de observación del individuo, no se produce el evento que se desea observar. La falta de datos puede darse por diversas razones, entre las cuales, puede ser que, hasta el momento de la finalización del estudio no ha ocurrido el evento de interés (siempre

en cuando el periodo de observación sea finito), el individuo abandone estudio o se haya producido otro evento que imposibilite que el suceso a observar ocurra. (García, 2014)

2.2.5.2. Censura por izquierda

Se produce cuando la primera observación que se realiza sobre el individuo, el evento que se desea observar ya ha ocurrido.

2.2.5.3. Censura por intervalo

Ocurre cuando solo se sabe que al individuo le ocurre el evento de interés en un intervalo de tiempo, es decir, entre en instante t_i y un tiempo t_k .

Los datos de supervivencia se presentan en forma (t_i, δ_i) donde t_i es el tiempo de observación y $\delta_i = 0$ si la observación es censurada y $\delta_i = 1$ cuando el evento de interés ocurre (deserción). (García , 2014)

2.3. Método no Paramétrico

2.3.1. Estimador de Kaplan-Meier

Kaplan- Meier (1958) es el método no paramétrico que se usa para estimar la función de supervivencia $S(t)$ basada en una muestra aleatoria independiente de tiempos de falla con censura, fundamentada en el principio de calcular la supervivencia. El método de construcción del estimador es un método no paramétrico ya que no asume ninguna estructura para la función de distribución de probabilidad del tiempo de vida.

Para realizar la estimación de la función de supervivencia el objetivo del estimador Kaplan-Meier es utilizar toda la información disponible de casos censurados y no censurados, en

cualquier instante de tiempo este estimador es obtenido de la multiplicación de una secuencia de probabilidades condicionales de supervivencias estimadas donde se obtiene probabilidades condicionadas estimadas a partir del número de casos observados en riesgo n_i y el número de “muertes” observadas d_i en un instante de tiempo y se calcula como $(n_i - d_i)/n_i$.

Sean $t_1 < t_2 < \dots < t_k$ los k distintos tiempos de falla observados en una muestra aleatoria de tamaño n . La posibilidad que ocurra mas de una falla en t_k esta permitido, por lo que denotaremos el número de fallas en ese tiempo por $d_i (d_i \geq 1)$. Para t fijo, el estimador de la función de supervivencia de Kaplan-Meier $\hat{S}(t)$ se calcula de la siguiente manera:

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

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 , es decir, individuos que no fallaron y no fueron censurados hasta el instante inmediatamente anterior a t_k .

De la ecuación 7 se puede observar. Cuando $t = 0$, $\hat{S}(t) = 1$; es decir, todos los individuos comienzan vivos en el estudio. (Puma , 2018)

2.3.2. Varianza del Estimador Kaplan-Meier

La varianza para los estimadores de la función de supervivencia es de gran importancia, por lo tanto, en este estudio se usará el método Delta el cual está basado en la expresión en series de Taylor de primer orden.

Método Delta

Usar una aproximación en series de Taylor para obtener una función lineal que aproxime una función más complicada como es el caso del estimador de la función de supervivencia es el propósito del método Delta.

Sea $f(x)$ una función de densidad de la probabilidad de una variable aleatoria X . La expansión en series de Taylor de primer orden alrededor de la media es:

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

Aplicando propiedades de varianza a la función $f(x)$.

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

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

Donde σ^2 es la varianza de la variable aleatoria X . como ejemplo se muestra la aplicación de propiedades de varianza para las siguientes funciones:

Función logaritmo

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

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

Función exponencial:

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

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

Varianza función de supervivencia

Calcularemos la varianza del estimador de Kaplan-Meier mediante el método Delta.

Aplicando logaritmo a ambos miembros de la ecuación (7).

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

realizando $\hat{p}_i = \frac{n_i - d_i}{n_i}$ te tiene:

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

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

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

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

Aplicando las propiedades de varianza 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 varianza de un constante es cero:

$$\text{Var}(\log(\hat{p}_i)) \cong \text{Var}(\hat{p}_i) \left[\frac{d}{dp_i}(\log(p_i)) \right]^2 \quad (10)$$

Suponiendo que $n_i \hat{p}_i$ tiene una Distribución Binomial con parámetros n y p_i es decir:

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

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

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

Simplificando n_i

$$\text{Var}(\hat{p}_i) = \frac{p_i(1 - p_i)}{n_i} \quad (11)$$

Reemplazando la ecuación (11) en (10)

$$\text{Var}(\log(\hat{p}_i)) \cong \frac{p_i(1-p_i)}{n_i} \times \frac{1}{p_i^2}$$

Simplificando p_i

$$\text{Var}(\log(\hat{p}_i)) \cong \frac{(1-p_i)}{n_i p_i}$$

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

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

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

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

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

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

Aplicando serie de Taylor a $\log \hat{S}(t)$ alrededor de $\hat{S}(t) = S(t)$ por tanto se tiene:

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

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

Aplicando derivada respecto a $S(t)$

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

Aplicando propiedades 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

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

Sustituyendo este resultado en la ecuación (12)

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

Finalmente hallamos la varianza del estimador de Kaplan-Meier

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

El error estándar de $\hat{S}(t)$ es la raíz cuadrada de la varianza

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

El intervalo de confianza para $S(t)$ al 95% este definido de la siguiente manera:

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

La mediana para el tiempo de supervivencia $S(t)$

$$S(t_{\text{mediana}}) = 0.50 \quad (15)$$

2.3.3. Comparación funciones de supervivencia

Es importante e ilustrativa la comparación gráfica de funciones de supervivencia, pero a veces esta comparación se vuelve muy subjetiva y en algunos casos difíciles de realizar. Para resolver este inconveniente se usa pruebas estadísticas que permiten identificar si las diferencias observadas entre las funciones de supervivencia son significativas como por ejemplo la prueba no paramétrica de Log-Rank.

2.3.3.1. Prueba de hipótesis para la función de supervivencia

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: Con un $p - value$ menor o igual a 0.05 se rechaza H_0 , se concluye que existe diferencia significativa y con un $p - value$ mayor que 0.05 no se rechaza H_0 , se concluye que no existe diferencia significativa.

2.3.3.2. Prueba de Log- Rank

La prueba de Log-Rank (Mantel y Haenzel) es un método que permite comparar las curvas de supervivencia de dos a más grupos de individuos.

En la mayoría de veces las pruebas estadísticas están basadas en tablas de contingencia de grupo por estado de cada tiempo de supervivencia, de la siguiente manera:

Tabla 3

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

Para comprobar las funciones de supervivencia entre grupos se da las siguientes hipótesis:

H_0 : La supervivencia de los grupos es la misma.

H_1 : La supervivencia de los grupos es diferente

El estadístico de prueba es la suma de las diferencias entre los eventos observados y el número esperado de muertes para todos los tiempos de supervivencia observados, dividido por la varianza del número de muertes para alguno de los dos grupos, por ejemplo, para el grupo 1 se tiene que el estimador del número de muertes es:

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

y 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 de Log-Rank para probar la hipótesis de igualdad entre las dos funciones de supervivencia está dado por:

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

donde Q se distribuye como una chi-cuadrada con un grado de libertad ($p = P(x^2(1)) \geq Q$).

Si se tuviera más de dos grupos se utiliza en forma similar al caso ya visto de solo dos grupos, a continuación, se muestra la tabla asociada a varias funciones de supervivencia base para el cálculo de los estadísticos de prueba para la igualdad de las funciones:

Tabla 4

Prueba de igualdad de la función de supervivencia de M grupos de observación t_i .

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

Primero se debe estimar el número esperado de muertes de cada grupo bajo la suposición de igual función de supervivencia, es decir:

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

teniendo varios grupos, ya no es una varianza la que se calcula sino una matriz de covarianzas del vector d_i . Para obtener la matriz de covarianzas se debe suponer que el número observado de eventos sigue una distribución hipergeométrica central multivariada y los términos de la diagonal de la matriz denotada por \hat{V}_i , se calculan de la siguiente manera:

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

Los componentes fuera de la diagonal se calculan como:

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

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

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

donde W_i es una matriz diagonal que tiene los pesos w_i , d_i es el vector que contiene la cantidad de muertes para cada grupo en el tiempo t_i , \hat{e}_i es el vector que contiene el número esperado de muertes para cada grupo en el tiempo t_i . Al igual que en el caso de dos grupos se tiene que Q para el caso de varios grupos, bajo la hipótesis nula de igualdad de funciones de supervivencia, se distribuye también con una Chi-cuadrado con $M-1$ grados de libertad ($p = P(x^2(1) \geq Q)$). (Barrera , 2008)

2.4. Método Semi Paramétrico

2.4.1. Modelo de Regresión de Cox

El modelo de riesgos proporcionales de Cox es más conocido como el modelo de regresión de Cox, este modelo permite estimar la relación que hay entre un conjunto de variables explicativas fijas llamadas covariables X_1, X_2, \dots, X_n y la respuesta o tiempo de supervivencia; o más bien, con la función de riesgo $h(t; X)$, que es la tasa instantánea del suceso de interés.

En el modelo de regresión de Cox (1972), la función de tasa de riesgo del tiempo de falla de un modelo con vectores de covariables dadas por X está definida de la siguiente manera:

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

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 $h(t; X)$ definido en la expresión (23) es un modelo semi-paramétrico, debido a que este modelo tiene una parte paramétrica y otra parte no paramétrica.

- a) La parte paramétrica del modelo es $\exp(\beta^T X)$, llamado función de riesgo relativo, la cual está claramente especificada los efectos relativos de los parámetros de regresión estimados sobre el riesgo.
- b) La parte no paramétrica del modelo es $h_0(t)$, la función de riesgo base, que es una función arbitraria y no especificada.

Asimismo, tal como se ha mencionado anteriormente, el modelo de regresión de Cox también es conocido como modelo de riesgos proporcionales de Cox, debido a que el cociente entre el riesgo para dos individuos con el mismo vector de covariable es constante e independiente en el tiempo, es decir

$$\frac{h(t; X_i)}{h(t; X_j)} = \frac{h_0(t)\exp(\beta^T X_i)}{h_0(t)\exp(\beta^T X_j)} = \exp[\beta^T (X_i - X_j)] \quad (24)$$

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 . (García , 2014)

Observación: El componente no-paramétrico absorbe a la constante β_0 por lo tanto no aparece en la componente paramétrica $\exp(X\beta_j^T)$.

Si un individuo se presenta al inicio de estudio con un riesgo igual a 4 veces el riesgo de otro individuo, entonces esta será la misma para cualquier tiempo t durante el periodo de estudio, significa que las tasas de falla son proporcionales, suposición que es básica para el uso del modelo de riesgos proporcionales de Cox.

En una regresión lineal la media de la variable respuesta se relaciona linealmente con los factores pronósticos y en el modelo de regresión de Cox el logaritmo del cociente de las funciones de riesgo es el que se relaciona linealmente con los factores pronósticos (covariables).

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:

$$S_0(t) = \exp \left(- \int_0^t h_0(t) dt \right).$$

Donde

$$\int_0^t h_0(t) dt = H_0(t) \Rightarrow S_0(t) = \exp(-H_0(t)) \quad (25)$$

$H_0(t)$ Es conocido como el riesgo base acumulado, de modo que:

$$\begin{aligned} S(t, X) &= \exp \left(- \int_0^t h_0(t) dt \exp(X\beta^T) \right) \\ &= \exp(-H_0(t) \exp(X\beta^T)) \end{aligned}$$

$$S(t, X) = (S_0(t))^{\exp(X\beta^T)} \quad (26)$$

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

$$\begin{aligned} h(t, X) &= \frac{f(t, X)}{S(t, X)} \Rightarrow f(t, X) = h(t, X) \cdot S(t, X) \\ &= h_0(t, X) \exp(X\beta^T) \cdot S(t, X) \\ &= h_0(t, X) \exp(X\beta^T) (S_0(t))^{\exp(X\beta^T)}, \end{aligned}$$

Donde $S_0(t)$ es la función de supervivencia base y la gráfica de $S_0(t)$ comparando con la gráfica $(S_0(t))^{\exp(X\beta^T)}$ son aproximadamente paralelas, por tanto, la suposición de riesgos proporcionales en el modelo de Cox se satisface. (Puma, 2018)

2.4.2. Estimación de los parámetros en el modelo de Cox

El método más utilizado para hacer la estimación de los parámetros de un modelo de regresión es el método de máxima verosimilitud que consiste en estimar los parámetros que maximice el logaritmo natural de la verosimilitud, en el caso del modelo de Cox este procedimiento no funciona fundamentalmente debido al desconocimiento de la función de riesgo base que aparece en el modelo de Cox.

Por eso Cox en (1972) desarrolla un método que permite estimar los parámetros de la regresión sin necesidad de asumir la forma particular para la función de riesgo base, plantea usar una expresión llamada “Método de máxima verosimilitud parcial” que depende solamente de los parámetros de interés. La función de verosimilitud parcial se escribe de la siguiente manera:

$$L(\beta) = \prod_{i=1}^k \frac{e^{X_{(i)}\beta^T}}{\sum_{j \in R(t_i)} e^{X_j\beta^T}} \quad (27)$$

(Barrera, 2008)

2.4.2.1. Estimación de β

El modelo de Cox se caracteriza por los coeficientes β que deben ser estimados por la observación. La presencia del componente no-paramétrico inválida el uso del método de máxima verosimilitud. Para estimar el vector de parámetros β Cox propuso el método de máxima verosimilitud parcial, donde la función de verosimilitud está dada por $L(\beta)$ escrita en la expresión (27) por que considera las probabilidades para aquellos individuos que son censurados. Se construye la función $L(\beta)$ a partir de una muestra de n individuos y que exista k ($k \leq n$)

fallas distintas en los tiempos t_1, t_2, \dots, t_k . Una forma simple de entender a la función de verosimilitud parcial es considerar el siguiente argumento condicional:

$$\begin{aligned}
 & P[\text{estudiante falla en } t_i / \text{del grupo de falla en } t_i] \\
 &= \frac{P(\text{estudiante falla en } t_i / \text{estaba vivo antes})}{P(\text{un fallo en } t_i \text{ del grupo en riesgo / estaba vivo antes})} \\
 &= \frac{h(t_i/X_i)}{\sum_{j \in R(t_i)} h(t_i/X_j)} \\
 &= \frac{h_0(t_i) \exp(X_i \beta^T)}{\sum_{j \in R(t_i)} h_0(t_i) \exp(X_j \beta^T)} \\
 &= \frac{\exp(X_i \beta^T)}{\sum_{j \in R(t_i)} \exp(X_j \beta^T)} \tag{28}
 \end{aligned}$$

Donde $R(t_i) = \{j: t_j > t_i\}$ es el conjunto de individuos en riesgo antes de t_i es decir, que están en observación y aún no han presentado el evento.

Ahora sí, la función de verosimilitud para β es el producto de diversas probabilidades, sobre todas las fallas $t_i, i = 1, 2, \dots, n$.

Esto es;

$$L(\beta) = \prod_{i=1}^n \left[\frac{\exp(X_i \beta^T)}{\sum_{j \in R(t_i)} \exp(X_j \beta^T)} \right]^{\delta_1} \tag{29}$$

Donde

$$\delta_1 = \begin{cases} 1, & \text{si } t_i \text{ no es censurado} & i = 1, 2, \dots, n \\ 0, & \text{si } t_i \text{ es censurado} & i = 1, 2, \dots, n \end{cases}$$

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

$$\ln[L(\beta)] = \sum_{i=1}^n \delta_1 \ln \left[\frac{\exp(X_i \beta^T)}{\sum_{j \in R(t_i)} \exp(X_j \beta^T)} \right]$$

$$\begin{aligned}
&= \sum_{i=1}^n \left[\delta_1 \ln \left(\exp(X_i \beta^T) - \sum_{j \in R(t_i)} \exp(X_j \beta^T) \right) \right] \\
l(\beta) = \ln[L(\beta)] &= \sum_{i=1}^n \delta_1 \left[\left(X_i \beta^T \right) - \ln \left[\sum_{j \in R(t_i)} \exp(X_j \beta^T) \right] \right] \quad (30)
\end{aligned}$$

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

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

Haciendo:

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

Obtenemos el siguiente sistema de ecuaciones por (ponderaciones).

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

A partir de este sistema de ecuaciones podemos obtener los valores estimados de $\hat{\beta}$. (Roque, 2009).

2.4.3. Estimación de la función de riesgo base $h_0(t)$

Para estimar la función de riesgo sea $\hat{h}_0(t)$ una constante entre tiempos de supervivencia no censurados. Primero se debe ordenar los tiempos de supervivencia no censurados en forma creciente:

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

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

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

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

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

Se desea 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:

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

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 Rivera & Niño Capcha, 2018)

2.4.4. Prueba de Wald

La prueba de Wald se calcula como la razón entre el coeficiente estimado y la desviación estándar de ese coeficiente como se muestra a continuación:

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

Bajo la hipótesis nula que el coeficiente es igual a cero, el estadístico de Wald se distribuye normal estándar, y para obtener el 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 se necesita hacer el test $H_0: \beta = \beta_0$ vs $H_1: \beta \neq \beta_0$, usamos el contraste de Wald, que bajo hipótesis nula, sigue una distribución Chi-cuadrada con p grados de libertad. (Puma, 2018)

2.5. DESERCIÓN ESTUDIANTIL

2.5.1. Definición de Deserción

La deserción estudiantil es el abandono total definitivo de las aulas de clase por diferentes motivos o razones y la no continuidad en la formación académica, que la sociedad quiere esperanzados en que termine felizmente los estudios universitarios (Gabriel et ál., 1999).

Un estudiante se considera desertor de programa si abandona un programa académico durante dos periodos consecutivos, se registra matriculado en otro programa sin que siga vinculado al primero o culmina su plan de estudios y no obtiene su grado durante más de dos periodos después de su culminación (Cruz y Ortega, 2008).

La deserción estudiantil es el abandono definitivo a una institución de educación superior en un momento determinado. El desarrollo de la deserción es todo un proceso, a veces es lento,

va creciendo y reforzándose en el interior del estudiante universitario, quien lo manifiesta en la deserción definitiva ya sea para bien o para mal del estudiante y su entorno.

2.5.2. Clases y actores de Deserción

Según Gabriel et ál. (1999). Existen varias clases de deserción en educación, no excluyentes entre sí:

- Deserción total: Abandono definitivo de la formación académica individual.
- Deserción discriminada por causas: según la causa de la decisión.
- Deserción por facultad (escuela o departamento): cambio facultad – facultad.
- Deserción por programa: cambio de programa en una misma facultad.
- Deserción a primer semestre de carrera: por inadecuada adaptación a la vida universitaria.
- Deserción acumulada: sumatoria de deserciones en una institución.

Los actores de la deserción son los siguientes.

- Desertores: estudiantes que se hayan retirado de su formación universitaria.
- Padres de familia de desertores.
- Excompañeros de estudio: alumnos del semestre del cual se retiró el desertor.
- Profesores: quienes acompañaron al desertor en el proceso educativo hasta el semestre de retiro.
- Directivas y administradores académicos.

2.5.3. Factores asociados.

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

2.5.3.1. Factores individuales

Los factores individuales son: edad, género, ocurrencia de una calamidad o un problema doméstico, facilidades de integración social del individuo, expectativas personales insatisfechas, entre otras:

2.5.3.2. Factores socioeconómicos

Los factores socioeconómicos son: estrato social del estudiante, situación laboral del estudiante, situación laboral de sus padres, si depende económicamente de alguien, si tiene personas a su cargo, nivel educativo de sus padres, entorno familiar (número de miembros en el hogar, estado civil de los padres de familia, etc.).

2.5.3.3. Factores académicos

Los factores académicos son: acceso a orientación profesional antes de ingresar a la institución educativa, tipo de colegio de la secundaria (público o privado), rendimiento académico del individuo, metodologías de estudio utilizadas, calificación en el examen de admisión, carga académica (número de créditos al semestre), descontento con el programa académico, etc.

2.5.3.4. Factores institucionales

Los factores institucionales (categoría más relacionada con la institución) son: normatividad académica, recursos con los que dispone la institución educativa, las relaciones del individuo con el profesorado y los demás estudiantes, calidad del programa en el que está inscrito, los programas anti-deserción o de acción afirmativa que implementen las instituciones educativas, etc.

DISEÑO METODOLÓGICO

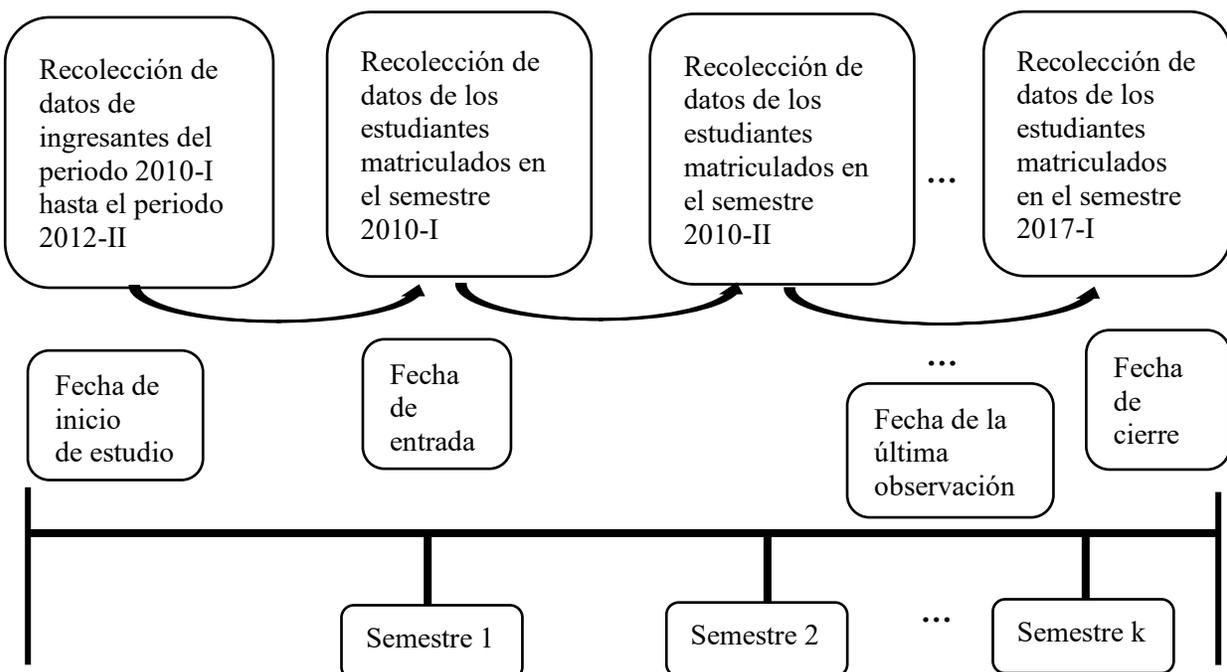
3.1. Tipo de Investigación

La investigación es de tipo descriptivo, explicativo, porque además de describir la deserción estudiantil, busca explicar que factores están asociados a la ocurrencia del evento.

El presente trabajo de investigación tiene un enfoque cuantitativo porque se utiliza la recolección de datos y el análisis de los mismos para contestar preguntas de investigación y probar hipótesis formuladas.

3.2. Diseño de Investigación

Es un estudio de tipo no experimental, longitudinal o evolutivo, retrospectivo, porque 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 después analizar, examinar y explicar cambios a través del tiempo en los estudiantes que ingresaron en el periodo 2010-I hasta el periodo 2012-II.



3.3. Población

La población para el presente estudio lo conforman 131 estudiantes que ingresaron a la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco que son registrados semestralmente durante los semestres 2010-I, 2010-II, 2011-I, 2011-II, 2012-I y 2012-II en el centro de cómputo de la UNSAAC.

Tabla 5

Número de estudiantes que ingresaron a la Escuela Profesional de Matemática según semestre de ingreso

Semestre de Ingreso	Número de estudiantes
2010-I	28
2010-II	8
2011-I	31
2011-II	16
2012-I	30
2012-II	18
Total	131

Criterios de inclusión:

1. 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.
2. Estudiantes con información completa requerida para el estudio.

Criterios de exclusión

3. 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.

3.4. Variables de estudio

3.4.1. *Variable dependiente*

- Deserción

3.4.2. *Variable independiente*

- Edad
- Sexo
- Tipo de colegio
- Tipo de modalidad de ingreso
- Ocupación del jefe de hogar
- Ingreso económico familiar
- Número de hermanos
- Situación familiar
- Lugar de procedencia
- El estudiante trabaja

3.5. Técnicas e instrumentos de recolección de datos

Técnica

Investigación documental de las fichas socioeconómicas (anexo 01) de los estudiantes que llenan al ingresar a la UNSAAC.

Instrumento

Se elaboro una ficha de registro de datos (anexo 02), en el que se recopilo la información correspondiente de los estudiantes que ingresaron a la Escuela Profesional de Matemática de la

Universidad Nacional de San Antonio Abad del Cusco, en los semestres 2010-I, 2010-II, 2011-I, 2011-II, 2012-I y 2012-II.

3.6. Procedimiento de recolección de datos

La recolección de datos para el presente trabajo de investigación se realizó de la siguiente manera.

1. Se solicitó al director de la Escuela Profesional de Matemática para que nos presente ante el director del Centro de Cómputo de la UNSAAC mediante una solicitud para que nos facilite el acopio de información para el desarrollo del presente trabajo de investigación.
2. Se solicitó la autorización al director del Centro de Cómputo de la UNSAAC para las facilidades de acopio de información de los ingresantes de la Escuela Profesional de Matemática a partir del semestre 2010-I hasta el 2012-II y los estudiantes matriculados del semestre 2010-I hasta el 2017-I, y la relación de estudiantes ingresantes y matriculados por cada semestre fueron entregados por el encargado del Centro de Cómputo de la UNSAAC (con el fin de mantener la originalidad y confidencialidad de los datos).
3. Se solicitó la autorización al Jefe de la Unidad de Bienestar Universitario de la UNSAAC (Área de Asistencia Social), para extraer la información socioeconómica de los estudiantes que ingresaron a la Escuela Profesional de Matemática en los semestres 2010-I, 2010-II, 2011-I, 2011-II, 2012-I y 2012-II. La información se recopiló en presencia del encargado del Área de asistencia Social y se mantuvo en total confidencialidad por los investigadores.

3.7. Técnicas de Procesamiento y Análisis de datos

Una vez obtenida la información de los datos para el presente trabajo de investigación los datos están ordenados en tablas de doble entrada en Excel, relacionadas a las variables explicativas incluidas en el estudio.

Seguidamente se procedió con el procesamiento en el paquete estadístico Rstudio:

Primero se realizó la estimación de la función de supervivencia mediante el estimador no paramétrico de Kaplan-Meier aplicado a cada variable independiente, en función de la variable tiempo (en semestres) y para comparar las curvas de supervivencia se utiliza la prueba de Test Long-Rank.

Segundo hacemos la estimación de los parámetros mediante el modelo Semi paramétrico (modelo de regresión de Cox) para identificar a aquellas variables independientes que tengan contribución significativa en la decisión de abandonar, desertar o retirarse de la Universidad.

3.8. Equipos y Software

Para el análisis de los datos del presente trabajo de investigación se utilizó una laptop Lenovo, para la organización de los datos se utilizó Excel 2013 y para el procesamiento de los datos se utilizó el Software estadístico Rstudio.

RESULTADOS

4.1. Resultado del Modelo de Kaplan Meier

En esta sección se presenta los resultados obtenidos de la aplicación del método de Kaplan-Meier para los estudiantes de la Escuela Profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco. Se extrae el tiempo de supervivencia de un estudiante, el tiempo en el cual un estudiante permanece en la universidad, la mortalidad se adecua a la deserción y la censura es equivalente a la salida del estudiante de la Universidad por causas diferentes a la deserción (por ejemplo, conclusión de la carrera o la graduación).

A continuación, se muestran tablas y figuras del software Rstudio que muestran los resultados de diferencias estadísticas y las curvas de supervivencia, en este presente trabajo se utilizó un nivel de confianza del 95% y con un nivel de significancia del 0.05.

4.1.1. Deserción Estudiantil en la Escuela Profesional de Matemática (EPM)

Tabla 6

Resumen de Kaplan Meier (EPM)

Total	N°. Eventos	N°. de Censurados	Eventos %	Censurados %	Media na	0.95 LCL	0.95 UCL
131	81	50	61.8%	38.2%	4	3	6

La tabla 6 se refiere a la totalidad de estudiantes analizados en la cohorte 2010-I hasta la cohorte 2012-II, se estudian en total a 131 estudiantes, donde 81 estudiantes desertaron durante el periodo de análisis y 50 estudiantes son censurados; los casos censurados se refieren a los estudiantes que abandonaron el estudio sin haber desertado como por ejemplo la conclusión o la graduación del estudiante y desertor se refiere a los estudiantes que abandonaron sus estudios Universitarios por diferentes causas.

Del 100% de estudiantes analizados en el estudio de la cohorte 2010-I hasta la cohorte 2012-II, al 61.8% de estudiantes analizados les ocurrió el evento de interés (deserción) y el 38.2% del resto de los estudiantes fueron censurados; los casos censurados se refieren a los estudiantes que abandonaron el estudio sin haber desertado como por ejemplo la conclusión o la graduación del estudiante y desertor se refiere a los estudiantes que abandonaron sus estudios Universitarios por diferentes causas.

La mediana de supervivencia es 4 (IC95% 3 - 6), lo que indica que al 4 semestre de la carrera universitaria la mitad de los estudiantes han desertado en la escuela profesional de Matemática.

De igual modo la Dirección de sistemas de información unidad de estadística (2016) y Dirección de sistemas de información unidad de estadística (2017) de la UNSAAC manifiesta que al 64% y 68.62% respectivamente se presenta la deserción en estudiantes de la escuela profesional de Matemática, la información proporcionada es similar a los resultados obtenidos en este trabajo de investigación, donde al 61.8% de estudiantes les ocurrió el evento de interés (deserción) en la Escuela Profesional de Matemática.

Tabla 7

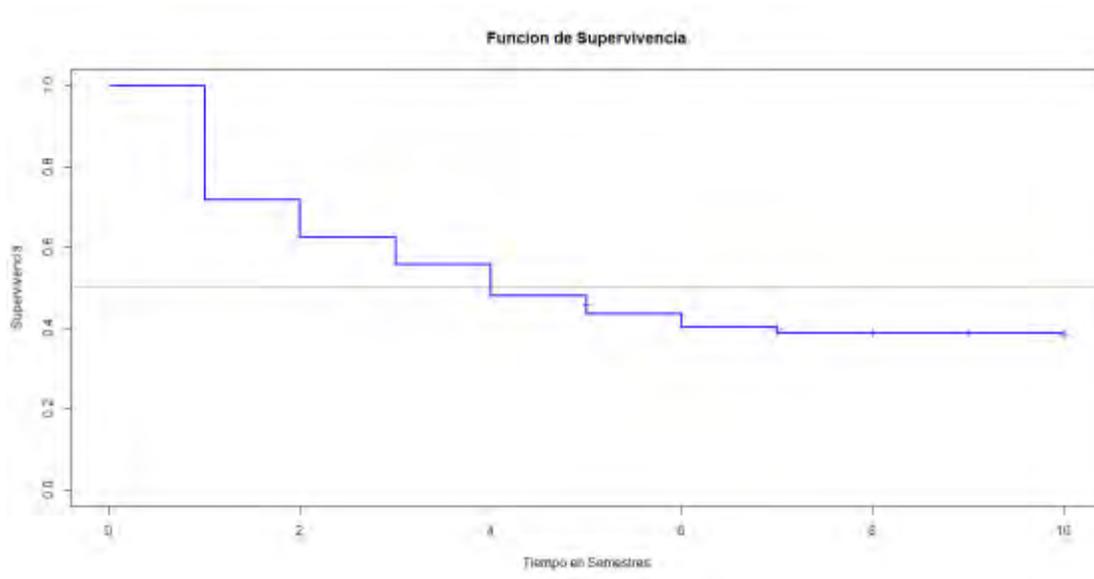
Los tiempos de supervivencia por semestres (EPM)

Tiempo	N°. Riesgo	N°. Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
0	0	0	1	1		
1	131	37	0.718	0.0393	0.644	0.799
2	94	12	0.626	0.0423	0.548	0.715
3	82	9	0.557	0.0434	0.478	0.649
4	73	10	0.481	0.0437	0.403	0.575
5	63	6	0.435	0.0433	0.358	0.529
6	55	4	0.403	0.043	0.327	0.497
7	51	2	0.388	0.0427	0.312	0.481
10	43	1	0.379	0.0427	0.304	0.472

En la tabla 7 se observa la estimación de la función de supervivencia utilizando el método de Kaplan-Meier respecto al tiempo en semestres, como podemos ver en la tabla en el primer semestre o semestre uno la probabilidad de que un estudiante sobreviva más de un semestre es 0.718 y la probabilidad de que un estudiante sobreviva más de cuatro semestres es 0.481 y la probabilidad de que un estudiante continúe sus estudios o sobreviva más de diez semestres es 0.379.

Figura 1

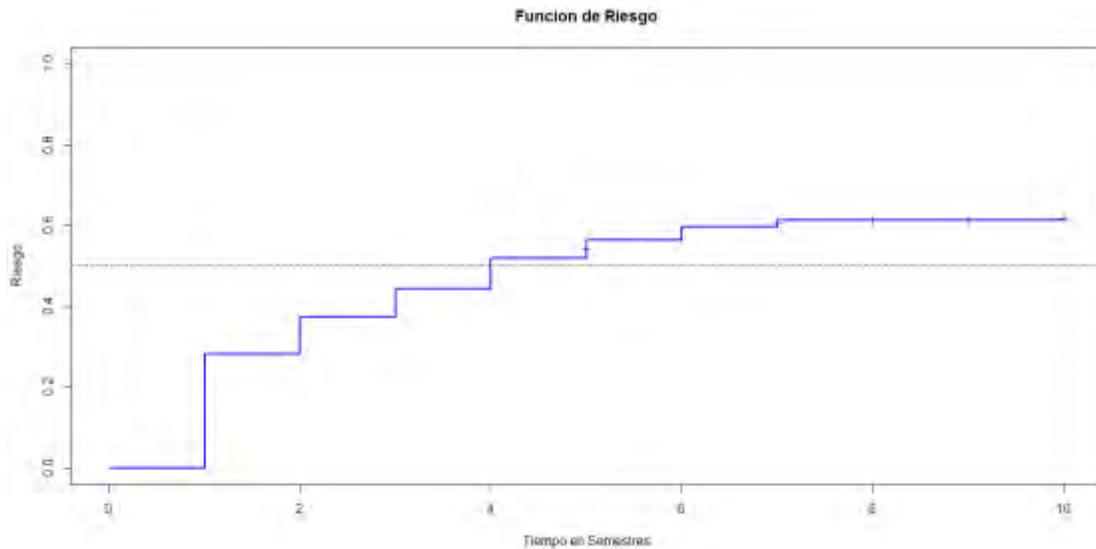
Función de supervivencia de la EPM



En la figura 1 se observa la curva de la función de supervivencia de los estudiantes de la Escuela Profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco, las líneas verticales indican los datos censurados. En el eje de las ordenadas se encuentra la supervivencia acumulada en términos de probabilidad (entre 0 y 1) y en el eje de las abscisas el tiempo en semestres. La gráfica nos explica que al transcurrir los semestres los estudiantes abandonan sus estudios Universitarios.

Figura 2

Función de riesgo de la EPM



En el Figura 2 se observa la curva de función de riesgo de los estudiantes de la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad del Cusco. En términos de probabilidad (entre 0 y 1) en el eje de las ordenadas se encuentra el riesgo acumulado y en el eje de las abscisas el tiempo en semestres.

4.1.1.1. Comparación de Curvas de Supervivencia.

Tabla 8

Resumen de Sexo

Sexo	N°. Total	N°. Eventos	N°. de Censurados	Eventos %	Censurado %
Mujer	39	22	17	56.4%	43.6%
Varón	92	59	33	64.1%	35.9%
Total	131	81	50	61.8%	38.2

En la tabla 8 se observa el resumen de la variable sexo (mujer y varón), del 100% de los estudiantes de la Escuela Profesional de Matemática de la cohorte 2010-I hasta la cohorte 2012-

II, durante el seguimiento del estudio el 56.4% del total de estudiantes mujeres le ocurrió el evento de interés (deserción) y 43.6% fueron censurados mientras tanto el 64.1% del total de estudiantes varones le ocurrió el evento de interés (deserción) y 35.9% son censurados. Se puede observar que a los varones les ocurrió más el evento de interés (deserción).

Según Puma (2018) en su trabajo de investigación indica que los estudiantes varones presentan mayor riesgo de deserción con respecto a las estudiantes mujeres con 61.0% y 51.6% respectivamente.

Asimismo, Barturén y Niño (2018) en su investigación menciona que un 29.5% de los estudiantes que abandonan la escuela profesional de Estadística fueron varones y 16.1% fueron mujeres.

Según Morales et al., (2017) en su investigación indica que las estudiantes varones presentaron mayor porcentaje de abandono de los estudios superiores de 28% con respecto a las mujeres 16%, se infiere que los estudiantes varones presentan una mayor tendencia hacia la desercion universitaria, resultado similar al que hemos obtenido en el que 64.1% de varones presentan mayor riesgo de deserción que las mujeres.

Tabla 9

Los tiempos de supervivencia, la proporción de estudiantes sobrevivientes en cada punto del tiempo en semestres para la variable sexo (mujer-varón)

Sexo=MUJER						
Tiempo	N°. Riesgo	N°. Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	39	10	0.744	0.0699	0.618	0.894
2	29	3	0.667	0.0755	0.534	0.832
3	26	1	0.641	0.0768	0.507	0.811
4	25	2	0.590	0.0788	0.454	0.766
5	23	2	0.538	0.0798	0.403	0.720
6	19	3	0.453	0.0809	0.320	0.643

10	13	1	0.419	0.0819	0.285	0.614
Sexo=VARON						
Tiempo	N°. Riesgo	N°. Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	92	27	0.707	0.0475	0.619	0.806
2	65	9	0.609	0.0509	0.517	0.717
3	56	8	0.522	0.0521	0.429	0.634
4	48	8	0.435	0.0517	0.344	0.549
5	40	4	0.391	0.0509	0.303	0.505
6	36	1	0.38	0.0506	0.293	0.494
7	35	2	0.359	0.0500	0.273	0.471

Tabla 10

Comparaciones Globales

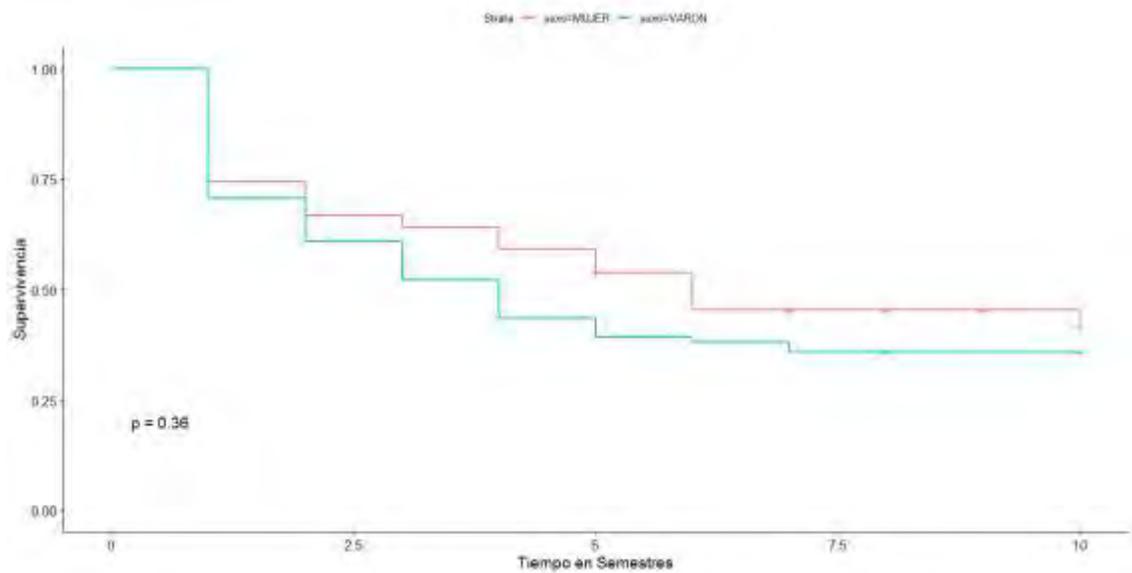
	Chi-cuadrado	gl	p-value.
Log-Rank	0.9	1	0.4

En la tabla 10, se muestra el resultado de la prueba estadística para la igualdad de las funciones de supervivencia, para comparar la función de supervivencia de la variable sexo (varón-mujer).

Usamos la prueba de Log-Rank con un valor estadístico de 0.9 y un nivel de significancia $\alpha = 0.05$, como $p - value = 0.4$ es mayor que $\alpha = 0.05$ no se rechaza (H_0). Se concluye que no existe diferencia significativa de la variable sexo (varón-mujer).

Figura 3

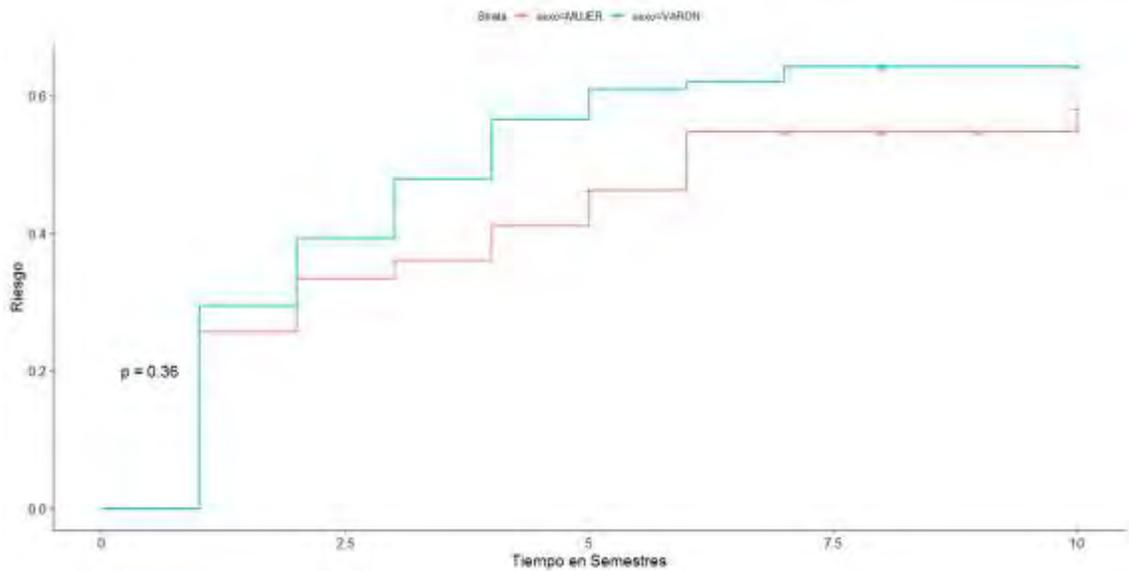
Función de supervivencia Sexo



En la figura 3 se observa la curva de la función de supervivencia de la variable sexo (mujer y varón), de los estudiantes de la Escuela Profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco, a pesar que no existe diferencia significativa entre las curvas de supervivencia de la variable sexo, se puede observar que hay una separación entre las curvas de supervivencia, aparentemente las estudiantes mujeres sobreviven más que los varones.

Figura 4

Función de riesgo de Sexo



En la figura 4 se muestran las curvas de la función de riesgo de la variable sexo (mujer y varón), de los estudiantes de la escuela profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco. Se observa que aparentemente los varones presentan mayor riesgo a desertar o abandonar sus estudios superiores y esta diferencia no es significativa.

Tabla 11

Resumen Tipo de Colegio

Tipo de colegio	N°. Total	N°. Eventos	N°. de Censurados	Eventos %	Censurado %
Estatal	110	65	45	59.1%	40.9%
Particular	21	16	5	76.2%	23.8%
Total	131	81	50	61.8%	38.2%

En la tabla 11 se observa el resumen tipo de colegio estatal y particular, del 100% de los estudiantes de la escuela profesional de Matemática de la cohorte 2010-I hasta la cohorte 2012-II, durante el periodo de seguimiento al 59.1% del total de estudiantes que provienen de colegios

estatales les ocurrió el evento de interés (deserción) y 40.9% fueron censurados, 76.2% del total de estudiantes que provienen de colegios particulares les ocurrió el evento de interés (deserción) y 23.8% fueron censurados. Aparentemente se puede observar que los estudiantes que provienen de colegios particulares les ocurrió más el evento de interés (deserción).

Según Puma (2018), al 63.7% de los estudiantes provenientes de colegios nacionales presentan mayor riesgo de deserción que al de colegios particulares (38.3%).

Asimismo Barturén y Niño (2018) indica que el 25% de estudiantes que abandonan la Escuela Profesional de Estadística provienen de colegios públicos y el 20.5% de colegios particulares.

La información proporcionada por los investigadores antes mencionadas son, contradictorios a los resultados obtenidos en este trabajo de investigación donde obtuvimos que el 76.2% de estudiantes provenientes de colegio particular presentan mayor riesgo de deserción que los estudiantes de colegio estatal (59.1%).

Tabla 12

Los tiempos de supervivencia, la proporción de estudiantes sobrevivientes en cada punto del tiempo en semestres para la variable tipo de colegio (estatal y particular)

Tipo de colegio = Estatal						
Tiempo	n. Riesgo	n. Evento	Supervivencia	std.err	L.I. 95%	L.S. 95%
1	110	30	0.727	0.0425	0.649	0.815
2	80	8	0.655	0.0453	0.571	0.750
3	72	8	0.582	0.0470	0.497	0.682
4	64	10	0.491	0.0477	0.406	0.594
5	54	5	0.445	0.0474	0.362	0.549
6	47	3	0.417	0.0471	0.334	0.520
7	44	1	0.408	0.0470	0.325	0.511
Tipo de colegio = Particular						
Tiempo	Nº. Riesgo	Nº. Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	21	7	0.667	0.1029	0.493	0.902

2	14	4	0.476	0.1090	0.304	0.746
3	10	1	0.429	0.1080	0.262	0.702
5	9	1	0.381	0.1060	0.221	0.657
6	8	1	0.333	0.1029	0.182	0.610
7	7	1	0.286	0.0986	0.145	0.562
10	5	1	0.229	0.0940	0.102	0.512

Tabla 13

Comparaciones Globales

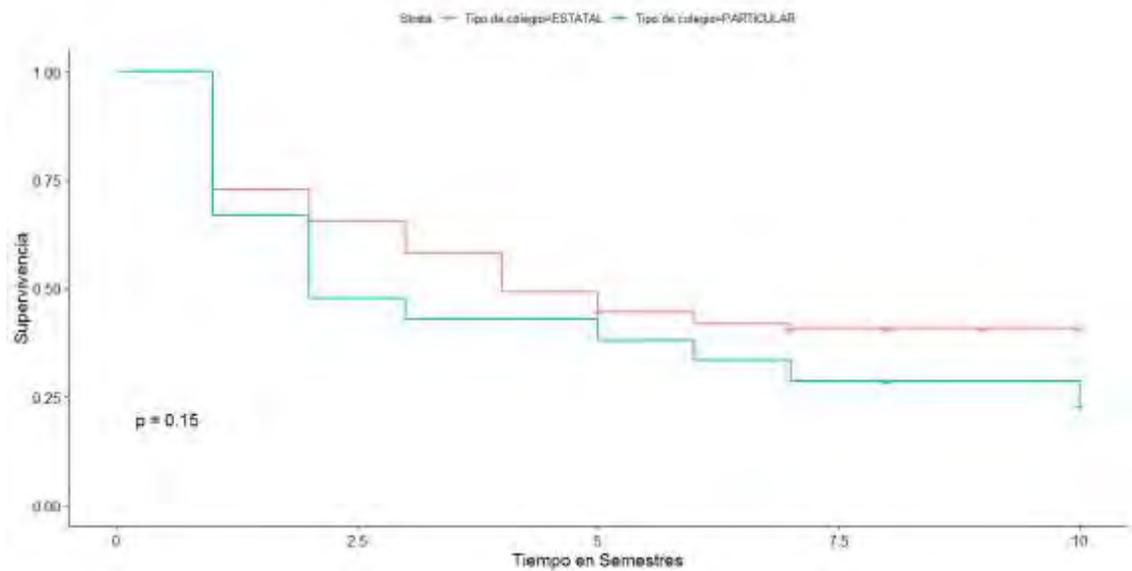
	Chi-cuadrado	gl	p-value.
Log-Rank	2.1	1	0.2

En la tabla 13, se muestra el resultado de la prueba estadística para la igualdad de las funciones de supervivencia, para comparar la función de supervivencia de la variable tipo de colegio estatal y particular.

Usamos la prueba de Log-Rank con un valor estadístico de 2.1 y un nivel de significancia $\alpha = 0.05$, como $p - value = 0.2$ es mayor que $\alpha = 0.05$ no se rechaza (H_0). Se concluye que no existe diferencia significativa en el tipo de colegio estatal y particular.

Figura 5

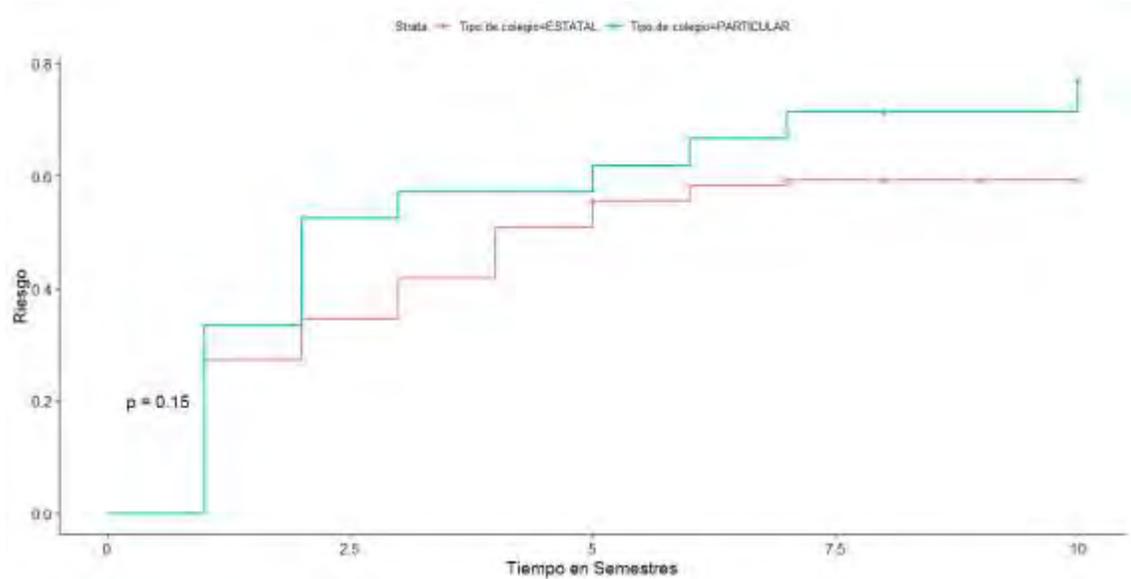
Función de supervivencia tipo de Colegio



En la figura 5 se presentan las curvas de la función de supervivencia de tipo de colegio estatal y particular de los estudiantes de la escuela profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco, a pesar que no existe diferencia significativa entre las curvas de supervivencia de tipo de colegio, se puede observar que hay una separación entre las curvas de supervivencia y aparentemente los estudiantes que provienen de colegio estatal sobreviven más que los estudiantes que provienen de colegio particular.

Figura 6

Función de Riesgo Tipo de Colegio



En la figura 6 se presentan las curvas de la función de riesgo de tipo de colegio estatal y particular de los estudiantes de la escuela profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco. Se observa que aparentemente los estudiantes que provienen de colegios particulares presentan mayor riesgo a desertar o abandonar sus estudios superiores y esta diferencia no es significativa.

Tabla 14

Resumen Tipo de Modalidad de Ingreso

Tipo de Modalidad de Ingreso	N°. Total	N°. Eventos	N°. de Censurados	Eventos %	Censurado %
CEPRU	14	7	7	50.0%	50.0%
Dirimencia	38	28	10	73.7%	26.3%
Ordinario	52	31	21	59.6%	40.4%
Primera Oportunidad	27	15	12	55.6%	44.4%
Total	131	81	50	61.8%	38.2%

En la tabla 14 se observa el resumen de tipo de modalidad de ingreso (CEPRU, dirimencia, ordinario y primera oportunidad), del 100% de los estudiantes de la Escuela Profesional de Matemática de la cohorte 2010-I hasta la cohorte 2012-II, durante el seguimiento del estudio al 50% del total de estudiantes que ingresaron por modalidad de CEPRU les ocurrió el evento de interés (deserción) y 50% fueron censurados, al 73.7% del total de estudiantes que ingresaron por modalidad de dirimencia les ocurrió el evento de interés (deserción) y 26.3% fueron censurados, al 59.6% del total de estudiantes que ingresaron por modalidad de ordinario les ocurrió el evento de interés (deserción) y 40.4% fueron censurados mientras al 55.6% del total de estudiantes que ingresaron por modalidad de primera oportunidad les ocurrió el evento de interés (deserción) y 44.4% fueron censurados. Aparentemente se puede observar que los estudiantes que ingresaron por modalidad de dirimencia les ocurrió más el evento de interés (deserción) con un 73.7%.

Tabla 15

Los tiempos de supervivencia, la proporción de estudiantes sobrevivientes en cada punto del tiempo en semestres para tipo de modalidad de ingreso (CEPRU, dirimencia, ordinario y primera oportunidad)

Tipo de modalidad de ingreso = CEPRU						
Tiempo	N°. Riesgo	N°.Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	14	1	0.929	0.0688	0.803	1.000
2	13	1	0.857	0.0935	0.692	1.000
3	12	1	0.786	0.1097	0.598	1.000
4	11	2	0.643	0.1281	0.435	0.950
5	9	2	0.500	0.1336	0.296	0.844
Tipo de modalidad de ingreso=dirimencia						
Tiempo	N°.Riesgo	N°.Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	38	11	0.711	0.0736	0.580	0.870
2	27	5	0.579	0.0801	0.441	0.759

3	22	5	0.447	0.0807	0.314	0.637
4	17	2	0.395	0.0793	0.266	0.585
5	15	3	0.316	0.0754	0.198	0.504
6	11	1	0.287	0.0738	0.173	0.475
10	9	1	0.255	0.0722	0.147	0.444
Tipo de modalidad de ingreso = ordinario						
Tiempo	Nº.Riesgo	Nº.Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	52	18	0.654	0.0660	0.537	0.797
2	34	4	0.577	0.0685	0.457	0.728
3	30	2	0.538	0.0691	0.419	0.693
4	28	3	0.481	0.0693	0.362	0.638
6	25	2	0.442	0.0689	0.326	0.600
7	23	2	0.404	0.0680	0.290	0.562
Tipo de modalidad de ingreso = primera oportunidad						
Tiempo	Nº.Riesgo	Nº.Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	27	7	0.741	0.0843	0.593	0.926
2	20	2	0.667	0.0907	0.511	0.870
3	18	1	0.630	0.0929	0.471	0.841
4	17	3	0.519	0.0962	0.361	0.746
5	14	1	0.481	0.0962	0.326	0.712
6	12	1	0.441	0.0962	0.288	0.676

Tabla 16

Comparaciones Globales

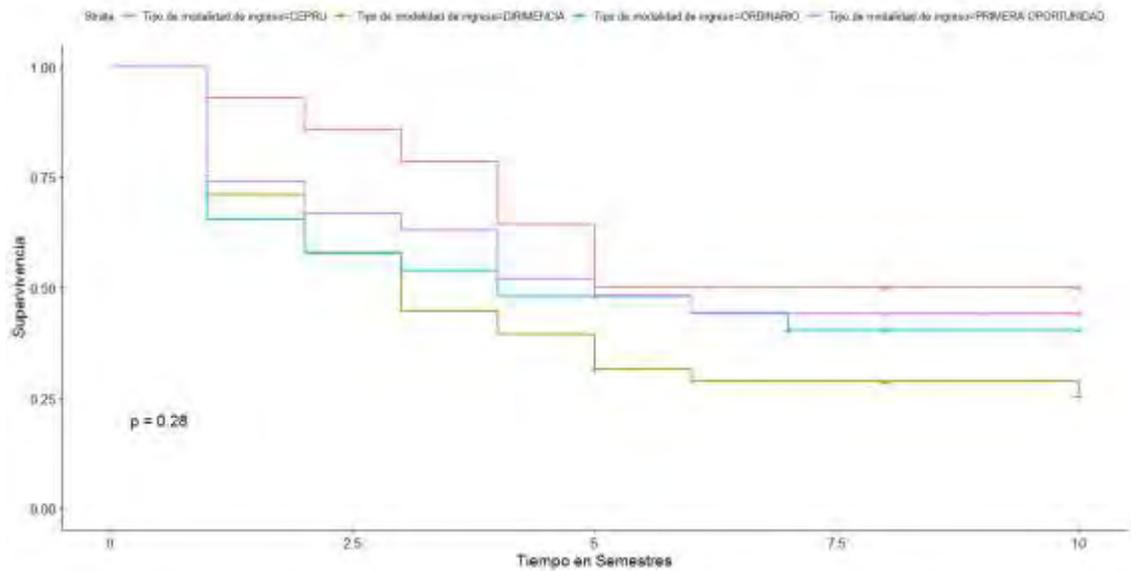
	Chi- cuadrado	Gl	p-value.
Log-Rank	3.8	3	0.3

En la tabla 16, se muestra el resultado de la prueba estadística para la igualdad de las funciones de supervivencia, para comparar la función de supervivencia tipo de modalidad de ingreso CEPRU, dirimencia, ordinario y primera oportunidad.

Usamos la prueba de Log-Rank con un valor estadístico de 3.8 y un nivel de significancia $\alpha = 0.05$, como $p - value = 0.3$ es mayor que $\alpha = 0.05$ no se rechaza (H_0). Se concluye que no existe diferencia significativa para el tipo de modalidad de ingreso CEPRU, dirimencia, ordinario y primera oportunidad.

Figura 7

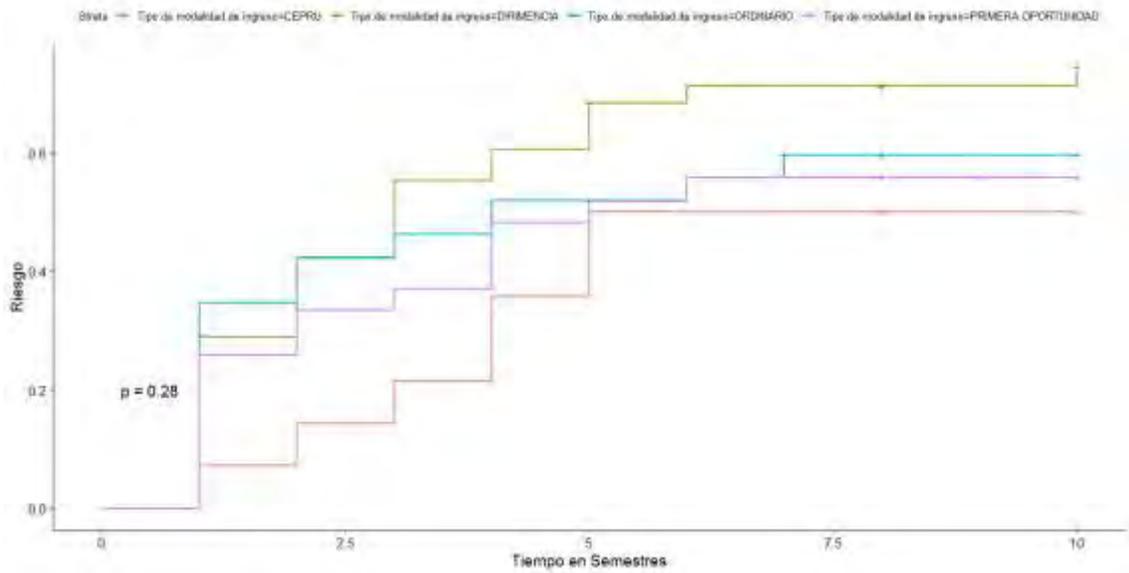
Función de supervivencia Tipo de Modalidad de Ingreso



En la figura 7 se observa la curva de la función de supervivencia para el tipo de modalidad de ingreso (CEPRU, dirimencia, ordinario y primera oportunidad) de los estudiantes de la Escuela Profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco, a pesar que no existe diferencia significativa entre las curvas de supervivencia de tipo de modalidad de ingreso, se puede observar que hay una separación entre las curvas de supervivencia y aparentemente los estudiantes que ingresan por CEPRU son los que sobreviven más que los estudiantes que ingresan por otras modalidades.

Figura 8

Función de Riesgo Tipo de Modalidad de Ingreso



En la figura 8 se muestran las curvas de la función de riesgo para el tipo de modalidad de ingreso (CEPRU, dirimencia, ordinario y primera oportunidad), de los estudiantes de la Escuela Profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco. Se observa que aparentemente los estudiantes que ingresan por la modalidad dirimencia tienen mayor riesgo a desertar o abandonar sus estudios superiores en comparación a los otros tipos de modalidad de ingreso y esta diferencia no es significativa.

Tabla 17

Resumen de Ocupación del Jefe de Hogar

Ocupación del jefe de hogar	Nº.Total	Nº. Eventos	Nº. de Censurados	Eventos %	Censurado %
No Profesional	111	65	46	58.6%	41.4%
Profesional	20	16	4	80.0%	20.0%
Total	131	81	50	61.8%	38.2%

En la tabla 17 se observa el resumen de ocupación del jefe de hogar (no profesional y profesional), del 100% de los estudiantes de la Escuela Profesional de Matemática de la cohorte 2010-I hasta la cohorte 2012-II, durante el periodo de seguimiento al 80% del total de estudiantes provenientes de padres profesionales le ocurrió el evento de interés (deserción) y el 20 % fueron censurados; al 58.6% del total de estudiantes provenientes de padres no profesionales le ocurrió el evento de interés (deserción) y el 41.4% fueron censurados.

Tabla 18

Los tiempos de supervivencia, la proporción de estudiantes sobrevivientes en cada punto del tiempo en semestres para la variable ocupación del jefe de hogar (no profesional y profesional)

Ocupación del jefe de hogar = no profesional						
Tiempo	N°.Riesgo	N°.Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	111	27	0.757	0.0407	0.681	0.841
2	84	9	0.676	0.0444	0.594	0.769
3	75	8	0.604	0.0464	0.519	0.702
4	67	10	0.514	0.0474	0.428	0.615
5	57	6	0.459	0.0473	0.376	0.562
6	50	4	0.423	0.0470	0.340	0.526
10	41	1	0.412	0.0469	0.330	0.515
Ocupación del jefe de hogar = profesional						
Tiempo	N°.Riesgo	N°.Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	20	10	0.5	0.112	0.3226	0.775
2	10	3	0.35	0.107	0.1926	0.636
3	7	1	0.3	0.102	0.1536	0.586
7	5	2	0.18	0.090	0.0676	0.480

Tabla 19*Comparaciones Globales*

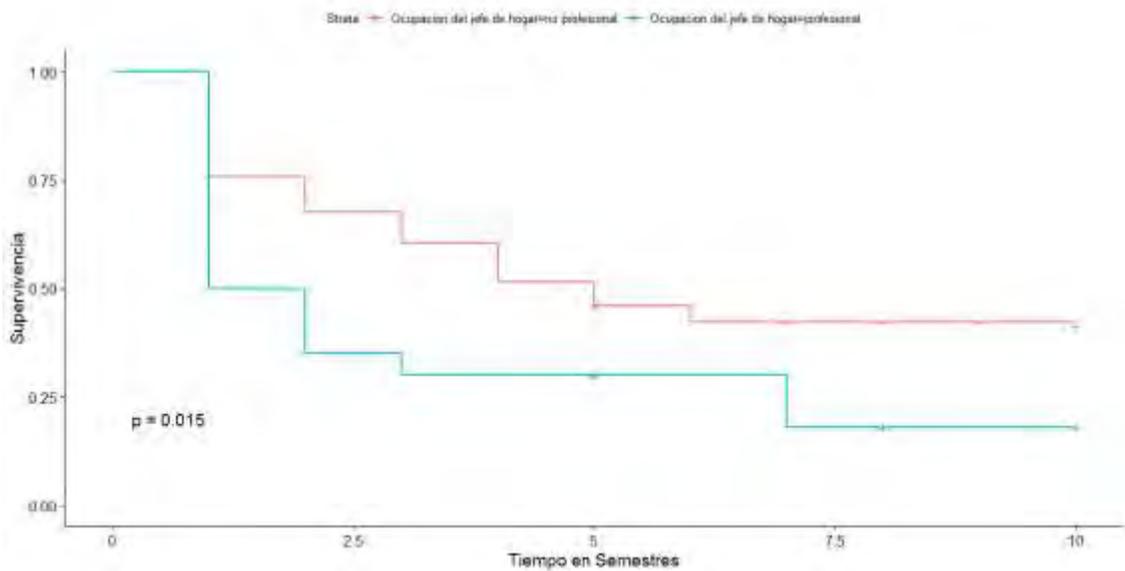
	Chi-cuadrado	Gl	p-value.
Log-Rank	5.9	1	0.02

En la tabla 19, se muestra el resultado de la prueba estadística para la igualdad de las funciones de supervivencia, para comparar la función de supervivencia de ocupación del jefe de hogar (no profesional y profesional).

Usamos la prueba de Log-Rank con un valor estadístico de 5.9 y un nivel de significancia $\alpha = 0.05$, como $p - value = 0.02$ es menor que $\alpha = 0.05$ se rechaza (H_0). Se concluye que existe diferencia significativa para ocupación del jefe de hogar (no profesional y profesional).

Figura 9

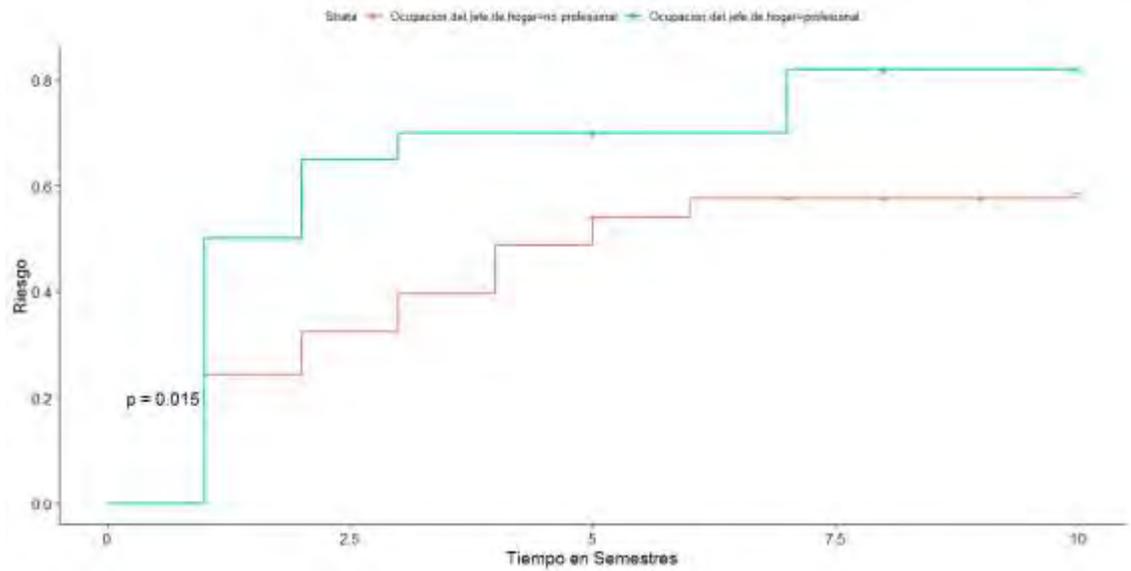
Función de Supervivencia Ocupación del jefe de hogar



En la figura 9, al observar la curva de la función de supervivencia para ocupación del jefe de hogar (no profesional y profesional), de los estudiantes de la escuela profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco, existe diferencia significativa entre las curvas de supervivencia, se puede observar que hay una separación entre las curvas de supervivencia y los estudiantes que provienen de jefe de hogar no profesional sobreviven más que los estudiantes que provienen de jefe de hogar profesionales.

Figura 10

Función de Riesgo Ocupación del jefe de hogar



En la figura 10 se muestran las curvas de la función de riesgo para ocupación del jefe de hogar (no profesional y profesional), de los estudiantes de la Escuela Profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco. Se observa que existe diferencia significativa entre las curvas de la función de riesgo y los estudiantes que tienen mayor riesgo a desertar o abandonar sus estudios superiores son los que provienen de jefe de hogar profesionales.

Tabla 20

Resumen de Ingreso Económico familiar

Ingreso Económico familiar	N°. Total	N°. Eventos	N°. de Censurados	Eventos %	Censurado %
Sueldo Alto	9	4	5	44.4%	55.6%
Sueldo Medio	42	29	13	69.0%	31.0%
Sueldo Básico	80	48	32	60.0%	40.0%
Total	131	81	50	61.8%	38.2%

En la tabla 20, se observa el resumen de ingreso económico familiar (sueldo alto, sueldo medio y sueldo básico), del 100% de los estudiantes de la Escuela Profesional de Matemática de la cohorte 2010-I hasta la cohorte 2012-II, durante el periodo de seguimiento al 69.0% de los estudiantes provenientes de padres con sueldo medio les ocurrió el evento de interés (deserción) y el 31.0% fueron censurados, al 60.0% de los estudiantes provenientes de padres con sueldo básico les ocurrió el evento de interés (deserción) y el 40.0% fueron censurados; y al 44.4% de los estudiantes provenientes de ingreso económico familiar con sueldo alto les ocurrió el evento de interés (deserción) y el 55.6% fueron censurados.

Tabla 21

Los tiempos de supervivencia, la proporción de estudiantes sobrevivientes en cada punto del tiempo en semestres para la variable ingreso económico familiar (sueldo alto, sueldo básico y sueldo medio)

Ingreso económico familiar = sueldo alto						
Tiempo	N°.Riesgo	N°.Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	9	1	0.889	0.105	0.706	1
3	8	1	0.778	0.139	0.549	1
6	6	1	0.648	0.165	0.393	1
7	5	1	0.519	0.176	0.267	1
Ingreso económico familiar = sueldo básico						
Tiempo	N°.Riesgo	N°.Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	80	23	0.713	0.0506	0.620	0.819
2	57	7	0.625	0.0541	0.527	0.741
3	50	4	0.575	0.0553	0.476	0.694
4	46	9	0.463	0.0557	0.365	0.586
5	37	3	0.425	0.0553	0.329	0.548
6	33	1	0.412	0.0551	0.317	0.536
10	29	1	0.398	0.0550	0.304	0.522
Ingreso económico familiar = sueldo medio						
Tiempo	N°.Riesgo	N°.Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	42	13	0.690	0.0713	0.564	0.845
2	29	5	0.571	0.0764	0.440	0.743

3	24	4	0.476	0.0771	0.347	0.654
4	20	1	0.452	0.0768	0.324	0.631
5	19	3	0.381	0.0749	0.259	0.560
6	16	2	0.333	0.0727	0.217	0.511
7	14	1	0.31	0.0713	0.197	0.486

Tabla 22

Comparaciones Globales

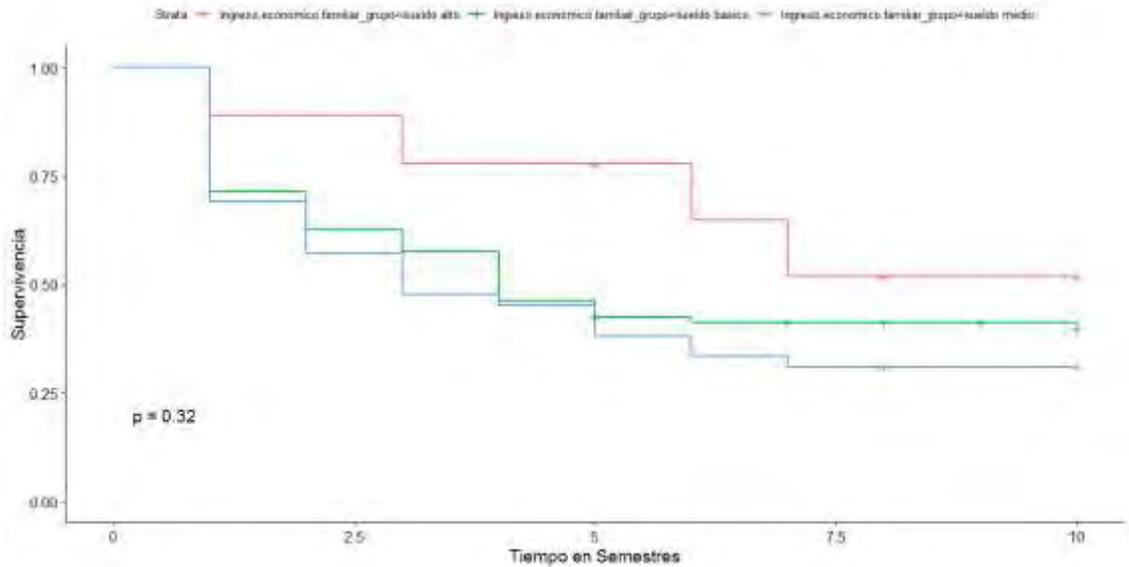
	Chi-cuadrado	G1	p-value
Log-Rank	2.3	2	0.3

En la tabla 22, se muestra el resultado de la prueba estadística para la igualdad de las funciones de supervivencia, para comparar la función de supervivencia de ingreso económico familiar (sueldo alto, sueldo básico y sueldo medio).

Usamos la prueba de Log-Rank con un valor estadístico de 2.3, y un nivel de significancia $\alpha = 0.05$, como $p - value = 0.3$ es mayor que $\alpha = 0.05$ no se rechaza la (H_0). Se concluye que no existe diferencia significativa para el ingreso económico familiar (sueldo alto, sueldo básico y sueldo medio).

Figura 11

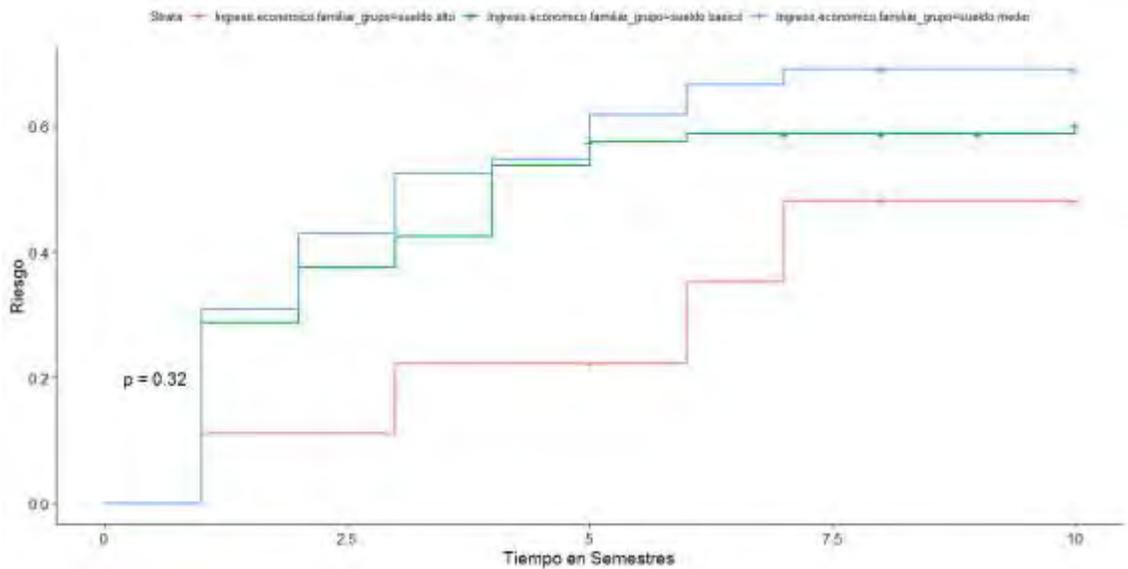
Función de Supervivencia Ingreso Económico familiar



En la figura 11, al observar la curva de la función de supervivencia de ingreso económico familiar (sueldo alto, sueldo básico y sueldo medio), de los estudiantes de la Escuela Profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco, a pesar que no existe diferencia significativa entre las curvas de supervivencia, se puede observar que hay una separación entre las curvas de supervivencia y aparentemente los estudiantes que provienen de ingreso económico familiar con sueldo alto sobreviven más que los estudiantes que provienen de ingreso económico familiar con sueldo básico y medio.

Figura 12

Función de Riesgo Ingreso Económico Familiar



En la figura 12 se muestran las curvas de la función de riesgo de ingreso económico familiar (sueldo alto, sueldo medio y sueldo básico), de los estudiantes de la Escuela Profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco. A pesar que no existe diferencia significativa entre las curvas de la función de riesgo, se observa que aparentemente los estudiantes que provienen de ingreso económico familiar con un sueldo medio son los que tienen mayor riesgo a desertar o abandonar sus estudios Universitarios.

Tabla 23

Resumen del Número de Hermanos

Número de Hermanos	Nº. Total	Nº. Eventos	Nº. de Censurados	Eventos %	Censurado %
0 a 1	38	23	15	60.5%	39.5%
2 a 3	62	38	24	61.3%	38.7%
más de 3	31	20	11	64.5%	35.5%
Total	131	81	50	61.8%	38.2%

En la tabla 23 se observa el resumen para el número de hermanos (0 a 1, 2 a 3 y más de 3), del 100% de los estudiantes de la Escuela Profesional de Matemática de la cohorte 2010-I hasta la cohorte 2012-II, durante el periodo de seguimiento al 64.5% del total de estudiantes con más de 3 hermanos le ocurrió el evento de interés (deserción) y el 35.5% fueron censurados, al 61.3% del total de estudiantes con (2 a 3) hermanos le ocurrió el evento de interés (deserción) y el 38.7% fueron censurados; y al 60.5% del total de estudiantes con (0 a 1) hermano le ocurrió el evento de interés (deserción) y el 39.5% fueron censurados.

Tabla 24

Los tiempos de supervivencia, la proporción de estudiantes sobrevivientes en cada punto del tiempo en semestres para la variable número de hermanos (0-1, 2-3 y más de 3)

Número de hermanos = 0-1						
Tiempo	Nº.Riesgo	Nº.Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	38	15	0.605	0.0793	0.468	0.782
2	23	5	0.474	0.0810	0.339	0.662
3	18	1	0.447	0.0807	0.314	0.637
4	17	1	0.421	0.0801	0.290	0.611
7	15	1	0.393	0.0795	0.264	0.584
Número de hermanos = 2-3						
Tiempo	Nº.Riesgo	Nº.Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	62	13	0.790	0.0517	0.695	0.898
2	49	5	0.71	0.0576	0.605	0.832
3	44	3	0.661	0.0601	0.553	0.790
4	41	7	0.548	0.0632	0.438	0.687
5	34	6	0.452	0.0632	0.343	0.594
6	27	3	0.401	0.0625	0.296	0.545
10	22	1	0.383	0.0622	0.279	0.527
Número de hermanos = más de 3						
Tiempo	Nº.Riesgo	Nº.Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	31	9	0.710	0.0815	0.567	0.889
2	22	2	0.645	0.0859	0.497	0.838
3	20	5	0.484	0.0898	0.336	0.696
4	15	2	0.419	0.0886	0.277	0.635
6	13	1	0.387	0.0875	0.249	0.603
7	12	1	0.355	0.0859	0.221	0.570

Tabla 25

Comparaciones Globales

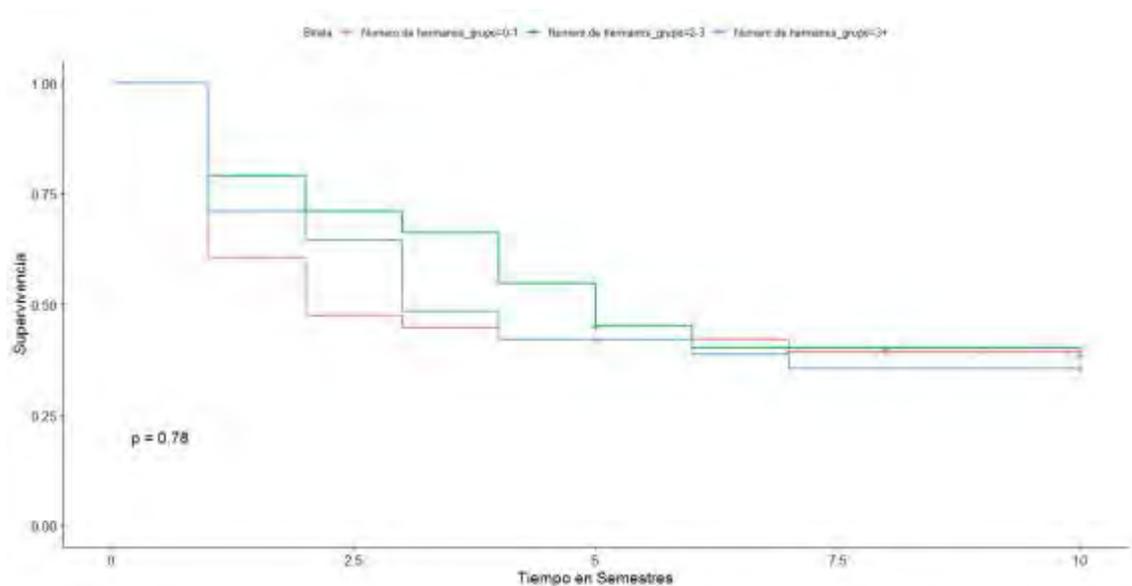
	Chi-cuadrado	Gl	p-value
Log-Rank	0.5	2	0.8

En la tabla 25, se muestra el resultado de la prueba estadística para la igualdad de las funciones de supervivencia, para comparar la función de supervivencia de número de hermanos (0-1, 2-3 y más de 3).

Usamos la prueba de Log-Rank con un valor estadístico de 0.5 y un nivel de significancia $\alpha = 0.05$, como $p - value = 0.8$ es mayor que $\alpha = 0.05$ no se rechaza (H_0). Se concluye que no existe diferencia significativa para el número de hermanos (0-1, 2-3 y más de 3).

Figura 13

Función de Supervivencia Número de Hermanos

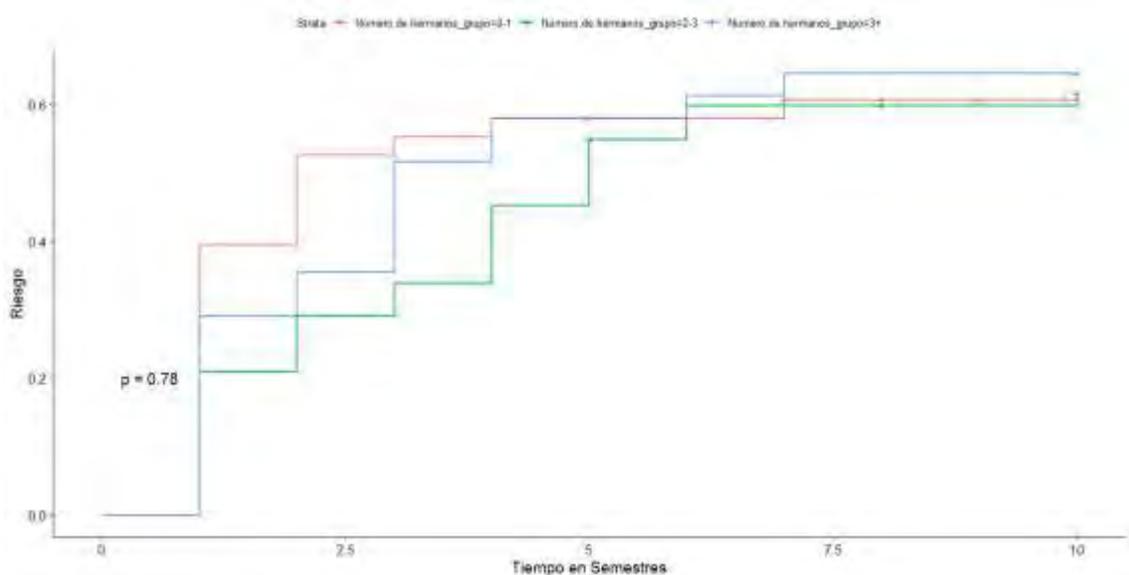


En la figura 13, al observar la curva de la función de supervivencia para el número de hermanos (0-1, 2-3 y más de 3), de los estudiantes de la Escuela Profesional de Matemática de la

Universidad Nacional San Antonio Abad del Cusco. A pesar que no existe diferencia significativa entre las curvas de supervivencia, se puede observar que hay una separación entre las curvas de supervivencia y aparentemente los estudiantes que tienen de 0 a 1 hermano son los que sobreviven más que los estudiantes que tienen 2 a 3 y más de 3 hermanos.

Figura 14

Función de Riesgo Número de Hermanos



En la figura 14, se muestran las curvas de la función de riesgo para el número de hermanos (0 a 1, 2 a 3 y más de 3), de los estudiantes de la Escuela Profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco. A pesar que no existe diferencia significativa entre las curvas de la función de riesgo, se observa que aparentemente los estudiantes que tienen más de 3 hermanos son los que tienen mayor riesgo a desertar o abandonar sus estudios Universitarios.

Tabla 26*Resumen de Situación Familiar*

Situación Familiar	Nº. Total	Nº. Eventos	Nº. de Censurados	Eventos %	Censurado %
Buena	41	27	14	65.9%	34.1%
Regular	71	40	31	56.3%	43.7%
Mala	3	2	1	66.7%	33.3%
Separados	16	12	4	75.0%	25.0%
Total	131	81	50	61.8%	38.2%

En la tabla 26 se observa el resumen de la situación familiar (buena, regular, mala y separados), del 100% de los estudiantes de la Escuela Profesional de Matemática de la cohorte 2010-I hasta la cohorte 2012-II, durante el periodo de seguimiento al 75.0% del total de estudiantes que provienen de situación familiar separados les ocurrió el evento de interés (deserción) y el 25.0% fueron censurados, al 66.7% del total de estudiantes que provienen de situación familiar mala les ocurrió el evento de interés (deserción) y el 33.3% fueron censurados, al 65.9% del total de estudiantes que provienen de situación familiar buena le ocurrió el evento de interés (deserción) y el 34.1% fueron censurados ; y al 56.3% del total de estudiantes que provienen de situación familiar regular le ocurrió el evento de interés (deserción) y el 43.7% fueron censurados.

Tabla 27

Los tiempos de supervivencia, la proporción de estudiantes sobrevivientes en cada punto del tiempo en semestres para la variable situación familiar (buena, regular, mala y separados)

Situación familiar= buena						
Tiempo	N°. Riesgo	N°. Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	41	10	0.756	0.0671	0.635	0.900
2	31	3	0.683	0.0727	0.554	0.841
3	28	3	0.61	0.0762	0.477	0.779
4	25	5	0.488	0.0781	0.356	0.668
5	20	3	0.415	0.0769	0.288	0.597
6	16	2	0.363	0.0755	0.241	0.546
7	14	1	0.337	0.0745	0.218	0.520
Situación familiar= mala						
Tiempo	N°. Riesgo	N°. Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
4	3	1	0.667	0.2720	0.2995	1
6	2	1	0.333	0.2720	0.0673	1
Situación familiar= regular						
Tiempo	N°.Riesgo	N°.Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	71	23	0.676	0.0555	0.576	0.794
2	48	6	0.592	0.0583	0.488	0.718
3	42	5	0.521	0.0593	0.417	0.651
4	37	3	0.479	0.0593	0.376	0.610
5	34	1	0.465	0.0592	0.362	0.597
6	33	1	0.451	0.0590	0.349	0.583
10	28	1	0.435	0.0591	0.333	0.567
Situación familiar= separados						
Tiempo	N°.Riesgo	N°.Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	16	4	0.750	0.108	0.5652	0.995
2	12	3	0.562	0.124	0.3651	0.867
3	9	1	0.500	0.125	0.3063	0.816
4	8	1	0.438	0.124	0.2510	0.763
5	7	2	0.312	0.116	0.1511	0.646

Tabla 28

Comparaciones Globales

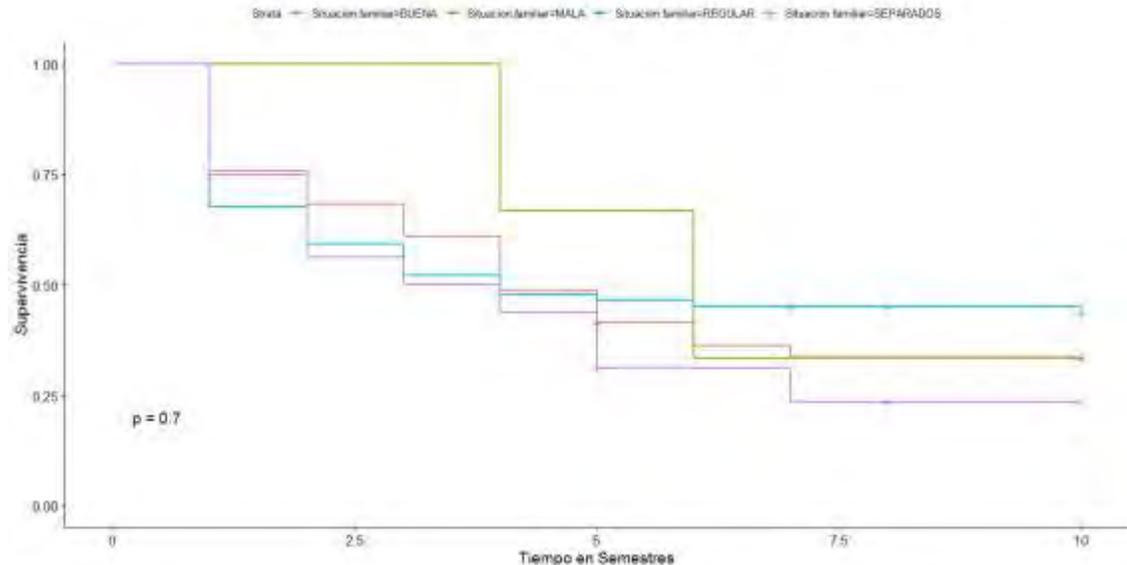
	Chi-cuadrado	Gl	p-value.
Log-Rank	1.4	3	0.7

En la tabla 28, se muestra el resultado de la prueba estadística para la igualdad de las funciones de supervivencia, para comparar la función de supervivencia de situación familiar (buena, regular, mala y separados).

Usamos la prueba de Log-Rank con un valor estadístico de 1.4 y un nivel de significancia $\alpha = 0.05$, como $p - value = 0.7$ es mayor que $\alpha = 0.05$ no se rechaza (H_0). Se concluye que no existe diferencia significativa para situación familiar buena, regular, mala y separados.

Figura 15

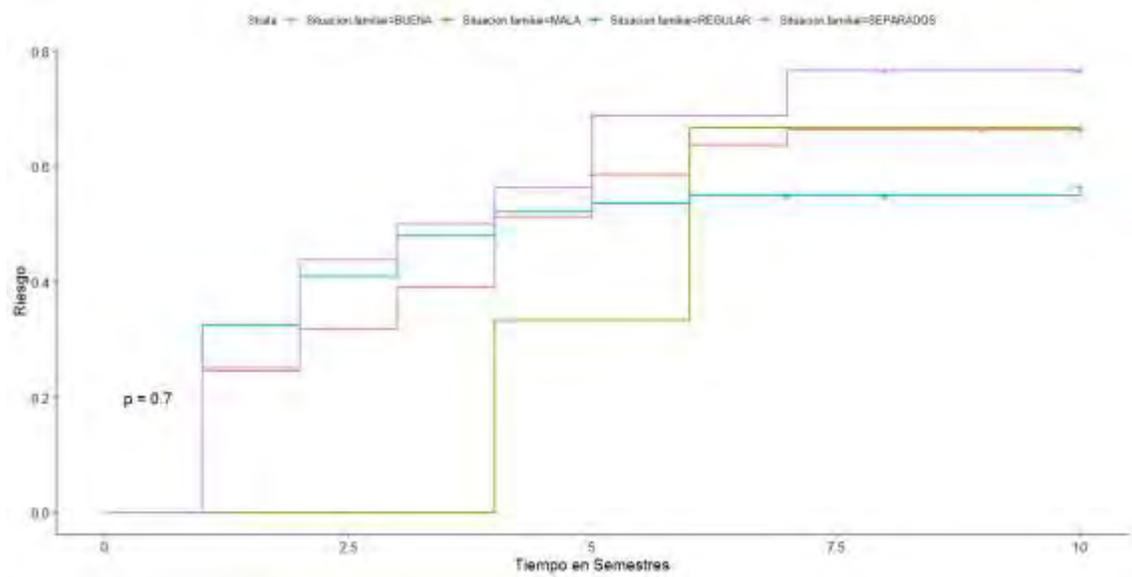
Función de Supervivencia Situación Familiar



En la figura 15, se observa la curva de la función de supervivencia para situación familiar (buena, mala, regular y separados), de los estudiantes de la Escuela Profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco. A pesar que no existe diferencia significativa entre las curvas de supervivencia, se puede observar que hay una separación entre las curvas de supervivencia y aparentemente los estudiantes que provienen de situación familiar regular son los que sobreviven más que los estudiantes que provienen de otras situaciones familiares.

Figura 16

Función de Riesgo Situación Familiar



En la figura 16, se observan las curvas de la función de riesgo para situación familiar (buena, regular, mala y separados), de los estudiantes de la Escuela Profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco. A pesar que no existe diferencia significativa entre las curvas de la función de riesgo, se observa que aparentemente los estudiantes que provienen de situación familiar separados son los que tienen mayor riesgo a desertar o abandonar sus estudios Universitarios.

Tabla 29

Resumen del Lugar de Procedencia

Lugar de Procedencia	N°. Total	N°. Eventos	N°. de Censurados	Eventos %	Censurado %
Cusco	61	37	24	60.7%	39.3%
Otros	70	44	26	62.9%	37.1%
Total	131	81	50	61.8%	38.2%

En la tabla 29, se observa el resumen de lugar de procedencia (Cusco y otros), del 100% de los estudiantes de la Escuela Profesional de Matemática de la cohorte 2010-I hasta la cohorte 2012-II, durante el periodo de seguimiento al 62.9% del total de estudiantes que proceden de otros lugares (provincias del departamento Cusco y departamentos del Perú) les ocurrió el evento de interés (deserción) y el 37.1% fueron censurados; y al 60.7% del total de estudiantes que proceden de la ciudad de Cusco les ocurrió el evento de interés (deserción) y 39.3% fueron censurados.

Según Puma (2018) indico que durante el período de análisis el 30.8% del total de estudiantes que provienen de provincia desertaron, mientras que el 62.0 % del total de los estudiantes de ciudad desertaron. Los resultados obtenidos son contradictorios porque al 62.9% provenientes de otros lugares son los presentan mayor riesgo de deserción con respecto a los que provienen de la ciudad de Cusco (60.7%).

Tabla 30

Los tiempos de supervivencia, la proporción de estudiantes sobrevivientes en cada punto del tiempo en semestres para la variable lugar de procedencia (Cusco y otros).

Lugar de procedencia= Cusco						
Tiempo	Nº.Riesgo	Nº.Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	61	17	0.721	0.0574	0.617	0.843
2	44	7	0.607	0.0625	0.496	0.742
3	37	3	0.557	0.0636	0.446	0.697
4	34	3	0.508	0.0640	0.397	0.650
5	31	2	0.475	0.0639	0.365	0.619
6	28	3	0.424	0.0635	0.317	0.569
7	25	2	0.391	0.0628	0.285	0.535
Lugar de procedencia= otros						
Tiempo	Nº.Riesgo	Nº.Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	70	20	0.714	0.0540	0.616	0.828
2	50	5	0.643	0.0573	0.540	0.766
3	45	6	0.557	0.0594	0.452	0.687

4	39	7	0.457	0.0595	0.354	0.590
5	32	4	0.400	0.0586	0.300	0.533
6	27	1	0.385	0.0582	0.286	0.518
10	22	1	0.368	0.0582	0.270	0.501

Tabla 31

Comparaciones Globales

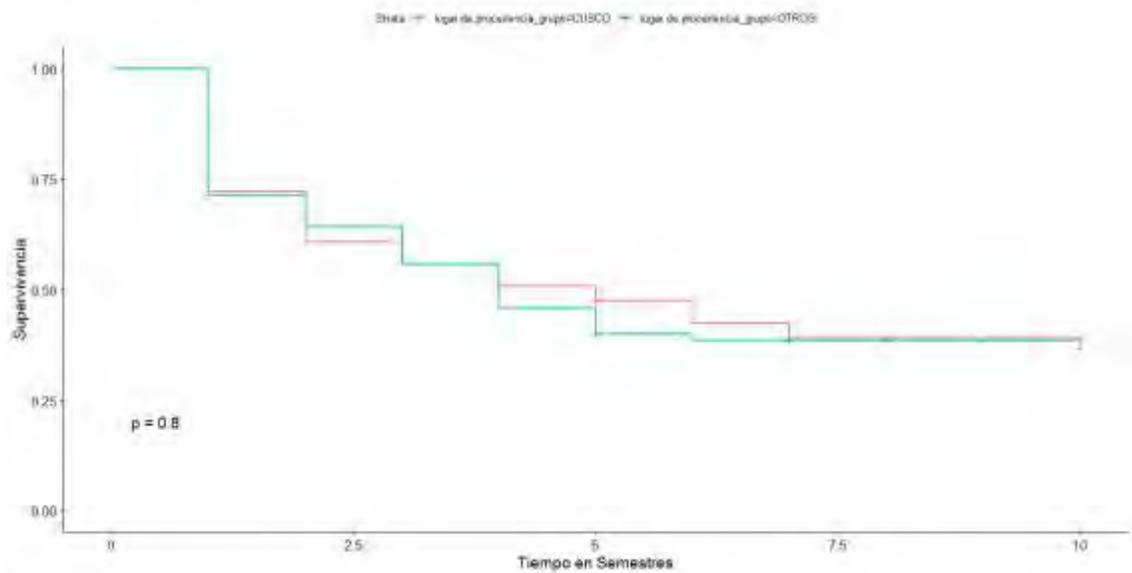
	Chi- cuadrado	Gl	p-value
Log-Rank	0.1	1	0.8

En la tabla 31, se muestra el resultado de la prueba estadística para la igualdad de las funciones de supervivencia, para comparar la función de supervivencia de lugar de procedencia (Cusco y otros).

Usamos la prueba de Log-Rank con un valor estadístico de 0.1 y un nivel de significancia $\alpha = 0.05$, como $p - value = 0.8$ es mayor que $\alpha = 0.05$ no se rechaza (H_0). Se concluye que no existe diferencia significativa para lugar de procedencia (Cusco y otros).

Figura 17

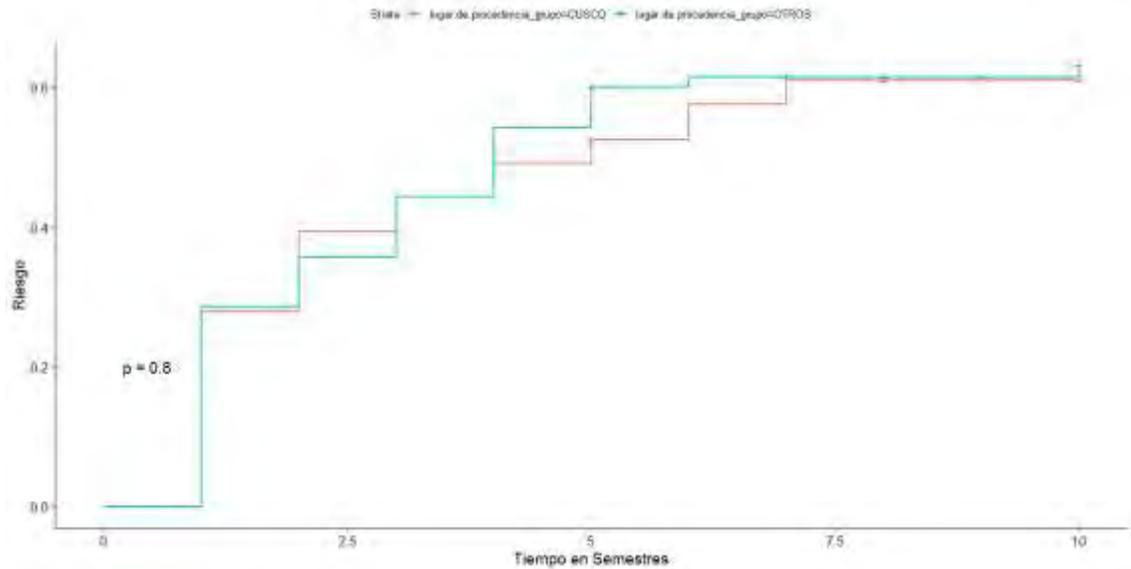
Función de Supervivencia Lugar de Procedencia



En la figura 17, se observa la curva de la función de supervivencia de lugar de procedencia (CUSCO y otros), de los estudiantes de la Escuela Profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco. A pesar que no existe diferencia significativa entre las curvas de supervivencia, se puede observar que hay una separación entre las curvas de supervivencia y aparentemente los estudiantes que proceden de la ciudad de Cusco son los que sobreviven más que los estudiantes que proceden de otros lugares.

Figura 18

Función de Riesgo Lugar de Procedencia



En la figura 18 se muestran las curvas de la función de riesgo de lugar de procedencia (Cusco y Otros), de los estudiantes de la Escuela Profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco. A pesar que no existe diferencia significativa entre las curvas de la función de riesgo, se observa que aparentemente los estudiantes que proceden de otros lugares son los que tienen mayor riesgo a desertar o abandonar sus estudios Universitarios.

Tabla 32

Resumen de Estudiante que Trabaja

EL Estudiante Trabaja	N°. Total	N°. Eventos	N°. de Censurados	Eventos %	Censurado %
No	107	71	36	66.4%	33.6%
Si	24	10	14	41.7%	58.3%
Total	131	81	50	61.8%	38.2%

En la tabla 32 se observa el resumen del estudiante que trabaja (no y sí), del 100% de los estudiantes de la Escuela Profesional de Matemática de la cohorte 2010-I hasta la cohorte 2012-II, durante el periodo de seguimiento al 66.4% del total de estudiantes que no trabajan les ocurrió el evento de interés (deserción) y el 33.6% fueron censurados; y al 41.7% del total de estudiantes que trabajan les ocurrió el evento de interés (deserción) y 58.3% fueron censurados.

Tabla 33

Los tiempos de supervivencia, la proporción de estudiantes sobrevivientes en cada punto del tiempo en semestres para la variable el estudiante trabaja (si, no)

El estudiante trabaja= no						
Tiempo	N°.Riesgo	N°.Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	107	32	0.701	0.0443	0.619	0.793
2	75	9	0.617	0.0470	0.531	0.716
3	66	8	0.542	0.0482	0.455	0.645
4	58	10	0.449	0.0481	0.364	0.553
5	48	6	0.393	0.0472	0.310	0.497
6	40	4	0.353	0.0464	0.273	0.457
7	36	1	0.343	0.0461	0.264	0.447
10	29	1	0.332	0.0460	0.253	0.435
El estudiante trabaja= si						
Tiempo	N°.Riesgo	N°.Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	24	5	0.792	0.0829	0.645	0.972
2	19	3	0.667	0.0962	0.502	0.885
3	16	1	0.625	0.0988	0.458	0.852
7	15	1	0.583	0.1006	0.416	0.818

Tabla 34

Comparaciones Globales

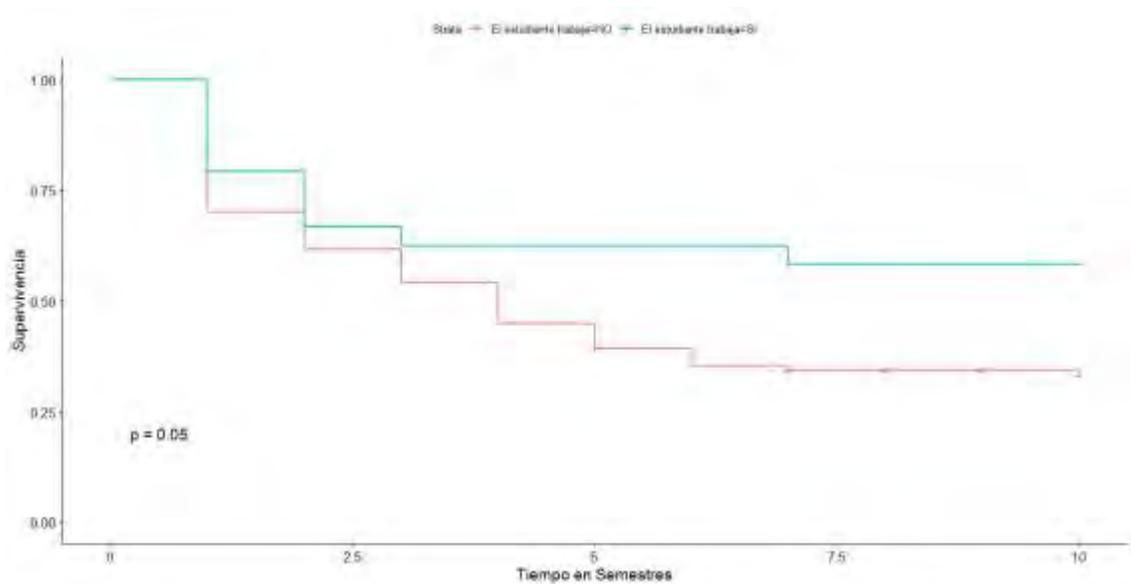
	Chi-cuadrado	Gl	p-value.
Log-Rank	3.9	1	0.05

En la tabla 34, se muestra el resultado de la prueba estadística para la igualdad de las funciones de supervivencia, para comparar la función de supervivencia del estudiante trabaja (no y sí).

Usamos la prueba de Log-Rank con un valor estadístico de 3.9 y un nivel de significancia $\alpha = 0.05$, como $p - value = 0.05$ es igual a $\alpha = 0.05$ se rechaza (H_0). Se concluye que existe diferencia significativa de la variable el estudiante trabaja (si, no).

Figura 19

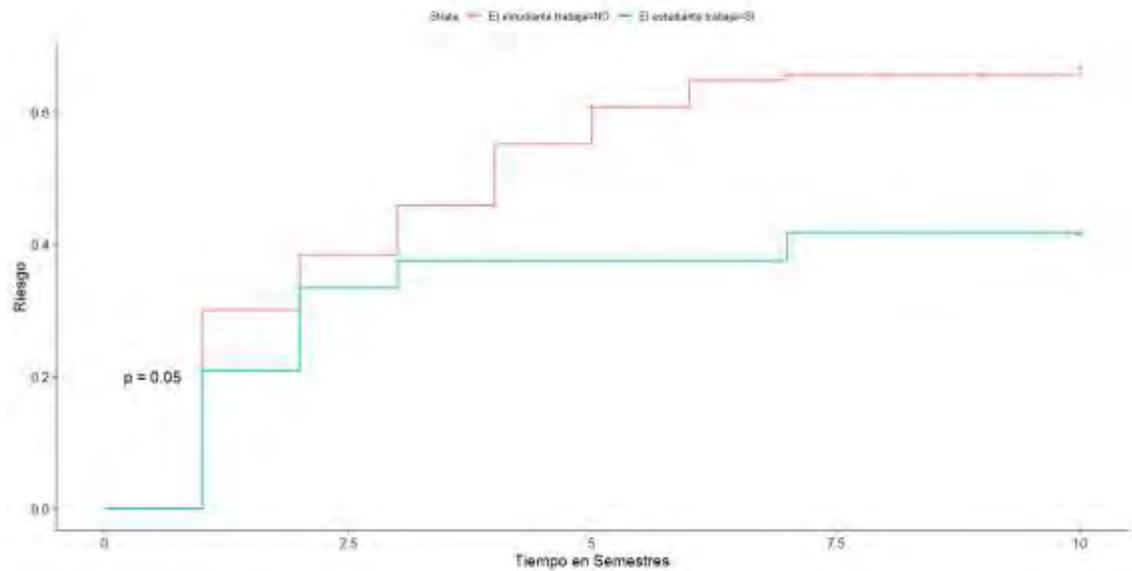
Función de Supervivencia el Estudiante Trabaja (no y sí)



En la figura 19, se observa la curva de la función de supervivencia para los estudiantes que si trabajan y no trabajan de la Escuela Profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco. Como existe diferencia significativa entre las curvas de supervivencia se concluye que los estudiantes que trabajan durante su formación profesional son los que sobreviven más que los estudiantes que no trabajan.

Figura 20

Función de Riesgo el Estudiante Trabaja (no y sí)



En la figura 20, se observa las curvas de la función de riesgo para los estudiantes que si trabajan y no trabajan de la Escuela Profesional de matemática de la Universidad Nacional San Antonio Abad del Cusco. Existe diferencia significativa entre las curvas de la función de riesgo y los estudiantes que no trabajan durante su formación profesional tienen mayor riesgo a desertar o abandonar sus estudios superiores que los estudiantes que trabajan.

Tabla 35

Resumen de Edad

Edad	N°. Total	N°. Eventos	N°. de Censurados	Eventos%	Censurado%
mayor o igual que 22 años	13	4	9	30.8%	69.2%
menor que 22 años	118	77	41	65.3%	34.7%
Total	131	81	50	61.8%	38.2%

En la tabla 35, se observa el resumen de la edad de ingreso a la Universidad (menor que 22 años y mayor o igual que 22 años), del 100% de los estudiantes de la Escuela Profesional de Matemática de la cohorte 2010-I hasta la cohorte 2012-II, durante el periodo de seguimiento al 65.3 % del total de estudiantes que ingresaron a la Universidad con edad menor a 22 años les ocurrió el evento de interés (deserción) y el 34.7% fueron censurados; y al 30.8% del total de estudiantes que ingresaron a la Universidad con edad mayor o igual a 22 años les ocurrió el evento de interés (deserción) y 69.2% fueron censurados.

Según Puma (2018), menciona que el 87.5% de los estudiantes de edad de 25 años a más son los que presentan mayor riesgo de deserción que los estudiantes de edad de 24 años a menos (57.1%). No presenta similitud con los resultados obtenido porque el 65.3% con edad menor que 22 años presentan mayor riesgo de deserción con respecto a estudiantes con mayor o igual que 22 años (30.8%).

Asimismo, Barturén y Niño (2018) menciona que el 37% de estudiantes con edad menor de 20 años desertan mas que los estudiantes con edad mayor o igual que 20 años. Resultado similar al que hemos obtenido al 65.3% con edad menor que 22 años presentan mayor riesgo de deserción con respecto a estudiantes con mayor o igual que 22 años (30.8%).

Tabla 36

Los tiempos de supervivencia, la proporción de estudiantes sobrevivientes en cada punto del tiempo en semestres para la variable edad (menor que 22 años y mayor o igual que 22 años)

Edad= mayor o igual de 22						
Tiempo	N. Riesgo	N. Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	13	2	0.846	0.100	0.671	1.000
2	11	1	0.769	0.117	0.571	1.000

3	10	1	0.692	0.128	0.482	0.995
Edad= menor de 22						
Tiempo	N. Riesgo	N. Evento	Supervivencia	Std.err	L.I. 95%	L.S. 95%
1	118	35	0.703	0.0420	0.626	0.791
2	83	12	0.602	0.0451	0.520	0.697
3	71	8	0.534	0.0459	0.451	0.632
4	63	10	0.449	0.0458	0.368	0.548
5	53	6	0.398	0.0451	0.319	0.497
6	45	4	0.363	0.0444	0.286	0.461
7	41	1	0.354	0.0442	0.277	0.452
10	35	1	0.344	0.0441	0.268	0.442

Tabla 37

Comparaciones Globales

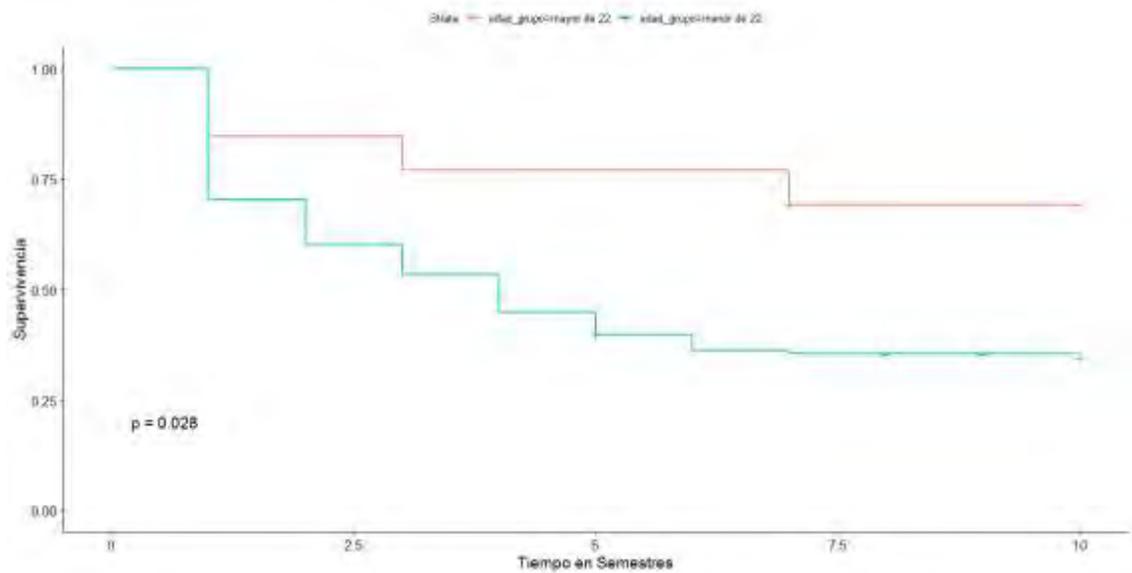
	Chi-cuadrado	Gl	p-value.
Log-Rank	4.8	1	0.03

En la tabla 37, se observa los resultados de la prueba estadística para la igualdad de las funciones de supervivencia, para comparar la función de supervivencia de edad (menor que 22 años y mayor o igual que 22 años).

Usamos la prueba de Log-Rank con un valor estadístico de 4.8 y un nivel de significancia $\alpha = 0.05$, como $p - value = 0.03$ es menor que $\alpha = 0.05$ se rechaza (H_0). Se concluye que existe diferencia significativa en la edad de ingreso a la Universidad de los estudiantes con menor de 22 años y mayor o igual de 22 años.

Figura 21

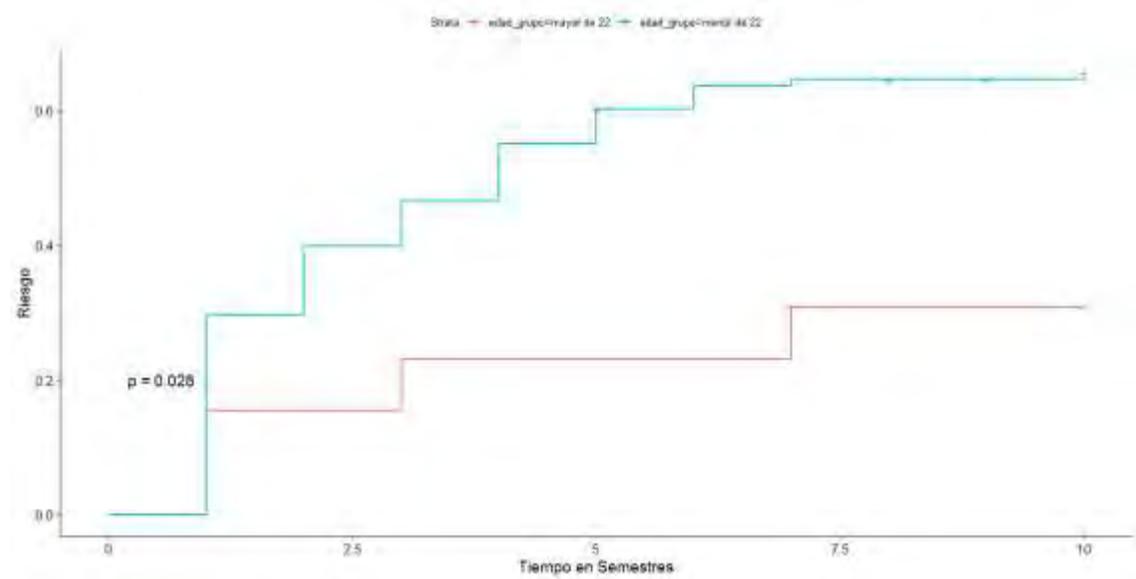
Función de Supervivencia Edad



En la figura 21, se observa la curva de la función de supervivencia para edad de ingreso a la Universidad (menor que 22 años y mayor o igual que 22 años), de los estudiantes de la Escuela Profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco. Como existe diferencia significativa entre las curvas de supervivencia se concluye que los estudiantes que ingresan a la Universidad con edad mayor o igual a 22 años son los que sobreviven más que los que ingresan con edad menor a 22 años.

Figura 22

Función de Riesgo Edad



En la figura 22, se observa la curva de la función de riesgo para los estudiantes que ingresan con edad (menor que 22 años y mayor o igual que 22 años) a la Escuela Profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco. Existe diferencia significativa entre las curvas de la función de riesgo y los estudiantes que ingresan con edad menor que 22 años tienen mayor riesgo a desertar o abandonar sus estudios superiores que los estudiantes que ingresan con edad mayor o igual que 22 años.

Tabla 38*Resumen de los resultados de Kaplan-Meier*

Nº. Total	VARIABLES	Categorías	Supervivencia	Deserción
131	Escuela Profesional	Matemática	38.2%	61.8
39	Sexo	Mujer	43.6%	56.4%
92	sexo	Varón	35.9%	64.1%
110	Tipo de Colegio	Estatad	40.9%	59.1%
21	Tipo de Colegio	Particular	23.8%	76.2%
14	Tipo de Modalidad de Ingreso	CEPRU	50.0%	50.0%
38	Tipo de Modalidad de Ingreso	Dirimencia	26.3%	73.7%
52	Tipo de Modalidad de Ingreso	Ordinario	40.4%	59.6%
27	Tipo de Modalidad de Ingreso	Primera oportunidad	44.4%	61.8%
111	Ocupación del jefe de hogar	No Profesional	41.4%	58.6%
20	Ocupación del jefe de hogar	Profesional	20.0%	80.0%
9	Ingreso económico familiar	Sueldo alto	55.6%	44.4%
42	Ingreso económico familiar	Sueldo medio	31.0%	69.0%
80	Ingreso económico familiar	Sueldo básico	40.0%	60.0%
38	Número de Hermanos	0 a 1	39.5%	60.5%
62	Número de Hermanos	2 a 3	38.7%	61.3%
31	Número de Hermanos	Más de 3	35.5%	64.5%
41	Situación Familiar	Buena	34.7%	65.9%
71	Situación Familiar	Regular	43.7%	56.3%
3	Situación Familiar	Mala	33.3%	66.7%
16	Situación Familiar	Separados	25.0%	75.0%
61	Lugar de Procedencia	Cusco	39.3%	60.7%
70	Lugar de Procedencia	Otros	37.1%	62.9%
107	El Estudiante Trabaja	No	33.6%	66.4%
24	El Estudiante Trabaja	Si	58.3%	41.7%
13	Edad	Mayor o Igual que 22	69.2%	30.8%
118	Edad	Menor que 22	34.7%	65.3%

1. Existe un 61.8% de deserción estudiantil del total de estudiantes de la Escuela Profesional de Matemática
2. Se observa que en las estudiantes mujeres, su nivel de supervivencia es de 43.6% y es mayor que a los estudiantes varones que tienen un nivel de supervivencia del 35.9%. Existe un 64.1% de deserción para los estudiantes varones y un 56.4% de deserción para las estudiantes mujeres, los que tienen mayor riesgo a desertar o abandonar sus estudios superiores son los estudiantes varones.

3. Se observa que los estudiantes que provienen de colegio estatal tienen un nivel de supervivencia 40.9% y es mayor con respecto a los estudiantes que provienen de colegio particular con 23.8%, los estudiantes que presentan mayor riesgo de desertar o abandonar sus estudios superiores son los estudiantes que provienen de colegio particular.
4. Existe un 73.7% de deserción para la modalidad de ingreso dirimencia, 61.8% de deserción para la modalidad de ingreso primera oportunidad, 59.6% de deserción para la modalidad de ingreso ordinario y 50.0% de deserción para la modalidad de ingreso CEPRU, los estudiantes que tienen mayor riesgo a desertar son los que ingresan por dirimencia y los estudiantes que sobreviven más son los que ingresan por CEPRU con un 50.0% de supervivencia.
5. Los estudiantes que provienen de jefe de hogar no profesional son los que sobreviven más con un 41.4% de supervivencia que los estudiantes que provienen de jefe de hogar profesional con un 20.0% de supervivencia y los estudiantes que provienen de jefe de hogar profesional tienen mayor riesgo de desertar o abandonar sus estudios superiores con un 80.0% de deserción que los que provienen de jefe de hogar no profesional con 58.6% de deserción.
6. Los estudiantes de ingreso económico familiar con sueldo medio con un 69% de deserción, seguida de los estudiantes de ingreso económico familiar con sueldo básico y sueldo alto 60% y 44.4% respectivamente. Los estudiantes de ingreso económico familiar con sueldo alto son los que sobreviven más con un 55.6% de supervivencia seguida de los estudiantes de ingreso económico familiar con sueldo básico y sueldo medio 40% y 31% respectivamente

7. Los estudiantes que tienen más de tres hermanos son los que desertan más con un 64.5% de deserción, seguida por los estudiantes que tienen 2 a 3 hermanos con un 62.3% de deserción y con un 60.5% de deserción para los alumnos con 0 a 1 hermano. Los estudiantes que tienen 0 a 1 hermano son los que sobreviven más con un 39.5% de supervivencia seguida por los alumnos que tienen 2 a 3 hermanos con un 38.7% de supervivencia y un 35.5% de supervivencia para los estudiantes que tienen más de 3 hermanos.
8. Los estudiantes con situación familiar separados tienen mayor riesgo a desertar con un 75.0% de deserción, seguido por alumnos con situación familiar mala, buena y regular con una deserción de 66.7%, 65.9% y 56.3% respectivamente. Los estudiantes que provienen de situación familiar regular son los que sobreviven más con un 43.7% de supervivencia seguida por la situación familiar buena, mala y separados con supervivencia de 34.7%, 33.3% y un 25% respectivamente.
9. Los estudiantes de lugar de procedencia otros son los que tienen mayor riesgo a desertar con un 62.9% de deserción, seguida por los alumnos de lugar de procedencia Cusco con un 60.7% y los alumnos de lugar de procedencia Cusco tienen un 39.3% de supervivencia y los alumnos de lugar de procedencia otros tienen un 37.1% de supervivencia, entonces los que sobreviven más son los estudiantes que provienen de la ciudad de Cusco mismo.
10. Los estudiantes que no trabajan tienen un 66.4% de deserción y los estudiantes que si trabajan tienen un 41.7% de deserción, los estudiantes que no trabajan tienen mayor riesgo de desertar. Los estudiantes que trabajan tienen un 58.3% de supervivencia y los estudiantes que no trabajan tienen un 33.6% de supervivencia, los que sobreviven más son los estudiantes que si trabajan.

11. Los estudiantes que tienen edad menor de 22 años son los que presentan mayor riesgo de desertar que los que tienen edad mayor o igual de 22 años con 65.3% y 30.8% de deserción respectivamente. Los estudiantes que tienen un 69.2% de supervivencia son los estudiantes de edad mayor o igual de 22 años y los estudiantes de edad menor de 22 años tiene 34.7% de supervivencia, los estudiantes que sobreviven más son los estudiantes de edad mayor o igual de 22 años.

4.2. Resultado del Modelo de Regresión de Cox

Esta sección tiene la finalidad de estimar, el efecto de las variables de estudio sobre la supervivencia de los estudiantes, a partir de una serie de variables consideradas en la base de datos de los alumnos de la Escuela Profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco. El modelo de regresión de Cox tiene como finalidad estimar la relación que hay entre un conjunto de variables explicativas, fijas de estudio en la probabilidad de supervivencia la deserción estudiantil.

Las variables que se analizan en esta investigación son los que están en la base de datos como: sexo, tipo de colegio, tipo de modalidad de ingreso, situación familiar, el estudiante trabaja, número de hermanos, edad, ingreso económico familiar, ocupación del jefe de hogar, lugar de procedencia. Los cálculos fueron desarrollados en el software RStudio.

Tabla 39*Coefficiente de la regresión de Cox para las variables en la ecuación*

Variables	β	$Exp(\beta)$	$se(\beta)$	z	p	Intervalo de Confianza	
						Inferior	Superior
sexo varón	0.783	2.189	0.289	2.715	0.007	1.244	3.854
Tipo de colegio particular	0.401	1.493	0.324	1.237	0.216	0.791	2.819
Tipo de modalidad de ingreso dirimencia	0.713	2.039	0.453	1.572	0.116	0.839	4.959
Tipo de modalidad de ingreso ordinario	0.687	1.987	0.448	1.533	0.125	0.826	4.782
Tipo de modalidad de ingreso primera oportunidad	0.197	1.128	0.479	0.413	0.680	0.477	3.114
Situación familiar mala	-0.244	0.783	0.770	-0.316	0.752	0.173	3.546
Situación familiar regular	-0.040	0.961	0.263	-0.153	0.879	0.574	1.607
Situación familiar separados	0.386	1.471	0.375	1.030	0.303	0.706	3.065
El estudiante trabaja si	-0.649	0.522	0.364	-1.781	0.075	0.256	1.067
Edad grupo menor que 22	1.510	4.523	0.605	2.495	0.013	1.383	14.814
Número de hermanos grupo 2-3	-0.436	0.646	0.301	-1.452	0.147	0.359	1.165
Número de hermanos grupo 3+	0.007	1.007	0.368	-0.026	0.985	0.489	2.072
Ingreso económico familiar grupo sueldo básico	1.787	5.970	0.624	2.865	0.004	1.759	20.268
Ingreso económico familiar grupo sueldo medio	1.497	4.467	0.585	2.560	0.010	1.420	14.052
Lugar de procedencia grupo otros	0.129	1.138	0.271	0.476	0.634	0.668	1.937
Ocupación del jefe de hogar grupo profesional	0.988	2.686	0.360	2.741	0.006	1.325	5.445

La tabla 39 nos muestra las estimaciones de los parámetros correspondientes a las variables: sexo varón, tipo de colegio particular, tipo de modalidad de ingreso dirimencia, tipo de modalidad de ingreso ordinario, tipo de ingreso primera oportunidad, situación familiar mala, situación familiar regular, situación familiar separados, el estudiante trabaja si, edad grupo menor de 22 años, número de hermanos grupo (2-3), número de hermanos grupo 3+, ingreso económico familiar grupo sueldo básico, ingreso económico familiar grupo medio, lugar de procedencia grupo otros, ocupación del jefe de hogar grupo profesional, de los estudiantes de la Escuela Profesional de Matemática de la Universidad Nacional San Antonio Abad del Cusco. Se observa que el coeficiente de las variables situación familiar mala, situación familiar regular, el estudiante trabaja si, número de hermanos grupo 2-3, es negativo, esto quiere decir que esas variables tienen un menor riesgo de desertar, mientras el coeficiente de las variables: sexo varón, tipo de colegio particular, tipo de modalidad de ingreso dirimencia, tipo de modalidad de ingreso ordinario, tipo de ingreso primera oportunidad, situación familiar separados, edad grupo menor de 22, número de hermanos 3+, ingreso económico del padre grupo sueldo básico, ingreso económico familiar grupo sueldo medio, lugar de procedencia grupo otros, ocupación del jefe de hogar grupo profesional, tienen coeficiente positivo, eso quiere decir que estas variables indican mayor riesgo de desertar.

En las variables: sexo varón, edad grupo menor que 22 años, ingreso económico familiar grupo sueldo básico, ingreso económico familiar grupo sueldo medio y ocupación del jefe de hogar grupo profesional, tiene una influencia significativa porque el p – valúe es menor que $\alpha = 0.05$ en la no permanencia de los estudiantes que ingresaron a la Escuela Profesional de Matemática en la cohorte 2010-I hasta la cohorte 2012-II.

Tabla 40

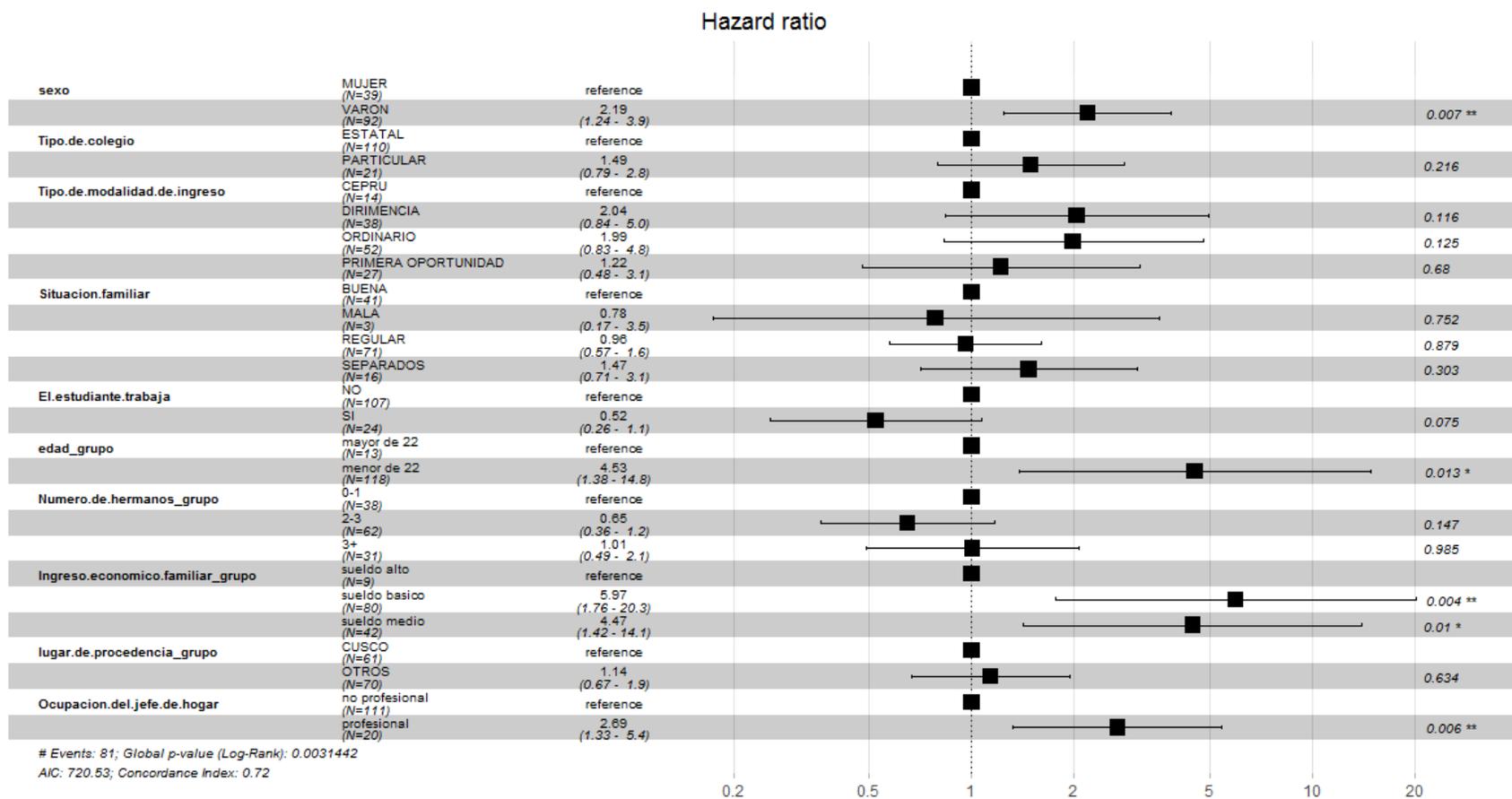
Comparación Global

	Chi-cuadrado	Gl	p-value.
Test de Wald	32.7	16	0.008

En la tabla 40, se observa el Test de Wald con un valor estadístico de 32.7 y un nivel de significancia de 0.05, como $p\text{-value} = 0.008$ menor que el nivel de significancia se rechaza la hipótesis nula H_0 y se concluye que en conjunto las variables independientes son estadísticamente significativas en la explicación de la variación de la variable dependiente (deserción).

Figura 23

Variables de la Regresión de Cox



Cada Hazard Ratio (HR) representa un riesgo de deserción que compara una instancia de una característica binaria con la instancia. Un $HR > 1$ indica mayor riesgo de desertar y si un $HR < 1$ indica una disminución del riesgo de desertar.

En la figura 23 se observa que:

El $HR > 1$ para la variable Sexo varón esto indica que los varones tienen mayor riesgo de desertar, su riesgo de desertar aumenta en un 2.19 veces que al de las mujeres.

Para la variable tipo de colegio particular el $HR > 1$ esto indica que los que provienen de colegios particulares tienen mayor riesgo a desertar, y su riesgo de desertar aumenta en un 1.49 veces al de los que provienen de colegios estatales.

Para la variable tipo de modalidad de ingreso dirimencia el $HR > 1$ esto indica que su riesgo de desertar aumenta en un 2.04 veces al que ingresa por CEPRU, el $HR > 1$ para los que ingresan por ordinario esto indica que su riesgo de desertar aumenta en un 1.99 veces al que ingresa por CEPRU y el $HR > 1$ para los que ingresan por primera oportunidad esto indica que su riesgo de desertar aumenta en un 1.22 veces al que ingresa por CEPRU.

Para la variable situación familiar mala el $HR < 1$ esto nos indica que la deserción disminuye en un 0.78 veces al de la situación familiar buena, el $HR < 1$ para la situación familiar regular esto indica que la deserción disminuye en un 0.96 veces al de la situación familiar buena, el $HR > 1$ para la situación familiar separados esto indica que el riesgo de desertar incrementa en un 1.47 veces al de la situación familiar buena.

Para el estudiante que si trabaja el $HR < 1$ esto indica que el riesgo de desertar disminuye en un 0.52 veces al estudiante que no trabaja.

Para la variable edad menor que 22 años el $HR > 1$ esto indica que el riesgo de desertar aumenta en un 4.53 veces al que ingresa con edad mayor o igual que 22 años.

Para la variable número de hermanos 2 a 3 el HR <1 esto indica que el riesgo de desertar disminuye en un 0.65 veces al del número de hermanos 0 a 1, el HR > 1 para número de hermanos 3 a más esto indica que el riesgo de desertar incrementa en un 1.01 veces al del número de hermanos 0 a 1.

Para la variable ingreso económico familiar sueldo básico el HR > 1 esto indica que el riesgo de desertar aumenta en un 5.97 veces al ingreso económico familiar sueldo alto; el HR > 1 para el ingreso económico familiar sueldo medio esto indica que el riesgo de desertar incrementa en un 4.47 veces al ingreso económico sueldo alto.

Para la variable lugar de procedencia otros el HR > 1 esto indica que el riesgo de desertar incrementa en un 1.14 veces al lugar de procedencia Cusco.

Para la variable ocupación del jefe de hogar profesional el HR > 1 esto indica que el riesgo de desertar incrementa en un 2.69 veces al de ocupación del jefe de hogar no profesional.

Tabla 41

Resumen de los resultados de la regresión de Cox

Variables	β	$Exp(\beta)$	p
Sexo varón	0.783	2.189	0.007
Tipo de colegio particular	0.401	1.493	0.216
Tipo de modalidad de ingreso dirimencia	0.713	2.039	0.116
Tipo de modalidad de ingreso ordinario	0.687	1.987	0.125
Tipo de modalidad de ingreso primera oportunidad	0.197	1.128	0.680
Situación familiar mala	-0.244	0.783	0.752
Situación familiar regular	-0.040	0.961	0.879
Situación familiar separados	0.386	1.471	0.303
El estudiante trabaja si	-0.649	0.522	0.075
Edad grupo menor de 22	1.510	4.523	0.013
Número de hermanos grupo 2-3	-0.436	0.646	0.147
Número de hermanos grupo 3+	0.007	1.007	0.985
Ingreso económico familiar grupo sueldo básico	1.787	5.970	0.004
Ingreso económico familiar grupo sueldo medio	1.497	4.467	0.010
Lugar de procedencia grupo otros	0.129	1.138	0.634

Ocupación del jefe de hogar grupo profesional	0.988	2.686	0.006
---	-------	-------	-------

Los factores asociados con la deserción estudiantil de los estudiantes de la Escuela Profesional de Matemática son las variables: sexo, edad, ingreso económico familiar y ocupación del jefe de hogar; es decir para los estudiantes varones el riesgo aumenta en un 2.189 veces de las estudiantes mujeres; los estudiantes que ingresaron a la universidad con edad menor que 22 años el riesgo aumenta en un 4.523 veces de desertar que los estudiantes que ingresaron con edad mayor o igual que 22 años, los estudiantes que provienen de un ingreso económico familiar sueldo básico el riesgo de desertar aumenta en un 5.970 veces y los alumnos que provienen de un ingreso económico familiar sueldo medio su riesgo de desertar aumenta en un 4.467 veces de que los estudiantes que provienen de un ingreso económico familiar alto; y los estudiantes que provienen de jefe de hogar profesional su riesgo de desertar aumenta en un 2.656 veces de los estudiantes que provienen de jefe de hogar no profesional.

CONCLUSIONES

1. En la Escuela Profesional de Matemática de la Facultad de Ciencias de la Universidad Nacional San Antonio Abad del Cusco, se ha observado un nivel de supervivencia del 38.2%, mientras el 61.8 % desertaron del total de los estudiantes que ingresaron en los semestres 2010-I, 2010-II, 2011-I, 2011-II, 2012-I y 2012-II, en un periodo de estudio de 10 semestres.
2. Los estudiantes de la Escuela Profesional de Matemática presentan tipo de censura a la derecha, porque se conoce el inicio del estudio (semestre de ingresó y matriculado) y no se conoce la culminación del semestre (porque no todos los estudiantes concluyen sus estudios en 10 semestres, concluyen en más de 10 semestres académicos).
3. Los factores asociados con la deserción estudiantil son las variables sexo, edad, ingreso económico familiar y ocupación del jefe de hogar ($p\text{-valué}<0,05$).
4. Los que presentan mayor riesgo de desertar o abandonar sus estudios superiores. Son los estudiantes varones, que ingresan a la Universidad con edad menor que 22 años, los que provienen de jefe de hogar profesional e ingreso económico familiar que tienen sueldo básico y medio.

RECOMENDACIONES

- Para investigaciones posteriores se recomienda considerar más variables que puedan pronosticar la probabilidad de desertar, tales como: factores académicos (orientación vocacional, número de créditos al semestre (carga académica) y calificación en el examen de admisión) y factores institucionales; ya que las bases teóricas recomiendan.
- Se recomienda al director de la Escuela Profesional de Matemática de la Universidad Nacional de San Antonio Abad de Cusco que es necesario brindar apoyo a los estudiantes que ingresan a la Escuela Profesional de Matemática con edad menor de 22 años y provienen de jefe de hogar profesional e ingreso económico familiar que tienen sueldo básico y medio, para evitar la deserción. Se propone que creen un programa anti- deserción en donde se dé a conocer los beneficios que ofrece la Universidad (vivienda Universitaria, intercambio estudiantil, Becas, etc.)
- Se recomienda a la Oficina de Unidad de Bienestar Universitario (Área de asistencia social) ordenar y organizar los datos de la ficha socioeconómica de manera digitalizada, así obtener información completa para investigaciones posteriores.

REFERENCIA

- Avendaño Garrido , M. L. (2013). *Modelos Generalizados de Riesgos Proporcionales para el Analisis de Supervivencia*. Tesis doctoral, Universidad Complutense de Madrid, Madrid. Obtenido de <http://eprints.ucm.es/22026/1/T34592.pdf>
- Barrera Rebellon, 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. Obtenido de <http://repositorio.utp.edu.co/dspace/bitstream/handle/11059/2043/3712913B272.pdf;sequence=1>
- Barturén Rivera , D. K., & Niño Capcha, 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*. Universidad Nacional Pedro Ruiz Gallo, Lambayeque, Lambayeque. Obtenido de <http://repositorio.unprg.edu.pe/bitstream/handle/UNPRG/3004/BC-TES-TMP-1823.pdf?sequence=1&isAllowed=y>
- Boj del Val, E. (Enero de 2017). *El Modelo de regresion de Cox*. Universidad de Barcelona. Obtenido de <http://diposit.ub.edu/dspace/bitstream/2445/49070/6/El%20modelo%20de%20Cox%20de%20riesgos%20proporcionales.pdf>
- Cruz Ordíñez, D. I., & Ortega Calpa, J. F. (2008). Analisis de la desercion 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 el sistema spadies. *Tesis de Licenciatura*. Universidad de Nariño, San Juan de Pasco.

- Díaz Peralta , C. (2008). Modelo conceptual para la Deserción Estudiantil Universitaria Chilena. *Estudios Pedagógicos XXXIV(2)*, 65-86. Obtenido de <http://revistas.uach.cl/pdf/estped/v34n2/art04.pdf>
- Dirección de sistemas de información, u. d. (2016). *compendio estadístico N°31*. Universidad Nacional de San Antonio Abad del Cusco, Cusco, Cusco. Obtenido de <http://conia.unsaac.edu.pe/images/oficinas/estadistica/CompendioEstadistico31.pdf>
- Dirección de sistemas de información, U. d. (2017). *Compendio estadístico N°32*. Universidad Nacional de San Antonio del Cusco, Cusco, Cusco. Obtenido de <http://ctic2009.unsaac.edu.pe/images/oficinas/estadistica/CompendioEstadistico32.pdf>
- Educación al Futuro.com. (2017). *Alta deserción Universitaria por falta de buena orientación vocacional*. [Publicación en línea] Junio 2017, Lima. Recuperado el 30 de Junio de 2020, de <https://educacionalfuturo.com/noticias/alta-desercion-universitaria-por-falta-de-buena-orientacion-vocacional/>
- Gabriel, Jaime, Páramo, Carlos, Arturo, & Correa. (Abril-Mayo de 1999). Deserción Estudiantil Universitaria Conceptualización. *Revista Universidad Eafit.*, 66-78. Obtenido de <file:///C:/Users/USUARIO/Downloads/1075-1-3396-1-10-20120703.pdf>
- García Hinojosa , C. P. (2014). *Estudio de Análisis de Supervivencia*. Doble Grado en Ingeniería Informática y Matemáticas, Universidad Autónoma de Madrid, Madrid. Obtenido de https://repositorio.uam.es/bitstream/handle/10486/661556/prueba_garcia_hinojosa_cristina_tfg.pdf?sequence=1
- Huaman Ojeda, Y., & Martínez Loayza, M. (2018). *Factores relacionados con la deserción de estudiantes de enfermería, Universidad Nacional de San Antonio Abad del Cusco*,

- semestre 2010-I al 2017-I*. Universidad Nacional de San Antonio Abad del Cusco, Cusco.
Obtenido de <http://repositorio.unsaac.edu.pe/handle/UNSAAC/3648>
- López Calviño, B. (2011-2012). *Estimador Presuavizado de Kaplan-Meier con indicadores de censura perdido aleatoriamente*. Master en técnicas Estadísticas, Universidad de Coruña, Coruña. Obtenido de
http://eio.usc.es/pub/mte/descargas/ProyectosFinMaster/Proyecto_419.pdf
- Ministerio de Educacion Nacional de Colombia. (2009). *Deserción estudiantil en la educacion superior colombiana*. Obtenido de
https://www.mineducacion.gov.co/sistemasdeinformacion/1735/articulos-254702_libro_desercion.pdf
- Mondragón Arbocó, J. A. (Septiembre de 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. Obtenido de
file:///H:/Libros%20para%20imprimir/MONDRAGON_JORGE_INTERVALOS_DE_CONFIANZA.pdf
- Puma Laura, C. V. (2018). *Análisis de supervivencia de la derescion en los estudiantes de la facultad de Ciencias Naturales y Formales a través del Modelo de Cox*. (Tesis de Licenciatura), Universidad nacional de San Agustín, Facultad de Ciencias Naturales y Formales, Arequipa. Obtenido de
<http://repositorio.unsa.edu.pe/bitstream/handle/UNSA/6483/MApulacv.pdf?sequence=1&isAllowed=y>
- Rodríguez Pineda, M., & Zamora Araya, J. A. (2014). *Análisis de la deserción en la Universidad Nacional desde una perspectiva longitudinal*. Quinto informe estado de la educación,

- Universidad Nacional de Costa Rica, Heredia . Obtenido de
https://estadonacion.or.cr/files/biblioteca_virtual/educacion/005/Magaly_Rodriguez_Analisis_de_la_desercion.pdf
- San Jose, B., Pérez , E., & Madero, R. (2009). *Hablemos de... Métodos estadísticos en estudios de supervivencia*, 55-59. Obtenido de www.apcontinuada.com/es/pdf/80000481/.
- Tineo Guevara , 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. Obtenido de
http://sisbib.unmsm.edu.pe/bibvirtualdata/monografias/basic/tineo_gf/intro.pdf
- Tinto, V. (s.f.). Definir la desercion: una cuestion de perspectiva. *Microsoft Word-DEFINIR LA DESERCIÓN.doc*, 43. Obtenido de
www.alfaguia.org/alfaguia/files/1342823160_52.pdf
- Vaira, S., Avila, O., Ricardi, P., & Bergesio, A. (2010). Desercion Universitaria. Un caso de estudio: variables que influyen y tiempo que demanda la toma de decisión. *FABICIB*, 14, 107-115. doi:<https://doi.org/10.14409/fabicib.v14i1.856>
- Vélez, A., & López Jiménez, D. F. (2004). Estrategias para vencer la deserción Universitaria. *Educación y Educadores*, 7, 177-203. Obtenido de
<https://educacionyeducadores.unisabana.edu.co/index.php/eye/article/view/555/648>
- Vergara Morales , J. R., Boj del Val, E., Barriga, O., & Díaz Larenas , C. (2017). Factores explicativos de la deserción de estudiantes de pedagogía. *Revista Complutense de Educacion*, 28(2), 609-630. Obtenido de
http://dx.doi.org/10.5209/rev_RCED.2017.v28.n2.50009

Vergel Ortega, M., Martínez Lozano, J. J., & Ibarquén Mondragón, E. (2016). Modelos estimados de análisis de supervivencia para el tiempo de permanencia de los estudiantes de la Universidad Francisco de Paula Santander. *Respuestas*, 21(2), 24-36. Obtenido de <https://revistas.ufps.edu.co/index.php/respuestas/article/view/775/754>

ANEXO

ANEXO 01

OFICINA DE BIENESTAR UNIVERSITARIO



CENTRO DE ASISTENCIA SOCIAL

FICHA SOCIOECONOMICA

CARRERA PROFESIONAL _____

CÓDIGO DEL ALUMNO _____

1. DATOS GENERALES

2. DATOS ACADÉMICOS

APELLIDOS Y NOMBRES : _____

LUGAR DE NACIMIENTO : _____

FECHA DE NACIMIENTO : DÍA _____ MES _____ AÑO _____

ESTADO CIVIL: SOLTERO () CASADO () D.N.I. _____

SEXO : F () M ()

DOMICILIO DEL ESTUDIANTE : _____

DOMICILIO DE LOS PADRES : _____

AÑO QUE CONCLUYO LA SECUNDARIA : _____

TIPO DE COLEGIO : ESTATAL () PARTICULAR () OTROS ()

AÑO QUE INGRESO A LA UNIVERSIDAD : _____

MODALIDAD DE INGRESO:
 PRIMERA OPORTUNIDAD () DIRIMENCIA () CEPRU () EXAMEN ORDINARIO ()
 PROFESIONAL () OTROS () EPECIFIQUE _____

3. COMPOSICION FAMILIAR

PARENTESCO	APELLIDOS Y NOMBRES	EDAD	EST. CIVIL	GRADO INST.	Ocupacion	CENTRO DE TRABAJO	INGRESO MENSUAL
PADRE							
MADRE							
HERMANOS							
CONYUGE							
HUOS							

4. SITUACION FAMILIAR

SITUACION FAMILIAR		
BUENA ()	MALA ()	REGULAR ()
SEPARADOS ()	OTROS ()	
ESPECIFIQUE: _____		

5. SITUACION ECONOMICA

¿DE QUIEN DEPENDES ECONOMICAMENTE?	
PADRE ()	MADRE () HNOS ()
PARIENTES () DE SI MISMO ()	
¿SI RECIBES PENSION DE TUS PADRES O PARIENTES CUANTO ES?	
Si/.....	
¿SI DEPENDES DE TI MISMO?	
¿Que ocupacion tienes? _____	
¿Que condicion laboral? _____	
¿Cuánto ganas? _____	

6. ALIMENTACION

DONDE TOMAS TUS ALIMENTOS	
EN CASA DE PADRES	()
CASA DE PARIENTES	()
PENSION	()
RESTAURANT	()
RESTAURANT POPULAR	()
TE PREPARAS TUS ALIMENTOS	()
OTROS: _____	

7. VIVIENDA

DURANTE TUS ESTUDIOS UNIVERSITARIOS, CON QUE PERSONAS VIVES ?			
PADRE ()	MADRE ()	HERMANOS ()	PARIENTES (()
AMIGOS ()	CONYUGE() SOLO ()		
CONDICION DE LA VIVIENDA:			
PROPIA ()	ALQUILADA()	CUIDANTE ()	ALOJADO ()
OTROS _____			
TIPO DE VIVIENDA:			
DPTO ()	CASA INDEPENDIENTE ()	HABITACION ()	
MATERIAL DE CONSTRUCCION:			
NOBLE ()	ADOBE ()	CELULAR: _____	
SERVICIOS:			
AGUA: ()	DESAGUE ()	LUZ ()	TELEF. FIJO: _____

8. SALUD

ESTAS ASEGURADO ? SI () NO ()	
ESSALUD ()	SIS () PARTICULAR ()
EL LUGAR DONDE ESTAS ASEGURADO ES?	
CUSCO ()	PROV. () DIST. ()
ESPECIFICA:	
.....	

PERSONAS CON QUIEN COMUNICARNOS EN CASO DE EMERGENCIA:

_____ APELLIDOS Y NOMBRES _____ DIRECCION _____ TELEFONO FIJO O CELULAR

DECLARO BAJO JURAMENTO QUE LOS DATOS PROPORCIONADOS EN LA FICHA SE AJUSTAN A LA VERDAD

_____ FIRMA DEL ALUMNO

	H U E LLA
--	--------------------

 _____ ASISTENTE SOCIAL CUSCO _____ DE _____ DEL 2015

ANEXO 02

FORMULARIO DE RECOLECCIÓN DE LA BASE DE DATOS PARA LOS ESTUDIANTES DE LA ESCUELA PROFESIONAL DE MATEMÁTICA DE LA CORTE 2010-I HASTA LA CORTE 2012-II DE LA UNIVERSIDAD NACIONAL DE SAN ANTONIO ABAD DEL CUSCO.

N°	Código del estudiante	Sexo	Edad	Tipo de colegio	Tipo de modalidad de ingreso	Estado civil	Ocupación del jefe de hogar	Ingreso económico familiar	Número de hermanos	Situación familiar	Lugar de procedencia	El estudiante trabaja
1												
2												
3												
4												
5												
6												
7												
8												
9												
10												
.												
.												
.												
131												