

PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR

FACULTAD DE ECONOMÍA

**MODELO DE PREDICCIÓN DE LA DESERCIÓN DE LOS
ESTUDIANTES DE LA PUCE SEDE QUITO COHORTES 2014-2017**

María José Beletanga Calderón
mbeletanga118@puce.edu.ec

Director: Edwin Buenaño
evbuenano@puce.edu.ec

Quito, abril 2021

Resumen

La presente investigación estudió las tasas de deserción de los estudiantes de la Pontificia Universidad Católica del Ecuador (PUCE) en Quito durante el período 2014-2019, donde se evaluaron las cohortes 2014-02 hasta 2017-02. Para este estudio, se utilizaron datos administrativos proporcionados por la Universidad y otras fuentes públicas adicionales. Esta investigación fue de tipo exploratoria y explicativa y el método de investigación fue cuantitativo. El objetivo fue determinar el perfil de los estudiantes que abandonaron la Universidad, así como predecir las tasas de riesgo de desertar de la Institución, dentro del período estudiado. Para esta tesis se utilizaron dos de los métodos de los Modelos de Supervivencia. En primer lugar, se aplicó el modelo de Kaplan-Meier y posteriormente, a partir de sus resultados, se probó con el Modelos de Riesgos Proporcionales de Cox y algunas variaciones de este modelo. Mediante estos métodos, se encontró que las variables de desempeño académico son las variables más relevantes a la hora de predecir la deserción de los estudiantes. Éstas, son: número total de créditos, promedio de calificaciones y número de materias perdidas. También hubo otros factores importantes como la edad, la puntuación de la Prueba de Admisión, el estado civil de los padres, el cambio de especialidad, el cambio en el estado socioeconómico y la duración de la especialidad. Los resultados obtenidos podrían ayudar a predecir la deserción de los estudiantes de la PUCE sede Quito en las próximas cohortes.

Palabras clave: Deserción, Modelos de Supervivencia, Modelos de Kaplan- Meier, Modelo de Riesgos Proporcionales de Cox, Función de supervivencia, Tasas de riesgo.

Abstract

The current study examined dropout rates of the students from the Pontifical Catholic University of Ecuador (PUCE) in Quito during the period 2014-2019, where the cohorts 2014-02 through 2017-02 were evaluated. For this research, administrative data provided by the University and other additional public sources was used. This study was exploratory and explanatory also, the research method applied was quantitative. The aim was to determine the profile of students who dropped out, as well as to predict the hazard rates of students' withdrawal from the Institution, within the studied period. For this thesis, two of the Methods of Survival Analysis were used. Firstly, the Kaplan-Meier model was applied, and then from its results, some variations of the Cox Proportional Hazards Models were also put into place. By means of these methods, some variables were found to be more relevant than others when predicting students' dropout. Those variables showed statistical significance related to the performance of students, such as: total number of credits, grade average and number of subjects failed. There were also other important factors as age, the admission test score, fathers' marital status, change of major, change in socio-economic status and length of the major. These variables could help to predict students' risk of abandoning PUCE-Quito in the upcoming cohorts.

Key words: Dropout, Survival Models, Kaplan-Meier Models, Cox Proportional Hazards Model, Survival Function, Hazard rates.

Agradezco a Dios, quien ha sido mi fortaleza en los momentos más difíciles.

A mis padres, Roberto y María Augusta, quienes me han apoyado siempre y a través de su ejemplo me alientan a dar lo mejor de mí.

A mi hermana menor María Emilia, quien en los días más difíciles me saca una sonrisa.

A mi tutor de tesis, Edwin Buenaño, quien con mucha paciencia me supo guiar de la mejor manera para la realización de este trabajo.

A mi mejor amiga Nicole Bernal, quien me supo entender en mis momentos de estrés.

A Andrés Buitrón y Bernardo Villegas quienes me ayudaron con los datos de esta disertación.

A Pauly Navarro, a quien agradezco por haberme motivado a seguir esta carrera.

Al resto de mi familia y amigos.

MODELO DE PREDICCIÓN DE LA DESERCIÓN DE LOS ESTUDIANTES DE LA PUCE SEDE QUITO COHORTES 2014-2017

INTRODUCCIÓN.....	8
METODOLOGÍA.....	10
Preguntas de Investigación	10
General	10
Específicas.....	10
Objetivos de la Investigación	10
General	10
Específicos.....	10
Descripción de la Metodología	10
Tipo de Investigación	10
Técnicas de Investigación	11
Fuentes de Información	11
Procedimiento Metodológico	11
FUNDAMENTACIÓN TEÓRICA	12
Capital Humano	12
Retornos de la Educación	13
Economía de la Educación	15
Función de Producción Educativa	16
Deserción	17
Modelos de Supervivencia	21
Funciones de Riesgo y Supervivencia	22
Modelo de Kaplan-Meier	23
Modelo de Riesgos Proporcionales de Cox	24
Prueba de Schoenfeld	25
Modelos Estratificados	26
Covariables de tiempo dependiente	26
Prueba de Residuos de Cox-Snell	26
VARIABLES DEL MODELO.....	27
CAPÍTULO I: PERFIL DE LOS ESTUDIANTES	29
Descripción de las Variables	29
Modelos de Supervivencia de Kaplan-Meier	36
Desertor	36
Sexo	38
Edad	39

Estado Civil del Estudiante	40
Residencia Previa	41
Periodo de Entrada	42
Tipo de Entrada a la Universidad	43
Puntaje de la Prueba de Admisión	44
Duración Ideal de la Carrera	45
Número Ideal de Créditos por Carrera	46
Facultad	47
Descuento Promedio por Semestre	48
Categoría Socioeconómica de la Pensión Diferenciada (SPD)	49
Cambio en la Categoría SPD	50
Tipo de Colegio	51
Nivel de Educación de la Madre	52
Nivel de Educación del Padre	53
Estado Civil de la Madre	54
Estado Civil del Padre	55
Cambio de Carrera	56
Promedio de Créditos por Semestre	57
Promedio General de Notas	59
Número de Materias Perdidas	60
Perfil de los estudiantes desertores	61
CAPÍTULO II: MODELO DE LA DESERCIÓN	63
Modelo 1: Modelo de Riesgos Proporcionales de Cox	63
Modelo 2: Modelo Estratificado de Cox	65
Modelo 3: Modelo de Covariables de Tiempo Dependiente	67
Modelo 4: Modelo de Covariables de Tiempo Dependiente Estratificado	69
Estadísticos de Bondad De Ajuste	70
Prueba de Residuos de Cox-Snell	71
Resultados	72
CONCLUSIONES	75
Recomendaciones	77
ANEXOS	83
ANEXO A	84
ANEXO B	86
ANEXO C	87
ANEXO D	88
ANEXO E	89

Índice de Tablas y Gráficos

Tabla 1: Variables utilizadas en modelos cuantitativos.....	20
Tabla 2: Tabla de variables.....	28
Tabla 3: Tabla Resumen de la variable dependiente.....	29
Tabla 4: Tabla resumen de las variables explicativas relacionadas con las características individuales de los estudiantes.....	30
Tabla 5: Tabla resumen de las variables explicativas relacionadas con las condiciones iniciales en la Universidad.....	31
Tabla 6: Tabla resumen de las variables explicativas relacionadas con los factores socioeconómicos y familiares.....	33
Tabla 7: Tabla resumen de las variables explicativas relacionadas con los factores de desempeño.....	35
Tabla 8: Resultados del Modelo 1.....	64
Tabla 9: Prueba de Schoenfeld del Modelo 1.....	65
Tabla 10: Resultados del Modelo 2.....	66
Tabla 11: Prueba de Schoenfeld del Modelo 2.....	67
Tabla 12: Resultados del Modelo 3.....	68
Tabla 13: Resultados del Modelo 4.....	69
Tabla 14: Estadísticos de bondad de ajuste.....	70
Gráfico 1: Curva de supervivencia basado únicamente en el número de desertores.....	36
Gráfico 2: Curva de riesgo basado en el número de desertores sin efecto en otras variables.....	37
Gráfico 3: Curva de riesgo basado en el sexo del estudiante.....	38
Gráfico 4: Curvas de supervivencia basadas en las categorías de la edad del estudiante.....	39
Gráfico 5: Curvas de supervivencia basadas en las categorías del estado civil del estudiante.....	40
Gráfico 6: Curvas de riesgo basado en la residencia previa a la universidad.....	41
Gráfico 7: Curvas de supervivencia basados en las categorías del periodo de entrada.....	42
Gráfico 8: Curva de supervivencia basado en las categorías de tipo de entrada a la Universidad.....	43
Gráfico 9: Riesgo acumulado de basado en las categorías de la Prueba de Admisión.....	44
Gráfico 10: Curva de riesgo basado en la duración ideal de la carrera.....	45
Gráfico 11: Curva de supervivencia basado en el número de créditos ideales por carrera.....	46
Gráfico 12: Curvas de riesgo basadas en la Facultad.....	47
Gráfico 13: Curvas de supervivencia de las categorías de descuento promedio por semestre.....	48
Gráfico 14: Curvas de riesgo basado en las categorías de Pensión Diferenciada.....	49
Gráfico 15: Curvas de riesgo basado en el cambio de categoría SPD.....	50
Gráfico 16: Curvas de riesgo basado en las categorías de tipo de colegio.....	51
Gráfico 17: Curvas de riesgo basado en las categorías del nivel de educación de la madre del estudiante.....	52
Gráfico 18: Curvas de riesgo acumulado basado en las categorías del nivel de educación del padre del estudiante.....	53
Gráfico 19: Curva de supervivencia basada en el estado civil de la madre del estudiante.....	54
Gráfico 20: Curva de supervivencia basada en el estado civil del padre del estudiante.....	55
Gráfico 21: Curvas de riesgo basado en las categorías de cambio de carrera.....	56
Gráfico 22: Gráfico de caja y bigotes basado del promedio de créditos por semestre.....	57
Gráfico 23: Curvas de supervivencia basadas en el promedio de créditos por semestre.....	57
Gráfico 24: Gráfico de caja y bigotes basado en el promedio general de notas de los estudiantes.....	59
Gráfico 25: Curvas de supervivencia basadas en las categorías del promedio general de notas.....	59
Gráfico 26: Curvas de riesgo basadas en las categorías del total de materias perdidas.....	60
Gráfico 27: Prueba Cox-Snell de los Modelos.....	71

INTRODUCCIÓN

La educación es un derecho universal e ineludible que juega un papel fundamental en la transformación social, en la configuración de la subjetividad individual y en el crecimiento y desarrollo de los países. Con el tiempo, se la ha reconocido como una de las prácticas sociales más importantes, a través de la cual toman forma los dominios de la política, la economía, la sociedad y la cultura (Lewis, 2020). Por este motivo, el objetivo de los diferentes actores sociales es lograr que este derecho se cumpla para toda la población.

A pesar de las diferentes estrategias y políticas planteadas, la educación universal presenta múltiples desafíos, uno de ellos es el abandono de los estudios o también llamado deserción. En América Latina aproximadamente veinte millones de jóvenes se inscribieron en un Instituto de Educación Superior (IES) en la última década; sin embargo, solo la mitad obtuvo un título (Banco Mundial, 2017). Las distintas investigaciones revelan que existe un número importante de estudiantes que no logran culminar sus estudios universitarios, generando un gran costo social (UNESCO 2004, citado por Díaz Peralta, 2008).

La deserción universitaria es un problema que abordan la mayoría de IES alrededor del mundo. Este fenómeno acarrea consecuencias negativas tanto para los estudiantes, como para las instituciones y sociedad en general. Según el Instituto Internacional para la Educación Superior en América Latina y el Caribe (2007), existen varias secuelas, tales como: la retroalimentación del círculo de la pobreza; el aumento de las tasas de desempleo y subempleo; la dependencia de la asistencia gubernamental; el descenso en los índices de eficiencia y calidad de los centros de estudios superiores; implicaciones económicas y financieras para las IES; consecuencias psicológicas y económicas para los desertores y sus familias; entre otras.

En el Ecuador, el abandono de la educación superior representa un gran obstáculo para el sistema educativo. Según un informe presentado por la Secretaría de Educación Superior, Ciencia y Tecnología (Senescyt) en 2019, aproximadamente 23.000 universitarios desertan cada año del sistema de educación superior (Machado, 2019). Además, según cifras presentadas por la misma institución, 26% de los estudiantes abandonan la carrera durante los primeros semestres (Rodríguez F. , 2019). Esta situación afecta directamente a las IES ya que reduce sus ingresos y disminuye su calificación de calidad, pues las tasas de retención y graduación son un indicador clave para medirla (Contrato Social, 2014). Por este motivo, múltiples establecimientos educativos se han enfocado en encontrar las causas de la deserción y una solución al problema.

La Pontificia Universidad Católica del Ecuador (PUCE) no es ajena a este fenómeno. Según la Dirección de Aseguramiento de Calidad, 25% de los estudiantes matriculados en 2015, desertaron para 2017. Por esta razón, se desea plantear un modelo que prediga la deserción de los estudiantes a través de la identificación de perfiles o características de estudiantes que representan mayor riesgo. La relevancia de esta investigación se apoya sobre el escaso análisis que existe sobre este fenómeno, tanto a nivel nacional como institucional; pues pretende crear consciencia sobre la deserción universitaria y busca fortalecer al sistema educativo. El objetivo de este trabajo de investigación es desarrollar un modelo cuantitativo, a partir de datos administrativos de la universidad y algunas fuentes externas que permitan

complementar dicha información. Para ello, se utilizarán técnicas de modelos de supervivencia. Este aporte pretende explicar el fenómeno de estudio y en el caso particular de la PUCE, tener mayor información para la toma de decisiones e implementación de políticas que reduzcan el abandono universitario. De esta manera, se generará un beneficio tanto para la Universidad, como para el alumnado y la sociedad en general.

Este trabajo de investigación cuenta con dos capítulos. Previo a la presentación de los dos capítulos se presenta la revisión bibliográfica que explica el concepto teórico de la deserción. Posteriormente, en el primer capítulo se realiza un análisis de estadística descriptiva y un análisis no paramétrico que describe el perfil de los estudiantes desertores. En el segundo capítulo y partiendo de la información obtenida en el primer capítulo, se construye un modelo que utiliza la metodología de los modelos de supervivencia semi-paramétricos para predecir la deserción en la PUCE Sede Quito. Finalmente se concluye y se presentan algunas recomendaciones para futuras investigaciones.

METODOLOGÍA

Preguntas de Investigación

General

- ¿Cómo se puede predecir la probabilidad de deserción de los estudiantes de la PUCE sede Quito?

Específicas

- ¿Cuáles son los perfiles de los estudiantes desertores de la PUCE sede Quito de las cohortes 2014 – 2017?
- ¿Cuáles son los factores determinantes que predicen la deserción de los estudiantes de la PUCE sede Quito según los registros administrativos de las cohortes 2014 - 2017?

Objetivos de la Investigación

General

- Analizar la probabilidad de desertar de los estudiantes de la PUCE sede Quito a través de un modelo de predicción.

Específicos

- Caracterizar y describir los perfiles de los estudiantes de las cohortes 2014-2017 que han desertado de la PUCE sede Quito.
- Identificar y analizar los factores que contribuyeron a la deserción universitaria en la PUCE sede Quito y construir un modelo cuantitativo que prediga la probabilidad de desertar de los estudiantes, basándose en información de los registros administrativos de las cohortes 2014 - 2017.

Descripción de la Metodología

Tipo de Investigación

Este estudio corresponde a una investigación de tipo exploratoria, ya que pretende determinar las variables que podrían predecir de mejor forma la deserción universitaria en los estudiantes de la PUCE sede Quito. También es explicativa, ya que intenta examinar la naturaleza de las relaciones y el efecto de las variables sobre el fenómeno estudiado. Las razones por las que se utilizan estos dos tipos de investigación se deben a que existen escasos conocimientos sobre el tema, se desea determinar las características del fenómeno y se pretende predecir relaciones causales.

Técnicas de Investigación

El método que se utiliza en la investigación es inductivo, ya que parte de un caso específico, como es el estudio de la PUCE sede Quito de las cohortes 2014-2017, para llegar a conclusiones generales sobre la deserción universitaria de dicha institución. Además, el enfoque de la investigación es cuantitativo, ya que se basa en diferentes cálculos estadísticos, la construcción de un modelo de supervivencia y la comprobación del mismo.

Fuentes de Información

La información proviene de los datos administrativos de la Dirección General de Estudiantes (DGE) y de la Dirección General de Admisiones (DGA) de la PUCE¹ sede Quito. Adicional a esta fuente, se obtuvo información del Ministerio de Educación y del Registro Civil. Con las fuentes adicionales se completó la información faltante en los registros administrativos de la Universidad y se añadieron otras variables sociodemográficas.

Procedimiento Metodológico

En la fundamentación teórica de la disertación, se construyó en base a una extensa revisión bibliográfica. El primer capítulo responde la primera pregunta y al primer objetivo planteado. En este capítulo se presentó una tabla de estadística descriptiva que resume la composición de la base de datos y se realizó un análisis no paramétrico de cada una de las variables. A través de este análisis se evaluó el perfil de los desertores y se seleccionaron las variables que serían utilizadas en el modelo. El segundo capítulo es la respuesta a la segunda pregunta específica, en donde se probaron diferentes modelos de supervivencia para predecir la deserción de los estudiantes. Es necesario recalcar que solo se presentaron los modelos con mejor ajuste.

Para la construcción de la base de datos, el análisis de estadística descriptiva y la modelización; se utilizó el software RStudio.

¹ La Pontificia Universidad Católica del Ecuador, es una institución privada situada en Ecuador. Una de sus sedes principales y más grande es la que se encuentra en Quito.

FUNDAMENTACIÓN TEÓRICA

Este apartado presenta una revisión bibliográfica de los diferentes conceptos y teorías relacionadas con la deserción. El objetivo es construir un marco teórico que explique el contexto, la importancia y la evolución de la teoría de la deserción. En primer lugar, se explica la teoría del capital humano, que fue la conjetura que dio inicio al estudio de la economía de la educación. A partir de la creación de esta rama, nació la importancia de evaluar los retornos de la educación y la función de producción educativa, la cual se basa en encontrar los resultados de la educación a partir del estudio de una serie de factores. Por último, se describe uno de estos productos, que es la deserción. Este concepto será el punto de partida de la investigación.

Capital Humano

La teoría del Capital Humano ha variado su significado y concepción a través de la historia. Es necesario reconocer que este término no fue usado hasta el siglo XX debido a los malentendidos causados por considerar al ser humano como capital. Los economistas clásicos como Adam Smith, Karl Marx y David Ricardo lo concebían como la capacidad de realizar trabajo manual en donde se requería pocos o nulos conocimientos y habilidades. Según ellos, todos los trabajadores tenían la misma dotación, por tanto, eran considerados sustitutos perfectos entre ellos, sin distinción (Schultz T. , 1961). Irvin Fisher (citado por Carnoy, 2013) intentó introducir el término, definiendo capital como cualquier activo que aumenta su valor a pesar de que este no pueda ser comprado, vendido o negociado en un mercado. Esta reformulación intentaba introducir las habilidades propias de la persona en la descripción de capital; sin embargo, el economista fue criticado y su teoría abandonada por varios años.

A partir de la década de 1950 se popularizó el término capital humano, sin embargo, no era totalmente aceptado en la academia, incluso en esta época fue considerado controversial. Los tres principales economistas que llegaron a desarrollar diferentes estudios empíricos y teóricos sobre el tema, fueron: Theodore Schultz (1902-1998), Jacob Mincer (1922-2006) y Gary Becker (1930-2014). Las diferentes investigaciones de estos autores generaron grandes aportes en el campo de la educación, economía laboral, políticas públicas y teorías de crecimiento económico.

Theodore Schultz (1961) resaltó la importancia de la inversión en capital humano. El economista recalcó que el crecimiento económico, más allá de deberse a la acumulación de tierras, al aumento de horas trabajadas o al incremento de maquinaria; se debe al capital humano, que es toda inversión que realizan las personas en ellas mismas. Schultz hizo referencia a la educación, salud, migraciones y capacitaciones. Este tipo de inversión aumenta la productividad de los trabajadores y las habilidades que éstos pueden tener, generando un aumento en la producción y por ende en los ingresos de las naciones.

Complementando la definición de Schultz, Jacob Mincer (1981) añadió que el capital humano también se refiere a todas las capacidades adquiridas que se desarrollan a través de la educación formal e informal en la escuela y en el hogar, y mediante la formación, la experiencia y la movilidad en el mercado laboral. La generación de capital humano implica no sólo la transmisión y encarnación del conocimiento, sino

también la producción de nuevo conocimiento. Estos, se convierten en la fuente de innovación y de cambio técnico que impulsan todos los factores de producción. Esta última función del capital humano genera crecimiento económico mundial independientemente de su ubicación geográfica o condiciones iniciales.

Asimismo, Gary Becker, siguiendo la misma línea de pensamiento de Schultz y Mincer, aseguró que la educación y la formación son las inversiones más importantes del capital humano. A través de varios estudios realizados en Estados Unidos, demostró que la educación secundaria y universitaria aumentan en gran medida los ingresos de las personas; incluso después de compensar los costos directos e indirectos, ajustar los antecedentes familiares y discernir las capacidades de cada individuo. Los ingresos de las personas con más educación casi siempre están por encima del promedio, aunque los beneficios son generalmente mayores en los países menos desarrollados (Becker, 1993).

El análisis del capital humano asume que la educación aumenta los ingresos y la productividad principalmente al proporcionar conocimientos, habilidades y una forma de analizar los problemas (Becker, 1993). Por el contrario, la teoría credencialista manifiesta que los títulos y la educación simplemente reafirman las habilidades ya existentes de los individuos. Según esta conjetura, se creó el modelo de señalización atribuido a Michael Spence (1973, citado por Page, 2010), en donde los graduados universitarios superan a los graduados de secundaria no porque los años de educación aumenten la productividad, sino porque los estudiantes más productivos van a la universidad. Una implicación del modelo es que los trabajadores más educados reciben un salario más alto porque la educación les proporciona una credencial, mas no por las habilidades adquiridas.

A pesar de la existencia del credencialismo o la teoría del “screening”, se ha realizado un mayor número de investigaciones que demuestran que los empleadores pagan salarios más altos a los trabajadores más educados porque su habilidad, conocimiento, productividad y experiencia son mayores que la de los trabajadores menos educados. Este gran aumento de los ingresos reales de los trabajadores, es explicado como un retorno a la inversión que se ha realizado. Por lo tanto, los salarios y los gastos en educación, también son una estimación de la inversión de capital humano (Schultz T. , 1961).

En resumen, la inversión en capital humano es un elemento fundamental en el proceso productivo y desarrollo de las naciones. En gran medida, este se ve representado por los niveles de educación alcanzados por cada uno de los individuos. Por lo tanto, es necesario estudiar el papel que juega la educación.

Retornos de la Educación

La educación, en todos los sentidos, es uno de los motores principales del desarrollo. Ningún país puede lograr un desarrollo económico sostenible sin una inversión sustancial en capital humano. La educación enriquece la comprensión de las personas sobre sí mismas y sobre el mundo, además de mejorar la calidad de vida y generar amplios beneficios sociales para las personas y la sociedad. La educación aumenta la productividad y la creatividad de las personas y promueve el espíritu empresarial y los

avances tecnológicos. Además, es un canal que posibilita la equidad a través de la distribución de ingresos y asegura el progreso económico y social (Ozturk, 2008).

Con la globalización, el desarrollo tecnológico y la economía de la información se han desarrollado dos teorías: la primera señala que la educación añade habilidades al trabajo y, por lo tanto, incrementa la capacidad laboral, intensificando la producción. Por otro lado, la segunda teoría señala que la educación mejora la capacidad del trabajador para la innovación, lo que aumenta su propia productividad y la de los otros trabajadores. Ambas teorías indican que una sociedad con mayor educación tiene más oportunidades para expresar sus capacidades creativas (Becker, 1993).

El apoyo a la escolarización y la creciente evidencia de los retornos de la educación como un determinante fundamental del crecimiento económico han elevado la importancia y visibilidad de la investigación sobre la formación de capital humano. Este tipo de políticas abordan las complicaciones introducidas por los innumerables e interrelacionados procesos de toma de decisiones de las familias, los maestros, los administradores y los formuladores de políticas. Se han utilizado una variedad de métodos para identificar relaciones causales, que van desde modelos basados en la maximización de la utilidad hasta enfoques experimentales y cuasi-experimentales (Meghir & Steven, 2011).

El interés por analizar y cuantificar los rendimientos de la educación ha dado origen a una extensa literatura en el ámbito internacional. La herramienta empírica utilizada en la mayoría de investigaciones ha sido la ecuación de ingresos de Mincer (1974), a través de la cual se estiman el impacto de un año adicional de estudios en los salarios de los individuos. La ecuación tradicional de Mincer, estima por Mínimos Cuadrados Ordinarios (MCO) un modelo semi-logarítmico, usando como variable dependiente el logaritmo de los ingresos y como variables independientes los años de educación y la experiencia laboral. Los datos utilizados para su estimación provienen tradicionalmente de datos transversales (Freire Seoane & Teijeiro Álvarez, 2010).

Una de las principales limitaciones del modelo de Mincer es que toma los retornos monetarios de la educación como el único beneficio que ésta ofrece, no obstante, diferentes estudios empíricos han demostrado que existen otras ventajas que no son cuantificables. Haveman y Wolfe (1984, citado por Dickson & Harmon, 2011) describen varios beneficios privados no monetarios que incluyen la mejora de la salud del individuo y de su familia en general, la fertilidad (lograr el tamaño deseado de la familia), mayor disfrute de otras actividades, eficiencia en la elección del consumidor, mayores tasas de ahorro y una mejor relación matrimonial. Por otro lado, los beneficios públicos no monetarios, incluyen la reducción del delito, la cohesión social, el cambio tecnológico y las donaciones caritativas. Asimismo, Dee (2004, citado por Dickson & Harmon, 2011) encuentra que la educación aumenta la participación electoral, la libertad de expresión y el conocimiento cívico. Además de los beneficios ya mencionados, Oreopoulos (2007, citado por Dickson & Harmon, 2011) encuentra que además de aumentar la riqueza de por vida en aproximadamente un 15%, mayor educación puede llevar a las personas a mejorar psicológicamente, ya que reduce la probabilidad de estar desempleado y de ser infeliz, a la vez que aumenta su sentido de autoestima y autoconciencia.

Según una investigación realizada por Psacharopoulos (1985), los retornos de la educación dependen de una serie de factores, que pueden ocasionar que no todos obtengan los mismos beneficios. En su estudio

se evalúa 61 países de los diferentes continentes y con diferente nivel de desarrollo. En esta investigación se reveló que los beneficios obtenidos en los países menos desarrollados son mayores en comparación a los países más desarrollados, lo cual es explicado por las tasas de escolarización y mercado laboral. Los diferentes modelos e investigaciones tienen distintos resultados, ya que dependen de las características endógenas del país. De esta manera se comprobó que los retornos monetarios pueden cambiar de acuerdo al sexo, raza, edad y ubicación geográfica.

Debido a la importancia de la educación dentro de la sociedad y al rol que ésta juega dentro de los aspectos económicos, se ha desarrollado una rama que estudia y analiza este campo.

Economía de la Educación

La economía se define como la asignación de recursos escasos para satisfacer las necesidades del ser humano. Tomando como referencia este concepto, se puede argumentar que la economía de la educación estudia la organización de la sociedad y la utilización de los recursos escasos para producir varios tipos de conocimiento y habilidades a través de la educación formal. Además, considera cómo estos elementos se distribuyen a los diferentes grupos en la sociedad. Cabe destacar que la economía de la educación es una rama multidisciplinaria que comprende una gama diversa de temas sociales y políticos. (Brewer & McEwan, 2010).

Según Carnoy (2006) la economía estudia a la educación en tanto mercancía de inversión y consumo. Debido a que la educación utiliza una gran cantidad de recursos, los economistas deben preocuparse por conocer el uso efectivo y eficiente de los recursos con el fin de alcanzar objetivos sociales implícitos y explícitos. Los beneficios de una población educada recaen sobre el conjunto de la sociedad, por tanto, esta es vista como una mercancía parcialmente pública.

Como se mencionó anteriormente, uno de los principales factores del capital humano es la educación. Mincer (1981) menciona que la inversión realizada en cada uno de los individuos, puede ser descrita dentro un ciclo de vida cronológico que coincide con las diferentes etapas de educación. Los recursos en el cuidado infantil se ven representado con el preescolar, a esto le sigue la educación formal que reciben los niños y adolescentes en las escuelas, y por último las inversiones en educación superior, movilidad en el mercado laboral, elección de trabajo, capacitación laboral y esfuerzo laboral que ocurren durante la vida adulta.

Si bien todas las etapas son importantes para el desarrollo de la persona, la educación superior y la formación laboral es una de las que más aporta a la economía. Por esta razón, es necesario estudiar los aportes, relaciones y fallas de la Economía de la Educación Superior. Este campo estudia las asignaciones de recursos para la producción de servicios educativos postsecundarios, la distribución de estos y sus efectos entre la población. Los temas que generalmente se consideran en esta área son: demanda de educación, financiamiento, gestión de instituciones, subvenciones y recursos humanos. La educación superior promueve beneficios para el individuo educado, para la eficiencia y la equidad de la sociedad en su conjunto. La provisión de educación superior es vista como una táctica clave para aumentar el crecimiento económico y disminuir las desigualdades de la sociedad (Mora & Villa, 2003).

lado, se valoran múltiples resultados (por ejemplo, habilidades básicas, habilidades vocacionales, creatividad y actitudes), que pueden acumularse de manera progresiva y discernirse muchos años en el futuro. Por otro lado, los insumos pueden ser difíciles de medir, y las dimensiones que se miden pueden no capturar todas sus características de manera adecuada. Además, existen insumos no escolares que influyen directamente en los resultados académicos, tales como el contexto familiar, la interacción con diferentes grupos de la sociedad, el ambiente en el que se desarrolla el estudiante, entre otros. Por lo tanto, el carácter disperso del proceso de educación obstaculiza su evaluación (Jencks et al., 1981).

La segunda perspectiva es aún más compleja, pues ésta incluye problemas de asimetrías de información y se relaciona directamente con la teoría del principal y agente. Uno de los principales problemas de este caso es la selección adversa, que sucede cuando las autoridades no conocen las habilidades de sus subordinados (coordinador escolar, profesores, etc.) quienes trabajan directamente con los alumnos. Otro gran problema es el riesgo moral, que se da cuando las personas encargadas de la educación persiguen objetivos diferentes a los planteados por las autoridades (Brewer & McEwan, 2010).

Con respecto a esta situación Carnoy (2013) menciona que existen diferentes tomadores de decisiones en la función de producción educativa. Primero se encuentran las autoridades e instituciones gubernamentales como es el Ministerio de Educación, en la siguiente posición jerárquica se encuentran las autoridades máximas de los centros educativos, tales como: los rectores, directores o decanos; posteriormente los profesores, y por último los padres y los mismos estudiantes. Los hacedores de política pueden no tener los mismos objetivos que el principal, el agente y los estudiantes; por ende, los resultados pueden verse afectados. Otro punto, que es necesario analizar es la minimización de costos y la maximización de calidad. El costo de la educación dependerá de los objetivos planteados por las instituciones (Carnoy, 2013).

Con todas las restricciones existentes, los economistas realizan varios supuestos simplificados para posibilitar el estudio de la función de producción educativa. Estas asunciones proporcionan un marco para examinar y explicar el comportamiento de los estudiantes. Para desarrollar cualquier tipo de modelo cuantitativo, es necesario encontrar un conjunto relevante de insumos que describan las características técnicas y las especificaciones del proceso. La mayoría de los estudios de funciones de producción miden la producción mediante las puntuaciones de las pruebas de rendimiento o de aptitud, actitudes de los estudiantes, tasas de asistencia, tasas de continuación, entre otros (Hanushek, 1987).

En la investigación que se presenta en este trabajo se pretende identificar los factores asociados a la deserción, los cuales son intrínsecos de la función de producción de la educación. Los insumos o factores utilizados se obtendrán de la información administrativa de los estudiantes. Por lo tanto, es necesario conocer al respecto de la teoría de la deserción.

Deserción

Tinto & Cullen (1973), proponen dos definiciones sobre la deserción. La primera se refiere a aquellas personas que dejan la universidad en la que están matriculados. El segundo concepto identifica como desertores, a aquellos individuos que nunca recibieron un título de ninguna institución de educación

superior. Por otro lado, Pervin et al. (1966) definen como desertor a cualquier estudiante que ha dejado la universidad por cualquier periodo de tiempo, sin importar el motivo. Este concepto no distingue entre los estudiantes que solicitan un cambio de institución, entre los que retornan después de un cierto número de semestres o entre los que han dejado la universidad por algún problema psicológico o de salud.

Definir la deserción como cualquier persona que abandona la universidad en la que está inscrita tiene, en su aplicación, fortalezas y debilidades. El principal beneficio radica en la facilidad con la que se pueden recopilar datos confiables y actualizados sobre la deserción en diversos tipos de instituciones de educación superior. Este tipo de datos es comúnmente utilizado en las investigaciones relacionadas con ciencias sociales y planificación institucional. Sin embargo, una de las debilidades, es que tiende a ignorar los fenómenos de transferencia entre universidades, por lo tanto, subestima el número real de desertores del sistema de educación superior (Tinto & Cullen, 1973).

Una vez establecida la noción de deserción, es necesario evaluar el proceso y factores que se encuentran implicados dentro de este fenómeno. Para explicar el abandono universitario se debe evaluar algunos modelos que han consolidado las bases teóricas y han dado paso a la creación de estudios cuantitativos. A través de la historia, se han desarrollado cuatro modelos teóricos que explican y predicen la deserción de la educación superior en las instituciones. En orden cronológico son: Spady (1970; 1971), Tinto (1975; 1997), Pascarella (1980) y Bean (1985) (citado por Nicoletti, 2019).

Según Spady (1970) la deserción es un proceso complejo que debe evaluarse de manera longitudinal con un enfoque interdisciplinario. El autor recalca que una manera factible de analizar la deserción es a través de la teoría del suicidio de Durkheim, la cual centra su atención en las causas sociales. Spady desarrolló el término “Congruencia Normativa”, que es la interacción entre los atributos de los estudiantes (es decir, disposiciones, intereses, actitudes y habilidades), las influencias, expectativas y demandas impuestas por diversas fuentes en el ámbito universitario (Spady, 1970). El “Modelo de Proceso de Deserción” (Model of the Dropout Process) se obtuvo después de un análisis de 109 investigaciones teóricas y empíricas desarrolladas entre 1940 y 1970. Este modelo sociológico integra variables del ámbito académico y social, en componentes generales o también llamados módulos. La principal conclusión a la que llega el modelo es que los estudiantes que logran adaptarse más fácilmente, tendrán una mejor interacción con compañeros, profesores y personal administrativo, lo que reduce su probabilidad de desertar. Cabe destacar que uno de los principales problemas de esta teoría fue la operacionalización de variables, ya que varias de ellas no eran medibles (Nicoletti, 2019).

Complementando y expandiendo la teoría de Spady, Tinto (1975) desarrolla un modelo teórico más amplio que explica el proceso de interacción entre los individuos y las instituciones. Éste es un modelo longitudinal que describe las diferentes condiciones bajo las cuales ocurre la deserción. En su modelo, Tinto resalta tres condiciones previas que tienen impacto directo e indirecto sobre el desempeño en la universidad. En primer lugar, se encuentran los atributos individuales (por ejemplo, sexo, raza, capacidad); en segundo lugar, están las experiencias preuniversitarias (por ejemplo, promedios de calificaciones, logros académicos y sociales); y, por último, los antecedentes familiares (por ejemplo, atributos de estatus social, climas de valores, climas esperados). Estas condiciones iniciales influyen y predicen el desarrollo de las expectativas y compromisos, lo que se traduce en un reflejo de las experiencias personales, decepciones y satisfacciones que los estudiantes viven en el contexto

universitario (Tinto, 1975). Son todas estas características, interacciones y compromisos los que permiten predecir la deserción de los estudiantes.

Pascarella (1980), también propuso un modelo longitudinal, este tomaba en cuenta las características previas de los estudiantes (antecedentes familiares, aptitudes, aspiraciones, logros académicos, etc.), los factores institucionales (cultura del profesorado; la organización, estructura e imagen institucional, estándares, etc.), el contacto informal con el profesorado (contexto, exposición, enfoque e impacto), otras experiencias de la universidad (cultura de los compañeros, clase, extracurriculares y pasatiempos) y resultados educativos (desempeño académico, desarrollo intelectual, aspiraciones profesionales, etc.); lo que permite al estudiante decidir si persiste o deserta. Según el autor, la evidencia sugiere que la relación estudiante-profesor fuera de las aulas y otras experiencias, pueden tener un impacto positivo y medible en varias facetas del desarrollo individual del estudiante. A diferencia de los otros autores, él manifestaba que estos grupos de variables deben ser observados por múltiples periodos de tiempo, ya que interactúan entre ellas de manera recíproca y dinámica. Este modelo sentó la base teórica y motivó a la investigación correlacional entre las variables (Pascarella E. , 1980).

Bean (1985), crea un nuevo modelo conceptual que se diferencia de los anteriores. Este puede predecir el síndrome de deserción, diferenciando el intento de irse, la disyuntiva de desertar y el momento en el que en verdad abandona la institución. Este modelo enfatiza en los factores académicos, socio-psicológicos y ambientales, que a su vez influyen en tres factores que resultan del proceso de socialización. El primer factor son las calificaciones universitarias que indican una evaluación externa positiva del comportamiento pasado de un estudiante. El segundo, es el ajuste institucional que indica la impresión subjetiva sobre el grado de adaptación propia sobre sí mismo, en base a las normas y valores de sus compañeros y mentores. Por último, el compromiso institucional indica el apego personal del estudiante a la institución. Finalmente, se evalúa cómo estos factores influyen sobre el síndrome de deserción. Cabe destacar que Bean, además de realizar un modelo teórico, lo complementó con un modelo cuantitativo que usaba las mismas variables expuestas e intentaba predecir el síndrome de deserción (Bean J. , 1985).

Por último, Tinto (1997) modifica y añade información a su modelo. Después de una serie de estudios y varias investigaciones experimentales realizadas por el autor, él asevera que es necesario considerar las características propias del aula a la que pertenecen los alumnos y la trayectoria de aprendizaje. De esta manera, Tinto expande su modelo y lo organiza a través de grandes categorías de variables, tales como: atributos previos a la entrada, los compromisos planteados para alcanzar sus objetivos, experiencia institucional, integración normativa (académica) y personal (social), esfuerzo de los estudiantes, resultados (aprendizaje), los compromisos finales y factores exógenos (comunidades externas) (Tinto, 1997). Este modelo analiza los factores personales, académicos y sociales de los estudiantes; realizando una valoración completa, que permite predecir con mayor certitud la deserción.

Los modelos presentados anteriormente son algunas construcciones conceptuales que han creado modelos teóricos de la deserción. Sin embargo, hay límites entre la teoría y la práctica, pues como se mencionó anteriormente, uno de los principales problemas es la cuantificación de variables, ya que varias de ellas no son medibles o no suelen estar disponibles. A pesar de las dificultades presentadas, existen varias investigaciones que han intentado predecir la deserción a través de modelos cuantitativos.

Como se puede observar en la Tabla 1, se presentan algunas variables que han sido utilizadas en diferentes estudios cuantitativos, por lo tanto, muestra los posibles factores que intervienen en el fenómeno de la deserción. En algunos de los casos los factores seleccionados se han obtenido de bases de datos administrativas y en otros se han realizado estudios experimentales, cuasiexperimentales o análisis cualitativos previos para obtener la información.

En resumen, no existe una única definición o caracterización de la deserción, pues puede tener múltiples connotaciones de acuerdo al comportamiento que ésta presente en los estudios que se realicen. Además, las diferentes investigaciones demuestran que existen varios factores que inciden sobre el fenómeno.

Modelos de Supervivencia

El análisis de supervivencia es un término genérico para una colección de métodos y modelos estadísticos que tienen como objetivo medir el tiempo antes de que un evento (muerte, infarto, boda, graduación, etc.) ocurra y establecer factores de riesgo que puedan explicar la causa (Aalen, Borgan, & Gjessing, 2008). Los modelos se dividen generalmente en modelos no paramétricos, semi-paramétricos y paramétricos. En esta sección solo se analizará la metodología del modelo no paramétrico de Kaplan-Meier y los modelos semi-paramétricos relacionados con el modelo de riesgos proporcionales de Cox. Sin embargo, antes de explicar cada uno de los modelos es necesario entender algunos conceptos clave.

Los modelos de supervivencia son modelos de regresión con estimadores de probabilidad que intentan calcular la tasa de riesgo de un evento a partir de una selección de factores. Por facilidad, el evento es definido como muerte o falla del individuo observado. La variable dependiente (resultado) es la tasa de riesgo o la probabilidad condicional de que un evento ocurra en un intervalo de tiempo particular (t). En otras palabras, se examina cuánto tiempo pasa hasta que ocurra el evento de interés. El objetivo de estos modelos no es solo examinar los efectos en el tiempo hasta que ocurre el evento, sino también evaluar la relación del tiempo de supervivencia con las variables explicativas. Las variables explicativas (covariables o variables independientes) son aquellos factores o características que generan un impacto sobre la variable dependiente. Estas variables pueden ser fijas o variables en el tiempo (Mills, 2011).

El tiempo que tarda antes de que ocurra un evento, se denomina tiempo de supervivencia, intervalo, tiempo de espera, período de riesgo, duración, etc. El eje del tiempo puede ser continuo o discreto, lo que dependerá de la frecuencia con la que se mida. Una escala continua es más precisa por lo tanto tiene unidades pequeñas (segundos, días, meses), mientras que el tiempo discreto utiliza unidades más grandes (meses, años, décadas), ya que solo se conoce que el evento ocurrió dentro de un intervalo de tiempo. Aunque la mayoría de los procesos ocurren en tiempo continuo, a menudo se miden en tiempo discreto (Singer & Willet, 2003).

Un factor distintivo de los modelos de supervivencia es la censura, que se da cuando se tiene información sobre el tiempo de supervivencia de un individuo, pero no se sabe el tiempo exacto de supervivencia (Kleinbaum y Klein, 2005, como se citó en Mills, 2011). Existen diferentes tipos: la censura por la derecha, que ocurre cuando el evento no es experimentado hasta la última observación y la censura de

intervalo que ocurre cuando no se conoce el momento exacto del evento. Otro fenómeno similar a la censura es el truncamiento, que es la ausencia de información dentro del tiempo de estudio (Mills, 2011).

Funciones de Riesgo y Supervivencia

Siguiendo a Rodríguez (2010), a continuación, se explica detalladamente las funciones de supervivencia y riesgo del análisis de supervivencia. Ambas funciones se construyen en función de T y t . T es una variable aleatoria no negativa que represente el tiempo de espera hasta que ocurra un evento y t representa el intervalo de tiempo en el que puede ocurrir el evento. T es una variable aleatoria continua con función de densidad de probabilidad (pdf): $f(t)$ y función de distribución acumulada (cdf): $F(t)$.

$$F(t) = \int_0^t f(x)dx = \Pr\{T \leq t\}$$

$F(t) = \Pr\{T \leq t\}$, representa la probabilidad de que el tiempo de supervivencia de T sea menor o igual al intervalo del tiempo de estudio (t). Por lo tanto, la función de supervivencia es el complemento de la función de distribución acumulada, ya que se desea encontrar la probabilidad de que el evento no haya ocurrido dentro de la duración t .

$$S(t) = \Pr\{T \geq t\} = 1 - F(t) = \int_t^{\infty} f(x)dx$$

Esta función expresa la probabilidad de que un tiempo de supervivencia T sea igual o mayor que algún tiempo t . $S(t)$ denota la proporción de sujetos que sobreviven más allá de t . En el origen $t = 0$, $S(0) = 1$, lo que significa que todos los sujetos del estudio están vivos. $S(t)$ es una función que disminuye con el tiempo, a medida que los sobrevivientes mueren.

Para cualquier punto de $F(t)$, la función de densidad de probabilidad es:

$$f(t) = \frac{dF(t)}{d(t)} = F'(t)$$

Esto implica que:

$$f(t) = \lim_{dt \rightarrow 0} \frac{F((t + dt) - F(t))}{dt}$$

$$f(t) = \lim_{dt \rightarrow 0} \frac{\Pr\{t \leq T \leq t + dt\}}{dt}$$

Se observa que $f(t)$ expresa la probabilidad instantánea incondicional. En otras palabras, describe la probabilidad no condicionada sobre covariables en cualquier instante t .

La función de riesgo o la tasa de ocurrencia del evento en la duración t es igual a la densidad de eventos en t , dividida por la probabilidad de sobrevivir hasta la duración máxima alcanzada sin experimentar el evento. Se expresa de la siguiente manera:

$$\lambda(t) = \frac{f(t)}{S(t)}$$

$$\lambda(t) = \lim_{dt \rightarrow 0} \frac{\Pr\{t \leq T \leq t + dt \mid T \geq t\}}{dt}$$

El numerador representa la probabilidad condicionada de que el evento ocurra en el intervalo $[t, t + dt)$ dado que no ha ocurrido antes, y el denominador es el ancho del intervalo. Dividiendo ambos, se obtiene

la tasa de ocurrencia de eventos por unidad de tiempo. Tomando el límite a medida que el ancho del intervalo desciende a cero, se obtiene la tasa instantánea de ocurrencia.

Es necesario recordar que:

$$S(t) = \int_t^{\infty} f(x) = - \int_0^t f(x)$$

$$f(x) = - \frac{dS(t)}{dt}$$

Por lo tanto, la función de riesgo se podría reorganizar:

$$\lambda(t) = \frac{f(t)}{S(t)} = \frac{- \frac{dS(t)}{dt}}{S(t)}$$

Al resolver esta ecuación se obtiene:

$$\lambda(t) = - \frac{d}{dt} \log S(t)$$

Si se integra de 0 a t y se introduce la condición de frontera $S(0) = 1$, se puede resolver la expresión anterior para obtener una fórmula que permita calcular la probabilidad de supervivencia en función del riesgo dentro del periodo de estudio t.

Cabe mencionar que esta sección de la ecuación representa el riesgo acumulado, $\Lambda(t)$, que es la suma de las tasas de riesgo desde el periodo 0 hasta el periodo t. Dada la función de riesgo, siempre se puede integrar ésta para obtener el riesgo acumulado y luego incorporarla en una función exponencial para obtener la función de supervivencia.

$$S(t) = \exp \left\{ - \int_0^t \lambda(x) dx \right\}$$

$$\Lambda(t) = \int_0^t \lambda(x) dx$$

La distribución de supervivencia más simple, se obtiene asumiendo un riesgo constante en el tiempo:

$$\lambda(t) = \lambda$$

La correspondiente función de supervivencia será:

$$S(t) = \exp\{-\lambda t\}$$

Esta distribución se llama distribución exponencial con parámetro λ . La densidad se puede obtener multiplicando la función de superviviente por el riesgo, de esta manera se obtiene:

$$f(t) = \lambda \exp\{-\lambda t\}$$

Por lo tanto, la media resultante es $1 / \lambda$.

Modelo de Kaplan-Meier

El estimador de Kaplan-Meier (KM) estima la función de supervivencia en el tiempo t. Goel, Khanna, & Kishore (2010) exponen tres supuestos que se deben considerar en el modelo:

- 1) Las observaciones censuradas tienen las mismas perspectivas de supervivencia que el resto de individuos no censurados.

- 2) Las probabilidades de supervivencia son las mismas para los sujetos reclutados al principio y al final del estudio.
- 3) El evento ocurre en un momento específico.

La precisión del modelo de Kaplan-Meier dependerá de la longitud de los intervalos de tiempo, pues con intervalos más cortos existe una mayor precisión. Mills (2011), indica que la función que permite calcular la probabilidad de supervivencia en el tiempo de falla t_j viene dado por:

$$S(t_j) = S(t_{j-1}) \times \Pr(T > t_j | T \geq t_j)$$

Donde t_j representa el tiempo actual y t_{j-1} el tiempo anterior.

La probabilidad de supervivencia de KM puede describirse:

$$S(t_{j-1}) = \prod_{i=1}^{j-1} \Pr(T > t_i | T \geq t_i)$$

Esta función representa la probabilidad acumulada de supervivencia. El factor distintivo de este modelo es que KM ajusta el tiempo de supervivencia estimado para tener en cuenta las observaciones censuradas por la derecha (Mills, 2011). Sin embargo, una limitación del modelo KM es que solo utiliza variables categóricas para su análisis y representación gráfica.

Por último, es necesario mencionar que el modelo KM permite realizar una comparación entre las curvas de supervivencia entre grupos y comprobar si éstas son estadísticamente idénticas a través de la prueba Log-Rank (LaMorte, 2016):

H0: Las curvas de supervivencia de los grupos son idénticas ($S_{1t} = S_{2t}$).

Ha: Las curvas de supervivencia de los grupos no son idénticas ($S_{1t} \neq S_{2t}$, en cualquier tiempo t).

Modelo de Riesgos Proporcionales de Cox

Mills (2011) señala que Cox (1972) introdujo una serie de modelos que se centran estrictamente en la función de riesgo. Para utilizar esta técnica es necesario tener en consideración tres supuestos principales:

1. Independencia de los tiempos de supervivencia entre los distintos individuos de la muestra.
2. Relación multiplicativa entre los predictores y el riesgo.
3. Una tasa de riesgo constante a lo largo del tiempo.

El modelo más simple es

$$\lambda_i(t|x_i) = \lambda_0(t)\exp\{x_i'\beta\}$$

Este modelo mide el riesgo de un individuo en el tiempo t con covariables x_i , sin incluir la constante. $\lambda_0(t)$ es una función de riesgo de referencia que describe el riesgo para los individuos con $x_i = 0$, mientras que $\exp\{x_i'\beta\}$ es el riesgo relativo, que se traducen en un aumento o reducción proporcional del riesgo asociado al conjunto de características x_i . Se asume que el efecto de las covariables x es el mismo en todos los tiempos t (Mills, 2011).

La tasa de riesgo (H) es la razón entre de la tasa de riesgo en el momento t y el riesgo de referencia.

$$H = \ln \left(\frac{\lambda_i(t|x_i)}{\lambda_0(t)} \right) = \{x_i' \beta\}$$

$$H = B_1 X_1 \dots B_k X_k$$

Como se indicó anteriormente, la función de supervivencia puede ser obtenida a través de la función de riesgos acumulados:

$$S_i(t|x_i) = S_0(t) \exp\{x_i' \beta\}$$

Donde $S_0(t) = \exp\{-\Lambda_0(t)\}$ representa la función de supervivencia inicial. Por tanto, el efecto de la covariable x_i aumenta el riesgo relativo $\exp\{x_i' \beta\}$ de la función de supervivencia.

Partiendo del modelo de riesgos proporcionales de Cox, se desprende una serie de modelos con diferentes modificaciones. Entre estos modelos se encuentra el modelo estratificado de Cox y el modelo de covariables de tiempo dependiente que se utilizan cuando se viola el supuesto de proporcionalidad.

Prueba de Schoenfeld

Para evaluar el cumplimiento de la condición de riesgos proporcionales en el modelo de Cox existen dos métodos diferentes. Uno de ellos consiste en analizar las curvas de Kaplan-Meier, en donde el supuesto se incumple si las categorías de las diferentes variables se entrecruzan o si cada una de las curvas crecen en proporciones diferentes a lo largo del periodo evaluado (Simon, 2019). El segundo método es el test de residuos de Schoenfeld. Esta prueba parte del supuesto de que el tiempo es un factor que afecta al modelo, ya sea en los coeficientes o en las covariables (van Houwelingen & Stijnen, 2014).

Esta prueba se puede utilizar para analizar la variabilidad de $\beta(t)$ y la validez del modelo con $\beta_j(t) \equiv \beta_j$ para cada covariable X_j . Para el conjunto de riesgo $R(t)$, el residual de Schoenfeld se define como:

$$score(t) = x_{ev}(t) - \bar{x}(\hat{B}, t)$$

$$x_{ev}(t) = \frac{\sum_i x_i * \Delta N_i(t)}{\Delta \bar{N}(t)}$$

$x_{ev}(t)$: es la media del valor de x de los individuos a los que les ha ocurrido el evento en el tiempo t (van Houwelingen & Stijnen, 2014).

Si el modelo de riesgos proporcionales es cierto, el valor esperado del $score(t)$ es igual a cero. La validez del modelo para la j -ésima covariable se puede verificar trazando el j -ésimo componente del $score(t)$ versus t . El valor medio del residual es igual a cero por construcción. Las tendencias temporales de los residuos indican una violación del supuesto de riesgos proporcionales (van Houwelingen & Stijnen, 2014). Este es el método gráfico que permite visualmente analizar si se cumple o no el supuesto de proporcionalidad. También, existen pruebas formales en donde se rechaza o no la hipótesis planteada (Park & Hendry, 2015):

H0: El modelo cumple con el supuesto de riesgos proporcionales.

Ha: El modelo no cumple con el supuesto de riesgos proporcionales.

Cabe mencionar que esta prueba analiza únicamente aquellas observaciones a las que les ha ocurrido el evento.

Modelos Estratificados

Cuando el supuesto de proporcionalidad no se cumple, el modelo de Cox puede extenderse a un modelo de Cox estratificado, considerando una variable de estratificación categórica, a la que se llamará G. El modelo estratificado permite que el riesgo de referencia dependa del estrato:

$$h(t|X, G) = h_{G0}(t)\exp\{x_i'\beta\}$$

El modelo es similar al modelo de riesgos proporcionales de Cox, la diferencia es que ahora los conjuntos de riesgos son específicos por estrato. La principal aplicación de este modelo es incluir covariables categóricas para las que no se cumplan el supuesto de riesgo proporcional. El supuesto principal es que el efecto de las covariables de X, sea el mismo en cada estrato. El modelo solo puede ser confiable si cada estrato tiene suficientes eventos para obtener una estimación estable del riesgo de referencia. Se permite que el efecto de una covariable sea diferente entre estratos cuando se agrega una interacción entre una covariable y la variable de estratificación (van Houwelingen & Stijnen, 2014).

Covariables de tiempo dependiente

El modelo de riesgos proporcionales de Cox considera explícitamente solo las covariables que se fijan en el tiempo, sin embargo, la naturaleza de este modelo de riesgos permite que las covariables cambian con el tiempo a partir de ciertas consideraciones. Este modelo toma en cuenta variables que pueden cambiar en el tiempo, tales como edad, tamaño, ingreso familiar, etc. (Mills, 2011).

Una de las principales modificaciones con respecto al modelo original de Cox es que las covariables están en función del tiempo, por lo tanto $x_i(t)$ representará el valor de un vector de covariables para el individuo i en el intervalo de tiempo t. Por lo tanto, el modelo de covariables de tiempo dependiente puede generalizarse a (Mills, 2011):

$$\lambda_i(t, x_i(t)) = \lambda_0(t) \exp\{x_i(t)'\beta\}$$

El cálculo de las funciones de supervivencia cuando existen covariables que varían en el tiempo es complicado, ya que es necesario especificar una ruta o trayectoria para cada variable y no se puede simplemente elevar la función de supervivencia básica a un exponente. Por lo general, se calcula únicamente el riesgo que cada una de las covariables produce (Mills, 2011).

Prueba de Residuos de Cox-Snell

Este análisis permite evaluar el ajuste global del modelo planteado. Al estimar el modelo de Cox, se obtiene la función de supervivencia $\hat{S}(t_i)$, la cual debería ser similar al verdadero valor de $S(t_i)$. Para obtener la diferencia entre la función estimada y la verdadera, se calculan los residuos Cox-Snell:

$$r_i = -\ln\hat{S}(t_i, X_i) = -\ln\left[\hat{S}_0(t)^{\{\exp(\hat{\beta}'X_i)\}}\right] = \{\exp(\hat{\beta}'X_i)\}[-\ln\hat{S}_0(t)]$$

Si el modelo apropiado se ajusta bien a los datos, entonces r_i tendrá para cada i un valor $\exp(1)$, es decir una distribución exponencial con razón o tasa de riesgo igual a 1. Al construir el gráfico de residuos, éste debería concordar con una recta de pendiente positiva igual a 1, esto significa una línea de 45°. De esta manera se comprueba si el modelo tiene un buen ajuste (Berrenechea López, 2018).

VARIABLES DEL MODELO

Una vez obtenida la información de las diferentes fuentes se tomaron en cuenta algunas consideraciones para seleccionar la temporalidad y la población de estudio. La temporalidad del análisis estará limitada por el período 2014-2019, en donde se utilizará la información desde el semestre 2014-02³ hasta el semestre 2019-01. La razón por la que se ha elegido este periodo de estudio es porque se cuenta con información completa que proviene de un mismo sistema. La información más actualizada pertenece a un sistema nuevo y no contiene todos los campos que se tenía hasta 2019-01, razón por la cual se descartó el periodo 2019-02. Una vez estipulada la temporalidad, se determinó que las cohortes de análisis serían desde 2014-02 hasta 2017-02, ya que es necesario observar el comportamiento de los estudiantes por lo menos durante cuatro semestres. Esta consideración es necesaria ya que la deserción ocurre con mayor frecuencia durante los primeros niveles.

En cuanto a la selección de la población de estudio, se determinó que se analizaría a todos los estudiantes de pregrado que se inscribieron por primera vez en un programa presencial de la PUCE sede Quito. Sin embargo, existen algunos grupos que presentaron características diferentes y que podrían causar errores en la investigación, razón por la cual se optó por retirarlos del estudio. Este fue el caso de los estudiantes de la Facultad de Ciencias Filosóficas y Teológicas y los de la Facultad de Medicina. La Facultad de Ciencias Filosóficas y Teológicas tenía gran ausencia de información y un plan de estudios más largo en el caso de Teología. Por otro lado, la Facultad de Medicina presentaba varias diferencias con respecto al resto, por ejemplo: calificaciones, número de semestres que deben cursar los estudiantes en un mismo año, créditos por semestres, número de materias, prácticas pre profesionales, entre otras.

Al identificar a los desertores se encontraron tres casos específicos. En el primer caso, se los categorizó como desertores a todos aquellos estudiantes que se retiraron por al menos un semestre. En el segundo, se los categorizó como desertores temporales a aquellos estudiantes que a pesar de haberse ausentado 2 períodos o más, regresaron a la universidad. El tercer caso, constituyó en aquellos estudiantes que nunca volvieron a la Universidad durante el período de análisis, a los estudiantes de este grupo se los nombró como desertores definitivos. Al contabilizar los casos, se encontró que el primer caso abarcaba a todos; en el segundo, los desertores temporales eran 79 personas, lo que representa menos del 1% de la población estudiada. Las características y comportamientos de estos grupos eran diferentes, por lo tanto, en esta investigación únicamente se trabaja con el grupo de desertores definitivos.

Una vez seleccionadas las características de la población se optó por crear dos tipos de bases. Para estudios transversales se utilizará la base de datos en donde cada alumno es una observación. Esta base contiene 6.127 observaciones y 37 variables. Para estudios longitudinales se utilizará la base de datos en donde los 6.127 alumnos aparecen repetidas veces de acuerdo al número de semestres en los que estuvieron inscritos. Esta base contiene 34.955 observaciones y 42 variables. La segunda base contiene un mayor número de variables ya que se adicionaron las variables que cambian en el tiempo. A continuación, se presenta una tabla que describe de manera general las variables que serán utilizadas:

³ Segundo semestre del año 2014. Los primeros periodos (-01) están comprendidos entre febrero y julio, mientras que los segundos periodos (-02) cubren los meses de agosto a enero. No existe una fecha exacta de entrada y de salida en cada uno de los semestres, por lo tanto, estos meses son simplemente una referencia.

Tabla 2: Tabla de variables

	Variables	Explicación de la Variable
id	Identificación del estudiante	Cédula de identidad del Estudiante
Variables explicativas	Características individuales de los estudiantes	
	Sexo	Sexo del estudiante.
	Edad del estudiante	La edad del estudiante al entrar a la Universidad.
	Estado civil del estudiante	El estado civil del estudiante al entrar a la Universidad.
	Residencia previa	Residencia previa a la entrada a la Universidad (Pichincha u otro lugar).
	Condiciones iniciales al entrar a la Universidad	
	Cohorte de entrada	Cohorte en la que el estudiante entró por primera vez a la Universidad.
	Periodo de entrada	Indica el periodo en el que el estudiante entró en el primer semestre. En el periodo 01 de febrero-julio o en el 02 de agosto-enero.
	Tipo de entrada	La forma en la que el estudiante entró a la Universidad.
	Puntaje de la Prueba de Admisión	El Puntaje obtenido en la Prueba de Admisión.
	Duración Ideal de la carrera	Cantidad de semestres que debería durar la carrera en la que se inscribió por primera vez el estudiante y que estudió durante el primer semestre.
	Número total de créditos de la carrera	Los créditos obligatorios que debería tomar el estudiante de acuerdo a la carrera en la que se inscribió y que estudió durante primer semestre.
	Facultad	Facultad a la que pertenece el estudiante en primer semestre.
	Factores Socioeconómicos y familiares	
	Descuento	El descuento total o promedio recibido por el estudiante en el pago de su matrícula durante sus estudios.
	Categoría Socioeconómica de la Pensión Diferenciada (SPD)	Categoría Socioeconómica de la Pensión Diferenciada, en la que fue colocado el alumno cuando ingresó a la Universidad.
	Cambio de la Categoría SPD	Variable binaria que indica si ha cambiado de categoría socioeconómica.
	Tipo de Colegio	El tipo de colegio del que el alumno se graduó.
	Nivel de educación de la Madre	El nivel máximo de educación alcanzado por la madre del estudiante.
	Estado Civil de la Madre	El estado civil de la madre del estudiante.
	Nivel de Educación del Padre	El nivel máximo de educación alcanzado por el padre del estudiante.
	Estado Civil del Padre	El estado civil del padre del estudiante.
	Factores de Desempeño Académico	
	Cambio de Facultad	Indica si el estudiante se ha cambiado al menos una vez de Facultad.
	Cambio de Carrera	Indica si el estudiante se ha cambiado al menos una vez de carrera.
	Créditos tomados	El promedio o el total de créditos tomados por el estudiante desde su primer semestre hasta el nivel máximo alcanzado
	Promedio de notas	Promedio general de las notas finales de cada una de las materias por semestre o del todas las materias cursadas.
	Materias perdidas	El promedio o el total acumulado de las materias perdidas.
Variable dependiente	Desertor	Estudiantes que abandonaron definitivamente sus estudios y que no volvieron a inscribirse en la Universidad durante todo el período de análisis.
	Tiempo	El número de semestres en el que el estudiante ha estado inscrito.

Elaborado por: Autor.

Nota: Para revisar los nombres de las variables de la base de datos revisar el Anexo A

CAPÍTULO I: PERFIL DE LOS ESTUDIANTES

A continuación, se presenta una visión general de la estructura de los datos y el comportamiento de los estudiantes, a través de un análisis de estadística descriptiva y modelos no paramétricos de Kaplan-Meier. El objetivo de este capítulo es construir el perfil de los estudiantes desertores a través de las características reveladas por estas metodologías.

La primera parte de este capítulo describe la composición de las bases de datos en función de los grandes grupos de variables que se han creado. Cabe mencionar que algunas de las variables han sido discretizadas y categorizadas para generar un mejor análisis. La segunda parte del capítulo se compone del análisis del modelo de supervivencia de Kaplan-Meier para cada una de las variables. Por último, se presenta un resumen con todas aquellas variables que tienen un efecto significativo en la deserción, con el fin de describir un potencial perfil de los desertores.

Descripción de las Variables

Tabla 3: Tabla Resumen de la variable dependiente

Variabes	Categorías / Medidas de posición	Número de Observaciones	Porcentaje
Desertor	0: Se ha inscrito en todos los semestres desde que inició la carrera	4542	74,1%
	1: Ha abandonado la universidad por al menos un semestre	1585	25,9%
Desertor Definitivo	0: No ha desertado definitivamente	5083	83,0%
	1: Desertó y no volvió a la universidad	1044	17,0%

Elaborado por: Autor

Al observar la Tabla 3 se puede ver que el 25,9% de los estudiantes que entraron entre agosto 2014 y agosto 2017 abandonaron la PUCE sede Quito por al menos un período y el 65,6% de este grupo desertó de manera definitiva. Los alumnos que abandonaron sus estudios de manera permanente representan el 17% de la población de estudio.

Tabla 4: Tabla resumen de las variables explicativas relacionadas con las características individuales de los estudiantes

Características individuales de los estudiantes				
VARIABLES	Categorías / Medidas de posición		Número de Observaciones	Porcentaje
Sexo	0: Hombre		2547	41,6%
	1: Mujer		3580	58,4%
Edad del estudiante	Mín: 17	Máx: 63		
	1er Cuartil: 18	3er Cuartil: 20		
	Mediana: 19	Media: 19,38		
	Valores ausentes: 8			
Edad del estudiante	17-18		2694	44,0%
	19-20		2413	39,4%
	Mayores de 20		1020	16,7%
Estado civil del estudiante	Otro		106	1,7%
	Soltero		6021	98,3%
Residencia previa	Desconocido		677	11,1%
	Otro lugar		834	13,6%
	Pichincha		4616	75,3%

Elaborado por: Autor

En cuanto a las características individuales de los estudiantes, se puede observar que el 58,4% del total de estudiantes son mujeres. La mínima edad con la que un estudiante ha entrado a la universidad es de 17 años y la máxima de 63 años. Esta variable se categorizó tomando en consideración estudios previos, tales como la investigación de Richardson y Woodley (2003) y Sheard (2009). En la variable categórica de edad se pudo observar que el 44% de estudiantes tienen entre 17 y 18 años cuando entra a la universidad y el 39,4% tienen entre 19 y 20 años; lo que significa que apenas el 16,7% son mayores a 20 años al entrar a la Universidad. Por otro lado, se puede ver que el 98,3% de los estudiantes son solteros al entrar a la Universidad. Por último, se observa que el 75,3% de los estudiantes residía en Pichincha previo a iniciar la universidad, el 13,6% provenían del extranjero u otras provincias y el 11,1% se desconoce.

Tabla 5: Tabla resumen de las variables explicativas relacionadas con las condiciones iniciales en la Universidad

Condiciones iniciales en la Universidad				
Variab les	Categorías / Medidas de posición		Número de Observaciones	Porcentaje
Cohorte de entrada	2014-02		1086	17,7%
	2015-01		443	7,2%
	2015-02		917	15,0%
	2016-01		551	9,0%
	2016-02		1243	20,3%
	2017-01		650	10,6%
	2017-02		1237	20,2%
Periodo de entrada	01: representa el semestre de febrero a julio		1651	27,0%
	02: representa el semestre de agosto a enero		4476	73,1%
Tipo de entrada	Desconocido		1167	19,1%
	Examen de Admisión		4484	73,2%
	Otro Tipo de Solicitud		476	7,8%
Puntaje de la Prueba de Admisión	Mín: 12	Máx: 70		
	1er Cuartil: 36	3er Cuartil: 47		
	Mediana: 42	Media: 41,95		
	Valores ausentes: 1407			
Puntaje de la Prueba de Admisión	0-36		1244	20,3%
	36-47		2385	38,9%
	47-70		1091	17,8%
	Desconocido		1407	23,0%
Duración ideal de la carrera	8		3132	51,1%
	9		1418	23,1%
	10		1577	25,7%
Número ideal de créditos por carrera	Mín: 153	Máx: 273		
	1er Cuartil: 188	3er Cuartil: 205		
	Mediana: 196	Media: 200,2		
	Valores ausentes: 0			
Facultad	Escuela de Bioanálisis		231	3,8%
	Escuela de Trabajo Social		41	0,7%
	Facultad de Arquitectura, Diseño y Artes		645	10,5%
	Facultad de Ciencias Administrativas y Contables		609	9,9%
	Facultad de Ciencias de la Educación		43	0,7%
	Facultad de Ciencias Exactas y Naturales		153	2,5%
	Facultad de Ciencias Humanas		539	8,8%
	Facultad de Comunicación, Lingüística y Literatura		474	7,7%
	Facultad de Economía		341	5,6%
	Facultad de Enfermería		928	15,2%
	Facultad de Ingeniería		748	12,2%
	Facultad de Jurisprudencia		538	8,8%
Facultad de Psicología		837	13,7%	

Elaborado por: Autor.

Al evaluar las condiciones iniciales se puede observar que a partir de 2016-01, la Universidad ha tenido un aumento progresivo de los alumnos que entran y por lo general un mayor número de alumnos entran en el segundo periodo. El 73% de los alumnos entran en el periodo 02 y tan solo el 27% entran en el periodo 01. Esta situación va ligada al calendario académico y graduación de los alumnos de colegios de la región Sierra, que es la misma zona en la que se encuentra ubicada la PUCE Sede Quito.

El 73,2% de los estudiantes entraron a la Universidad a través del Examen de Admisión⁴, el 7,8% entran por otros medios⁵; y, debido a limitaciones de las bases de datos de la Universidad, se desconoce la manera de entrada del 19,1% de los alumnos. De los 4.855⁶ alumnos que rindieron el examen, solo se tiene el puntaje de 4.720, el resto de estudiantes tiene información desconocida. La variable Puntaje de la Prueba de Admisión se categorizó de acuerdo a los cuartiles de toda la población. Una vez categorizada esta variable, se pudo observar que del total de alumnos que entran a la universidad el 20,3% obtuvo una nota baja en la Prueba de Admisión, que varía entre 0 y 36 punto; el 38,9% obtuvo una nota regular que varía entre 36 y 47 puntos y el 17,8% obtuvo una nota alta que varía entre 47 y 70 puntos. El 23% de los puntajes se desconoce. Cabe mencionar que la prueba es sobre 70, por lo tanto, esta ha sido la máxima nota alcanzada y la mínima es de 12. En general la media y mediana de los puntajes obtenidos es de 42.

Al analizar la duración ideal de la carrera, se puede observar que el 51,1% de estudiantes siguen una carrera de 8 semestres. En varios casos esta decisión no es tomada por el estudiante, sino que se debe a la oferta de la Universidad, que es la encargada de asignar la duración de las carreras. La variación de créditos obligatorios entre carreras tiene una dispersión grande, pues el mínimo de créditos de una carrera es de 154 y el máximo de 273. La media de créditos que los estudiantes deben tomar es de 200. Para categorizar esta variable se tomó en cuenta la distribución de las observaciones y las medidas de posición.

En cuanto a la variable Facultad se puede observar que las Facultades⁷ con mayor número de estudiantes son la Facultad de Enfermería con 15,2%, la Facultad de Psicología con el 13,7% del alumnado y la Facultad de Ingeniería con 12,2%. Las Facultades con menor número de estudiantes son: la Facultad de Ciencias de la Educación con un 0,7%, la Escuela de Trabajo Social con 0,7% y la Facultad de Ciencias Exactas y Naturales con el 2,5% del total de estudiantes.

⁴ Los estudiantes de Psicología deben tomar un examen diferente, sin embargo, en la investigación no se hará ninguna consideración especial ni distinción.

⁵ Las otras formas de entrar son: cambio de sede, cambio de universidad, cambio de universidad psicología, inicio de carrera psicología, curso de nivelación, estudiantes especiales sin convenio, inicio carrera PUCE-SENESCYT e inicio con título universitario.

⁶ Algunos estudiantes que entraron de otra manera también tienen un puntaje de la Prueba de Admisión ya que en algunos casos dieron la misma prueba o una similar.

⁷ Se ha retirado a la Facultad de Medicina y a la Facultad de Ciencias Teológicas y Filosóficas.

Tabla 6: Tabla resumen de las variables explicativas relacionadas con los factores socioeconómicos y familiares

Factores Socioeconómicos y Familiares				
VARIABLES	Categorías / Medidas de posición		Número de Observaciones	Porcentaje
Descuento Acumulado	Mín: 0	Máx: 39161		
	1er Cuartil: 2915	3er Cuartil: 9025		
	Mediana: 5119	Media: 6575		
	Valores ausentes: 0			
Descuento promedio por semestre	Mín: 0	Máx: 7768,2		
	1er Cuartil: 734,5	3er Cuartil: 1472		
	Mediana: 1068	Media: 1174,7		
	Valores ausentes: 0			
Descuento promedio por semestre	0-800 USD\$		1748	28,5%
	800-1600 USD\$		3138	51,2%
	1600-3500 USD\$		1130	18,4%
	3500-8000 USD\$		111	1,8%
Categoría Socioeconómica de la Pensión Diferenciada	A		946	15,4%
	B		1719	28,1%
	C		1263	20,6%
	D		718	11,7%
	E		498	8,1%
	F		983	16,0%
Cambio de la Categoría Socioeconómica	0: No ha cambiado		5252	85,7%
	1: Sí ha cambiado		875	14,3%
Tipo de Colegio	Desconocido		106	1,7%
	Extranjero		21	0,3%
	Privado		3861	63,0%
	Público		1266	20,7%
	Semi-Público		873	14,3%
Nivel de educación de la Madre	Desconocido		514	8,4%
	Primer		539	8,8%
	Segundo		2610	42,6%
	Tercer		2464	40,2%
Estado Civil de la Madre	Casada		3720	60,7%
	Desconocido		701	11,4%
	Divorciada		1009	16,5%
	Soltera		527	8,6%
	Viuda		170	2,8%
Nivel de educación del Padre	Desconocido		692	11,3%
	Primer		509	8,3%
	Segundo		2088	34,1%
	Tercer		2838	46,3%
Estado Civil del Padre	Casado		4064	66,3%
	Desconocido		919	15,0%
	Divorciado		889	14,5%
	Soltero		191	3,1%
	Viudo		64	1,0%

Elaborado por: Autor.

En cuanto a los factores socioeconómicos y familiares, existen algunas variables que conforman este grupo. Al analizar el descuento que recibe el estudiante se tomaron en consideración diferentes formas de presentar la variable. Al tomar en cuenta el descuento acumulado se observa que el mínimo de

descuento que puede recibir un estudiante es \$0 y lo máximo que ha recibido es de \$30.161. En general la media del descuento total recibido por los alumnos es de \$6.575. Al analizar el descuento promedio que recibe el estudiante en cada semestre se observa que el mínimo es igualmente de \$0 y el máximo es de \$7.768,20. El descuento promedio que reciben los estudiantes por semestre es de \$1.174,70. Al categorizar esta variable, se puede ver que el 51,2% de estudiantes reciben un descuento entre \$800 y \$1.600. Los descuentos que los estudiantes reciben pueden deberse a la categoría socioeconómica en la que fue colocado el estudiante y/o becas recibidas. Para categorizar esta variable se analizó la distribución de los datos del descuento promedio por semestre, se conversó con una persona que trabaja en la DGE y se realizaron diferentes pruebas.

En cuanto a la categoría socioeconómica de la pensión diferenciada, aproximadamente el 49% de los estudiantes están concentrados en categoría B o C. Por otro lado, en la categoría más baja, la A, se concentra el 15,4% y en la más alta, la F, el 16,6%. Al analizar esta variable en el tiempo, se puede ver que el 14,3% de los estudiantes reciben un cambio de categoría. Es preciso mencionar que la mayoría de estudiantes que reciben el cambio, son colocados en una categoría inferior.

En cuanto al tipo de colegio, se observa que el 63% de los estudiantes provenían de un colegio privado, el 20,7% de un colegio público, el 14,3% de un colegio semi-público que puede ser un colegio municipal o fiscomisional, el 0,3% de un colegio extranjero y el 1,7% es desconocido.

Al analizar las características familiares, se puede observar que el 40,2% de los estudiantes tienen madres que han alcanzado un tercer nivel o superior⁸ de educación y el 46,3% tienen padres que han alcanzado el tercer nivel o superior. De igual manera, en cuanto al estado civil de la madre se observa que el 60,7% de los estudiantes tienen madres casadas⁹ y el 16,5% tiene madres divorciadas. Por otro lado, al observar el estado civil del padre, se puede ver que el 66,3% de los estudiantes tienen padres casados, el 14,5% divorciados y el 15% es desconocido.

⁸ En la categoría de tercer nivel se incluye a aquellas madres y padres que han alcanzado también el cuarto nivel.

⁹ En la categoría de casados se incluye unión libre.

Tabla 7: Tabla resumen de las variables explicativas relacionadas con los factores de desempeño académico.

Factores de Desempeño Académico			
Variab les	Categorías / Medidas de posición	Número de Observaciones	Porcentaje
Cambio de Facultad	0: No se cambió	5809	94,8%
	1: Sí se cambió	318	5,2%
Cambio de Carrera	0: No se cambió	5470	89,3%
	1: Sí se cambió	657	10,7%
Total de créditos tomados	Mín: 3	Máx: 304	
	1er Cuartil: 98	3er Cuartil: 197	
	Mediana: 142	Media: 141,8	
	Valores ausentes: 0		
Promedio de créditos por semestre	Mín: 3	Máx: 39	
	1er Cuartil: 22,5	3er Cuartil: 27,2	
	Mediana: 25,17	Media: 24,72	
	Valores ausentes: 0		
Promedio de créditos por semestre	Demasiados: [30-40]	275	4,5%
	Normal: [19-30)	5418	88,4%
	Pocos: [1-19)	434	7,1%
Promedio General de Notas	Mín: 4,4	Máx: 50	
	1er Cuartil: 34,45	3er Cuartil: 40,7	
	Mediana: 37,77	Media: 37,04	
	Valores ausentes: 0		
Promedio General de Notas	Inferior al mínimo: [0-29,5)	518	8,5%
	Promedio Bajo: [29,5-36)	1663	27,1%
	Promedio Bueno: [36-44)	3583	58,5%
	Promedio Excelente: [44-50]	363	5,9%
Total de Materias perdidas	Mín: 0	Máx: 29	
	1er Cuartil: 0	3er Cuartil: 4	
	Mediana: 1	Media: 2,67	
	Valores ausentes: 0		
Total de Materias perdidas	0: Ninguna materia perdida	2269	37,0%
	1: Ha perdido una materia	922	15,1%
	2: Ha perdido dos materias	669	10,9%
	Múltiples: Perdió más de 2	2267	37,0%

Elaborado por: Autor.

Al evaluar los factores de desempeño se observa que el 5,2% del total de estudiantes se cambian de Facultad y el 10,7% de carrera, lo que indica que el 5,5% de los estudiantes se cambian de carrera, pero no de Facultad. Estas dos variables tienen una correlación muy alta, por ende, para los siguientes análisis únicamente se incluirá la variable cambio de carrera.

Al analizar el total de créditos tomados por un estudiante, se puede observar que el mínimo de créditos tomados es de 3 y el máximo es de 304, que pudo darse si el estudiante se cambió de carrera o si perdió varias materias. En cuanto al promedio de créditos tomados por semestre se observa que el máximo de

créditos que un estudiante ha tomado es de 36 créditos y el mínimo de 3 créditos, mientras que la media es de 25 créditos. Tomando esto en consideración las mallas de las diferentes materias, se categorizó la variable de acuerdo a la comparación de mallas de distintas carreras y a la distribución de los datos, lo que revela que el 7,1% de los estudiantes toman pocos créditos, esto significa entre 1 y 19 créditos. El 88,4% toma un número regular de créditos, entre 19 y 30 créditos y solo el 4,5% toman entre 30 y 40 créditos.

En cuanto a las calificaciones, se observa que el promedio general más bajo es de 4,4, mientras que el promedio de notas más alto es de 50, que es la nota máxima que un estudiante puede obtener y la media es 37,04. Para categorizar esta variable se tomó en cuenta las notas más altas, más bajas y las medias de cada Facultad, también se consideró las medidas de posición. Una vez categorizada la variable se obtuvo que el 27,1% de los estudiantes tienen un promedio bajo que oscila entre 29,5 y 36 puntos: mientras que el 58,5% tiene un promedio entre 36 y 44 puntos. Apenas el 5,9% de los estudiantes alcanzan un promedio de 44 puntos o superior.

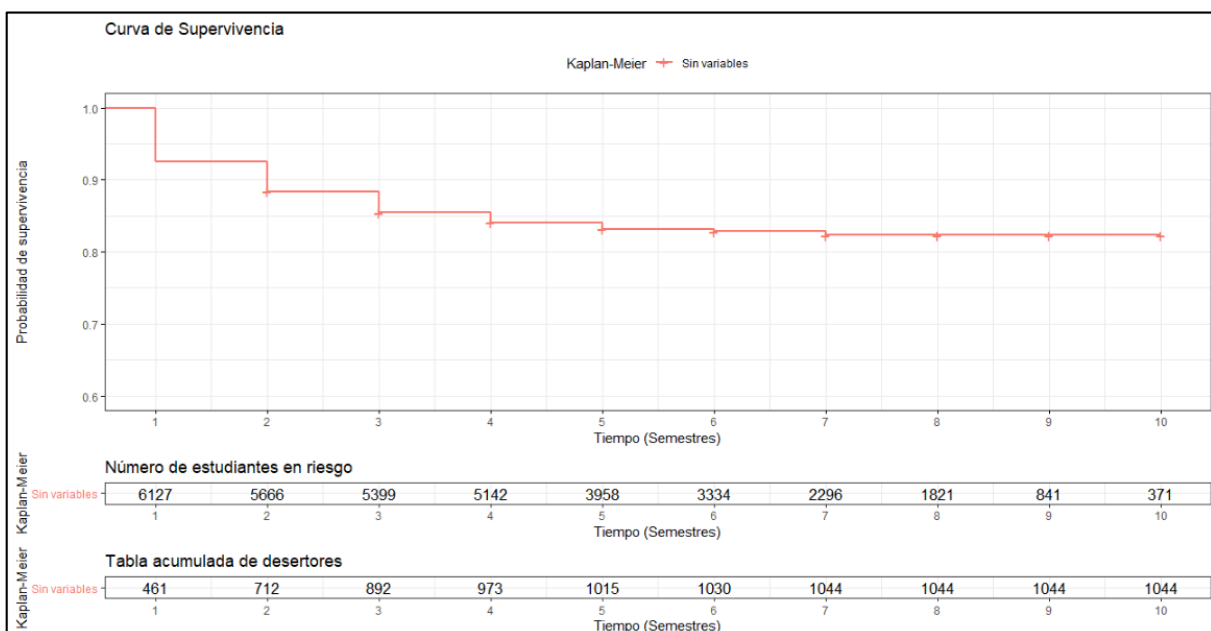
Por último, en cuanto a las materias perdidas, se observa que el máximo de materias que un estudiante ha llegado a perder durante sus estudios es de 29 materias. Para categorizar esta variable se tomó en cuenta la distribución de los datos. La variable categórica de materias perdidas muestra que el 37% de los estudiantes nunca han perdido una materia; el 15,1% han perdido una materia, el 10,9% han perdido 2 materias y 37% han perdido más de dos materias.

Modelos de Supervivencia de Kaplan-Meier

Desertor

Desertor es la variable dependiente que se pretende estudiar y modelar. Antes de evaluar la interacción con el resto de variables es necesario entender su comportamiento.

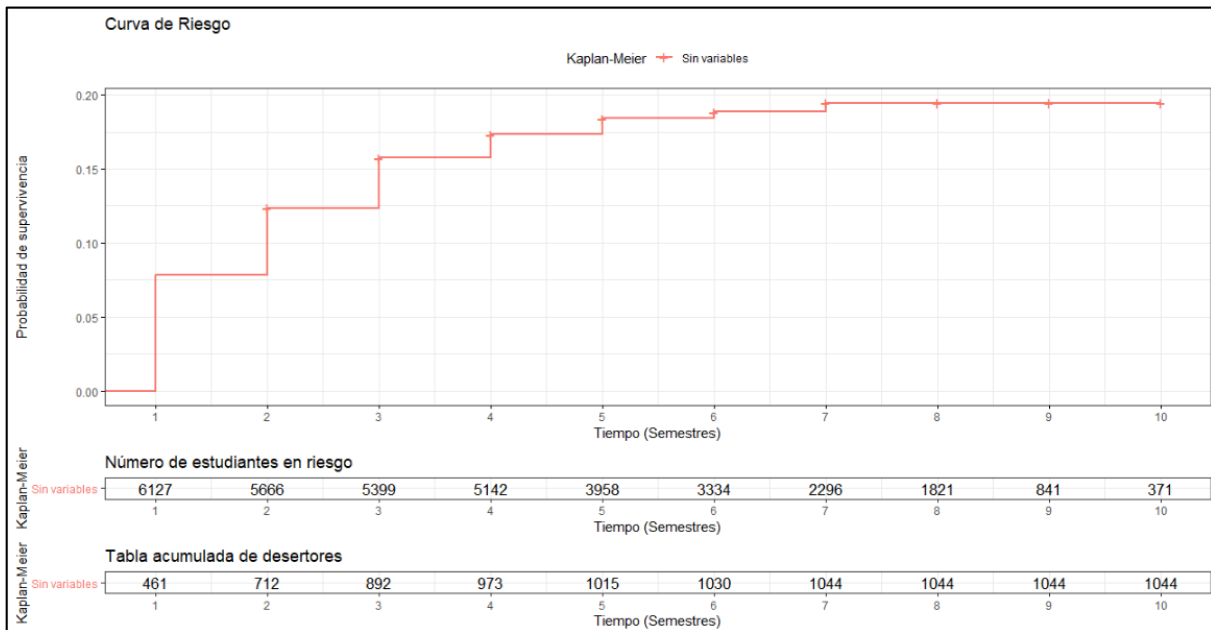
Gráfico 1: Curva de supervivencia basado únicamente en el número de desertores



Elaborado por: Autor.

Al analizar este Gráfico 1, se observa que la deserción ocurre durante los primeros niveles. Según la tabla acumulada de desertores, se observa que después del primer semestre 44% de los estudiantes identificados como desertores abandonan la PUCE sede Quito, después del segundo semestre el 68% y hasta el tercer semestre el 85% ya han abandonado la Universidad. A partir del cuarto semestre la probabilidad de que un estudiante deserte cae considerablemente, y a partir del séptimo semestre es nula.

Gráfico 2: Curva de riesgo basado en el número de desertores sin efecto en otras variables

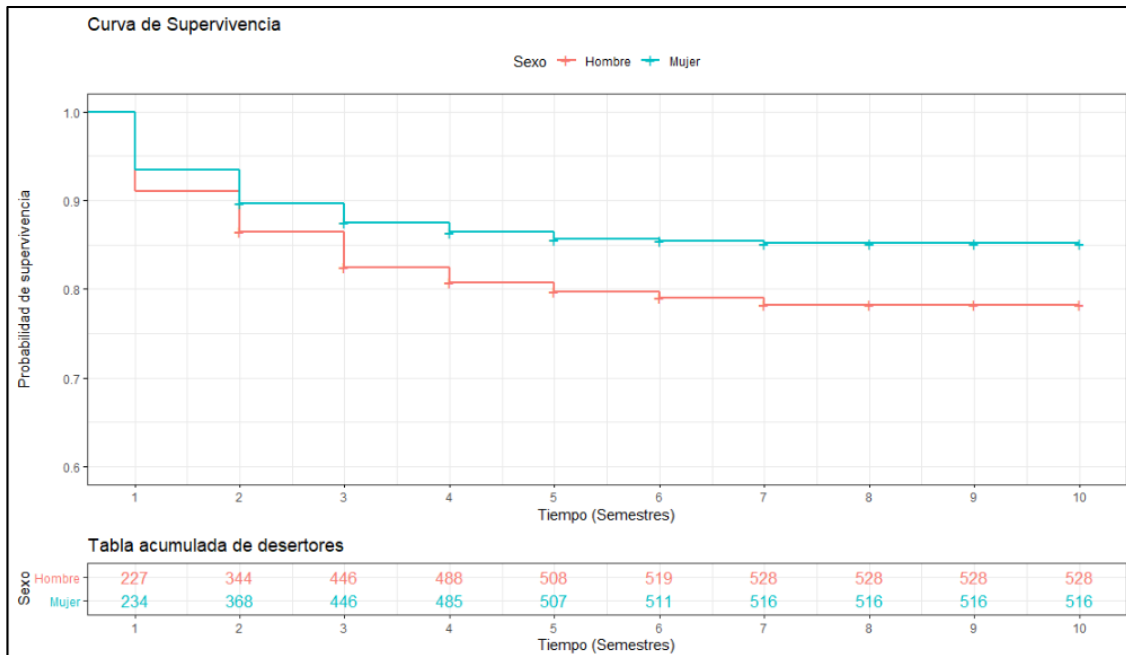


Elaborado por: Autor

La curva de riesgo acumulado y curvas de supervivencia se construyen en base a la misma información, sin embargo, la curva de riesgos es la representación contraria de la curva de supervivencia. En el gráfico se observa que el mayor riesgo de desertar ocurre después del primer semestre, donde la tasa de riesgo es del 7,5%. En el segundo semestre el riesgo acumulado aumenta a 12,5%, lo que significa que la tasa de riesgo es del 5%. De esta manera se observa que la tasa de riesgo acumulada máxima se alcanza en el séptimo semestre, donde ésta es del 19%.

Sexo

Gráfico 3: Curva de riesgo basado en el sexo del estudiante



Elaborado por: Autor

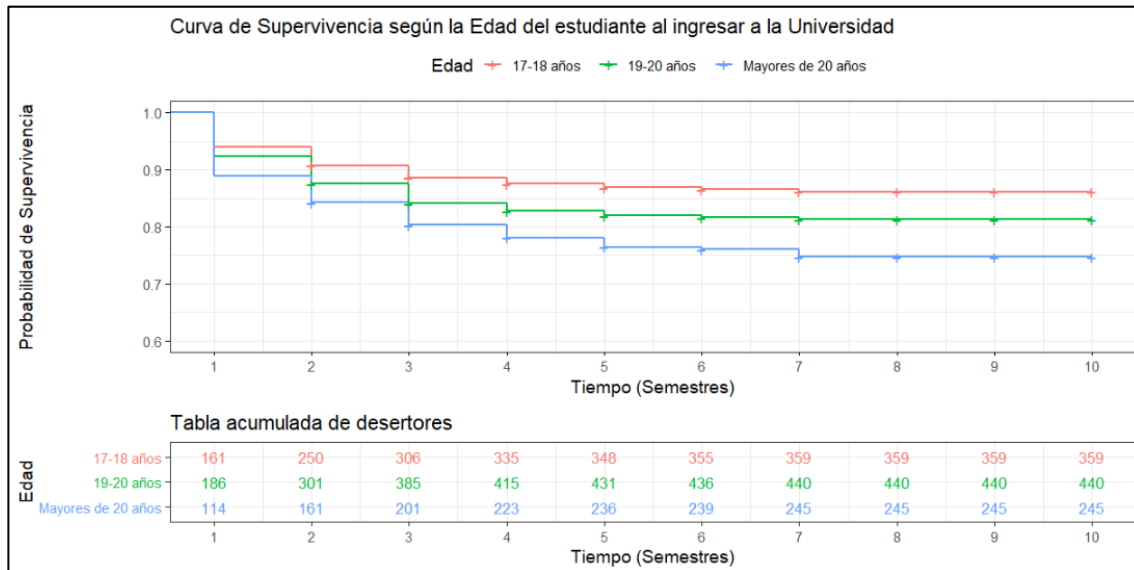
Cuando se tienen dos o más categorías, se utiliza la prueba de rango logarítmico (Log-Rank Test) para comprobar si las curvas de supervivencia entre los grupos son idénticas o no. La hipótesis nula plantea que no existen diferencias entre las categorías, mientras que la hipótesis alternativa plantea diferencias en la deserción por categorías¹⁰.

El estadístico Log Rank calculado a partir de la tabla acumulada del gráfico anterior es aproximadamente 0 (8e-11), por lo cual se rechaza la hipótesis nula, lo que permite evidenciar que existen diferencias significativas en la deserción en función del sexo. La probabilidad de supervivencia de las mujeres es mayor que la de los hombres. La probabilidad de sobrevivir más allá del primer semestre es del 91% para los hombres y del 94% para las mujeres. A partir del séptimo semestre la probabilidad acumulada de supervivencia de los hombres es del 78%, mientras que para las mujeres es del 85%.

¹⁰ En caso de requerir más información revisar la sección de los Modelos de Kaplan Meier en Modelos de Supervivencia.

Edad

Gráfico 4: Curvas de Supervivencia basadas en las categorías de la edad del estudiante.

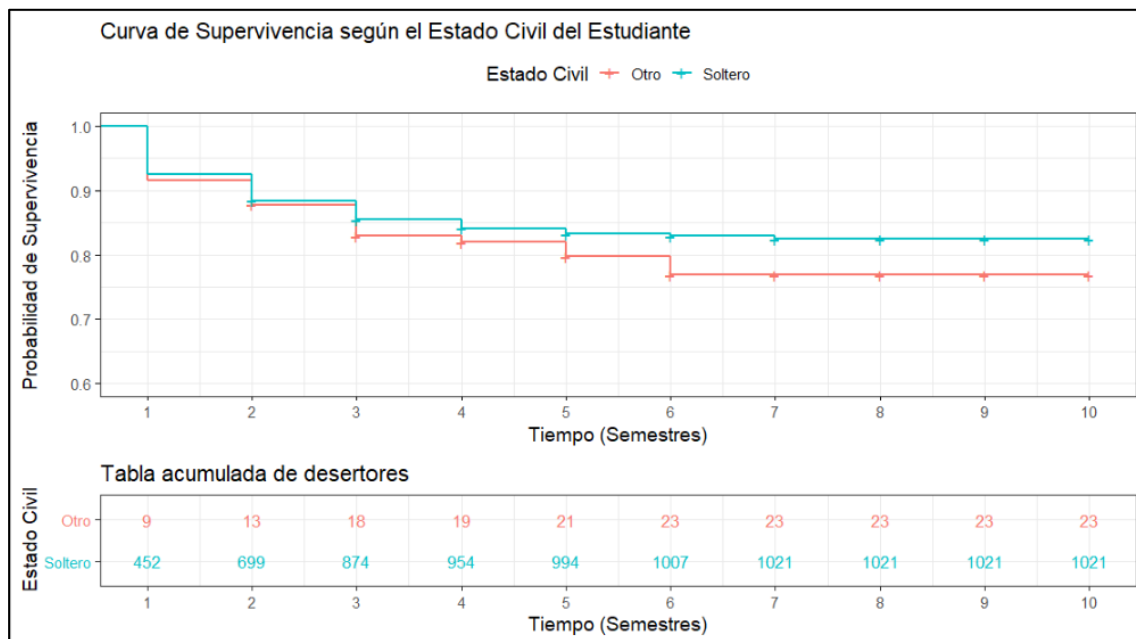


Elaborado por: Autor

El estadístico Log Rank es aproximadamente 0 ($1e-14$); por lo que se rechaza la hipótesis nula, lo que significa que existen diferencias significativas entre las curvas de deserción en función de la edad del estudiante. Mientras mayor edad tiene el estudiante al entrar a la universidad, la probabilidad acumulada de sobrevivir disminuye. La probabilidad acumulada de permanecer en la universidad para los alumnos entre 17 y 18 años es del 86,2%, para los alumnos que ingresan con 19 y 20 años es del 81,4%; mientras que para los alumnos que ingresan teniendo una edad superior a 20 años es del 74,5%.

Estado Civil del Estudiante

Gráfico 5: Curvas de supervivencia basadas en las categorías del estado civil del estudiante.

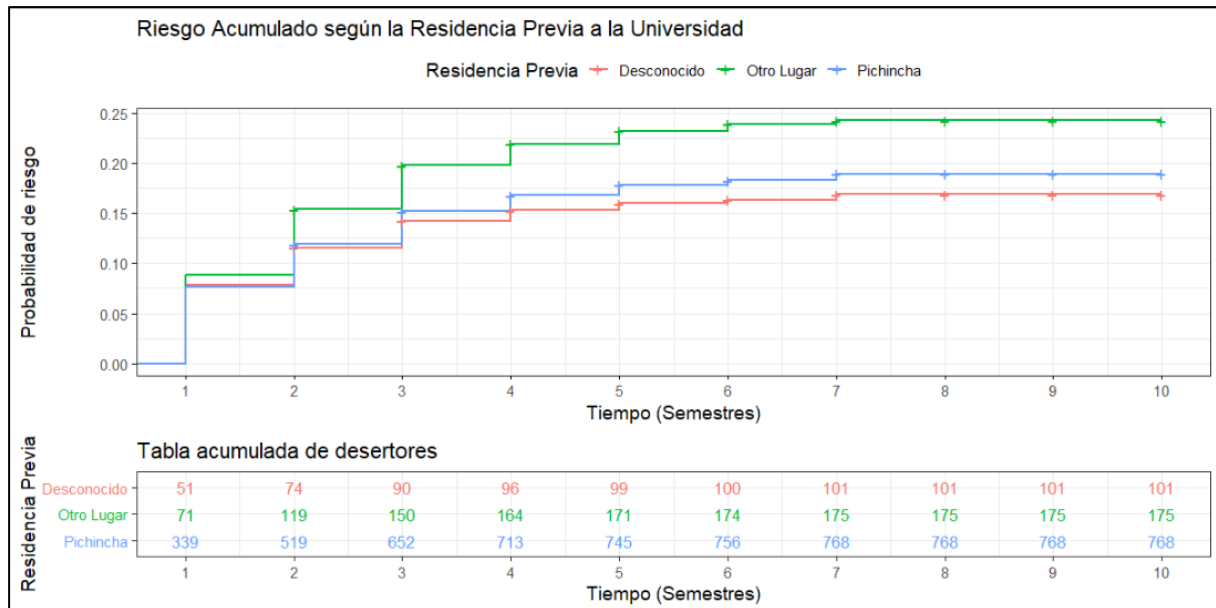


Elaborado por: Autor

El estadístico Log Rank es de 0,2; por lo que no se rechaza la hipótesis nula, lo que significa que no existe una diferencia significativa en la deserción en función del estado civil de estudiante. El comportamiento de ambas curvas de supervivencia es muy similar hasta el quinto semestre, periodo a partir del cual existe una pequeña diferencia entre ambas curvas. En sexto semestre la probabilidad acumulada de sobrevivir para los estudiantes que son solteros es del 82%, mientras que para aquellos estudiantes con un estado civil diferente es del 77%.

Residencia Previa

Gráfico 6: Curvas de riesgo acumulado basado en la residencia previa a la universidad.



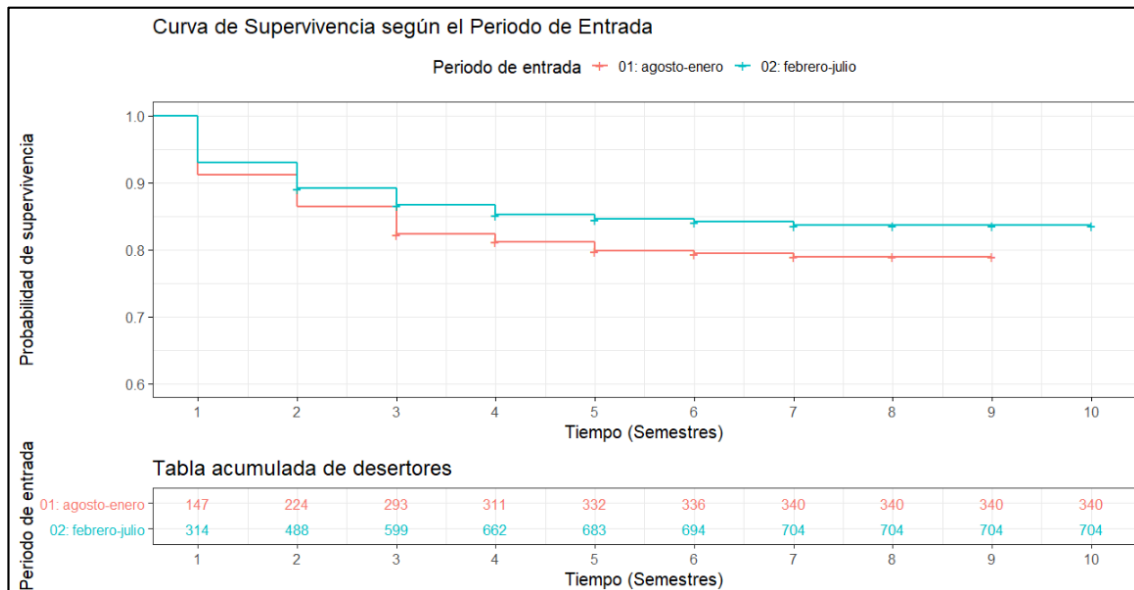
Elaborado por: Autor

El estadístico Log Rank es 0,003; por lo que se rechaza la hipótesis nula, lo que demuestra diferencias significativas en la deserción en función del lugar de residencia previo a la universidad. El comportamiento de las curvas es similar para los estudiantes de Pichincha y para los estudiantes que se desconoce su lugar previo de residencia, sin embargo, aquellos estudiantes que provienen de otras provincias o del extranjero tienen un comportamiento diferente. La probabilidad acumulada de riesgo a partir del séptimo semestre es del 22% para los estudiantes que provenían de otro lugar, del 16% para los que se desconoce la residencia previa y del 17% para los estudiantes que venían de Pichincha.

Los estudiantes con mayor riesgo de desertar son aquellos que provienen de una ciudad, provincia o país diferente al lugar en el que se encuentra la PUCE sede Quito. Esto puede deberse a que los alumnos deben enfrentar un nuevo ambiente totalmente solos. Si bien los padres pueden ayudar en la manutención de sus hijos, al estar lejos no pueden brindarles un apoyo moral o ayudarlos en las dificultades que se les presentan (Sosu & Pheunpha, 2019).

Periodo de Entrada

Gráfico 7: Curvas de supervivencia basados en las categorías del periodo de entrada.



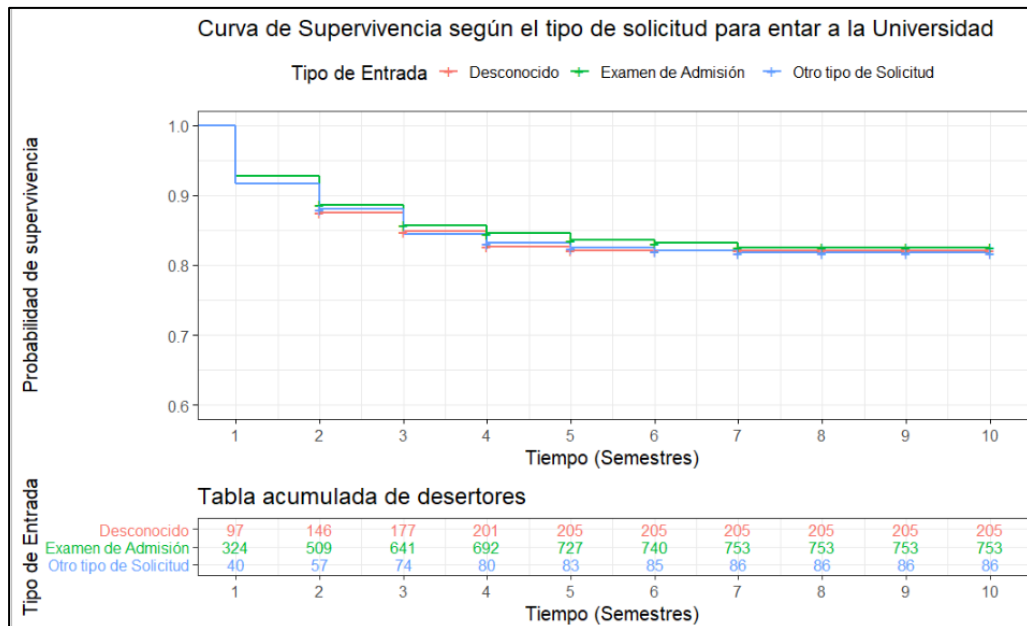
Elaborado por: Autor

El estadístico Log Rank tiene una probabilidad crítica de 0,000025; por tal motivo se rechaza la hipótesis nula y se determina que existe una diferencia significativa en la deserción en función del periodo de entrada. Se observa que los alumnos que entran en el segundo periodo del año (agosto-enero), tienen menos probabilidad de desertar en comparación a los que inician su carrera en el primer periodo (febrero a julio). En el primer semestre la probabilidad de permanecer en la universidad es prácticamente la misma, pues para los que entraron en el primer periodo es del 91% y para los del segundo periodo es del 93%. En el séptimo semestre la probabilidad acumulada de permanecer en la universidad es del 79% para los alumnos que entraron en el primer periodo y del 84% para los del segundo periodo.

Una de las razones por las que más estudiantes entran en el primer período en comparación al segundo es por el calendario académico. La PUCE sede Quito se encuentra en Pichincha, una provincia de la región Sierra, en donde los estudiantes de los colegios se gradúan entre junio y julio, por tanto, se alinea con la entrada a la universidad entre agosto y septiembre. Una posible explicación de por qué los estudiantes que entran en el segundo periodo tienen mayor probabilidad de desertar, puede deberse a que los estudiantes pierden el ritmo de estudio al no inscribirse inmediatamente en la universidad después de graduarse del colegio.

Tipo de Entrada a la Universidad

Gráfico 8: Curva de supervivencia basado en las categorías de tipo de entrada a la Universidad



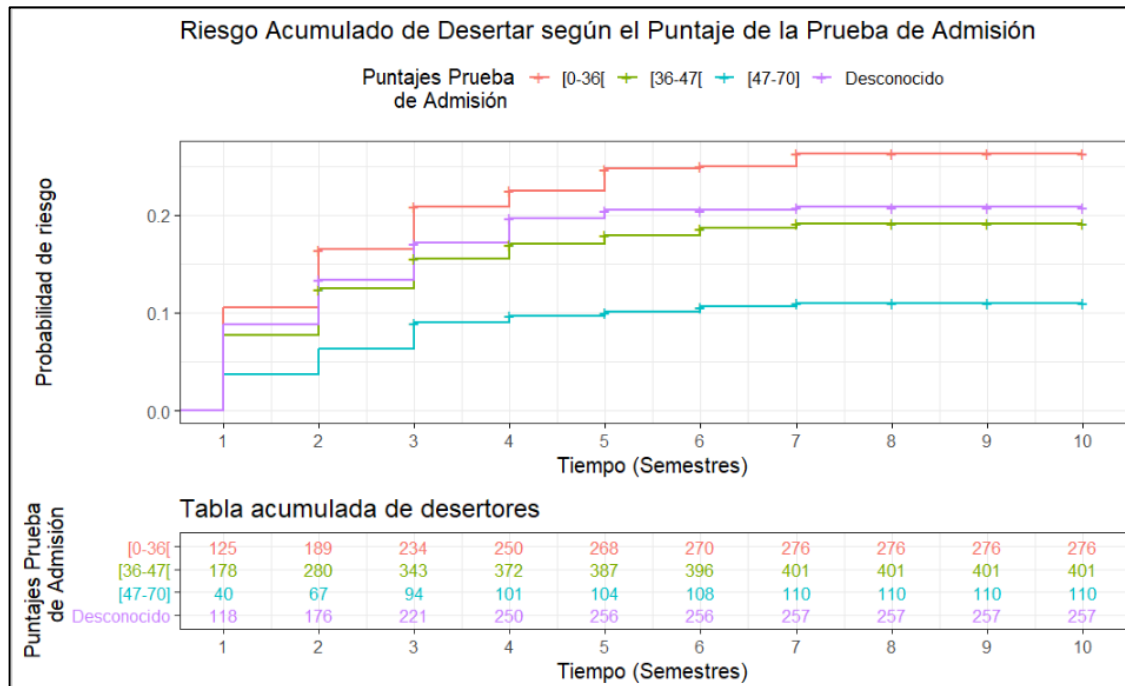
Elaborado por: Autor

El estadístico Log Rank es de 0,5; por lo que no se rechaza la hipótesis nula, lo que significa que no existe una diferencia significativa en la deserción en función de los tipos de entrada a la universidad. La probabilidad acumulada de supervivencia a partir del séptimo semestre es de aproximadamente 82% para las tres categorías.

Esta variable puede demostrar el mismo comportamiento para todas sus variables, ya que la forma de entrar a la universidad puede cambiar en cierto sentido en los procesos que deben seguir los estudiantes, sin embargo, en la mayoría de casos los aspirantes deben rendir la Prueba de Admisión.

Puntaje de la Prueba de Admisión

Gráfico 9: Riesgo acumulado de basado en las categorías de la Prueba de Admisión.



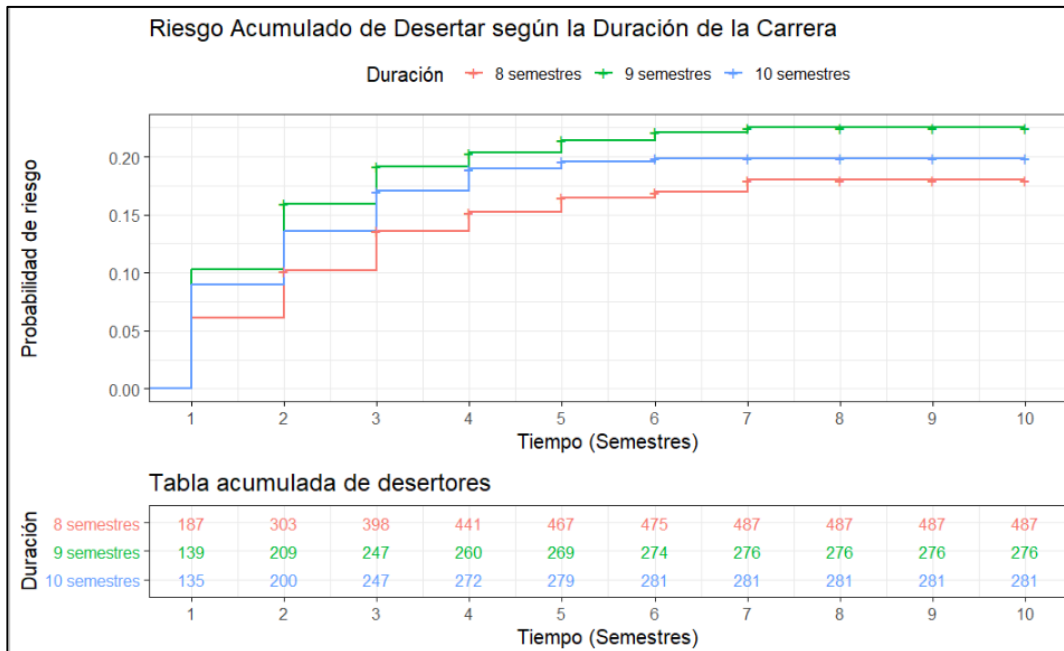
Elaborado por: Autor

El estadístico Log Rank es aproximadamente 0 ($1e-13$); por lo que se rechaza la hipótesis nula, lo que evidencia diferencias significativas en la deserción en función del puntaje obtenido en la prueba de admisión. En el Gráfico 9, se puede observar que mientras más bajo es el puntaje, mayor es el riesgo de desertar. La principal diferencia se presenta en el primer semestre donde el riesgo de desertar para los estudiantes que obtuvieron un puntaje alto, que varía entre 47 y 70 puntos, es del 3,7%, mientras que el riesgo para aquellos que obtuvieron un puntaje entre 0 y 36 puntos, es del 10%. La probabilidad de riesgo acumulado para los estudiantes que obtuvieron un puntaje entre 0 y 36 puntos es del 23%; para aquellos que se desconoce el puntaje, es del 19%; para los que tienen un puntaje entre 36 y 47 puntos es del 17% y para los que obtuvieron los puntajes más alto, entre 47 y 70 puntos, es del 11%.

Varios estudios han demostrado que el puntaje de los exámenes de admisión puede ser un buen estimador del desempeño del estudiante. Estas pruebas por lo general demuestran los conocimientos iniciales con los que entra un estudiante a la universidad, sin embargo, esta variable no es necesariamente considerada un factor determinante de la deserción, ya que, con el tiempo la proyección puede cambiar (Ameri et al., 2016; Juajibioy, 2016; da Costa et al., 2018; Thammasiri et al., 2014).

Duración Ideal de la Carrera

Gráfico 10: Curva de riesgo basado en la duración ideal de la carrera



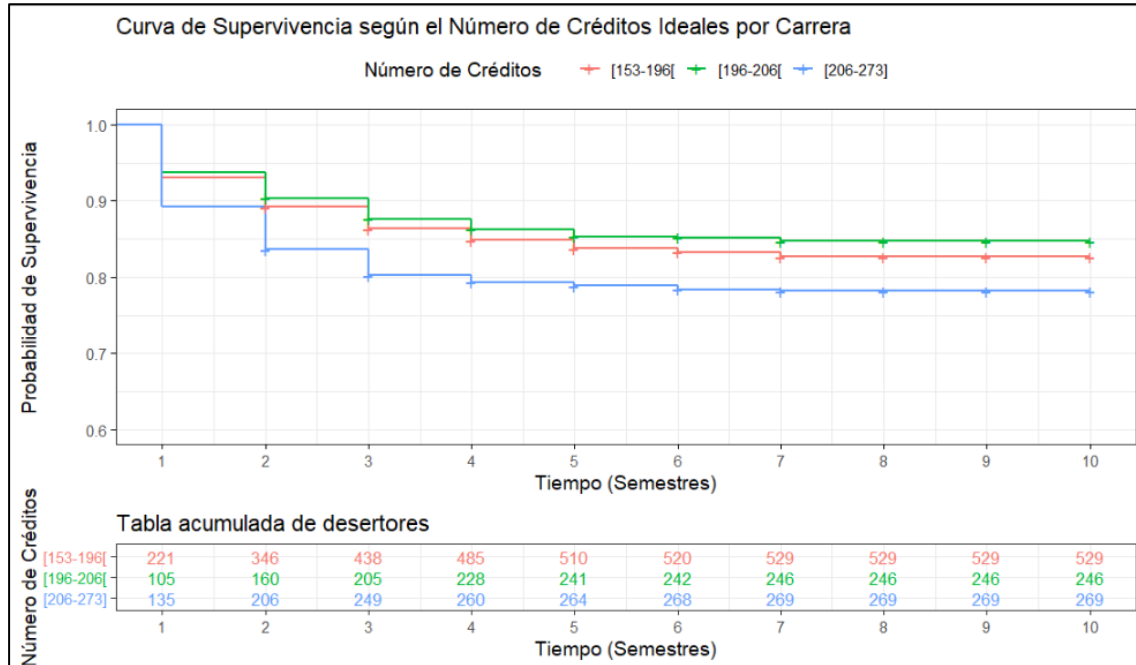
Elaborado por: Autor

El estadístico Log Rank calculado a partir de la tabla acumulada del gráfico anterior es de 0,002, por lo que se rechaza la hipótesis nula, lo que permite evidenciar que existen diferencias significativas en la deserción en función de la duración ideal de la carrera. Como se observa en el gráfico, los estudiantes con mayor riesgo de desertar son aquellos que estudian una carrera de 9 semestres, donde la tasa de riesgo acumulado alcanza el 23%; mientras que los estudiantes con menor riesgo son aquellos que estudian una carrera de 8 semestres, donde la probabilidad acumulada de desertar es del 18%.

Número Ideal de Créditos por Carrera

Para categorizar esta variable se tomó en cuenta la distribución de los datos y las medidas de posición.

Gráfico 11: Curva de supervivencia basado en el número de créditos ideales por carrera



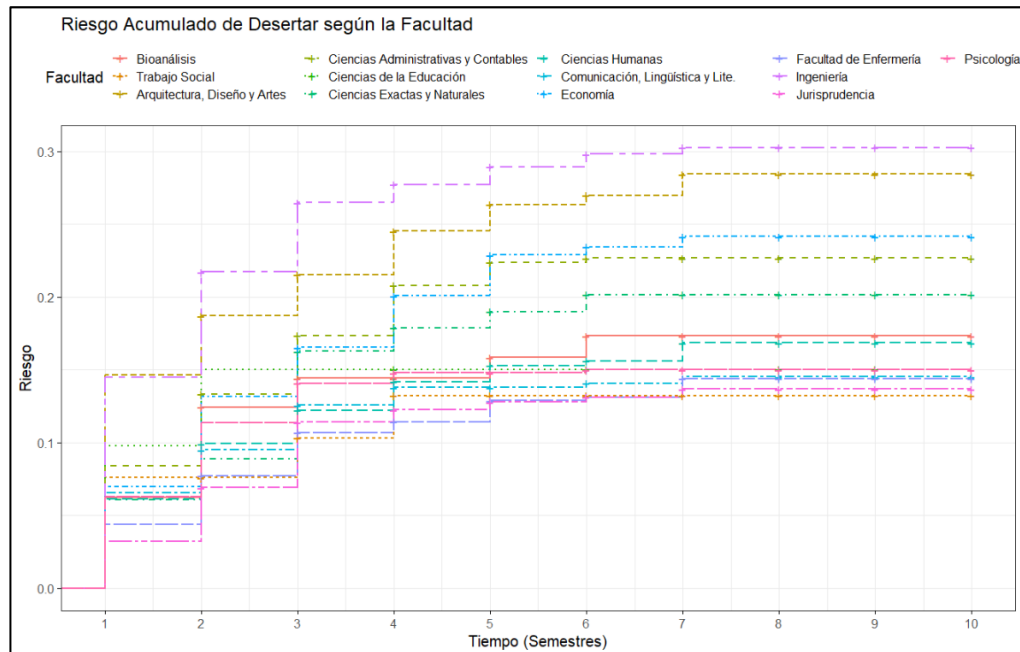
Elaborado por: Autor

El estadístico Log Rank es de 0,000003; por lo que se rechaza la hipótesis nula, lo que permite evidenciar que existen diferencias significativas en la deserción en función del número de créditos de la carrera. Sin embargo, no hay una relación exacta entre el número de créditos ideales que debe tomar el estudiante y la probabilidad de supervivencia. Los estudiantes con mayor probabilidad de sobrevivir son aquellos que toman entre 196 y 206 créditos. La probabilidad acumulada de sobrevivir para este grupo es del 85%. Para los estudiantes que deben tomar entre 153 y 196 créditos, la probabilidad acumulada de supervivencia es del 83%. Por último, para los estudiantes que deben tomar entre 206 y 273 créditos, la probabilidad acumulada de supervivencia es del 78%.

En resumen, se observa que el estudiante tiene menor probabilidad de supervivencia cuando debe tomar un mayor número de créditos. Cuando deben tomar un rango moderado o menor a este, la probabilidad de supervivencia es similar.

Facultad

Gráfico 12: Curvas de riesgo basadas en la Facultad.



Nota: No se incluyó la tabla acumulada de desertores, ya que esta variable contiene múltiples categorías.

Elaborado por: Autor

El estadístico Log Rank es aproximadamente 0 ($2e-16$); por lo que se rechaza la hipótesis nula, lo que significa que las curvas de riesgo en función de la Facultad a la que pertenece el estudiante son significativamente diferentes. La principal diferencia se da en el primer semestre, donde los estudiantes corren mayor o menor riesgo de acuerdo a la Facultad a la que pertenecen.

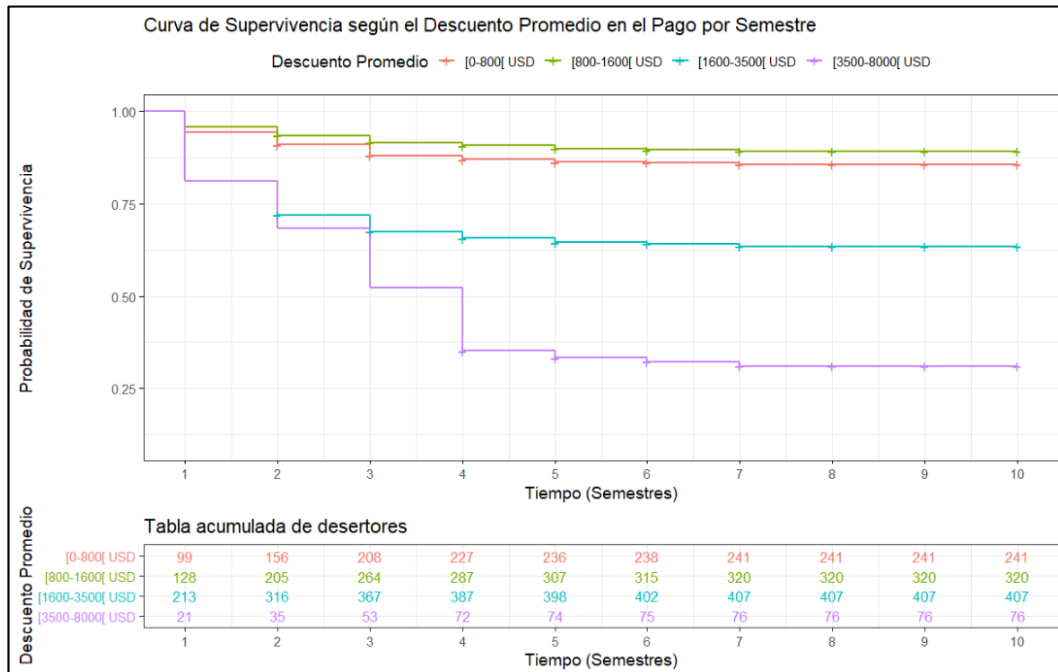
Al observar el gráfico se puede ver que la Facultad con mayor riesgo es Ingeniería, una con un riesgo muy bajo es la de Jurisprudencia y una con un riesgo intermedio es la de Ciencias Exactas y naturales. Al compara estas tres Facultades, se puede observar que, en primer semestre, los estudiantes de la Facultad de Ingeniería tienen una probabilidad de desertar del 13%, mientras que los estudiantes de Ciencias Exactas y Naturales tienen una tasa de riesgo del 6% y los estudiantes de Jurisprudencia del 3%. El riesgo acumulado de desertar para las mismas Facultades es del 30%, 17% y del 13% respectivamente.

Al observar el gráfico, se puede ver que todas aquellas carreras que tienen un riesgo de desertar mayor al 20%, son aquellas carreras que contienen análisis más cuantitativos dentro de sus mallas de estudio y donde el uso de las matemáticas es fundamental. Además, el título de estas carreras les permite a los estudiantes trabajar en profesiones menos filantrópicas. Por otro lado, las Facultades con un riesgo acumulado menor al 20%, son aquellas carreras donde al graduarse los estudiantes llegan a trabajar directamente con personas y por su bienestar. Es necesario mencionar que alrededor del mundo existen múltiples estudios que demuestran que las carreras de Ingeniería son las que presentan las mayores tasas de deserción (Meyer & Marx, 2014; Paura & Arhipova, 2014; Salas-Morera et al., 2018).

Descuento Promedio por Semestre

Se utilizó la variable categórica de descuento promedio por semestre ya que el descuento total no toma en consideración el número de semestres que ha cursado el estudiante, por lo tanto, el verdadero efecto podría verse afectado

Gráfico 13: Curvas de supervivencia de las categorías de descuento promedio por semestre.



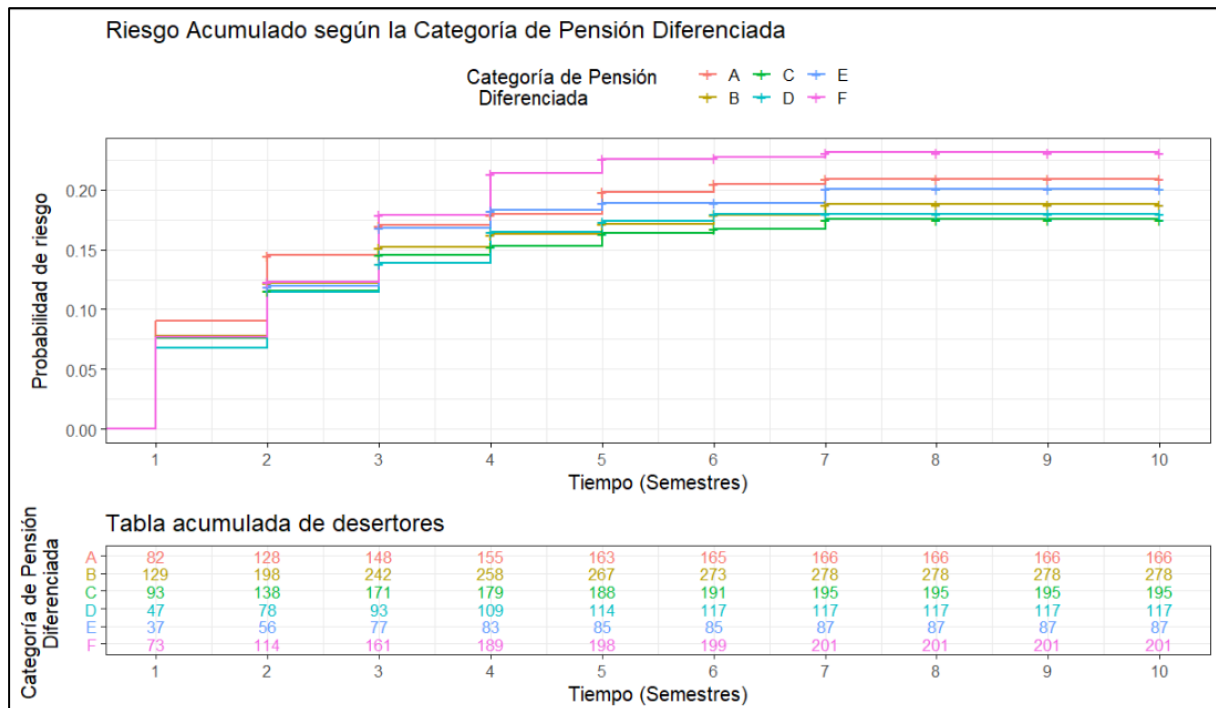
Elaborado por: Autor.

El estadístico Log Rank es aproximadamente 0 ($2e-16$); por lo que se rechaza la hipótesis nula, lo que demuestra diferencias significativas en la deserción en función del descuento promedio por semestre. En base al gráfico se observa que la supervivencia acumulada a partir del séptimo semestre, que es cuando la curva se estabiliza, difiere en gran medida para cada categoría. Para los estudiantes que reciben entre 3500 y 8000 USD de descuento, la probabilidad de supervivencia acumulada es del 31%, para los que reciben un descuento entre 1600 y 3500 USD es del 64%, para aquellos estudiantes que reciben un descuento entre 800 y 1600 USD es del 89% y para los que reciben un descuento entre 0 y 800 USD es del 86%. No existe un patrón de comportamiento, sin embargo, se observa que los estudiantes con mayor riesgo de desertar son aquellos que reciben un mayor descuento.

El comportamiento de estas curvas es diferente al esperado, pues se esperaría que mientras mayor sea el descuento, menor sea el riesgo de desertar. No obstante, las personas encargadas de las becas y admisiones de la DGE y la DGA comentaron que las becas en donde se cubre el 75% o el 100% del costo de la carrera por semestre son las becas SNNA, que son otorgadas a personas de escasos recursos. En estos casos las personas pueden tender a desertar debido a que no tienen los conocimientos necesarios o debido a que no logran adaptarse y socializar con sus compañeros.

Categoría Socioeconómica de la Pensión Diferenciada (SPD)

Gráfico 14: Curvas de riesgo acumulado basado en las Categorías de Pensión Diferenciada.



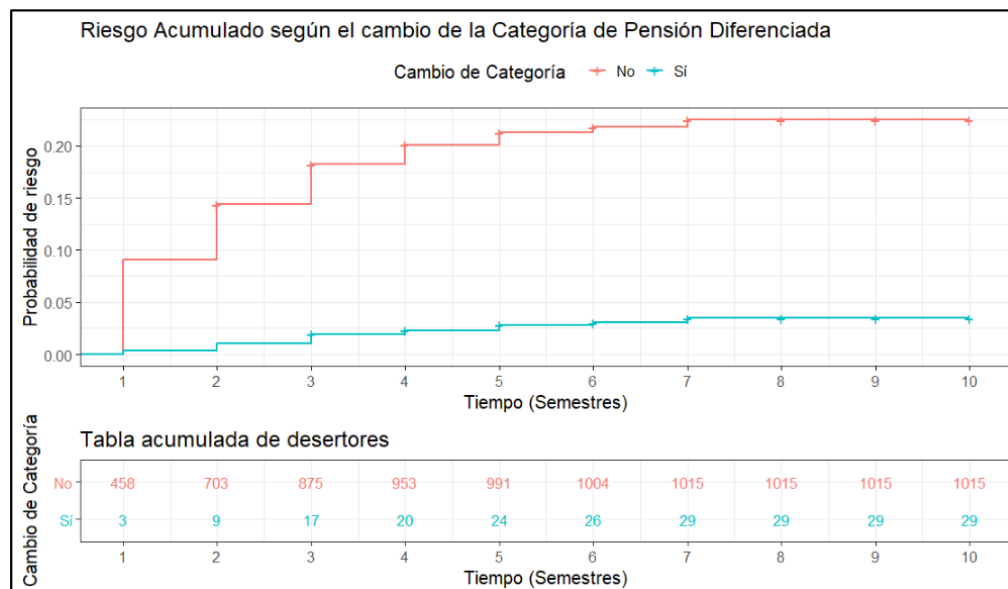
Elaborado por: Autor

El estadístico Log Rank es de 0,07; por lo que no se rechaza la hipótesis nula, lo que significa que no existe una diferencia significativa en la deserción en función de la categoría socioeconómica de la pensión diferenciada. Las curvas de supervivencia de todas las categorías tienen un comportamiento similar, sin embargo, la curva de supervivencia de la categoría F, presenta el mayor riesgo en comparación al resto de categorías, con una probabilidad acumulada de desertar del 21%. La categoría con menor probabilidad de desertar es la C, que al final del séptimo semestre llega a tener un riesgo acumulado del 16%. El comportamiento similar de las curvas no permite evaluar si esta variable predice o no la deserción.

Es sorprendente encontrar que la variable que identifica la categoría socioeconómica de los estudiantes no es significativa, ya que existen otros estudios que demuestran lo contrario, como es el caso de la investigación de Juajibioy (2016). En este caso, la diferencia entre las distintas curvas es mínima, la única que muestra una diferencia significativa es la categoría F. La explicación del comportamiento de los estudiantes que se encuentran en esta categoría, puede deberse a que los estudiantes no reciben ningún tipo de ayuda o descuento en su pensión, por lo tanto, al analizar su costo de oportunidad, los estudiantes pueden pensar que les da lo mismo el estudiar en la PUCE sede Quito o en cualquier otra universidad privada del Ecuador.

Cambio en la Categoría SPD

Gráfico 15: Curvas de riesgo acumulado basado en el cambio de categoría SPD.

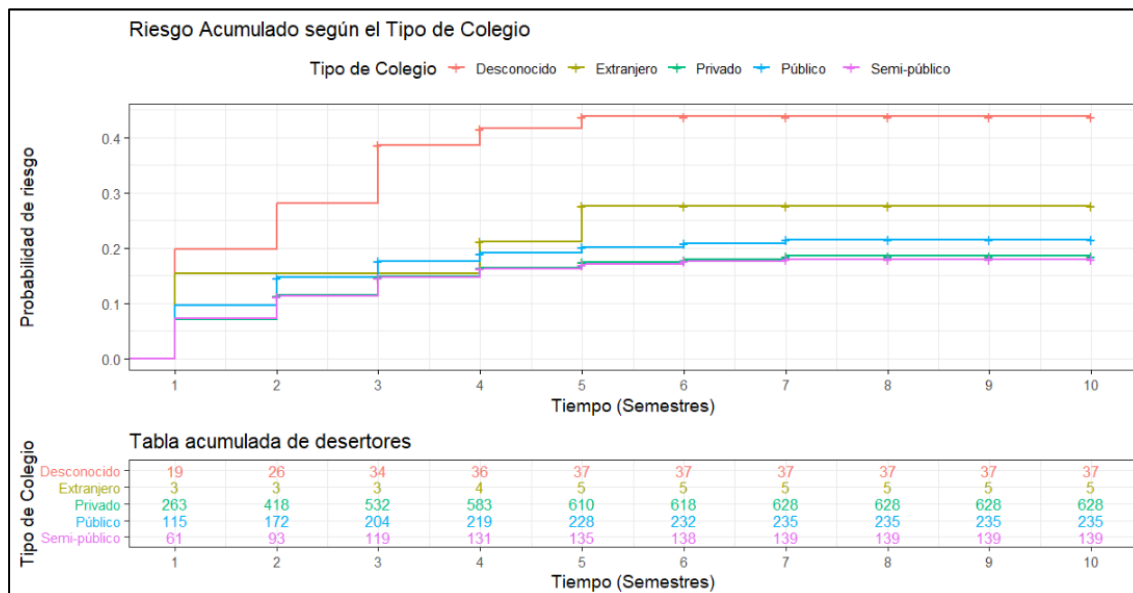


Elaborado por: Autor.

El estadístico Log Rank es aproximadamente 0 ($2e-16$); por lo que se rechaza la hipótesis nula, lo que demuestra diferencias significativas en la deserción en función del cambio la categoría socioeconómica de la pensión diferenciada. Existe una diferencia significativa entre los estudiantes que reciben un cambio y los que no. La diferencia se da principalmente en el primer semestre donde los estudiantes que no reciben este cambio alcanzan una tasa de riesgo del 9% mientras que la tasa de riesgo para aquellos estudiantes sí recibieron el cambio es inferior al 1%. La tasa acumulada de riesgo para aquellos que no cambiaron de categorías es del 20%, mientras que para los que sí recibieron el cambio, es del 3,4%.

Tipo de Colegio

Gráfico 16: Curvas de riesgo basado en las categorías de tipo de colegio.



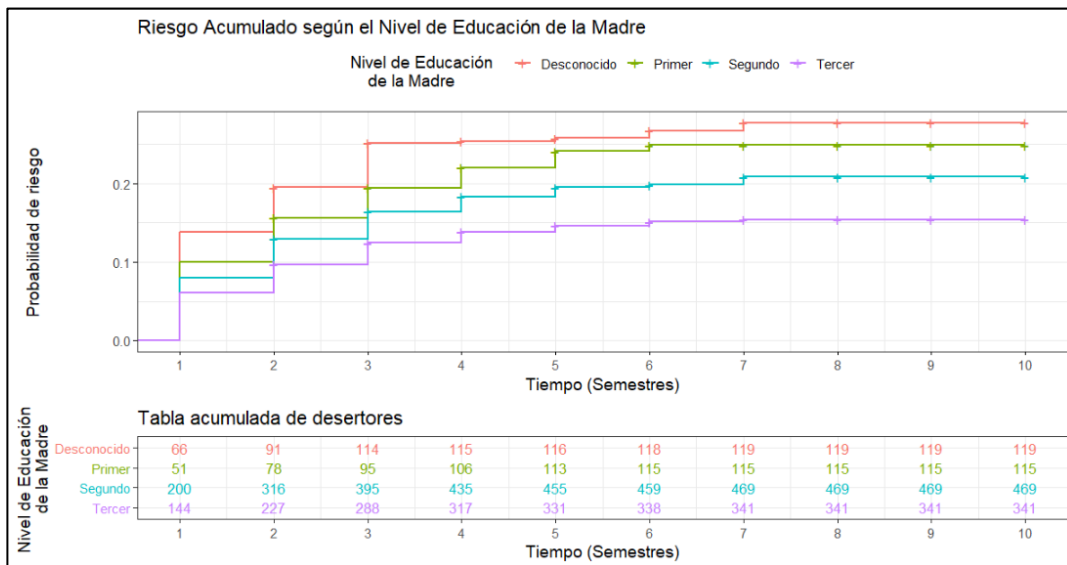
Elaborado por: Autor

El estadístico Log Rank es 0,000001; por lo que se rechaza la hipótesis nula, lo que demuestra diferencias significativas en la deserción en función del tipo de colegio del que se graduó el estudiante. Las curvas de riesgo acumulado del colegio privado y semi-público tienen un comportamiento casi idéntico, sin embargo, el resto de categorías difieren significativamente. En primer semestre la tasa de riesgo es del 7% para los estudiantes que provenían de colegios privados y semi-públicos; para los estudiantes de colegios públicos es del 9%; para los estudiantes de colegios extranjeros es del 14% y para aquellos que no se conoce el colegio es del 18%. A partir del séptimo semestre, el riesgo acumulado para los estudiantes de un colegio semi-público es del 16%, para los de colegio privado del 17%, para los de colegio público del 19%, para los de colegio extranjero del 24% y para aquellos que se desconoce el colegio es del 36%.

Cabe mencionar que la información de los colegios se obtuvo del Ministerio de Educación. La categoría “desconocido”, se les otorgó a todos los alumnos que no tenían ningún registro. Este error se pudo deber a un error del sistema y en el peor de los casos a una situación de fraude. Por lo tanto, se desconoce las causas exactas para cada uno de los casos en donde se omite esta información; sin embargo, llama la atención que esta sea la categoría con mayor probabilidad de deserción.

Nivel de Educación de la Madre

Gráfico 17: Curvas de riesgo acumulado basado en las categorías del nivel de educación de la madre del estudiante.



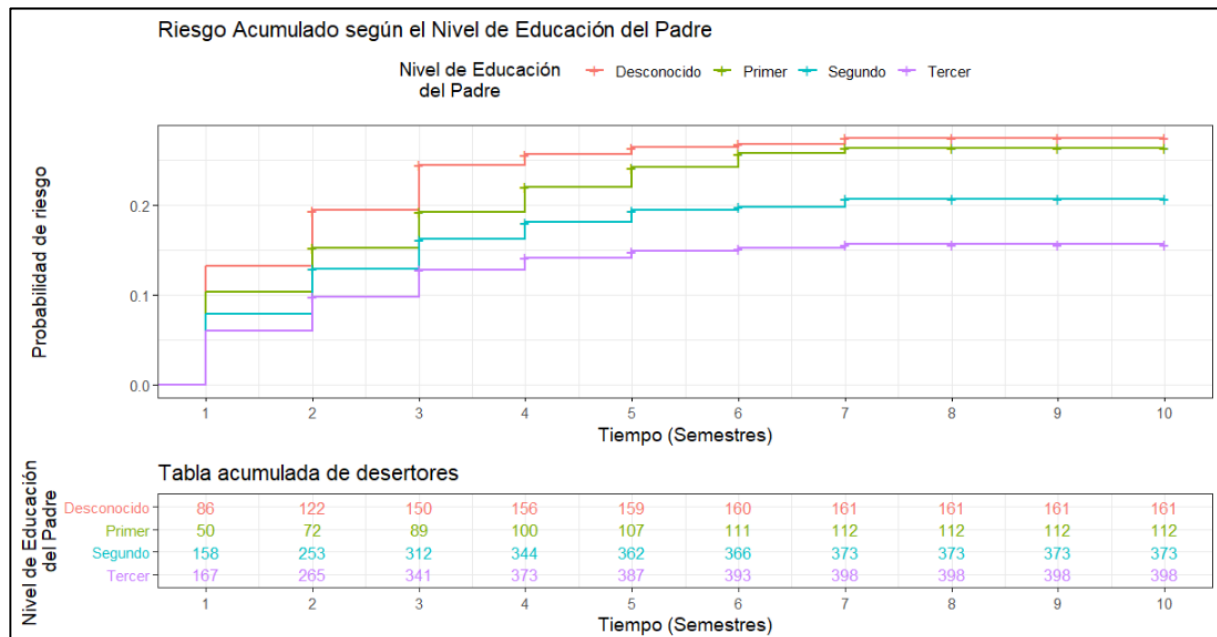
Elaborado por: Autor

El estadístico Log Rank es de 0,000000002; por lo cual se rechaza la hipótesis nula, lo que permite evidenciar que existen diferencias significativas en la deserción en función del nivel de educación de la madre del estudiante. En base al gráfico, se puede observar que mientras mayor es el grado de educación de la madre, existe menor probabilidad de desertar para el estudiante.

Las mayores tasas de riesgo ocurren en los primeros niveles. La probabilidad de riesgo de desertar para los estudiantes con madres con tercer nivel de educación es del 6%, para aquellos que tienen madres con un segundo nivel de educación es del 8%, mientras que para los estudiantes con madres con un primer nivel de educación es del 10% y para los que tienen madres con un nivel de educación desconocido es del 13%. Conforme los estudiantes avanzan en sus estudios, el riesgo disminuye. El riesgo se reduce en menor proporción para aquellos estudiantes con madres que tienen un nivel de educación desconocido. A partir de séptimo semestre el riesgo acumulado de desertar para los estudiantes con madres que han alcanzado un tercer nivel es del 14%, para los de segundo nivel es del 19%, para los del primer nivel es del 22% y para todos aquellos que se desconoce el nivel de educación de sus madres es del 24%.

Nivel de Educación del Padre

Gráfico 18: Curvas de riesgo acumulado basado en las categorías del nivel de educación del padre del estudiante.



Elaborado por: Autor

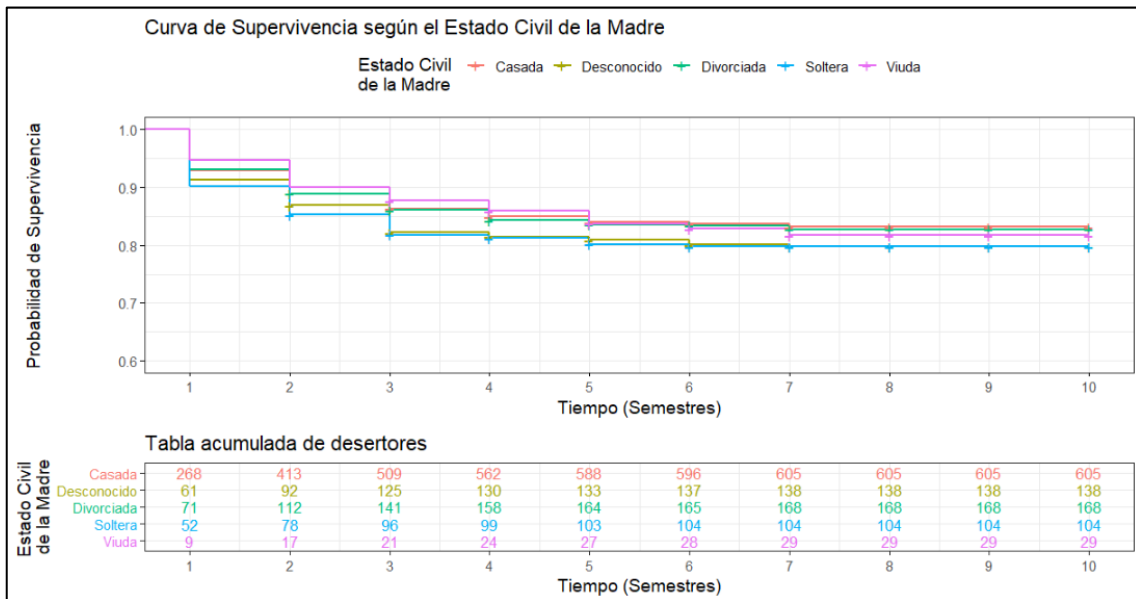
El estadístico Log Rank es aproximadamente 0 ($5e-11$); por lo que se rechaza la hipótesis nula, lo que demuestra diferencias significativas en la deserción en función del nivel de educación del padre del estudiante. Mientras mayor es el nivel de educación alcanzado por el padre, menor es la probabilidad de desertar del estudiante. La mayor tasa de riesgo alcanzada para todos los estudiantes, ocurre en primer nivel, aunque en diferentes magnitudes. En primer nivel, el riesgo para todos aquellos estudiantes con padres que han alcanzado un tercer nivel de educación es del 6%, mientras que para los estudiantes con padres que han alcanzado un segundo nivel es del 8%. Para los estudiantes con padres que han alcanzado un primer nivel, la tasa de riesgo es del 10% y de aquellos con un nivel de educación desconocido es del 12%. El riesgo acumulado para los estudiantes con padres que han alcanzado un tercer nivel es del 14%, para los de segundo nivel es del 19%, para los del primer nivel es del 23% y para todos aquellos que se desconoce es del 24%.

El comportamiento de las curvas de riesgos en función del nivel de educación de los padres, es muy similares al de las curvas de riesgos en función del nivel de educación de las madres de los estudiantes.

El nivel de educación de los padres y las madres influye en varios sentidos en el riesgo de desertar de los estudiantes. Mientras mayor es el nivel de formación, mayor apoyo y ayuda podrán brindar a sus hijos. También puede influir en la economía familiar, pues mientras más alto sea el nivel de educación, mayor es la probabilidad de tener un sueldo alto, por ende, el riesgo de desertar debido a problemas económicos se reduce. Por último, los mismos estudiantes pueden sentirse motivados a alcanzar al menos el mismo nivel de educación de sus padres, por lo tanto, mientras mayor sea el nivel de educación de los padres, menor riesgo de desertar tendrá el estudiante (Aina, 2005; Farah & Upadhyay, 2017; Bird, 2018).

Estado Civil de la Madre

Gráfico 19: Curva de supervivencia basada en el estado civil de la madre del estudiante.

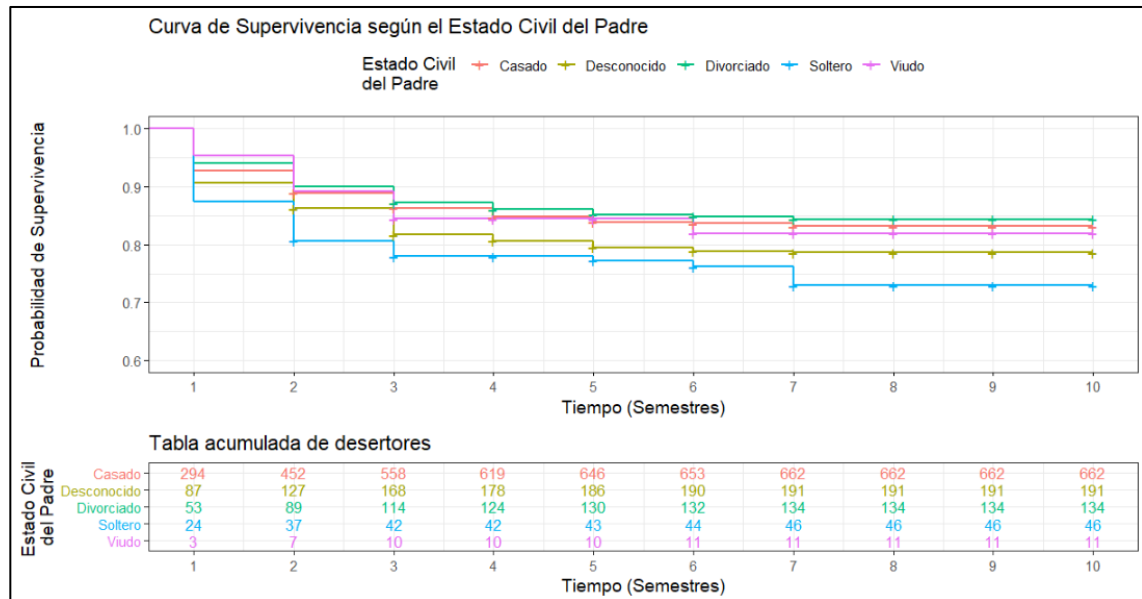


Elaborado por: Autor

El estadístico Log Rank es de 0,07; por lo que no se rechaza la hipótesis nula, lo que significa que no existe una diferencia significativa en la deserción en función del estado civil de la madre del estudiante. El comportamiento de las curvas de supervivencia de los diferentes grupos es muy similar. La probabilidad acumulada de supervivencia a partir de séptimo semestre varía entre 80 y 83% para cualquiera de las categorías.

Estado Civil del Padre

Gráfico 20: Curva de supervivencia basada en el estado civil del padre del estudiante.



Elaborado por: Autor

El estadístico Log Rank es 0,0003; por lo que se rechaza la hipótesis nula, lo que significa que existen diferencias significativas entre las curvas de deserción en función del estado civil del padre del estudiante. Las curvas de supervivencia según el estado civil del padre muestran diferentes comportamientos de acuerdo a las categorías. La curva de supervivencia de los estudiantes con padres divorciados, casados y viudos es similar, y la de los estudiantes con padres solteros o un estado civil desconocido, difiere en gran medida. La probabilidad acumulada de supervivencia para los estudiantes con padres divorciados es del 84%, para estudiantes con padres casados es del 83% y para los que tienen padres viudos es del 82%. Para los estudiantes con padres que se desconoce el estado civil es del 79% y para aquellos que tienen padres solteros es del 73%.

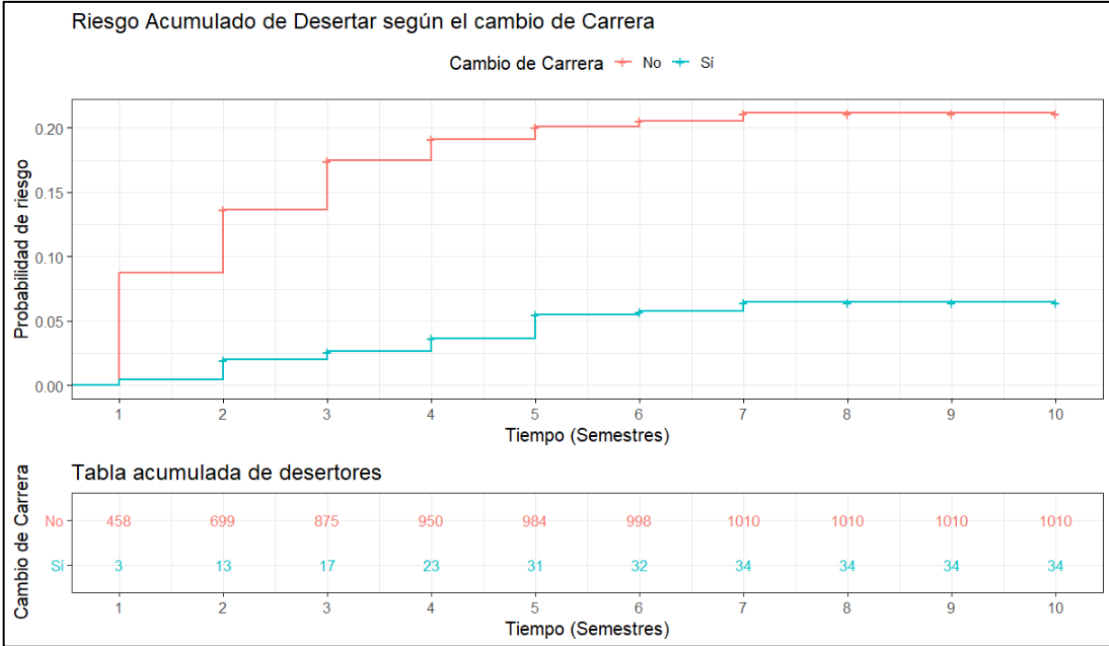
El estado civil del padre tiene relevancia sobre la deserción, a diferencia del estado civil de la madre. Esto se puede deber a la configuración de los hogares, pero también puede tener gran relación con las leyes ecuatorianas. De acuerdo al art. 4 del Código de la Niñez y Adolescencia – Ley 100 – RO 737 (2003):

Tienen derecho a reclamar alimentos [...] los adultos o adultas hasta la edad de 21 años que demuestren que se encuentran cursando estudios en cualquier nivel educativo que les impida o dificulte dedicarse a una actividad productiva y carezcan de recursos propios y suficientes; [...]

A partir de esta ley se puede deducir que todos aquellos estudiantes que tienen padres divorciados tienen menor riesgo de desertar debido a que sus progenitores tienen la obligación de mantenerlos durante sus estudios. Por otro lado, los padres de los estudiantes con estado civil soltero, probablemente no tienen ninguna relación con sus hijos y esta es la razón por la que estos alumnos tienen mayor riesgo de desertar.

Cambio de Carrera

Gráfico 21: Curvas de riesgo basado en las categorías de cambio de carrera.



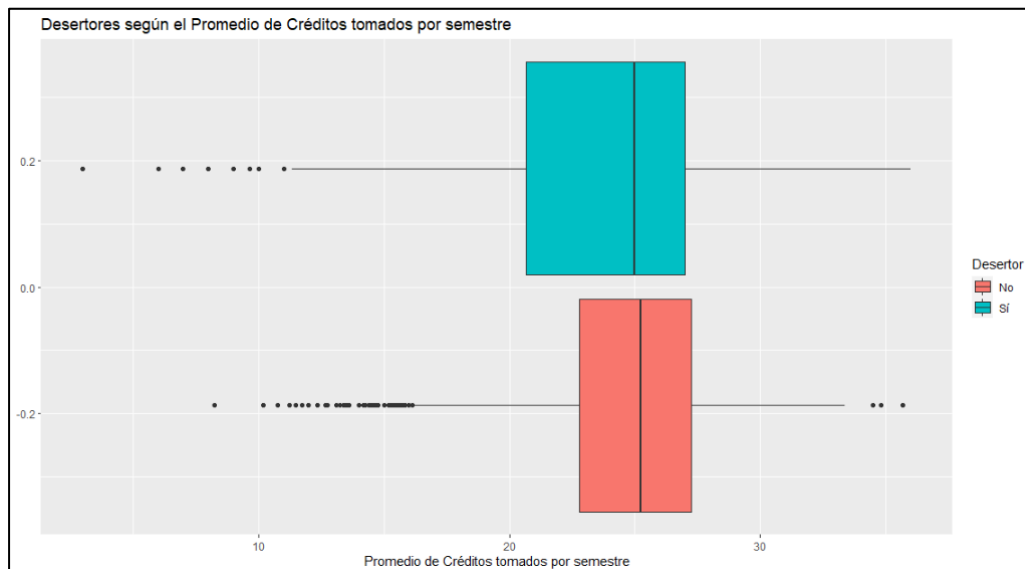
Elaborado por: Autor

El estadístico Log Rank es aproximadamente $0 (2e-16)$; por lo que se rechaza la hipótesis nula, lo que demuestra diferencias significativas en la deserción en función del cambio de carrera. Las curvas de supervivencia de los estudiantes que se cambian de carrera y de los que no se cambian tienen una tasa de crecimiento diferente. El riesgo acumulado para los estudiantes que no se cambian de carrera, a partir del séptimo semestre es del 19,4%, mientras que para los estudiantes que sí se cambian de carrera es del 6,4%.

La explicación que se le puede dar a este comportamiento es que una vez que el estudiante se cambia de carrera, se siente bien, tiene motivación de estudiar y le agrada el ambiente en el que se encuentra; por tanto, tienen menor riesgo de desertar.

Promedio de Créditos por Semestre

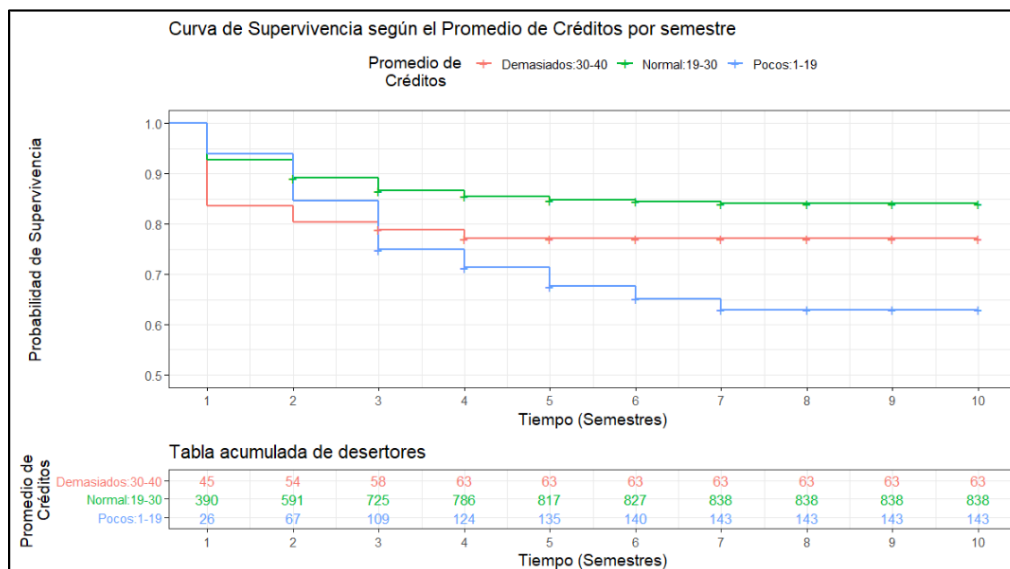
Gráfico 22: Gráfico de caja y bigotes basado del promedio de créditos por semestre.



Elaborado por: Autor.

En el Gráfico 22 se puede observar que los estudiantes que desertan tienen un rango de créditos más amplio en comparación a aquellos estudiantes que no desertan. El 50% de los estudiantes toman en promedio 25 créditos, sin importar si han sido desertores o no.

Gráfico 23: Curvas de supervivencia basadas en el promedio de créditos por semestre



Elaborado por: Autor.

El estadístico Log Rank es aproximadamente $0 (2e-16)$; por lo que se rechaza la hipótesis nula, lo que significa que existen diferencias significativas entre las curvas de deserción en función del promedio de créditos tomados por el estudiante.

Las curvas de las tres categorías en las que se ha dividido la variable promedio de créditos por semestre tienen un comportamiento distinto entre ellas y con respecto al resto de variables. Los estudiantes que toman pocos créditos, entre 1 y 19 créditos son los estudiantes con menor riesgo de desertar en primer semestre, pues en este nivel, la probabilidad de supervivencia es del 94%. A partir del tercer semestre, los estudiantes que toman pocos créditos llegan a ser los estudiantes con mayor probabilidad de desertar, pues la probabilidad de riesgo acumulado es del 75%. A partir del séptimo semestre la probabilidad acumulada de supervivencia para los estudiantes que toman pocos créditos es del 63%.

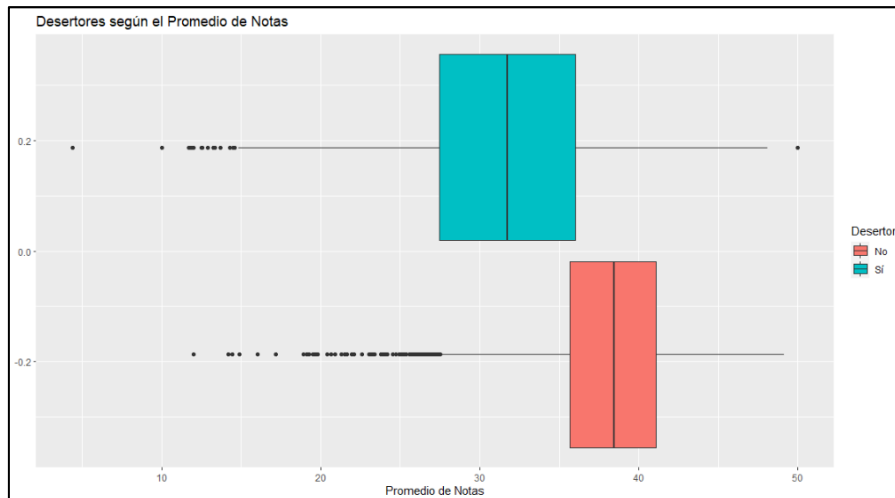
Por otro lado, los estudiantes que toman entre 30 y 40 créditos, han sido categorizados como estudiantes que toman demasiados créditos. Este grupo presenta un mayor riesgo durante el primer semestre, pues la probabilidad de supervivencia después del primer semestre es del 84%, sin embargo, conforme el estudiante avanza en sus estudios, la tasa de riesgo disminuye y a partir del cuarto semestre la probabilidad acumulada de permanecer en la universidad es del 77%.

El grupo de estudiantes que toman entre 19 y 30 créditos, que es el número de créditos promedio recomendado por la Universidad, mantienen durante todos los periodos la mayor probabilidad de supervivencia. Este grupo tiene un 93% de probabilidad de sobrevivir después del primer semestre y a partir del séptimo semestre la probabilidad acumulada de supervivencia es del 84%.

El tomar un menor número de créditos conforme el estudiante avanza en sus estudios puede deberse a que el estudiante ha perdido varias materias, por tanto, no puede tomar el número sugerido de créditos. Esta situación va muy ligada al desempeño académico y puede ocasionar la deserción del alumno. Por otro lado, los estudiantes que toman demasiadas materias, pueden tener alguna necesidad que les obliga a tomar más créditos, como igualarse en la carrera o alguna situación específica, sin embargo, en estos casos el estudiante tiene una mayor probabilidad de desertar. Esto se debe a múltiples factores, tales como no convivir con sus compañeros, estrés debido a la carga de materias, dificultad para organizarse, entre otros (Tinto, 1975).

Promedio General de Notas

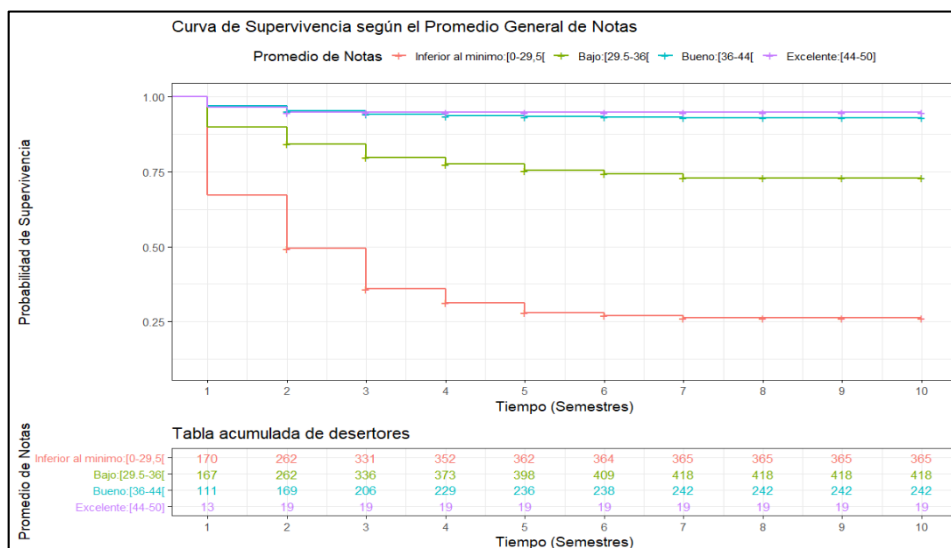
Gráfico 24: Gráfico de caja y bigotes basado en el promedio general de notas de los estudiantes.



Elaborado por: Autor

Se puede observar que los estudiantes que tienen un promedio de calificaciones alto, tienen menor riesgo de desertar que aquellos estudiantes con promedios más bajos. Las notas de los estudiantes que desertan se encuentran acumuladas entre 27 y 36 puntos, donde la mediana de calificaciones es de 32. Por otro lado, las notas de los estudiantes que no desertan se encuentran acumulado entre 36 y 41 puntos, donde la mediana de las notas es de 39.

Gráfico 25: Curvas de supervivencia basadas en las categorías del promedio general de notas.



Elaborado por: Autor.

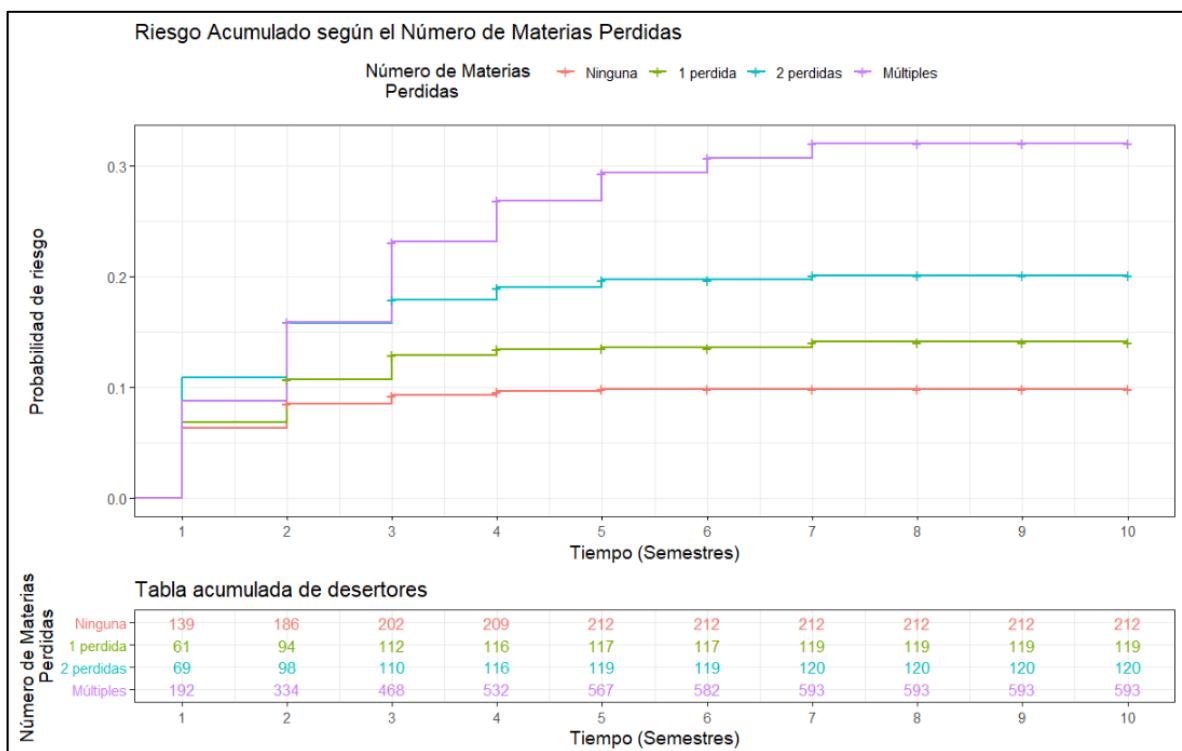
El estadístico Log Rank es aproximadamente 0 ($2e-16$); por lo que se rechaza la hipótesis nula, lo que significa que existen diferencias significativas entre las curvas de deserción en función del promedio general de notas. En base al gráfico se observa que mientras más alto es el promedio general de notas, mayor es la probabilidad de supervivencia del estudiante. La probabilidad acumulada de sobrevivir cuando el estudiante tiene un promedio inferior al mínimo es del 26%, cuando el promedio es bajo, la

probabilidad de supervivencia es del 73%, cuando tiene un promedio bueno es del 93% y cuando tiene un promedio alto, es del 95%.

El promedio general de notas es una de las principales variables de desempeño, por ende, puede ser una variable relevante para pronosticar la deserción. Sin embargo, esta variable también refleja la satisfacción del estudiante y cómo se puede sentir. Según Tinto (1975), las notas pueden ser un reflejo de la adaptación del estudiante.

Número de Materias Perdidas

Gráfico 26: Curvas de riesgo basadas en las categorías del total de materias perdidas.



Elaborado por: Autor.

El estadístico Log Rank es aproximadamente 0 ($<2e-16$); por lo que se rechaza la hipótesis nula, lo que permite evidenciar que existen diferencias significativas en la deserción en función del número total de materias perdidas. Se observa que mientras más materias ha perdido el estudiante, mayor es el de riesgo de desertar. Los estudiantes que no han perdido ninguna materia durante toda su carrera, tienen una tasa de riesgo acumulado del 10%. Cuando el estudiante ha perdido una materia, el riesgo acumulado de desertar es del 14%, mientras que cuando han perdido dos materias, es del 20%; por otro lado, cuando han perdido más de dos materias, es del 33%.

Perfil de los estudiantes desertores

A continuación, se realizará un resumen tomando en consideración el análisis no paramétrico de los individuos, donde se evalúo la deserción en función de cada una de las variables.

Al evaluar las características intrínsecas de los estudiantes se pudo observar que los hombres tienen un mayor riesgo de desertar en comparación a las mujeres. También se puede ver que mientras más edad tiene el estudiante al iniciar sus estudios en la universidad, mayor es el riesgo de desertar. Por otro lado, se observó que el estado civil del estudiante no influye en la deserción. Por último, al observar si la residencia previa influye, se observó que los estudiantes que residían en Pichincha, previo a la entrada a la universidad, tenían menor riesgo de desertar en comparación a los estudiantes que provenían de otras provincias o del extranjero. Cabe mencionar que el 75% de los estudiantes provienen de Pichincha, por lo tanto, el efecto puede no ser claro.

En cuanto a las características previas o iniciales a la entrada a la universidad, se observa que existe una mayor probabilidad de permanecer en la universidad cuando el estudiante entra por primera vez a la PUCE sede Quito en el semestre que inicia en agosto. Cabe mencionar que un mayor número de alumnos optan por entrar en el segundo periodo. Por otro lado, el tipo de entrada no influye en la deserción de los estudiantes, sin embargo, el puntaje obtenido en la Prueba de Admisión que deben rendir sí influye. Mientras mayor es el puntaje obtenido, el estudiante tiene menor probabilidad de desertar.

Otro grupo de variables evaluadas fueron las condiciones iniciales del estudiante al entrar a la Universidad, como la duración ideal de la carrera, el número de créditos totales y la Facultad a la que entra el estudiante por primera vez. Después de realizar el análisis no paramétrico se observó que el número de créditos totales de la carrera y la duración ideal de la carrera en número de semestres tienen un comportamiento similar y son factores que representan lo mismo. Los estudiantes que persiguen una carrera de 8 semestres al iniciar sus estudios, son los estudiantes con menor probabilidad de desertar con respecto al resto de categorías. Sin embargo, aquellos estudiantes que siguen una carrera de 10 semestres tienen menos riesgo de desertar que aquellos que siguen una de 9. Por lo tanto, no existe una tendencia clara que demuestre que el seguir una carrera más corta o más larga puede aumentar o disminuir el riesgo de desertar. Debido a que existen 43 carreras distintas, se optó por evaluar únicamente la Facultad a la que pertenece el estudiante, en donde se observó que sí existe una diferencia. La Facultad con mayor probabilidad de supervivencia es la Escuela de Bioanálisis y en segundo lugar la Facultad de Jurisprudencia, mientras que la Facultad con menor probabilidad de supervivencia es la Facultad de Ingeniería, seguida por la Facultad de Arquitectura.

Se observa que el descuento que recibe el estudiante en el pago del semestre sí influye en la deserción. En este caso, los estudiantes que reciben un descuento entre 800 y 1.600 USD es el grupo con mayor probabilidad de permanecer en la Universidad. Los estudiantes que reciben un descuento entre 0 y 800 USD, tienen un comportamiento similar, siendo la segunda categoría con menor riesgo de deserta. Sin embargo, se observó que los estudiantes que reciben un descuento entre 3.500 y 8.000 USD son los estudiantes con mayor probabilidad de desertar; sin embargo, como se señaló anteriormente, esto puede estar explicado por las características académicas y socioeconómicas de los estudiantes con becas SNNA.

En cuanto al grupo de variables que analizan la situación socioeconómica y familiar de estudiante, se puede observar que algunas de ellas influyen en gran medida en la deserción. En la categoría socioeconómica en la que es colocado el alumno, no existe una diferencia significativa entre las curvas de riesgo, por tanto, no se puede decir que esta variable influya en la deserción. Sin embargo, si el estudiante recibió un cambio de categoría, que por lo general es un descuento mayor en el pago de su matrícula, tiene menor riesgo de desertar. Por otro lado, el tipo de colegio sí influye en la deserción, pues los estudiantes que provienen de colegios particulares, públicos y semi-públicos tienen menor probabilidad de desertar en comparación al resto de categorías. Los estudiantes con mayor riesgo de desertar son aquellos de los que no se conoce el colegio. Los estudiantes de colegios extranjeros, también tienen una gran probabilidad de desertar, sin embargo, estos casos son muy escasos, razón por la que el efecto puede no ser claro. Por último, las características del padre y de la madre que resultaron significativas en el análisis de la deserción fueron el nivel de educación, en donde se observó que mientras mayor es el nivel de educación alcanzado, mayor es la probabilidad de supervivencia del alumno. En cuanto al estado civil, se observó que el status del padre sí influye en la deserción, mientras que el de la madre no es significativo.

Por último, al evaluar el desempeño y el proceder de los estudiantes en la Universidad, se pudo ver que todos aquellos estudiantes que se cambian de carrera tienen menos riesgo de desertar en comparación a los que no se cambian. Al evaluar el promedio de créditos que toma el estudiante por semestre, se observó que los estudiantes tienen mayor probabilidad de permanecer en la institución cuando toman el número de créditos sugeridos por la Facultad, que varía entre 19 y 30 de acuerdo al semestre. Cuando toman demasiados o muy pocos créditos tienen mayor riesgo de desertar. Al evaluar el desempeño se pudo ver que el efecto del promedio general de notas es positivo, pues mientras mejores sean las notas, mayor probabilidad de supervivencia tiene el estudiante. En el caso del número de materias perdidas, el efecto fue negativo, se demostró que mientras más materias haya perdido el estudiante, menor probabilidad de permanecer en la PUCE sede Quito tiene.

CAPÍTULO II: MODELO DE LA DESERCIÓN

El análisis de Kaplan-Meier, utilizado en el capítulo anterior, permite estudiar la deserción a través de un análisis univariado de covariables de tipo categóricas. Debido a estas limitaciones, en este capítulo se utiliza el modelo de riesgos proporcionales de Cox y sus diferentes variaciones. Este tipo de modelos permiten una evaluación multidimensional de la deserción. Esta técnica permite observar la interacción entre distintas variables explicativas, así como la inclusión de variables de tipo continuas.

Los resultados obtenidos en el capítulo I junto con la revisión bibliográfica fueron la base para seleccionar un conjunto de variables que se utilizan en el modelo de Cox. Es decir, en este capítulo, se usan aquellas variables que en el modelo Kaplan-Meier mostraron diferencias significativas en la prueba estadística Log-rank test. Por lo tanto, a excepción de las variables: estado civil del estudiante, tipo de entrada a la universidad, categoría socioeconómica de la pensión diferenciada y el estado civil de la madre; las demás variables son consideradas para el análisis

En total se realizaron 4 modelos, los cuales se presentan en las Tabla 8, 10, 12 y 13. El primer modelo, corresponde al Modelo de Riesgos Proporcionales de Cox. El segundo, es un Modelo estratificado de Cox. El tercero corresponde a un modelo de Cox de covariables de tiempo dependiente y por último el cuarto es un Modelo de Cox de covariables de tiempo dependiente estratificado. Una consideración que se debe tener presente es que para todos los modelos se tomó como categoría de referencia a la menos riesgosa, según los resultados obtenidos en el modelo de Kaplan-Meier.

Modelo 1: Modelo de Riesgos Proporcionales de Cox

El Modelo de Riesgos Proporcionales Cox permite medir la tasa de riesgo de los diferentes factores y cómo la interacción entre ellos puede cambiar el efecto y/o la significancia de los mismos. La Tabla 8 presenta los resultados del primer modelo, en donde se evalúa principalmente las tasas de riesgo y la significancia de cada una de las variables.

Tabla 8: Resultados del Modelo 1

<i>Variables Explicativas</i>	<i>coef</i>	<i>exp(coef)</i>	<i>se(coef)</i>	<i>z</i>	<i>Pr(> z)</i>	<i>Sig.</i>
<i>Sexo: Hombre</i>	-0,0064	0,9936	0,0693	-0,0920	0,9267	
<i>Edad</i>	0,0284	1,0290	0,0120	2,3630	0,0182	*
<i>Residencia Previa: Desconocida</i>	-0,2600	0,7711	0,1101	-2,3620	0,0182	*
<i>Residencia Previa: Otro lugar</i>	-0,2465	0,7815	0,0993	-2,4840	0,0130	*
<i>Periodo Cohorte: 01</i>	-0,0133	0,9868	0,0736	-0,1810	0,8564	
<i>Puntaje Prueba Admisión: [0-36]</i>	0,3432	1,4090	0,1205	2,8470	0,0044	**
<i>Puntaje Prueba Admisión: [36-47]</i>	0,2736	1,3150	0,1102	2,4840	0,0130	*
<i>Puntaje Prueba Admisión: Desconocido</i>	-0,2576	0,7729	0,1234	-2,0880	0,0368	*
<i>Duración ideal: 9 semestres</i>	-0,5674	0,5670	0,1881	-3,0170	0,0026	**
<i>Duración ideal: 10 semestres</i>	-0,3135	0,7309	0,1569	-1,9980	0,0457	*
<i>Facultad: Administración</i>	0,7564	2,1310	0,2227	3,3970	0,0007	***
<i>Facultad: Arquitectura</i>	0,9037	2,4690	0,1775	5,0920	0,0000	***
<i>Facultad: Bioanálisis</i>	0,7011	2,0160	0,2852	2,4590	0,0140	*
<i>Facultad: C.Exactas & C.Naturales</i>	0,4306	1,5380	0,2336	1,8430	0,0653	.
<i>Facultad: C.Humanas</i>	0,6386	1,8940	0,1794	3,5600	0,0004	***
<i>Facultads: Comunicación</i>	0,6164	1,8520	0,1829	3,3700	0,0008	***
<i>Facultad: Economía</i>	0,5567	1,7450	0,1745	3,1900	0,0014	**
<i>Facultad: Educación</i>	0,7805	2,1820	0,4394	1,7760	0,0757	.
<i>Facultad: Enfermería</i>	-0,1286	0,8794	0,1630	-0,7890	0,4302	
<i>Facultad: Ingeniería</i>	1,1800	3,2530	0,2391	4,9330	0,0000	***
<i>Facultad: Psicología</i>	0,6451	1,9060	0,2268	2,8440	0,0045	**
<i>Facultad: Trabajo Social</i>	-0,1347	0,8740	0,4709	-0,2860	0,7748	
<i>Descuento Promedio</i>	0,0004	1,0000	0,0000	13,0790	< 2e-16	***
<i>Cambio SPD: No</i>	1,6020	4,9610	0,1909	8,3880	< 2e-16	***
<i>Tipo de Colegio: Desconocido</i>	0,4464	1,5630	0,1776	2,5130	0,0120	*
<i>Tipo de Colegio: Extranjero</i>	0,5613	1,7530	0,5892	0,9530	0,3408	
<i>Tipo de Colegio: Público</i>	-0,1860	0,8303	0,0853	-2,1820	0,0291	*
<i>Tipo de Colegio: Semi-Público</i>	0,0773	1,0800	0,0965	0,8010	0,4233	
<i>Nivel Educación Madre: Desconocido</i>	0,1130	1,1200	0,1568	0,7200	0,4713	
<i>Nivel Educación Madre: Primero</i>	-0,0666	0,9356	0,1276	-0,5210	0,6021	
<i>Nivel Educación Madre: Segundo</i>	0,0879	1,0920	0,0755	1,1640	0,2445	
<i>Nivel Educación Padre: Desconocido</i>	0,1682	1,1830	0,2035	0,8260	0,4086	
<i>Nivel Educación Padre: Primero</i>	0,0254	1,0260	0,1257	0,2020	0,8395	
<i>Nivel Educación Padre: Segundo</i>	0,0671	1,0690	0,0776	0,8650	0,3873	
<i>Estado Civil Padre: Desconocido</i>	-0,1175	0,8892	0,1688	-0,6960	0,4865	
<i>Estado Civil Padre: Divorciado</i>	-0,1726	0,8415	0,0964	-1,7910	0,0733	.
<i>Estado Civil Padre: Soltero</i>	0,1997	1,2210	0,1555	1,2840	0,1990	
<i>Estado Civil Padre: Viudo</i>	-0,1215	0,8856	0,3179	-0,3820	0,7022	
<i>Cmbio de Carrera: No</i>	1,2850	3,6160	0,1794	7,1650	0,0000	***
<i>Rango Créditos: [30-40]</i>	0,5610	1,7520	0,1350	4,1550	0,0000	***
<i>Rango Créditos: [1-19]</i>	0,2295	1,2580	0,0987	2,3240	0,0201	*
<i>Promedio General Notas</i>	-0,2083	0,8119	0,0057	-36,5180	< 2e-16	***
<i>Total Materias Perdidas</i>	-0,2000	0,8187	0,0130	-15,3510	< 2e-16	***

Elaborado por: Autor.

Nota: (.) significativa al 90% de confianza. (*) significativa al 95% de confianza. (**) significativa al 99% de confianza. (***) significativa al 99,9% de confianza.

Nota: Para comprobar la ecuación del modelo se puede consultar el Anexo B.

Este modelo debe cumplir con la condición de riesgos proporcionales, lo que significa que las tasas de riesgo deben mantenerse estables en el tiempo. Para comprobar si este supuesto se cumple se utiliza la prueba de Schoenfeld. Es necesario recordar que cuando se rechaza la hipótesis nula, la condición de proporcionalidad no se cumple, mientras que cuando no se rechaza sí se cumple.

Tabla 9: Prueba de residuos de Schoenfeld del Modelo 1

Variables	chisq	df	p	H0	Condición
Sexo	17,84	1	0,0000	Rechazo la H0	No Cumple
Edad	7,91	1	0,0049	Rechazo la H0	No Cumple
Residencia Previa	7,47	2	0,0239	Rechazo la H0	No Cumple
Periodo Cohorte	2,88	1	0,0896	No rechazo la H0	Cumple
Puntaje Prueba Admisión	2,08	3	0,5554	No rechazo la H0	Cumple
Duración ideal	12,29	2	0,0021	Rechazo la H0	No Cumple
Facultad	26,02	12	0,0107	Rechazo la H0	No Cumple
Descuento Promedio	45,17	1	0,0000	Rechazo la H0	No Cumple
Cambio SPD	8,55	1	0,0035	Rechazo la H0	No Cumple
Tipo de Colegio	3,62	4	0,4592	No rechazo la H0	Cumple
Nivel Educación Madre	4,06	3	0,2546	No rechazo la H0	Cumple
Nivel Educación Padre	3,95	3	0,2669	No rechazo la H0	Cumple
Estado Civil Padre	3,66	4	0,4544	No rechazo la H0	Cumple
Cambio de Carrera	16,97	1	0,0000	Rechazo la H0	No Cumple
Rango Créditos	117,07	2	< 2e-16	Rechazo la H0	No Cumple
Promedio General Notas	249,58	1	< 2e-16	Rechazo la H0	No Cumple
Total Materias Perdidas	461,92	1	< 2e-16	Rechazo la H0	No Cumple
Global	622,64	43	< 2e-16	Rechazo la H0	No Cumple

Elaborado por: Autor.

Nota: Ver el Anexo B si se desea analizar el gráfico.

Como se observa en la Tabla 9, once de las diecisiete variables no cumplen con la condición de proporcionalidad. Al comparar estos resultados con los gráficos del Modelos de Kaplan Meier, se puede observar que todas aquellas variables que incumplen con la condición también lo hacen a través del método gráfico, pues las curvas de las categorías se entrecruzan, sin mantener una tasa proporcional de crecimiento/decrecimiento. Conjuntamente, al evaluar el modelo global, este tampoco cumple. Si bien no es necesario que todas las variables cumplan con la condición de proporcionalidad, sí es necesario que la mayoría de ellas lo hagan. Debido a la violación del supuesto, se puede concluir que este modelo no es adecuado para estimar la deserción de los estudiantes de la PUCE sede Quito.

Modelo 2: Modelo Estratificado de Cox

Debido al incumplimiento de la condición de proporcionalidad se optó por probar con una variación del modelo anterior, el modelo estratificado de Cox. Este modelo transfiere el efecto de las variables estratificadas al riesgo de referencia. Las covariables seleccionadas para estratificar deben tener ciertas características. En primer lugar, no deben cumplir con la condición de riesgos proporcionales, deben ser variables categóricas y deben tener características similares dentro de cada categoría, pero diferentes

entre los grupos. A pesar de existir múltiples variables que incumplen la condición, se estratificó la variable Facultad porque ésta muestra una diferencia significativa entre sus categorías, pues como se observa en el gráfico de Kaplan Meier, el riesgo de cada Facultad cambia en cada semestre, por este motivo se encuentran diferentes cruces entre las curvas. Además, existen varios estudios que comprueban que existen características propias de cada Facultad que pueden contribuir a la deserción de los estudiantes (Ver, Restaino,2008; Hedge & Prageeth, 2018).

A continuación, en la Tabla 10 se presentan los resultados del modelo estimado. Este tendrá una evaluación y comparación con respecto al anterior modelo, pues de esta manera se podrá evaluar si la nueva técnica mejora o no la predicción de la deserción.

Tabla 10: Resultados del Modelo 2

<i>Variables Explicativas</i>	<i>coef</i>	<i>exp(coef)</i>	<i>se(coef)</i>	<i>z</i>	<i>Pr(> z)</i>	
<i>Sexo: Hombre</i>	-0,0001	0,9999	0,0695	-0,0010	0,9989	
<i>Edad</i>	0,0311	1,0320	0,0119	2,6080	0,0091	**
<i>Residencia Previa: Desconocida</i>	-0,2498	0,7790	0,1104	-2,2630	0,0236	*
<i>Residencia Previa: Otro lugar</i>	-0,2432	0,7841	0,0996	-2,4430	0,0146	*
<i>Periodo Cohorte: 01</i>	-0,0179	0,9823	0,0737	-0,2420	0,8086	
<i>Puntaje Prueba Admisión: [0-36[</i>	0,3526	1,4230	0,1209	2,9170	0,0035	**
<i>Puntaje Prueba Admisión: [36-47[</i>	0,2714	1,3120	0,1105	2,4570	0,0140	*
<i>Puntaje Prueba Admisión: Desconocido</i>	-0,2278	0,7963	0,1236	-1,8430	0,0654	.
<i>Duración ideal: 9 semestres</i>	-0,5878	0,5555	0,1891	-3,1080	0,0019	**
<i>Duración ideal: 10 semestres</i>	-0,3199	0,7262	0,1588	-2,0140	0,0440	*
<i>Descuento Promedio</i>	0,0004	1,0000	0,0000	12,9610	< 2e-16	***
<i>Cambio SPD: No</i>	1,5990	4,9490	0,1910	8,3710	< 2e-16	***
<i>Tipo de Colegio: Desconocido</i>	0,4362	1,5470	0,1790	2,4370	0,0148	*
<i>Tipo de Colegio: Extranjero</i>	0,5430	1,7210	0,5896	0,9210	0,3571	
<i>Tipo de Colegio: Público</i>	-0,1945	0,8232	0,0859	-2,2660	0,0235	*
<i>Tipo de Colegio: Semi-Público</i>	0,0612	1,0630	0,0968	0,6320	0,5272	
<i>Nivel Educación Madre: Desconocido</i>	0,1288	1,1370	0,1566	0,8230	0,4107	
<i>Nivel Educación Madre: Primero</i>	-0,0789	0,9241	0,1283	-0,6150	0,5384	
<i>Nivel Educación Madre: Segundo</i>	0,0852	1,0890	0,0756	1,1270	0,2599	
<i>Nivel Educación Padre: Desconocido</i>	0,1782	1,1950	0,2051	0,8690	0,3850	
<i>Nivel Educación Padre: Primero</i>	0,0452	1,0460	0,1259	0,3590	0,7196	
<i>Nivel Educación Padre: Segundo</i>	0,0671	1,0690	0,0776	0,8640	0,3875	
<i>Estado Civil Padre: Desconocido</i>	-0,1192	0,8877	0,1707	-0,6980	0,4852	
<i>Estado Civil Padre: Divorciado</i>	-0,1779	0,8370	0,0965	-1,8430	0,0653	.
<i>Estado Civil Padre: Soltero</i>	0,1995	1,2210	0,1557	1,2820	0,2000	
<i>Estado Civil Padre: Viudo</i>	-0,1908	0,8263	0,3229	-0,5910	0,5545	
<i>Cambio de Carrera: No</i>	1,2680	3,5550	0,1798	7,0530	0,0000	***
<i>Rango Créditos: [30-40]</i>	0,5763	1,7790	0,1354	4,2550	0,0000	***
<i>Rango Créditos: [1-19]</i>	0,2111	1,2350	0,0996	2,1200	0,0340	*
<i>Promedio General Notas</i>	-0,2094	0,8111	0,0059	-35,7070	< 2e-16	***
<i>Total Materias Perdidas</i>	-0,2016	0,8174	0,0131	-15,3570	< 2e-16	***

Elaborado por: Autor.

Nota: (.) significativa al 90% de confianza. (*) significativa al 95% de confianza. (**) significativa al 99% de confianza. (***) significativa al 99,9% de confianza

Nota: Para revisar la ecuación del modelo se puede consultar el Anexo C.

Al evaluar las tasas de riesgo de este modelo, se observa que no existe una gran variación de los coeficientes ni en la significancia de las variables. Después de aplicar la estratificación, el modelo debería cumplir con la condición de proporcionalidad, por tanto, es necesario aplicar la prueba de Schoenfeld al Modelo 2.

Tabla 11: Prueba de residuos de Schoenfeld del Modelo 2

Variabes	chisq	df	p	H0	Condición
Sexo	19,17	1	0,0000	Rechazo la H0	No Cumple
Edad	8,36	1	0,0038	Rechazo la H0	No Cumple
Residencia Previa	6,51	2	0,0386	Rechazo la H0	No Cumple
Periodo Cohorte	2,91	1	0,0879	No rechazo la H0	Cumple
Puntaje Prueba Admisión	3,46	3	0,3256	No rechazo la H0	Cumple
Duración ideal	17,01	2	0,0002	Rechazo la H0	No Cumple
Descuento Promedio	45,60	1	0,0000	Rechazo la H0	No Cumple
Cambio SPD	8,01	1	0,0046	Rechazo la H0	No Cumple
Tipo de Colegio	3,46	4	0,4836	No rechazo la H0	Cumple
Nivel Educación Madre	5,32	3	0,1497	No rechazo la H0	Cumple
Nivel Educación Padre	4,46	3	0,2162	No rechazo la H0	Cumple
Estado Civil Padre	2,95	4	0,5656	No rechazo la H0	Cumple
Cambio de Carrera	19,03	1	0,0000	Rechazo la H0	No Cumple
Rango Créditos	122,57	2	< 2e-16	Rechazo la H0	No Cumple
Promedio General Notas	236,08	1	< 2e-16	Rechazo la H0	No Cumple
Total Materias Perdidas	436,35	1	< 2e-16	Rechazo la H0	No Cumple
Global	571,96	31	< 2e-16	Rechazo la H0	No Cumple

Elaborado por: Autor.

Nota: Ver el Anexo C si se desea analizar el gráfico.

Al observar la Tabla 11 de resultados de la prueba de residuos de Schoenfeld, se observa que persiste el incumplimiento de la condición de proporcionalidad en múltiples variables y en el modelo global. Por este motivo, este modelo tampoco permite estimar el riesgo de deserción de manera adecuada. Es necesario mencionar que se realizaron múltiples pruebas en donde se estratificaron diferentes variables para ver si el modelo mejoraba, sin embargo, el incumplimiento de la condición persistía. La situación presentada puede deberse a que la naturaleza propia de los datos incumple con la condición de proporcionalidad. Por este motivo es necesario probar otro modelo.

Modelo 3: Modelo de Covariables de Tiempo Dependiente

Debido al incumplimiento de la condición de riesgos proporcionales, se probó con el modelo de covariables de tiempo dependiente, el cual no está sujeto al supuesto de proporcionalidad. Para este modelo se utilizará información longitudinal, ya que este es uno de los requerimientos. En este caso, las variables con este tipo de información son: promedio general de notas por semestre y materias perdidas por semestre. A continuación, se presentan los resultados del Modelo 3.

Tabla 12: Resultados del Modelo 3

<i>Variables Explicativas</i>	<i>coef</i>	<i>exp(coef)</i>	<i>se(coef)</i>	<i>z</i>	<i>Pr(> z)</i>	
<i>Sexo: Hombre</i>	-0,0641	0,9379	0,0689	-0,9390	0,3477	
<i>Edad</i>	0,0389	1,0400	0,0122	3,0330	0,0024	**
<i>Residencia Previa: Desconocida</i>	-0,2028	0,8164	0,1097	-1,8810	0,0599	.
<i>Residencia Previa: Otro lugar</i>	-0,1815	0,8340	0,0966	-1,8910	0,0586	.
<i>Periodo Cohorte: 01</i>	-0,0034	0,9966	0,0725	-0,0460	0,9631	
<i>Puntaje Prueba Admisión: [0-36]</i>	0,2973	1,3460	0,1188	2,5310	0,0114	*
<i>Puntaje Prueba Admisión: [36-47]</i>	0,2435	1,2760	0,1100	2,2520	0,0243	*
<i>Puntaje Prueba Admisión: Desconocido</i>	-0,1509	0,8599	0,1226	-1,2180	0,2230	
<i>Duración ideal: 9 semestres</i>	-0,3832	0,6817	0,1780	-2,0750	0,0380	*
<i>Duración ideal: 10 semestres</i>	-0,0948	0,9096	0,1529	-0,5600	0,5755	
<i>Facultad: Administración</i>	0,7375	2,0910	0,2199	3,2090	0,0013	**
<i>Facultad: Arquitectura</i>	0,6589	1,9330	0,1796	3,3610	0,0008	***
<i>Facultad: Bioanálisis</i>	0,8594	2,3620	0,2801	3,0810	0,0021	**
<i>Facultad: C.Exactas & C.Naturales</i>	0,3850	1,4700	0,2362	1,6690	0,0950	.
<i>Facultad: C.Humanas</i>	0,7613	2,1410	0,1821	4,1720	0,0000	***
<i>Facultad: Comunicación</i>	0,7948	2,2140	0,1830	4,4970	0,0000	***
<i>Facultad: Economía</i>	0,6837	1,9810	0,1762	3,9030	0,0001	***
<i>Facultad: Educación</i>	0,7988	2,2230	0,4379	2,0070	0,0448	*
<i>Facultad: Enfermería</i>	0,0473	1,0480	0,1629	0,2870	0,7743	
<i>Facultad: Ingeniería</i>	1,2290	3,4190	0,2315	5,0880	0,0000	***
<i>Facultad: Psicología</i>	0,7513	2,1200	0,2240	3,1750	0,0015	**
<i>Facultad: Trabajo Social</i>	-0,3480	0,7061	0,4721	-0,8610	0,3893	
<i>Descuento Promedio</i>	0,0005	1,0000	0,0000	12,2870	< 2e-16	***
<i>Cambio SPD: No</i>	1,8270	6,2170	0,1903	10,3230	< 2e-16	***
<i>Tipo de Colegio: Desconocido</i>	0,4835	1,6220	0,1802	2,7320	0,0063	**
<i>Tipo de Colegio: Extranjero</i>	0,4619	1,5870	0,5888	0,8310	0,4060	
<i>Tipo de Colegio: Público</i>	-0,2022	0,8170	0,0852	-2,3410	0,0193	*
<i>Tipo de Colegio: Semi-Público</i>	0,0139	1,0140	0,0965	0,1510	0,8796	
<i>Nivel Educación Madre: Desconocido</i>	0,2828	1,3270	0,1516	1,8890	0,0589	.
<i>Nivel Educación Madre: Primero</i>	-0,1659	0,8471	0,1280	-1,2030	0,2290	
<i>Nivel Educación Madre: Segundo</i>	0,0973	1,1020	0,0746	1,3100	0,1903	
<i>Nivel Educación Padre: Desconocido</i>	0,2508	1,2850	0,1962	1,2900	0,1969	
<i>Nivel Educación Padre: Primero</i>	-0,1856	0,8306	0,1241	-1,3310	0,1832	
<i>Nivel Educación Padre: Segundo</i>	0,0415	1,0420	0,0764	0,5650	0,5718	
<i>Estado Civil Padre: Desconocido</i>	-0,1903	0,8267	0,1632	-1,2480	0,2119	
<i>Estado Civil Padre: Divorciado</i>	-0,2149	0,8066	0,0961	-2,2780	0,0227	*
<i>Estado Civil Padre: Soltero</i>	0,2597	1,2970	0,1546	1,8480	0,0646	.
<i>Estado Civil Padre: Viudo</i>	0,0059	1,0060	0,3139	0,0210	0,9829	
<i>Cambio de Carrera: No</i>	1,8360	6,2690	0,1785	10,3890	< 2e-16	***
<i>Rango Créditos: [30-40]</i>	0,6837	1,9810	0,1345	5,3300	0,0000	***
<i>Rango Créditos: [1-19]</i>	0,2155	1,2400	0,0993	2,2260	0,0260	*
<i>Promedio General Notas</i>	-0,0896	0,9143	0,0053	-16,076	< 2e-16	***
<i>Total Materias Perdidas</i>	0,1679	1,1830	0,0236	6,7950	0,0000	***

Elaborado por: Autor.

Nota: (.) significativa al 90% de confianza. (*) significativa al 95% de confianza. (**) significativa al 99% de confianza. (***) significativa al 99,9% de confianza.

Nota: Para comprobar la ecuación del modelo se puede consultar el Anexo D.

En este modelo se pudo observar que la mayoría de variables resultaron significativas y en la mayoría de casos se obtuvo un comportamiento similar al obtenido en las curvas de Kaplan Meier, sin embargo, en algunos casos la magnitud del riesgo aumentó y en otros casos disminuyó.

Modelo 4: Modelo de Covariables de Tiempo Dependiente Estratificado

Es necesario mencionar que varias fuentes bibliográficas establecen que se puede utilizar una o varias metodologías al estimar un modelo de supervivencia semiparamétricos, todo dependerá del tipo de información que se tenga (Date, 2020). Finalmente, se probó un último modelo, el Modelo 4, en donde se combinó la técnica de covariables de tiempo dependiente y la estratificación de la variable Facultad. Aunque el Modelo 3 parecía tener las condiciones necesarias para ser el modelo que mejor se ajustaba al estudio de la deserción en la PUCE sede Quito, se optó por probar con una combinación de ambas técnicas debido a que la variable Facultad tiene variaciones significativas entre sus diferentes categorías.

Tabla 13: Resultados del Modelo 4

<i>Variables Explicativas</i>	<i>coef</i>	<i>exp(coef)</i>	<i>se(coef)</i>	<i>z</i>	<i>Pr(> z)</i>	
<i>Sexo: Hombre</i>	-0,0506	0,9506	0,0693	-0,7480	0,4544	
<i>Edad</i>	0,0393	1,0400	0,0122	3,1550	0,0016	**
<i>Residencia Previa: Desconocida</i>	-0,1912	0,8260	0,1100	-1,7880	0,0738	.
<i>Residencia Previa: Otro lugar</i>	-0,1870	0,8295	0,0969	-1,9350	0,0530	.
<i>Periodo Cohorte: 01</i>	-0,0089	0,9912	0,0727	-0,1190	0,9054	
<i>Puntaje Prueba Admisión: [0-36]</i>	0,2961	1,3450	0,1191	2,4970	0,0125	*
<i>Puntaje Prueba Admisión: [36-47]</i>	0,2506	1,2850	0,1103	2,2910	0,0220	*
<i>Puntaje Prueba Admisión: Desconocido</i>	-0,1391	0,8702	0,1230	-1,1100	0,2671	
<i>Duración ideal: 9 semestres</i>	-0,3886	0,6780	0,1790	-2,1210	0,0339	*
<i>Duración ideal: 10 semestres</i>	-0,0911	0,9129	0,1548	-0,5520	0,5808	
<i>Descuento Promedio</i>	0,0005	1,0000	0,0000	12,2620	< 2e-16	***
<i>Cambio SPD: No</i>	1,8250	6,2040	0,1904	10,3570	< 2e-16	***
<i>Tipo de Colegio: Desconocido</i>	0,5179	1,6790	0,1812	3,0310	0,0024	**
<i>Tipo de Colegio: Extranjero</i>	0,4781	1,6130	0,5892	0,8620	0,3887	
<i>Tipo de Colegio: Público</i>	-0,1936	0,8239	0,0856	-2,2450	0,0248	*
<i>Tipo de Colegio: Semi-Público</i>	0,0026	1,0030	0,0968	0,0290	0,9772	
<i>Nivel Educación Madre: Desconocido</i>	0,2508	1,2850	0,1517	1,6780	0,0934	.
<i>Nivel Educación Madre: Primero</i>	-0,1918	0,8254	0,1286	-1,4060	0,1596	
<i>Nivel Educación Madre: Segundo</i>	0,0721	1,0750	0,0749	0,9810	0,3267	
<i>Nivel Educación Padre: Desconocido</i>	0,2984	1,3480	0,1982	1,5700	0,1164	
<i>Nivel Educación Padre: Primero</i>	-0,1597	0,8524	0,1247	-1,1720	0,2414	
<i>Nivel Educación Padre: Segundo</i>	0,0515	1,0530	0,0766	0,7060	0,4800	
<i>Nivel Educación Padre: Desconocido</i>	-0,2251	0,7985	0,1654	-1,5110	0,1309	
<i>Estado Civil Padre: Divorciado</i>	-0,2190	0,8034	0,0964	-2,3210	0,0203	*
<i>Estado Civil Padre: Soltero</i>	0,2753	1,3170	0,1549	2,0010	0,0455	*
<i>Estado Civil Padre: Viudo</i>	-0,0493	0,9519	0,3185	-0,1720	0,8634	
<i>Cmbio de Carrera: No</i>	1,8370	6,2800	0,1793	10,2970	< 2e-16	***
<i>Rango Créditos: [30-40]</i>	0,7137	2,0420	0,1358	5,5830	0,0000	***
<i>Rango Créditos: [1-19]</i>	0,2444	1,2770	0,0994	2,5600	0,0105	*
<i>Promedio General Notas</i>	-0,0887	0,9151	0,0056	-15,4470	< 2e-16	***
<i>Total Materias Perdidas</i>	0,1742	1,1900	0,0248	6,7800	0,0000	***

Elaborado por: Autor.

Nota: (.) significativa al 90% de confianza. (*) significativa al 95% de confianza. (**) significativa al 99% de confianza. (***) significativa al 99,9% de confianza.

Nota: Para comprobar la ecuación del modelo se puede consultar el Anexo E.

La significancia de las variables permaneció igual a la del Modelo 3 y los coeficientes de las tasas de riesgo cambiaron muy poco en algunos casos y en otros permanecieron igual.

Estadísticos de Bondad De Ajuste

Tabla 14: Estadísticos de bondad de ajuste

	Modelo 1	Modelo 2	Modelo 3	Modelo 4
Razón de verosimilitud	2064 ($p < 2e-16$)	1943 ($p < 2e-16$)	2002 ($p < 2e-16$)	1882 ($p < 2e-16$)
Concordancia	0,868	0,866	0,846	0,839
AIC	15943,45	11126,62	16004,72	11118,22

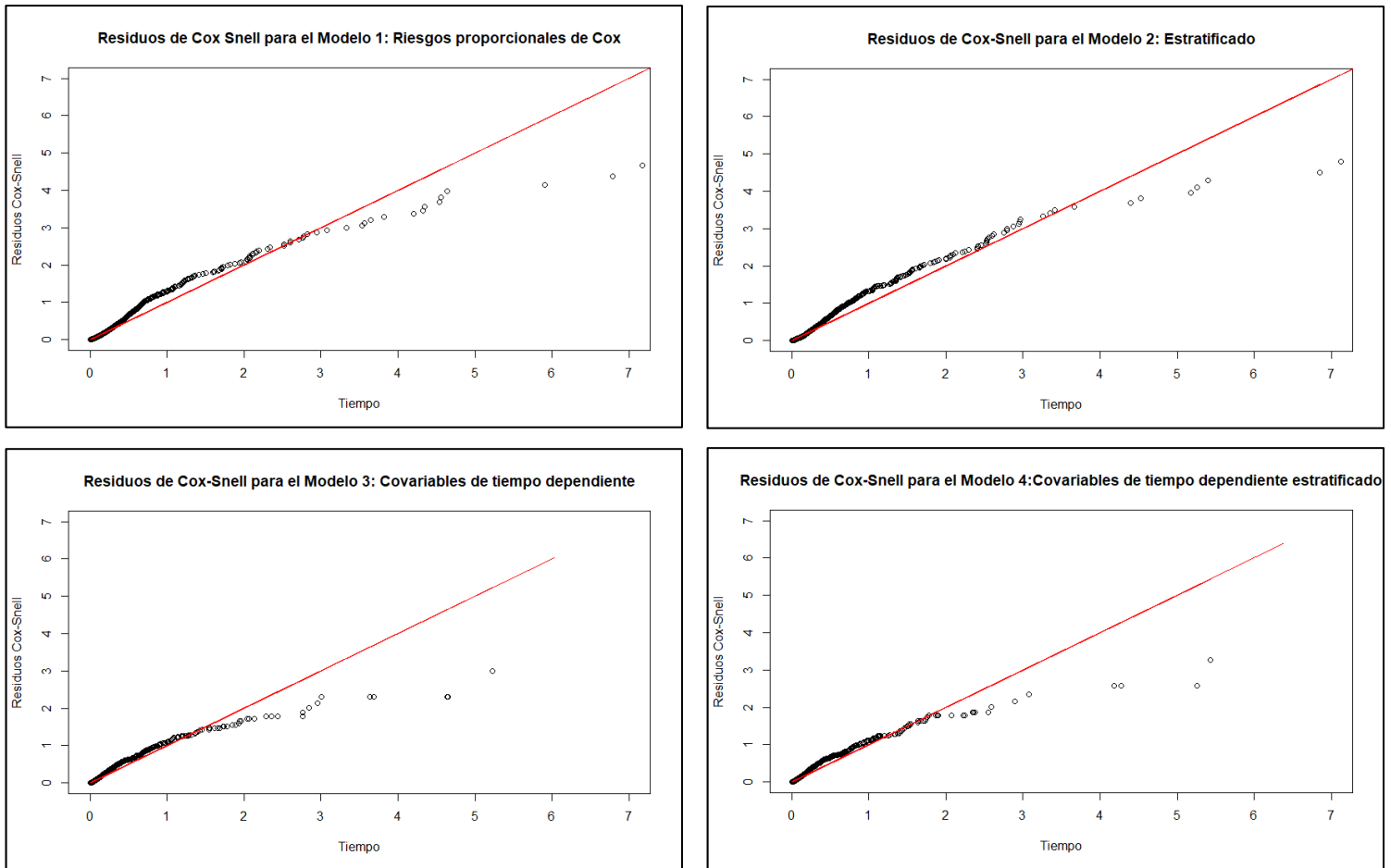
Elaborado por: Autor.

El p-valor de la razón de verosimilitud nos muestra que todos los modelos son globalmente significativos. Por otro lado, la concordancia es un indicador que calcula la correspondencia entre las respuestas observadas y predichas. Como se puede apreciar en la Tabla 14, todos tienen valores cercanos por encima de 83,9%, siendo el Modelo 1 el que presenta el valor más alto de 86,8%. Finalmente, el criterio de información de Akaike (AIC) es mejor mientras menor sea su valor, en este caso el menor valor lo presenta el Modelo 4.

En conclusión, todos los modelos presentan buenos indicadores de ajuste; sin embargo, los Modelos 1 y 2 violan supuestos importantes, mientras que los Modelos 3 y 4 no requieren el cumplimiento de dicho supuesto.

Prueba de Residuos de Cox-Snell

Gráfico 27: Prueba Cox-Snell de los Modelos.



Elaborado por: Autor.

La prueba de residuos de Cox Snell¹¹ permite comprobar el ajuste del modelo, mostrando que mientras los residuos más se acercan a la línea de 45°, mejor predecirá el modelo. Como se observa en el Gráfico 27, los cuatro modelos son muy parecidos y existe una mayor aglomeración de los residuos en cada uno de los modelos durante los primeros semestres, lo cual concuerda con la información que se tiene, en donde se muestra que aproximadamente el 70% de los estudiantes deserten durante los primeros dos semestres.

Al observar el Modelo 1 y el Modelo 2, se puede ver que tienen una superposición perfecta en cero, sin embargo, antes de llegar al primer semestre, los residuos se alejan de la recta. El gráfico de residuos de Cox-Snell demuestra que los Modelo 1 y 2 pueden predecir la deserción por periodos más largos que los Modelo 3 y 4, sin embargo, el pronóstico no será preciso. Los residuos del Modelo 3 y 4 se ajustan de manera casi perfecta a la recta de 45°. Sin embargo, una de las principales diferencias entre ambos

¹¹ Para mayor información por favor consultar en la Sección de Modelos de Supervivencia ubicada en la Fundamentación Teórica.

modelos es que el Modelo 4 mantiene la tendencia y claramente se puede ver que la sobreposición de los residuos sobre la curva permanece hasta el segundo periodo, mientras que la del Modelo 3 solo se mantiene durante el primer periodo. El ajuste de ambos modelos es muy bueno, no obstante, se selecciona el Modelo 4 ya que este puede predecir con mayor precisión durante un periodo de tiempo más largo. Es necesario mencionar, que tanto el Modelo 3 como el Modelo 4 pueden predecir la deserción únicamente a corto plazo (durante los primeros semestres).

Resultados

A continuación, se evaluará los resultados del Modelo 4 y se analizará aquellas variables que resultaron significativas al 95% de confianza.

En el grupo de características del estudiante, la única variable que resultó significativa fue la variable edad. Se observa que por cada año adicional que el estudiante tenga cuando ingresa a la universidad, el riesgo de desertar aumenta en 4%. El riesgo puede estar ligado a otros factores que ocasionaron que los estudiantes entren con más años a la Universidad y decidan desertar. Diferentes estudios han comprobado que existe una mayor probabilidad de desertar cuando el estudiante es mayor de 19 años al momento de inscribirse por primera vez en una universidad (Miller, 2019). Esta situación puede deberse a una serie de factores tales como: el factor psicológico, pues al no convivir con gente de su edad, el estudiante puede no adaptarse de manera adecuada; la situación económica, pues probablemente el estudiante debe trabajar para poder pagar sus estudios; otras obligaciones como hijos o algún familiar a cargo; múltiples cambios de carrera; otros intereses; etc. (Miller, 2019; Sandoval-Palis, Naranjo, Vidal, & Gilar-Corbi, 2020).

En cuanto a las condiciones iniciales, cuando el estudiante ha obtenido un puntaje entre 0-36 puntos en la prueba de admisión, la probabilidad de desertar del estudiante aumenta en un 35% con respecto a aquellos alumnos que han obtenido un puntaje superior a 47 puntos. Cuando el estudiante ha obtenido un puntaje entre 36-47 puntos en la prueba de admisión, la probabilidad de desertar del estudiante aumenta en un 29% con respecto a aquellos alumnos que han obtenido un puntaje entre 47-70 puntos. La categoría puntaje de admisión desconocido no es significativa. Esta variable demuestra que el conocimiento adquirido previamente al inicio de la universidad sí es importante. La relevancia de este factor puede tener una gran influencia durante los primeros años.

Por otro lado, al evaluar la duración ideal de la carrera se observa que la tasa de riesgo de desertar para los alumnos que siguen una carrera de 9 semestres es 32% menor a aquellos que siguen una carrera con una duración de 8 semestres. Para los alumnos que siguen una carrera de 10 semestres, el riesgo de desertar es 9% menor en comparación a los estudiantes que siguen una carrera de 8 semestres. La duración ideal es una variable significativa al 95% de confianza para los estudiantes que siguen una carrera de 9 semestres y no es significativa para aquellos estudiantes que siguen una carrera de 10 semestres. La significancia y los efectos de estas categorías muestran un comportamiento diferente al modelo de Kaplan Meier. Esto podría demostrar que no existe un impacto claro de la duración ideal de la carrera sobre la deserción.

Al analizar los factores socioeconómicos y familiares, resalta el hecho de que la variable descuento promedio sea significativa pero que su efecto sea nulo. Por cada dólar adicional de descuento que recibe el estudiante en el pago por semestre, la probabilidad de desertar aumenta en 0%, lo que significa que el monto de descuento que recibe el estudiante por semestre no genera menor o mayor riesgo. Es necesario mencionar que esta variable también fue categorizada y se vio como cambiaba su efecto en el modelo, sin embargo, el efecto fue similar en algunas de sus categorías y perdió significancia. Por este motivo se decidió mantener la variable continua. En comparación al modelo de Kaplan Meier, el comportamiento cambió y esto pudo darse por la interacción con otras variables, pues hay otros factores que tienen mayor peso y que anulan su efecto sobre la deserción.

Por otro lado, se observó que los estudiantes que no obtienen un cambio en la Categoría Socioeconómica de Pensión Diferenciada, tienen una probabilidad de desertar 6,20 veces mayor, frente a los estudiantes que sí reciben un cambio en la categoría de pensión diferenciada. Esta variable demuestra que el recibir una degradación de la categoría socioeconómica se ve reflejada en un beneficio monetario que reduce la probabilidad de desertar de los estudiantes. Otra explicación para el comportamiento de esta variable podría ser que el cambio de categoría fue otorgado a alumnos que verdaderamente demostraban necesidad económica y estos alumnos piensan que en otras universidades no tendrán el mismo beneficio, por lo tanto, no desertan de la PUCE sede Quito.

En cuanto al tipo de colegio, se puede ver que cuando se desconoce el tipo de colegio del que se graduó el alumno, el riesgo de desertar aumenta en un 68% en comparación a los estudiantes que se graduaron en colegios privados. Cuando el estudiante se gradúa en un colegio público, el riesgo de desertar es 18% menor al de los alumnos que se graduaron en colegios privados. Estas dos categorías resultaron significativas, sin embargo, las categorías de colegio extranjero y semi-público no lo fueron. El tipo de colegio desconocido debería ser una alerta para las autoridades de la universidad, ya que la ausencia de esta información puede deberse a problemas más serios. En cuanto a los estudiantes que provienen de colegios públicos, puede que ellos tengan menor probabilidad de desertar, ya que valoran en mayor medida la educación privada y el desertar podría representar un costo de oportunidad más alto.

Las categorías divorciado y soltero en la variable estado civil del padre resultaron significativas, por lo tanto, cuando el padre del alumno está divorciado, el riesgo de desertar del estudiante es 20% menor que la de los estudiantes con padres casados. Cuando el padre del alumno es soltero, el riesgo de desertar del estudiante es 32% mayor que la de los estudiantes con padres casados. La relevancia de estas variables se mantiene igual que en el Capítulo I, por lo tanto, la interacción entre variables no afecta su comportamiento.

Por último, en cuanto a las variables de desempeño, se observa que los estudiantes que no se cambian de carrera, tienen un riesgo de desertar 6,28 veces mayor que los estudiantes que sí se cambian de carrera. El resultado demuestra que el cambio de carrera influye negativamente en la deserción. Por otro lado, en cuanto al rango promedio de créditos que toma el estudiante por semestre, se demuestra que cuando toma demasiados créditos, entre 30 y 40, el riesgo de desertar es 2,04 veces mayor al de los estudiantes que toman entre 19 y 30 créditos por semestre. Cuando el estudiante toma pocos créditos, entre 1 y 19 por semestre, el riesgo de desertar es 28% mayor que la de los estudiantes que toman entre 19 y 30. Es necesario reconocer que esta variable era continua, sin embargo, fue necesario categorizarla

para visualizar su efecto. Al utilizar la variable continua el efecto se anula debido a que el riesgo de desertar aumenta cuando el estudiante toma pocos o demasiados créditos.

Al analizar el promedio de notas por semestre, se observa que por cada punto adicional que el estudiante obtenga en el promedio de notas por semestre, el riesgo de desertar disminuye en 8%. Por lo tanto, se puede observar que mientras menor es el promedio de notas por semestre, mayor es la probabilidad de desertar. Al mismo tiempo, si se analiza el número de materias perdidas por semestre, se observa que por cada materia adicional que pierde el alumno por semestre, la probabilidad de desertar aumenta en 19%. El comportamiento de esta variable se mantiene y demuestra que puede tener un efecto grande en la deserción. Estas variables son definitivamente uno de los mejores estimadores del desempeño del estudiante y pueden llegar a ser buenos predictores de la deserción.

Finalmente, es necesario mencionar que la variable Facultad no fue tomada en cuenta, ya que al estratificar el modelo pasó a formar parte del riesgo de referencia. Además, existen algunas variables que al introducirlas en el modelo perdieron significancia y en algunos casos sus efectos cambiaron en comparación a los resultados obtenidos en el modelo de Kaplan-Meier. Las variables que no resultaron significativas fueron: el sexo del estudiante, la residencia previa, el periodo de entrada, el nivel de educación de la madre y el nivel de educación del padre. Estos factores se relacionan con las características de los estudiantes y las condiciones iniciales al iniciar la Universidad, por tanto, se deduce que la interacción entre variables reduce la relevancia de estas variables.

CONCLUSIONES

Esta investigación se destinó al estudio de la deserción en la PUCE sede Quito. Para realizarla se utilizó la metodología de los Modelos de Supervivencia, en donde los principales métodos usados fueron el Modelo de Kaplan-Meier y algunas variantes del Modelo de Riesgo Proporcional de Cox. Estos procedimientos fueron las herramientas principales para analizar la probabilidad de desertar de los estudiantes. A través de estos métodos se establecieron los principales factores que influyeron en el abandono de los estudiantes de las cohortes 2014-02 hasta 2017-02. A través de estos modelos se logró identificar las características de los potenciales desertores y predecir el abandono a través de los perfiles de los estudiantes que presentan un mayor riesgo.

Al igual que en diversas investigaciones sobre la temática, en este estudio también se comprueba que el mayor porcentaje de deserción ocurre durante los primeros semestres. En este caso, el porcentaje de alumnos que desertan definitivamente de la institución constituye el 17%; de este grupo el 70% abandonan la Universidad durante el primer y segundo semestre. Los resultados se anulan totalmente después del sexto semestre. Una de las principales explicaciones de este comportamiento se debe al costo de oportunidad que representa el retirarse de la carrera, pues mientras más alto es el semestre del que deserta el estudiante, más dinero y tiempo se habrá invertido.

En este trabajo de investigación, se encontraron distintas dimensiones que contribuyen a la deserción de los estudiantes. Entre los grupos más importantes de variables se encuentran las características propias de los estudiantes, las condiciones iniciales al entrar a la Universidad, los factores económicos y las variables de desempeño académico. Cada una de ellas contribuye en mayor o menor medida a la deserción y construye el perfil de los desertores.

En primera instancia se encontró que el grupo de variables de desempeño académico fueron las variables más relevantes en la construcción del modelo. Todas las variables que conformaban este grupo resultaron significativas. En los resultados se pudo observar que el cambio de carrera y el promedio de notas tienen una relación negativa sobre la deserción, pues mientras mejor sea el promedio y cuando el estudiante se ha cambiado de carrera, la probabilidad de desertar disminuye. Por otro lado, el número de materias perdidas tienen un efecto positivo, pues mientras mayor es el número de materias perdidas, mayor probabilidad de desertar tiene el estudiante. Por último, el número de créditos tomados por semestre tiene un efecto difuso, pues, cuando los estudiantes toman demasiados o muy pocos créditos, tienen mayor probabilidad de desertar, frente a los que toman el número de créditos sugerido. La relevancia de estas variables se debió principalmente a que la deserción y el desempeño son temas muy relacionados.

En cuanto a las características propias del estudiante, se identificó a la variable edad como la única que tiene un efecto claro sobre la deserción. Por cada año adicional que el estudiante tenga cuando entra a la universidad, la probabilidad de desertar aumenta en un 4%, tomando como edad base los 17 - 18 años. Por otro lado, fue sorprendente encontrar que las variables sexo, residencia previa y estado civil del estudiante; no tienen ningún efecto sobre la deserción. Cabe mencionar que hay múltiples estudios que indican que las mujeres tienen una menor probabilidad de desertar en comparación a los hombres. Al realizar un análisis univariado se pudo comprobar esto, sin embargo, al realizar el análisis multivariado

se comprobó que este efecto se anula cuando se toma en consideración otros factores. De igual manera, la residencia previa del estudiante, es una variable relevante cuando únicamente se analiza este factor, sin embargo, pierde relevancia cuando se la estudia junto al resto de covariables.

Al analizar las variables socioeconómicas, se encontraron diferentes resultados. En primer lugar, se observó que la variable Categoría Socioeconómica de la Pensión Diferenciada (SPD) no fue relevante en el estudio, pues todas las categorías tenían un comportamiento similar que no permitía identificar a los estudiantes con mayor riesgo de desertar. Este es un hecho sorprendente ya que la bibliografía y otros estudios han demostrado que esta variable influye en la deserción. Una de las principales razones por las que puede no existir ningún efecto, puede estar relacionado con una inadecuada clasificación. Por otro lado, al observar el cambio de categoría SPD se pudo observar que cuando el estudiante recibe una degradación de su categoría tiene menor riesgo de desertar. Por último, en cuanto a los descuentos promedio que recibe el estudiante por semestre, se observó que en el modelo multivariado el efecto fue nulo. En cambio, en el modelo de Kaplan-Meier se pudo ver que mientras mayor es el descuento, mayor es la probabilidad de desertar, no obstante, cuando el estudiante recibe un descuento pequeño tiene más probabilidades de quedarse en la Universidad. Esto se puede deber a que el estudiante siente que recibe un beneficio que podría no recibir en ningún otro lado.

En cuanto a la variable tipo de colegio, en verdad no se puede asegurar que tenga un efecto sobre la deserción. Sin embargo, es necesario mencionar que la categoría de colegio desconocido sí es una variable que representa una mayor riesgo. Cabe mencionar que no existe una explicación clara para que el colegio del estudiante no se encuentre publicado en las páginas oficiales de las entidades pertinentes. Otra subcategoría de este grupo de variables son los aspectos familiares, en donde se observó que el nivel de educación del padre y de la madre influyen en la deserción, principalmente durante los primeros semestres. De igual manera, varios estudios muestran que mientras mayor sea el nivel de educación de los padres, menor riesgo de desertar tiene el alumno. Sin embargo, al analizarlo con el resto de variables, pierden relevancia. Por otro lado, el estado civil de la madre no tiene relevancia, mientras que el del padre sí. Esta variable resultó significativa tanto en el análisis univariado como en el multivariado. En ambos modelos se mostró que los hijos de padres divorciados tienen menos probabilidad de desertar y los hijos de padres con status soltero, tienen un mayor riesgo de abandonar la Institución.

En cuanto al grupo de variables relacionadas con las condiciones iniciales de la Universidad, se observó que las variables duración de la carrera y periodo de entrada, no tienen un efecto claro, por ende, no se puede decir que explican la deserción. Por otro lado, en el modelo no paramétrico se pudo observar que la variable Facultad puede llegar a tener un gran efecto sobre la deserción. En este estudio, se pudo comprobar que aquellas Facultades con mayor contenido matemático, tienden a ser las carreras con mayor probabilidad de deserción, mientras que aquellas ciencias experimentales y sociales con objetivos más filantrópicos, tienden a ser de las Facultades con menores tasas de deserción. Es necesario mencionar que en el modelo multivariado no se calcularon las tasas de riesgo de esta variable, ya que se estratificó y por lo tanto su efecto pasó a ser parte del riesgo de referencia.

Por último, es necesario mencionar que además de estas variables, pueden existir otras que no han sido contempladas. Sin embargo, en esta investigación se ha utilizado la información que la universidad tenía disponible. Cabe mencionar que la deserción no se podrá pronosticar nunca de manera exacta, ya que

en la mayoría de casos depende de factores que son complejos de medir, tales como los factores psicológicos, de personalidad o sentimientos de los estudiantes. Como se indicó desde el inicio de esta investigación, la deserción es un fenómeno multidisciplinario que puede depender de una infinidad de factores.

Recomendaciones

Una de las principales limitaciones del estudio fue la ausencia de información. Los ámbitos que presentaron una mayor ausencia de información fueron las variables relacionadas con la información previa la entrada a la Universidad y los factores socioeconómicos y familiares. Para futuras investigaciones se recomienda obtener la información académica previa de los estudiantes, como por ejemplo el promedio de notas con el que se graduó del colegio y/o las notas obtenidas en el Examen Nacional de Evaluación Educativa “Ser Bachiller”. Además, se recomienda obtener mayor información socioeconómica y características personales del estudiante tales como el pago total por semestre, salario promedio de sus padres, sector o barrio donde habita, etnia, número de miembros del hogar, etc. Esta información permitirá analizar si estas características influyen en la deserción y permiten predecirla con mayor precisión.

Para próximas indagaciones sobre este tema, se recomienda hacer un estudio longitudinal que contemple un periodo más largo, en el cual se tome en cuenta la información de todos aquellos estudiantes que ya han terminado su programa de estudios. De igual manera, se sugiere analizar las características de cada una de las Facultades por separado y crear un modelo para cada una, ya que como se mencionó anteriormente, los estudiantes de cada una de las Facultades pueden tener distintas características. Estas diferencias pueden ocasionar variaciones en los resultados y la significancia de variables puede cambiar entre las Facultades analizadas.

En cuanto a recomendaciones metodológicas, se aconseja probar con otras técnicas estadísticas y algoritmos de Machine Learning. Además, se sugiere en futuras ocasiones acompañar el estudio con metodologías cualitativas, en donde se interactúe con los desertores. Este estudio podría estar acompañado con teorías de otras áreas, como por ejemplo de psicología, antropología y pedagogía. Estos estudios complementarios pueden construir un perfil más preciso del desertor.

Esta investigación propone un modelo de predicción de la deserción universitaria y de ninguna manera afirma que sus resultados son universales para todo tipo de casos. Sin embargo, se espera que esta investigación sea el punto de partida para otros estudios similares que se realicen en la PUCE sede Quito y en general en el Ecuador.

BIBLIOGRAFÍA

- Aalen, O., Borgan, Ø., & Gjessing, H. (2008). An introduction to survival and event history analysis. In O. O. Aalen, Ø. Borgan, & H. K. Gjessing, *Survival and Event History Analysis*. Nueva York: Springer. doi:https://doi.org/10.1007/978-0-387-68560-1_1
- Aina, C. (2005, agosto). Parental background and college drop-out. Evidence from Italy. (U. o. Pavia, Ed.) Retrieved from https://www.researchgate.net/publication/228599916_Parental_background_and_college_drop-out_Evidence_from_Italy
- Alban Taípe, M. S., & Sanchez, D. M. (2018, Diciembre 28). Prediction of university dropout through technological factors: a case study in Ecuador. *Prediction of university dropout through technological factors: a case study in Ecuador*, 39(52), 8.
- Alvarez, Y. L. (2016, junio 29). *Aplicación de un modelo de sobrevida para el análisis de la deserción en los programas de ingeniería electromecánica y diseño industrial de la Universidad Pedagógica y Tecnológica de Colombia*.
- Ameri, S., Chinnam, R., Fard, M., & Reddy, C. (2016). Survival Analysis based Framework for Early Prediction of Student Dropouts. *CIKM'16: ACM Conference on Information and Knowledge Management*. Nueva York: Association for Computing Machinery. doi:ISBN: 978-1-4503-4073-1
- Banco Mundial. (2017, Mayo 17). *Banco Mundial*. Retrieved Mayo 21, 2020, from Graduarse: solo la mitad lo logra en América Latina: <https://www.bancomundial.org/es/news/feature/2017/05/17/graduating-only-half-of-latin-american-students-manage-to-do-so>
- Bean, J. (1985, Primavera). Interaction Effects Based on Class Level in an Explanatory Model of College Student Dropout Syndrome. *American Educational Research Journal*, 22(1), 35-64. doi: American Educational Research Journal
- Bean, J., & Eaton, S. (2001, Mayo 01). The Psychology Underlying Successful Retention Practices. *Journal of College Student Retention: Research, Theory & Practice*, 3, 73-89.
- Becker, G. (1993). *Human Capital. A theoretical and Empirical Analysis, with Special Reference to Education*. Estados Unidos: The University of Chicago Press. doi:ISBN: 0-226-04120-4
- Beer, C., & Lawson, C. (2017, junio). The problem of student attrition in higher education: An alternative perspective. *Journal of Further and Higher Education*, 41, 773-784. doi:<https://doi.org/10.1080/0309877X.2016.1177171>
- Berens, J., Schneider, K., Gortz, S., Oster, S., & Burghoff, J. (2019, diciembre 23). Early Detection of Students at Risk-Predicting Student Dropouts Using Administrative Student Data from German Universities and Machine Learning Methods. *Journal of Educational Data Mining*, 11(3), 1-41. doi:<https://doi.org/10.5281/zenodo.3594771>
- Berrenechea López, L. (2018). *Técnicas no paramétricas y modelos de regresión para datos de tiempo de vida*. Retrieved enero 2021, from Trabajo Fin de Grado.
- Bird, G. (2018, febrero 08). The Impact of Parents' Education Levels. *Inside Higher ED*. Retrieved 2021, from [https://www.insidehighered.com/news/2018/02/08/students-postsecondary-education-arcs-affected-parents-college-backgrounds-study#:~:text=A%20person's%20parental%20education%20levels,of%20college%2C%20the%20study%20found.&text=The%20children%20of%20parents%](https://www.insidehighered.com/news/2018/02/08/students-postsecondary-education-arcs-affected-parents-college-backgrounds-study#:~:text=A%20person's%20parental%20education%20levels,of%20college%2C%20the%20study%20found.&text=The%20children%20of%20parents%20)
- Bonaldo, L., & Nobre Pereira, L. (2016). Dropout: Demographic Profile of Brazilian University Students. *Procedia - Social and Behavioral Sciences*, 228, 138-143. doi:<https://doi.org/10.1016/j.sbspro.2016.07.020>
- Bound, J., & Turner, S. (2011). Dropouts and Diplomas: The Divergence in Collegiate Outcomes. In E. Hanushek, S. Machin, & L. Woessmann, *Handbook in Economics of Education* (pp. 603-609). Amsterdam: Elsevier. doi:ISBN: 978-0-444-53444-6
- Brewer, D., & McEwan, P. (2010). *Economics of Education*. Oxford, UK: ELSEVIER - ACADEMIC PRESS.

- Cameron, A. C., & Trivedi, P. K. (2005). Binary Outcomes Models. In A. Cameron, & P. Trivedi, *MICROECONOMETRICS* (pp. 463-474). New York, Estados Unidos: CAMBRIDGE University Press.
- Carnoy, M. (2006). Capítulo I. La economía de la educación en una economía global. In M. Carnoy, *Economía de la educación*. Barcelona: Editorial UOC.
- Carnoy, M. (2013). *Economía de la educación*. (M. Castells, Trans.) Barcelona, España: Editorial UOC. doi:ISBN: 84-9788-585-X
- CEAACES. (2016). *Proceso de Evaluación, Acreditación y Recategorización 2016 Informe Final*. Cuenca: Universidad Católica de Cuenca. Retrieved noviembre 2020
- CEAACES. (2020). *Modelo de Evaluación Institucional para los Institutos Superiores Técnicos y Tecnológicos en Proceso de Acreditación*. Consejo de Aseguramiento de la Calidad de la Educación Superior, Quito.
- Chavez, M., & Zurita, G. (2005). ESTUDIO ESTADÍSTICO ACERCA DE LA DESERCIÓN ESTUDIANTIL DE LA ESPOL: EL CASO DE LAS FACULTADES DE INGENIERÍAS. *ESTUDIO ESTADÍSTICO ACERCA DE LA DESERCIÓN ESTUDIANTIL DE LA ESPOL: EL CASO DE LAS FACULTADES DE INGENIERÍAS*. Guayaquil, Ecuador: ESPOL.
- Congreso Nacional. (2003, enero 03). CÓDIGO DE LA NIÑEZ Y ADOLESCENCIA. *Registro Oficial No. 737*, Artículo 4. Ecuador. Retrieved febrero 2021, from https://www.igualdad.gob.ec/wp-content/uploads/downloads/2019/12/codigo_ninez_adolescencia_nov2019.pdf
- Contrato Social. (2014). Cuaderno del Contrato Social por la Educación. *Deserción, calidad y reforma universitaria. Apuntes para el debate, Cuaderno 10*. Cevallos Tejada, Francisco;. Retrieved marzo 2021
- Cortez, N., Aguila, A., Gonzalez, M., & Muñoz, Z. (2016). Perfil de los estudiantes que desertan de la Universidad Austral de Chile entre los años 2010-2015. *IV Conferencia Latinoamericana sobre el Abandono en la Educación Superior*.
- Da Costa, F., De Souza Bispo, M., & De Faria Pereira, R. (2018). Dropout and retention of undergraduate students in management: a study at a Brazilian Federal University. *RAUSP Management Journal*, 53(1), 74-85. doi:<https://doi.org/10.1016/j.rauspm.2017.12.007>
- Date, S. (2020, diciembre 11). *The Stratified Cox Proportional Hazards Regression Model*. Retrieved marzo 2021, from <https://towardsdatascience.com/the-stratified-cox-proportional-hazards-regression-model-fa1fa5de2bb1>
- Díaz Peralta, C. (2008). MODELO CONCEPTUAL PARA LA DESERCIÓN ESTUDIANTIL UNIVERSIDAD CHILENA. *SciELO*, 34(2), 65-86. doi:<http://dx.doi.org/10.4067/S0718-07052008000200004>
- Díaz Peralta, C. (2008). MODELO CONCEPTUAL PARA LA DESERCIÓN ESTUDIANTIL UNIVERSITARIA CHILENA. *Estudios Pedagógicos (Valdivia)/ SciELO*, 34(2), 65-86.
- Farah, N., & Upadhyay, M. (2017). How are school dropouts related to household characteristics? Analysis of survey data from Bangladesh. (C. Elliot, Ed.) *Cogent Economics & Finance*, 5(1). doi:<https://www.tandfonline.com/doi/full/10.1080/23322039.2016.1268746>
- Federal Education Policy Project - SIFEPP. (2009). *States' Impact on Federal Education Policy Project -Federal Education Policy and the States, 1945-2009: A Brief Synopsis*. Nueva York.
- Fernández, X., & Silva, E. (2014). Deserción estudiantil universitaria en el primer semestre. El caso de una institución de educación superior ecuatoriana. *Cuadernos del contrato social por la educación*, 35-48.
- Ferreira, M. M., Avitabile, C., Botero Alvarez, J., Haimovich Paz, F., & Urzua, S. (2017). At a Crossroads: Higher Education in Latin America and the Caribbean. *Directions in Development - Human Development*.
- Freire Seoane, M., & Teijeiro Álvarez, M. (2010). Las ecuaciones de Mincer y las tasas de rendimiento de la educación en Galicia. España: Universidad de Coruña. Retrieved from <https://ideas.repec.org/h/aec/ieed05/05-14.html>
- Gallegos, J., Campos, N., Canales, K., & González, E. (2017, diciembre). Factores Determinantes en la Deserción Universitaria. Caso Facultad de Ciencias Económicas y Administrativas de la Universidad Católica de la Santísima Concepción (Chile). *La Serena - SciELO*, 11(3). doi:<http://dx.doi.org/10.4067/S0718-50062018000300011>

- Goel, M. K., Khanna, P., & Kishore, J. (2010). Understanding survival analysis: Kaplan-Meier estimate. (U. n. Health, Ed.) *International Journal of Ayurveda Research*, 1(4). doi:10.4103/0974-7788.76794
- Goldin, C. (2014). Human Capital. *Handbook of Cliometrics*.
- González Fiegehen, L. E. (2008, enero). Deserción en educación superior en América Latina y el Caribe 2008-16. *ResearchGate*. Retrieved mayo 06, 2020, from <https://www.researchgate.net/publication/275275484>
- Guimarães, J., Sampaio, B., & Sampaio, Y. (2010, Junio). What is behind University Dropout Decision in Brazil? A Bivariate Probability Model. *The Empirical Economics Letters*, 9(6). doi:ISSN 1681 8997
- Hanushek, E. (1987). Educational Production Functions. In G. Psacharopoulos, *Economics of Education*. ELSEVIER. doi:<https://doi.org/10.1016/C2013-0-03790-5>
- Hegde, V., & Prageeth, P. (2018, Junio). Higher education student dropout prediction and analysis through educational data mining. (IEEE, Ed.) *2nd International Conference on Inventive Systems and Control (ICISC)*. doi:10.1109/ICISC.2018.8398887
- Hidalgo Torres, S. E. (2018, Marzo). DIAGNÓSTICO DE LA DESERCIÓN ESTUDIANTIL EN LA CARRERA DE ARQUITECTURA DE LA PUCESI. (A. L. Tocain, Ed.) Quito, Ecuador: PUCE.
- IESALC. (2007). *INFORME SOBRE LA EDUCACIÓN SUPERIOR EN AMÉRICA LATINA Y EL CARIBE 2000-2005. La metamorfosis de la educación superior*. UNESCO, Instituto Internacional de la UNESCO para la Educación Superior en América Latina y el Caribe, Caracas. doi:ISBN -980-6556-19-4
- IESALC-UNESCO. (2002). *LA EDUCACIÓN SUPERIOR EN COLOMBIA*.
- Ishitani, T., & DesJardins, S. (2002, agosto). A Longitudinal Investigation of Dropout from College in the United States. *Journal of College Student Retention: Research, Theory & Practice*, 4(2), 173-201. doi:<https://doi.org/10.2190/V4EN-NW42-742Q-2NTL>
- Jencks, C., Smith, M., Acland, H., Bane, M., Cohen, D., Gintis, H., . . . Michelson, S. (1981). *Inequality: A Reassessment of the Effect of Family and Schooling in America*. Nueva York, Estados Unidos: Harper. Retrieved 2020
- Johnes, G., Johnes, J., Agasisti, T., L., & Lopez-Torres, L. (2017). *Handbook of Contemporary Education Economics*. Cheltenham, Glos, UK: Edward Elgar Publishing Limited.
- Juajibioy, J. C. (2016, octubre 24). Study of University Dropout Reason Based on Survival Model. *Scientific Research Publishing - Open urnal of Statistics*, 908-916. doi:10.4236/ojs.2016.65075
- LaMorte, W. (2016, Junio 3). *Survival Analysis*. (Boston University School of Public Health) Retrieved enero 2021, from Comparing Survival Curves: https://sphweb.bumc.bu.edu/otlt/mph-modules/bs/bs704_survival/BS704_Survival5.html
- Lewis, N. (2020). Education. In A. Kobayashi, *International Encyclopedia of Human Geography* (Segunda Edición ed., pp. 69-76). Canada. Retrieved octubre 2020, from <https://doi.org/10.1016/B978-0-08-102295-5.10255-0>
- Machado, J. (2019, Diciembre 02). 23.000 universitarios abandonan cada año la educación superior Para hacer uso de este contenido cite la fuente y haga un enlace a la nota original en Primicias.ec: <https://www.primicias.ec/noticias/sociedad/agustin-alban-senescyt-ser-bachiller-universida>. *PRIMICIAS*, pp. <https://www.primicias.ec/noticias/sociedad/agustin-alban-senescyt-ser-bachiller-universidades/>.
- Meghir, C., & Steven, R. (2011). Econometric Methods for Research in Education - Volume 3. In E. Hanushek, S. Machin, & L. Woessman, *Handbook of Economics*. Holanda: ELSEVIER. doi:10.1016/S0169-7218(11)03001-2
- Meyer, M., & Marx, S. (2014, septiemmbre). Engineering Dropouts: A Qualitative Examination of Why Undergraduates Leave Engineering. *Journal of Engineering Education*, 103(4). doi:10.1002/jee.20054
- Miller, C. (2019, noviembre 06). *Educatedata.org*. Retrieved febrero 2021, from College Dropout Rates: <https://educationdata.org/college-dropout-rates#:~:text=Students%20under%20the%20age%20of,%2D%20or%20four%2Dyear%20colleges>.

- Mills, M. (2011). The fundamentals of survival and event history analysis. In M. Mills, *Introducing Survival and Event History Analysis* (pp. 1-6). Londres, Inglaterra: SAGE Publications. Retrieved 2021
- Min, Y., Zhang, G., Long, R., Aderson, T., & Ohland, M. (2011, abril). Nonparametric Survival Analysis of the Loss Rate of Undergraduate Engineering Students. *Journal of Engineering Education*, 100(2), 349–373. Retrieved 2020
- Mincer, J. (1981). Human Capital and Economic Growth. *National Bureau of Economic Research (NBER) WORKING PAPER SERIES*. Retrieved octubre 2020
- Miranda, J. (2011, Noviembre 30). La función de producción educativa: Una perspectiva schumpeteriana para el análisis de la eficacia de la educación en el Caribe colombiano. *ECONOMICAS CUC*, 32(1), 199-208.
- Mora, J.-G., & Villa, L. (2003). The Economics of Higher Education. In R. Begg, *The Dialogue between Higher Education Research and Practice* (pp. 121-134). Dordrecht: Kluwer Academic Publishers.
- Murillo, D., & de Clunie, G. (2018). VIII Congreso CLABES-VIII Conferencia Latinoamericana sobre el Abandono de la Educación Superior. In D. Murillo, D. Saavedra, & M. Fernandez (Ed.), *VIII Conferencia Latinoamericana sobre el Abandono de la Educación Superior - CLABES DIGITED*. Ciudad de Panamá: Universidad Tecnológica de Panamá. Retrieved from <http://clabes-alfaguia.org/clabes-2018/>
- Nicoletti, M. (2019). Revisiting the Tinto's Theoretical Dropout Model. *Higher Education Studies*, 9(3), 52-64. doi:10.5539/hes.v9n3p52
- Ozturk, I. (2008, Mayo 28). The Role of Education in Economic Development: A Theoretical Perspective. (C. University, Ed.) *SSRN*, 7. doi:<http://dx.doi.org/10.2139/ssrn.1137541>
- Park, S., & Hendry, D. (2015). Reassessing Schoenfeld residual tests of proportional hazards in political science event history analyses. *LSE Research Online*, 59. doi:10.1111/aips.12176
- Pascarella, E. (1980). Student-Faculty Informal Contact and College Outcomes. *Review of Educational Research*, 50(4), 545-595. doi:<https://doi.org/10.2307/1170295>
- Pascarella, E., & Terenzini, P. (2005). How College Affects Students: A Third Decade of Research V.2.
- Paura, L., & Arhipova, I. (2014). Cause Analysis of students' dropout rate in higher education study program. *Procedia Social and Behavioral Science*, 109, 1282-1286. Retrieved 2020, from <https://cyberleninka.org/article/n/234446/viewer>
- Perchinunno, P., Bilancia, M., & Vitale, D. (2019, Diciembre 19). A Statistical Analysis of Factors Affecting Higher Education Dropouts. *Social Indicators Research*.
- Pérez Benítez, E. J. (2014, diciembre). Determinantes socioeconómicos del avance y deserción del nivel primario al secundario en el sistema educativo del Ecuador: ENEMDU 2012. *Determinantes socioeconómicos del avance y deserción del nivel primario al secundario en el sistema educativo del Ecuador: ENEMDU 2012*. Quito: Repositorio PUCE.
- Proyecto ALFA GUIA. (2013). Marco Conceptual sobre el Abandono.
- PUCE. (2020, Junio 24). *IGNATIUS-Sistema de Información PUCE*. Retrieved from Dirección de Informática y Dirección de Aseguramiento de la Calidad.: <https://www.puce.edu.ec/ignatius/>
- Restaino, M. (2008). Dropping out of University of Salerno: A Survival Approach. (D. d. Statistiche, Ed.) *Ideas*. Retrieved 2021, from https://ideas.repec.org/p/sep/wpaper/3_193.html
- Rodriguez C., J. (2016, Marzo). Determinantes de la deserción escolar en la educación secundaria asociados a las características de las instituciones educativas en el Ecuador. *Determinantes de la deserción escolar en la educación secundaria asociados a las características de las instituciones educativas en el Ecuador*. (P. Hermidas, Ed.) Quito, Ecuador: PUCE.
- Rodríguez, F. (2019, diciembre 15). *Diario Expreso*. Retrieved from La deserción universitaria en el país alcanza el 26 %: <https://www.expreso.ec/guayaquil/desercion-universitaria-pais-alcanza-26-1456.html#:~:text=En%20Ecuador%2C%20un%20informe%20publicado,carrera%20en%201os%20primeros%20semestres.>
- Rodríguez, G. (2010). *Generalized Linear Statistical Models*. (P. University, Producer) Retrieved 2020, from POP 507 / ECO 509 / WWS 509 - Princeton University: <https://data.princeton.edu/wws509/notes/c7.pdf>

- Salas-Morera, L., Cejas-Molina, M., Olivares-Olmedilla, J., Garcia-Hernandez, L., & Palomo-Romero, J. (2018, diciembre). Factors Affecting Engineering Students Dropout: A Case Study. *International Journal of Engineering Education*, 35(1), 156-167. Retrieved 2021
- Sandoval-Palis, I., Naranjo, D., Vidal, J., & Gilar-Corbi, R. (2020, noviembre 10). Early Dropout Prediction Model: A Case Study of University Leveling Course Students. *MDPI*. doi:10.3390/su12229314
- Schultz, T. (1961, Marzo). Investment in Human Capital. *The American Economic Review*, 51(1), 1-17. Retrieved Octubre 2020, from <http://links.jstor/sci=0002-8282%28196103%2951%3A1%3C1%3AIIHC%3E2.0.CO%3B2-4>
- Schultz, T. (1987). Education and Population Quality. In G. Psacharopoulos, & G. Psacharopoulos (Ed.), *Economics of Education: Research and Studies* (The World Bank ed.). Washington DC, Estados Unidos: Pergamon Press.
- Simon, S. (2019). *The Analysis Factor*. Retrieved from The Proportional Hazard Assumption in Cox Regression: <https://www.theanalysisfactor.com/assumptions-cox-regression/>
- Singer, J., & Willet, J. (2003). 1. A Framework for Investigating Change over Time. In J. D. Singer, & J. B. Willet, *Applied Longitudinal Data Analysis* (pp. 3-15). Oxford, Inglaterra: Oxford University Press. Retrieved 2020
- Sosu, E., & Pheunpha, P. (2019, febrero 12). Trajectory of University Dropout: Investigating the Cumulative Effect of Academic Vulnerability and Proximity to Family Support. *Frontiers in education*, 4.
- Spady, W. (1970). Dropouts from higher education: An interdisciplinary review and synthesis. *Interchange*, 1, 64-85. doi: <https://doi.org/10.1007/BF02214313>
- Thammasiri, D., Delen, D., Meesad, P., & Nihat, K. (2014, febrero 1). A critical assessment of imbalanced class distribution problem: The case of predicting freshmen student attrition. *Expert Systems with Applications*, 41(2), 321-330. doi:<https://doi.org/10.1016/j.eswa.2013.07.046>
- Tinto, V. (1975). Dropout from Higher Education: A Theoretical Synthesis of Recent Research. *Review of Educational Research Winter*, 45(1), 89-125.
- Tinto, V. (1993). *Leaving College: Rethinking the Causes and Cures of Student Attrition*. Estados Unidos: University of Chicago Press. doi:ISBN: 9780226804491
- Tinto, V. (1997). Classrooms as Communities: Exploring the Educational Character of Student Persistence. *The Journal of Higher Education*, 68(6), 599-623. doi:<https://doi.org/10.2307/2959965>
- Tinto, V., & Cullen, J. (1973, Junio 30). *Dropout in Higher Education: A Review and Theoretical Synthesis of Recent Research*. (C. University, Ed.) Retrieved Mayo 05, 2020, from <https://files.eric.ed.gov/fulltext/ED078802.pdf>
- UCLA. (2018). *Institute of Digital Research & Education Statistical Consulting*. Retrieved 2021, from Survival Analysis with STATA: <https://stats.idre.ucla.edu/stata/seminars/stata-survival/>
- UNESCO. (2019). *UNESCO*. Retrieved octubre 2020, from La educación transforma vidas: <https://es.unesco.org/themes/education>
- United Nations. (s.f.). *ACADEMIC IMPACT*. Retrieved mayo 20, 2020, from Higher Education: <https://academicimpact.un.org/content/higher-education>
- van Houwelingen, H., & Stijnen, T. (2014). Cox Regression Model. In J. P. Klein, H. C. van Houwelingen, J. G. Ibrahim, & T. H. Scheike, *Handbook of Survival Analysis* (pp. 5-23). Boca Raton, Florida, Estados Unidos: Taylor & Francis Group. Retrieved enero 2021
- Viteri Castro, D. A., & Uquillas Narvaez, M. A. (2011). Estudio sobre la deserción estudiantil en la Pontificia Universidad Católica del Ecuador - Matriz, en los niveles 1ro, 2do y 3ero de todas las Facultades y Escuelas del primer semestre del año académico 2007-2008. *Estudio sobre la deserción estudiantil en la Pontificia Universidad Católica del Ecuador - Matriz, en los niveles 1ro, 2do y 3ero de todas las Facultades y Escuelas del primer semestre del año académico 2007-2008*. (M. Moscoso, Ed.) Quito, Ecuador.
- Woodhall, M. (1987). Human capital Concepts. In G. Psacharopoulos, *Economics of Education: Research and Studies* (p. 21). Washington DC: PERGAMON PRESS.

ANEXOS

ANEXO A

Variables de la base de datos

	Nombre de la base	Variables	Explicación de la Variable	Tipo de Variable
id	ALU_DNIALU	Identificación del estudiante	Cédula de identidad del Estudiante	Identificación
Variables explicativas	Características individuales			
	MUJER	Sexo	Sexo del estudiante.	Catógorica
	estadocivil_ESTUDIANTE2	Estado civil del estudiante	El estado civil del estudiante al entrar a la universidad.	Catógorica
	edad_estudiante2	Edad del estudiante	La edad del estudiante al entrar a la universidad.	Continua
	PICHINCHA	Residencia previa	Residencia previa a la entrada a la Universidad (Pichincha o un lugar diferente).	Catógorica
	Características académicas			
	Nueva_etiqueta_COHORTE	Cohorte de entrada	Cohorte en la que el estudiante entró por primera vez a la Universidad	Catógorica
	Año_COHORTE	Año de entrada	Año en el que el estudiante entró a la Universidad	Catógorica
	Periodo_COHORTE	Periodo de entrada	Indica el periodo en el que el estudiante entró en el primer semestre. En el periodo de febrero-julio o en el de agosto-enero.	Catógorica
	semestre_max_alcanzado	Último semestre inscrito	Número de semestre en el que ha estado inscrito el estudiante según las notas registradas hasta 2019-01.	Catógorica
	Nueva_etiqueta_SEMESTRE	Semestres cursados	Cada uno de los semestres en los que estuvo inscrito el estudiante.	Catógorica
	DUR_IDEAL	Duración de la carrera	Cantidad de semestres que debería durar la carrera en la que se inscribió por primera vez el estudiante y estudio durante el primer semestre.	Catógorica
	CRED_IDEAL_OBLIG	Total de créditos de la carrera	Los créditos que debería tomar el estudiante de acuerdo a la carrera en la que se inscribió por primera vez y que estudió durante primer semestre.	Continua
	TIPO_ENTRADA	Tipo de entrada	La forma en la que el estudiante entró a la Universidad.	Catógorica
	PN	Puntaje de Admisión	El Puntaje obtenido en la Prueba de Admisión	Continua
	FACULTADES	Facultad	Facultad a la que pertenece el estudiante.	Catógorica
	cambio_FACULTAD	Cambio de Facultad	Indica si el estudiante se ha cambiado al menos una vez de Facultad.	Catógorica
cambio_CARRERA_2	Cambio de Carrera	Indica si el estudiante se ha cambiado al menos una vez de carrera.	Catógorica	
suma_creditos_total	Total de créditos tomados	Total de créditos tomados por el estudiante desde su primer semestre.	Continua	

	creditos_semestre	Créditos tomados por semestre	Número de créditos que el estudiante curso durante cada uno de los semestres en el que estuvo inscrito.	Continua
	PROMEDIO_creditos	Promedio de créditos por semestre	El promedio de créditos de créditos por semestre.	Continua
	Materias_perdidas	Materias perdidas por semestre	Número de materias que el estudiante perdió durante cada uno de los semestres en el que estuvo inscrito	Discreta
	total_materias_perdidas	Total de Materias perdidas	Total de materias perdidas desde que el estudiante inició sus estudios.	Continua
	promedio	Promedio de notas por semestre	Promedio de las notas finales de cada una de las materias por semestre.	Continua
	PROMEDIO_notas	Promedio general de calificaciones	Promedio de notas de todos los semestres. Promedio general de las notas.	Continua
Características socioeconómicas y familiares				
	SPD	Categoría Socioeconómica	Categoría socioeconómica de la Pensión diferenciada, en la que fue colocado el alumno cuando ingresó a la Universidad.	Categórica
	cambio_SPD	Cambio de la Categoría Socioeconómica	Variable binaria que indica si ha cambiado de categoría socioeconómica.	Categórica
	TOTAL_DESCUENTO_PAGO	Descuento total en la matrícula	El descuento total que ha recibido el estudiante durante sus estudios.	Continua
	DESCUENTO_PROMEDIO	Descuento promedio por semestre	Es el total de descuento promedio que recibe el estudiante en el pago del semestre.	Continua
	TIPO_COLEGIO	Tipo de Colegio	El tipo de colegio del que el alumno se graduó.	Categórica
	Nivel_educacion_Madre	Nivel de educación de la madre	El nivel máximo de educación alcanzado por la madre del estudiante.	Categórica
	estadocivil_MADRE	Estado Civil de la Madre	El estado civil de la madre del estudiante.	Categórica
	Nivel_educacion_Padre	Nivel de Educación del Padre	El nivel máximo de educación alcanzado por el padre del estudiante.	Categórica
	estadocivil_PADRE	Estado Civil del Padre	El estado civil del padre del estudiante.	Categórica
Variable dependiente	DESERTOR	Desertor	Estudiantes que abandonaron la Universidad y no vuelve durante todo el período de análisis.	Categórica
	tstart	Tiempo de inicio	Variable que indica el periodo que indica el inicio del periodo que cursa el estudiante	Discreta
	tstop	Tiempo final	Variable que indica el final del periodo que cursa el estudiante	Discreta

Elaborado por: Autor.

ANEXO B

Modelo1: Modelo de riesgos proporcionales de Cox:

$$\lambda_i(t|x_i) = \lambda_0(t)\exp\{x_i'\beta\}$$

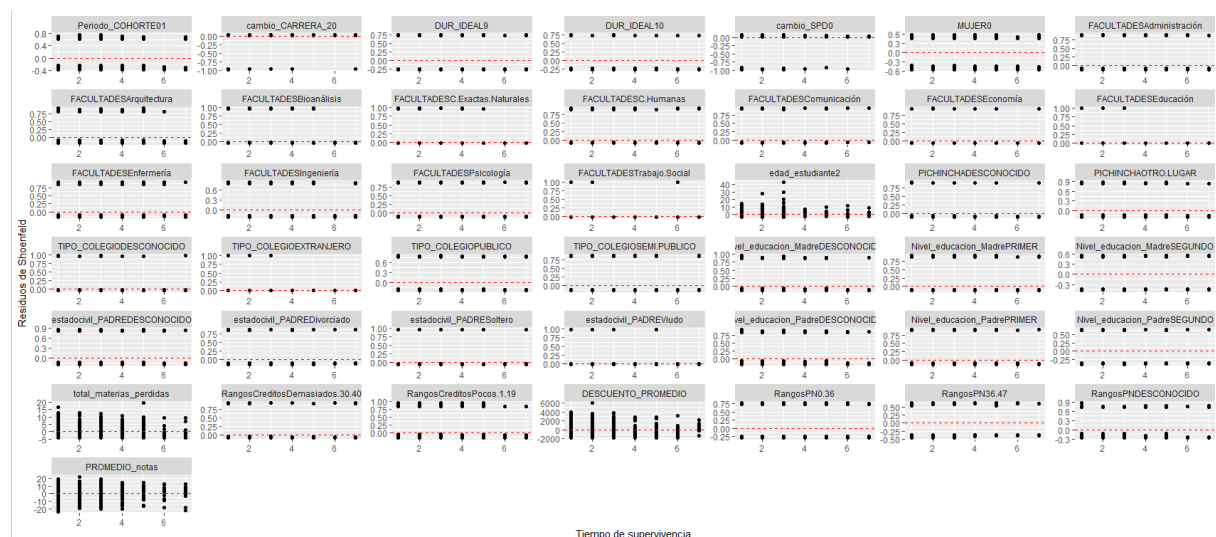
$$\text{tasa de riesgo} = \ln\left(\frac{\lambda_i(t|x_i)}{\lambda_0(t)}\right) = \{x_i'\beta\}$$

$$\text{tasa de riesgo} = B_1X_1 \dots B_kX_k$$

tasa de riesgo

$$\begin{aligned} &= B_1\text{Periodo Cohorte}_{01} + B_2\text{cambioCARRERA}_{20} + B_3\text{cambioCARRERA}_{20} + B_4\text{cambioSPD}_0 \\ &+ B_5\text{MUJER}_0 + B_6\text{FACULTADES}_{\text{Administración}} + B_7\text{FACULTADES}_{\text{Arquitectura}} + B_8\text{FACULTADES}_{\text{Bioanálisis}} \\ &+ B_{10}\text{FACULTADES}_{\text{C.Exactas\&Naturales}} + B_{11}\text{FACULTADES}_{\text{C.Humanas}} + B_{12}\text{FACULTADES}_{\text{Comunicación}} \\ &+ B_{13}\text{FACULTADES}_{\text{Economía}} + B_{14}\text{FACULTADES}_{\text{Educación}} + B_{15}\text{FACULTADES}_{\text{Enfermería}} + B_{16}\text{FACULTADES}_{\text{Ingeniería}} \\ &+ B_{17}\text{FACULTADES}_{\text{Psicología}} + B_{18}\text{FACULTADES}_{\text{Trabajo Social}} + B_{20}\text{edadestudiante}_2 \\ &+ B_{21}\text{PICHINCHA}_{\text{DESCONOCIDO}} + B_{22}\text{PICHINCHA}_{\text{Otro Lugar}} + B_{23}\text{TipoColegio}_{\text{Desconocido}} \\ &+ B_{24}\text{TipoColegio}_{\text{Extranjero}} + B_{25}\text{TipoColegio}_{\text{Publico}} + B_{26}\text{TipoColegio}_{\text{Semi-Publico}} \\ &+ B_{27}\text{Nivel Educacion Madre}_{\text{Desconocido}} + B_{28}\text{Nivel Educacion Madre}_{\text{Primer}} \\ &+ B_{28}\text{Nivel Educacion Madre}_{\text{Segundo}} + B_{30}\text{Nivel Educacion Padre}_{\text{Desconocido}} \\ &+ B_{31}\text{Nivel Educacion Padre}_{\text{Primer}} + B_{32}\text{Nivel Educacion Padre}_{\text{Segundo}} \\ &+ B_{33}\text{Estado Civil Padre}_{\text{Desconocido}} + B_{34}\text{Estado Civil Padre}_{\text{Divorciado}} + B_{35}\text{Estado Civil Padre}_{\text{Soltero}} \\ &+ B_{36}\text{Estado Civil Padre}_{\text{Viudo}} + B_{37}\text{Promedio Creditos} + B_{38}\text{Descuento Promedio} \\ &+ B_{39}\text{total materias perdidas} + B_{40}\text{Promedio Notas} + B_{41}\text{Rango PN}_{0-36} + B_{42}\text{Rango PN}_{36-47} \\ &+ B_{43}\text{Rango PN}_{\text{Desconocido}} + B_{44}\text{Credito Ideal Obligatorio} \end{aligned}$$

Prueba de residuos de Schoenfeld Modelo 1



ANEXO C

Modelo 2: Modelo Estratificado de Cox

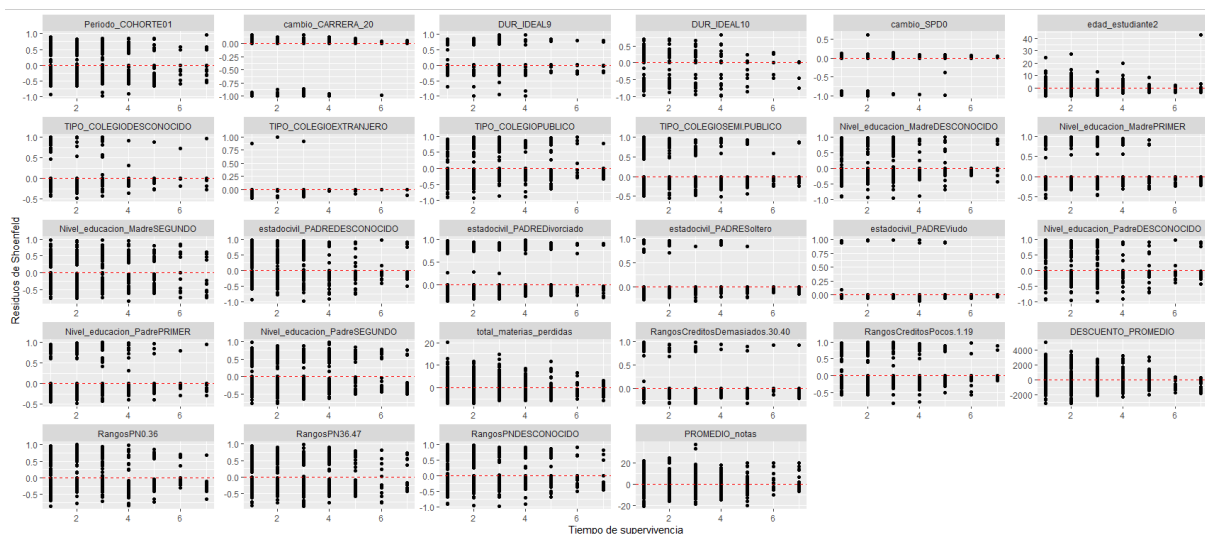
$$h(t|X, G) = h_{G0}(t) \exp\{x'_i \beta\}$$

$$\ln\left(\frac{h(t|X, G)}{h_{G0}(t)}\right) = \{x'_i \beta\}$$

$$\text{tasa de riesgo} = \ln\left(\frac{h(t|X, FACULTADES)}{h_{FACULTADES, Mujer, Pichincha}(t)}\right) = \{x'_i \beta\}$$

$$\begin{aligned} \text{tasa de riesgo} = & B_1 \text{Periodo Cohorte}_{01} + B_2 \text{DuracionIdeal}_9 + B_2 \text{DuracionIdeal}_{10} \\ & + B_3 \text{cambioCARRERA}_{20} + B_4 \text{cambioCARRERA}_{20} \\ & + B_5 \text{cambioSPD}_0 + B_7 \text{edadestudiante}_2 + B_{10} \text{TipoColegio}_{\text{Desconocido}} \\ & + B_{11} \text{TipoColegio}_{\text{Extranjero}} + B_{12} \text{TipoColegio}_{\text{Publico}} + B_{13} \text{TipoColegio}_{\text{Semi-Publico}} \\ & + B_{14} \text{Nivel Educacion Madre}_{\text{Desconocido}} + B_{15} \text{Nivel Educacion Madre}_{\text{Primer}} \\ & + B_{16} \text{Nivel Educacion Madre}_{\text{Segundo}} + B_{17} \text{Nivel Educacion Padre}_{\text{Desconocido}} \\ & + B_{18} \text{Nivel Educacion Padre}_{\text{Primer}} + B_{19} \text{Nivel Educacion Padre}_{\text{Segundo}} \\ & + B_{20} \text{Estado Civil Padre}_{\text{Desconocido}} + B_{21} \text{Estado Civil Padre}_{\text{Divorciado}} \\ & + B_{22} \text{Estado Civil Padre}_{\text{Soltero}} + B_{23} \text{Estado Civil Padre}_{\text{Viudo}} \\ & + B_{24} \text{Rango de Créditos}_{30-40} + B_{25} \text{Rango de Créditos}_{1-19} + B_{26} \text{Descuento Promedio} \\ & + B_{27} \text{total materias perdidas} + B_{28} \text{Promedio Notas} + B_{29} \text{Rango PN}_{0-36} \\ & + B_{30} \text{Rango PN}_{36-47} + B_{31} \text{Rango PN}_{\text{Desconocido}} \end{aligned}$$

Prueba de residuos de Schoenfeld Modelo 2



ANEXO D

Modelo 3: Modelo de Cox con covariables de tiempo dependiente

$$\lambda_i(t, x_i(t)) = \lambda_0(t) \exp\{x_i(t)' \beta\}$$
$$\text{tasa de riesgo} = \ln \left(\frac{\lambda_i(t|x_i(t))}{\lambda_0(t)} \right) = \{x_i' \beta + x_i'(t) \beta\}$$
$$\text{tasa de riesgo} = B_1 X_1 + B_2 X_2(t) \dots B_k X_k$$

tasa de riesgo

$$\begin{aligned} &= B_1 \text{Periodo Cohorte}_{01} + B_2 \text{cambioCARRERA2}_0 + B_3 \text{DuracionIdeal}_9 + B_4 \text{DuracionIdeal}_{10} \\ &+ B_5 \text{cambioSPD}_0 \\ &+ B_6 \text{MUJER}_0 + B_7 \text{FACULTADES}_{\text{Administración}} + B_8 \text{FACULTADES}_{\text{Arquitectura}} + B_9 \text{FACULTADES}_{\text{Bioanálisis}} \\ &+ B_{10} \text{FACULTADES}_{\text{C.Exactas\&Naturales}} + B_{11} \text{FACULTADES}_{\text{C.Humanas}} + B_{12} \text{FACULTADES}_{\text{Comunicación}} \\ &+ B_{13} \text{FACULTADES}_{\text{Economía}} + B_{14} \text{FACULTADES}_{\text{Educación}} + B_{15} \text{FACULTADES}_{\text{Enfermería}} + B_{16} \text{FACULTADES}_{\text{Ingeniería}} \\ &+ B_{17} \text{FACULTADES}_{\text{Psicología}} + B_{18} \text{FACULTADES}_{\text{Trabajo Social}} + B_{19} \text{edadestudiante2} \\ &+ B_{20} \text{PICHINCHA}_{\text{DESCONOCIDO}} + B_{21} \text{PICHINCHA}_{\text{Otro Lugar}} + B_{22} \text{TipoColegio}_{\text{Desconocido}} \\ &+ B_{23} \text{TipoColegio}_{\text{Extranjero}} + B_{24} \text{TipoColegio}_{\text{Publico}} + B_{25} \text{TipoColegio}_{\text{Semi-Publico}} \\ &+ B_{26} \text{Nivel Educacion Madre}_{\text{Desconocido}} + B_{27} \text{Nivel Educacion Madre}_{\text{Primer}} \\ &+ B_{28} \text{Nivel Educacion Madre}_{\text{Segundo}} + B_{29} \text{Estado Civil Padre}_{\text{Desconocido}} \\ &+ B_{30} \text{Estado Civil Padre}_{\text{Divorciado}} + B_{31} \text{Estado Civil Padre}_{\text{Soltero}} + B_{32} \text{Estado Civil Padre}_{\text{Viudo}} \\ &+ B_{33} \text{Nivel Educacion Padre}_{\text{Desconocido}} + B_{34} \text{Nivel Educacion Padre}_{\text{Primer}} \\ &+ B_{35} \text{Nivel Educacion Padre}_{\text{Segundo}} + B_{36} \text{Rango de Créditos}_{30-40} + B_{37} \text{Rango de Créditos}_{1-19} \\ &+ B_{38} \text{Descuento Promedio} + B_{39} \text{Materias perdidas}(t) + B_{40} \text{Rango PN}_{0-36} + B_{41} \text{Rango PN}_{36-47} \\ &+ B_{43} \text{Rango PN}_{\text{Desconocido}} + B_{43} \text{Promedio Notas}(t) \end{aligned}$$

ANEXO E

Modelo 4: Modelo de Cox con covariables de tiempo dependiente estratificado:

$$\ln\left(\frac{h(t|X(t), G)}{h_{G_0}(t)}\right) = \{x'_i\beta + x'_i(t)\beta\}$$

$$\text{tasa de riesgo} = \ln\left(\frac{h(t|X(t), FACULTADES)}{h_{FACULTADES}(t)}\right) = \{x'_i\beta\}$$

$$\begin{aligned} \text{tasa de riesgo} = & B_1 \text{Periodo Cohorte}_{01} + B_2 \text{cambioCARRERA}_{20} + B_3 \text{DuracionIdeal}_9 \\ & + B_4 \text{DuracionIdeal}_{10} + B_5 \text{cambioSPD}_0 + B_6 \text{MUJER}_0 + B_7 \text{edadestudiante}_2 \\ & + B_8 \text{PICHINCHA}_{DESCONOCIDO} + B_9 \text{PICHINCHA}_{Otro Lugar} + B_{10} \text{TipoColegio}_{DESCONOCIDO} \\ & + B_{11} \text{TipoColegio}_{Extranjero} + B_{12} \text{TipoColegio}_{Publico} + B_{13} \text{TipoColegio}_{Semi-Publico} \\ & + B_{14} \text{Nivel Educacion Madre}_{DESCONOCIDO} + B_{15} \text{Nivel Educacion Madre}_{Primer} \\ & + B_{16} \text{Nivel Educacion Madre}_{Segundo} + B_{17} \text{Estado Civil Padre}_{DESCONOCIDO} \\ & + B_{18} \text{Estado Civil Padre}_{Divorciado} + B_{19} \text{Estado Civil Padre}_{Soltero} \\ & + B_{20} \text{Estado Civil Padre}_{Viudo} + B_{21} \text{Nivel Educacion Padre}_{DESCONOCIDO} \\ & + B_{22} \text{Nivel Educacion Padre}_{Primer} + B_{23} \text{Nivel Educacion Padre}_{Segundo} \\ & + B_{24} \text{Materias perdidas}(t) + B_{25} \text{Rango de Créditos}_{30-40} \\ & + B_{26} \text{Rango de Créditos}_{1-19} + B_{27} \text{Descuento Promedio} + B_{28} \text{Rango PN}_{0-36} \\ & + B_{29} \text{Rango PN}_{36-47} + B_{30} \text{Rango PN}_{DESCONOCIDO} + B_{31} \text{Promedio Notas} \end{aligned}$$