

# VARIABLES QUE INFLUYEN EN LA FINALIZACIÓN DE LA MAESTRÍA EN ESTADÍSTICA DE LA UNIVERSIDAD DE LOS ANDES (VENEZUELA)



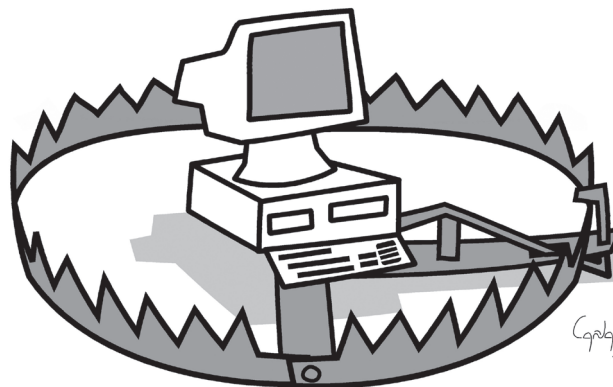
VARIABLES AFFECTING THE COMPLETION OF THE MASTER OF STATISTICS AT THE UNIVERSITY OF LOS ANDES (MÉRIDA-VENEZUELA)

VARIÁVEIS QUE AFETAM A CONCLUSÃO DA PÓS-GRADUAÇÃO EM ESTATÍSTICA DA UNIVERSIDADE DE LOS ANDES (MÉRIDA, VENEZUELA)

JESÚS ALBERTO PEÑA GUILLÉN  
japg@ula.ve  
japguillen@gmail.com

RAFAEL EDUARDO BORGES PEÑA  
borgesr@ula.ve y  
borgesr@gmail.com

Universidad de Los Andes  
Facultad de Farmacia y Bionálisis  
MÉRIDA, estado Mérida.  
Venezuela



Fecha de recepción: 19/01/2012  
Fecha de corrección: 03/02/2012  
Fecha de aceptación: 11/09/2012

## Resumen

Se muestran los resultados de determinar la función del tiempo en que los estudiantes de la Maestría en Estadística finalizan sus estudios y los factores asociados a un menor tiempo de graduación. La información se tomó de los expedientes de los 225 estudiantes que ingresaron a la maestría desde la primera cohorte (1973) hasta los de la cohorte del 2002, cuyo seguimiento fue hasta septiembre de 2004. La metodología utilizada fue un análisis de supervivencia clásico, que consistió en estimar la función del tiempo que tarda un individuo para graduarse, utilizando el estimador de Kaplan y Meier, y el ajuste de modelos para explicar el riesgo de graduarse, mediante el modelo de Cox. Los resultados muestran que las variables que contribuyen a que el tiempo de graduación sea más corto son: poseer una profesión en el área científica, contar con el beneficio de la beca, tener un dominio instrumental del idioma inglés y tomar la maestría por motivos de investigación.

**Palabras clave:** deserción en estudios de postgrado, análisis de supervivencia, estimador de Kaplan y Meier, modelo de Cox.

## Abstract

*We present the results that determine the function of the time expended by the students on the Master Degree Program in Statistics, from the beginning of their studies until they get graduated and, to identify the factors associated to a lower time of graduation. The study was conducted using the academic records of 225 students from the master program between the first cohort (1973) and the 2002 cohort, with observations until September 2004. The methodology used was a classical survival analysis, including the estimation of the graduation time by the Kaplan-Meier estimator and the fitting of Cox regression models to explain the hazard of getting the grade. The results show that the covariates that contribute to a lesser time of graduation are: a profession in a scientific area, to have an economic aid, the instrumental knowledge of English language and, taking the master program by a research motivation.*

**Keywords:** Desertion of graduate studies, Survival analysis, Kaplan-Meier estimator, Cox model.

## Resumo

*Os resultados determinam a função de tempo em que os alunos do Mestrado em estatística, terminando seus estudos e os fatores associaram com menor tempo de graduação. A informação foi tirada da registros de 225 alunos que entraram para o mestre da primeira coorte (1973) até a coorte de 2002, cujo seguimento foi até setembro de 2004. A metodologia utilizada foi a análise de sobrevivência clássica, que consistia em estimar a função desde o momento em que um indivíduo leva para formar, usando o estimador de Kaplan e Meier e o ajuste de modelos para explicar o risco de se formar, por meio do modelo de Cox. Os resultados mostram que as variáveis que contribuem para o menor tempo de graduação são: possuir uma profissão na área científica, tendo o benefício da bolsa de estudos, tendo um domínio instrumental da língua inglesa e ter Mestrado por razões de investigação.*

**Palavras-chave:** desercão em pós-graduação, análise de sobrevivência, o estimador de Kaplan e Meier, modelo de Cox.

## INTRODUCCIÓN



Todo proceso de enseñanza-aprendizaje implica desafíos para el estudiante, algunos lamentablemente no logran alcanzar el objetivo planteado y la población estudiantil venezolana no escapa de esta realidad, sea a causa de bajo rendimiento, deserción, abandono de las actividades académicas o, simplemente, los estudiantes no finalizan el proceso educativo que dieron inicio como parte forjadora de su futuro académico y social.

Ahora bien, ¿qué variables pudieran influir para que un estudiante no finalice su plan de estudios? La no finalización del plan de estudios por parte de la comunidad estudiantil ocurre en todos los ámbitos de la educación, por diversas causas, desde preescolar hasta los estudios de cuarto nivel, que constituye el área académica abordada en la presente investigación y donde a través de la metodología estadística Análisis de Supervivencia se da respuesta a la pregunta de investigación planteada.

En relación con ello, el hecho de que un estudiante no finalice sus estudios de postgrado en el nivel de maestría pudiera estar ligado a factores de orden académico, social y personal. En la presente investigación se consideran más factores de orden académico que de orden personal del estudiante, como tiempo de permanencia en la maestría, profesión, la investigación como motivo por el cual el estudiante ingresó a la maestría en estadística, dominio de un idioma extranjero a nivel de lectura, entre otros.

Por tanto, el problema de la finalización de las tesis de maestría es común y ha sido abordado por diversos autores, desde distintas perspectivas.

Al respecto, Bloom (1981) presenta una aproximación bastante detallada del fenómeno, que es tratado desde una perspectiva asociada a la búsqueda de factores relacionados a la no culminación de la tesis.

Algunas investigaciones relacionadas con la no culminación o retardo de las tesis de maestría, como en el trabajo de Latone y Browne (2001), se han enfocado hacia la búsqueda de los factores asociados a estos dos fenómenos. Entre los factores asociados a la no culminación o retardo en la culminación de las tesis de maestrías se mencionan los siguientes:

1. Según Evans (2002), citado por Carlino (2005), el tiempo de dedicación del tesista a las labores vinculadas a los estudios de maestría y a la elaboración de la tesis está vinculado a la ausencia de becas que permitan realizar las actividades asociadas a estas con una dedicación a tiempo completo.
2. Así mismo Seagram, Gould y Pike (1998), citados por Latona y Browne (2001), coinciden con Delamont y Atkinson (2001) en que la tardanza en la identificación del tema a investigar pudiera estar asociada a la experiencia previa que tiene el tesista en investigación y en los tópicos relacionados con la maestría.
3. Por otro lado, Delamont y Atkinson (2001) establecen la escasa posibilidad que se les da a los tesistas de integrarse a grupos de investigación consolidados.
4. Otro factor muy importante considerado en los trabajos citados por Carlino (2005): Cryer (1996), Dinham y Scott (1999) y Marshall (1993), se refiere al número de publicaciones derivadas de las tesis, lo que constituye un elemento que influye en la motivación del tesista y del tutor.
5. Además, en el trabajo publicado por Rosas, Flores y Valarino (2006), se hace referencia en opinión de los estudiantes de postgrado a la relación de trabajo entre el tutor y el tesista, donde se destacan variables como experiencia investigativa, aporte de ideas e información acerca de líneas de investigación factibles, entre otras variables.

En atención a lo anterior, el factor común de estos trabajos de investigación referentes a la culminación de los postgrados es la investigación o experiencia en investigación como motivo fundamental por el cual un estudiante realiza estudios de cuarto nivel en maestría o doctorado.

En esta investigación se aborda el problema desde el punto de vista de la búsqueda de factores que influyen en la culminación de la tesis y en el tiempo para la culminación de la misma, utilizando la metodología estadística denominada “Análisis de Supervivencia” (Collett, 2003; Hosmer y Lemeshow, 1999 y, Klein y Moeschberger, 1997).

Cabe considerar que existen algunos antecedentes del uso de las técnicas estadísticas de análisis de supervivencia en el ámbito educativo, con especial énfasis en el problema de la deserción en las carreras de pregrado. Entre estos trabajos destacan los de Giovagnoli (2002) y el de Castaño y colaboradores (Castaño et al., 2006), en los cuales se plantean modelos para tiempos discretos basados principalmente en el modelo de Prentice y Gloeckler (1978) y el modelo propuesto por Meyer (1990). En otros trabajos, metodológicamente más cercanos al nuestro, debido a los modelos trabajados, se presentan estimaciones a través del estimador de Kaplan y Meier (1958) y modelos para tiempos continuos, destacándose el uso del modelo de regresión de Cox (1972), entre estos trabajos se pueden mencionar el de Furtado (2003) y el de Lavado y Gallegos (2005), donde además del modelo de regresión de Cox, se plantean modelos de falla acelerado basado en la distribución Log – Logística.

En este trabajo se presentan por una parte, resultados correspondientes a la estimación de la función del tiempo que tarda en graduarse un estudiante de la maestría en estadística, obtenida mediante el estimador de Kaplan y Meier y por otra, resultados de ajustes del modelo de regresión de Cox, el cual ha sido validado de acuerdo a los métodos propuestos por Therneau y Grambsch (2000).

## 1. OBJETIVO Y METODOLOGÍA

El objetivo del presente trabajo es identificar los factores que están asociados con la culminación exitosa de los estudios de maestría, tomando como caso particular la Maestría en Estadística, de la Universidad de Los Andes, Mérida (Venezuela).

Los objetivos secundarios son proveer una metodología que permita identificar el tiempo de graduación de los estudiantes y proponer modelos que permitan la identificación de factores asociados con el tiempo de graduación en la Maestría.

Se aplicó la metodología estadística que forma parte de una gran área de la estadística a la cual se le ha dado el nombre de Análisis de Supervivencia. Dentro de esta área se ha utilizado la estimación de la función de supervivencia a través del estimador de Kaplan y Meier (1958) y el ajuste de los modelos semiparamétricos de riesgo pro-

porcional de Cox (1972), para identificar los factores que influyen en el riesgo de graduarse y por ende en el tiempo de culminación de la tesis de maestría.

Los datos corresponden a 225 estudiantes de la Maestría en Estadística, que funciona adscrita al Instituto de Estadística Aplicada y Computación (IEAC), ubicado en la Facultad de Ciencias Económicas y Sociales (FA-CES) de la Universidad de Los Andes (ULA), Núcleo Mérida (Venezuela), que ingresaron al programa de maestría entre septiembre de 1973 (primera cohorte) y septiembre de 2002, extendiéndose el estudio hasta septiembre de 2004 (último registro de un estudiante graduado).

El esquema que los estudiantes deben seguir en el programa de la Maestría en Estadística es el siguiente:

Una vez finalizada la escolaridad el estudiante debe presentar un trabajo de Asesoría Estadística a alguna institución pública o privada, que puede ser la misma Universidad de Los Andes, el Instituto Autónomo Hospital Universitario de Los Andes (IAHULA), el Instituto Nacional de Estadística (INE, Mérida), entre otros. Este trabajo para muchos estudiantes implica la realización de una “minitesis”, debido al tiempo de dedicación requerido, teniendo la ventaja de que el estudiante da inicio a un trabajo de investigación donde está involucrada la estadística desde un punto de vista más práctico que teórico y además, en algunos casos tienen incluso la oportunidad de publicar estos trabajos de investigación en revistas del ámbito estadístico, ciencias sociales o ciencias de la salud. Posteriormente a la presentación de la Asesoría Estadística viene la elaboración y presentación del proyecto de tesis para finalmente presentar la tesis y graduarse. Pues bien, algunos estudiantes después de finalizada la escolaridad cumplen con todo el proceso explicado anteriormente, otros se pierden del postgrado sin justificación alguna, otros llegan sólo hasta la presentación de la Asesoría Estadística y otros incluso sólo hasta la presentación del proyecto de tesis y, no logran finalmente graduarse.

De acuerdo a Castaño, Gallón, Gómez y Vásquez (2008), en una investigación basada en registros de estudiantes de pregrado de las carreras de Ingeniería y Ciencias Económicas de la Universidad de Antioquia, Colombia, referente a los factores que determinan la decisión por parte del estudiante de desertar o graduarse, se identificaron cuatro factores: institucionales, socioeconómicos, académicos e individuales. Estos fueron analizados en función de los modelos de duración de riesgo proporcional en tiempo discreto de Prentice y Gloeckler (1978), debido a una alta presencia de observaciones censuradas y empatadas respecto al evento *desertar*. En cambio, para el presente trabajo se usaron registros de los estudiantes del postgrado de la Maestría en Estadística de la Universidad de Los Andes, Mérida (Venezuela), lo cual facilitó en cierta forma el seguimiento de los estu-

diantes hasta la ocurrencia del evento graduarse, debido a que no se tiene una alta proporción de observaciones empatadas y por lo cual se hizo uso de los modelos de riesgo proporcional de Cox (1972). A pesar de tratarse de grupos de estudiantes con formación académica diferente, coincidimos en algunos resultados en función de los factores mencionados. Respecto al factor institucional, se tiene que el beneficio de la beca es determinante al momento de observar la ocurrencia del evento graduarse. Así mismo ocurrió con el factor socioeconómico, la situación laboral y la dependencia económica son elementos importantes para el caso de que los estudiantes de pregrado logren graduarse o desertar, aunque en el caso de los estudiantes de postgrado, esas variables se traducen a la profesión del estudiante que es un ente que le proporciona estabilidad laboral y en consecuencia dependencia económica. Sin embargo, en los factores individuales coincidimos en que la edad de inicio en el postgrado y el género (femenino) son variables importantes respecto a la ocurrencia del evento graduarse.

El procesamiento de los datos se llevó a cabo mediante el paquete *survival* (S original by Terry Therneau and ported by Thomas Lumley, 2004) utilizando el lenguaje de procesamiento estadístico R (R Development Core Team, 2004).

## 2. ELEMENTOS BÁSICOS DEL ANÁLISIS DE SUPERVIVENCIA

Los métodos para analizar datos de supervivencia están referidos a ensayos clínicos, estudios epidemiológicos y otras áreas de investigación referidas a la salud de pacientes humanos o pruebas con animales experimentales de laboratorio, donde el tiempo de supervivencia es la variable dependiente y de mayor importancia en el estudio.

En relación con ello, para el registro del tiempo de supervivencia se requiere que el investigador defina un punto de partida para el estudio y a partir de allí, se mide el tiempo en que ocurre un evento definido previo al estudio, el tiempo se mide en función de los segundos, minutos, horas, días, semanas, meses o años, hasta observar el evento de interés.

Con respecto a la presente investigación, el tiempo de supervivencia se define como el tiempo en meses que el estudiante permaneció en la maestría de estadística hasta obtener el grado de Magister Scientiae en Estadística.

Ahora bien, según Lee (2003) la característica a observar o evento de interés puede ser el desarrollo de una enfermedad, respuesta al tratamiento médico, recaída o muerte. En este caso, el evento de interés a observar es que el estudiante adquiera el grado de Magister Scientiae

en Estadística o equivalentemente el evento a observar es, Graduarse.

Siguiendo este mismo orden de ideas, el objetivo del análisis de supervivencia consiste en hacer un seguimiento a los estudiantes desde que se inician en la maestría de estadística (en una cohorte particular) hasta que se presente el evento graduarse. Esta metodología estadística le permite al investigador predecir la probabilidad de supervivencia del estudiante en función de una serie de factores de índole académico que pudieran influir en el tiempo de permanencia en la maestría. Similarmente, se determina el riesgo de pronóstico relativo al evento a ser observado, en general, al desarrollo de la ocurrencia del evento en el tiempo.

Así pues, de acuerdo a Kleinbaum (2005), si en el transcurso del tiempo no se observa la ocurrencia del evento de interés, se tiene un tiempo de supervivencia incompleto, el resultado obtenido hasta ese momento particular, también es conocido como observación censurada o tiempo de supervivencia censurado, en caso contrario, se tiene el tiempo de supervivencia real. Por ejemplo, para los estudiantes que lograron finalizar la maestría en estadística, sus tiempos de permanencia en la institución son tiempos de supervivencia reales, mientras que, para los estudiantes que no lograron graduarse en la maestría, bien sea por retiro voluntario, abandono, finalización sólo de la escolaridad o debido a factores de naturaleza individuales, económicos e institucionales, que además, no fueron considerados en esta investigación, sus tiempos de supervivencia son censurados o incompletos.

Por otra parte, Lee y Wang (2003) indican que los tiempos de supervivencia censurados se clasifican en, censura a la derecha, censura a la izquierda o censura por intervalo. Esta situación depende de la naturaleza de los datos, de las unidades estadísticas de la investigación y de los objetivos del trabajo. De ellas, la de mayor relevancia en las investigaciones de este orden es la censura por la derecha, que a su vez, se clasifica en tres subgrupos, censuras de tipo I, II y III:

La Censura Tipo I es aplicable cuando el estudio se refiere a grupos de individuos que se incorporan en un mismo instante al estudio, en los cuales se observa de manera individual la no ocurrencia del evento a lo largo del tiempo, pero fijando el período de tiempo de observación. Así, los individuos que superan el período de tiempo establecido sin experimentar el evento de interés, su tiempo de supervivencia es al menos la longitud del periodo de tiempo prefijado.

Mientras que, la Censura Tipo II es más aplicable en estudios donde las unidades de observación son animales de laboratorio, debido a que, consiste en observar en qué



tiempo un subgrupo ( $k$ ) del total de individuos ( $n$ ) experimentan el evento de interés, con la finalidad de terminar el estudio en ese instante de tiempo. Los restantes ( $n - k$ ) individuos les corresponden tiempos censurados. En este caso, a diferencia de la Censura Tipo I, no se tiene un intervalo de tiempo fijo para la observación de la ocurrencia del evento de interés.

Y la Censura Tipo III, es recomendada para estudios clínicos y epidemiológicos debido a que el período de estudio es fijo y los pacientes se incorporan al mismo en diferentes instantes de tiempo durante ese período establecido. Algunos individuos experimentan el evento antes de finalizar el estudio; sus tiempos de supervivencia son conocidos. Algunos se retiran antes de finalizar el estudio y se consideran perdidos en el seguimiento, sus tiempos de supervivencia van desde que ingresaron al estudio hasta el último registro. Otros se alivian al final del estudio o permanecieron todo el período de estudio y no experimentaron el evento, sus tiempos de supervivencia son al menos el periodo de tiempo definido.

Respecto a la presente investigación, se observaron tiempos de supervivencia censurados a la derecha, basados en censuras tipo I y II, debido a que no se estableció un periodo de tiempo fijo en el cual el estudiante de la maestría pudiera experimentar el evento graduarse, aunque la amplitud del tiempo fue finita.

**Funciones de tiempo de supervivencia:** Los datos de tiempo de supervivencia miden el tiempo en que un cierto evento de interés ocurre, como por ejemplo, graduarse en la maestría de estadística. El tiempo de finalización de la maestría es una variable aleatoria y a su vez forma una distribución de probabilidad, conocida como, distribución de probabilidad de los tiempos de supervivencia, que se caracteriza por las siguientes funciones matemáticas: (1) Función de Supervivencia,  $S(t)$ , (2) Función de densidad de probabilidad,  $f(t)$ , y (3) Función de riesgo,  $h(t)$ .

**Función de Supervivencia:** Sea  $T$  una variable aleatoria que denota el tiempo de supervivencia. La función de supervivencia  $S(t)$ , se define como la probabilidad de que un individuo sobreviva al menos el instante  $t$ , esto es:

$$S(t) = P(T > t) \quad (1)$$

Ahora bien,  $F(t)$  es la función de distribución acumulada de  $T$ , esto es:  $F(t) = P(T < t)$ , la probabilidad de que el evento ocurra como máximo en el instante  $t$ . En consecuencia, ambas funciones  $S(t)$  y  $F(t)$  están relacionadas, mediante,  $S(t) = 1 - F(t)$ .

Para representar gráficamente  $S(t)$ , Berkson (1942) citado por Lee y Wang (2003) recomendó un gráfico llamado, Curva de Supervivencia, que no es más que una función del tiempo decreciente que indica si la probabilidad de supervivencia observada es baja o alta (Gráfico No. 1).

**Función de densidad de probabilidad:** Para toda variable aleatoria existe su correspondiente función de densidad de probabilidad,  $f(t)$ , que se define como el límite de la probabilidad de que un individuo experimente el evento de interés en el intervalo  $t, t + dt$  por unidad de tiempo con amplitud  $dt$  y está dado por la siguiente expresión:

$$f(t) = \lim_{dt \rightarrow 0} \frac{P(t \leq T < t + dt)}{dt} \quad (2)$$

La representación gráfica de  $f(t)$  se conoce como curva de densidad, que posee las siguientes propiedades: (1)  $f(t)$  es no negativa para todo valor  $t$  no negativo,  $f(t)$  es cero en caso contrario y (2) el área entre la curva de densidad y el eje del tiempo  $t$  es igual a uno.

**Función de riesgo:** Es el límite de la probabilidad de que un individuo experimente el evento de interés en un corto periodo de tiempo  $t, t + dt$ , dado que ha sobrevivido al instante  $t$ , esto es:

$$h(t) = \lim_{dt \rightarrow 0} \frac{P(t \leq T < t + dt | T \geq t)}{dt} \quad (3)$$

A partir de (3) se tiene que:

$$h(t) = \frac{f(t)}{S(t)} = \frac{f(t)}{1 - F(t)} \quad (4)$$

$h(t)$  también es conocida como tasa de fracaso instantánea, fuerza de mortalidad, tasa de mortalidad condicional y tasa específica de estudiantes graduados en el IEAC.

Si el estudiante de la maestría en estadística supera por ejemplo la escolaridad, nos podemos preguntar por el riesgo de graduarse en el periodo de tiempo siguiente, cuya respuesta está dada por  $h(t)$ .

Según Lee y Wang (2003), la función de riesgo  $h(t)$  puede incrementarse, disminuir, permanecer constante o indicar un proceso a un más complejo. Este es un supuesto estadístico importante al momento de elegir el modelo

de supervivencia que mejor se ajuste a los datos del estudio, como es el caso del Modelo de riesgo proporcional de Cox.

A partir de la función anterior (4), se obtiene la función de riesgo acumulada  $H(t)$ :

$$H(t) = \int_0^t h(w)dw = -\ln S(t) \quad (5)$$

Lo que implica que de ésta nueva expresión (5), se obtiene la probabilidad de supervivencia  $S(t)$  hasta  $t$  como una función de la función de riesgo acumulado, en cada uno de los tiempos  $t$  observados, como:

$$S(t) = \exp\{-H(t)\} = \exp\left\{-\int_0^t h(w)dw\right\} \quad (6)$$

En consecuencia, mediante las expresiones (5) y (6) se tiene que: para  $t = 0$ ,  $S(t) = 1$  y  $H(t) = 0$ , al inicio del periodo de seguimiento todos los individuos tienen la misma probabilidad de supervivencia, cuyo riesgo acumulado de que ocurra el evento es nulo. Para  $t = \infty$ ,  $S(t) = 0$ ,  $H(t) = \infty$ , a medida que se incrementa el tiempo de supervivencia la probabilidad de supervivencia disminuye y por lo cual, el riesgo de que ocurra el evento es inevitable.

Respecto a los estadísticos descriptivos de la variable aleatoria  $T$ , se tienen, media, mediana y varianza de supervivencia, cuyos detalles pueden ser revisados en Andersen y colaboradores (Andersen et al., 1982), Gill (1983), Andersen y Borgan (1985), Lee y Wang (2003), Pepe y Fleming (1989, 1991) y en Gill y Johansen (1990).

Función de Supervivencia de Kaplan y Meier: Es un método no paramétrico propuesto por Kaplan y Meier (1958), también conocido como fórmula de Producto Límite, en presencia de datos censurados, donde no se toma en consideración el impacto de las covariables presentes en el estudio, respecto al tiempo de supervivencia. El estimador se define como:

$$\hat{S}_{km}(t) = \prod_{t_{(i)} \leq t} \frac{r(t_{(i)}) - d(t_{(i)})}{r(t_{(i)})} \quad (7)$$

Donde  $t_{(i)}$  son los tiempos de ocurrencia del evento de interés,  $r(t_{(i)})$  es el conjunto de individuos en riesgo de experimentar el evento cuyos tiempos de supervivencia son al menos  $t_{(i)}$  y,  $d(t_{(i)})$  es el número de ocurrencias observadas en cada uno de los tiempos  $t_{(i)}$ .

La varianza de la expresión (7) está dada por la fórmula de Greenwood (1926):

$$\hat{V}\hat{a}r(\hat{S}_{km}(t)) = \hat{S}_{km}^2(t) \sum_{t_{(i)} \leq t} \frac{d(t_{(i)})}{r(t_{(i)}) [r(t_{(i)}) - d(t_{(i)})]} \quad (8)$$

Para la obtención de los intervalos de confianza del 95% para la función de supervivencia estimada, se obtiene de la forma tradicional, mediante:

$$\hat{S}_{km}(t) \pm 1.96ee(\hat{S}_{km}(t)) \quad (9)$$

donde  $ee(\hat{S}_{km}(t))$  es el error estándar de la estimación del estimador de la función de supervivencia de Kaplan y Meier.

Para ver más detalles de la teoría del estimador de Kaplan y Meier pueden ser revisados en Andersen y colaboradores (Andersen et al., 1982), Gill (1983), Andersen y Borgan (1985), Pepe y Fleming (1989, 1991) y en Gill y Johansen (1990).

**Modelo de Regresión de Cox:** El modelo de regresión de Cox (1972) es un método semiparamétrico usado para modelar variables relacionadas al tiempo de supervivencia. Su mayor aplicación ha sido en el área médica, sin embargo, para el presente trabajo se propone su aplicación en el área educativa.

Según Lee (2003) el modelo de riesgo proporcional de Cox, supone que existe un conjunto de covariables  $X = (x_1, x_2, \dots, x_p)^t$  que pudieran presentar cierto efecto en la variable tiempo de ocurrencia del evento de interés. La Regresión de Cox estima la tasa de riesgo a lo largo del tiempo ("Hazard Ratio" (HR)) de que ocurra el evento de interés condicionado a las covariables.

A continuación se muestran las expresiones matemáticas para obtener las funciones de Riesgo y Supervivencia de Cox:

Función de riesgo de Cox: En el modelo de regresión de Cox, el riesgo para el  $j$ -ésimo individuo con el conjunto de características  $X_j(t)$  en el tiempo  $t$ , está dado por:

$$h(t, X_j(t)) = h_0(t) \exp\{\beta' X_j(t)\} \quad (10)$$

Donde:  $X_j(t)$  es el vector de  $p$  covariables,  $h_0(t)$  es llamada función de riesgo base, describe el riesgo para los individuos con la característica  $X_j(t) = 0$ , y sirve como

punto de referencia. La función  $h_0(t)$  se estima sin hacer alusión a alguna distribución teórica definida, por lo cual representa la parte no paramétrica del modelo. La expresión  $\exp(\beta' X_j(t))$  es el riesgo relativo (risk score), indica el incremento o disminución del riesgo, asociado al conjunto de características  $X_j(t)$ , representa la parte paramétrica del modelo.  $\beta$  es el vector de parámetros o coeficientes de la regresión que no incluye el término del intercepto, debido a que es absorbido por  $h_0(t)$  y el término “riesgo” esta asociado al fracaso u ocurrencia del evento de interés.

Sobre esa base de ideas, se tiene que, si el signo del coeficiente es positivo (negativo), significa que la variable asociada a éste, manteniendo fijas las demás variables, tiene un efecto positivo (negativo) respecto al riesgo de que ocurra el evento de interés, mientras que los coeficientes expresados en la forma de riesgo relativo,  $\exp(\cdot)$ , su interpretación depende de la naturaleza de la variable, continua o categórica. Si la variable del modelo es continua,  $\exp(\cdot)$ , indica cuantas veces crece (decrece) el riesgo ante un aumento en una unidad de la variable regresora. Si la variable del modelo es categórica, donde los individuos quedan clasificados por grupos definidos,  $\exp(\cdot)$ , indica la cuantificación del riesgo de que ocurra el evento respecto a la categoría de base, manteniendo constantes el resto de características de los individuos. Además, para seleccionar las variables que resultan estadísticamente significativas del modelo de regresión de Cox se hace en base al p\_valor del estadístico z cuyo valor absoluto debe ser mayor a 1,96 ( $p < 0,05$ ).

De acuerdo a Kleinbaum (2005), el supuesto estadístico más importante que debe cumplir la regresión de Cox para su aplicación es el supuesto de proporcionalidad de riesgos, de ahí el nombre, modelo de riesgo proporcional de Cox. Sin embargo, este también puede verse afectado por la presencia de observaciones influenciales, la significancia de las variables y la presencia de multicolinealidad.

En este caso particular estudiamos el evento graduarse en la maestría de estadística y supongamos que los estudiantes se clasificaron en dos grupos A y B, entonces debe cumplirse que la razón de las tasas de riesgo entre ambos grupos debe ser constante a lo largo del tiempo de seguimiento. Esto significa que, si ambos grupos de estudiantes son observados durante 5 años con un HR en el grupo A de 3 respecto al grupo B, su HR debe ser aproximadamente igual a lo largo de los 5 años de seguimiento. En consecuencia, la razón del riesgo para dos individuos con el mismo vector de covariables no depende del tiempo  $t$ , y por tanto el riesgo de que ocurra el evento de interés, es constante para cada uno de los tiempos de supervivencia observados, tal como se indica en las siguientes expresiones matemáticas:

$$\hat{h}(t, X_j^*(t), X_i(t)) = \frac{h(t, X_j(t))}{h(t, X_i(t))} = \frac{h_0(t) \exp(\beta' X_j(t))}{h_0(t) \exp(\beta' X_i(t))} = \exp\{\beta' (X_j(t) - X_i(t))\} \quad (11)$$

Función de Supervivencia de Cox: A partir de la función de riesgo  $h(t, X_j(t))$ , se obtiene la función de supervivencia de Cox, esto es:

$$-\ln \hat{S}(t) = -\exp(\beta' X_j(t)) \ln \hat{S}_0(t) \quad (12)$$

de donde se obtiene, la siguiente expresión para la función de supervivencia de Cox:

$$\hat{S}_{Cox}(t, X_j(t)) = \hat{S}_0(t)^{\exp(\beta' X_j(t))} \quad (13)$$

Para la estimación de los coeficientes  $\beta$ , vector de parámetros de la regresión, se basa en el método de máxima verosimilitud, Andersen y colaboradores (Andersen et al., 1982), Gill (1983), Andersen y Borgan (1985), Pepe y Fleming (1989, 1991) y en Gill y Johansen (1990).

### 3. EVALUACIÓN DE LOS COEFICIENTES DEL MODELO DE REGRESIÓN DE COX

Según Lee (2001), una vez que se ha ajustado un modelo de regresión de Cox, debe ser evaluado para medir su significancia estadística, para esto existen tres pruebas de hipótesis que son asintóticamente equivalentes.

Test de razón de verosimilitud parcial (RVP): Esta prueba es obtenida mediante:

$$RVP = 2 \left\{ \log(L(\hat{\beta}_0)) - \log(L(\hat{\beta})) \right\} \quad (14)$$

donde  $\hat{\beta}_0$  son los coeficientes estimados de un primer modelo, llamado modelo nulo, y  $\hat{\beta}$  son los coeficientes estimados luego de ajustar el modelo a otras covariables no incluidas en el anterior. Bajo la hipótesis nula de que el coeficiente o los coeficientes son cero, este estadístico sigue una distribución Chi cuadrado con un grado de libertad, si se refiere sólo a un coeficiente o,  $m$  grados de libertad, si se refiere a  $m$  coeficientes del modelo.

Test de razón de Wald: La segunda prueba es el test de Wald ( $W$ ), éste estadístico proporciona una prueba por variable en vez de una medida global de significancia estadística. El estadístico se obtiene mediante la siguiente expresión:

$$W = (\hat{\beta} - \hat{\beta}_0)' \hat{\Sigma}^{-1} (\hat{\beta} - \hat{\beta}_0) \quad (15)$$

donde  $\hat{\Sigma}_{\beta}$  es la matriz de varianzas y covarianzas estimada. Este estadístico bajo la hipótesis nula de que los coeficientes son cero, se ajusta a la distribución normal estándar.

Test de Puntajes (score test). La tercera prueba es el test de puntajes, definido como  $U'U$  donde  $U$  es el vector de derivadas del  $\log(L(\beta))$  de orden, dado por:

$$U(\beta) = \sum_{i=1}^n \int_0^{\infty} [X_i(t) - \bar{X}(\beta, t)] dN_i(t) \quad (16)$$

e  $I$  es la matriz de información, dada por:

$$I(\beta) = \sum_{i=1}^n \frac{\sum_j Y_j(t) \exp(\beta' X_j(t)) [X_i(t) - \bar{X}(t)] [X_i(t) - \bar{X}(\beta, t)]}{\sum_j Y_j(t) \exp(\beta' X_j(t))} dN_i(t) \quad (17)$$

donde  $\bar{X}(\beta, t)$  es la media ponderada de las covariables para aquellos individuos que aun están en riesgo en el tiempo  $t$ , cuya expresión es:

$$\bar{x}(\beta, t) = \frac{\sum Y_j(t) \exp(\beta' X_j(t)) X_i(t)}{\sum Y_j(t) \exp(\beta' X_j(t))} \quad (18)$$

Lo anterior fue propuesto por Cox, y después fue demostrado por Efron (1977) y Cox y Oakes (1984), que la verosimilitud parcial contiene casi toda la información acerca de los  $\beta$ .

#### 4. RESULTADOS Y DISCUSIÓN

Durante el período de tiempo de 30 años (1973 – 2002) ingresaron al Instituto de Estadística Aplicada y Computación (IEAC) 225 estudiantes, de los cuales 117 (52%) lograron finalizar la escolaridad y de estos, sólo 77 lograron graduarse mostrándose una baja tasa de finalización de la Maestría. Ahora bien, de la totalidad de estudiantes que ingresaron a la maestría, el 34.2% logró finalizar el postgrado y, el 65.8% se retiró o abandonó durante la escolaridad o abandonó después de finalizada la escolaridad (Peña, 2006).

En la maestría de estadística se presenta el fenómeno de que una cierta proporción (34.2%) de los que finalizaron la escolaridad no logran finalizar la tesis de grado, el 65.8% de los estudiantes que finalizaron la escolaridad se graduaron y, el 34.2% restante abandonó el postgrado (Peña, 2006).

Ahora bien, para el grupo de estudiantes graduados, el tiempo de permanencia en la maestría de estadística oscilo entre 33 y 190 meses (entre 3 y 16 años aproximadamente), como el tiempo dedicado a la escolaridad de la maestría consta aproximadamente de 24 meses, entonces el tiempo para graduarse después de finalizada la escolaridad oscilo entre 9 y 166 meses (13.8 años), sin embargo, algo que influye en un tiempo muy prolongado para finalizar el postgrado es que algunos estudiantes al completar la escolaridad ingresan al campo laboral y, bajo estas condiciones es aun más difícil presentar la tesis y graduarse.

Otro factor que influye notablemente en la finalización de la maestría en estadística, es la profesión del estudiante que ingresa al postgrado. En el período de tiempo de 30 años (1973-2002) ingresaron ingenieros (la mayoría ingenieros de sistemas), Ciencias Sociales (Estadísticos y Economistas), Ciencias Exactas (Matemática, Física, Química y Biología), Humanidades (Educación Mención Matemática), Áreas de la salud (Farmacia, Bioanálisis y Medicina) entre otras profesiones. Cada uno de estos profesionales presentó diferentes grados de dificultad al inicio del postgrado, durante el postgrado o al momento de llevar a cabo el trabajo de tesis. Por ejemplo, los profesionales del área de la salud mostraron el mayor grado de dificultad debido a que de ese grupo, ninguno logró finalizar el postgrado, el riesgo de retirarse o abandonar la maestría fue 3.9 veces mayor respecto a profesionales de áreas científicas como Ingeniería de Sistemas, Ciencias o Estadísticos (Peña, 2006).

Al mismo tiempo que, los profesionales de áreas humanísticas, que en su mayoría de los que ingresaron a la maestría eran Licenciados en Educación Mención Matemática, mostraron un incremento del 78% en el riesgo de retirarse del postgrado respecto a profesionales de áreas científicas como Ingeniería de Sistemas, Ciencias o Estadísticos; aunque se tiene un 56% de probabilidad de que un egresado de Humanidades como Educación Mención Matemática no logre retirarse o abandonar la maestría en Estadística (Peña, 2006).

Por lo tanto, los profesionales que mostraron la mayor probabilidad de finalizar la maestría en estadística y por ende el menor grado de dificultad al momento de elaborar la tesis fueron, los ingenieros de sistemas, ciencias y estadísticos. Un elemento muy importante en estos grupos de profesionales es la preparación matemática y computación, específicamente en programación. Los profesionales de áreas científicas evidenciaron la mejor preparación en matemática y computación que los egresados de humanidades o el área de la salud. Los egresados de Humanidades, Licenciados en Educación Mención Matemática, mostraron una preparación en matemática equivalente a los profesionales de áreas científicas pero, menor preparación en computación ó programación.



La influencia de la variable profesión, también se observa en otros postgrados y otros países como es el caso de algunas universidades de Argentina y de habla inglesa, que según Paula Carlino (2005), la profesión, la prolongación del tiempo en la elaboración de la tesis y un número reducido de publicaciones del área son factores que definen la finalización o no de un postgrado por parte de sus estudiantes. Factores que en nuestro caso particular también se mostraron muy significativos con respecto a la finalización de la maestría en Estadística.

Por otra parte, independientemente de la profesión del estudiante que ingresó a la maestría en estadística fue que el alumno mostrara evidencia de dominio del idioma inglés al menos a nivel técnico de lectura, contara con el beneficio de la beca y tuviera experiencia previa con la estadística, la computación o ambas, estos factores garantizaron la finalización de la maestría y además, disminuyeron el riesgo de retirarse del postgrado en un 80% respecto aquellos estudiantes que no mostraban las características mencionadas (Peña, 2006).

Efectivamente, el hecho de que el estudiante no abandone o presente un retiro formal de la maestría, incrementa la probabilidad de finalizar el postgrado y necesariamente de graduarse, tomando en consideración las covariables que pudieran incidir en la obtención del grado, como se muestra a continuación:

Estimación de la función de supervivencia mediante el estimador de Kaplan y Meier (KM) para el evento graduarse: De los 225 estudiantes que al menos han estado inscritos para cursar la maestría en estadística, 77 lograron graduarse y los restantes 148 abandonaron o presentaron un retiro formal. De los estudiantes graduados, hubo algunos (6) que sus tiempos de supervivencia estuvieron entre los 135 y 190 meses (entre 11 y 16 años aproximadamente de permanencia en la Maestría), estos datos para algunos modelos resultaron influyentes, por lo cual se tomó la decisión de eliminarlos y así garantizar modelos mejor ajustados.

La función de supervivencia estimada de KM, fue obtenida a partir de 219 estudiantes que al menos estuvieron inscritos para cursar la maestría en estadística. De la Tabla 1 se observa que el evento graduarse ocurrió en 71 estudiantes, la mediana de supervivencia fue de 61 meses, en el 50% de los estudiantes ocurrió el evento graduarse como máximo a los 61 meses de permanencia en la maestría que se corresponde a un poco más de cinco (5) años.

Además, se tiene un 95% de confianza de que el tiempo de supervivencia de los estudiantes para obtener el grado de Magister Scientiae en Estadística oscila entre

59 y 69 meses respectivamente (entre 5 y 6 años aproximadamente), tal como se indica en la Tabla 1.

**Tabla 1**  
Estadísticas básicas de la función de supervivencia de Kaplan y Meier para el Evento graduarse en la maestría de estadística.

N	eventos	media	ee(media)	Mediana	LCI (0.95)	LCS (0.95)
219	71	62.77	2.77	61	59	69

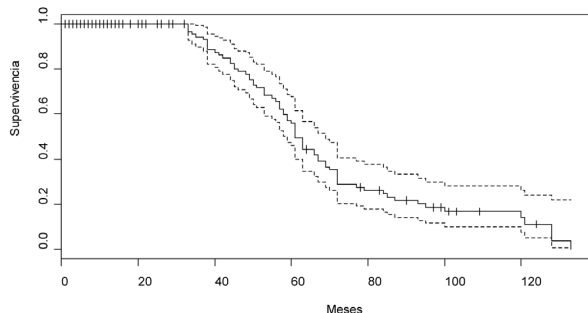
Donde:  
**n:** Número de estudiantes.  
**eventos:** Número total de estudiantes graduados.  
**media:** Media aritmética del tiempo de supervivencia.  
**ee(media):** Error estándar de la media de supervivencia.  
**mediana:** Mediana de supervivencia.  
**LCI (0.95):** Límite de confianza inferior del 95% para la mediana de supervivencia.  
**LCS (0.95):** Límite de confianza superior del 95% para la mediana de supervivencia.

El tiempo medio en que los estudiantes de la maestría finalizaron la tesis fue a los 62.8 meses de permanencia en el postgrado (5.2 años), entonces el tiempo medio dedicado a la tesis después de finalizada la escolaridad fue de 38.8 meses (3.2 años). El rango de tiempo de permanencia en la maestría para entregar la tesis se encuentra entre 33 y 190 meses, cuya mediana fue de 61 meses ó de 5 años y un mes. El 50% de los estudiantes que finalizaron la escolaridad de la maestría en estadística permanecieron máximo 5 años y un mes hasta el momento en que entregaron la tesis. Del grupo de estudiantes que logró graduarse, 67 (87%) eran de áreas profesionales científicas (ingenieros, ciencias y estadísticos), 10 (13%) eran del área humanística, educación mención matemática y ninguno del área de la salud logró graduarse.

Del grupo de estudiantes que se graduaron, 53 (68.8%) tenían el beneficio de la beca y además, mostraron evidencias de experiencia previa con la estadística, la computación o ambas. También de ese mismo grupo, 74 estudiantes (96.1%) mostraron evidencias del dominio del idioma inglés al menos a nivel técnico de lectura y, 37 estudiantes (48.1%) ingresaron a la maestría con motivo de que la estadística sería una herramienta útil en los proyectos y trabajos de investigación.

El Gráfico No. 1 (función de supervivencia) es la curva de supervivencia de los estudiantes que ingresaron a la maestría de estadística durante los años 1973 y 2004. Se observa un patrón constante durante los primeros 30 meses de permanencia en la maestría, siendo durante ese intervalo de tiempo donde se tiene la mayor proporción de observaciones censuradas (líneas verticales), bien sea abandono o retiro formal de la maestría, luego entre los 30 y 100 meses de permanencia en la maestría la tendencia es decreciente, así como también disminuyó la proporción de tiempos censurados después de finalizada la escolaridad.

Gráfico No. 1. Función de Supervivencia estimada de Kaplan y Meier  
Evento: Graduarse en el IEAC, 1973 - 2004



**Modelo de riesgo proporcional de Cox para el evento graduarse:** Se obtuvieron dos modelos de regresión de riesgo proporcional de Cox, para el evento graduarse en la maestría de estadística.

En las Tablas 2 y 3 se muestran las covariables del Modelo 1 de Regresión de Cox, que resultaron significativas al 5% ( $p < 0.05$ ), así como el exponencial de los coeficientes (Hazard Ratio) y sus intervalos de confianza del 95%.

Tabla 2.  
Estimación de los coeficientes para el Modelo 1 de Regresión de Cox  
Evento graduarse en la maestría de estadística.

Covariable	coef	exp(coef)	ee(coef)	z	p
Beca	0.684	1.982	0.260	2.64	0.0084
Edad (33)	-0.607	0.545	0.290	-2.10	0.0360
Investigación	0.710	2.034	0.261	2.72	0.0065

Donde:  
**coef:** Coeficiente estimado mediante el Modelo 1 de Cox.  
**exp(coef):** Exponente del coeficiente (Hazard Ratio).  
**ee(coef):** Error estándar del coeficiente.  
**z:** Estadístico contraste para la significación del coeficiente.  
**p:** p\_valor o probabilidad de significancia del coeficiente.

Tabla 3  
Exponencial de los coeficientes para el Modelo 1 de Regresión de Cox  
Evento graduarse en la maestría de estadística.

Covariable	exp(coef)	exp(-coef)	LCI (95%)	LCS (95%)
Beca	1.982	0.505	1.192	2.296
Edad (33)	0.545	1.835	0.309	0.961
Investigación	2.034	0.492	1.220	3.394

De acuerdo al Modelo 1 de Cox el evento graduarse en la maestría de estadística depende de las covariables beca, edad media al inicio de la maestría (33 años) e investigación como motivo de ingreso al postgrado.

Para los estudiantes que tenían el beneficio de la beca durante su tiempo de permanencia en la maestría, el riesgo de graduarse se incrementó en un 98.2% respecto aquellos estudiantes que no contaban con este beneficio. La probabilidad de graduarse para los estudiantes becados en la maestría de estadística aumento a partir de un 19.2%.

La edad de los estudiantes al ingresar a la maestría, también mostro un efecto significativo respecto al riesgo de

graduarse. Para los estudiantes cuyas edades al momento de ingresar a la maestría eran en promedio 33 o más años, el riesgo de graduarse disminuyó en un 45.5% respecto al grupo de estudiantes más jóvenes. Los estudiantes de mayor edad tienden a permanecer más tiempo en la maestría, mientras que los estudiantes más jóvenes presentan un mayor riesgo de graduarse. La probabilidad de graduarse para los estudiantes que para el momento de ingreso a la maestría tenían en promedio 33 años o más años de edad, disminuyó desde un 69.13% hasta un 3.8%.

Otra variable que el modelo 1 arrojó estadísticamente significativa fue, investigación, algunos estudiantes de los que ingresaron a la maestría de estadística lo hizo con la intención de que esta serviría como herramienta adicional para sus trabajos de investigación. Si el motivo de ingreso a la maestría era la investigación, esta variable incremento el riesgo de graduarse en 2.03 veces respecto al grupo de estudiantes que ingreso por otro motivo. La probabilidad de graduarse para estos estudiantes aumento desde un 22%. Estos resultados coinciden en cierta forma con el trabajo de Delamont, S. y Atkinson, P. (2001) citados por Paula Carlini (2005) donde se indica que, si los estudiantes de postgrado pertenecen o se integran a un grupo de investigación esta condición es definitiva a la hora de finalizar la tesis. En nuestra Universidad este comportamiento se observa generalmente en los estudiantes de la maestría que se desempeñan como profesores cuya categoría era Asistente o inferior y, además ingresaron al postgrado con motivo de trabajo de ascenso, esta condición pues les permite integrarse a un grupo de investigación del postgrado o de la facultad en la cual se desempeña como profesor y, a su vez, le permite finalizar la maestría en estadística.

En las Tablas 4 y 5 se muestran las covariables del Modelo 2 de Regresión de Cox, que resultaron significativas al 5% ( $p < 0.05$ ), así como el exponencial de los coeficientes (Hazard Ratio) y sus intervalos de confianza del 95%.

Tabla 4  
Estimación de los coeficientes para el Modelo 2 de Regresión de Cox  
Evento graduarse en la maestría de estadística.

Covariable	coef	exp(coef)	ee(coef)	z	p
Sexo	-0.598	0.55	0.260	-2.298	0.0220
Nacionalidad	1.039	2.83	0.375	2.768	0.0056
Prof. Otra Univ	0.274	1.32	0.422	0.649	0.5200
Prof. ULA, Vzla	0.840	2.32	0.286	2.939	0.0033

Donde:  
**coef:** Coeficiente estimado mediante el Modelo 2 de Cox.  
**exp(coef):** Exponente del coeficiente (Hazard Ratio).  
**ee(coef):** Error estándar del coeficiente.  
**z:** Estadístico contraste para la significación del coeficiente.  
**p:** p\_valor o probabilidad de significancia del coeficiente.

Tabla 5  
Exponencial de los coeficientes para el Modelo 2 de Regresión de Cox  
Evento graduarse en la maestría de estadística.

covariable	exp(coef)	exp(-coef)	LCI (95%)	LCS (95%)
Genero	0.55	1.818	0.330	0.916
Nacionalidad	2.83	0.354	1.354	5.899
Prof. Otra Univ.	1.32	0.760	0.575	3.010
Prof. ULA, Vzla.	2.32	0.432	1.323	4.059

De acuerdo al Modelo 2 de Cox el evento graduarse en la maestría de estadística depende de las covariables género, nacionalidad y profesor universitario.

El género es una variable significativa para explicar el evento graduarse en la maestría. Si el estudiante es de género masculino el riesgo de graduarse disminuye en un 45% respecto a un estudiante de género femenino. En consecuencia, los estudiantes de género femenino presentan un mayor riesgo de graduarse, mientras que los estudiantes de género masculino, presentan una mayor supervivencia en la maestría. La probabilidad de que un estudiante de género masculino lograra graduarse en la maestría disminuye desde un 67% hasta un 8.5%.

Para los estudiantes de nacionalidad Venezolana, el riesgo de graduarse se incrementa en 2.8 veces respecto a los estudiantes Extranjeros. La probabilidad de que un estudiante venezolano finalice la maestría en estadística se incrementa desde un 35%.

Si el estudiante para el momento en que desempeña sus estudios de maestría labora como profesor de otra universidad, diferente a la ULA, el riesgo de graduarse se incrementa en un 32% respecto al grupo de estudiantes que durante su permanencia en el postgrado no se desempeñan como profesores. Sin embargo, este es un resultado no significativo a un nivel del 5%, esto es, el riesgo de graduarse entre estudiantes profesores de otra universidad es igual al riesgo de graduarse de los estudiantes que no laboran como profesores.

Para los estudiantes de la maestría que se desempeñaban como profesores de la Universidad de Los Andes (ULA, Mérida) el riesgo de graduarse es 2.3 veces mayor respecto al grupo de estudiantes que no se desempeñaban como profesores. Si un estudiante de la maestría para el momento que cursa estos estudios se desempeña como profesor universitario en la ULA, la probabilidad de graduarse se incrementa desde un 32%.

Ahora bien, al comparar el riesgo de graduarse de estudiantes que se desempeñen como profesores de otra universidad respecto a estudiantes que se desempeñen como profesores de la ULA, se tiene que, si el estudiante se desempeña como profesor de otra universidad el riesgo de graduarse disminuye en un 63.21% respecto al grupo de estudiantes que se desempeñan como profesores la ULA, la probabilidad de graduarse para los estudiantes en estas condiciones disminuye desde un 91.1% hasta un 52%.

#### 4.1. Validez de los modelos de Regresión de Cox

Las Tablas 6 y 7 muestran que los modelos de regresión de Cox son estadísticamente significativos al 0.1% ( $p$

$< 0.001$ ), de acuerdo a los tres criterios que allí se indican.

**Tabla No. 6**  
Pruebas para medir la significancia estadística del Modelo 1 de Regresión de Cox. Evento graduarse en la maestría de estadística.

Test	Estadístico	Grados de Libertad	p
Test de razón de verosimilitud	16.6	3	0.000866
Test de Wald	15.4	3	0.001490
Test de puntajes (logrank)	15.8	3	0.001270

**Tabla 7**  
Pruebas para medir la significancia estadística del Modelo 2 de Regresión de Cox. Evento graduarse en la maestría de estadística.

Test	Estadístico	Grados de Libertad	p
Test de razón de verosimilitud	16.4	4	0.00254
Test de Wald	16.4	4	0.00250
Test de puntajes (logrank)	16.9	4	0.00201

#### de Regresión de Cox

En las Tablas 8 y 9 se verifica el cumplimiento del supuesto de riesgo proporcional tanto en forma general como para cada una de las covariables de los modelos 1 y 2 de Regresión de Cox, se verifica bajo la hipótesis nula, el riesgo de que ocurra el evento graduarse se mantiene constante en cada uno de los tiempos de supervivencia observados, para cada una de las covariables de los Modelos 1 y 2 de Regresión de Cox.

**Tabla 8**  
Test de riesgos proporcionales para el Modelo 1 de Regresión de Cox para el Evento graduarse en la maestría de estadística.

Covariables	Rho	chi2	p
Beca	-0.0238	0.03910	0.843
Edad (33)	0.0348	0.08654	0.769
Investigación	-0.0080	0.00478	0.945
GLOBAL	NA	0.11975	0.989

**Tabla 9**  
Test de riesgos proporcionales para el Modelo 2 de Regresión de Cox para el Evento graduarse en la maestría de estadística.

Covariables	Rho	chi2	p
Sexo	-0.0514	0.2058	0.6501
Nacionalidad	-0.2135	3.9907	0.0458
Prof. Otra Univ	-0.0899	0.6900	0.4062
Prof. ULA, Vzla	-0.0232	0.0489	0.8250
GLOBAL	NA	4.6861	0.3211

Al comparar cada uno de los  $p$ \_valor con un nivel de significancia del 5%, se observa que para cada una de las covariables de los Modelos 1 y 2 de Regresión de Cox, no se tienen evidencias suficientes como para rechazar el cumplimiento del supuesto de riesgo proporcional ( $p > 0.05$ ).

Del  $p$ \_valor global para ambos modelos de Regresión de Cox, se tiene que para los modelos en términos generales se satisface el supuesto de riesgo proporcional ( $p > 0.05$ ).

## 5. CONCLUSIÓN

La profesión del estudiante que ingresa a la Maestría de Estadística es una primera variable importante, debido a que garantiza finalizar el postgrado. Por ejemplo, la maestría esta mejor orientada para profesionales de áreas científicas y humanísticas, como ingenieros de sistemas, matemáticos, físicos, químicos, biólogos y educación mención matemática. Esto es debido a la preparación previa en matemática y computación que requieren los estudiantes de esta maestría.

Es importante que el estudiante, al momento de ingresar a la maestría en estadística, cuente con el beneficio de la beca (el postgrado requiere dedicación exclusiva), muestre evidencia de dominio del idioma inglés al menos en nivel técnico de lectura y, que presenten experiencia previa a nivel básico de estadística, computación o ambas.

Uno de los motivos por los cuales se ingresa a la maestría en estadística debería ser la investigación, ya que actualmente la estadística forma parte de cualquier trabajo o proyecto de investigación. La estadística es la parte de la ciencia, que da la última palabra en cuanto al éxito o fracaso de un proyecto o trabajo de investigación.

En consecuencia, la institución responsable de la formación de estadísticos en el nivel de postgrado debería promover el dictado de cursos de estadística básicos,

enfocados al manejo de metodologías estadísticas y herramientas computacionales necesarias para cualquier investigador de otras áreas del conocimiento que requieran del uso de la estadística. ©

## AGRADECIMIENTOS

Al CDCHTA (ULA) por el financiamiento otorgado al proyecto de código FA-366-05-09-B. Al Instituto de Estadística Aplicada y Computación (IEAC) bajo la coordinación de la profesora Elizabeth Torres, Facultad de Ciencias Económicas y Sociales (FACES) de la Universidad de Los Andes (ULA), Mérida (Venezuela).

**Jesús Alberto Peña Guillón**, Licenciado en Educación Mención Matemática. Magister scientarum en Estadística. Profesor Asociado de la Universidad de Los Andes, Facultad de Farmacia y Bioanálisis. Cátedra: Bioestadística. PPI, Nivel A.

**Rafael Eduardo Borges Peña**, Licenciado en Ciencias Estadísticas. Magister scientarum en Estadística Aplicada. Actualmente es estudiante de Doctorado en Salud Pública, Mención Bioestadística. Profesor Asociado de la Universidad de Los Andes, Facultad de Ciencias Económicas y Sociales Catedra: Estadística Aplicada. PPI, Nivel B.

## BIBLIOGRAFÍA

- Andersen, Per Kragh y Borgan, Ærnulf. (1985). Counting Process Model for Life History Data: A Review. *Scandinavian Journal of Statistics*, 12: 97 – 158.
- Andersen, Per Kragh, Borgan, Ærnulf., Gill, Richard y Keiding, Niels. (1982). Linear Nonparametric Test for Comparison of Counting Processes with Applications to Survival Data (with discussion). *International Statistical Review*, 50: 219 – 258.
- Andersen, Per Kragh, Borgan, Ærnulf, Gill, Richard y Keiding, Niels. (1993). *Statistical Models Based on Counting Processes*. N.Y.: Springer – Verlag.
- Borges Peña, Rafael Eduardo. (2002). Análisis de Supervivencia Aplicado a un caso de Diálisis Renal: Diálisis Peritoneal en el Hospital Clínico Universitario de Caracas y Hemodiálisis en el Hospital de Clínicas Caracas, 1980 – 2000. Tesis de M.Sc. en Estadística Aplicada, Mérida, Venezuela: Instituto de Estadística Aplicada y Computación, Universidad de los Andes. (Disponible en: <http://tesis.saber.ula.ve>).
- Borges Peña, Rafael Eduardo. (2004). VI Jornadas Académicas: Análisis de Supervivencia Básico utilizando el lenguaje R. Área de Postgrado en Estadística y Actuarial. Escuela de Estadística y Ciencias Actuariales. Facultad de Ciencias Económicas y Sociales. Universidad Central de Venezuela. Caracas – Venezuela.
- Bloom, Lynn Z. (1981). Why Graduate Students Can't Write: Implications of Research on Writing Anxiety for Graduate Education. *Journal of Advanced Composition*, 2 (1-2): 103 – 117.
- Breslow, Norman E. (1974). Covariance analysis of censored survival data. *Biometrics*, 30: 89 – 100.
- Carlino, Paula. (2005). ¿Por qué no se completan las tesis en los postgrados? Obstáculos percibidos por maestrandos en curso y Magistri exitosos. *Educere – Tránsito de lo publicado*, 30: 415 – 420.



## BIBLIOGRAFÍA

- Castaño, Elkin, Gallón, Santiago, Gómez, Karoll y Vásquez, Johanna. (2004). Deserción estudiantil universitaria: una aplicación de modelos de duración. *Lecturas de Economía*, 60: 39 – 65.
- Castaño, Elkin, Gallón, Santiago, Gómez, Karoll y Vásquez, Johanna. (2006). Análisis de los factores asociados a la deserción y graduación estudiantil universitaria. *Lecturas de Economía*, 65: 9 – 36.
- Castaño, Elkin, Gallón, Santiago, Gómez, Karoll y Vásquez, Johanna. (2008). Análisis de los factores asociados a la deserción estudiantil en la Educación Superior: un estudio de caso. *Revista de Educación*, 345: 255 – 280.
- Cox, David Roxbee (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society: Series B*, 34: 187 – 220.
- Cox, David Roxbee y Oakes, David (1984). *Analysis of Survival Data*. Londres: Chapman & Hall.
- Collett, David (2003). *Modelling Survival Data in Medical Research*. 2da Edición. Boca Ratón: Chapman & Hall.
- Delamont, Sara y Atkinson, Paul (2001). Doctoring Uncertainty: Mastering Craft Knowledge. *Social Studies of Science*, 31(1): 87-107.
- Efron, Brad (1977). The efficiency of Cox's likelihood function for censored data. *Journal of American Statistical Association*, 72: 557 – 565.
- Fleming, Thomas R. y Harrington, David P (1991). *Counting Processes and Survival Analysis*. N.Y.: John Wiley & Sons, Inc.
- Furtado, Magdalena (2003). Trayectoria Educativa de los jóvenes: El problema de la deserción. Cuaderno de Trabajo Serie Aportes TEMS. N°22. Comisión de Transformación de la Educación Media Superior en el Uruguay /ANEP – Codicen
- Gill, Richard D. (1983). Large Sample Behavior of the Product – limit Estimator on the Whole Line. *Annals of Statistics*, 11: 49 – 58.
- Gill, Richard D. y Johansen, Soren (1990). A Survey of Product – integration with a View Towards applications in Survival Analysis. *Annals of Statistics*, 18: 1501 – 1550.
- Giovagnoli, Paula I (2002). Determinantes de la deserción y graduación universitaria: Una aplicación utilizando modelos de duración. Departamento de Economía, Facultad de Ciencias Económicas, Universidad de la Plata. Documento de trabajo Nro. 37. Argentina.
- Greenwood, Major (1926). The natural duration of cancer. *Reports on Public Health and Medical Subjects*, 33: 1 – 26, Londres: Her Majesty's Stationery Office.
- Harrington, David P. and Fleming, Thomas R. (1982). A class of rank test procedures for censored survival data. *Biometrika*, 69: 553 – 566.
- Hosmer, David W. y Lemeshow, Stanley (1999). *Applied Survival Analysis: Regression Modeling of Time to Event Data*. N.Y.: Jhon Wiley & Sons, Inc.
- Johnson, Norman L. y Kotz, Samuel (1997). *Discrete Multivariate Distributions*. New York. Jhon Wiley & Sons, Inc.
- Kaplan, Edward L. y Meier, Paul (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53: 457 – 481.
- Klein, John P. y Moeschberger, Melvin L. (1997). *Survival Analysis: Techniques for Censored and Truncated Data*. New York: Springer – Verlag.
- Kleinbaum, David G. (2005). *Survival Analysis: A self – learning Text*. Segunda edición. New York: Springer – Verlag.
- Latona, Kris y Browne, Mairead (2001). Factors Associated with Completion of Research Higher Degrees. Higher Education Series Report No. 37. Australia: Department of Education, Training and Youth Affairs.
- Lavado, Pablo y Gallegos, José (2005). La dinámica de la deserción escolar en el Perú: Un enfoque usando modelos de duración. Centro de Investigación de la Universidad del Pacífico, Lima, Perú.
- Lee, Elisa T. y Wang, John W. (2003). *Statistical Methods for Survival Data Analysis*. New Jersey.: Wiley – Interscience.

## BIBLIOGRAFÍA

- Mantel, Nathan (1966). Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemotherapy Reports*, 50: 163 – 170.
- Mathsoft, Inc. (1999). *S – PLUS 2000 Guide to Statistics, Volume 2*. Data Analysis Products Division, Mathsoft, Seattle, WA.
- Meyer, Bruce D. (1990). Unemployment insurance and unemployment spells. *Econometría*, 58 (4): 757 – 782.
- Pepe, Margaret S. y Fleming, Thomas R. (1989). Weighted Kaplan – Meier Statistic: A Class of Distance test for Censored Survival Data. *Biometrics*, 45: 497 – 507.
- Pepe, Margaret S. y Fleming, Thomas R. (1991). Weighted Kaplan – Meier Statistic: Large Sample and Optimality Considerations. *Journal of the Royal Statistical Society, Serie B*, 53: 341 – 352.
- Peña Guillén, Jesús Alberto. (2006). Aplicación del Análisis de Supervivencia a los estudiantes de la Maestría en Estadística, 1973 – 2004. Tesis de M.Sc. en Estadística, Mérida, Venezuela: Instituto de Estadística Aplicada y Computación, Universidad de Los Andes.
- Peto, Richard y Peto, Julian (1972). Asymptotically efficient rank invariant test procedures (with discussion). *Journal of the Royal Statistical Society: Series A*, 135: 195 – 206.
- Prentice, Ross L. y Gloeckler, Lynn A. (1978). Regression analysis of grouped survival data with application to breast cancer data. *Biometrics*, 34: 57 – 67.
- R Development Core Team (2004). *R: A lenguaje and environment for statistical computing*. Viena, Austria: R Foundation for Statistical Computing.
- Rosas, Ana Karina; Flores, Daniela y Valarino, Elizabeth (2006). Rol del tutor de tesis: Competencias, Condiciones Personales y Funciones. *Revista de Investigación y Postgrado, Universidad Pedagógica Experimental Libertador*, 21 (1), 153 – 185. Caracas – Venezuela.
- S original by Terry Therneau and ported by Thomas Lumley (2004). survival: Survival analysis, including penalized likelihood. R package version 2.13 – 2.
- Schoenfeld, David (1982). Partial residuals for the proportional hazard regression model. *Biometrika*, 69: 239 – 241.
- Therneau, Terry M., Grambsch, Patricia M. y Fleming, T.R. (1990). Martingale – based residuals for survival models. *Biometrika*, 77: 147-160.
- Therneau, Terry M. y Grambsch, Patricia M. (2000). *Modeling Survival Data: Extending the Cox Model*. N.Y.: Springer-Verlag.



Los científicos también aclaran por qué muchos amores de verano se desvanecen cuando acaba el estío, al ser como «un espejismo», con el consiguiente riesgo de que uno de los dos se quede «colgado».

Viene de la pág.310

Nicole Prashak-Rieder y Matthaues Willeit, del Centro de Adicción y Salud Mental de la Universidad de Toronto, comprobaron con tep, tomografía por emisión de positrones, la existencia de mayores cantidades del transportador de serotonina en época estival prácticamente en todas las áreas cerebrales.

Continúa en la pág.330