

## Tamaño de la Muestra en Análisis de Sobrevida

JAVIER A. CASTAÑEDA \*

ADRIANA PÉREZ †

JACKY F.A. GIL ‡

---

### Abstract

This article presents some formulae and tables of easy use to handle the minimum sample size requirements in survival analysis. Sample size formulae are presented to estimate a specific hazard ratio, to compare two hazard ratios under several scenarios of censoring and follow up as well as to test the proportional hazard ratio assumption. A discussion about the advantages and disadvantages on the use of the presented formulae is presented.

*Key Words:* Type I error, Power, Hazard function, Censoring, Follow up.

### Resumen

Este artículo presenta algunas expresiones y tablas de fácil manejo para la determinación del tamaño de la muestra necesario en la estimación de una tasa de riesgo específica y para comparar las tasas de

---

\*Estadístico. Joven investigador Unidad de Epidemiología Clínica y Bioestadística Facultad de Medicina. Pontificia Universidad Javeriana; Carrera 7 No 40-62, HUSI, Piso 2, Bogotá, Colombia. E-mail:javier.castaneda@javeriana.edu.co

†Profesora Asociada, Unidad de Epidemiología Clínica y Bioestadística, Facultad de Medicina. Pontificia Universidad Javeriana. Profesora Asistente, Departamento de Matemáticas y Estadística. Universidad Nacional de Colombia, Bogotá, Colombia

‡Instructor Asistente. Unidad de Epidemiología Clínica y Bioestadística, Facultad de Medicina. Pontificia Universidad Javeriana, Bogotá, Colombia

riesgo entre dos grupos para varios escenarios de censura y seguimiento en un análisis de sobrevida. Se discuten algunas ventajas y desventajas en el empleo de las fórmulas presentadas.

*Palabras Claves* Error tipo I, potencia, tasa de riesgo, censura, seguimiento.

## 1 Introducción

Un adecuado tamaño de la muestra es fundamental para garantizar la validez y solidez de las conclusiones generadas a partir de un estudio. Para su cálculo, los investigadores deben considerar la magnitud de la diferencia entre las estimaciones de los parámetros de interés, los supuestos propios de dichos estimadores, el diseño del estudio y las pruebas estadísticas involucradas en la fase de análisis. Este artículo se centra en la determinación del número de observaciones requeridas bajo un proceso de muestreo aleatorio simple, cuando se planea un análisis de sobrevida. La decisión final sobre el tamaño de la muestra se toma teniendo en cuenta la factibilidad del estudio con base en los recursos disponibles.

En estudios donde se involucra un análisis de sobrevida, la variable de interés es el tiempo transcurrido hasta la ocurrencia del evento y no únicamente el número de eventos reportados. Como eventos se pueden entender: la muerte, el daño de una pieza de una máquina, la aparición de una enfermedad, la falla de un circuito, la recaída de una enfermedad, etc. En estos estudios, los datos son tradicionalmente analizados mediante el uso de tablas de vida cuando no se presenta censura. El análisis por tablas de vida involucra el tiempo al evento para cada observación y provee estimadores más efectivos del tiempo de sobrevida, en comparación con los estimadores generados mediante la proporción de sobrevivientes después de un tiempo determinado, Lachin (1981) y Elandt-Johnson y Johnson (1980). Cuando el estudio presenta censura o no la presenta, el análisis de los datos se puede realizar usando modelos de riesgos proporcionales de Cox, Kleinbaum (1995), como también mediante el uso de métodos no paramétricos, Lee (1992).

En este artículo se evalúan métodos paramétricos bajo el supuesto que el tiempo de sobrevida se distribuye exponencialmente con tasa de riesgo  $\lambda$ . Con el ánimo de tener una traducción adecuada al español del término *hazard ratio*, aquí nos referiremos a esta expresión como tasa de riesgo. Se denota la proporción de sobrevivientes  $P_s(t)$  para algún tiempo  $t$ , como  $P_s(t) = \exp(-\lambda t)$ . Bajo éste supuesto el  $\log [P_s(t)]$  es linealmente decreciente en el tiempo con

pendiente  $\lambda$ . La función de riesgo esta dada en la ecuación (1) e indica para una unidad de tiempo dado el potencial instantáneo de la ocurrencia del evento de interés, dado que el individuo ha sobrevivido hasta el instante  $t$ , Kleinbaum (1995).

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

Las fórmulas que se muestran son empleadas para determinar el tamaño de la muestra requerido al realizar una prueba de hipótesis sobre una tasa de riesgo específica, cuando se comparan las tasas de riesgo para dos grupos y cuando se emplea la prueba log-rank para comparar dos tasas de riesgo. Todas estas fórmulas tienen como supuesto específico que la distribución de la función de sobrevivencia es exponencial. Autores interesados en comparaciones para más de dos grupos pueden consultar Toulbee y Symons (1993).

Este artículo presenta resultados parciales del proyecto desarrollado por Pérez et al, (2001), el cual contemplo el cálculo de los tamaños de la muestra para niveles de error tipo I de 0.01 y 0.05; en conjunto con niveles de error tipo II de 0.1 y 0.2.

Se presentan algunas fórmulas de uso frecuente en estudios de sobrevivencia y algunas tablas generadas por su cálculo, para niveles de error tipo I y tipo II de 0.05 y 0.20, respectivamente, que ofrecen a los investigadores no especializados en estadística un fácil acceso a la determinación del mínimo tamaño de la muestra requerido en su estudio.

En la sección 2 se presenta la notación empleada; en la sección 3 se presenta la revisión de las fórmulas implementadas, en la sección 4 se muestran ejemplos y tablas del mínimo tamaño de la muestra para las formulas implementadas y en la sección 5 se presenta una breve discusión.

## 2 Notación

Se empleará la siguiente notación a lo largo del presente artículo. El nivel de error tipo I se notará como  $\alpha$ , el nivel de error tipo II se notará como  $\beta$ ,  $Z_{1-\alpha/2}$  y  $Z_{1-\beta}$  corresponden a los valores críticos de la distribución normal estándar a los niveles  $1 - \alpha/2$  y  $1 - \beta$  respectivamente,  $\Phi(\cdot)$  identifica la función de la distribución normal estándar. En las fórmulas donde se comparan dos grupos,  $Q_t$  representa la proporción de individuos en el grupo tratamiento con respecto al total de individuos en el estudio de interés,  $Q_c$  representa la proporción de individuos en el grupo control con respecto al total de individuos en el

estudio ( $Q_t = 1 - Q_c$ ),  $\lambda_t$  y  $\lambda_c$  son las tasas de riesgo estimadas de los grupos tratamiento y control respectivamente,  $\bar{\lambda} = Q_t\lambda_t + Q_c\lambda_c$  representa el promedio ponderado de las tasa de riesgo, el mínimo tamaño de la muestra total requerido en el estudio se representa por  $n$ ,  $n_c = nQ_c$  representa el número de individuos en el grupo control y  $n_t = n - n_c$  representa el número de individuos en el grupo tratamiento.

### 3 Algunas fórmulas para calcular el tamaño de la muestra

En la sección 3.1 se presenta la fórmula para el cálculo del tamaño de la muestra cuando se desea detectar una tasa de riesgo específica, mediante el uso de la determinación del número de eventos o desenlaces. En la sección 3.2 se presentan las fórmulas más frecuentemente usadas cuando se desea comparar las tasas de riesgo entre dos grupos que denotaremos tratamiento y control, bajo diferentes escenarios. En las sub-secciones 3.2.1 a 3.2.4 se presentan las fórmulas para calcular el tamaño de la muestra cuando: (3.2.1) ambos grupos son independientes sin presencia de censura, (3.2.2) ambos grupos independientes con censura, (3.2.3) ambos grupos independientes con censura progresiva y (3.2.4) ambos grupos independientes con reclutamiento limitado y censura. Finalmente, en la sección 3.3 se presenta la fórmula para determinar el tamaño de la muestra basado en el mínimo número de eventos o desenlaces ha observarse durante un estudio, cuando se desea evaluar la diferencia de las tasas de riesgo en ambos grupos usando la prueba log-rank.

#### 3.1 Tamaño de la muestra requerido para detectar una tasa de riesgo específica

Esta fórmula se basa en el mínimo número de eventos o desenlaces ha observarse en un estudio para estimar una tasa de riesgo específica  $\lambda_0$  y fue reportada por Law (1996). Como es de esperarse, se debe especificar el nivel de error tipo I y tipo II a probar la hipótesis con referencia a la tasa de riesgo específica. Se denotara la hipótesis de interés como  $H_0 : \lambda = \lambda_0$  vs  $H_1 : \lambda \neq \lambda_0$ . En la ecuación (2) se presenta la fórmula para determinar el número mínimo de desenlaces ( $m$ ) requeridos para probar esta hipótesis para un nivel de potencia  $1 - \beta$ , con un nivel de confianza  $1 - \alpha$ . En forma viceversa, en la ecuación (3) se presenta la fórmula para calcular la potencia con base en el número de desenlaces de interés calculados de la ecuación (2).

$$m = \left[ \frac{(1 + \lambda_0)}{(1 - \lambda_0)} \right]^2 [Z_{1-\alpha/2} + Z_{1-\beta}]^2 \quad (2)$$

$$1 - \beta = \Phi(Z_{1-\beta}) \quad (3)$$

donde

$$Z_{1-\beta} = \sqrt{\frac{m}{\left(\frac{1+\lambda_0}{1-\lambda_0}\right)^2} - Z_{1-\alpha/2}^2}$$

Esto quiere decir que el tamaño de la muestra ( $n$ ) requerido se debe calcular ajustando por la probabilidad ( $u$ ) del evento de interés en la población. En otras palabras  $n = m/u$ .

### 3.2 Comparación de dos tasas de riesgo

Si en un estudio de sobrevida se cuenta con dos grupos independientes y los individuos de ambos grupos son seguidos durante un tiempo  $t$  hasta la ocurrencia del evento de interés, se puede estar interesado en probar la hipótesis de igualdad de las tasas de riesgo  $H_0 : \lambda_t = \lambda_c$  vs  $H_1 : \lambda_t \neq \lambda_c$ . Esta hipótesis puede ser probada usando el estadístico  $Z = (\lambda_t - \lambda_c)/S$ , el cual bajo la hipótesis nula se distribuye normal estándar, Lachin (1981).

El tiempo transcurrido por la unidad de interés, entre el ingreso al estudio hasta la ocurrencia del evento o hasta el momento cuando se detiene el estudio, se conoce con el nombre de tiempo de sobrevida. Usualmente, el periodo de observación llega a su fin y no todas las unidades del estudio han presentado el evento de interés. Este fenómeno, se denomina censura, y se define como aquella observación donde el evento de interés no ha ocurrido al momento de terminar el periodo de observación, Cox y Oakes (1984) y Hosmer y Lemeshow (1999). La censura en los estudios puede ocurrir por: finalización del estudio, por pérdida de seguimiento de las unidades del estudio, por muerte o por causa de egreso ajena al estudio, entre otras.

En las siguientes fórmulas para el tamaño de la muestra, el tiempo transcurrido desde el ingreso al estudio hasta la observación del evento se asume independiente y exponencialmente distribuido. De igual manera, la pérdida por seguimiento o censura se asume también independiente y exponencialmente distribuida, Lakatos (1992). La censura puede ser a la derecha, la más común de las censuras, o censura a la izquierda, Kleinbaum (1995).

### 3.2.1 Tamaño de muestra para dos grupos independientes sin censura

La ecuación (4), reportada por Lachin (1981), se utiliza para calcular el tamaño de la muestra cuando se desea comparar las tasas de riesgo entre dos grupos (tratamiento y control) para un nivel de potencia  $1 - \beta$  y un nivel de confianza  $1 - \alpha$ . Esta fórmula se construyó bajo el supuesto que las unidades del estudio serán seguidas todo el tiempo requerido hasta la ocurrencia del evento de interés, es decir, no se presentará censura en el estudio. Despejando de la ecuación (4), se puede obtener a la ecuación (5) donde se presenta la fórmula para calcular la potencia de la prueba, con un tamaño de la muestra determinado.

$$n = \left\{ \frac{Z_{1-\alpha/2} \sqrt{g(\bar{\lambda}) (Q_t^{-1} + Q_c^{-1})} + Z_{1-\beta} \sqrt{g(\lambda_t) Q_t^{-1} + g(\lambda_c) Q_c^{-1}}}{|\lambda_t - \lambda_c|} \right\}^2 \quad (4)$$

donde

$$g(\lambda_t) = \lambda_t^2, g(\lambda_c) = \lambda_c^2 \quad \text{ó} \quad g(\bar{\lambda}) = \bar{\lambda}^2$$

$$1 - \beta = \Phi(Z_{1-\beta}) \quad (5)$$

donde

$$Z_{1-\beta} = \frac{\sqrt{n} |\lambda_t - \lambda_c| - Z_{1-\alpha/2} \sqrt{g(\bar{\lambda}) (Q_t^{-1} + Q_c^{-1})}}{\sqrt{g(\lambda_t) Q_t^{-1} + g(\lambda_c) Q_c^{-1}}}$$

La ecuación (4) raramente se puede aplicar porque es difícil seguir a la totalidad de las unidades del estudio hasta el evento de interés.

### 3.2.2 Tamaño de la muestra para dos grupos independientes con censura

Un acercamiento más realista que el presentado en la ecuación (4) es el presentado en la ecuación (6), Lachin (1981).

$$n = \left\{ \frac{Z_{1-\alpha/2} \sqrt{g(\bar{\lambda}) (Q_t^{-1} + Q_c^{-1})} + Z_{1-\beta} \sqrt{g(\lambda_t) Q_t^{-1} + g(\lambda_c) Q_c^{-1}}}{|\lambda_t - \lambda_c|} \right\}^2 \quad (6)$$

donde  $g(\lambda) = \frac{\lambda^2}{(1+\exp(-\lambda T))}$ , para  $\lambda = \lambda_t, \lambda_c$  ó  $\bar{\lambda}$ .

Esta fórmula se utiliza para calcular el tamaño de la muestra cuando se desean comparar las tasas de riesgo para dos grupos independientes bajo el supuesto que todas las unidades del estudio se siguen desde su inicio hasta un tiempo fijo  $T$ , para un nivel de potencia  $1 - \beta$  y un nivel de confianza  $1 - \alpha$ . El seguimiento puede empezar teniendo todas las unidades del estudio al mismo instante de inicio o cada unidad entra en el estudio a diferentes tiempos en forma aleatoria. Es decir, se espera seguir todas las unidades del estudio por un tiempo  $T$  fijo, independiente del momento de entrada al estudio. Debido a que las unidades pueden estar en diferentes momentos en el tiempo, la duración del estudio se convierte en una variable aleatoria, por que se desconoce el momento en que se ingresa la última unidad al estudio (ver figura 1).

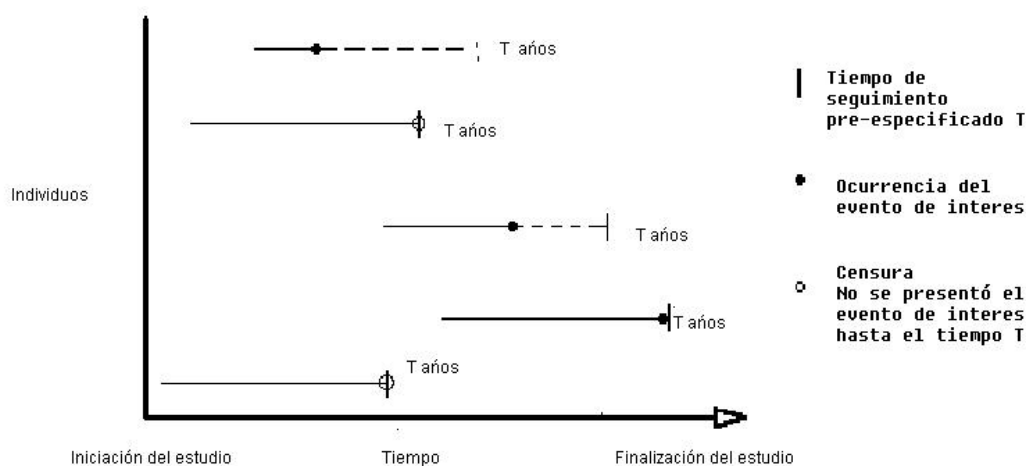


Figura 1: Todos los individuos son seguidos exactamente por  $T$  años después de entrar al estudio.

La ecuación (7) se utiliza para determinar la potencia del estudio cuando el tamaño de la muestra se calcula usando la ecuación (6).

$$1 - \beta = \Phi(Z_{1-\beta}) \quad (7)$$

donde

$$Z_{1-\beta} = \frac{\sqrt{n}|\lambda_t - \lambda_c| - Z_{1-\alpha/2}\sqrt{g(\bar{\lambda})(Q_t^{-1} + Q_c^{-1})}}{\sqrt{g(\lambda_t)Q_t^{-1} + g(\lambda_c)Q_c^{-1}}}$$

### 3.2.3 Tamaño de la muestra para dos grupos independientes con censura progresiva

Otra situación se presenta cuando las unidades del estudio inician en algún momento entre el tiempo  $(0, T]$  fijo. El tiempo  $T$  es fijado por los investigadores quienes establecen el momento exacto del inicio del estudio y el momento exacto de finalización del mismo.

Las unidades censuradas en el tiempo  $(0, T]$  se conocen como censura progresiva porque no todas las unidades entraron ni se censuraron al mismo tiempo. Implicando que las unidades entran en forma progresiva y no todos los individuos son seguidos por el mismo tiempo. Por lo tanto, la censura ocurre en forma aleatoria, pero se puede pensar que la probabilidad de censura es mayor en las unidades quienes entran en un momento cercano al tiempo  $T$  en comparación con las unidades quienes entraron en un momento cercano al tiempo 0 (ver figura 2). La ecuación (8) se utiliza para calcular el tamaño de la muestra de este escenario, buscando obtener un nivel de potencia  $1 - \beta$  y un nivel de confianza  $1 - \alpha$ . Similarmente para determinar la potencia del estudio se utiliza la ecuación (9), reemplazando el tamaño de la muestra obtenido bajo el uso de la ecuación (8).

$$n = \left\{ \frac{Z_{1-\alpha/2}\sqrt{g(\bar{\lambda})(Q_t^{-1} + Q_c^{-1})} + Z_{1-\beta}\sqrt{g(\lambda_t)Q_t^{-1} + g(\lambda_c)Q_c^{-1}}}{|\lambda_t - \lambda_c|} \right\}^2 \quad (8)$$

donde  $g(\lambda) = \frac{\lambda^3 T}{(\lambda T - 1 + \exp(-\lambda T))}$ , para  $\lambda = \lambda_t, \lambda_c$  ó  $\bar{\lambda}$ .

$$1 - \beta = \Phi(Z_{1-\beta}) \quad (9)$$

donde

$$Z_{1-\beta} = \frac{\sqrt{n}|\lambda_t - \lambda_c| - Z_{1-\alpha/2}\sqrt{g(\bar{\lambda})(Q_t^{-1} + Q_c^{-1})}}{\sqrt{g(\lambda_t)Q_t^{-1} + g(\lambda_c)Q_c^{-1}}}$$



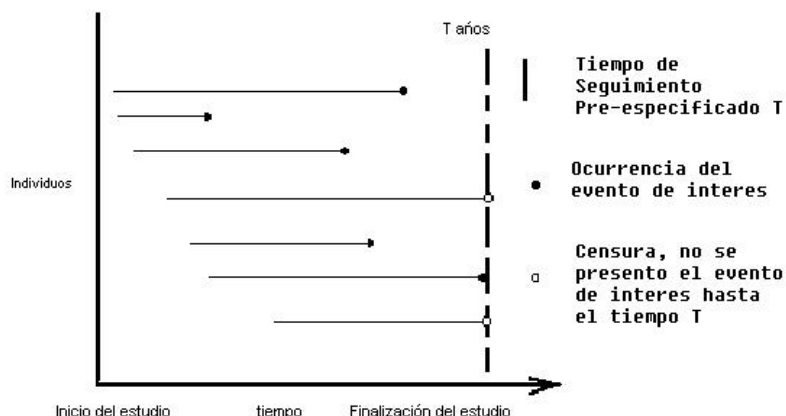


Figura 2: Los individuos entran al estudio de forma progresiva entre el tiempo  $(0, T)$  y son seguidos hasta el momento planeado para el fin del seguimiento

### 3.2.4 Tamaño de la muestra para dos grupos independientes con reclutamiento limitado y censura

Usualmente es deseable reclutar individuos para el estudio en un determinado intervalo de tiempo definido  $(0, T_0)$ , antes de iniciar el período de seguimiento. A partir de  $T_0$  hasta el tiempo  $T$  se realiza el seguimiento de los individuos reclutados, Lachin (1981). Esto implica que la censura ocurre entre el tiempo  $T_0$  y  $T$  (ver figura 3). Para calcular el tamaño de la muestra se requiere otra fórmula que involucre estos tiempos de reclutamiento y seguimiento; se presenta en la ecuación (10), con el supuesto  $T > T_0$ . Esta fórmula permite determinar el mínimo tamaño de la muestra requerido para obtener un nivel de potencia  $1 - \beta$  y un nivel de confianza  $1 - \alpha$  en el estudio. Para determinar la potencia del estudio se utiliza la ecuación (11), reemplazando el tamaño de la muestra obtenido bajo el uso de la ecuación (10).

$$n = \left\{ \frac{Z_{1-\alpha/2} \sqrt{g(\bar{\lambda}) (Q_t^{-1} + Q_c^{-1})} + Z_{1-\beta} \sqrt{g(\lambda_t) Q_t^{-1} + g(\lambda_c) Q_c^{-1}}}{|\lambda_t - \lambda_c|} \right\}^2 \quad (10)$$

donde  $g(\lambda) = \frac{\lambda^3 T}{\lambda T_0 - [\exp(-\lambda(T-T_0)) - \exp(-\lambda T)]}$ , para  $\lambda = \lambda_t, \lambda_c$  ó  $\bar{\lambda}$ .

$$1 - \beta = \Phi(Z_{1-\beta}) \tag{11}$$

donde

$$Z_{1-\beta} = \frac{\sqrt{n} |\lambda_t - \lambda_c| - Z_{1-\alpha/2} \sqrt{g(\bar{\lambda}) (Q_t^{-1} + Q_c^{-1})}}{\sqrt{g(\lambda_t) Q_t^{-1} + g(\lambda_c) Q_c^{-1}}}$$

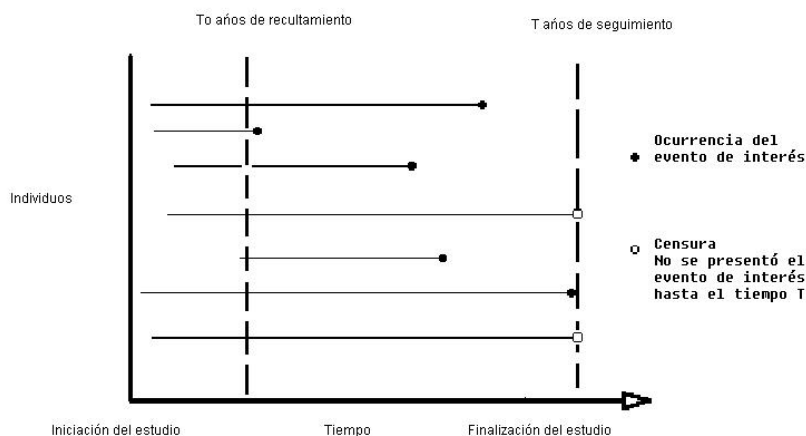


Figura 3: A partir de  $T_0$  se realiza el seguimiento de todos los individuos reclutados hasta el tiempo  $T$ , establecido para el seguimiento en el estudio, donde  $T > T_0$

### 3.3 Tamaño de la muestra para evaluar la hipótesis de tasas de riesgo proporcionales

La fórmula descrita en esta subsección se basa en el mínimo número de eventos o desenlaces requeridos cuando se desea comparar las tasas de riesgo entre dos grupos, por ejemplo un grupo tratamiento y un grupo control, usando el modelo de riesgos proporcionales de Cox y específicamente mediante la prueba log-rank, Cox (1982).

La prueba log-rank permite probar la hipótesis nula que la tasa de riesgo de uno de los dos grupos, por ejemplo el grupo control, es igual a  $k$ -veces la tasa de riesgo en el otro grupo, por ejemplo el grupo tratamiento,  $H_0 : \lambda_t = k\lambda_c$ . Usualmente la hipótesis alterna es que la tasa de riesgo en el grupo tratamiento es distinta de  $k$ -veces la tasa de riesgo en el grupo control,  $H_1 : \lambda_t \neq k\lambda_c$ . Equivalentemente, estas hipótesis se pueden expresar en términos de la razón de las tasas y por esto reciben el nombre de tasas de riesgos proporcionales,  $H_0 : \frac{\lambda_t}{\lambda_c} = k$  vs  $H_1 : \frac{\lambda_t}{\lambda_c} \neq k$ .

La ecuación (12) reportada por Yatemán y Skene (1992), permite calcular el mínimo número de eventos o desenlaces ( $m$ ) que deben ser observados durante el estudio para probar la hipótesis mencionada para un nivel de potencia  $1 - \beta$  y un nivel de confianza  $1 - \alpha$ .  $\lambda$  representa la razón entre las tasas de riesgo de ambos grupos, por ejemplo la del grupo tratamiento al grupo control,  $\lambda = \lambda_t/\lambda_c$ . Posteriormente, el tamaño de la muestra requerido ( $n$ ) se debe determinar ajustando, el número de eventos calculados en la ecuación (12), por la probabilidad ( $u$ ) del desenlace de interés en la población. En otras palabras  $n = m/u$ . Despejando de la ecuación (12), se puede obtener la ecuación (13) donde se presenta la fórmula para calcular la potencia de la prueba basado en el número de desenlaces de interés proveniente de la ecuación (12).

$$m = \frac{(Z_{1-\alpha/2} + Z_{1-\beta})^2}{(Q_t Q_c L N^2 \lambda)} \quad (12)$$

$$1 - \beta = \Phi(Z_{1-\beta}) \quad (13)$$

donde

$$Z_{1-\beta} = \sqrt{m(Q_t Q_c L N^2 \lambda)} - Z_{1-\alpha/2}$$

Se debe tener claro que en estudios donde se plantea este tipo diseño el tiempo de la duración del estudio es una variable aleatoria. Esto implica que se debe contar con los recursos suficientes para desarrollar el estudio hasta la ocurrencia del evento de interés en la totalidad de los pacientes requeridos y estimados en la formula enunciada.

## 4 Ejemplos

### 4.1 Tamaño de la muestra para detectar una tasa de riesgo específica

Suponga que se desea estimar la tasa de riesgo de muerte de pacientes quienes son hospitalizados en una clínica en Bogotá; se desea estimar la tasa de riesgo del evento muerte durante el primer mes después del diagnóstico. Estudios previos en otras clínicas similares de la ciudad indican que la tasa de riesgo de interés oscila entre 0.3 y 0.35. Se conoce que la probabilidad de muerte de esa clínica es 0.02. Los investigadores están dispuestos a cometer niveles de error tipo I y II de 0.05 y 0.20 respectivamente. Reemplazando los datos en la ecuación (2) se puede determinar que se necesitan 28 muertes y 34 muertes para las hipótesis de que la tasa de riesgo es de 0.30 y 0.35 respectivamente. De igual manera, mediante el uso de la tabla 1, se pueden encontrar estos valores requeridos.

Esto implica que para obtener este número de desenlaces se deben diagnosticar y seguir en el tiempo un total de 1400 y 1700 individuos respectivamente.

$\lambda_0$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55
m	10	12	15	18	22	28	34	43	55	71	94

$\lambda_0$	0.60	0.65	0.70	0.75	0.8	0.85	0.9	0.95	1.0	2	3
m	126	175	253	385	636	1194	2834	11939	.	71	32

$\lambda_0$	4	5	6	7	8	9	10
m	22	18	16	14	13	13	12

Tabla1. Mínimo número de eventos de interés o desenlaces requeridos ( $m$ ) para detectar una tasa de riesgo específica ( $\lambda_0$ ), con un nivel de error tipo I de 0.05 y un nivel de error tipo II de 0.20.

### 4.2 Tamaño de la muestra para dos grupos independientes sin censura

Mediante un ensayo clínico se está interesado en comparar las tasas de riesgo entre un tratamiento nuevo y el tratamiento convencional para un tipo de cáncer. Además se cuenta con todo el tiempo necesario para el seguimiento de los pacientes. Por medio de una revisión de la literatura y una muestra piloto se estimaron las tasas de riesgo de los grupos que recibieron el nuevo tratamiento y el tratamiento convencional, como 0.25 y 0.3, respectivamente. Se

desea determinar el tamaño de la muestra necesario para probar la hipótesis de igualdad de las tasas de riesgo, a dos colas, con un nivel de confianza del 95 % y una potencia del 80 %. Se espera contar con igual proporción de individuos en los dos grupos. Los investigadores confían en que no se presente censura en el estudio.

Reemplazando los datos en la ecuación (4) se puede determinar que se necesitan 953 individuos, es decir aproximadamente 477 individuos en cada grupo. De igual manera, mediante el uso de la tabla 2, es posible encontrar el tamaño de la muestra requerido en la intersección de la fila de la tasa de riesgo del tratamiento convencional ( $\lambda_c$ ) de 0.3 y la columna de la tasa de riesgo del nuevo tratamiento ( $\lambda_t$ ) de 0.25, encontrándose el valor de 953 individuos.

		$\lambda_t$																		
$\lambda_c$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1.00
0.05	..	509	175	100	70	54	44	38	33	30	27	25	23	22	21	20	19	18	17	17
0.10	509	..	889	274	147	97	72	57	48	41	36	33	30	27	26	24	23	22	21	20
0.15	175	889	..	1304	383	197	127	92	72	59	50	43	39	35	32	29	28	26	24	23
0.20	100	274	1304	..	1756	501	251	159	113	87	70	59	51	45	40	37	34	31	29	28
0.25	70	147	383	1756	..	2247	628	310	193	136	103	83	69	59	52	46	42	38	35	33
0.30	54	97	197	501	2247	..	2777	766	374	230	161	121	96	80	68	59	52	47	43	40
0.35	44	72	127	251	628	2777	..	3349	914	442	270	187	140	111	91	77	67	59	53	48
0.40	38	57	92	159	310	766	3349	..	3964	1074	516	313	215	160	126	103	87	75	66	59
0.45	33	48	72	113	193	374	914	3964	..	4623	1244	594	358	245	182	143	116	98	84	74
0.50	30	41	59	87	136	230	442	1074	4623	..	5327	1426	677	407	277	205	160	130	109	93
0.55	27	36	50	70	103	161	270	516	1244	5327	..	6078	1620	766	458	311	229	178	145	121
0.60	25	33	43	59	83	121	187	313	594	1426	6078	..	6878	1826	861	513	347	255	198	160
0.65	23	30	39	51	69	96	140	215	358	677	1620	6878	..	7727	2044	961	571	385	282	218
0.70	22	27	35	45	59	80	111	160	245	407	766	1826	7727	..	8626	2276	1066	632	426	311
0.75	21	26	32	40	52	68	91	126	182	277	458	861	2044	8626	..	9578	2520	1178	696	468
0.80	20	24	29	37	46	59	77	103	143	205	311	513	961	2276	9578	..	10582	2778	1296	764
0.85	19	23	28	34	42	52	67	87	116	160	229	347	571	1066	2520	10582	..	11640	3049	1419
0.90	18	22	26	31	38	47	59	75	98	130	178	255	385	632	1178	2778	11640	..	12752	3334
0.95	17	21	24	29	35	43	53	66	84	109	145	198	282	426	696	1296	3049	12752	..	13921
1.00	17	20	23	28	33	40	48	59	74	93	121	160	218	311	468	764	1419	3334	13921	..

Tabla2. Mínimo tamaño de muestra requerido para observar diferencias entre dos grupos independientes sin censura, a un nivel de error tipo I de 0.05 y un nivel de error tipo II de 0.2, para pruebas a dos colas cuando  $Q_n = Q_c = 0,5$ , para diferentes escenarios de tasas de riesgo en ambos grupos ( $\lambda_c$ : tasa de riesgo de interés en uno de los grupos y  $\lambda_t$ : tasa de riesgo de interés en el otro grupo).

### 4.3 Tamaño de la muestra para dos grupos independientes con censura

Retomando el ejemplo anterior, los investigadores después de consultar al estadístico, optan por una aproximación más realista aceptando la posible presencia de censura en el estudio. Adicionalmente, los investigadores limitaron el tiempo de seguimiento de cada individuo a dos años desde el momento del ingreso del individuo al estudio. El tamaño de la muestra requerido reemplazando estos datos en la ecuación (6) genera un total de 2247 individuos, es decir 1124 individuos por grupo. De igual manera, mediante el uso de la tabla 3, es posible encontrar el tamaño de la muestra requerido en la intersección de la fila de la tasa de riesgo del tratamiento convencional ( $\lambda_c$ ) de 0.3 y la columna de la tasa de riesgo del nuevo tratamiento ( $\lambda_t$ ) de 0.25, encontrándose el valor de 2247 individuos.

	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50	0,55	0,60	0,65	0,70	0,75	0,80	0,85	0,90	0,95	1,00
0,05	.	73	34	25	20	18	17	16	15	14	14	14	13	13	13	13	13	12	12	12
0,10	73	.	199	73	46	34	28	25	22	20	19	18	17	17	16	16	15	15	15	14
0,15	34	199	.	387	128	73	52	41	34	30	27	25	23	21	20	19	19	18	18	17
0,20	25	73	387	.	639	199	108	73	56	46	39	34	31	28	26	25	23	22	21	20
0,25	20	46	128	639	.	953	285	150	99	73	59	49	43	38	34	31	29	27	26	25
0,30	18	34	73	199	953	.	1329	387	199	128	94	73	61	52	46	41	37	34	32	30
0,35	17	28	52	108	285	1329	.	1769	505	255	162	116	90	73	62	54	48	43	40	37
0,40	16	25	41	73	150	387	1769	.	2271	639	318	199	141	108	88	73	63	56	50	46
0,45	15	22	34	56	99	199	505	2271	.	2836	788	387	240	169	128	103	86	73	64	57
0,50	14	20	30	46	73	128	255	639	2836	.	3464	953	464	285	199	150	120	99	84	73
0,55	14	19	27	39	59	94	162	318	788	3464	.	4155	1133	548	334	232	174	138	113	96
0,60	14	18	25	34	49	73	116	199	387	953	4155	.	4908	1329	639	387	267	199	157	128
0,65	13	17	23	31	43	61	90	141	240	464	1133	4908	.	5725	1541	736	444	305	225	177
0,70	13	17	21	28	38	52	73	108	169	285	548	1329	5725	.	6604	1769	841	505	345	255
0,75	13	16	20	26	34	46	62	88	128	199	334	639	1541	6604	.	7546	2012	953	570	387
0,80	13	16	19	25	31	41	54	73	103	150	232	387	736	1769	7546	.	8550	2271	1071	639
0,85	13	15	19	23	29	37	48	63	86	120	174	267	444	841	2012	8550	.	9618	2546	1197
0,90	12	15	18	22	27	34	43	56	73	99	138	199	305	505	953	2271	9618	.	10748	2836
0,95	12	15	18	21	26	32	40	50	64	84	113	157	225	345	570	1071	2546	10748	.	11941
1,00	12	14	17	20	25	30	37	46	57	73	96	128	177	255	387	639	1197	2836	11941	.

Tabla3. Mínimo tamaño de la muestra requerido para observar diferencias entre dos grupos independientes con censura, a un nivel de error tipo I de 0.05 y un nivel de error tipo II de 0.2, para pruebas a dos colas con tiempo de seguimiento  $T = 2$  y  $Q_n = Q_c = 0,5$ . Para diferentes escenarios de tasas de riesgo en ambos grupos ( $\lambda_c$ : tasa de riesgo de interés en uno de los grupos y  $\lambda_t$ : tasa de riesgo de interés en el otro grupo)

#### 4.4 Tamaño de la muestra para dos grupos independientes con censura progresiva

Siguiendo con el ejemplo anterior, los investigadores necesitaron limitar el estudio para obtener los recursos financieros para realizarlo y deciden seguir a cada individuo durante cinco años desde el momento de financiamiento del estudio. El tamaño de la muestra requerido reemplazando estos datos en la ecuación (8) genera un total de 2082 individuos, es decir 1041 individuos por grupo. De igual manera, mediante el uso de la tabla 4, es posible encontrar el tamaño de la muestra requerido en la intersección de la fila de la tasa de riesgo del tratamiento convencional ( $\lambda_c$ ) de 0.3 y la columna de la tasa de riesgo del nuevo tratamiento ( $\lambda_t$ ) de 0.25, encontrándose el valor de 2082 individuos.

	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50	0,55	0,60	0,65	0,70	0,75	0,80	0,85	0,90	0,95	1,00
0,05	.	427	190	87	62	48	40	35	31	28	26	24	23	22	21	20	19	18	18	17
0,10	427	.	768	241	130	88	66	53	45	39	35	32	29	27	25	24	23	22	21	20
0,15	190	768	.	1155	344	179	117	86	68	56	48	42	38	35	32	30	28	26	25	24
0,20	87	241	1155	.	1593	460	234	149	107	84	68	58	50	45	40	37	34	32	30	28
0,25	62	130	344	1593	.	2082	589	294	185	131	101	82	69	59	52	47	43	39	36	34
0,30	48	88	179	460	2082	.	2625	131	360	224	158	120	96	80	69	60	54	49	45	41
0,35	40	66	117	234	589	2625	.	3223	888	433	266	186	140	112	93	79	69	61	55	50
0,40	35	53	86	149	294	731	3223	.	3876	1058	512	313	217	163	129	106	90	78	69	62
0,45	31	45	68	107	185	360	888	3876	.	4587	1243	597	363	250	187	147	120	102	88	77
0,50	28	39	56	84	131	224	433	1058	4587	.	5356	1443	689	416	285	212	166	136	114	98
0,55	26	35	48	68	101	158	266	512	1243	5356	.	6183	1657	888	474	323	239	187	152	127
0,60	24	32	42	58	82	120	186	313	597	1443	6183	.	7071	1886	893	535	364	266	209	169
0,65	23	29	38	50	69	96	140	217	363	689	1657	7071	.	8018	2131	1005	600	407	299	232
0,70	22	27	35	45	59	80	112	163	250	416	788	1886	8018	.	9027	2391	1124	669	452	331
0,75	21	25	32	40	52	69	93	129	187	285	474	893	2131	9027	.	10096	2666	1250	741	500
0,80	20	24	30	37	47	60	79	105	147	212	323	535	1005	2391	10096	.	11227	2966	1383	818
0,85	19	23	28	34	43	54	69	90	120	166	239	364	600	1124	2666	11227	.	12420	3252	1522
0,90	18	22	26	32	39	49	61	78	102	136	187	268	407	669	1250	2966	12420	.	13675	3584
0,95	18	21	25	30	36	45	55	69	88	114	152	209	299	452	741	1383	3252	13675	.	14992
1,00	17	20	24	28	34	41	50	62	77	98	127	169	232	331	500	818	1522	3584	14992	.

Tabla4. Mínimo tamaño de la muestra requerido para observar diferencias entre dos grupos independientes con censura progresiva, a un nivel de error tipo I de 0.05 y un nivel de error tipo II de 0.2, para pruebas a dos colas con tiempo de duración del estudio  $T = 5$  y  $Q_n = Q_c = 0,5$ . Para diferentes escenarios de tasas de riesgo en ambos grupos ( $\lambda_c$ : tasa de riesgo de interés en uno de los grupos y  $\lambda_t$ : tasa de riesgo de interés en el otro grupo).

### 4.5 Tamaño de la muestra para dos grupos independientes con reclutamiento limitado y censura

El consultor estadístico, les recomienda a los investigadores del ejemplo anterior que consideren la opción de tener un período de reclutamiento de individuos, antes de iniciar la aleatorización al tratamiento y esto les permitiría evaluar rápidamente la factibilidad del estudio. Los investigadores aceptan la sugerencia y deciden reclutar los individuos durante un año y al segundo año iniciar la aleatorización al tratamiento. El mejor escenario es que los individuos reciban el tratamiento durante un máximo de dos años. Es decir, los investigadores reportan a su fuente financiadora un tiempo máximo del estudio de 3 años.

	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50	0,55	0,60	0,65	0,70	0,75	0,80	0,85	0,90	0,95	1,00
0,05	,	667	228	130	90	69	56	47	41	37	33	31	28	26	25	24	22	21	21	20
0,10	667	,	1153	354	188	124	91	72	60	51	45	40	36	33	31	29	27	26	24	23
0,15	228	1153	,	1675	489	290	180	115	89	73	61	53	47	42	38	35	33	31	29	27
0,20	130	354	1675	,	2233	633	316	198	140	107	86	72	62	54	48	44	40	37	34	32
0,25	90	188	489	2233	,	2829	787	387	239	168	127	101	84	71	62	55	50	45	41	38
0,30	69	124	290	633	2829	,	3464	951	462	283	196	147	117	96	81	70	62	58	50	46
0,35	56	91	180	316	787	3464	,	4139	1125	541	329	227	169	133	109	92	79	70	62	56
0,40	47	72	115	198	387	951	4139	,	4855	1308	625	378	259	182	150	122	103	88	77	69
0,45	41	60	88	140	239	462	1125	4855	,	5613	1504	715	429	283	216	169	137	114	98	86
0,50	37	51	73	107	168	283	541	1309	5613	,	6415	1710	809	483	325	241	188	152	127	108
0,55	33	45	61	86	127	196	329	625	1504	6415	,	7260	1927	908	540	366	268	208	168	139
0,60	31	40	53	72	101	147	227	378	715	1710	7260	,	8151	2155	1012	601	405	296	229	184
0,65	28	36	47	62	84	117	169	259	429	809	1927	8151	,	9088	2395	11221	664	446	326	251
0,70	26	33	42	54	71	96	133	192	293	483	908	2155	9088	,	10072	2647	1236	730	490	356
0,75	25	31	38	48	62	81	109	150	216	328	540	1012	2395	10072	,	11104	2911	1356	799	535
0,80	24	29	35	44	55	70	92	122	169	241	366	601	1121	2647	11104	,	12185	3188	1482	871
0,85	22	27	33	40	50	62	79	103	137	188	268	405	664	1236	2111	12185	,	13316	3477	1613
0,90	21	26	31	37	45	58	70	88	114	152	208	296	446	730	1358	3188	13316	,	14497	3779
0,95	21	24	29	34	41	50	62	77	98	127	188	229	328	490	799	1482	3477	14497	,	15730
1,00	20	23	27	32	38	45	56	69	86	108	139	184	251	356	535	871	1613	3779	15730	,

Tabla5. Mínimo tamaños de la muestra requerido para observar diferencias entre dos grupos independientes con reclutamiento limitado y censura, a un nivel de error tipo I de 0.05 y un nivel de error tipo II de 0.2, para pruebas a dos colas con tiempo de reclutamiento  $T_0 = 1$ , tiempo de seguimiento  $T = 2$ ,  $Q_n = Q_c = 0,5$  y para diferentes escenarios de tasas de riesgo en ambos grupos ( $\lambda_c$ : tasa de riesgo de interés en uno de los grupos y  $\lambda_t$ : tasa de riesgo de interés en el otro grupo).



El tamaño de la muestra requerido reemplazando estos datos en la ecuación (10) genera un total de 2829 individuos, es decir 1415 individuos por grupo. De igual manera, mediante el uso de la tabla 5, es posible encontrar el tamaño de la muestra requerido en la intersección de la fila de la tasa de riesgo del tratamiento convencional ( $\lambda_c$ ) de 0.3 y la columna de la tasa de riesgo del nuevo tratamiento ( $\lambda_t$ ) de 0.25, encontrándose el valor de 2829 individuos.

#### 4.6 Tamaño de la muestra para evaluar la hipótesis de tasas de riesgo proporcionales

Se tiene un estudio donde el evento de interés es el daño de una pieza mecánica de un motor. Se cuenta con una nueva pieza mecánica cuya producción es económica y se quiere comparar con una pieza mecánica convencional de producción más costosa; además se tiene el supuesto que las piezas mecánicas tienen tasas de riesgo proporcionales. La tasa de riesgo estimada con la pieza mecánica convencional es de 0.32 y la tasa de riesgo estimada con la nueva pieza mecánica es de 0.4. Se desea estimar el mínimo número de piezas en las que se debe observar el evento de interés (daño) para comparar las tasas de riesgo. Los investigadores con anterioridad definieron que del total de las piezas seguidas el 40% deben corresponder a las piezas mecánicas nuevas y el restante 60% deben ser las piezas convencionales.

Reemplazando estos datos en la ecuación (10) se obtiene un total de 657 piezas, es decir 263 piezas mecánicas nuevas y 395 piezas mecánicas convencionales. De igual manera, mediante el uso de la tabla 6, es posible encontrar el número de piezas en las que se debe observar el evento de interés, correspondiente al lugar donde  $\lambda = \lambda_t/\lambda_c=0.80$ .

Tabla6. Mínimo número de eventos o desenlaces ( $m$ ) que deben observarse para evaluar la hipótesis de tasas de riesgo proporcionales en dos grupos, a un nivel de error tipo I de 0.05 y un nivel de error tipo II de 0.2, para pruebas a dos colas y proporciones en los grupos de  $Q_t=0.60$  y  $Q_c=0.40$ .

$\lambda_0$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55
m	4	7	10	13	17	22	30	39	52	69	92

$\lambda_0$	0.60	0.65	0.70	0.75	0.8	0.85	0.9	0.95	1.0	2	3
m	126	177	258	396	657	1239	2946	12431	.	69	28

$\lambda_0$	4	5	6	7	8	9	10
m	18	13	11	9	8	7	7

## 5 Discusión

Consideramos que el presente artículo aporta una amplia revisión sobre el cálculo del tamaño de la muestra cuando se planea un análisis de sobrevida. Por tanto, constituye un valioso documento tanto para el profesional en estadística como para todos aquellos profesionales involucrados en estudios de esta área de trabajo. Las personas con conocimientos matemáticos seguramente encontrarán en las fórmulas presentadas una herramienta de trabajo para resolver la pregunta de “cuántos sujetos son requeridos para mi estudio?”. Por otra parte, las personas que no deseen emplear las ecuaciones reportadas pueden responder sus preguntas de tamaño de la muestra haciendo uso de las tablas mostradas. Adicionalmente, los ejemplos presentados para cada uno de las fórmulas reportadas serán de gran utilidad en el uso de las mismas.

Los cálculos anteriormente presentados en las tablas 2, 3, 4 y 5 evidencian que los mayores tamaños de la muestra son requeridos en los escenarios donde se presenta algún tipo de censura. Igualmente, la disminución en la magnitud de la diferencia entre las tasas de riesgo contrastadas en la hipótesis de estudio, implica un incremento en el mínimo tamaño de la muestra requerido. De otra parte, a medida que los valores de  $\lambda_c$  y  $\lambda_t$  se acercan simultáneamente al valor de 1, el tamaño de la muestra aumenta rápidamente.

Las fórmulas del tamaño de la muestra reportadas en las ecuaciones (4), (6), (8) y (10) son simétricas respecto a los valores de las tasas de riesgo cuando las proporciones de asignación en ambos grupos son iguales ( $Q_c = 0,5$  y  $Q_n = 0,5$ ). Esto permite que el cálculo del tamaño de la muestra cuando  $\lambda_c = a$  y  $\lambda_t = b$  sea el mismo que para el caso donde  $\lambda_c = b$  y  $\lambda_t = a$ . Lo anterior permite usar las tablas con los tamaños de la muestra presentados en la parte superior de la diagonal, siempre y cuando  $Q_c = 0,5$  y  $Q_n = 0,5$ .

Cuando el tiempo de seguimiento en el estudio está fijo, es decir, cuando se limita el seguimiento a un número determinado de años por decir 1, 2, 3, 4, ó 5, el número de individuos necesarios en el estudio se reduce a medida que el tiempo de seguimiento aumenta.

*Agradecimientos: El autor principal desea expresar su agradecimiento a Colciencias y a la Pontificia Universidad Javeriana por apoyarlo mediante el programa de Jóvenes Investigadores. Esta investigación fue apoyada por la FPIT (120-104R-0101-106), Colciencias (120-104R-0101-103) y la PUJ (120-104R0101-200).*

## Referencias

- [1] Cox, D. R. y Oakes, D. (1984) Analysis of survival data. Chapman & Hall, New York
- [2] Elandt-Johnson R.C. y Johnson N.L. (1980) Survival models and data analysis. John Wiley, New York.
- [3] Hosmer, D. W. y Lemeshow, S. (1999) Applied survival analysis. John Wiley & Sons, New York
- [4] Kleinbaum, D.G. (1995) Survival Analysis. Springer-Verlag, New York.
- [5] Lachin, J.M. (1981) Introduction to sample size determination and power analysis for clinical trials. Controlled Clinical Trials. Vol 2, 93-113.
- [6] Lakatos, E. (1992) A comparison of sample size methods for the logrank statistic. Statistics in Medicine Vol 11, 179-191.
- [7] Law, M. G. (1996) Sample size calculation for within-patient comparisons with a binary or survival endpoint. Controlled Clinical Trials Vol. 17, 221-225.
- [8] Lee, E.T. (1992) Statistical methods for survival data analysis. John Wiley, New York.
- [9] Perez, A., Gil, J.F.A., y Castañeda, J. (2001). Informe final del proyecto "Modelamiento y simulación del tamaño de la muestra en algunos diseños de investigación en salud". Pontificia Universidad Javeriana, Biblioteca Unidad de Epidemiología Clínica y Bioestadística. Bogotá, Colombia.
- [10] Taulbee, J, y Symons, M. (1983). Sample size and duration for cohort studies of survival time with covariables. Biometrics. Vol 39, 351-360.
- [11] Yateman, N. A. y Skene, A. M. (1992) Sample size for proportional hazards survival studies with arbitrary patient entry and loss to follow up distributions. Statistics in Medicine. Vol. 11, 1103-1113.