Revista Colombiana de Estadística
Junio 2011, volumen 34, no. 1, pp. 197 a 209

Comparación de intervalos de confianza para la función de supervivencia con censura a derecha

Comparison of Confidence Intervals for the Survival Function in the Presence of Right Censoring

Javier Ramírez[a]

Departamento de Matemáticas y Estadística, Facultad de Ciencias Básicas, Universidad de Córdoba, Montería, Colombia

Resumen

En este trabajo se comparan métodos para encontrar intervalos de confianza para la función de supervivencia, como los métodos de remuestreo *Bootstrap* aplicado a los estimadores de Kaplan-Meier y Nelson-Aalen. También, mediante las transformaciones log, $\log(-\log)$ y arc sen que pueden resultar en muchos casos más efectivos. Además, se muestra el comportamiento que presentan los intervalos de confianza no paramétricos frente a los paramétricos.

***Palabras clave*:** bootstrap, censura a derecha, estimador Kaplan-Meier, función de supervivencia, intervalos de confianza.

Abstract

This work compares methods to find confidence interval for the survival function such as the resampling methods *Bootstrap*, applied to the Kaplan-Meier and Nelson-Aalen estimators. Also through log, $\log(-\log)$ and arcsin transformations that can result more effectives in many cases. The behavior of nonparametric confidence intervals against parametric ones is also shown.

***Key words*:** Bootstrap, Confidence intervals, Kaplan-Meier estimator, Survival function, Right censoring.

1. Introducción

En la actualidad existen diferentes métodos para encontrar intervalos de confianza para la función de supervivencia en un tiempo de interés, con censura a

[a]Profesor asistente. E-mail: javierramirez@sinu.unicordoba.edu.co

derecha, tales como los métodos tradicionales utilizando los estimadores de Kaplan-Meier y Nelson-Aalen, también mediante remuestreo *Bootstrap* aplicado a estos estimadores y a través de las transformaciones log, $\log(-\log)$ y $\arcsin$, que pueden resultar en muchos casos más efectivos. Así, este trabajo compara de manera simultánea todos estos métodos utilizando diferentes porcentajes de censura y tamaños de muestra, a través de diferentes modelos generadores. Además, proporciona criterios para establecer qué intervalos de confianza no paramétricos utilizar para estimar la función de supervivencia en un tiempo de interés, con censura a derecha y se muestra el comportamiento de los estimadores no paramétricos frente a los paramétricos.

En la sección 2 se muestran los estimadores no paramétricos para la función de supervivencia utilizados en este estudio, como son los estimadores de Kaplan-Meier y Nelson-Aalen. En la sección 3 se presentan los intervalos de confianza utilizados en la comparación como son los tradicionales, *Bootstrap* y mediante transformaciones log, $\log(-\log)$ y $\arcsin$, luego en la sección 4 se presentan los criterios de comparación de los intervalos de confianza y finalmente en la sección 5 se muestra una descripción de los escenarios de simulación, así como sus resultados y conclusiones en la sección 6.

Al final del documento, en los anexos, se encuentran los resultados mediante el índice de comparación de intervalos de confianza propuesto por Correa & Sierra (2003), con base en las distribuciones Weibull y Exponencial para los tiempos de falla/censura.

2. Estimadores utilizados

2.1. Estimador de Kaplan-Meier

Kaplan & Meier (1958) propusieron una modificación de $\widehat{S}(t)$ a la cual denominaron, estimador del producto límite (EPL) de la función de supervivencia. En efecto, supónganse que existen observaciones de n individuos y que hay $k \leq n$ tiempos distintos en los cuales la muerte ocurre, esto es, $t_1 < t_2 < \cdots < t_k$.

Se admite la posibilidad de tener más de una muerte en t_j y d_j representará el número de muertes en t_j. Además, existen los tiempos de censura t_c para individuos cuyo tiempo de vida no es observado.

El estimador del producto límite, $\widehat{S}(t)$, se define como:

$$\widehat{S}(t) = \prod_{j:t_j<t} \frac{n_j - d_j}{n_j} \tag{1}$$

donde n_j es el número de individuos en riesgo en t_j, es decir, el número de individuos vivos y no censurados justo antes de t_j. Cualquier individuo con tiempo de censura registrado igual a t_j será incluido en el conjunto de n_j individuos en riesgo en t_j, como individuos que murieron en t_j. Esta convención es razonable, puesto que un individuo censurado en el tiempo t_c casi ciertamente sobrevive después de t_c.

2.1.1. Varianza del estimador de Kaplan-Meier

Para evaluar los resultados eficazmente cuando se usa el estimador del producto límite, es conveniente tener un estimador de la varianza de $\widehat{S}(t)$, una de las utilizadas en este trabajo es la fórmula de Greenwood (1926), como

$$\widehat{Var}(\widehat{S}(t)) = \widehat{S}^2(t) \sum_{j:t_j<t} \frac{d_j}{n_j[n_j - d_j]} \tag{2}$$

2.2. Estimador de Nelson-Aalen

Nelson (1969), sugieren otra alternativa para estimar la función de supervivencia basándose en la estimación de la función Hazard Acumulada como

$$\widehat{H}(t) = \sum_{j:t_j<t} \frac{d_j}{n_j} \tag{3}$$

este estimador es muy utilizado en los casos cuando se tienen tamaños de muestra pequeños, donde

$$\widehat{S}(t) = \exp(-\widehat{H}(t))$$

2.2.1. Varianza del estimador de Nelson-Aalen

Una forma de calcular la varianza la función de supervivencia, basándose en la función Hazard Acumulada es

$$\widehat{Var}(\widehat{H}(t)) = \sum_{j:t_j<t} \frac{d_j(n_j - d_j)}{n_j^3} \tag{4}$$

otra forma alternativa para varianza es

$$\widehat{Var}(\widehat{H}(t)) = \sum_{j:t_j<t} \frac{d_j}{n_j^2} \tag{5}$$

Aalen & Johansen (1978) proponen una alternativa para la estimación de la varianza de $\widehat{S}(t)$, como

$$\widehat{Var}(\widehat{S}(t)) = \widehat{S}^2(t) \sum_{j:t_j<t} \frac{d_j}{n_j^2} \tag{6}$$

3. Intervalos de confianza

3.1. Intervalo de confianza tradicional Kaplan-Meier

Teniendo en cuenta la normalidad asintótica de los estimadores de máxima verosimilitud, los intervalos de $100(1-\alpha)\,\%$ de confianza de la función de supervivencia en cada tiempo fijo $t_i = 0.2$ pueden calcularse de la siguiente manera

$$\widehat{S}(0.2) \pm Z_{(1-\frac{\alpha}{2})}\widehat{S}(0.2)\sqrt{\sum_{j:t_j<0.2} \frac{d_j}{n_j[n_j - d_j]}} \tag{7}$$

donde $Z_{(1-\alpha/2)}$ es el valor que se excede con probabilidad $(1 - \alpha/2)$ para una distribución normal estándar.

3.2. Intervalo de confianza tradicional Nelson-Aalen

Un intervalo de confianza $100(1 - \alpha)\,\%$ para $\widehat{S}(0.2)$, mediante el estimador de Nelson-Aalen, está dado por

$$\widehat{S}(0.2) \pm Z_{(1-\frac{\alpha}{2})}\sqrt{Var[\widehat{S}(0.2)]} \tag{8}$$

donde, $Var[\widehat{S}(0.2)]$, está dado en la ecuación 6

3.3. Método de remuestreo *Bootstrap* aplicado a datos de supervivencia

Un problema involucra determinar los límites de confianza para la función de supervivencia teórica o parámetros que describen esta función. Akritas (1986), propone utilizar el método *Bootstrap* para estimar la supervivencia utilizando el estimador de Kaplan-Meier.

Para estimar los intervalos de confianza para función de supervivencia utilizando el estimador de Kaplan-Meier a través del intervalo de confianza (7), mediante el remuestreo *Bootstrap*, consiste en lo siguiente:

1. Dada la muestra de tamaño n, estimar $\widehat{S}(0.2)$ mediante (1). La distribución de esta muestra se considera equivalente a la distribución de la población y $\widehat{S}(0.2)$ es el estimador muestral del parámetro poblacional $S(0.2)$.
2. Generar B muestras *Bootstrap* de tamaño n mediante muestreo con reemplazamiento de la muestra original, asignando a cada tiempo una probabilidad $1/n$ y calcular los correspondientes valores $\widehat{S}(0.2)^{*1}, \widehat{S}(0.2)^{*2}, \ldots, \widehat{S}(0.2)^{*B}$ para cada una de las B muestras *Bootstrap*.
3. Estimar el error estándar del parámetro estimado $\widehat{S}(0.2)$ calculando la desviación estándar de las B réplicas *Bootstrap*.

Así, obtenemos que el error estándar es

$$\sigma^*_{S(0.2)} = \sqrt{\frac{\sum_{b=1}^{B}\left(S(0.2)^{b*} - \overline{\widehat{S}}(0.2)^*\right)^2}{(B-1)}} \tag{9}$$

donde $\overline{\widehat{S}}(0.2)^*$ corresponde al promedio de la estimación de la función de supervivencia evaluada en el tiempo $t = 0.2$ de las muestras *Bootstrap*.

3.4. Intervalos de confianza mediante transformaciones

En muchos casos resulta de interés calcular intervalos de confianza mediante transformaciones, como son log, $\log(-\log)$ y arc sen, con el hecho principal ya sea de simetrizar la distribución de un parámetro cualquiera, estabilizar la varianza, etc.

3.4.1. Intervalos de confianza mediante la transformación log

Estos intervalos de confianza fueron sugeridos inicialmente por Kalbfleisch & Prentice (1980); luego un intervalo de confianza $100(1-\alpha)\,\%$, mediante esta transformación, para $S(0.2)$, es

$$\widehat{S}(0.2)\exp\left\{\frac{\pm Z_{(1-\frac{\alpha}{2})}\sigma_s(0.2)}{\widehat{S}(0.2)}\right\} \tag{10}$$

donde

$$\sigma_s^2(0.2) = \frac{\widehat{Var}[\widehat{S}(0.2)]}{\widehat{S}^2(0.2)} \tag{11}$$

es decir que $\sigma_s(0.2)$ corresponde a la raíz de la suma en la formula de Greenwood (1926) (2).

3.4.2. Intervalos de confianza mediante la transformación $\log(-\log)$

Para encontrar intervalos de confianza $100(1-\alpha)\,\%$ mediante esta transformación, se deben encontrar los límites

$$\widehat{S}(0.2)\exp\left\{\pm Z_{(1-\dfrac{\alpha}{2})}\frac{\sigma_s(0.2)}{\log(\widehat{S}(0.2))}\right\} \tag{12}$$

los intervalos de confianza mediante la transformación $\log(-\log)$ son muy utilizados en la práctica, debido a sus propiedades asintóticas.

3.4.3. Intervalos de confianza mediante la transformación arc sen

Otra alternativa para calcular intervalo de confianza $100(1-\alpha)\,\%$ para la función de supervivencia es mediante la transformación arc sen, sugeridos inicialmente por Nair (1984), dado por

$$\operatorname{sen}^2\left\{\text{máx}\left\langle 0, \operatorname{arc\,sen}(\sqrt{\widehat{S}(0.2)}) \pm 0.5Z_{(1-\frac{\alpha}{2})}\sigma_s(0.2)\sqrt{\frac{\widehat{S}(0.2)}{1-\widehat{S}(0.2)}}\right\rangle\right\} \tag{13}$$

4. Criterio de comparación de los intervalos de confianza

Hay dos conceptos importantes que se deben de considerar al evaluar los intervalos de confianza: la precisión indicada por la longitud del intervalo y la probabilidad de cobertura $P(LI \leq S(t) \leq LS)$. Estos criterios no se pueden analizar por separado porque de poco nos sirve un intervalo con probabilidad de cobertura alta si su longitud es muy grande o un intervalo con una longitud muy pequeña, pero con probabilidad de cobertura muy baja.

Idealmente se quiere ver que los intervalos sean cortos y tengan probabilidad de cobertura muy cercana al nivel de confianza nominal, que los procedimientos que se utilicen para construir los intervalos de confianza den intervalos tales que sus longitudes sean pequeñas, pero diferentes de cero y que la probabilidad de cobertura no sea inferior al nivel de confianza nominal, Correa & Sierra (2003).

En este trabajo se calcula para cada método las tasas de error (TE), la longitud promedio del intervalo de confianza (LPI) para la función de supervivencia y la propuesta de Correa & Sierra (2003), del índice de comparación de intervalos de confianza (I), el cual tiene en cuenta simultáneamente el nivel de confianza real (NR), el nivel de confianza nominal (NN), la longitud promedio del intervalo (LPI) como

$$I = \frac{2 - LPI}{2} \times \frac{NR}{NN} \tag{14}$$

donde el nivel de confianza real (NR) corresponde a la proporción de intervalos simulados que cubre el verdadero valor de $S(t)$. Por lo tanto, mientras mayor sea el índice, mejor será el método, luego:

$$T.E = \frac{\#\ de\ I.C\ que\ no\ cubren\ el\ verdadero\ valor\ S(t)}{N}$$

$$NR = 1 - T.E$$

$$LPI = \sum_{i=1}^{N} \frac{(LS_i - LI_i)}{N}$$

donde N corresponde al número de simulaciones, LS y LI representa el límite superior e inferior respectivamente.

5. Estudio de simulación

En esta sección se presentan los resultados obtenidos al comparar los intervalos de confianza para la función de supervivencia en un tiempo de interés $t_i = 0.2$, este tiempo de interés fue escogido teniendo en cuenta la literatura y con el propósito de ser comparables con los resultados de Borgan & Liestøl (1990), mediante los estimadores de Kaplan-Meier y Nelson-Aalen, a través de los métodos de remuestreo *Bootstrap* aplicado a los estimadores, transformaciones log, $\log(-\log)$ y $\operatorname{arc\,sen}$. Estos métodos se compararon utilizando un algoritmo en R Development Core Team (2008).

5.1. Resultados simulación

Se utilizaron diferentes combinaciones de modelos generadores de los tiempos de supervivencia (Time) y de censura (Cens), como son el modelo $\exp(\lambda)$ y $Weib(\lambda, \beta)$, donde $\lambda = 2$, $\lambda = 1$ y $\beta = 0.5$ con porcentajes de censura tipo I, de 0 %, 15 %, 25 %, 35 %, 45 %, 55 % y tamaños de muestra $n = 25, 50, 75$ y 100.

La razón de utilizar los porcentajes de censura y los tamaños de muestra se debe a la finalidad de comparar los resultados con los resultados de algunos autores referenciados. Con el propósito de abreviar los títulos de las tablas se presenta las siguientes siglas, I.C: Intervalos de confianza, TE: Tasas de Error, LPI: Longitud promedio del intervalo (valores en paréntesis), NN: Nivel de confianza nominal, T.KM: I.C mediante el estimador de Kaplan-Meier, Boot.KM y Boot.NA: corresponden a los I.C de los estimadores de Kaplan-Meier y Nelson-Aalen, mediante el remuestreo *Bootstrap*. Adicionalmente log, $\log(-\log)$ y arc sen corresponden a las transformaciones de las funciones de supervivencia definidas en (10), (12) y (13), respectivamente.

Para determinar la efectividad de utilizar intervalos de confianza no paramétricos para la función de supervivencia frente a los paramétricos, los valores de referencia para el caso exponencial $\exp(2)$ la $S(t) = 0.67$ y $\exp(1)$ la $S(t) = 0.82$, mientras que para el caso $weib(1, 0.5)$ la $S(t) = 0.64$. Además, para el caso $\exp / \exp$ la $LPI = 0.53$, $\exp / weib$ la $LPI = 0.53$, $weib/ \exp$ la $LPI = 0.32$ y para $weib/weib$ la $LPI = 0.32$, los resultados para $\widehat{S}_{KM}(0.2)$ y $\widehat{S}_{NA}(0.2)$ correspondiente al estimador de Kaplan-Meier y Nelson-Aalen, respectivamente, son:

5.2. Resultados para $n = 25$

TABLA 1: TE (LPI) para un NN de 95 % con 0 % de censura y $n = 25$.

Estimador $\widehat{S}_{KM}(0.2)$					
Time/Cens	T.KM	Boot.KM	log	$\log(-\log)$	arc sen
$\exp / \exp$	0.05(0.37)	0.08(0.39)	**0.02**(0.63)	0.08(0.37)	0.05(0.36)
$\exp / weib$	0.03(0.32)	0.09(0.32)	**0.01**(0.42)	0.07(0.33)	0.07(0.31)
$weib/ \exp$	0.10(0.37)	0.08(0.38)	**0.02**(0.64)	0.05(0.37)	0.06(0.36)
$weib/weib$	0.10(0.37)	0.08(0.38)	**0.02**(0.64)	0.05(0.37)	0.06(0.36)
Estimador $\widehat{S}_{NA}(0.2)$					
Time/Cens	T.NA	Boot.NA	log	$\log(-\log)$	arc sen
$\exp / \exp$	0.05(0.36)	0.11(0.36)	**0.02**(0.62)	0.08(0.37)	0.05(0.36)
$\exp / weib$	0.03(0.31)	0.10(0.30)	**0.01**(0.41)	0.07(0.33)	0.07(0.30)
$weib/ \exp$	0.08(0.37)	0.10(0.36)	**0.02**(0.62)	0.05(0.37)	0.04(0.36)
$weib/weib$	0.08(0.37)	0.10(0.36)	**0.02**(0.62)	0.05(0.37)	0.04(0.36)

En los resultados para $n = 25$ se nota que cuando no hay observaciones censuradas, los I.C mediante la transformación $\log(S(t))$, presentan mejor comportamiento, esto se debe a que resultan ser más amplios, independientemente del estimador que se utilice y el modelo generador. Además a medida que se aumenta el porcentaje de censura los I.C para la función de supervivencia mediante la transformación $\log(-\log(S(t)))$ resultan ser mejores.

Tabla 2: TE y LPI de NN 95 % con 25 % de censura para $n = 25$.

Estimador $\widehat{S}_{KM}(0.2)$					
Time/Cens	T.KM	Boot.KM	log	log(− log)	arc sen
exp / exp	0.11(0.35)	0.16(0.37)	**0.01**(0.52)	0.02(0.36)	0.08(0.34)
exp / *weib*	0.06(0.29)	0.14(0.31)	0.05(0.36)	**0.04**(0.31)	0.05(0.29)
weib/ exp	0.17(0.35)	0.11(0.32)	0.11(0.53)	**0.04**(0.36)	0.11(0.34)
weib/*weib*	0.15(0.36)	0.13(0.33)	0.12(0.54)	**0.05**(0.36)	0.12(0.34)
Estimador $\widehat{S}_{NA}(0.2)$					
Time/Cens	T.NA	Boot.NA	log	log(− log)	arc sen
exp / exp	0.11(0.35)	0.17(0.34)	0.11(0.51)	**0.02**(0.36)	0.09(0.34)
exp / *weib*	0.057(0.29)	0.15(0.29)	0.05(0.35)	**0.02**(0.32)	0.05(0.28)
weib/ exp	0.19(0.35)	0.16(0.33)	0.12(0.51)	**0.04**(0.36)	0.11(0.34)
weib/*weib*	0.17(0.35)	0.17(0.34)	0.13(0.53)	**0.05**(0.37)	0.12(0.34)

Tabla 3: TE y LPI de NN 95 % con 45 % de censura para $n = 25$.

Estimador $\widehat{S}_{KM}(0.2)$					
Time/Cens	T.KM	Boot.KM	log	log(− log)	arc sen
exp / exp	0.31(0.33)	0.33(0.34)	0.26(0.43)	**0.08**(0.34)	0.21(0.32)
exp / *weib*	0.17(0.26)	0.17(0.29)	0.17(0.31)	**0.04**(0.30)	0.14(0.26)
weib/ exp	0.39(0.33)	0.41(0.35)	0.34(0.44)	**0.15**(0.34)	0.33(0.32)
weib/*weib*	0.37(0.34)	0.37(0.36)	0.34(0.46)	**0.14**(0.35)	0.30(0.33)
Estimador $\widehat{S}_{NA}(0.2)$					
Time/Cens	T.NA	Boot.NA	log	log(− log)	arc sen
exp / exp	0.32(0.32)	0.36(0.32)	0.29(0.42)	**0.08**(0.34)	0.23(0.31)
exp / *weib*	0.17(0.26)	0.18(0.27)	0.17(0.30)	**0.04**(0.30)	0.14(0.26)
weib/ exp	0.43(0.32)	0.43(0.33)	0.35(0.43)	**0.15**(0.34)	0.34(0.31)
weib/*weib*	0.39(0.33)	0.42(0.34)	0.35(0.45)	**0.14**(0.36)	0.32(0.32)

5.3. Resultados para $n = 50$

Tabla 4: TE y LPI de NN 95 % con 0 % de censura para $n = 50$.

Estimador $\widehat{S}_{KM}(0.2)$					
Time/Cens	T.KM	Boot.KM	log	log(− log)	arc sen
exp / exp	0.06(0.26)	0.04(0.27)	**0.02**(0.42)	0.08(0.26)	0.06(0.26)
exp / *weib*	0.04(0.22)	0.06(0.22)	**0.02**(0.28)	0.07(0.22)	0.08(0.22)
weib/ exp	0.08(0.26)	0.13(0.27)	**0.03**(0.41)	0.03(0.26)	0.05(0.26)
weib/*weib*	0.08(0.26)	0.13(0.27)	**0.03**(0.41)	0.03(0.26)	0.05(0.26)
Estimador $\widehat{S}_{NA}(0.2)$					
Time/Cens	T.NA	Boot.NA	log	log(− log)	arc sen
exp / exp	0.06(0.26)	0.06(0.26)	**0.02**(0.41)	0.05(0.26)	0.06(0.26)
exp / *weib*	0.04(0.22)	0.07(0.21)	**0.02**(0.28)	0.07(0.23)	0.04(0.21)
weib/ exp	0.08(0.26)	0.12(0.26)	**0.03**(0.41)	0.03(0.26)	0.04(0.26)
weib/*weib*	0.08(0.26)	0.12(0.26)	**0.03**(0.41)	0.03(0.26)	0.04(0.26)

Tabla 5: TE y LPI de NN 95 % con 25 % de censura para $n = 50$.

Estimador $\widehat{S}_{KM}(0.2)$					
Time/Cens	T.KM	Boot.KM	log	log(− log)	arc sen
exp / exp	0.17(0.25)	0.20(0.25)	0.11(0.35)	**0.08**(0.25)	0.15(0.24)
exp / *weib*	0.11(0.20)	0.17(0.21)	0.11(0.25)	**0.03**(0.21)	0.10(0.20)
weib/ exp	0.29(0.25)	0.26(0.24)	0.21(0.35)	**0.16**(0.25)	0.26(0.24)
weib/*weib*	0.27(0.25)	0.27(0.25)	0.19(0.36)	**0.16**(0.25)	0.23(0.25)
Estimador $\widehat{S}_{NA}(0.2)$					
Time/Cens	T.NA	Boot.NA	log	log(− log)	arc sen
exp / exp	0.17(0.25)	0.23(0.24)	0.12(0.34)	**0.09**(0.25)	0.16(0.24)
exp / *weib*	0.11(0.20)	0.17(0.20)	0.11(0.24)	**0.03**(0.21)	0.10(0.20)
weib/ exp	0.32(0.25)	0.25(0.30)	0.23(0.35)	**0.16**(0.25)	0.27(0.24)
weib/*weib*	0.29(0.25)	0.25(0.31)	0.21(0.35)	**0.16**(0.26)	0.25(0.25)

Tabla 6: TE y LPI de NN 95 % con 45 % de censura para $n = 50$.

Estimador $\widehat{S}_{KM}(0.2)$					
Time/Cens	T.KM	Boot.KM	log	log(− log)	arc sen
exp / exp	0.58(0.23)	0.59(0.24)	0.46(0.29)	**0.39**(0.23)	0.50(0.22)
exp / *weib*	0.33(0.18)	0.31(0.19)	0.32(0.21)	**0.09**(0.20)	0.23(0.18)
weib/ exp	0.73(0.23)	0.80(0.23)	0.62(0.29)	**0.57**(0.23)	0.67(0.22)
weib/*weib*	0.67(0.24)	0.74(0.24)	0.55(0.31)	**0.48**(0.24)	0.61(0.23)
Estimador $\widehat{S}_{NA}(0.2)$					
Time/Cens	T.NA	Boot.NA	log	log(− log)	arc sen
exp / exp	0.59(0.23)	0.60(0.23)	0.47(0.29)	**0.40**(0.23)	0.53(0.22)
exp / *weib*	0.34(0.18)	0.32(0.19)	0.33(0.21)	**0.09**(0.20)	0.25(0.18)
weib/ exp	0.74(0.23)	0.81(0.22)	0.65(0.29)	**0.57**(0.23)	0.71(0.22)
weib/*weib*	0.69(0.23)	0.74(0.24)	0.57(0.30)	**0.49**(0.24)	0.63(0.23)

Tabla 7: TE y LPI de NN 95 % con 0 % de censura para $n = 100$.

Estimador $\widehat{S}_{KM}(0.2)$					
Time/Cens	T.KM	Boot.KM	log	log(− log)	arc sen
exp / exp	0.06(0.18)	0.04(0.18)	**0.01**(0.28)	0.06(0.18)	0.06(0.18)
exp / *weib*	0.05(0.15)	0.07(0.15)	**0.02**(0.19)	0.06(0.16)	0.05(0.15)
weib/ exp	0.09(0.19)	0.19(0.18)	**0.04**(0.28)	0.06(0.18)	0.06(0.18)
weib/*weib*	0.09(0.19)	0.19(0.18)	**0.04**(0.28)	0.06(0.18)	0.06(0.18)
Estimador $\widehat{S}_{NA}(0.2)$					
Time/Cens	T.NA	Boot.NA	log	log(− log)	arc sen
exp / exp	0.06(0.18)	0.04(0.18)	**0.01**(0.28)	0.04(0.19)	0.06(0.18)
exp / *weib*	0.05(0.15)	0.07(0.15)	**0.02**(0.19)	0.06(0.16)	0.05(0.15)
weib/ exp	0.09(0.18)	0.18(0.18)	**0.04**(0.28)	0.06(0.19)	0.09(0.18)
weib/*weib*	0.09(0.18)	0.18(0.18)	**0.04**(0.28)	0.06(0.19)	0.09(0.18)

5.4. Resultados para $n = 100$

Se nota que los resultados para $n = 50$ son similares a $n = 25$. Es de resaltar que los I.C para la función de supervivencia mediante el remuestreo *Bootstrap* para Kaplan-Meier y Nelson-Aalen arrojan resultados inapropiados, esto se debe a que se utilizó la opción type = “norm”, el cual es sensible. Por otra parte en el último escenario la comparación de estos intervalos de confianza para la función

Tabla 8: TE y LPI de NN 95 % con 25 % de censura para $n = 100$.

Estimador $\widehat{S}_{KM}(0.2)$					
Time/Cens	T.KM	Boot.KM	log	log(− log)	arc sen
exp / exp	0.36(0.17)	**0.23**(0.19)	**0.23**(0.24)	0.26(0.18)	0.31(0.17)
exp /*weib*	0.19(0.14)	0.23(0.15)	0.14(0.17)	**0.09**(0.15)	0.14(0.14)
weib/ exp	0.60(0.17)	0.50(0.17)	**0.44**(0.24)	0.50(0.18)	0.57(0.17)
weib/*weib*	0.54(0.18)	0.56(0.19)	**0.38**(0.25)	0.45(0.18)	0.51(0.18)
Estimador $\widehat{S}_{NA}(0.2)$					
Time/Cens	T.NA	Boot.NA	log	log(− log)	arc sen
exp / exp	0.37(0.17)	0.33(0.20)	**0.24**(0.24)	0.27(0.18)	0.32(0.17)
exp /*weib*	0.20(0.14)	0.21(0.21)	0.15(0.17	**0.09**(0.15)	0.14(0.14)
weib/ exp	0.62(0.17)	0.62(0.22)	**0.45**(0.24)	0.52(0.18)	0.59(0.17)
weib/*weib*	0.56(0.18)	0.62(0.22)	**0.40**(0.24)	0.46(0.18)	0.53(0.18)

Tabla 9: TE y LPI de NN 95 % con 45 % de censura para $n = 100$.

Estimador $\widehat{S}_{KM}(0.2)$					
Time/Cens	T.KM	Boot.KM	log	log(− log)	arc sen
exp / exp	0.89(0.16)	0.89(0.16)	**0.80**(0.20)	0.83(0.16)	0.87(0.16)
exp /*weib*	0.53(0.13)	0.54(0.13)	0.45(0.15)	**0.35**(0.13)	0.48(0.13)
weib/ exp	0.98(0.16)	0.94(0.16)	**0.93**(0.20)	0.96(0.16)	0.97(0.16)
weib/*weib*	0.94(0.17)	0.87(0.17)	**0.88**(0.21)	0.91(0.17)	0.93(0.16)
Estimador $\widehat{S}_{NA}(0.2)$					
Time/Cens	T.NA	Boot.NA	log	log(− log)	arc sen
exp / exp	0.90(0.16)	0.91(0.16)	**0.81**(0.20)	0.83(0.16)	0.88(0.16)
exp /*weib*	0.54(0.13)	0.56(0.13)	0.51(0.15)	**0.35**(0.13)	0.48(0.13)
weib/ exp	0.98(0.16)	**0.93**(0.16)	0.94(0.20)	0.96(0.16)	0.98(0.16)
weib/*weib*	0.95(0.16)	**0.87**(0.16)	0.89(0.21)	0.91(0.17)	0.94(0.16)

de supervivencia en muestras de tamaño $n = 100$, los resultados respaldan lo mencionado con los demás tamaños de muestra, sin embargo es de resaltar que a medida que se aumenta el porcentaje de censura al caso más extremo, las tasas de error son bastantes altas en particular al aumentar de 35 % a 45 % observaciones censuradas en la muestra simulada, resultando similar a los resultados de Borgan & Liestøl (1990) para $t = 0.2$, sin embargo los autores no incluyen los I.C mediante la transformación $\log(-\log)$ que en este trabajo resultan ser mejores.

Por otra parte los resultados de los I.C mediante la transformación $\log(-\log)$ cuando se presenta altos porcentajes de censura coinciden con los resultados de Anderson, Bernstein & Pike (1982). Teniendo en cuenta los resultados anteriores en algunos casos se presentan confusiones para la escogencia del IC que presenta mejores resultados por lo que se implementó la metodología de Correa & Sierra (2003) para comparar dichos intervalos a través del índice propuesto por dichos autores, modificando el nivel de confianza nominal (NN $= 0.9, 0.95$ y 0.99), lo anterior se muestra tomando la distribución de fallas/censura exp / exp, esto con el fin de presentar el funcionamiento del índice presentado en el anexo.

Es de resaltar que mientras mayor sea el índice mejor será el intervalo de confianza, sin embargo éste índice fue propuesto para comparar intervalos de confianza para diferencia de proporciones, lo que cambiaría el rango de valores resultantes, pero la analogía de la interpretación se mantiene. Cabe resaltar que se realizaron 2000 simulaciones para la comparación y $B = 1000$ remuestras *Bootstrap*, estos valores se consideraron teniendo en cuenta que no se presentaron cambios significativos para un número mayor de simulaciones en las estimaciones.

6. Conclusiones

Cuando no se presentan observaciones censuradas en los datos los I.C mediante la transformación log poseen menor tasas de error, independientemente de la distribución de falla/censura y tamaños de muestra, como también el estimador utilizado, esto se debe a que resultan ser más amplios.

A medida que se aumenta el porcentaje de censura, los IC mediante la transformación $\log(-\log)$ resultan ser más efectivos, en general para los modelos generadores, tamaños de muestra y estimador de supervivencia utilizado.

Cuando las falla/censura se distribuyen exp / exp los resultados de la estimación de $S(0.2)$ resultan ser mejores que las demás combinaciones de distribución de falla/censura, en particular utilizando el estimador de KM.

[Recibido: septiembre de 2010 — Aceptado: marzo de 2011]

Referencias

Aalen, O. & Johansen, S. (1978), 'An empirical transition matrix for nonhomogeneus Markov chains based on censored observations', *Scandinavian Journal of Statistics* **5**(3), 141–150.

Akritas, M. (1986), 'Bootstrapping the Kaplan-Meier estimator', *American Statistical Association* **81**(396), 1032–1038.

Anderson, J., Bernstein, L. & Pike, M. (1982), 'Confidence intervals for probabilities of survival and quantiles in life-table analysis', *Biometrics* **38**(2), 407–416.

Borgan, Ø. & Liestøl, K. (1990), 'A note on confidence intervals and bands for the survival function based on transformations', *Scandinavian Journal of Statistics* **17**(1), 35–41.

Correa, J. & Sierra, E. (2003), 'Intervalos de confianza para la comparación de dos proporciones', *Revista Colombiana de Estadística* **26**(1), 61–75.

Greenwood, M. (1926), 'The natural duration of cancer', *Reports on Public Health and Medical Subjects* (33), 1–26.

Kalbfleisch, J. & Prentice, R. (1980), *The Statistical Analysis of Failure Time Data*, Wiley, New York, United States.

Kaplan, E. & Meier, P. (1958), 'Estimation from incomplete observations', *American Statistical Association* **53**(282), 457–481.

Nair, V. (1984), 'Confidence bands for survival functions with censored data', *Technometrics* **26**(3), 265–275.

Nelson, W. (1969), 'Hazard plotting for incomplete failure data', *Journal of Quality Technology* **61**(1), 27–52.

R Development Core Team (2008), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
*http://www.R-project.org

Apéndice
Tablas

TABLA 10: Índice de IC para $\widehat{S}_{KM}(t)$ con 15 % de censura.

NN	Trad.K-M	Boot.K-M	log	log(− log)	arc sen
			n=25		
90 %	0.85	0.84	0.87	0.89	0.84
95 %	0.78	0.78	0.79	0.81	0.80
99 %	0.73	0.73	0.73	0.74	0.73
			n=50		
90 %	0.83	0.86	0.89	0.90	0.84
95 %	0.82	0.82	0.85	0.87	0.83
99 %	0.80	0.79	0.81	0.82	0.81
			n=75		
90 %	0.83	0.85	0.90	0.88	0.86
95 %	0.83	0.83	0.87	0.87	0.85
99 %	0.82	0.83	0.84	0.85	0.84
			n=100		
90 %	0.89	0.82	0.90	0.85	0.85
95 %	0.84	0.83	0.85	0.88	0.85
99 %	0.83	0.82	0.83	0.86	0.84

Tabla 11: Índice de IC para $\widehat{S}_{NA}(t)$ con 15 % de censura.

NN	Trad.N-A	Boot.N-A	log	log(− log)	arc sen
			n=25		
90 %	0.80	0.81	0.87	0.88	0.85
95 %	0.78	0.77	0.79	0.81	0.79
99 %	0.73	0.75	0.73	0.74	0.73
			n=50		
90 %	0.83	0.86	0.89	0.90	0.84
95 %	0.82	0.85	0.84	0.86	0.82
99 %	0.79	0.80	0.81	0.82	0.81
			n=75		
90 %	0.82	0.86	0.89	0.90	0.82
95 %	0.82	0.82	0.87	0.87	0.84
99 %	0.82	0.82	0.84	0.85	0.83
			n=100		
90 %	0.88	0.80	0.89	0.89	0.80
95 %	0.84	0.74	0.85	0.88	0.85
99 %	0.82	0.71	0.84	0.86	0.84