

*MODELO DE REGRESIÓN BIRNBAUM-SAUNDERS
EN LA PRESENCIA DE EFECTOS NO
PARAMÉTRICOS Y CENSURA NO INFORMATIVA*

Laura Marcela Rúa Yáñez
Matemático
Código: 2949873115



UNIVERSIDAD NACIONAL DE COLOMBIA
FACULTAD DE CIENCIAS
DEPARTAMENTO DE ESTADÍSTICA
BOGOTÁ, D.C.
MAYO DE 2018

*MODELO DE REGRESIÓN BIRNBAUM-SAUNDERS
EN LA PRESENCIA DE EFECTOS NO
PARAMÉTRICOS Y CENSURA NO INFORMATIVA*

LAURA MARCELA RÚA YÁNEZ

MATEMÁTICO

CÓDIGO: 2949873115

DISERTACIÓN PRESENTADA PARA OPTAR AL TÍTULO DE
MAGISTER EN CIENCIAS ESTADÍSTICA

DIRECTOR

LUIS HERNANDO VANEGAS PENAGOS, PH.D.

DOCTOR EN CIENCIAS ESTADÍSTICA



UNIVERSIDAD NACIONAL DE COLOMBIA

FACULTAD DE CIENCIAS

DEPARTAMENTO DE ESTADÍSTICA

BOGOTÁ, D.C.

MAYO DE 2018

Título en español

Modelo de regresión Birnbaum-Saunders en la presencia de efectos no paramétricos y censura no informativa. 2018. Tesis (Maestría)-Universidad Nacional de Colombia, 2018.

Title in English

Birnbaum-Saunders regression model under the presence of non-parametric effects and non-informative censoring. 2018. Thesis (master's degree)-Universidad Nacional de Colombia, 2018.

Resumen: El principal objetivo de este trabajo es proporcionar una versión más flexible del modelo de regresión Log-Birnbaum-Saunders bajo la presencia de censura aleatoria y no informativa, donde se asume que los efectos de algunas variables explicativas son no lineales pero su forma funcional es desconocida.

La flexibilidad que proporciona el modelo que se propone radica en su capacidad para describir la mediana de la distribución del tiempo de vida mediante un conjunto de covariables haciendo uso de una suma de funciones arbitrarias, cuya forma funcional es estimada de los datos (observaciones), reduciendo así la posibilidad de error de especificación, lo que permite estimar los parámetros de forma robusta bajo la presencia de extremos o outliers, incluyendo observaciones censuradas no informativas.

Abstract: The main objective of this work is to provide a more flexible version of the Log-Birnbaum-Saunders regression model under the presence of random and non-informative censoring, where it is assumed that the effects of some explanatory variables are nonlinear but their functional form is unknown .

The flexibility provided by the proposed model lies in its ability to describe the median life-time distribution by a set of covariates that make use of a sum of arbitrary functions, which is estimated from the data (observations), thus reducing the possibility of specification error, which allows the estimation of robust parameters under the presence of extremes values or outliers, including non-informative censored observations.

Palabras clave: Log-Birnbaum-Saunders, estimación robusta, censura no informativa, tiempo de vida, efectos no-paramétricos.

Keywords: Log-Birnbaum-Saunders, robust estimation, non-informative censoring, life time, non-parametric effects.

Nota de aceptación

Trabajo de tesis

Aprobado

“Mención Meritoria o Laureada”

Jurado

Jurado

Jurado

Director
Luis Hernando Vanegas Penagos

Bogotá, D.C., Mayo de 2018

Dedicado a

Dedicado a mi madre quién con su perseverancia y amor siempre tiene una voz de aliento para apoyarme en cada proyecto que emprendo.

Agradecimientos

Hoy que he culminado un proyecto más en mi vida puedo decir con certeza que mi crecimiento profesional y personal ha sido gratificante y doy gracias por aquellas personas que siempre han tenido una voz de aliento y de apoyo cuando más lo he necesitado:

A mi madre **Yolanda** y a mi hermano **Stheven** quienes desde siempre han sido mi motor y mi más valioso apoyo sobre todo cuando he intentado desfallecer.

A mis grandes amigos **Alvaro, Estefanny, Yurle, Luis, Saira, Liz** y **Leidy** porque siempre han estado de manera incondicional, gracias pues han sido parte fundamental de mi vida.

A mi director de tesis **Phd. Luis Hernando Vanegas** quién con su sabiduría y paciencia me forjó como estudiante y me convirtió en una mejor profesional. Gracias infinitamente por creer en mis capacidades y en mis ganas de ser mejor cada día.

A los profesores del programa de la Maestría en Ciencias Estadística quienes me enseñaron disciplina, amor profesional, análisis crítico, resistencia y sobre todo el poder de esta ciencia, en especial a los profesores **Ramón Giraldo** y **Alvaro Montenegro**, agradezco infinitamente su interferencia en mi crecimiento profesional.

A la **Universidad Nacional de Colombia sede Bogotá** por recibirme en su seno y haberme dado la oportunidad de vivir una de las mejores experiencias profesionales y además el poder compartirlas como aquellas personas que llenan mi entorno diario.

Índice general

Índice general	I
Índice de tablas	III
Índice de figuras	IV
1. INTRODUCCIÓN	1
2. DISTRIBUCION BIRNBAUM-SAUNDERS PARA DATOS CENSURADOS	4
2.1. Distribución Birnbaum-Saunders bajo el supuesto de daño acumulado	7
2.2. Modelo de regresión Log-Birnbaum-Saunders en la presencia de censura no informativa a la derecha	10
2.3. Distribución Birnbaum-Saunders-t	14
2.3.1. Algoritmo Newton-Raphson	17
2.3.2. Inferencia Estadística	18
3. REGRESIÓN BIRNBAUM-SAUNDERS CENSURADO CON EFECTOS NO PARAMÉTRICOS	20
3.1. Formulación del Modelo de regresión Log-Birnbaum-Saunders con efectos no paramétricos y censura no informativa	23
3.2. Estimación de parámetros del modelo de regresión Log-Birnbaum-Saunders con efectos no paramétricos y censura no informativa	25
3.2.1. Algoritmo Newton-Raphson	26
3.2.2. Selección de los parámetros de suavización	27
3.2.3. Inferencia estadística	27
3.3. Estudio de simulación	28
3.3.1. Resultados	29
3.4. Aplicación	37

A. Algoritmos	44
A.1. Datos	44
A.2. Implementación del algoritmo de la sección 3.3	45
A.3. Algoritmo de Regresión log-Birnbaum-Saunders censurado paramétrico y no paramétrico con censura no informativa.	47
A.3.1. Datos de fatiga de bajo ciclo para una prueba de tensión controlada de 26 especímenes cilíndricos a base de níquel.	47
A.4. Algoritmo de Regresión log-Birnbaum-Saunders-t(4) censurado paramétrico y no paramétrico con censura no informativa.	49
A.4.1. Datos de fatiga de bajo ciclo para una prueba de tensión controlada de 26 especímenes cilíndricos a base de níquel.	49
Trabajo futuro	52
Bibliografía	53

Índice de tablas

3.1. Estimaciones de los $\hat{\beta}$'s usando los modelos descritos anteriormente por los autores: Nelson(1984), Meeker & Escobar(1998) y Heuchenne & Van Keilegom(2007)	38
3.2. Estimaciones de los $\hat{\beta}$'s usando los modelos descritos anteriormente por los autores: Nelson(1984), Meeker & Escobar(1998) y Heuchenne & Van Keilegom(2007)	38
3.3. (4) Estimaciones del modelo Log-Bs censurado paramétrico y con efectos no paramétricos	39
3.4. Estimaciones del modelo Log-Bs-t(4) censurado paramétrico y con efectos no paramétricos	39
3.5. Valores de los criterios AIC y BIC para los modelos Log-BS Censurado y Log-BS Censurado con efectos no paramétricos	41
3.6. Valores de los criterios AIC y BIC para los modelos Log-BS-t(4) Censurado y Log-BS-t(4) Censurado con efectos no paramétricos	41
3.7. Valores de los criterios AIC y BIC para los modelos estimados por Meeker & Escobar(1998) y Nelson(1984)	42
3.8. Valores de los criterios AIC y BIC para los modelos Log-BS Censurado y Log-BS-t Censurado con efectos no paramétricos comparables con los calculados por Meeker & Escobar(1998) y Nelson(1984)	42
3.9. Valores de los criterios AIC y BIC para los modelos Log-BS-t(4) Censurado y Log-BS-t(4) Censurado con efectos no paramétricos comparables con los calculados por Meeker & Escobar(1998) y Nelson(1984)	42
A.1. Ciclos de fatiga de especímenes a base de Níquel bajo un nivel de Pseudo-estrés (En unidades de miles de ciclos hasta la falla.)	44

Índice de figuras

2.1. Densidad Birnbaum-Saunders para diferentes valores de α y η	9
2.2. (2) Densidad Birnbaum-Saunders-t para diferentes valores de α y ν con $\eta = 1$	15
3.1. Curva de estimación simulada del modelo Log-BS censurado con efectos no paramétricos (A) y Log-BS-t censurado con efectos no paramétricos (B):($\alpha = 0.5, \gamma = 0.003$)	29
3.2. Curva de estimación simulada del modelo Log-BS censurado con efectos no paramétricos (A) y Log-BS-t censurado con efectos no paramétricos (B) ($\alpha = 0.5, \gamma = 0.00003$)	30
3.3. Curva de estimación simulada del modelo Log-BS censurado con efectos no paramétricos (a) y Log-BS-t censurado con efectos no paramétricos (b) ($\alpha = 1.5, \gamma = 0.003$)	31
3.4. Curva de estimación simulada del modelo Log-BS censurado con efectos no paramétricos (A) y Log-BS-t censurado con efectos no paramétricos ($\alpha = 1.5, \tau = 0.0003$)	32
3.5. Curva de estimación simulada del modelo Log-BS censurado con efectos no paramétricos (A) y Log-BS-t censurado con efectos no paramétricos ($\alpha = 2.0, \tau = 0.003$)	33
3.6. Curva de estimación simulada del modelo ((A) : $\alpha = 1.0, \tau = 0.0003$); ((B) : $\alpha = 1.5, \tau = 0.0003$)	34
3.7. Curva de estimación simulada del modelo Log-BS censurado con efectos no paramétricos (A) y Log-BS-t censurado con efectos no paramétricos ($\alpha = 2.0, \tau = 0.00003$)	35
3.8. Curva de estimación simulada del modelo Log-BS censurado con efectos no paramétricos (A) y Log-BS-t censurado con efectos no paramétricos ($\alpha = 1.0, \tau = 0.003$)	36
3.9. Logaritmo del número de ciclos antes de la falla contra el Pseudoestress. . .	37
3.10. Ajuste del modelo Log-Bs Censurado para los datos de fatiga en la prueba de tensión (A) y Ajuste del modelo Log-Bs Censurado con efectos no paramétricos para los datos de fatiga en la prueba de tensión (B)	40

3.11. Ajuste del modelo Log-Bs-t Censurado para los datos de fatiga en la prueba de tensión (A) y Ajuste del modelo Log-Bs-t Censurado con efectos no paramétricos para los datos de fatiga en la prueba de tensión (B)	40
---	----

INTRODUCCIÓN

Los modelos de regresión lineales y no lineales proporcionan un marco rico y flexible que se adapta a las necesidades de muchos analistas en los distintos campos de aplicación, sin embargo no siempre lo hacen de forma adecuada y esto impide que se pueda dar una solución conveniente frente a las distintas situaciones que se puedan presentar. Para solucionar las limitaciones frente a la modelación correcta de los datos se han propuesto varias extensiones de los modelos clásicos de regresión, relajando los supuestos de los errores y de la componente sistemática que forma el modelo, en particular, el supuesto de los errores simétricos. Este supuesto ha sido de gran estudio especialmente en la última década. Uno de los principales atractivos de relajar los supuestos de los modelos clásicos de regresión es la posibilidad de permitir distribuciones para los errores aleatorios con colas ligeras o colas pesadas. Esta flexibilidad puede enriquecer el análisis de tiempos de vida, los cuales se llevan a cabo usando distribuciones estrictamente positivas y asimétricas. Sin embargo, en algunas ocasiones, el deterioro de elementos (por ejemplo, componentes, sistemas, subsistemas, especímenes, unidades, entre otros) no se mide en términos cronológicos, si no por medio de otras variables aleatorias, por ejemplo, el nivel de degradación de un elemento, la flexibilidad de un adhesivo, el número de ciclos hasta que un material falle debido a la fatiga (en este caso fatiga = vida).

En 1969, motivados por los problemas de vibración encontrados en los nuevos aviones comerciales y por los problemas de fatiga de materiales, Birnbaum y Saunders (1969) presentaron un modelo probabilístico ingenioso para tiempos de vida de estructuras bajo carga dinámica bajo el supuesto de daño acumulado. La distribución Birnbaum-Saunders se derivó a partir de un modelo que muestra que las fallas se deben al desarrollo y crecimiento de una fractura dominante. Así, la distribución Birnbaum-Saunders describe totalmente el tiempo de falla que ha transcurrido cuando cierta clase de daño acumulado excede un umbral ω . Aunque esta distribución se conoce como el modelo de Birnbaum-Saunders, un trabajo anterior atribuido a Freudenthal & Shinozuca (1961) introdujo un modelo similar con una parametrización diferente, además Fréchet (1927) presentó algunas ideas análogas relacionadas con la de Birnbaum-Saunders pero también con una parametrización distinta a la considerada por ellos.

Los modelos de regresión Log-Birnbaum-Saunders se aplican basados en la distribución Sinh-Normal descrita por Rieck & Nedelman (1991), como un caso particular de la distribución (BS), de acuerdo a estos autores, la distribución Sinh-normal es simétrica, y para un α muy pequeño, converge a una distribución normal, a su vez, presenta grados de curtosis más grandes y pequeños que la distribución normal, y admite unimodalidad y bimodalidad.

Estos modelos se han aplicado extensivamente en estudios de datos de fatiga de materiales en los cuales, el tiempo total hasta la falla puede estar relacionada con algún tipo de daño acumulado. Sin embargo, las aplicaciones de esta distribución no se reducen al estudio de falla de materiales. Por ejemplo, en el campo médico, es de interés estudiar diversas enfermedades en las que el daño acumulado causado por diversos factores de riesgo que pueden conducir a una degradación y esta a su vez conduce a un proceso de fatiga (Barros (2008)).

En años recientes los modelos de fatiga (tiempos de falla) o modelos de log-localización-escala con una distribución de error específica y censura aleatoria no informativa han sido objeto de estudio por diversos investigadores como Wood(2006), Lin(2009), Vane-gas(2015a), entre otros, han propuesto diversas estrategias para introducir las distribuciones asimétricas en los modelos de regresión, en los que se incluye la presencia de efectos no paramétricos. Generalmente se ha modelado un solo parámetro con la distribución de fatiga pero no se admite la presencia de efectos no paramétricos en sus componentes sistemáticas.

Este documento presenta una versión más flexible del modelo de regresión Log-Birnbaum-Saunders en la presencia de censura aleatoria y no informativa (véase, por ejemplo, Leiva(2016)), donde se asume que los efectos de algunas variables explicativas son no lineales pero sus formas funcionales son desconocidas.

La flexibilidad proporcionada por el modelo propuesto radica en su capacidad para describir la mediana de la distribución del tiempo de vida (fatiga) mediante un conjunto de covariables usando una suma de funciones arbitrarias, cuya forma funcional es estimada de los datos, reduciendo así la posibilidad del error de especificación.

Este documento se encuentra organizado como sigue:

En el capítulo 2 se presenta la distribución Birnbaum-Saunders, sus características más importantes y el modelo de regresión log-Birnbaum-saunders para datos censurados que permite el análisis de datos de fatiga(tiempo de vida), además del estimador de máxima verosimilitud propuesto para este modelo(Véase, Leiva (2016)). Adicionalmente, se considera la distribución Birnbaum Saunders- t como una distribución más robusta (posee colas más pesadas) que la distribución Birnbaum-Saunders, la cual según Barros M (2008) produce estimaciones de parámetros que son menos sensibles a los datos atípicos en comparación a la distribución Birnbaum-Saunders, así mismo, se presenta el modelo de regresión para esta distribución y el estimador de máxima verosimilitud.

En el capítulo 3 se propone un modelo más flexible que el modelo de regresión Log-Birnbaum-Saunders para datos censurados, el cuál asume la presencia de efectos no lineales cuya forma funcional es desconocida pero se puede aproximar usando splines cúbicos naturales (Green y Silverman(1994)) o P-splines cúbicos (Eilers y Marx(1996)). Los parámetros se estiman por el método de máxima verosimilitud penalizada usando algo-

ritmos como Newton-Raphson. Adicionalmente, se presenta una aplicación basada en un conjunto de datos presentados y analizados por Nelson(1984), posteriormente por Meeker & Escobar (1998) y finalmente por Heuchene & Vankeilegom (2007). Se hicieron las correspondientes estimaciones de los datos basados en el método propuesro y finalmente se analizaron los resultados.

CAPÍTULO 2

DISTRIBUCION BIRNBAUM-SAUNDERS PARA DATOS CENSURADOS

DISTRIBUCION BIRNBAUM-SAUNDERS PARA DATOS CENSURADOS

La distribución Birnbaum-Saunders (Birnbaum & Saunders (1969)) es una alternativa para modelar datos de tiempo de vida o datos de fatiga de un material en los cuales, el tiempo total hasta la falla puede estar relacionado con algún tipo de daño acumulado.

El daño acumulado puede ser causado por diversos factores de riesgo que pueden conducir a una degradación, y esta a su vez a un proceso de fatiga.

Se considera esta distribución ya que ella fué construída bajo el supuesto de daño acumulado, y posee dos propiedades importantes: cerrada bajo cambios de escala y cerrada bajo recíprocos; propiedades que según Puig(2008) son deseables en distribuciones que se usan para describir datos obtenidos como el cociente de cantidades positivas.

Birnbaum & Saunders (1969) consideran muestras de materiales que se someten cíclicamente al estrés, asumiendo que la falla en el material puede ser atribuída al inicio, crecimiento y desarrollo final de una grieta o daño dominante, es decir, se asume que la falla se debe a la fatiga del material. También se asume que la grieta o el daño en el material aumenta en cada ciclo en una cantidad que se considera aleatoria debido a las fluctuaciones en el material, así como en la magnitud del estrés a la que es sometido. De la misma forma, se supone que el aumento en la extensión de la grieta debido al j -ésimo ciclo, denotado M_j es una variable aleatoria con media δ y varianza σ^2 , $j = 1, 2, \dots$. Entonces el daño acumulado en el material hasta el ciclo t , denotado $W(t)$, se puede escribir como:

$$W(t) = \sum_{j=1}^t M_j$$

Por lo tanto, la distribución de T , el tiempo de vida del material, es una representación del número de ciclos necesarios para forzar el daño de la grieta hasta exceder su valor crítico ω . Suponiendo que T es bastante grande para usar el teorema central del límite, la distribución de $W(t)$ puede ser aproximada usando la distribución normal.

2.1. Distribución Birnbaum-Saunders bajo el supuesto de daño acumulado

Como una distribución de fatiga, el modelo Birnbaum-Saunders considera un espécimen expuesto a una secuencia de t cargas cíclicas, denotado por $\{l_i, i = 1, 2, \dots, t, t \in \mathbb{R}\}$. El esquema de la carga puede ser descrita como sigue:

$$\underbrace{l_1, l_2 \cdots l_t}_{\text{Primer Ciclo}} \quad \underbrace{l_{t+1}, l_{t+2} \cdots l_{2t}}_{\text{Segundo Ciclo}} \quad \cdots \quad \underbrace{l_{jt+1}, l_{jt+2} \cdots l_{j+1t}}_{(j+1) \text{ésimo Ciclo}}, \quad (2.1.1)$$

donde $l_{jt+i} = l_{kt+i}$, para $j \neq k$.

El proceso de fatiga que utilizó Birnbaum y Saunders (1969a) se basó en lo siguiente:

- Un material está sujeto a cargas cíclicas o choques repetitivos, que producen una grieta o desgaste en el material.
- La falla en el material ocurre cuando el tamaño de la grieta del material excede un nivel de resistencia (umbral), denotado por ω .
- La secuencia de cargas impuestas al material es la misma de un ciclo a otro.
- Durante un ciclo de carga, la extensión de la grieta causada por la carga l_i , X_i , es una variable aleatoria cuya distribución está gobernada por todas las cargas L_j , para $i < j$, con $i = 1, \dots, m$, y por la extensión de la grieta que lo precede
- El tamaño total de la grieta causada por el ciclo k , Y_k , es una variable aleatoria que sigue una distribución de la media μ y la varianza σ^2 , ambas finitas.
- El tamaño de las grietas (Y_k) en diferentes ciclos es independiente.

Por tanto, basado en los 6 supuestos anteriores de la distribución Birnbaum-Saunders se define la distribución de T como sigue:

definición 2.1.1. *Sea t_1, t_2, \dots, t_n muestras aleatorias de materiales que se someten cíclicamente al estrés donde la distribución de T es el tiempo de vida del material que representa el número de ciclos necesarios para forzar el daño de la grieta hasta exceder su valor crítico ω .*

La distribución de $W(t)$ puede ser aproximada usando la distribución normal, por lo tanto, la función de distribución acumulada de T , $F_T(t)$, puede ser expresada como:

$$\begin{aligned} F_T(t) &= P[T \leq t] = P[W(t) \geq \omega] \\ &= 1 - \Phi \left[\frac{\omega - t\delta}{\sigma\sqrt{t}} \right] \\ &= \Phi \left[\frac{t\delta - \omega}{\sigma\sqrt{t}} \right] \quad (\text{Por la simetría de la distribución normal}) \\ &= \Phi \left[\frac{\sqrt{t}\delta}{\sigma} - \frac{\omega}{\sigma\sqrt{t}} \right], \end{aligned}$$

siendo $\Phi(\cdot)$ la función de distribución acumulada de la normal estándar. Con $\alpha = \sigma/\sqrt{\delta\omega}$ y $\eta = \omega/\delta$ se tiene que $F_T(t)$ puede escribirse como:

$$F_T(t) = \Phi \left[\frac{1}{\alpha} \left(\sqrt{\frac{t}{\eta}} - \sqrt{\frac{\eta}{t}} \right) \right], \quad t > 0 \quad (2.1.2)$$

Si la función de distribución acumulada de T se puede escribir como 2.1.2, entonces T sigue una distribución Birnbaum-Saunders y se denota $T \sim BS(\alpha, \eta)$, donde $\alpha > 0$ y $\eta > 0$ son los parámetros de forma y escala, respectivamente.

Desmond (1985) proporcionó una derivación más general de esta distribución basada en un modelo biológico. También reforzó la justificación física para el uso de esta distribución relajando algunas suposiciones consideradas anteriormente por Birnbaum y Saunders (1969). A partir de $F_T(t)$ se puede encontrar la función de densidad de T, $f_T(t)$, la cual se puede expresar en la siguiente forma:

$$f_T(t) = \phi \left[\frac{1}{\alpha} \left(\sqrt{\frac{t}{\eta}} - \sqrt{\frac{\eta}{t}} \right) \frac{1}{2\alpha\eta} \left(\sqrt{\frac{\eta}{t}} - \sqrt{\frac{\eta^3}{t^3}} \right) \right], \quad t > 0, \quad (2.1.3)$$

siendo $\phi(\cdot)$ la función de densidad de la distribución normal estándar.

Entre las propiedades más importantes de $T \sim BS(\alpha, \eta)$ (ver Birnbaum & Saunders(1969)), con parámetros de forma y escala dados por $\alpha > 0$ y η , respectivamente se tienen:

$$(B1) \quad E [T] = \eta \left(1 + \frac{1}{2}\alpha^2 \right)$$

$$(B2) \quad Var [T] = (\alpha\eta)^2 \left(1 + \frac{5}{4}\alpha^2 \right)$$

$$(B3) \quad \eta \text{ es la mediana de T}$$

$$(B4) \quad cT \sim BS(\alpha, c\eta), \text{ con } c > 0 \text{ una constante;}$$

$$(B5) \quad T^{-1} \sim BS(\alpha, \eta^{-1})$$

$$(B6) \quad \frac{1}{\alpha} \left(\sqrt{\frac{t}{\eta}} - \sqrt{\frac{\eta}{t}} \right) \sim N(0, 1)$$

$$(B7) \quad \text{El percentil } 100(q)\% \text{ de T está dada por } t(q) = \eta \left[\alpha z_q + (\alpha^2 z_q^2 + 4)^{1/2} \right]^2 / 4, \text{ es decir,}$$

$$P [T \geq t(q)] = q, \text{ en que } z_q \text{ es el cuantil } 100(q)\% \text{ de la distribución normal estándar } (N(0, 1)).$$

Aunque la distribución Birnbaum-Saunders (BS) fué desarrollada para describir fatiga de materiales, se ha aplicado en otras áreas, por ejemplo en muchos problemas médicos como enfermedades cardiacas crónicas y diferentes tipos de cáncer, el daño se acumula causando que varios factores de riesgo se presenten. Esta degradación lleva a un proceso de fatiga cuyos tiempos de propagación pueden ser adecuadamente modelados por la distribución

Birnbaum-Saunders (Véase Barros (2008)). Leiva (2016) aplicó la distribución Birnbaum-Saunders para modelar tiempos de sobrevivencia en pacientes con melanoma múltiple. Podlaski (2008) usa la distribución Birnbaum-Saunders para caracterizar los diámetros de la distribución de una floresta natural. Leiva (2011a) desarrollan una metodología de gráficos de control basados en la distribución Birnbaum-Saunders, en que el monitoreo del tiempo de vida se realiza usando el tiempo acumulado o el tiempo medio hasta la ocurrencia de una serie de fallas.

La figura (2.1) representa el comportamiento de $f_T(t)$ para diferentes valores de α y η . Se puede observar que la distribución de T presenta diferentes niveles de asimetría. De hecho, para η fijo, la distribución de T tiende a ser simétrica a medida que el valor de α tiende a cero. Además para α fijo, la varianza de T aumenta a medida que η aumenta.

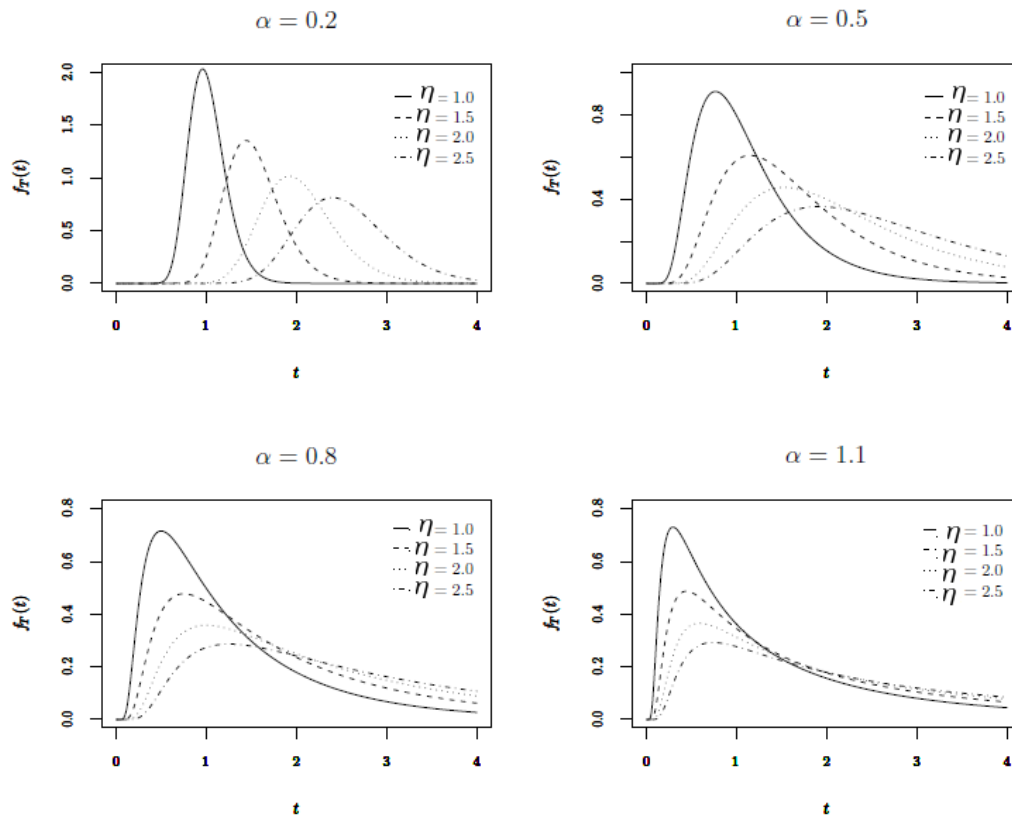


FIGURA 2.1. Densidad Birnbaum-Saunders para diferentes valores de α y η

2.2. Modelo de regresión Log-Birnbaum-Saunders en la presencia de censura no informativa a la derecha

Una manera de llevar a cabo el análisis de datos usando la distribución Birnbaum-Saunders es por medio de la distribución Sinh-Normal descrita por *Rieck & Nedelman (1991)*.

De acuerdo a estos autores, la distribución Sinh-Normal es simétrica *Rieck & Nedelman (1991)* y, para α muy pequeño, converge a una distribución normal. presenta grados de curtosis más grandes y pequeños que la distribución normal, y admite unimodalidad y bimodalidad.

La distribución Sinh-normal se obtiene de la transformación:

$$Y = \mu + \sigma \sinh^{-1} \left(\frac{\alpha Z}{2} \right),$$

donde $Z \sim N(0, 1)$, $\alpha > 0$ es el parámetro de escala, $\mu \in \mathbb{R}$ es un parámetro de localización. En este caso se usa la notación $Y \sim SHN(\mu, \sigma, \alpha)$. La función de densidad de probabilidad de Y está dada por:

$$f_Y(y; \alpha, \mu, \sigma) = \phi \left(\frac{2}{\alpha} \sinh \left(\frac{y - \mu}{\sigma} \right) \right) \frac{2 \cosh((y - \mu)/\sigma)}{\alpha \sigma}, \quad y \in \mathbb{R}. \quad (2.2.1)$$

La función de distribución acumulativa de Y está definida por:

$$F_Y(y; \alpha, \mu, \sigma) = \Phi \left(\frac{2}{\alpha} \sinh \left(\frac{y - \mu}{\sigma} \right) \right), \quad y \in \mathbb{R}. \quad (2.2.2)$$

La función cuantil de Y está definida como:

$$y(q; \alpha, \mu, \sigma) = F_Y^{-1}(q; \alpha, \mu, \sigma) = \mu + \sigma \sinh^{-1} \left(\frac{\alpha z(q)}{2} \right), \quad 0 \leq q \leq 1, \quad (2.2.3)$$

donde $z(q)$ es el $q \times 100$ -ésimo cuantil de $Z \sim N(0, 1)$ y $F_Y^{-1}(\cdot)$ es la función inversa de $F_Y(\cdot)$. La media y la varianza de Y se puede obtener usando la función generadora de momentos dada por:

$$m(r) = \exp(r\mu) \left(\frac{K_{r\sigma+1/2}(1/\delta^2)}{2K_{1/2}(1/\delta^2)} \right), \quad (2.2.4)$$

donde $E(Y) = \mu$ y $K_\lambda(\cdot)$ es la función de Bessel modificada del tercer tipo con la cual se obtuvieron los momentos de todos los órdenes para la distribución Birnbaum-Saunders dada por:

$$K_\lambda(u) = \frac{1}{2} \left(\frac{u}{2} \right)^\lambda \int_0^\infty v^{-\lambda-1} \exp \left(-v - \frac{u^2}{4v} \right).$$

La función de confiabilidad y la tasa de falla de $Y \sim SHN(\mu, \sigma, \alpha)$ están dadas por:

$$R_Y(y; \alpha, \mu, \sigma) = \Phi\left(-\frac{2}{\alpha} \sinh\left(\frac{y-\mu}{\sigma}\right)\right), \quad y \in \mathbb{R}$$

$$h_Y(y; \alpha, \mu, \sigma) = \frac{2\phi\left(\frac{2}{\alpha} \sinh\left(\frac{y-\mu}{\sigma}\right) \cosh\left(\frac{y-\mu}{\sigma}\right)\right)}{\alpha\sigma\Phi\left(-\frac{2}{\alpha} \sinh\left(\frac{y-\mu}{\sigma}\right)\right)} \quad y \in \mathbb{R},$$

respectivamente.

Rieck (1989) mostró que una distribución sinh-normal con parámetros α , μ y σ es:

- (a) Simétrica alrededor de la media μ
- (b) Unimodal para $\alpha < 2$ y su curtosis es más pequeña que la curtosis de una distribución normal.
- (c) Platicúrtica para $\alpha = 2$
- (d) Bimodal para $\alpha > 2$, es decir, cuando α aumenta, la distribución se convierte en bimodal y su curtosis es mayor que de la distribución normal. Adicionalmente, si $Y \sim SHN(\mu, \sigma, \alpha)$ entonces, cuando α se aproxima a cero

$$\frac{2(Y - \mu)}{\alpha\sigma} \sim N(0, 1).$$

Rieck & Nedelman(1991) probaron que si $T \sim (\alpha, \eta)$, entonces $Y = \log(T) \sim SHN(\mu, \sigma = 2, \alpha)$, donde $\mu = \log(\eta)$. Por esta razón la distribución Sinh-normal también se conoce como **Distribución Log-Birnbaum-Saunders**, esto es, si se basa en la ecuación 2.2.4, $Y = \log(T)$ y la relación:

$$M_Y(r) = E(\exp(rY)) = E(T^r), \quad (2.2.5)$$

los momentos de cualquier orden de la distribución Birnbaum-Saunders pueden computarse ((*Rieck(1999)*)).

En conclusión, la estimación de los parámetros de la distribución Birnbaum-Saunders y la generación de los números aleatorios de esta, pueden obtenerse de forma más eficiente por la distribución Sinh-Normal.

Formulación del modelo de regresión Log-Birnbaum-Saunders Censurado

definición 2.2.1.

Sean t_1, \dots, t_n las mediciones de una característica de interés realizadas en n sujetos, se asume que son realizaciones de n variables aleatorias independientes, denotadas T_1, \dots, T_n , que siguen distribución Birnbaum-Saunders con parámetros de forma y escala dados por $\alpha > 0$ y $\eta_k > 0$, respectivamente. Entonces se asume que T_k se obtiene de acuerdo al siguiente mecanismo:

$$T_k = \eta_k \varepsilon_k, \quad k = 1, \dots, n, \quad (2.2.6)$$

donde:

- $\varepsilon_1, \dots, \varepsilon_n$ son variables aleatorias independientes e idénticamente distribuidas tales que $\varepsilon_k \sim \text{BS}(\alpha, 1)$.
- Se asume que η_k está relacionada con un conjunto de variables explicativas en la siguiente forma:

$$\log(\eta_k) = \beta_1 x_{k1} + \dots + \beta_p x_{kp} = \mathbf{x}_k^\top \boldsymbol{\beta}, \quad k = 1, \dots, n, \quad (2.2.7)$$

donde:

- $\mathbf{x}_k = (x_{k1}, \dots, x_{kp})^\top$ es un vector con los valores que asumen p variables explicativas, continuas y/o discretas, cuando son medidas en el sujeto o unidad experimental k .
- $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ es un vector de parámetros desconocidos que deben ser estimados.

En la escala logarítmica el modelo se puede escribir en la siguiente forma:

$$Y_k = \mu_k + \epsilon_k, \quad k = 1, \dots, n, \quad (2.2.8)$$

donde

- $Y_k = \log(T_k)$
- $\mu_k = \log(\eta_k) = \mathbf{x}_k^\top \boldsymbol{\beta}$
- $\epsilon_1, \dots, \epsilon_n$ son variables aleatorias independientes e idénticamente distribuidas tales que $\log(\xi_k) = \epsilon_k$ y $\epsilon_k \sim \text{SH}(0, 2, \alpha)$

Siendo así, $E(\epsilon_k) = 0$, $\text{Var}(\epsilon_k) = w(\alpha)$, y la distribución de ϵ_k es simétrica alrededor de 0, lo cual implica que $E(Y_k) = \mu_k = \mathbf{x}_k^\top \boldsymbol{\beta}$, $\text{Var}(Y_k) = w(\alpha)$, y la distribución de Y_k es simétrica alrededor de $\mu_k = \mathbf{x}_k^\top \boldsymbol{\beta}$.

Para el modelo (2.2.7) se asume la presencia de censura(a la derecha) no informativa, es decir, cada sujeto tiene un tiempo de censura que es independiente del momento de falla. El tiempo de supervivencia observado es el mínimo entre el momento de la censura y de la

falla. Por tanto, se asume la presencia de variables aleatorias independientes C_1, C_2, \dots, C_n mutuamente independientes cuya función de densidad de probabilidad no dependen de β ni de α y se encuentran definidas como sigue:

definición 2.2.2.

Sean C_1, C_2, \dots, C_n variables aleatorias continuas e independientes que representan tiempos de censura. Las funciones de densidad de probabilidad de las C_k 's no dependen de β ni de α . Entonces, los datos observados denotados $(t_1, \delta_1), \dots, (t_n, \delta_n)$, se asumen como realizaciones de las variables aleatorias bivariadas $(T_1, \bar{\delta}_1), \dots, (T_n, \bar{\delta}_n)$, las cuales se generaron por:

$$T_k = \min(T_k, C_k) \quad y \quad \bar{\delta}_k = \mathbf{I}(T_k > C_k),$$

donde $\mathbf{I}(\cdot)$ es la función indicadora.

En consecuencia, el estimador de máxima verosimilitud de $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \alpha)^\top$, denotado $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}^\top, \hat{\alpha})^\top$, se puede escribir como $\hat{\boldsymbol{\theta}} = \operatorname{argmax} \ell(\boldsymbol{\theta})$, en que $\ell(\boldsymbol{\theta})$ se reduce a:

$$\ell(\boldsymbol{\theta}) = C + \sum_{k=1}^n (1 - \delta_k) \log(\xi_{k1}) - \frac{1}{2} \sum_{k=1}^n (1 - \delta_k) \xi_{k2}^2 + \sum_{k=1}^n \delta_k \log[1 - \Phi(\xi_{k2})], \quad (2.2.9)$$

donde:

- C es una constante.
- $\xi_{k1} = \frac{2}{\alpha} \cosh(u_k)$.
- $\xi_{k2} = \frac{2}{\alpha} \sinh(u_k)$.
- $u_k = (y_k - \mathbf{x}_k^\top \boldsymbol{\beta})/2$.
- δ_k es una variable indicadora que informa la presencia de censura a la derecha no informativa (vea Klein y Moeschberger (1997, capítulo 3))

2.3. Distribución Birnbaum-Saunders-t

La distribución de Birnbaum-Saunders- t con ν grados de libertad puede considerarse como una alternativa para el estudio de datos de tiempo de vida (fatiga) en lugar de la distribución de Birnbaum-Saunders ya que esta produce estimaciones de parámetros que son menos sensibles a los datos atípicos; ver Leiva (2008a), Barros (2008) y Paula (2012). La diferencia radica en que esta, usa la distribución t -student con ν grados de libertad para describir el daño acumulado, mientras que la distribución Birnbaum-Saunders usa la distribución normal para el mismo fin.

A diferencia de la distribución Birnbaum-Saunders, la distribución Birnbaum-Saunders- t describe el daño acumulado hasta el ciclo t , denotado como $W(t)$, usando la distribución t -student con ν grados de libertad.

La función de distribución acumulada de la variable aleatoria T , la cual describe el tiempo de vida del material es:

$$F_T(t) = \Phi_t \left[\frac{1}{\alpha} \left(\sqrt{\frac{t}{\eta}} - \sqrt{\frac{\eta}{t}} \right) \right], \quad t > 0, \quad (2.3.1)$$

siendo $\Phi_t(\cdot)$ la función de distribución acumulada t -student con ν grados de libertad.

La función de densidad de T , $f_T(t)$, se puede expresar como:

$$f_T(t) = \phi_t \left[\frac{1}{\alpha} \left(\sqrt{\frac{t}{\eta}} - \sqrt{\frac{\eta}{t}} \right) \right] \frac{1}{2\alpha\eta} \left(\sqrt{\frac{\eta}{t}} - \sqrt{\frac{\eta^3}{t^3}} \right), \quad t > 0, \quad (2.3.2)$$

siendo $\phi(\cdot)$ la función de densidad de la distribución t -Student con ν grados de libertad.

Entre las propiedades estadísticas más importantes de $T \sim BS-t(\alpha, \eta, \nu)$ con parámetros de forma y escala dados por $\alpha > 0$ y η_k , respectivamente se pueden citar las siguientes:

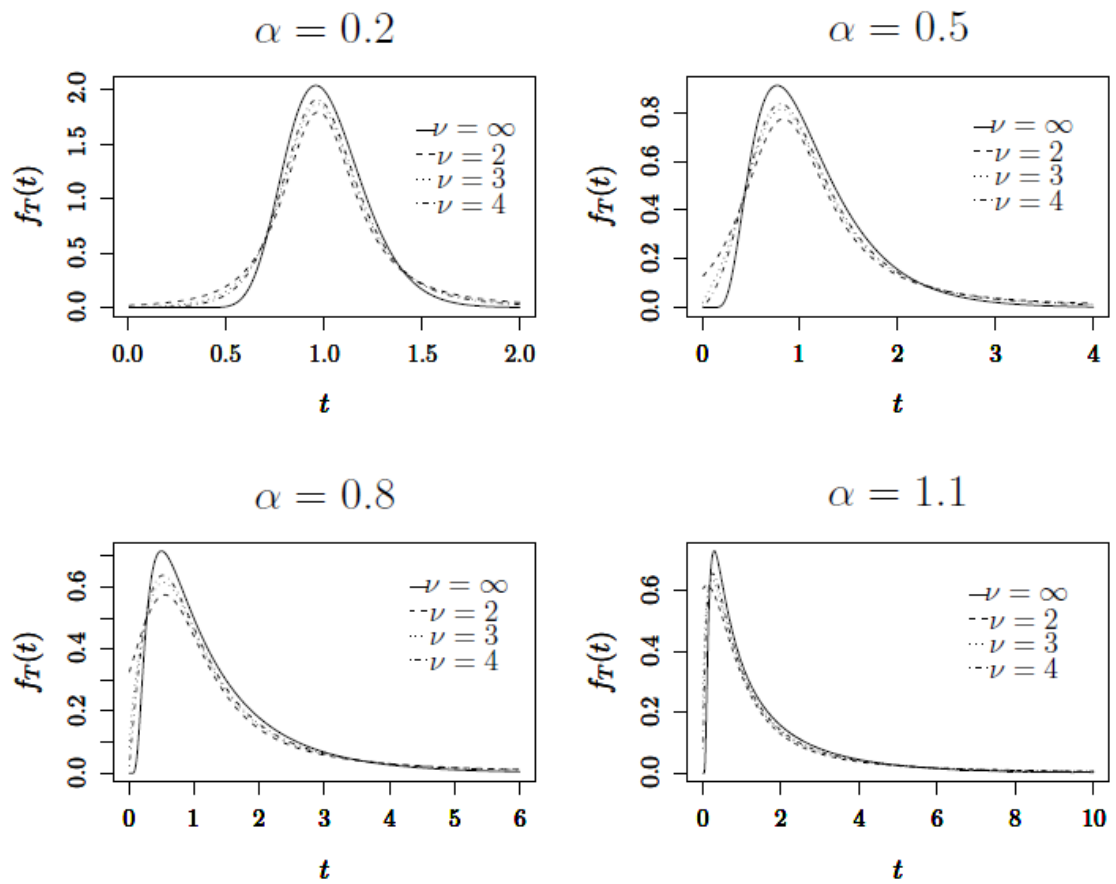
(C1) η es la mediana de T

(C2) $cT \sim BS-t(\alpha, c\eta, \nu)$ donde $c > 0$ es una constante

(C3) $T^{-1} \sim BS-t(\alpha, \eta^{-1}, \nu)$

(C4) $\frac{1}{\alpha} \left(\sqrt{\frac{t}{\eta}} - \sqrt{\frac{\eta}{t}} \right) \sim t\text{-student}(\nu)$

(C5) El percentil $100(q)\%$ de T es $t(q) = \eta \left[\alpha z_q + (\alpha^2 z_q^2 + 4)^{1/2} \right]^2 / 4$, es decir, $P[T \geq t(q)] = q$, siendo z_q el percentil $100(q)\%$ de la distribución t -Student con ν grados de libertad.

FIGURA 2.2. (2) Densidad Birnbaum-Saunders-t para diferentes valores de α y ν con $\eta = 1$

Formulación del modelo de regresión Log-Birnbaum-Saunders-t Censurado

definición 2.3.1.

Sean t_1, \dots, t_n las mediciones de una característica de interés realizadas en n sujetos, las cuales son realizaciones de n variables aleatorias independientes, denotadas T_1, \dots, T_n , que siguen distribución Birnbaum-Saunders-t teniendo parámetros de forma, escala, y grados de libertad dados por $\alpha > 0$, $\eta_k > 0$, y ν , respectivamente. Se asume que T_k se obtiene de acuerdo al siguiente mecanismo:

$$T_k = \eta_k \varepsilon_k, \quad k = 1, \dots, n, \quad (2.3.3)$$

donde:

- $\varepsilon_1, \dots, \varepsilon_k$ son variables aleatorias independientes e idénticamente distribuidas tales que $\varepsilon_k = \log(\xi_k)$ y $\varepsilon_k \sim \text{BS-t}(\alpha, 1, \nu)$.
- Se asume que η_k está relacionada con un conjunto de variables explicativas como sigue:

$$\log(\eta_k) = \beta_1 x_{k1} + \dots + \beta_p x_{kp} = \mathbf{x}_k^\top \boldsymbol{\beta}, \quad k = 1, \dots, n, \quad (2.3.4)$$

donde:

- $\mathbf{x}_k = (x_{k1}, \dots, x_{kp})^\top$ es un vector con los valores que asumen p variables explicativas, continuas y/o discretas, cuando son medidas en el sujeto o unidad experimental k .
- $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ es un vector de parámetros desconocidos que deben ser estimados.

En la escala logarítmica el modelo se puede escribir en la siguiente forma:

$$Y_k = \mu_k + \epsilon_k, \quad k = 1, \dots, n, \quad (2.3.5)$$

donde

- $Y_k = \log(T_k)$
- $\mu_k = \log(\eta_k) = \mathbf{x}_k^\top \boldsymbol{\beta}$
- $\epsilon_1, \dots, \epsilon_n$ son variables aleatorias independientes e idénticamente distribuidas tales que $\log(\xi_k) = \epsilon_k$ y $\epsilon_k \sim \text{SH-t}(0, 2, \alpha, \nu)$

Siendo así, $E(\epsilon_k) = 0$, $\text{Var}(\epsilon_k) = w(\alpha, \nu)$, y la distribución de ϵ_k es simétrica alrededor de 0, lo cual implica que $E(Y_k) = \mu_k = \mathbf{x}_k^\top \boldsymbol{\beta}$, $\text{Var}(Y_k) = w(\alpha, \nu)$, y la distribución de Y_k es simétrica alrededor de $\mu_k = \mathbf{x}_k^\top \boldsymbol{\beta}$.

Por tanto el estimador de máxima verosimilitud suponiendo censura no informativa como la definida en la sección (2.2) de $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \alpha)^\top$, denotado $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}^\top, \hat{\alpha})^\top$, se puede escribir como $\hat{\boldsymbol{\theta}} = \operatorname{argmax} \ell(\boldsymbol{\theta})$, en que $\ell(\boldsymbol{\theta})$ se reduce a:

$$\ell(\boldsymbol{\theta}) = C + \sum_{k=1}^n (1 - \delta_k) \log(\xi_{k1}) - \left[\frac{\nu + 1}{2} \right] \sum_{k=1}^n (1 - \delta_k) \log(\nu + \xi_{k2}^2) + \sum_{k=1}^n \delta_k \log[1 - \Phi_t(\xi_{k2})], \quad (2.3.6)$$

donde:

1. C es una constante.
2. $\xi_{k1} = \frac{2}{\alpha} \cosh(u_k)$.
3. $\xi_{k2} = \frac{2}{\alpha} \sinh(u_k)$.
4. $u_k = (y_k - \mathbf{x}_k^\top \boldsymbol{\beta})/2$
5. δ_k es una variable indicadora que informa la presencia de censura a la derecha no informativa (vea Klein y Moeschberger (1997, capítulo 3)).

Si en los modelos de regresión Log-Birnbaum-Saunders y Log-Birnbaum-Saunders-t censurados, la matriz Hessiana de los estimadores de máxima verosimilitud $\ell(\boldsymbol{\theta})$ es definida negativa para todo $\boldsymbol{\theta}$, entonces la función $\ell(\boldsymbol{\theta})$ es estrictamente cóncava y $\hat{\boldsymbol{\theta}}$ es la única solución al sistema de $p + 1$ ecuaciones no lineales dado por $\mathbf{U}(\hat{\boldsymbol{\theta}}) = \mathbf{0}$. el gradiente de $\ell(\boldsymbol{\theta})$ también conocido como la función de score, es dado por:

$$\mathbf{U}(\boldsymbol{\theta}) = \nabla \ell(\boldsymbol{\theta}) = \left(\frac{\partial \ell(\boldsymbol{\theta})}{\partial \beta_1}, \dots, \frac{\partial \ell(\boldsymbol{\theta})}{\partial \beta_p}, \frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha} \right)^\top. \quad (2.3.7)$$

En general, el sistema de ecuaciones $\mathbf{U}(\hat{\boldsymbol{\theta}}) = \mathbf{0}$ no tiene solución cerrada, por tanto, la estimación del vector de parámetros ($\boldsymbol{\theta}$) debe obtenerse mediante métodos numéricos, como, por ejemplo por el algoritmo de Newton-Raphson.

2.3.1. Algoritmo Newton-Raphson

La idea de este algoritmo es maximizar iterativamente aproximaciones cuadráticas de la función $\ell(\boldsymbol{\theta})$ para los modelos de regresión Log-Birnbaum-saunders y Log-birnbaum-Saunders-t presentdos en las secciones anteriores, hasta encontrar el valor de $\hat{\boldsymbol{\theta}}$. Para esto, se calcula una aproximación de $\ell(\boldsymbol{\theta})$ basada en series de Taylor de segundo orden en torno de, digamos, $\boldsymbol{\theta}^{[t]}$, obteniendo lo siguiente:

$$\ell(\boldsymbol{\theta}) \approx \ell^*(\boldsymbol{\theta}) = \ell(\boldsymbol{\theta}^{[t]}) + \mathbf{U}^\top(\boldsymbol{\theta}^{[t]})(\boldsymbol{\theta} - \boldsymbol{\theta}^{[t]}) - \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\theta}^{[t]})^\top \mathbf{J}(\boldsymbol{\theta}^{[t]})(\boldsymbol{\theta} - \boldsymbol{\theta}^{[t]}), \quad (2.3.8)$$

donde $\mathbf{J}(\boldsymbol{\theta}) = -\nabla^2 \ell(\boldsymbol{\theta})$ es la matriz de información observada de Fisher. Si $\mathbf{J}(\boldsymbol{\theta})$ es definida positiva para todo $\boldsymbol{\theta} \in \Omega$, entonces $\ell(\boldsymbol{\theta})$ es estrictamente cóncava, $\mathbf{J}(\boldsymbol{\theta}^{[t]})$ es definida positiva, $\mathbf{J}^{-1}(\boldsymbol{\theta}^{[t]})$ existe y $\ell^*(\boldsymbol{\theta})$ también es estrictamente cóncava, pues $\nabla^2 \ell^*(\boldsymbol{\theta}) = \nabla^2 \ell(\boldsymbol{\theta})$.

Por lo tanto, derivando $\ell^*(\boldsymbol{\theta})$ e igualando a 0 se obtiene que su único máximo global, denotado $\boldsymbol{\theta}^{[t+1]}$, satisface la siguiente ecuación:

$$0 = \mathbf{U}(\boldsymbol{\theta}^{[t]})\mathbf{J}(\boldsymbol{\theta}^{[t]})(\boldsymbol{\theta}^{[t+1]} - \boldsymbol{\theta}^{[t]}). \quad (2.3.9)$$

Siendo así, $\boldsymbol{\theta}^{[t+1]}$ puede expresarse como:

$$\boldsymbol{\theta}^{[t+1]} = \boldsymbol{\theta}^{[t]} + \mathbf{J}^{-1}(\boldsymbol{\theta}^{[t]})\mathbf{U}(\boldsymbol{\theta}^{[t]}). \quad (2.3.10)$$

Este resultado motiva el siguiente algoritmo:

Paso 0: Inicializar el contador en $t = 0$, fijar $\epsilon > 0$ y especificar un valor inicial, $\boldsymbol{\theta}^{[0]}$.

Paso 1: Calcular $\boldsymbol{\theta}^{[t+1]} = \boldsymbol{\theta}^{[t]} + \mathbf{J}^{-1}(\boldsymbol{\theta}^{[t]})\mathbf{U}(\boldsymbol{\theta}^{[t]})$

Paso 2: Calcular $\delta^{(t+1)} = \delta(\boldsymbol{\theta}^{[t]}, \boldsymbol{\theta}^{[t+1]})$

Paso 3: Actualizar el contador haciendo $t = t + 1$

Paso 4: Repetir los pasos 1, 2 y 3 hasta que $\delta^{(t)} < \epsilon$.

Paso 5: Hacer $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}^{[t]}$.

2.3.2. Inferencia Estadística

La inferencia para muestras grandes de $\boldsymbol{\theta}$ se puede obtener a partir de la normalidad asintótica del estimador $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \alpha)^\top$. Cuando $n \rightarrow \infty$, se tiene que:

$$\hat{\boldsymbol{\theta}} \sim N_{p+2}(\boldsymbol{\theta}, \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}), \quad (2.3.11)$$

donde $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}$ es la matriz asintótica de varianzas y covarianzas de tamaño $(p+1) \times (p+1)$ de $\hat{\boldsymbol{\theta}}$. Observe que $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}$ puede aproximarse por $-\ddot{\ell}^{-1}$, donde $-\ddot{\ell}$ es la matriz de información observada de Fisher de tamaño $(p+1) \times (p+1)$ que se obtiene de la matriz Hessiana de segundas derivadas dada por:

$$\ddot{\ell} = \begin{pmatrix} \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \alpha^2} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \alpha \partial \beta} \\ \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \beta \partial \alpha} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \beta \partial \beta^\top} \end{pmatrix} = \begin{pmatrix} \ddot{\ell}_{\alpha\alpha} & \ddot{\ell}_{\alpha\beta} \\ \ddot{\ell}_{\beta\alpha} & \ddot{\ell}_{\beta\beta} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n \frac{(1-3\xi_{i2}^2)}{\alpha^2} & k^\top X \\ X^\top k & X^\top V X \end{pmatrix}, \quad (2.3.12)$$

donde $\mathbf{V} = \text{diag}\{v_1(\boldsymbol{\theta}), \dots, v_n(\boldsymbol{\theta})\}$ y $k = (k_1, \dots, k_n)^\top$, con

$$v_i(\boldsymbol{\theta}) = \frac{1}{4} \left(\text{sech} \left(\frac{y_i - \mu_i}{2} \right) \right)^2 - \frac{1}{\alpha^2} \cosh(y_i - \mu_i) \quad (2.3.13)$$

$$k_i = -\frac{2}{\alpha^3} \sinh(y_i - \mu_i), \quad i = 1, 2, \dots, n. \quad (2.3.14)$$

Así, basado en el resultado dado en 2.3.11, una región de confianza aproximada del $(1 - \gamma) \times 100\%$ para el parámetro $\boldsymbol{\theta}$ se obtiene de:

$$R \equiv \left\{ \boldsymbol{\theta} \in \mathbb{R}^{p+1} : \left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}} \right)^\top \hat{\Sigma}^{-1} \left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}} \right) \leq \chi_{1-\rho}^2(p+1) \right\},$$

donde $\chi_{1-\rho}^2(p+1)$ denota el cuantil $(1 - \rho) \times 100$ -ésimo de la distribución central chicuadrado con $p+2$ grados de libertad y $\hat{\Sigma}^{-1}$ es una estimación de Σ la cuál, se puede obtener de la matriz de información observada basada en la ecuación 2.3.12.

CAPÍTULO 3

REGRESIÓN BIRNBAUM-SAUNDERS CENSURADO CON EFECTOS NO PARAMÉTRICOS

REGRESIÓN BIRNBAUM-SAUNDERS CENSURADO CON EFECTOS NO PARAMÉTRICOS

En este capítulo se propone una versión más flexible del modelo de regresión log-Birnbaum-Saunders y Log-Birnbaum-Saunders-t censurados expuestos en el capítulo 2. En principio, el modelo propuesto se expone únicamente bajo la distribución Birnbaum-Saunders pero sin pérdida de generalidad este modelo se puede proponer de la misma forma bajo la distribución Birnbaum-Saunders-t, en este caso, se asume que los efectos de algunas variables explicativas son no lineales pero sus formas funcionales son desconocidas. Se define un modelo de regresión que permita ajustar los datos de tiempo de vida o fatiga de un material usando funciones suaves pero arbitrarias que se pueden aproximar usando *splines* cúbicos naturales (Green y Silverman (1994)) ó P-splines cúbicos (Eilers y Marx(1996)). Estas funciones se penalizan con la intención de que el algoritmo devuelva el spline cúbico natural ideal y óptimo que permita una correcta aproximación de las funciones, por lo que el estimador de máxima verosimilitud para el modelo es penalizado, además, se hace una descripción detallada del estimador de parámetros que para este caso es el estimador de máxima verosimilitud penalizada y se propone un algoritmo para encontrar las estimaciones de la función de score por medio del algoritmo de Newton-Raphson.

Adicionalmente, la sección (3.3) presenta un estudio de simulación con el objetivo de evaluar la efectividad del procedimiento de estimación para recuperar el verdadero valor de μ tanto en modelos Log-Birnbaum-Saunders y Log-Birnbaum-Saunders-t. En la sección (3.4) estudia un conjunto de datos de bajo ciclo presentados por Nelson (1984) para la tensión de 26 especímenes a base de níquel, se analiza estadísticamente y se comparan los 4 modelos estudiados con el fin de verificar la ganancia estadística que se obtiene al usar tanto los modelos paramétricos basados en las distribuciones Birnbaum-Saunders y la distribución Birnbaum-Saunders-t y el modelo propuesto con las distribuciones Birnbaum-Saunders y la distribución Birnbaum-Saunders-t.

3.1. Formulación del Modelo de regresión Log-Birnbaum-Saunders con efectos no paramétricos y censura no informativa

definición 3.1.1.

Sean t_1, \dots, t_n mediciones de una característica de interés realizadas en n sujetos o unidades experimentales, las cuales se asume que son realizaciones de n variables aleatorias, independientes, continuas y estrictamente positivas, denotadas T_1, T_2, \dots, T_n , que representan el tiempo de vida T de las n realizaciones y siguen distribución Birnbaum-Saunders con función de densidad de probabilidad (2.1.3), teniendo parámetros de forma y escala dados por $\alpha > 0$ y η_k , respectivamente.

Se asume que T_k se obtiene de acuerdo al siguiente mecanismo:

$$T_k = \eta_k \xi_k, \quad k = 1, \dots, n, \quad (3.1.1)$$

donde $\xi_1, \xi_2, \dots, \xi_k$ son un conjunto de errores multiplicativos independientes e idénticamente distribuidos tales que $\xi_k \sim BS(\alpha, 1)$ y cuya función de densidad se puede escribir como:

$$f(\xi) = \frac{1}{2\xi\sqrt{2\pi}} \left[\sqrt{\xi} + \frac{1}{\sqrt{\xi}} \exp \left\{ -\frac{1}{2\alpha^2} \left(\frac{1}{\sqrt{\xi}} - \sqrt{\xi} \right)^2 \right\} \right], \quad \xi > 0, \quad (3.1.2)$$

siendo $\alpha > 0$ un parámetro estrictamente positivo. Además, se asume que η_k está relacionada con un conjunto de variables explicativas en la siguiente forma:

$$\log(\eta_k) = \mathbf{x}_k^\top \boldsymbol{\beta} + g_1(s_{k1}) + \dots + g_q(s_{kq}), \quad k = 1, \dots, n, \quad (3.1.3)$$

donde:

- $\mathbf{x}_k = (x_{k1}, \dots, x_{kp})^\top$ es un vector con los valores que asumen p variables explicativas, continuas y/o discretas, cuando son medidas en el sujeto o unidad experimental k ,
- $(s_{k1}, \dots, s_{kq})^\top$ son q variables explicativas continuas asociadas al sujeto o unidad experimental k
- $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ es un vector de parámetros desconocidos que deben ser estimados.
- $g_1(\cdot), \dots, g_q(\cdot)$ son funciones suaves pero arbitrarias que se pueden aproximar usando *splines* cúbicos naturales (Green y Silverman (1994)) ó P-splines cúbicos (Eilers y Marx(1996)).

Las funciones $g_1(\cdot), \dots, g_q(\cdot)$ permiten aproximar la forma funcional de las variables explicativas de las cuales se sospechan que la relación con la variable de respuesta es no lineal. Por lo tanto, para $r = 1, \dots, q$, el vector $(g_r(s_{1r}), \dots, g_r(s_{nr}))^\top$ se puede escribir como $\mathbf{N}_r \boldsymbol{\tau}_r$, donde \mathbf{N}_r es una matriz conocida de dimensión $n \times p_r$, y $\boldsymbol{\tau}_r$ es un vector de parámetros desconocidos de dimensión p_r que deben ser estimados.

Por ejemplo, para aproximar la función $g_r(\cdot)$ usando *splines* cúbicos naturales se consideran los p_r valores diferentes en el conjunto $s_{1r}, s_{2r}, \dots, s_{nr}$, donde s_{kr} es el valor que toma el k -ésimo individuo en la covariable s_r , los cuales se denotan como: $s_{r(1)} < s_{r(2)} < \dots < s_{r(p_r)}$.

El elemento (ij) de la matriz \mathbf{N}_r es 1 si $s_{ir} = s_{r(j)}$ y 0 en otros casos. Siendo así, $\mu_k = \log(\eta_k)$ se puede reescribir en la siguiente forma

$$\mu_k = \mathbf{x}_k^\top \boldsymbol{\beta} + \mathbf{m}_{k1}^\top \boldsymbol{\tau}_1 + \dots + \mathbf{m}_{kq}^\top \boldsymbol{\tau}_q. \quad (3.1.4)$$

donde \mathbf{m}_{kr}^\top es la k -ésima fila de \mathbf{N}_r .

En la escala logarítmica el modelo se puede escribir en la siguiente forma:

$$Y_k = \boldsymbol{\mu}_k + \epsilon_k, \quad k = 1, \dots, n, \quad (3.1.5)$$

donde:

- $Y_k = \log(T_k)$,
- $\boldsymbol{\mu}_k = \log(\eta_k) = \mathbf{X}\boldsymbol{\beta} + \mathbf{N}_1\boldsymbol{\tau}_1 + \mathbf{N}_2\boldsymbol{\tau}_2 + \dots + \mathbf{N}_q\boldsymbol{\tau}_q$,
- $\epsilon_1, \dots, \epsilon_n$ son variables aleatorias independientes e idénticamente distribuidas tales que $\epsilon_k \sim \text{SH}(0, 2, \alpha)$.

Siendo así, $E(\epsilon_k) = 0$, $\text{Var}(\epsilon_k) = w(\alpha)$, y la distribución de ϵ_k es simétrica alrededor de 0, lo cual implica que $E(Y_k) = \boldsymbol{\mu}_k$, $\text{Var}(Y_k) = w(\alpha)$, y la distribución de Y_k es simétrica alrededor de $\boldsymbol{\mu}_k$.

3.2. Estimación de parámetros del modelo de regresión Log-Birnbaum-Saunders con efectos no paramétricos y censura no informativa

Si $\boldsymbol{\tau} = (\boldsymbol{\tau}_1^\top, \dots, \boldsymbol{\tau}_q^\top)^\top$ es el vector de parámetros desconocidos tal que $(g_r(s_{1r}), \dots, g_r(s_{nr}))^\top$ se puede escribir como $\mathbf{N}_r \boldsymbol{\tau}_r$, el estimador de máxima verosimilitud penalizada del vector de parámetros $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\tau}^\top, \alpha)^\top$, denotado $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}^\top, \hat{\boldsymbol{\tau}}^\top, \hat{\alpha})^\top$, se puede escribir como $\hat{\boldsymbol{\theta}} = \operatorname{argmax} PL(\boldsymbol{\theta})$, en que:

$$PL(\boldsymbol{\theta}) = L(\boldsymbol{\theta}) + P(\boldsymbol{\theta})$$

donde

$$L(\boldsymbol{\theta}) = C + \sum_{k=1}^n (1 - \delta_k) \log(\xi_{k1}) - \frac{1}{2} \sum_{k=1}^n (1 - \delta_k) \xi_{k2}^2 + \sum_{k=1}^n \delta_k \log[1 - \Phi(\xi_{k2})],$$

y

$$P(\boldsymbol{\theta}) = -\frac{1}{2} \sum_{r=1}^q \lambda_r \boldsymbol{\tau}_r^\top \mathbf{M}_r \boldsymbol{\tau}_r,$$

donde:

- C es una constante.
- $y_k = \log(T_k)$.
- $\xi_{k1} = \frac{2}{\alpha} \cosh(u_k)$.
- $\xi_{k2} = \frac{2}{\alpha} \sinh(u_k)$.
- $u_k = (y_k - \boldsymbol{\mu}_k^\top) / 2$
- δ_k es una variable indicadora que informa la presencia de censura a la derecha no informativa (vea Klein y Moeschberger (1997, capítulo 3)).
- $\mathbf{M}_1, \dots, \mathbf{M}_q$ son matrices de penalización conocidas.
- $\lambda_1 > 0, \dots, \lambda_q > 0$ son parámetros de suavización.

Si la matriz Hessiana de $PL(\boldsymbol{\theta})$ es definida negativa para todo $\boldsymbol{\theta}$, entonces la función $PL(\boldsymbol{\theta})$ es estrictamente cóncava, y $\hat{\boldsymbol{\theta}}$ es la única solución al sistema de $p + 1 + p_1 + \dots + p_q$ ecuaciones no lineales dado por $\mathbf{U}(\hat{\boldsymbol{\theta}}) = 0$, si ella existe, en que $\mathbf{U}(\boldsymbol{\theta})$ es el gradiente de $PL(\boldsymbol{\theta})$, también conocido como la función de score, es decir, $\mathbf{U}(\boldsymbol{\theta}) = \nabla PL(\boldsymbol{\theta})$.

El sistema de ecuaciones $\mathbf{U}(\hat{\boldsymbol{\theta}}) = 0$ no tiene solución cerrada, por lo que estimación del vector de parámetros $\boldsymbol{\theta}$ se obtiene mediante métodos numéricos como el algoritmo Newton-Raphson.

$$\mathbf{U}(\boldsymbol{\theta}) = \begin{bmatrix} \frac{1}{2} \overline{\mathbf{X}}^\top \mathbf{S} \\ -\frac{n}{\alpha} + \frac{1}{\alpha} \sum_{i=1}^n \xi_{i2}^2 \end{bmatrix} - \overline{\mathbf{M}}\boldsymbol{\theta}, \quad (3.2.1)$$

en que:

- $\overline{\mathbf{X}} = [\mathbf{X}, \mathbf{N}_1, \dots, \mathbf{N}_q]$,
- $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_n)^\top$,
- $\mathbf{S} = (S_1, \dots, S_n)^\top$, $S_k = \xi_{i1}\xi_{i2} - \xi_{i2}/\xi_{i1}$,
- $\overline{\mathbf{M}} = \{\mathbf{0}, \lambda_1 \mathbf{M}_1, \dots, \lambda_q \mathbf{M}_q, \mathbf{0}\}$.

3.2.1. Algoritmo Newton-Raphson

Se calcula una aproximación de $PL(\boldsymbol{\theta})$ basada en series de Taylor de segundo orden en torno de, digamos, $\boldsymbol{\theta}^{[l]}$, obteniendo lo siguiente:

$$PL(\boldsymbol{\theta}) \approx PL^*(\boldsymbol{\theta}) = PL(\boldsymbol{\theta}^{[l]}) + \mathbf{U}^\top(\boldsymbol{\theta}^{[l]})(\boldsymbol{\theta} - \boldsymbol{\theta}^{[l]}) - \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\theta}^{[l]})^\top \mathbf{J}(\boldsymbol{\theta}^{[l]})(\boldsymbol{\theta} - \boldsymbol{\theta}^{[l]}),$$

donde:

- $\mathbf{J}(\boldsymbol{\theta}) = -\nabla^2 PL(\boldsymbol{\theta})$ es la matriz de información observada de Fisher.
- Si $\mathbf{J}(\boldsymbol{\theta})$ es definida positiva para todo $\boldsymbol{\theta}$, entonces $PL(\boldsymbol{\theta})$ es estrictamente cóncava, $\mathbf{J}(\boldsymbol{\theta}^{[l]})$ es definida positiva, $\mathbf{J}^{-1}(\boldsymbol{\theta}^{[l]})$ existe y $PL^*(\boldsymbol{\theta})$ también es estrictamente cóncava, pues $\nabla^2 PL^*(\boldsymbol{\theta}) = \nabla^2 PL(\boldsymbol{\theta})$.

Por lo tanto, derivando $PL^*(\boldsymbol{\theta})$ e igualando a $\mathbf{0}$ se obtiene que su único máximo global, denotado $\boldsymbol{\theta}^{[l+1]}$, satisface la siguiente ecuación:

$$\mathbf{0} = \mathbf{U}(\boldsymbol{\theta}^{[l]}) - \mathbf{J}(\boldsymbol{\theta}^{[l]})(\boldsymbol{\theta}^{[l+1]} - \boldsymbol{\theta}^{[l]}).$$

Siendo así, $\boldsymbol{\theta}^{[l+1]}$ puede expresarse como

$$\boldsymbol{\theta}^{[l+1]} = \boldsymbol{\theta}^{[l]} + \mathbf{J}^{-1}(\boldsymbol{\theta}^{[l]})\mathbf{U}(\boldsymbol{\theta}^{[l]}).$$

Este resultado motiva el siguiente algoritmo:

Paso 0: Inicializar el contador en $l = 0$, fijar $\varepsilon > 0$ y especificar un valor inicial, $\boldsymbol{\theta}^{[0]}$.

Paso 1: Calcular $\boldsymbol{\theta}^{[l+1]} = \boldsymbol{\theta}^{[l]} + \mathbf{J}^{-1}(\boldsymbol{\theta}^{[l]})\mathbf{U}(\boldsymbol{\theta}^{[l]})$

Paso 2: Calcular $\kappa^{(l+1)} = \kappa(\boldsymbol{\theta}^{[l]}, \boldsymbol{\theta}^{[l+1]})$

Paso 3: Actualizar el contador haciendo $l = l + 1$

Paso 4: Repetir los pasos 1, 2 y 3 hasta que $\kappa^{(l)} < \varepsilon$.

Paso 5: Hacer $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}^{[l]}$.

3.2.2. Selección de los parámetros de suavización

El parámetro de suavización $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_q)^\top$ permite suavizar el spline cúbico natural con el que se aproximan las funciones $g(s_k)$. El valor de $\boldsymbol{\lambda}$ se puede seleccionar usando el siguiente criterio (ver Green & Silverman (1994))

$$\hat{\boldsymbol{\lambda}} = \operatorname{argmin} \operatorname{AIC}(\boldsymbol{\theta}),$$

donde

$$\operatorname{AIC}(\boldsymbol{\theta}) = -2L(\hat{\boldsymbol{\theta}}) + 2(p + \operatorname{gl}_1 + \dots + \operatorname{gl}_r), \quad (3.2.2)$$

en que $\operatorname{gl}_r = \operatorname{traza}[(\mathbf{N}_r^\top \mathbf{N}_r + \lambda_r \mathbf{M}_r)^{-1} (\mathbf{N}_r^\top \mathbf{N}_r)]$ es el número de grados de libertad gastados al aproximar $g_r(\cdot)$.

3.2.3. Inferencia estadística

1. Para grandes tamaños de la muestra $\hat{\boldsymbol{\theta}} \sim N(\boldsymbol{\theta}, \mathbf{J}^{-1}(\boldsymbol{\theta}))$, es decir, para n suficientemente grande $\hat{\boldsymbol{\theta}}$ es un estimador aproximadamente insesgado de $\boldsymbol{\theta}$, cuya matriz de varianzas y covarianzas aproximada es $\operatorname{Var}(\hat{\boldsymbol{\theta}}) = \mathbf{J}^{-1}(\boldsymbol{\theta})$.
2. La matriz de varianzas y covarianzas de $\hat{\boldsymbol{\theta}}$ para muestras grandes se puede estimar consistentemente usando la matriz $\mathbf{J}^{-1}(\hat{\boldsymbol{\theta}})$.
3. la regla de decisión del test de Wald para juzgar el sistema de hipótesis dado por $H_0 : \theta_j = 0$ versus $H_1 : \theta_j \neq 0$ es “Rechazar H_0 al nivel de significancia aproximado de $100(\gamma)\%$ si $|Z| > Z_{1-\frac{\gamma}{2}}$ o $Z^2 > \chi_{1-\gamma}^2(1)$ ”, donde $Z = \hat{\theta}_j / \operatorname{ee}(\hat{\theta}_j)$, en que $\operatorname{ee}(\hat{\theta}_j)$ es la raíz cuadrada del j -ésimo elemento de la diagonal principal de la matriz $\mathbf{J}^{-1}(\hat{\boldsymbol{\theta}})$.

3.3. Estudio de simulación

En este caso, se describe un estudio de simulación con el objetivo de evaluar la efectividad del procedimiento de estimación para recuperar el verdadero valor de μ tanto en modelos log-Birnbaum-Saunders y log-Birnbaum-Saunders-t en la presencia de censura a la derecha no informativa ante los efectos no paramétricos de una variable explicativa como sigue:

- (1) se generaron r muestras independientes e idénticamente distribuidas de tamaño $n = 1000$, cada una de las cuales corresponde a una realización del modelo log-Birnbaum-Saunders que considera en este caso una variable explicativa no lineal s , donde el tiempo de vida del individuo k , $k = 1, \dots, 1000$ puede ser descrito como:

$$Y_k = \beta_1 + \beta_2 x_k + g(s_k) + \epsilon_k,$$

- (2) Los tiempos de censura se generaron independientes, siguiendo una distribución exponencial con media γ . Los valores de los parámetros para el experimento fueron $\beta_1 = \frac{1}{2}$, $\beta_2 = 4$, $g(s_k) = \frac{5s_k}{0.3 + s_k} - 2$.
- (3) Los valores de s_k se generaron a partir de la construcción de la matriz de base \mathbf{N} y la matriz de penalización haciendo uso de la función `ncs` del paquete `SSYM` (Vanegas (2016)) disponible en el software R-project. La función `ncs` permite aproximar una función suave utilizando una spline cúbica natural.

Para el parámetro de forma α y la media de la censura γ :

- (a) se tomaron varios valores con el fin de analizar el comportamiento de la estimación $g(s_k)$ dentro del modelo, los valores para el parámetro de forma fueron $(\alpha = 0.5; 1; 1.5; 2)$ y para la censura $\gamma = (0.00003; 0.0003; 0.003)$
- (b) La variable explicativa, x , se generó siguiendo una distribución uniforme en el intervalo $(0, 1)$
- (c) El parámetro de suavización λ se estimó con la función la sección (3.2.2).
- (d) En cada una de las $r = 10000$ réplicas del método los parámetros fueron estimados para el modelo log-Birnbaum-Saunders y log-Birnbaum-Saunders-t Censurado usando el método de máxima verosimilitud penalizada.

3.3.1. Resultados

Las gráficas (3.1), (3.2), (3.3), (3.4), (3.5), (3.6), (3.7) y (3.8) representan las estimaciones de $g(s)$ bajo los modelos Log-BS censurado y Log-BS- t censurado ambos con efectos no paramétricos, la función de color negro representa el verdadero valor de la curva de regresión mientras que las bandas de confianza en las gráficas, la de color rojo (Log-BS) y las de color azul (Log-BS-t) representan los cuantiles 2.5% y 97.5% de las estimaciones de los parámetros por el método de máxima verosimilitud penalizada.

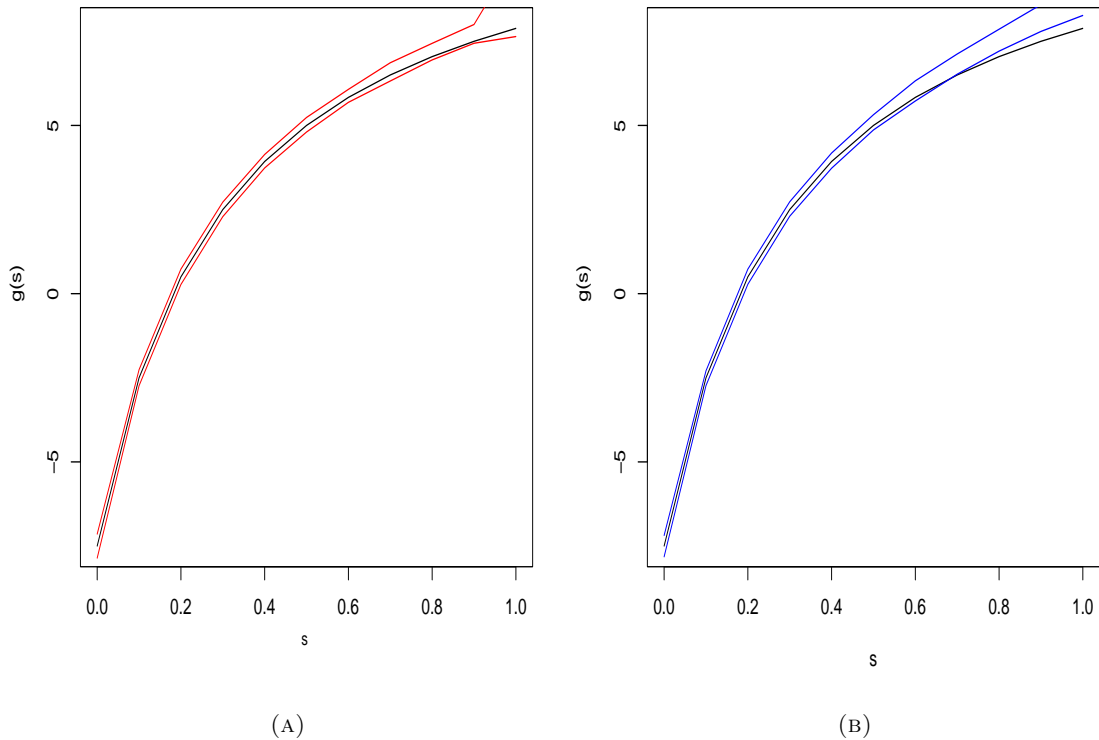


FIGURA 3.1. Curva de estimación simulada del modelo Log-BS censurado con efectos no paramétricos (A) y Log-BS-t censurado con efectos no paramétricos (B):($\alpha = 0.5, \gamma = 0.003$)

Las gráficas (A) y (B) de la figura (3.1) representan estimaciones del modelo de regresión Log-BS y Log-BS-t ambos no paramétricos y con censura simulados en este caso. Las gráficas (A) y (B) representan estimaciones con $\alpha = 0.5$ y $\tau = 0.003$ (esto equivale al 4.1 % de datos censurados), note que en el gráfico (A) que contiene el verdadero valor de la curva se encuentra aproximadamente dentro de las bandas estimadas (las bandas representan los percentiles 2.5 % y 97.5 % de todas las R curvas estimadas en la simulación), lo que significa que la estimación recupera el verdadero valor de la curva. El gráfico (B) indica que si se aumenta el valor de α con el mismo porcentaje de censura, recuperar el verdadero valor de la curva es más complejo de estimar ya que hay una parte de la estimación que no la está recuperando.

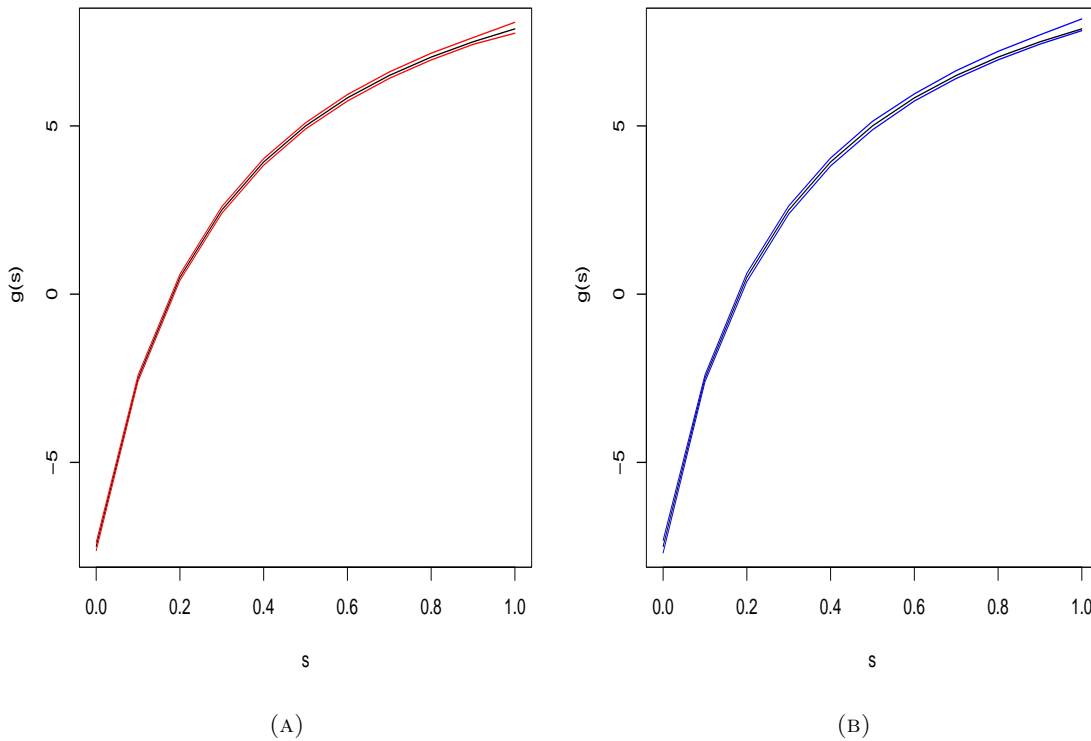


FIGURA 3.2. Curva de estimación simulada del modelo Log-BS censurado con efectos no paramétricos (A) y Log-BS-t censurado con efectos no paramétricos (B) ($\alpha = 0.5, \gamma = 0.00003$)

Las gráficas (A) y (B) de la figura (3.2) representan estimaciones del modelo de regresión Log-BS y Log-BS-t ambos no paramétricos y con censura simulados en este caso. Las gráficas (A) y (B) representan estimaciones con $\alpha = 0.5$ y $\tau = 0.00003$ (esto equivale al 3.4 % de datos censurados), note que en el gráfico (A) que en este caso, ambas simulaciones están recuperando el verdadero valor de la curva, es decir, el verdadero valor se encuentra aproximadamente dentro de las bandas de confianza estimadas (las bandas representan los percentiles 2.5 % y 97.5 % de todas las R curvas estimadas en la simulación), lo que significa que en este caso ambos modelos están recuperando el verdadero valor de la curva. Tenga en cuenta que en este caso, el parámetro de escala es el mismo pero se hizo una disminución de la censura y permitió una mejor aproximación, esto indica que a menor censura, mejor precisión de las estimaciones.

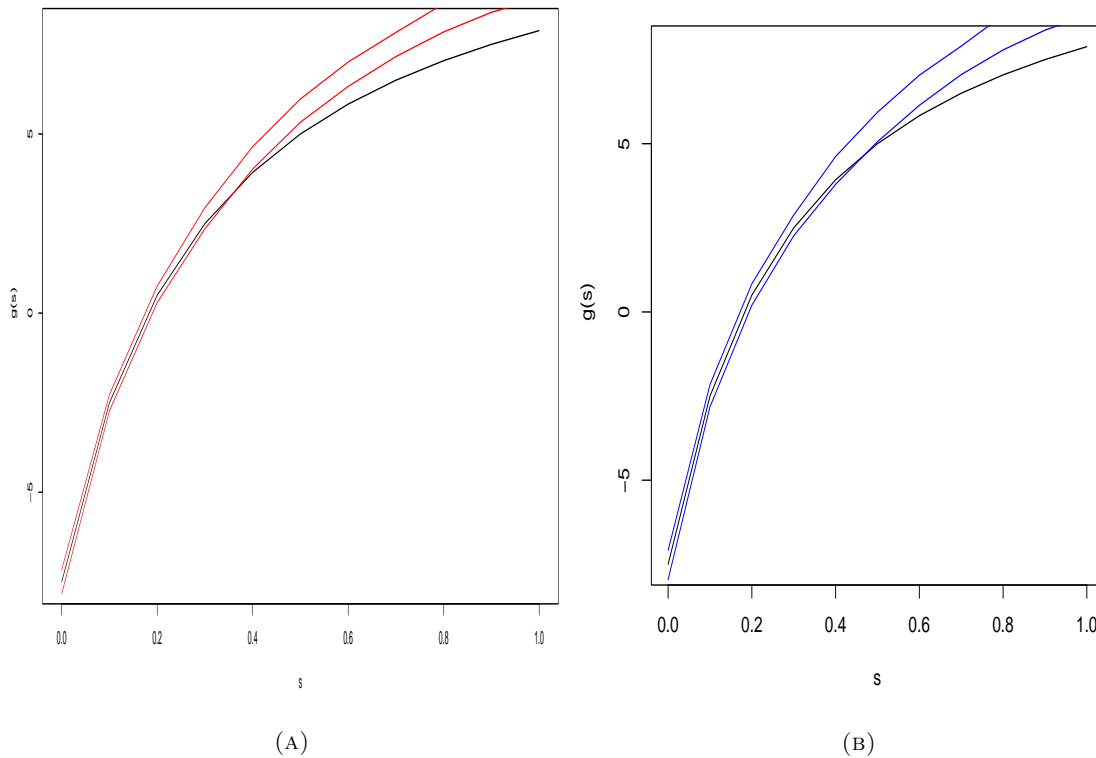


FIGURA 3.3. Curva de estimación simulada del modelo Log-BS censurado con efectos no paramétricos (a) y Log-BS-t censurado con efectos no paramétricos (b) ($\alpha = 1.5, \gamma = 0.003$)

La figura (3.3) muestra las estimaciones de la función para valores altos del parámetro de escala y censura del 43 %, en ambos casos (tanto para el modelo (A) como para el (B)) el método no está recuperando por completo el verdadero valor de la curva, cabe recalcar que el porcentaje de censura y los valores de $\alpha = 1.5$ son altos lo que dificulta la convergencia adecuada del modelo, para mejorar la estimación se debería tomar un valor más alto de n y aumentar las réplicas, de esta forma se obtendría una mejor estimación del modelo.

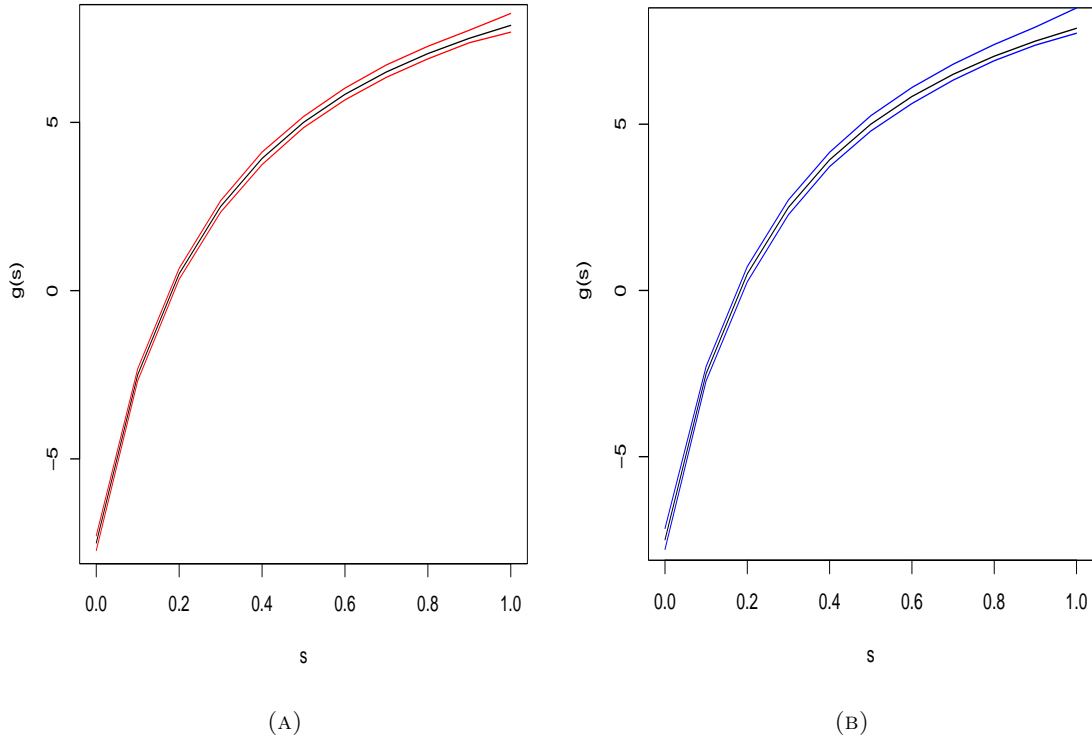


FIGURA 3.4. Curva de estimación simulada del modelo Log-BS censurado con efectos no paramétricos (A) y Log-BS-t censurado con efectos no paramétricos ($\alpha = 1.5, \tau = 0.0003$)

Note que al igual que en la figura (3.3) el valor del parámetro de escala es $\alpha = 1.5$ en las gráficas (A) y (B) respectivamente pero en este caso la censura disminuyó de un 42 % a un 30 %, lo que significa que para valores grandes de α entre menos censura se tenga, el método se aproxima al verdadero valor de la curva y para estas simulaciones, el modelo Log-BS-Censurado con efectos no paramétricos (A) tiene mejor estimación, es decir, las R curvas estimadas en el modelo devuelve aproximadamente el verdadero valor de la curva.

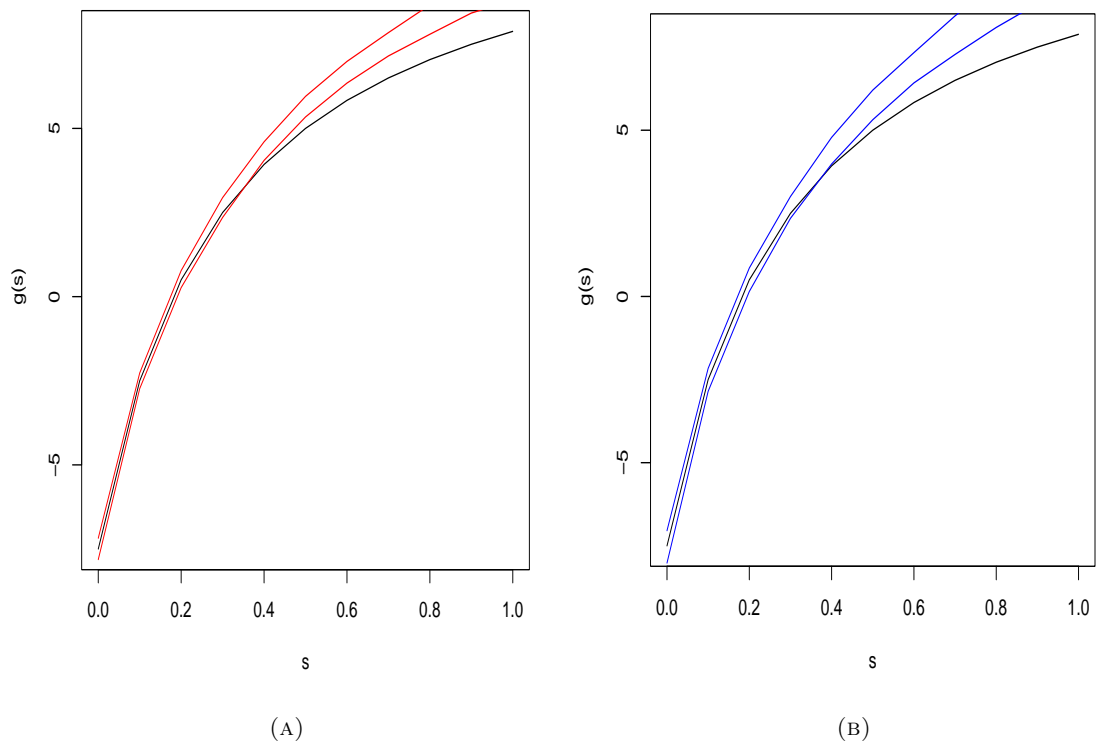


FIGURA 3.5. Curva de estimación simulada del modelo Log-BS censurado con efectos no paramétricos (A) y Log-BS-t censurado con efectos no paramétricos ($\alpha = 2.0, \tau = 0.003$)

En las las gráficas (A) y (B) de la figura (3.5) el valor del parámetro de escala en este caso es $\alpha = 2.0$ pero en este caso la censura es de un 50 %, note que con respecto a la figura (3.4) se aumentó la censura y el parámetro de escala, los resultados muestran que las estimaciones están muy cerca al verdadero valor de la curva, sin embargo, la curva verdadera no está contenida totalmente en las bandas, observe que en la gráfica (B) el verdadero valor de la curva está sobresaliendo más que en la gráfica (A) lo que significa que a pesar de que las estimaciones no son significativamente eficientes, las r curvas de (A) producen bandas de estimaciones con mejor aproximación. Se espera que si se aumenta el n y las repeticiones, las estimaciones mejoren y contengan al verdadero valor de la curva.

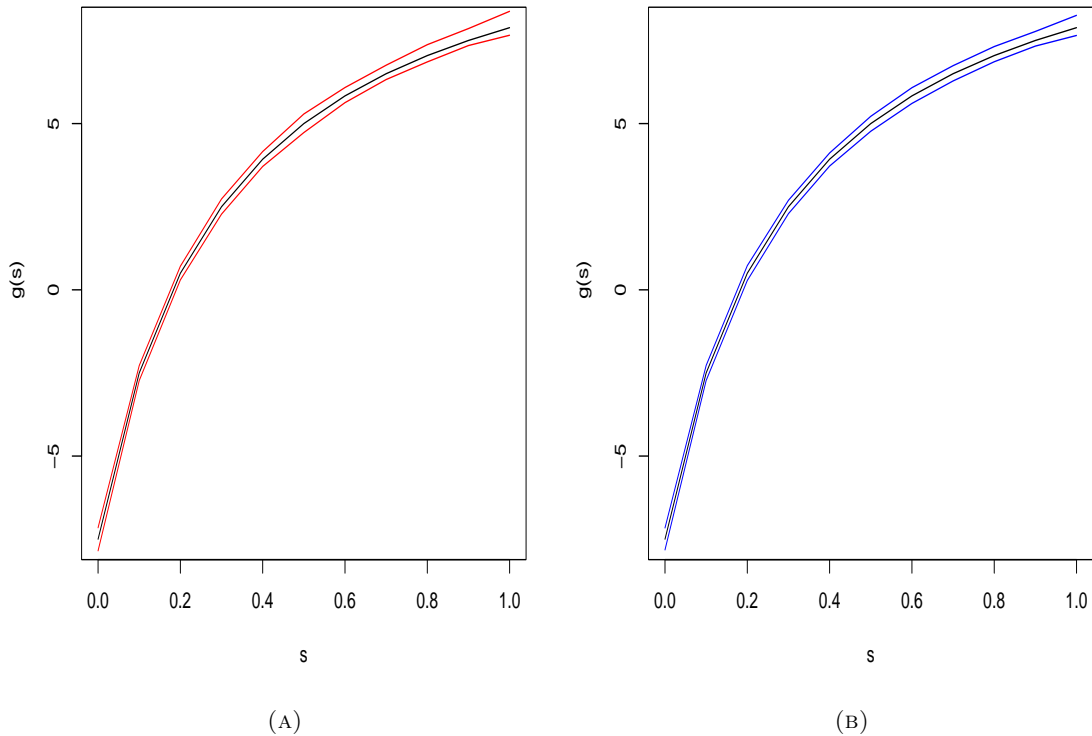


FIGURA 3.6. Curva de estimación simulada del modelo ((A) : $\alpha = 1.0, \tau = 0.0003$); ((B) : $\alpha = 1.5, \tau = 0.0003$)

Como se indicó en la figura (3.5) para un parámetro de escala alto se recomienda disminuir la censura ya que con ambos valores significativos el modelo tiene problemas para estimar el verdadero valor de la curva a pesar de que las repeticiones y réplicas se aumenten. En este caso el parámetro de escala tomó un valor de $\alpha = 2.0$ pero la censura se disminuyó de un 30 % a un 5.6 %, tanto (A) como (B) muestran aproximaciones estadísticamente significativas, ambos modelos brindan aproximaciones igual de buenas, tanto como las estimaciones para un parámetro de escala $\alpha = 0.5$.

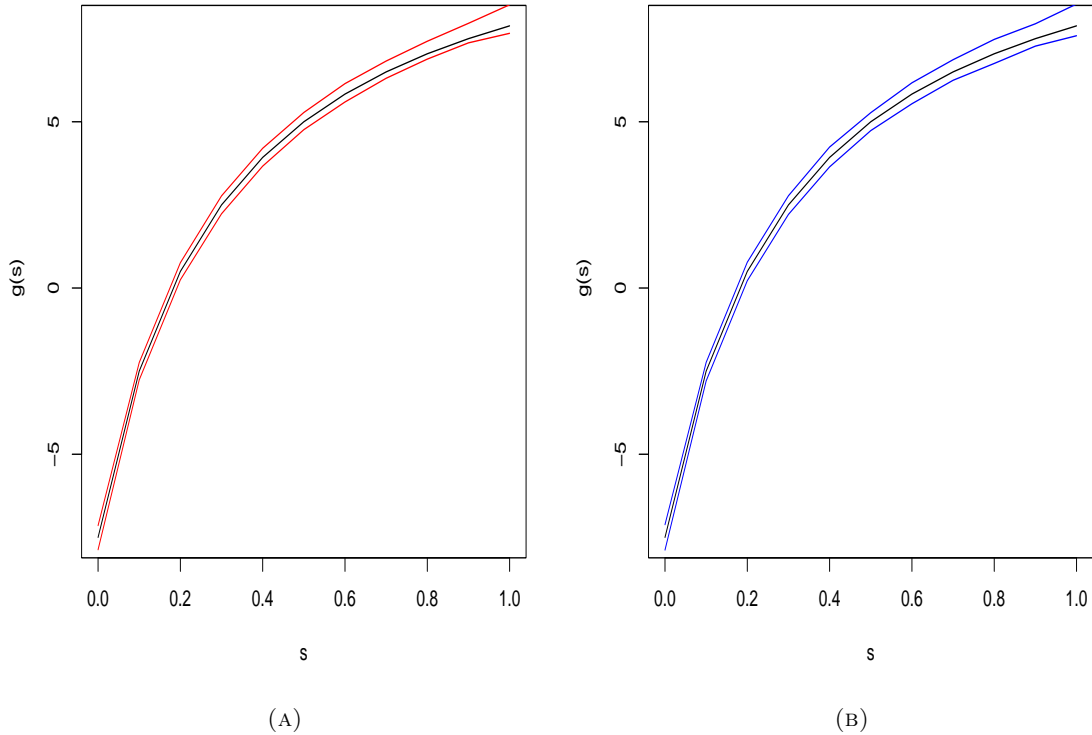


FIGURA 3.7. Curva de estimación simulada del modelo Log-BS censurado con efectos no paramétricos (A) y Log-BS-t censurado con efectos no paramétricos ($\alpha = 2.0, \tau = 0.00003$)

La figura (3.7) muestra las estimaciones de las r curvas con parámetro de escala $\alpha = 1.0$ y censura del 3.1 %, la gráfica (B) permite observar una mejor estimación del verdadero valor de la curva con respecto a la gráfica (A), observe que las bandas de confianza (estimadas) de la gráfica (B) contiene a la función de color negro (la verdadera función) con mayor confiabilidad que (A), lo que indica en este caso que el modelo Log-BS-t censurado con efectos no paramétricos tiene mejor estimación que el Log-BS.

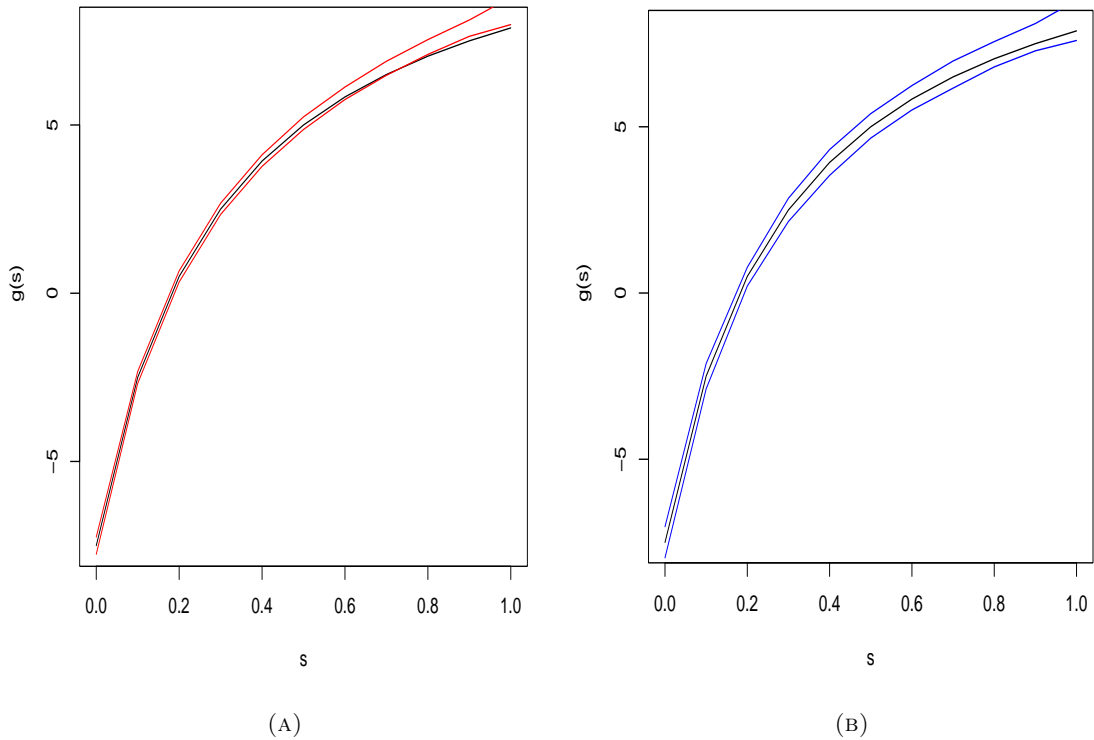


FIGURA 3.8. Curva de estimación simulada del modelo Log-BS censurado con efectos no paramétricos (A) y Log-BS-t censurado con efectos no paramétricos ($\alpha = 1.0, \tau = 0.003$)

En este caso, se hizo un aumento de la censura con respecto a la figura (3.7), pasó de 3.1% a 5.6% con el mismo parámetro de escala $\alpha = 1.0$, se puede observar que con el aumento de censura, en la gráfica (A) la curva se sale de las bandas de confianza estimadas, es decir, el modelo no está devolviendo el verdadero valor de la curva, mientras que en la gráfica (B), a pesar de que la estimación no es realmente buena, la curva se encuentra en el límite de las bandas de confianza, esto es, las bandas de confianza encierran el verdadero valor de la curva.

3.4. Aplicación

- (1) Se presenta un conjunto de datos de fatiga de bajo ciclo para una prueba de tensión controlada en 26 especímenes cilíndricos a base de níquel. Los datos fueron originalmente analizados por Nelson(1984), también por Meeker & Escobar (1998) y por Cédric Heuchenne & Ingrid Van Keilegom (2007)
- (2) La figura (3.9) muestra el logaritmo del número de ciclos antes de la falla contra el Pseudo-estrés (El módulo de Young multiplicado por la tensión)
- (3) Se observaron cuatro datos censurados: una observación se censura si ocurre una falla en el radio, soldadura o hilos (la censura proviene de impurezas o vacíos) o si no ocurre ninguna falla. Se puede observar que el Pseudo-estrés tiene un efecto no lineal sobre el logaritmo de los ciclos.

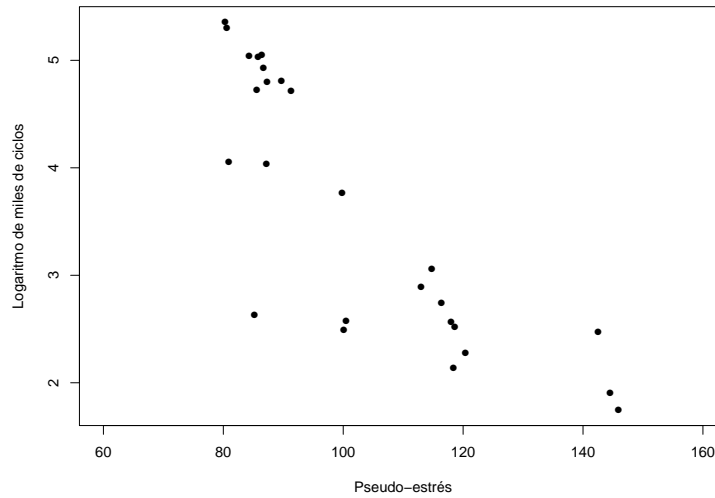


FIGURA 3.9. Logaritmo del número de ciclos antes de la falla contra el Pseudoestress.

El modelo a estimar está dado por:

$$Y_k = \mu_k + \epsilon_k, \quad k = 1, \dots, 26, \quad (3.4.1)$$

Los dos predictores a usar son:

i $\mu_k = \beta_1 + \beta_2 x_k$ (Forma lineal)

ii $\mu_k = g(x_k)$ (Forma no lineal)

A su vez, se consideran dos variables aleatorias para modelar los errores:

i $\epsilon_k \sim SH(0, 2, \alpha)$

ii $\epsilon_k \sim SH - t(0, 2, \alpha, 4),$

donde $k = 1, 2, \dots, 26$ son los 26 especímenes cilíndricos a base de níquel, Y_k representa el logaritmo natural del número de ciclos antes de la falla contra el Pseudo-estrés, μ_k es el predictor del modelo, x_k es el nivel de Pseudo-estrés del k -ésimo espécimen cilíndrico, $k = 1, 2, \dots, 26$ y ϵ_k son los errores aleatorios independientes e idénticamente distribuidos.

Así, se estudian los cuatro modelos posibles para ajustar el número de ciclos hasta la falla. A su vez, se consideran dos variables aleatorias para modelar los errores:

- i $\epsilon_k \sim SH(0, 2, \alpha)$
- ii $\epsilon_k \sim SH - t(0, 2, \alpha, 4)$

Así, se estudian los cuatro modelos posibles para ajustar el número de ciclos hasta la falla.

La siguiente tabla muestra las estimaciones de los parámetros β_1 y β_2 de acuerdo a los modelos propuestos por los autores:

1. Nelson W. (1984) $\mu(S) = \gamma_1 + \gamma_2[LPS - A] + \gamma_3[LPS - A]^2$.
2. Meeker & Escobar (1998) $\log[t_p(x)] = y_p(x) = \beta_0 + \beta_1x + \beta_2x^2 + \Phi^{-1}(p)\sigma$.
3. Heuchenne & Van Keilegom (2007) $\log Y = \beta_0 + \beta_1 \log(x - \gamma) + \sigma(X)\varepsilon$.

Estimaciones			
Autores	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$
Nelson (1984)	4.482	-7.012	19.962
Meeker & Escobar (1998)	217.61	-85.52	8.48
Heuchenne & Van Keilegom(2007)	9.243	-1.722	-

TABLA 3.1. Estimaciones de los $\hat{\beta}$'s usando los modelos descritos anteriormente por los autores: Nelson(1984), Meeker & Escobar(1998) y Heuchenne & Van Keilegom(2007)

Estimaciones			
Autores	$ee(\hat{\beta}_0)$	$ee(\hat{\beta}_1)$	$ee(\hat{\beta}_2)$
Nelson (1984)	0.62	0.884	6.64
Meeker & Escobar (1998)	62.1	26.53	2.83
Heuchenne & Van Keilegom(2007)	3.05	0.64	-

TABLA 3.2. Estimaciones de los $\hat{\beta}$'s usando los modelos descritos anteriormente por los autores: Nelson(1984), Meeker & Escobar(1998) y Heuchenne & Van Keilegom(2007)

Heuchenne & Van Keilegom hicieron las comparaciones con la estimación del parámetro $\hat{\gamma} = 71.1792$ con un modelo propuesto por ellos mismos en 2005b cuya estimación fué de $\hat{\gamma} = 65.273$ por lo tanto es improbable estadísticamente hacer las comparaciones del modelo propuesto por Heuchenne & Van Keilegom (2007) y el que se presenta en este documento.

A continuación se muestran las estimaciones obtenidas por los modelos Log-Birnbaum-Saunders censurado paramétrico y no paramétrico y los modelos Log-Birnbaum-Saunders-t paramétrico y no paramétrico:

Modelo Log-Birnbaum-Saunders Censurado						
*	$\hat{\beta}_1$	$\hat{\beta}_2$	ee($\hat{\beta}_1$)	ee($\hat{\beta}_2$)	$\hat{\alpha}$	ee($\hat{\alpha}$)
Paramétrico	8.94286	-0.05179	0.76409	0.00736	0.77526	0.11857
No Paramétrico	-	-	-	-	0.71474	0.11357

TABLA 3.3. (4) Estimaciones del modelo Log-Bs censurado paramétrico y con efectos no paramétricos

Modelo Log-Birnbaum-Saunders-t (4) Censurado						
*	$\hat{\beta}_1$	$\hat{\beta}_2$	ee($\hat{\beta}_1$)	ee($\hat{\beta}_2$)	$\hat{\alpha}$	ee($\hat{\alpha}$)
Paramétrico	9.68269	-0.05761	0.70951	0.00678	0.56360	0.11663
No Paramétrico	-	-	-	-	0.35694	0.09334

TABLA 3.4. Estimaciones del modelo Log-Bs-t(4) censurado paramétrico y con efectos no paramétricos

Note que el valor de $\hat{\alpha}$ de los modelos Log-Birnbaum-Saunders censurado y Log-Birnbaum-Saunders-t censurado ambos con efectos no paramétricos es menor que el estimado por el modelo de regresión Log-Birnbaum-Saunders censurado y Log-Birnbaum-Saunders censurado-t paramétrico, lo que significa que el ajuste del modelo propuesto estadísticamente es más óptimo. La significancia de los $\hat{\beta}'s$ depende del modelo paramétrico, sin embargo no significa que si las estimaciones de los $\hat{\beta}'s$ son no significativo en el modelo Log-Birnbaum-Saunders censurado (o Log-Birnbaum-Saunders-t censurado), también lo sea en el modelo Log-Birnbaum-Saunders censurado con efectos no paramétricos (o en su defecto en el modelo Log-Birnbaum-Saunders censurado-t con efectos no paramétricos), ya que esto indica que si hay relación entre la variable explicativa y la de respuesta, solo que la relación es no lineal. Para estar seguro de la ganancia del modelo que se propone, se calculan los criterios de información bayesiana (BIC) y el criterio de Akaike (AIC)

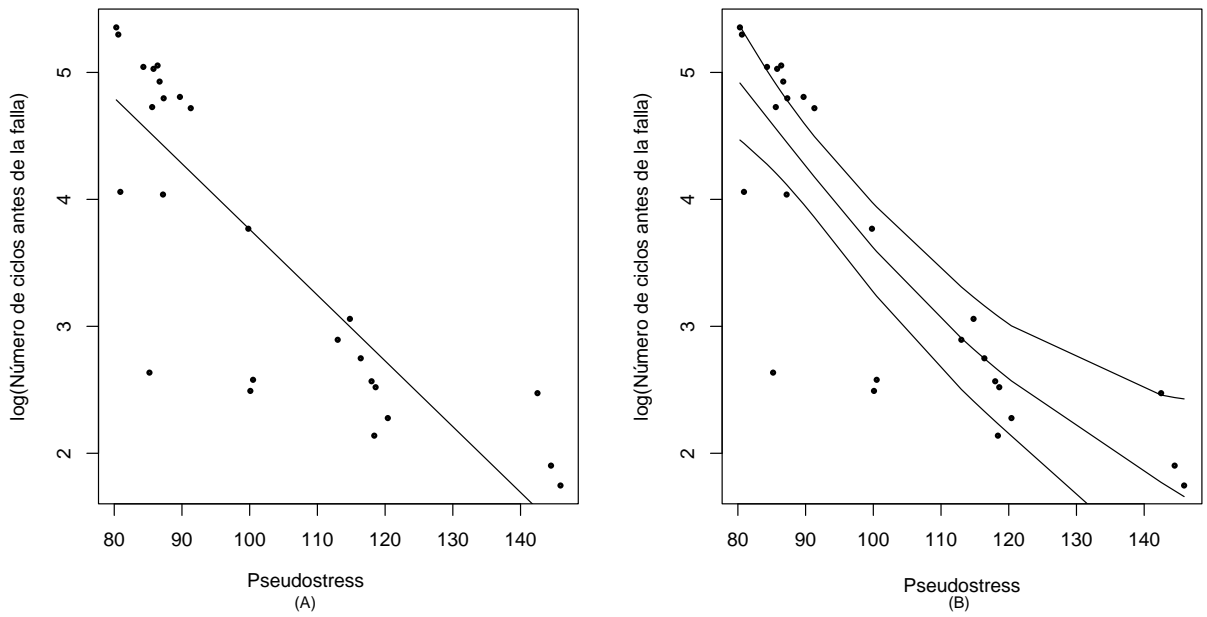


FIGURA 3.10. Ajuste del modelo Log-Bs Censurado para los datos de fatiga en la prueba de tensión (A) y Ajuste del modelo Log-Bs Censurado con efectos no paramétricos para los datos de fatiga en la prueba de tensión (B)

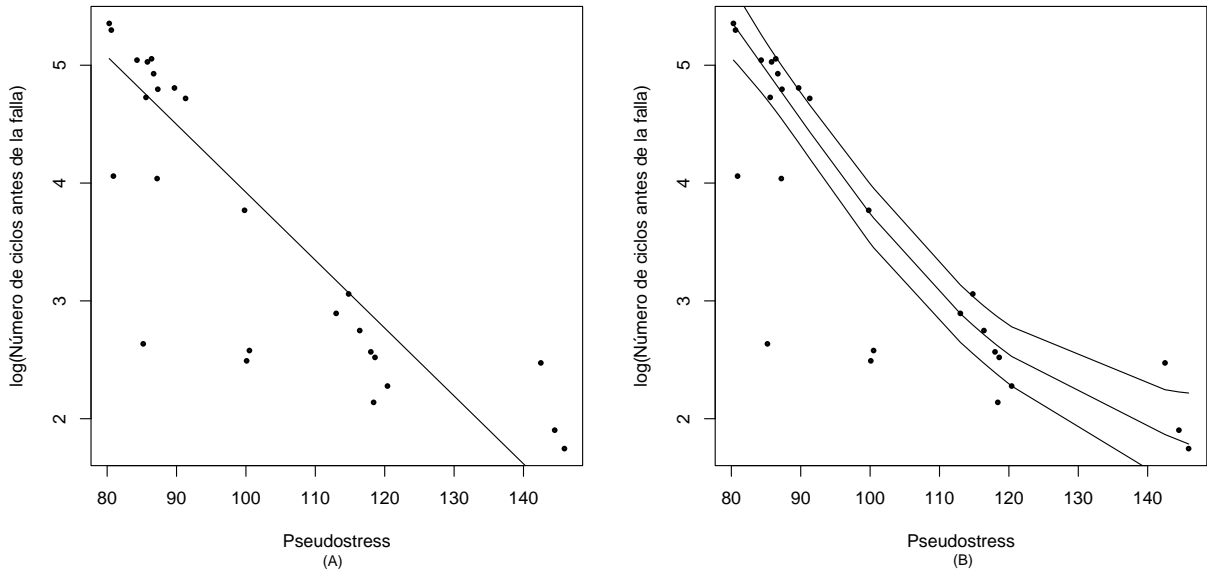


FIGURA 3.11. Ajuste del modelo Log-Bs-t Censurado para los datos de fatiga en la prueba de tensión (A) y Ajuste del modelo Log-Bs-t Censurado con efectos no paramétricos para los datos de fatiga en la prueba de tensión (B)

las Figuras (3.10) y (3.11) muestran el ajuste para los datos de fatiga en la prueba de tensión de los modelos Log-BS censurado, Log-BS Censurado con efectos no paramétricos, Log-BS-t(4) censurado y Log-BS-t(4) Censurado con efectos no paramétricos, note que el ajuste de ambos modelos (por gráfica) es distinto, esto se debe a que el modelo propuesto (gráfica B) posee efectos no paramétricos los cuales hacen complicado la comparación de los residuales con respecto a los modelos (gráfica A) paramétricos. Mientras en la figura (3.10(B)) y (3.11(B)) se comparan el ajuste con bandas de confianza de los cuantiles 2.5 % y 97.5 %, las figuras (3.10(A)) y (3.11(A)) permiten observar el ajuste del modelo con la curva de la distribución normal teórica, por lo que se procede a calcular el BIC (Criterio de información bayesiano) y el AIC (Criterio de información de Akaike) para hacer las comparaciones de forma óptima.

Las tablas 3.5 y 3.6 presentan los valores obtenidos del BIC y del AIC para los modelos Log-BS Censurado, Log-BS Censurado con efectos no paramétricos , Log-BS-t(4) Censurado y Log-BS-t(4) Censurado con efectos no paramétricos, note que para ambos criterios, los modelos Log-BS censurado y Log-BS-t(4) censurado ambos con efectos no paramétricos brindan el mejor ajuste en los datos si se compara con los modelos paramétricos, con un BIC de 14.112 y un AIC de 9.935 (ver en la tabla 3.5), el modelo Log-BS censurado con efectos no paramétricos describe mejor los datos analizados en este caso, y con un BIC de 3.029 y un AIC de -1.147 (ver en la tabla 3.6), el modelo Log-Bs-t(4) describe mejor los datos analizados en este caso. Por tanto se puede concluir que el modelo Log-Birnbaum-Saunders-t Censurado con efectos no paramétricos y el Log-Birnbaum-SaundersCensurado con efectos no paramétricos pueden ser usados para describir la relación entre el número de ciclos antes de la falla y el nivel de Pseudoestres

Modelo Log-Birnbaum-Saunders Censurado		
*	AIC	BIC
Paramétrico	12.56	16.34
No Paramétrico	9.935	14.112

TABLA 3.5. Valores de los criterios AIC y BIC para los modelos Log-BS Censurado y Log-BS Censurado con efectos no paramétricos

Modelo Log-Birnbaum-Saunders-t (4) Censurado		
*	AIC	BIC
Paramétrico	10.30	14.7
No Paramétrico	-1.147	3.029

TABLA 3.6. Valores de los criterios AIC y BIC para los modelos Log-BS-t(4) Censurado y Log-BS-t(4) Censurado con efectos no paramétricos

La tabla (3.7), muestran las diferentes estimaciones que obtuvo Nelson(1984) y Meeker & Escobar(1998). Para hacer comparables estas estimaciones con las obtenidas en la tabla (3.5) y la tabla (3.6) se debe sumar 187.3116 a los valores de AIC y BIC de los modelos BS y BS-t censurados paramétricos, ya que la escala obtenida en estos modelos es logarítmica, estas estimaciones se encuentran en las tablas (3.8) y (3.9).

Estimaciones		
Autores	AIC	BIC
Nelson (1984)	494.55	489.35
Meeker & Escobar (1998)	292.5	287.66

TABLA 3.7. Valores de los criterios AIC y BIC para los modelos estimados por Meeker & Escobar(1998) y Nelson(1984)

Modelo Log-Birnbaum-Saunders Censurado		
*	AIC	BIC
Paramétrico	199.87	203.65
No Paramétrico	197.94	201.42

TABLA 3.8. Valores de los criterios AIC y BIC para los modelos Log-BS Censurado y Log-BS-t Censurado con efectos no paramétricos comparables con los calculados por Meeker & Escobar(1998) y Nelson(1984)

Modelo Log-Birnbaum-Saunders-t (4) Censurado		
*	AIC	BIC
Paramétrico	197.61	202.01
No Paramétrico	186.16	190.34

TABLA 3.9. Valores de los criterios AIC y BIC para los modelos Log-BS-t(4) Censurado y Log-BS-t(4) Censurado con efectos no paramétricos comparables con los calculados por Meeker & Escobar(1998) y Nelson(1984)

De las tablas anteriores se pueden observar que al comparar las estimaciones de los modelos paramétricos Log-Birnbaum-Saunders y Log-Birnbaum-Saunders-t, los criterios AIC y BIC muestran que efectivamente existe una ganancia estadística en cuanto a las estimaciones por este modelo con diferencia de los modelos propuestos por Meeker & Escobar (1998) y Nelson (1984), esto es, las estimaciones son estadísticamente más óptimas. Ahora, si a su vez se comparan con el modelo propuesto, se puede observar que los criterios AIC y BIC son mucho más óptimos que si se asume que el efecto de la variable explicativa es lineal, en otras palabras, parece que el ajuste no lineal es mucho más óptimo por lo que la relación de los $\hat{\beta}'s$ con respecto al número de ciclos antes de la falla es no lineal.

APENDICE

APÉNDICE A

Algoritmos

A.1. Datos

A continuación se presentan los datos utilizados para el estudio del nivel de Pseudo-estrés con respecto al número de ciclos antes de la falla. Los casos marcados con F son fallas y los casos marcados con C son censura. (Tomado de *Nelson(1990)*, página 272)

Pseudo-estrés	K-Ciclos	Estado	Pseudo-estrés	K-Ciclos	Estado
80.3	211.629	F	99.8	43.331	F
80.6	200.027	F	100.1	12.076	F
80.8	57.923	C	100.5	13.181	F
84.3	155.000	F	114.8	21.300	F
85.2	13.949	F	114.8	21.300	F
85.6	112.968	C	116.4	15.616	F
85.8	152.680	F	118.0	13.030	F
86.4	156.725	F	118.4	8.489	F
86.7	138.114	C	118.6	12.434	F
87.2	56.723	F	120.4	9.750	F
87.3	121.075	F	142.5	11.865	F
89.7	122.372	C	144.5	6.705	F
91.3	112.002	F	145.9	5.733	F

TABLA A.1. Ciclos de fatiga de especímenes a base de Níquel bajo un nivel de Pseudo-estrés (En unidades de miles de ciclos hasta la falla.)

A.2. Implementación del algoritmo de la sección 3.3

```

library(numDeriv)
library(Formula)
library(normalp)
library(GIGrvg)
library(survival)
library(ssym)

n <- 1000
## 0.5, 1, 1,5, 2
alpha <- 1.5
x <- sort(round(runif(n),digits=1))
eta <- 0.5 + 4*((5*x)/(0.3+x) -2)
rep <- 5000
temp <- ncs(x,all.knots=TRUE)
N <- attr(temp,"N")
K <- attr(temp,"K")
sBS <- function(z,alpha) 1-pnorm((2/alpha)*sinh(z))
dBS <- function(z,alpha){
xi1 <- (2/alpha)*cosh(z)
xi2 <- (2/alpha)*sinh(z)
log(xi1) - xi2^2/2
}
loglikBS <- function(theta){
g <-theta[1:ncol(N)]
eta <- N%*%g
z <- (y - eta)/2
alpha <- exp(theta[ncol(N)+1])
-sum(dBS(z,alpha)*(1-delta) + log(sBS(z,alpha))*delta) +(1/2)*lambda*t(g)%*%K%*%g
}
loglikBS2 <- function(theta){
g <- theta[1:ncol(N)]
eta <- N%*%g
z <- (y - eta)/2
alpha <- theta[ncol(N)+1]
-sum(dBS(z,alpha)*(1-delta) + log(sBS(z,alpha))*delta) +(1/2)*lambda*t(g)%*%K%*%g
}
resultsBS <- matrix(0,rep,length(unique(x)))
i <- 1
while(i <= rep){
y <- eta + asinh((alpha/2)*rnorm(n))*2
## 0.00003, 0.0003, 0.003
rate <- 0.0003 ##69% de censura
c <- log(rexp(n,rate=rate))
delta <- ifelse(c<=y,1,0)
y_k <- ifelse(c<=y,c,y)
gcv <- function(lambda){
H <- N%*%solve(t(N)%*%N + exp(lambda)*K)%*%t(N)

```



```
sum((y_k - H%*%y_k)^2)/(1 - mean(diag(H)))^2
}
ps <- optim(0,gcv,method="BFGS")
lambda <- exp(ps$par)
g0 <- solve(t(N)%*%N + lambda*K)%*%t(N)%*%y_k
salidaBS <- try(optim(c(g0,log(alpha)), loglikBS, method="BFGS"),silent=TRUE)
if(is.list(salidaBS)){
  if(salidaBS$convergence==0){
    resultsBS[i,] <- unique(N%*%salidaBS$par[1:ncol(N)])
    i <- i + 1
  }
}
}
}
bandas <- apply(resultsBS,2,quantile,probs=c(0.025,0.975))
plot(x,eta, type="l")
lines(unique(x),bandas[1,],col="red")
lines(unique(x),bandas[2,],col="red")
```

A.3. Algoritmo de Regresión log-Birnbaum-Saunders censurado paramétrico y no paramétrico con censura no informativa.

A.3.1. Datos de fatiga de bajo ciclo para una prueba de tensión controlada de 26 especímenes cilíndricos a base de níquel.

```

niquel <- read.table("niquel.txt",header=TRUE)
attach(niquel)
x <-Pseudostress
y <- log(k_cycles)
X <- cbind(1,Pseudostress)
delta<-status
X0 <- X[delta==0,]
beta <- solve(t(X0)%*%X0)%*%t(X0)%*%y[delta==0]
alpha <- 1

sBS <- function(z,alpha) 1-pnorm((2/alpha)*sinh(z))
dBS <- function(z,alpha){
xi1 <- (2/alpha)*cosh(z)
xi2 <- (2/alpha)*sinh(z)
-log(sqrt(2*pi)*2) + log(xi1) - xi2^2/2
}

loglikBS <- function(theta){
p <- length(theta) - 1
z <- (y - X%*%theta[1:p])/2
alpha <- exp(theta[p+1])
-sum(dBS(z,alpha)*(1-delta) + log(sBS(z,alpha))*delta)
}
CIBS <- function(theta){
p <- length(theta) - 1
z <- (y - X%*%theta[1:p])/2
alpha <- theta[p+1]
AIC <- -2*sum(dBS(z,alpha)*(1-delta) + log(sBS(z,alpha))*delta)
+ 2*(p+1)
BIC <- -2*sum(dBS(z,alpha)*(1-delta) + log(sBS(z,alpha))*delta)
+ log(length(y))*(p+1)
list(AIC=AIC, BIC=BIC)
}

loglikBS2 <- function(theta){
p <- length(theta) - 1
z <- (y - X%*%theta[1:p])/2
alpha <- theta[p+1]
-sum(dBS(z,alpha)*(1-delta) + log(sBS(z,alpha))*delta)
}

```

```

salidaBS <- try(optim(c(beta,log(alpha)), loglikBS, method="BFGS"),silent=TRUE)
parBS <- salidaBS$par
parBS[length(parBS)] <- exp(salidaBS$par[length(parBS)])
varBS <- diag(solve(hessian(loglikBS2, parBS)))
cbind(parBS,sqrt(varBS),parBS/sqrt(varBS),round(1-pchisq(parBS^2/varBS,1),5))
x2 <- seq(min(x),max(x),length=100)
par(mfrow=c(1,2))
plot(x,y,cex=0.3,lwd=3,xlab="Pseudostress",
ylab="log(Número de ciclos antes de la falla)",main="")
lines(x2,parBS[1] + parBS[2]*x2)
CIBS(parBS)

temp <- ncs(x,all.knots=TRUE)
N <- attr(temp,"N")
K <- attr(temp,"K")
loglikBS <- function(theta){
g <-theta[1:ncol(N)]
eta <- N**g
z <- (y - eta)/2
alpha <- exp(theta[ncol(N)+1])
-sum(dBS(z,alpha)*(1-delta) + log(sBS(z,alpha))*delta) + (1/2)*lambda*t(g)**K**g
}
CIBS <- function(theta){
g <-theta[1:ncol(N)]
eta <- N**g
z <- (y - eta)/2
alpha <- theta[ncol(N)+1]
gl <- sum(diag(solve(t(N)**N+lambda*K)**t(N)**N))
AIC <- -2*sum(dBS(z,alpha)*(1-delta) + log(sBS(z,alpha))*delta) + 2*(gl+1)
BIC <- -2*sum(dBS(z,alpha)*(1-delta) + log(sBS(z,alpha))*delta)
+ log(length(y))*(gl+1)
list(AIC=AIC, BIC=BIC)
}

loglikBS2 <- function(theta){
g <-theta[1:ncol(N)]
eta <- N**g
z <- (y - eta)/2
alpha <- theta[ncol(N)+1]
-sum(dBS(z,alpha)*(1-delta) + log(sBS(z,alpha))*delta) + (1/2)*lambda*t(g)**K**g
}

lambda <- 50000
g0 <- solve(t(N)**N + lambda*K)**t(N)**y
salidaBS <- try(optim(c(g0,log(alpha)), loglikBS, method="BFGS"),silent=TRUE)
parBS <- salidaBS$par

```

```

parBS[length(parBS)] <- exp(salidaBS$par[length(parBS)])
varBS <- solve(hessian(loglikBS2, parBS))
cbind(parBS[length(parBS)],sqrt(varBS[length(parBS),length(parBS)]))
varBS <- varBS[-length(parBS),-length(parBS)]
plot(x,y,cex=0.3,lwd=3,xlab="Pseudostress",
ylab="log(Número de ciclos antes de la falla)",main="")
ids <- sort(x,index.return=TRUE)$ix
lines(x[ids],N%%parBS[1:ncol(N)][ids])
lines(x[ids],N%%parBS[1:ncol(N)] - qnorm(0.975)*sqrt(diag(N%%varBS%%t(N))[ids])
lines(x[ids],N%%parBS[1:ncol(N)] + qnorm(0.975)*sqrt(diag(N%%varBS%%t(N))[ids])
CIBS(parBS)

```

A.4. Algoritmo de Regresión log-Birnbaum-Saunders-t(4) censurado paramétrico y no paramétrico con censura no informativa.

A.4.1. Datos de fatiga de bajo ciclo para una prueba de tensión controlada de 26 especímenes cilíndricos a base de níquel.

```

nu <- 4
sBSt <- function(z,alpha) 1-pt((2/alpha)*sinh(z),nu)
dBSt <- function(z,alpha){
xi1 <- (2/alpha)*cosh(z)
xi2 <- (2/alpha)*sinh(z)
log(gamma((nu+1)/2)) -log(2*sqrt(nu*pi)*gamma(nu/2))
+ log(xi1) - ((nu + 1)/2)*(log(xi2^2 + nu) - log(nu))
}

loglikBSt <- function(theta){
p <- length(theta) - 1
z <- (y - X%%theta[1:p])/2
alpha <- exp(theta[p+1])
-sum(dBSt(z,alpha)*(1-delta) + log(sBSt(z,alpha))*delta)
}

CIBSt <- function(theta){
p <- length(theta) - 1
z <- (y - X%%theta[1:p])/2
alpha <- theta[p+1]
AIC <- -2*sum(dBSt(z,alpha)*(1-delta) + log(sBSt(z,alpha))*delta) + 2*(p+1)
BIC <- -2*sum(dBSt(z,alpha)*(1-delta)
+ log(sBSt(z,alpha))*delta) + log(length(y))*(p+1)
list(AIC=AIC, BIC=BIC)
}

loglikBSt2 <- function(theta){
p <- length(theta) - 1

```

```

z <- (y - X%%theta[1:p])/2
alpha <- theta[p+1]
-sum(dBSt(z,alpha)*(1-delta) + log(sBSt(z,alpha))*delta)
}

salidaBSt <- try(optim(c(beta,log(alpha)), loglikBSt, method="BFGS"),silent=TRUE)
parBSt <- salidaBSt$par
parBSt[length(parBSt)] <- exp(salidaBSt$par[length(parBSt)])
varBSt <- diag(solve(hessian(loglikBSt2, parBSt)))
cbind(parBSt,sqrt(varBSt),parBSt/sqrt(varBSt),round(1-pchisq(parBSt^2/varBSt,1),5))
x2 <- seq(min(x),max(x),length=100)
par(mfrow=c(1,2))
plot(x,y,cex=0.3,lwd=3,xlab="Pseudostress",
ylab="log(Número de ciclos antes de la falla)",main="")
lines(x2,parBSt[1] + parBSt[2]*x2)
CIBSt(parBSt)

temp <- ncs(x,all.knots=TRUE)
N <- attr(temp,"N")
K <- attr(temp,"K")
loglikBSt <- function(theta){
g <- theta[1:ncol(N)]
eta <- N%%g
z <- (y - eta)/2
alpha <- exp(theta[ncol(N)+1])
-sum(dBSt(z,alpha)*(1-delta) + log(sBSt(z,alpha))*delta) + (1/2)*lambda*t(g)%%K%%g
}
CIBSt <- function(theta){
g <-theta[1:ncol(N)]
eta <- N%%g
z <- (y - eta)/2
alpha <- theta[ncol(N)+1]
gl <- sum(diag(solve(t(N)%%N+lambda*K)%%t(N)%%N))
AIC <- -2*sum(dBSt(z,alpha)*(1-delta) + log(sBSt(z,alpha))*delta) + 2*(gl+1)
BIC <- -2*sum(dBSt(z,alpha)*(1-delta) +
log(sBSt(z,alpha))*delta) + log(length(y))*(gl+1)
list(AIC=AIC, BIC=BIC)
}

loglikBSt2 <- function(theta){
g <- theta[1:ncol(N)]
eta <- N%%g
z <- (y - eta)/2
alpha <- theta[ncol(N)+1]
-sum(dBSt(z,alpha)*(1-delta) + log(sBSt(z,alpha))*delta) + (1/2)*lambda*t(g)%%K%%g
}

lambda <- 50000

```

```
g0 <- solve(t(N)**N + lambda*K)**t(N)**y
salidaBSt <- try(optim(c(g0,log(alpha)), loglikBSt, method="BFGS"),silent=TRUE)
parBSt <- salidaBSt$par
parBSt[length(parBSt)] <- exp(salidaBSt$par[length(parBSt)])
varBSt <- solve(hessian(loglikBSt2, parBSt))
cbind(parBSt[length(parBSt)],sqrt(varBSt[length(parBSt),length(parBSt)]))
varBSt <- varBSt[-length(parBSt),-length(parBSt)]
plot(x,y,cex=0.3,lwd=3,xlab="Pseudostress",
ylab="log(Número de ciclos antes de la falla)",main="")
ids <- sort(x,index.return=TRUE)$ix
lines(x[ids],N**parBSt[1:ncol(N)][ids])
lines(x[ids],N**parBSt[1:ncol(N)] - qnorm(0.975)*sqrt(diag(N**varBSt**t(N)))[ids])
lines(x[ids],N**parBSt[1:ncol(N)] + qnorm(0.975)*sqrt(diag(N**varBSt**t(N)))[ids])
CIBSt(parBSt)
```

Trabajo futuro

- El modelo de Regresión Log-Birnbaum-Saunders con efectos no paramétricos y censura no informativa brinda estimaciones robustas y siempre se aproxima al valor real siempre y cuando el parámetro de escala sea pequeño y el porcentaje de censura no sea grande mayor al 40 % de los datos o cuando el parámetro de escala es grande y el porcentaje de censura es pequeño, es decir, si se tiene porcentajes de censura mayores al 50 % y un parámetro de escala grande el modelo converge pero no arroja estimaciones que se aproximen al verdadero valor de la curva, por tanto sería importante seguir estudiando este tipo de modelos con la convicción de obtener mejores aproximaciones a la hora de relajar los supuestos del modelo, tener censura en los datos y adicional a ello que las variables explicativas de las cuales se sospechen que posean un efecto no lineal.

Bibliografía

- [1] BALARKRISHMAN N., LEIVA V., SANHUEZA A. AND VILCA F (2009), *Estimation in the Birnbaum-Saunders distribution based on scale-mixture of normals and the em-algorithm*. Statistics and Operations Research Transactions(SORT), **33**, 171 – 192
- [2] BARROS M, PAULA G.A AND LEIVA V. (2008), *A new class of survival regression models with heavy-tailed errors: robustness and diagnostics*. life time Data Analysis, **14**, 316-332.
- [3] BARROS M, PAULA G.A AND LEIVA V. (2009), *An R implementation for generalized Birnbaum-Saunders distributions*. Computational Statistics and Data Analysis **53**, 1511– 1528
- [4] BIRNBAUM Z.W AND SAUNDERS S.C (1969), *A new family of distributions*. Journal of Applied Probability **6(2)**, 319– 327.
- [5] CORDEIRO G.M AND ANDRADE M.G, *Transformed symmetric models*. Statistical Modelling, **11**, 371 – 388.
- [6] DESMOND, A.F. (1985), *Stochastic models of failure in random environments*. Canadian Journal of Statistics **13**, 171 - 183.
- [7] EILERS, P.H. AND MARX, B.D. (1996) , *Flexible Smoothing with B-Splines and Penalties*. Statistical Science, **11**, 89-121.
- [8] FR'ECHET, M. (1927), *Sur la loi de probabilit'e de l'ecart maximum*. Ann. de la Soc. polonaisede Math. **6**, 93 - 116.
- [9] FREUDENTHAL A.M & SHINOZUKA M.(1961) *Structural safety under conditions of ultimate load failure and fatigue*. Columbia University
- [10] GREEN P.J AND SILVERMAN B.W (1994), *Nonparametric regression and generalized linear models*.
- [11] HEUCHENNE C. AND VAN KEILEGOM I.(2007), *Nonlinear Regression with Censored Data*. Technometrics, **49(1)**, 34 – 44.
- [12] JONES M.C AND ARNOLD B.C (2008), *Distributions that are both log-symmetric and R-symmetric*. Electronic Journal of Statistics, **2(1)**, 1300 – 1208.
- [13] KLEIN J., MOESCHBERGER M. (1997), *Survival Analysis: Techniques for Censored and Truncated Data*. Springer, New York.

-
- [14] LEIVA V., BARROS M., PAULA G., SANHUEZA A.(2008A) *Generalized Birnbaum - Saunders distribution applied to air pollutant concentration*. *Environmetrics* **19**, 235-249.
- [15] LEIVA, V., ATHAYDE, E., AZEVEDO, C., MARCHANT, C., (2011A). *Modeling wind energy flux by a Birnbaum-Saunders distribution with unknown shift parameter*. *J. Appl. Stat.* **38**, 2819 - 2838.
- [16] LEIVA V. ATHADEYDE E., ACEVEDO C., MARCHANT C. (2011B), *Modeling wind energy flux by a Birnbaum - Saunders distribution with unknown shift parameter*. *J. Appl. Stat.* **38**, 2819-2838.
- [17] LEIVA V. (2016), *The Birnbaum-Saunders distribution*. Adolfo Ibañez University. Viña del Mar, Chile.
- [18] LIN JG., XIE FC., XIE BC. (2009), *Statisticak diagnostics for skew-t-normal non-linear models*. *Communications in statistics Simulation and Computation*, **38**, 2096-2110.
- [19] MCCULLAGH P. AND NELDER J.A (1989), *Generalized Linear Models*. Chapman and Hall (Second Edition).
- [20] MEEKER W.Q & ESCOBAR L.A (1998), *Statiscal Methods for Reblity Data*.New york: Wiley.
- [21] MYERS R.H, MONTGOMERY D.C, VINING G.G AND ROBINSON T.J (2010) , *Generalized Linear Models with applications in engineering and sciences*. John Wiley and sSons, INC., PUBLICATION (Second Edition).
- [22] NELSON W.B (1984), *Acelerated Testing Statistical Models, Test Plans, and Data Analysis*. A jhon Wiley \$ sons, INC. Publication.
- [23] NG H.K.T, KUNDU D. AND BALAKRISHANN N. (2003), *Modified moment estimation for the two-parameter Birnbaum-Saunders distribution*. *Computational Statistics and Data Anlalysis*, **43**, 283 –298.
- [24] PAULA G.A, LEIVA V., BARROS M. AND LIU S (2012), *Robust statistical modeling using Birnbaum-Saunders-t distribution splied to insurance*. *Applied Stochastic Models in Business and Industry*, **28**, 16 – 34.
- [25] PAULA G.A (2013), *On diagnostic in double generalized models*. *Computational Statistics and Data Analysis*, **8**, 44 – 51.
- [26] POLADSKI R (2008) *Characterization of diameter distribution data in near-natural forests using the Birnbaum-Saunders distribution*. *Can.J.For.Res.* **18**, 518-527
- [27] PUIG P. (2008) *A note on the harmonic law: A two-parameter family of distributiones for ratios*. *Statistics and probability Letters.* **78**, 320-326.
- [28] RIECK J.R (1989), *Statistical Analysis for the Birnbaum-Saunders Fatigue Life Distribution*. Ph.D. thesis, Department of Mathematical Sciences, Clemson University, Clemson.

-
- [29] RIECK J.R (1999), *A moment-generating function with application to the Birnbaum-Saunders distribution*. Communications in Statistics. Theory and Methods, **28**,2213 –2222.
- [30] RIECK J.R AND NEDELMAN J.R (1991), *A log linear model for the Birnbaum-Saunders distribution*,Technometrics,**33**, 51 – 60.
- [31] VANEGAS L.H, RONDON L.M AND CYSNEROS F.J (2011), *Diagnostic procedures in Birnbau,-Saunders nonlinear regression models*. Reporte interno de investigacion, Universidad Nacional de Colombia, 1 – 20.
- [32] VANEGAS L.H AND PAULA G.A (2014), *Fitting Semiparametric Symmetric Regression Models*.
- [33] VANEGAS L.H AND PAULA G.A (2015A), *A semiparametric approach for joint modeling of median and skewness*. TEST, **24**, 110 –135.
- [34] VANEGAS L.H AND PAULA G.A (2016B), *Log-symmetric distributions: Statistical properties and parameter estimation*.Brazilian Journal of Probability and Statistics, **30**, 106 – 220
- [35] WOOD S. (2006), *Generalized Additive Models: an introduction with R*. Chapman and Hall-RC.
- [36] XIE F.CH AND WEIB B.CH (2007), *Diagnostics analysis for log-Birnbaum Saunders regression models*. Computational Statistics and Data Analysis, **51**, 4692 – 4706.