

---

# Modelo de regresión Birnbaum-Saunders en la presencia de efectos no paramétricos y censura no informativa

## Birnbaum-Saunders regression model under the presence of non-parametric effects and non-informative

Laura Marcela Rúa Yáñez<sup>a</sup>  
laura.rua@docentes.umb.edu.co

Luis Hernando Vanegas Penagos<sup>b</sup>  
lhvanegasp@unal.edu.co

---

### 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.

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

### 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.

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

---

<sup>a</sup>Universidad Manuela Beltrán

<sup>b</sup>Universidad Nacional de Colombia

## 1. Introducción

La distribución Birnbaum-Saunders (1969a) se basa en el daño acumulativo que produce la fatiga en los materiales, este argumento es la regla de Miner (1945). Birnbaum-Saunders (1968) obtuvieron una interpretación probabilística de esta regla. Por lo tanto, la distribución de fatiga-vida se derivó de un modelo que muestra que el tiempo total transcurrido hasta que el daño acumulativo, producido por el desarrollo y crecimiento de una grieta dominante, sobrepasa un valor umbral y hace que el material del espécimen falle. Desmond (1985) reforzó la justificación para el uso de esta distribución al relajar algunos supuestos iniciales hechos por Birnbaum-Saunders (1969a). Además, una desviación más general de la distribución de Birnbaum-Saunders fue proporcionada por Desmond (1985) basada en un modelo biológico introducido por Cramer (1946).

Aunque la distribución fatiga-vida se conoce como el modelo de Birnbaum-Saunders, un trabajo anterior atribuido a Freudenthal y Shinozuka (1961) introdujo un modelo similar con una parametrización diferente. Además, el trabajo de Fréchet (1927) presentó algunas ideas similares relacionadas con el Birnbaum-Saunders, pero también con una parametrización diferente a la considerada por Birnbaum y Saunders (1969a).

Los modelos de regresión de log-Birnbaum-Saunders (Rieck y Nedelman (1991)) se han aplicado extensivamente en estudios de fatiga de materiales en los que 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 puede conducir a una degradación que conduce a un proceso de fatiga. En años recientes, los modelos de tiempos de falla acelerada o modelos de log-localización-escala con una distribución de error específica y censura aleatoria no informativa han sido objeto de estudio, diversos investigadores como Wood (2006), Lin (2009), Vanegas (2015), 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 modela un solo parámetro de la distribución de tiempo de fallo y éste, no admite la presencia de efectos no paramétricos en sus componentes sistemáticas.

Este artículo 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.

## 2. La Distribución Birnbaum-Saunders

### 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:

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

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:

- (a) Un material está sujeto a cargas cíclicas o choques repetitivos, que producen una grieta o desgaste en el material.
- (b) La falla en el material ocurre cuando el tamaño de la grieta del material excede un nivel de resistencia (umbral), denotado por  $\omega$ .
- (c) La secuencia de cargas impuestas al material es la misma de un ciclo a otro.
- (d) 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
- (e) El tamaño total de la grieta causada por el ciclo  $k$ ,  $Y_k$ , es una variable aleatoria que sigue una distribución de la media  $\mu$  y la varianza  $\sigma^2$ , ambas finitas.
- (f) 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:

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  $\omega$ .

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:

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

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:

$$F_T(t) = \Phi \left[ \frac{1}{\alpha} \left( \sqrt{\frac{t}{\eta}} - \sqrt{\frac{\eta}{t}} \right) \right], \quad t > 0 \tag{2}$$

Si la función de distribución acumulada de  $T$  se puede escribir como [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:

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

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  $\eta$ , 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)  $\eta$  es la mediana de T

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

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

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

(B7) El percentil 100(q) % de T está dada por  $t(q) = \eta \left[ \alpha z_q + (\alpha^2 z_q^2 + 4)^{1/2} \right]^2 / 4$ , es decir,  $P[T \geq t(q)] = q$ , en que  $z_q$  es el cuantil 100(q) % 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 [1] representa el comportamiento de  $f_T(t)$  para diferentes valores de  $\alpha$  y  $\eta$ . Se puede observar que la distribución de T presenta diferentes niveles de asimetría. De hecho, para  $\eta$  fijo, la distribución de T tiende a ser simétrica a medida que el valor de  $\alpha$  tiende a cero. Además para  $\alpha$  fijo, la varianza de T aumenta a medida que  $\eta$  aumenta.

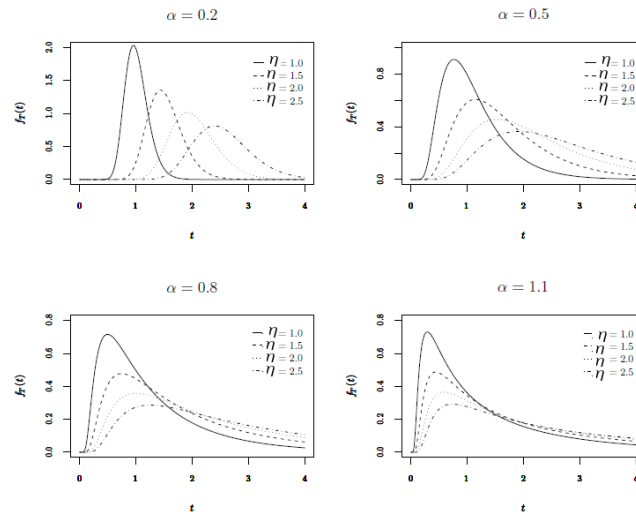


Figura 1: Densidad Birnbaum-Saunders para diferentes valores de  $\alpha$  y  $\eta$

### 3. Formulación del modelo

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 [3], teniendo parámetros de forma y escala dados por  $\alpha > 0$  y  $\eta_k$ , respectivamente.

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

$$T_k = \eta_k \xi_k, \quad k = 1, \dots, n, \quad (4)$$

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:

$$f(\xi) = \frac{1}{2\xi\sqrt{2\pi}} \left[ \sqrt{\xi} + \frac{1}{\sqrt{\xi}} \exp \left\{ -\frac{1}{2\alpha^2} \left( \frac{1}{\sqrt{\xi}} - \sqrt{\xi} \right)^2 \right\} \right], \quad \xi > 0, \quad (5)$$

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

$$\log(\eta_k) = \mathbf{x}_k^\top \boldsymbol{\beta} + g_1(s_{k1}) + \dots + g_q(s_{kq}), \quad k = 1, \dots, n, \quad (6)$$

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  $k$

- $\beta = (\beta_1, \dots, \beta_p)^\top$  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  $(i, j)$  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

$$\mu_k = \mathbf{x}_k^\top \boldsymbol{\beta} + \mathbf{m}_{k1}^\top \boldsymbol{\tau}_1 + \dots + \mathbf{m}_{kq}^\top \boldsymbol{\tau}_q. \quad (7)$$

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:

$$Y_k = \boldsymbol{\mu}_k + \epsilon_k, \quad k = 1, \dots, n, \quad (8)$$

donde:

- $Y_k = \log(T_k)$ ,
- $\mu_k = \log(\eta_k) = \mathbf{X}\boldsymbol{\beta} + \mathbf{N}_1\boldsymbol{\tau}_1 + \mathbf{N}_2\boldsymbol{\tau}_2 + \dots + \mathbf{N}_q\boldsymbol{\tau}_q$ ,
- $\epsilon_1, \dots, \epsilon_n$  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  $\epsilon_k$  es simétrica alrededor de 0, lo cual implica que  $E(Y_k) = \mu_k$ ,  $\text{Var}(Y_k) = w(\alpha)$ , y la distribución de  $Y_k$  es simétrica alrededor de  $\boldsymbol{\mu}_k$ .

### 3.1. 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  $\hat{\boldsymbol{\theta}} = \text{argmax } PL(\boldsymbol{\theta})$ , 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.
- $y_k = \log(T_k)$ .
- $\xi_{k1} = \frac{2}{\alpha} \cosh(u_k)$ .
- $\xi_{k2} = \frac{2}{\alpha} \sinh(u_k)$ .
- $u_k = (y_k - \boldsymbol{\mu}_k^\top) / 2$
- $\delta_k$  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.

$$\mathbf{U}(\boldsymbol{\theta}) = \begin{bmatrix} \frac{1}{2} \overline{\mathbf{X}}^\top \mathbf{S} \\ -\frac{n}{\alpha} + \frac{1}{\alpha} \sum_{i=1}^n \xi_{i2}^2 \end{bmatrix} - \overline{\mathbf{M}} \boldsymbol{\theta}, \quad (9)$$

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. 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  $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}^{[l]}$ .

### 3.3. 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

$$\operatorname{AIC}(\boldsymbol{\theta}) = -2L(\hat{\boldsymbol{\theta}}) + 2(p + \operatorname{gl}_1 + \dots + \operatorname{gl}_r), \quad (10)$$

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  $g_r(\cdot)$ .



### 3.4. Inferencia estadística

1. Para grandes tamaños de la muestra  $\hat{\theta} \sim N(\theta, \mathbf{J}^{-1}(\theta))$ , es decir, para  $n$  suficientemente grande  $\hat{\theta}$  es un estimador aproximadamente insesgado de  $\theta$ , cuya matriz de varianzas y covarianzas aproximada es  $\text{Var}(\hat{\theta}) = \mathbf{J}^{-1}(\theta)$ .
2. La matriz de varianzas y covarianzas de  $\hat{\theta}$  para muestras grandes se puede estimar consistentemente usando la matriz  $\mathbf{J}^{-1}(\hat{\theta})$ .
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  $Z^2 > \chi^2_{1-\gamma}(1)$ ”, donde  $Z = \hat{\theta}_j / \text{ee}(\hat{\theta}_j)$ , en que  $\text{ee}(\hat{\theta}_j)$  es la raíz cuadrada del  $j$ -ésimo elemento de la diagonal principal de la matriz  $\mathbf{J}^{-1}(\hat{\theta})$ .

## 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 [2] 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 observar que el Pseudo-estrés tiene un efecto no lineal sobre el logaritmo de los ciclos.

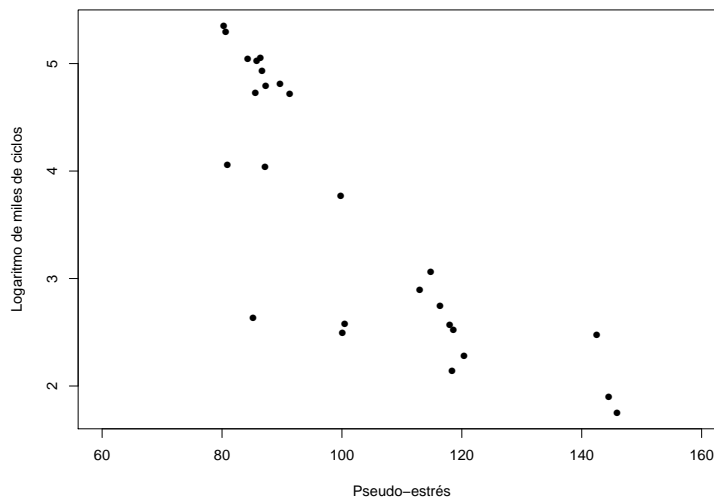


Figura 2: Logaritmo del número de ciclos antes de la falla contra el Pseudoestress.

El modelo a estimar está dado por:

$$Y_k = \mu_k + \epsilon_k, \quad k = 1, \dots, 26, \tag{11}$$

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,  $\mu_k$  es el predictor del modelo,  $x_k$  es el nivel de Pseudo-estrés del  $k$ -ésimo espécimen cilíndrico,  $k = 1, 2, \dots, 26$  y  $\epsilon_k$  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  $\beta_1$  y  $\beta_2$  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)  $\log[t_p(x)] = y_p(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \Phi^{-1}(p)\sigma$ .
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 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)

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.

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

Tabla 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)

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						
*	$\hat{\beta}_1$	$\hat{\beta}_2$	$ee(\hat{\beta}_1)$	$ee(\hat{\beta}_2)$	$\hat{\alpha}$	$ee(\hat{\alpha})$
<b>Paramétrico</b>	8.94286	-0.05179	0.76409	0.00736	0.77526	0.11857
<b>No Paramétrico</b>	-	-	-	-	0.71474	0.11357

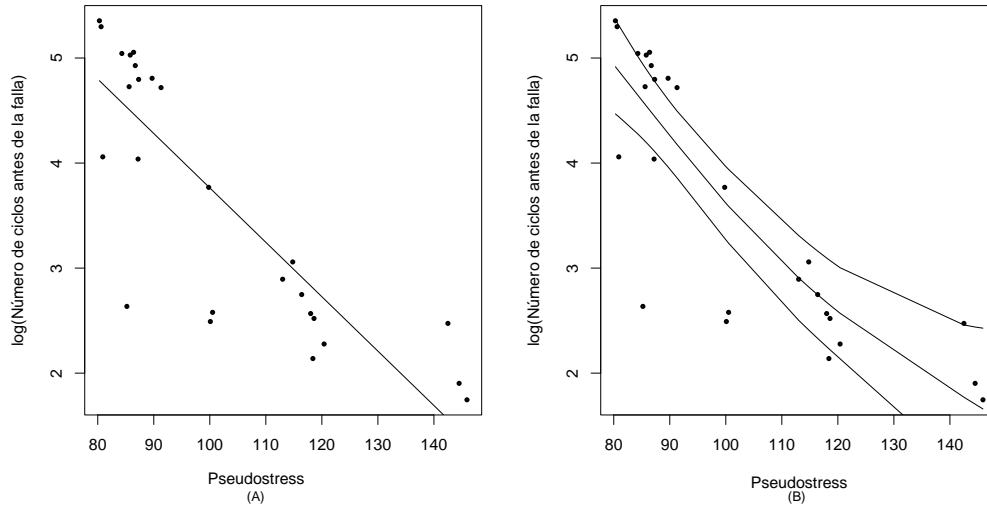
Tabla 3: Estimaciones del modelo Log-Bs censurado paramétrico y con efectos no paramétricos

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

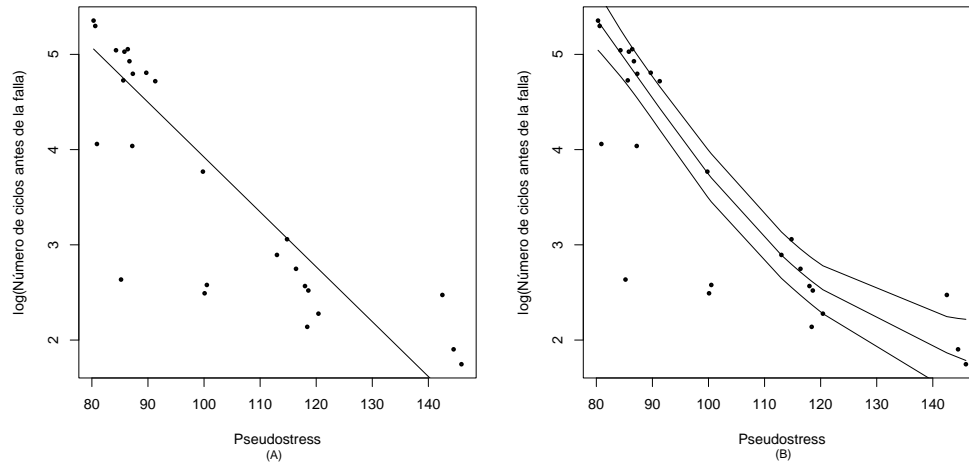
Tabla 4: Estimaciones del modelo Log-Bs-t(4) censurado paramétrico y con efectos no paramétricos

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)

las Figuras [3(a)] y [3(b)] de la gráfica [3] 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(b)] se comparan el ajuste con bandas de confianza de los cuantiles 2.5% y 97.5%, las figuras [3(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.



((a))



((b))

Figura 3: **3(a)**: 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).

**3(b)** 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 tablas [5] y [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 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 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-Saunders Censurado 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

<b>Modelo Log-Birnbaum-Saunders Censurado</b>		
<b>*</b>	<b>AIC</b>	<b>BIC</b>
<b>Paramétrico</b>	12.56	16.34
<b>No Paramétrico</b>	9.935	14.112

Tabla 5: Valores de los criterios AIC y BIC para los modelos Log-BS Censurado y Log-BS Censurado con efectos no paramétricos

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

Tabla 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 [7], muestran las diferentes estimaciones que obtuvo Nelson(1984) y Meeker & Escobar(1998). Para hacer comparables estas estimaciones con las obtenidas en la tabla (5) y la tabla (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 (8) y (9).

<b>Estimaciones</b>		
<b>Autores</b>	<b>AIC</b>	<b>BIC</b>
<b>Nelson (1984)</b>	<b>494.55</b>	<b>489.35</b>
<b>Meeker &amp; Escobar (1998)</b>	<b>292.5</b>	<b>287.66</b>

Tabla 7: Valores de los criterios AIC y BIC para los modelos estimados por Meeker & Escobar(1998) y Nelson(1984)

De las tablas anteriores [8] y [9] 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  $\beta$ 's con respecto al número de ciclos antes de la falla es no lineal.

<b>Modelo Log-Birnbaum-Saunders Censurado</b>		
<b>*</b>	<b>AIC</b>	<b>BIC</b>
<b>Paramétrico</b>	199.87	203.65
<b>No Paramétrico</b>	197.94	201.42

Tabla 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)

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

Tabla 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)

## 5. Conclusiones

- En este documento se introdujo un nuevo modelo de regresión Log-Birnbaum-Saunders para la estimación de parámetros con efectos no paramétricos y censura (a la derecha) no informativa, donde la forma no paramétrica es asumida por  $g(s_k)$ . Las sistemáticas en los análisis de datos de tiempo de vida (fatiga) y por tanto brinda una estimación más robusta cuando las variables explicativas poseen un efecto no lineal.
- El análisis de los datos de fatiga para una prueba de tensión controlada en 26 especímenes cilíndricos a base de níquel presentada en la sección (3) muestran una muy buena aproximación aplicando el modelo propuesto en este documento y haciendo la comparación con la estimación sin asumir que la variable explicativa tiene un efecto no lineal, es decir, usando el modelo de regresión Log-BS censurado se puede verificar que el modelo efectivamente brinda mejores estimaciones cuando las variables explicativas tienen un efecto no lineal sobre el  $\log(y)$ . Además si se hace la comparación con las estimaciones expuestas en Cédric Heuchenne & Ingrid Van Keilegom (2007) y en Meeker, W.Q & Escobar, L.A. (1998) se puede verificar que el modelo efectivamente brinda estimaciones más robustas con este método que con los propuestos

## Referencias

- S. A. Balarkrishnan N., Leiva V. and V. F. 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, 2009.
- P. G. Barros M and L. V. An r implementation for generalized birnbaum-saunders distributions. *Computational Statistics and Data Analysis*, 53:1511–1528, 2009.
- H. C. and V. K. I. Nonlinear regression with censored data. *Technometrics*, 49(1):34–44, February 2007.
- X. F.CH and W. B.CH. Diagnostics analysis for log-birnbaum-saunders regression models. *Computational Statistics and Data Analysis*, 51:4692–4706, 2007.
- P. G.A. On diagnostic in double generalized models. *Computational Statistics and Data Analysis*, 8:44–51, 2013.
- C. G.M and A. M.G. Transformed symmetric models. *Statistical Modelling*, 11:371–388, 2011.
- R. J.R. A moment-generating function with application to the birnbaum-saunders distribution. *Communications in Statistics. Theory and Methods*, 28:2213–2222, 1999.
- R. J.R and N. J.R. A log linear model for the birnbaum-saunders distribution. *Technometrics.*, 33: 51–60, 1991.
- V. L.H and P. G. A semiparametric approach for joint modeling of median and skewness. *TEST*, 24:110–135, 2015a.
- V. L.H and P. G. Log-symmetric distributions: Statistical properties and parameter estimation. *Brazilian Journal of Probability and Statistics*, 30:196–220, 2016b.
- V. L.H and P. G.A. *Fitting Semiparametric Symmetric Regression Models*. 2014.
- J. M.C and A. B.C. Distributions that are both log-symmetric and r-symmetric. *Electronic Journal of Statistics*, 2:1300–1208, 2008.
- V. G. Myers R.H, Montgomery D.C and R. T.J. *Generalized Linear Models with applications in engineering and sciences*. John Wiley and Sons, INC., PUBLICATION, second edition, 2010.
- K. D. Ng H.K.T and B. N. Modified moment estimation for the two-parameter birnbaum-saunders distribution. *Computational Statistics and Data Analysis*, 43:283–298, 2003.
- B. M. Paula G.A, Leiva V. and L. S. Robust statistical modeling using birnbaum-saunders-t distribution applied to insurance. *Applied Stochastic Models in Business and Industry*, 28:16–34, 2012.
- G. P.J and S. B.W. *Nonparametric regression and generalized linear models*. Chapman and Hall, 1994.
- W. S. *Generalized Additive Models: an introduction with R*. Chapman and Hall-RC, 2006.
- R. L. Vanegas L.H and C. F.J. Diagnostic procedures in birnbaum-saunders nonlinear regression models. *Reporte interno de investigación, Universidad Nacional de Colombia*, pages 1–20, February 2011.
- L. Victor. *The Birnbaum-Saunders distribution*. Academic Press is an imprint of Elsevier, 2016.
- M. W.Q and E. L. A. *Statistical Methods for Reliability Data*. New York: Wiley, 1998.
- B. Z.W and S. S.C. A new family of distributions. *Journal of Applied Probability*, 6(2):319–327, August 1969.