# PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

## ESCUELA DE POSGRADO

# MODELO DE REGRESIÓN LINEAL CON CENSURA BASADO EN UNA DISTRIBUCIÓN SENH-NORMAL/INDEPENDIENTE: UNA PERSPECTIVA FRECUENTISTA

### TESIS PARA OPTAR POR EL GRADO ACADÉMICO DE MAGÍSTER EN ESTADÍSTICA


**AUTOR:**

Max Walter Alonzo Huaman

**ASESOR:**

Luis Enrique Benites Sánchez


**MIEMBROS DEL JURADO:**

Dr. Cristian Luis Bayes Rodríguez

Dr. Luis Enrique Benites Sánchez

Dra. Zaida Jesus Quiroz Cornejo

Lima, Mayo 2021

# Resumen

En esta tesis se estudia el modelo de regresión lineal para datos censurados considerando una distribución senh-normal/independiente para los errores desde un enfoque frecuentista. Este trabajo considera la revisión de la teoría existente, la construcción del nuevo modelo, estimación de parámetros, estudios de simulación para recuperar los parámetros del modelo y la aplicación a un conjunto de datos reales.

**Palabras-clave:** senh-normal/independiente, censura, inferencia frecuentista.

# Abstract

In this thesis, the linear regression model for censored data is studied considering a sinh-normal / independent distribution for errors from a frequentist approach. This paper considers the revision of the existing theory, the construction of the new model, estimation of parameters, simulation studies to retrieve the parameters of the model and the application to a set of real data.

**Keywords:** sinh-normal/independent, censor, frequentist inference.

# Dedicatoria

Dedicado a mis hijos (Luciana, Diego y Camila), esposa Yesi, hermanos (María, Clara, Jhon, Libia, Mirtha y George) y mis padres (Esteban y Genoveva)

# Agradecimentos

A todas las personas que contribuyeron con su granito de arena en la culminación de la tesis, en especial a mi familia, asesor de tesis, compañeros de la maestría y amigos.

# Índice general

# Índice de figuras

# Índice de cuadros

# Capítulo 1

# Introducción

## 1.1. Consideraciones Preliminares

Un supuesto del modelo de regresión lineal es que el error sigue una distribución normal. Sin embargo, se sabe que existen fenómenos que no satisfacen esta suposición. Johnson (1949) mencionó que es natural y conveniente construir distribuciones no normales transformando una variable aleatoria normal. Algunos ejemplos son las distribuciones Birnbaum-Saunders (BS) (Birnbaum y Saunders, 1969), Senh-Normal (SN) (Rieck y Nedelman, 1991) y log-normal (Johnson y Kotz, 1970).

La distribución BS surgió en el contexto de ingeniería de materiales, ya que la distribución es apropiada para describir procesos de degradación acumulativa, actualmente la distribución ha venido siendo aplicado en otras áreas del conocimiento, para más detalles de las aplicaciones se puede ver Leiva et al. (2008) y Leiva et al. (2010) donde podrán ser encontrados algunos ejemplos de aplicaciones (o referencias bibliográficas) en ciencias de la salud, ambiental, forestal, demográficas, actuarial, financiera, entre otras.

Lemonte y Cordeiro (2009) propusieron un modelo de regresión no lineal BS, extendiendo el trabajo de Rieck y Nedelman (1991). Luego, Paula et al. (2011) realizaron un ajuste robusto y un análisis de diagnóstico para las distribuciones BS, log-BS y generalizaciones de la distribución de t-student, y desarrollaron el modelo de regresión BS-t de student. A su vez, Vilca et al. (2015) siguiendo las ideas presentadas en el articulo de Lemonte y Cordeiro (2009) propusieron un modelo de regresión no lineal basado en la distribucion Senh-Normal/Independiente (Senh-NI) y en Vilca et al. (2017) desarrollaron una alternativa bayesiana de dicho modelo. La ventaja de este modelo de regresión propuesto por Vilca et al. (2017) es su capacidad de trabajar con conjuntos de datos que presentan observaciones atípicas, esto se debe a que la distribución Senh-NI tiene colas más pesadas que con la distribución SN. Sin embargo, estos modelos de regresión no pueden ser utilizados cuando se tienen observaciones censuradas.

Por otro lado, se han realizado estudios considerando modelos de regresión basados en extensiones de la distribución log-BS para datos con observaciones censuradas. Por ejemplo Barros et al. (2008), propusieron una nueva clase de modelos de regresión para datos de tiempos de vida, donde los errores siguen una distribución log-BS-t-Student (log-BST). Los autores discutieron y realizaron un análisis de diagnóstico desde una perspectiva frecuentista. Además, utilizaron un conjunto de datos con observaciones censuradas y realizaron un análisis de sensibilidad para comparar aspectos de robustez de los estimadores de los parámetros de los modelos log-BS y log-BS de regresión. Leiva et al. (2007) realizó un análisis de diagnóstico para un modelo de regresión log-BS para datos censurados considerando un tipo de censura no informativa. Luego, realizaron un análisis de diagnóstico de influencia para el modelo de regresión log-BS para datos censurados basado en el modelo de eliminación de casos, una prueba de hipótesis para observaciones atípicas y discutieron una prueba de score para la homegenidad del parámetro de forma. Posteriormente, Mikhail et al. (2013) desarrollaron una prueba de bondad de ajuste Chi-cuadrado modificada basada en las estadísticas de Nikulin-Rao-Robson para la familia de distribuciones de Birnbaum-Saunders generalizada para datos censurados por la derecha con parámetros desconocidos utilizando la estimación por máxima verosimilitud. En los últimos años Lachos et al. (2017), realizaron un enfoque bayesiano de un modelo de regresión con datos censurados de la clase de mixtura de escala normal usaron el algoritmo de muestreo de Gibbs y Metropolis-Hastings en la que obtuvieron las estimaciones de los parámetros. Además, presentaron algunas discusiones sobre la selección de modelos para comparar los modelos ajustados y desarrollaron un análisis de diagnóstico de influencia considerando la eliminación de casos para la distribución a posteriori conjunta basada en la divergencia de Kullback-Leibler.

En este trabajo, tomando como referencia las artículos Leiva et al. (2007) y Vilca et al. (2015), se desarrollará un modelo de regresión basado en la distribución Senh-NI para datos censurados considerando un tipo de censura no informativa desde un enfoque frecuentista. La ventaja de esta distribución simétrica es que posee colas pesadas permitiendo modelar datos con observaciones atípicas.

## 1.2. Objetivos de la tesis

El objetivo general de esta tesis es proponer un modelo de regresión lineal censurado usando la distribución Senh-NI desde una perspectiva frecuentista, incluyendo la metodología para la estimación de parámetros y su aplicación a conjuntos de datos reales. De manera específica:

- Proponer e implementar en R un método para estimar los parámetros del modelo basado en optimización numérica.
- Desarrollar estudios de simulación que permitan ilustrar el desempeño del modelo propuesto.
- Aplicar el modelo a un conjunto de datos reales que permitan ilustrar la metodología utilizada.

## 1.3. Organización del Trabajo

En el Capítulo 2, se presentan los conceptos preliminares asociados al modelo de regresión Senh-Normal/Independiente, así se revisa la literatura acerca de distribuciones Birbaum-Saunders, Senh-Normal, Senh-Normal/Independiente, regresión lineal, censura no informativa y criterios de selección de modelos. En el Capítulo 3 se estudia el modelo propuesto, modelo de regresión lineal con censura basado en una distribución Senh-Normal/Independiente desde una perspectiva frecuentista, considerando el algoritmo para la estimación de sus parámetros. En el Capítulo 4 se presentan los resultados del estudio de simulación para la recuperación de parámetros, estudiar la precisión de las estimaciones de los parámetros, consistencia de las estimaciones de los errores estándar y propiedad de consistencia de los estimadores. En el Capítulo 5 se muestra la aplicación del modelo a un conjunto de datos reales. En el Capítulo 6 se presentan las conclusiones producto del estudio realizado. También se incluye un apéndice donde se realizan demostraciones matemáticas y el código R utilizado.

# Capítulo 2

# Conceptos Preliminares

## 2.1. Distribución Birnbaum-Saunders

Birnbaum y Saunders (1969) desarrollaron una importante distribución de tiempos de vida a partir de un problema de fatiga de materiales. Esta distribución posee muchas características y propiedades atractivas. Tal como una relación funcional muy simple entre la BS y las distribuciones normales. Por estas razones, la distribución BS se ha convertido en un modelo comúnmente utilizado en diferentes áreas de la ciencia y la ingeniería, ver, por ejemplo Leiva et al. (2007). Birnbaum y Saunders (1969) mostraron que sobre ciertas condiciones la función de distribución acumulada (fda) de una variable aleatoria $T$ real no negativa, que representa el tiempo total hasta que ocurra una falla es dada por

$$F_T(t;\alpha,\beta) = P(T \leq t) = \Phi\left[\frac{1}{\alpha}\left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right)\right], \quad t > 0, \tag{2.1}$$

donde se dice que $T$ tiene una distribución BS con parámetro de localización $\alpha > 0$ y parámetro de escala $\beta > 0$ denotada por $T \sim \mathrm{BS}(\alpha,\beta)$, $\Phi(\cdot)$ es la función de distribución acumulada de una distribución normal. Una propiedad importante en la construcción de esta distribución es que la variable aleatoria

$$Z = \frac{1}{\alpha}\left(\sqrt{\frac{T}{\beta}} - \sqrt{\frac{\beta}{T}}\right) \sim \mathrm{N}(0,1).$$

Así la distribución BS está estrechamente relacionada con la distribución normal a través de la relación

$$T = \beta\left[\frac{\alpha Z}{2} + \sqrt{\left(\frac{\alpha Z}{2}\right)^2 + 1}\right]^2,$$

donde $Z \sim \mathrm{N}(0,1)$. Considerando la fda de la variable aleatoria $T$ dada en (2.1), su correspondiente función de densidad de probabilidad (fdp) es dada por

$$f_T(t) = \phi(a_t(\alpha,\beta))A_t(\alpha,\beta),$$

donde $\phi(\cdot)$ es la fdp de la distribución normal estándar, $a_t(\alpha,\beta)$ y $A_t(\alpha,\beta)$ son dadas respectivamente por

$$a_t(\alpha,\beta)=\frac{1}{\alpha}\left(\sqrt{\frac{t}{\beta}}-\sqrt{\frac{\beta}{t}}\right) \quad \text{y} \quad A_t(\alpha,\beta)=\frac{d}{dt}a_t(\alpha,\beta)=\frac{t^{-3/2}(t+\beta)}{2\alpha\beta^{1/2}}.$$

La distribución BS presenta las siguientes propiedades. Si $T\sim \mathrm{BS}(\alpha,\beta)$ entonces tenemos:

(i) $cT\sim \mathrm{BS}(\alpha,c\beta), c>0$;

(ii) $\dfrac{1}{T}\sim \mathrm{BS}\left(\alpha,\frac{1}{\beta}\right)$.

El valor esperado y la varianza de la distribución BS son:

(i) $\mathrm{E}(T)=\beta\left(1+\frac{\alpha^2}{2}\right)$,

(ii) $\mathrm{Var}(T)=(\alpha\beta)^2\left(1+\frac{5\alpha^2}{4}\right)$.

## 2.2. Distribución Senh-Normal

De acuerdo con Rieck (1989), una variable aleatoria $Y$ que tiene una distribución SN está relacionada con un modelo normal a través de la representación estocástica $Y=\mu+\sigma\,\mathrm{arcsinh}(\alpha Z/2)$, donde $Z\sim\ \mathrm{N}(0,1)$, $\alpha>0$ es el parámetro de forma, $\mu\in\mathbb{R}$ es el parámetro de localización, y $\sigma>0$ es el parámetro de escala. En este caso la notación es $Y\sim\ \mathrm{SN}(\alpha,\mu,\sigma)$. La fdp de $Y$ es simétrica alrededor de $\mu$, unimodal si $\alpha\leq 2$ y bimodal si $\alpha>2$. La fdp de $Y$ presentada en (2.2) es dado por:

$$f_Y(y)=\frac{2}{\alpha\sigma\sqrt{2\pi}}\cosh\left(\frac{y-\mu}{\sigma}\right)\exp\left(-\frac{2}{\alpha^2}\,\mathrm{senh}^2\left[\frac{y-\mu}{\sigma}\right]\right), y\in\mathbb{R}, \tag{2.2}$$

donde $\mu=\log(\beta)$. En este caso decimos que $Y$ tiene una distribución Senh-Normal (SN), la cual es denotada por $\mathrm{SN}(\alpha,\mu,\sigma)$. Sin embargo la fdp mostrado en (2.2) se puede re-expresar como:

$$f_Y(y)=\frac{1}{\sigma}\phi(\xi_{2y})\xi_{1y},$$

donde $\phi(\cdot)$ es la fdp de la distribución normal estándar, $\xi_{1y}$ y $\xi_{2y}$ son dadas por

$$\begin{aligned}\xi_{1y}&=\xi_1(y;\alpha,\mu,\sigma)=\frac{2}{\alpha}\cosh\left(\frac{y-\mu}{\sigma}\right),\\ \xi_{2y}&=\xi_2(y;\alpha,\mu,\sigma)=\frac{2}{\alpha}\,\mathrm{senh}\left(\frac{y-\mu}{\sigma}\right).\end{aligned} \tag{2.3}$$

Mientras que la fda de la distribución SN se expresa de la siguiente forma:

$$F_Y(y)=\Phi\left(\frac{2}{\alpha}\mathrm{senh}\left(\frac{y-\mu}{\sigma}\right)\right),$$

donde $\Phi(\cdot)$ es la fda de la distribución normal estándar. Algunas propiedades de la distribución SN son:

(i) $Z = \frac{2}{\alpha}\text{senh}\Big(\frac{Y-\mu}{\sigma}\Big)$,

(ii) $a + bY \sim \text{SN}(\alpha, a + b\mu, |b|\sigma)$, para todo $a, b \in \mathbb{R}$ y $b \neq 0$.

El valor esperado y la varianza de la distribución SN son:

(i) $\text{E}(Y) = \mu$,

(ii) $\text{Var}(Y) = \sigma^2\eta(\alpha)$, con $\eta(\alpha) = (\alpha^2/4)(1 - \alpha^2/4)$.

## 2.3. Distribución Senh-Normal/Independiente

Vilca et al. (2015) siguiendo las ideas de Balakrishnan et al. (2009), extendieron la distribución SN reemplazando la distribución $\text{N}(0,1)$ por la distribución normal/independiente (NI). Consideraron la variable aleatoria $Z = U^{-1/2}\,Z_0$ donde $Z_0 \sim \text{N}(0,1)$, en lugar de la distribución normal estándar, entonces $Y$ puede ser representada estocásticamente en términos de $Z$ como

$$Y = \mu + \sigma\,\text{arcsinh}\Big(\frac{\alpha\,Z}{2}\Big) = \mu + \sigma\,\text{arcsinh}\Big(\frac{\alpha\,U^{-1/2}Z_0}{2}\Big). \tag{2.4}$$

En este caso $Y$ tiene una distribución Senh-Normal/Independiente (Senh-NI) con parámetro de forma $\alpha > 0$, parámetro de localización $\mu$ y parámetro de escala $\sigma > 0$. La notación $Y \sim \text{Senh-NI}(\alpha, \mu, \sigma = 2; H)$ es usada para esta distribución, $H$ es la función de distribución acumulada indexada a la variable $U$. Cuando $U = 1$, la distribución Senh-NI se reduce a la distribución SN (Rieck, 1989). Usando el teorema de cambio de variable, la función de densidad de probabilidad (fdp) de $Y$ es

$$f_Y(y) \;=\; \frac{1}{\sigma}\,\phi_{\text{NI}}(\xi_{2y})\,\xi_{1y} = \frac{1}{\sigma}\int_0^\infty \phi\big(\xi_{2y}; 0, u^{-1}\big)\,dH(u)\;\xi_{1y}, \quad y \in \mathbb{R}, \tag{2.5}$$

donde $\phi_{\text{NI}}(\cdot)$ es la fdp de $Z$, $H(\cdot)$ es la función de densidad acumulada de $U$ (indexada por el escalar o vector de parámetros $\boldsymbol{\nu}$), $\xi_{1y} = \xi_1(y; \alpha, \mu, \sigma) = \frac{2}{\alpha}\cosh\left(\frac{y-\mu}{\sigma}\right)$, $\xi_{2y} = \xi_2(y; \alpha, \mu, \sigma) = \frac{2}{\alpha}\,\text{senh}\left(\frac{y-\mu}{\sigma}\right)$ y $\phi(\cdot; 0, u^{-1})$ denota la función de densidad $\text{N}(0, u^{-1})$.

La fda de la distribución Senh-NI viene dada por

$$F_Y(y) \;=\; \Phi_{\text{NI}}\left(\frac{2}{\alpha}\,\text{senh}\left(\frac{y-\mu}{\sigma}\right)\right). \tag{2.6}$$

Al igual que la distribución SN, si $Y \sim \text{Senh-NI}(\alpha, \mu, \sigma; H)$, entonces $a{+}bY \sim \text{Senh} - \text{NI}(\alpha, \text{a}{+}\text{b}\mu, |\text{b}|\sigma)$, para todo $a, b \in \mathbb{R}$ y $b \neq 0$. Este resultado muestra que la distribución Senh-NI es cerrada bajo cualquier transformación lineal. La demostración se muestra en el anexo A.5. La función generatriz de momentos de $Y$ queda expresada de la siguiente forma:

$$M_Y(s) = \frac{\exp(s\mu)}{\alpha} \int_0^{\infty} u^{1/2} \frac{\exp(u/\alpha^2)}{\sqrt{2\pi}} \left[\mathbb{K}_{\frac{\sigma s-1}{2}}\left(\frac{u}{\alpha^2}\right) + \mathbb{K}_{\frac{\sigma s+1}{2}}\left(\frac{u}{\alpha^2}\right)\right] dH(u),$$

donde $\mathbb{K}_v(.)$ es la función Bessel modificadas de tercera especie. De esta forma los momentos de $Y$ pueden obtenerse usando métodos de aproximación para algunos casos especiales.

Balakrishnan et al. (2009) muestra algunos casos especiales de la familia Senh-NI considerando como base las distribuciones normal contaminada (CN), slash (SL) y $t$-Student (St). Las distribuciones resultantes son las llamadas Senh-normal contaminada (Senh-NC), Senh-slash (Senh-SL) y Senh-$t$-Student (Senh-St).

Para obtener la función de densidad de la distribución Senh-NC se considera que $U$ tiene una fdp dada por $h_U(u) = v\mathbb{I}_{\{\gamma\}}(u) + (1-v)\mathbb{I}_{\{1\}}(u)$, $0<v<1,0<\gamma < 1, \mathbb{I}_A(.)$ es una función indicadora del conjunto $A$ y $\nu = (v, \gamma)^{\tau}$, entonces la fdp de $Y$ queda de la forma que sigue

$$f_Y(y) = (1/\sigma)[v\sqrt{\gamma}\phi(\sqrt{\gamma}\xi_{2y}) + (1-v)\phi(\xi_{2y})]\xi_{1y}, \quad y \in \mathbb{R}.$$

Para la distribución Senh-SL, $U \sim \text{Beta}(\nu, 1)$ y entonces la fdp de $Y$ es

$$f_Y(y) = (1/\sigma)[v \int_0^1 u^{v-1}\phi(\xi_{2y}; 0, u^{-1})du]\xi_{1y}, \quad y \in \mathbb{R}.$$

Finalmente, para el caso de la distribución Senh-St, se considera $U \sim \text{Gamma}(\nu/2, \nu/2)$, entonces la fdp de $Y$ es

$$f_Y(y) = k(v)v^{v/2}/\sigma(v + \xi_{2y}^2)^{-((v+1)/2)}\xi_{1y}, \quad y \in \mathbb{R}, \quad k(v) = \Gamma((v+1)/2)/[\sqrt{\pi}\Gamma(v/2)].$$

Para $\alpha \leq 2$ la Senh-NI es unimodal, mientras que si $\alpha > 2$ tiende a ser bimodal. La Figura 2.1 muestra densidades de la SN y Senh-NI para $\alpha = 1{,}5, \alpha = 4, \mu = 0$ y $\sigma = 2$, donde también se aprecia que las distribuciones Senh-St ($\nu = 3$), Senh-NC ($\nu = 0{,}1$ y $\gamma = 0{,}1$), Senh-SL ($\nu = 3$) tienen colas más pesadas si las comparamos con la distribución SN. Desde que el modelo de regresión es nuestro principal interés, en el Teorema 1.1 de Rieck y Nedelman (1991) los autores demostraron que si $T \sim \text{BS}(\alpha, \beta)$ entonces $Y = \log(T) \sim \text{SN}(\alpha, \mu, \sigma = 2)$, debido a este resultado la distribución SN también es llamada distribución log-BS. Por lo tanto siguiendo el resultado de Rieck y Nedelman (1991) se considerará $\sigma = 2$ para el desarrollo de la presente tesis.

## 2.4. Censura no Informativa

Según Kleinbaum y Klein (2011), la censura se da cuando el evento de interés no tiene lugar durante la duración del estudio, así por ejemplo la muerte de un paciente después del tiempo observado, más aún, sea cuando el individuo no presenta el evento de interés durante el tiempo de duración del estudio. De acuerdo a Bogaerts et al. (2017), los supuestos que se

(a)

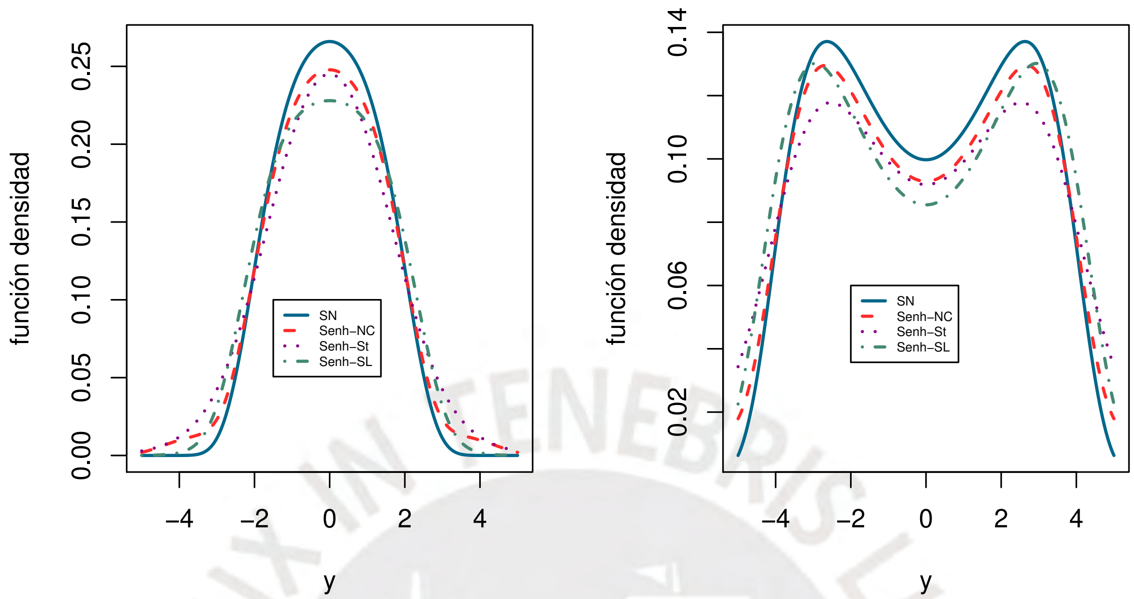


Figura 2.1: Gráfico de funciones de densidades para las distribuciones SN, Senh-NC, Senh-St y Senh-SL con $\alpha = 1{,}5$ (izquierda), $\alpha = 4$ (derecha), $\mu = 0$ y $\sigma = 2$. Con Senh-St ($\nu = 3$), Senh-NC ($\nu = 0{,}1$ y $\gamma = 0{,}1$), Senh-SL ($\nu = 3$)

usan cuando se trata con censuras son: censura independiente, se da cuando los tiempos de censura son independientes para cada entidad; censura aleatoria, donde los tiempos de censura no dependen ni obedecen a ningún comportamiento; y la censura no informativa, donde la distribución de los tiempos de supervivencia no proporciona información sobre la distribución de los tiempos de censura, y viceversa. Según Klein y Moeschberger (2003) si asumimos censura aleatoria hacia la derecha y tomamos una muestra de los tiempos de sobrevivencia independiente e idénticamente distribuidas (i.i.d), tal que $(y_1, \delta_1), \ldots, (y_n, \delta_n)$, donde $y_i = \min(t_i, c_i)$ es el mínimo entre el tiempo de supervivencia y el tiempo de censura, y donde $\delta_i = I(t_i \leq c_i)$ es el indicador del evento. Entonces, se define $T$ como la variable aleatoria del tiempo del evento con fdp $f(\cdot)$ y función de supervivencia $S(\cdot)$, mientras que $C$ es la variable aleatoria del tiempo de censura con fdp $g(\cdot)$ y supervivencia $G(\cdot)$. Asumiendo independencia entre $T$ y $C$, tenemos que

$$Pr[T = y_i, C > y_i] = G(y_i)f(y_i).$$

Del mismo modo, la contribución a la función de verosimilitud de los datos censurados $(y_i, 0)$ se puede expresar como

$$Pr[C = y_i, T > y_i] = S(y_i)g(y_i).$$

Por lo tanto, la función de verosimilitud para los datos completos se puede escribir como

$$L = \prod_{i=1}^{n} [G(y_i)f(y_i)]^{\delta_i} [S(y_i)g(y_i))]^{1-\delta_i}.$$

Ahora, supongamos que la distribución de $C$ no depende de los parámetros de la distribución de $T$. Entonces, los factores $G(y_i)^{\delta_i} g(y_i)^{1-\delta_i}$ son no informativos y se pueden eliminar:

$$L \propto \prod_{i=1}^{n} [f(y_i)]^{\delta_i} [S(y_i))]^{1-\delta_i}.$$

Esta es la verosimilitud habitual cuando se trata de datos de supervivencia. En términos generales, la independencia entre $T$ y $C$ le permite separar la contribución conjunta de $T$ y $C$ en sus contribuciones marginales, mientras que la suposición de censura no informativa podemos desconsiderar $g(\cdot)$ y $G(\cdot)$.

## 2.5. Criterio de Selección de Modelos

Los criterios que serán usados para seleccionar modelos son el Criterio de Información de Akaike (AIC) desarrollado por Akaike (1974) y el Criterio de Información Bayesiana (BIC) desarrollado por Schwarz (1978). El AIC evalúa el ajuste del modelo y el número de parámetros, por ello el modelo será más penalizado a medida que tenga un menor ajuste y mayor número de parámetros. Por tal, el modelo a seleccionar debe ser aquel que ofrece el menor AIC, queda definido como

$$AIC = 2k - 2\ell(\widehat{\theta}),$$

donde $k$ es el número de parámetros del modelo y $\ell(\theta)$ es el logaritmo de la verosimilitud del modelo. El criterio BIC funciona de manera similar al AIC, con la diferencia que penaliza el número de variables utilizando el $\log(n)$ en lugar de usar el valor fijo 2, donde $n$ es el tamaño de la muestra, así:

$$BIC = \log(n)k - 2\ell(\widehat{\theta}).$$

## Capítulo 3

# Modelo de regresión lineal con censura basado en una distribución Senh-Normal/Independiente

### 3.1. Modelo de regresión lineal con censura basado en Senh-Normal/ Independiente

Se considera el modelo de regresión lineal Senh-Normal/Independiente (Senh-NI-ML) siguiendo una propuesta similar desarrollada por Vilca et al. (2015). Así, tenemos el modelo de regresión lineal

$$Y_i = \mathbf{x}_i^\top \boldsymbol{\beta} + \varepsilon_i, \; i = 1, \ldots, n, \tag{3.1}$$

donde $Y_i$ es el $i$-ésimo individuo, $\mathbf{x}_i$ es un vector $m \times 1$ de variables explicativas conocidas asociadas con la $i$-ésima respuesta observable de $Y_i$ y $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_m)^\top$ es un vector de parámetros lineales desconocidos. Además, se asume que

$$\varepsilon_i \sim \text{Senh} - \text{NI}(\alpha, 0, 2; H), \; i = 1, \ldots, n. \tag{3.2}$$

Por la linearidad de la distribución Senh-NI, tenemos que $Y_i \sim \text{Senh-NI}(\alpha, \mathbf{x}_i^\top \boldsymbol{\beta}; H)$ y de (2.5) tenemos que su función de densidad es

$$f_{Y_i}(y_i) \;\; = \;\; \frac{1}{2}\, \phi_{\text{NI}}(\xi_{2i})\, \xi_{1i}, \tag{3.3}$$

donde $\xi_{1i} = \xi_1(y_i; \alpha, \boldsymbol{\beta})$ y $\xi_{2i} = \xi_2(y_i; \alpha, \boldsymbol{\beta})$ están definidos en (2.3), con $\sigma = 2$ y $\mu_i = \mathbf{x}_i^\top \boldsymbol{\beta}$ (para no recargar la notación aquí se usa $\xi_{1i}$ y $\xi_{2i}$ en lugar de $\xi_{1y_i}$ y $\xi_{2y_i}$, respectivamente). Se asume censura no informativa y que los tiempos de vida observados y los tiempos censurados son independientes. Las letras $D$ y $C$ denotan conjuntos de individuos para los cuales $y_i$ es el logaritmo de tiempos de vida o logaritmo de tiempo de censura, respectivamente. Cabe mencionar que $\boldsymbol{\nu}$ asociada a la distribución acumulada $H(\cdot)$ mostrada en la Subsección 2.3 será considera como un valor fijo y no como un parámetro a estimar, debido a los problemas de convergencia que se tenía al momento de considerarlo como parámetro en el modelo.

## 3.2. Estimación de los parámetros

La función de verosimilitud del modelo dado en (3.1) para $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \alpha)^\top$, siguiendo las ideas de Vilca et al. (2015) consideraremos el parámetro $\boldsymbol{\nu}$ como conocido, es

$$\ell(\boldsymbol{\theta}) \quad \propto \quad \sum_{i \in D} \ell_i(\boldsymbol{\theta}) + \sum_{i \in C} \ell_{ic}(\boldsymbol{\theta}),$$

donde $\ell_i(\boldsymbol{\theta}) = \log f_Y(y_i; \boldsymbol{\theta})$, $\ell_{ic}(\boldsymbol{\theta}) = \log S_Y(y_i; \boldsymbol{\theta})$ y $S_Y(y_i; \boldsymbol{\theta}) = 1 - F_Y(y; \boldsymbol{\theta})$ es la función de supervivencia, con $f_Y(y_i; \boldsymbol{\theta})$ dado en (3.3) y $F_Y(y; \boldsymbol{\theta})$ en (2.6). Entonces, la función de verosimilitud para $\boldsymbol{\theta}$ puede ser expresada como

$$\ell(\boldsymbol{\theta}) \quad \propto \quad \sum_{i \in D} [\log \phi_{\mathrm{NI}}(\xi_{2i}) + \log(\xi_{1i})] + \sum_{i \in C} \log(1 - \Phi_{\mathrm{NI}}(\xi_{2i})).$$

Los elementos de la función score, cuya demostración se encuentra en el anexo A.3, dados por:

$$\begin{aligned} U_\alpha(\boldsymbol{\theta}) &= \frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha} = \frac{1}{\alpha}\left[\sum_{i \in D}\left(\frac{\xi_{2i}^2 A_i(3/2)}{A_i(1/2)} - 1\right) + \sum_{i \in C} \xi_{2i} h(\xi_{2i})\right], \\ U_{\boldsymbol{\beta}}(\boldsymbol{\theta}) &= \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} = \frac{1}{2}\left[\sum_{i \in D} \xi_{2i}\left(\frac{\xi_{1i} A_i(3/2)}{A_i(1/2)} - \frac{1}{\xi_{1i}}\right)\frac{\partial \mu_i}{\partial \boldsymbol{\beta}} + \sum_{i \in C}\left(\xi_{1i} h(\xi_{2i}) \frac{\partial \mu_i}{\partial \boldsymbol{\beta}}\right)\right], \end{aligned}$$

donde $\phi_{\mathrm{NI}}(\cdot)$ y $\Phi_{\mathrm{NI}}(\cdot)$ denotan la fdp y la fda de la distribución NI, respectivamente, mientras que $h(\xi_{2i})$ y $A_i(w)$ tienen las siguientes expresiones

$$\begin{aligned} h(\xi_{2i}) &= \frac{\phi_{\mathrm{NI}}(\xi_{2i})}{1 - \Phi_{\mathrm{NI}}(\xi_{2i})}, \\ A_i(w) &= \int_0^\infty u^w exp\left(\frac{-u}{2}\xi_{2i}^2\right) dH(u). \end{aligned} \tag{3.4}$$

La estimación por máxima verosimilitud de los coeficientes y el parámetro de forma son soluciones de las ecuaciones $U_\alpha(\boldsymbol{\theta}) = 0$ y $U_{\boldsymbol{\beta}}(\boldsymbol{\theta}) = 0$. Sin embargo, tales ecuaciones no presentan soluciones analíticas y por tal es necesario el uso de métodos iterativos para encontrar las soluciones. El algoritmo Broyden–Fletcher–Goldfarb–Shanno (BFGS), Press et al. (1992), será usado para maximizar la función de log-verosimilitud $\ell(\boldsymbol{\theta})$. Como valor inicial para $\alpha$ se considera el estimador obtenido bajo el caso no censurado y para $\boldsymbol{\beta}$ los estimadores por mínimos cuadrados ordinarios. Algunos casos especiales del modelo de regresión propuesto son como sigue:

- Para $U = 1$ obtenemos el modelo de regresión de Leiva et al. (2007)
- Para $U \sim \text{Gamma}(\nu/2, \nu/2)$ obtenemos el modelo de regresión de Paula et al. (2011)
- Para $U = 1$ y $C = \varnothing$ las ecuaciones (3.1) y (3.2) nos llevan al modelo de Rieck y

Nedelman (1991).

- Cuando $U \sim \text{Gamma}(\nu/2, \nu/2)$ obtendremos el modelo Senh-St (Senh-t-student).
- Considerando $U \sim \text{Beta}(\nu, 1)$ obtendremos el modelo Senh-NC (Senh-slash)
- Se tendrá que si la fdp de $U$ es $h_U(u) = \nu\mathcal{I}_{\{\gamma\}}(u) + (1-\nu)\mathcal{I}_{\{1\}}(u)$, $0 < \nu < 1$, $0 < \gamma < 1$, $\mathcal{I}_{\{A\}}(\cdot)$ denota la función indicada del conjunto $A$ y $\boldsymbol{\nu} = (\nu, \gamma)^\top$ obtendremos el modelo Senh-NC (Senh-normal contaminada).

A partir de (2.4) se puede obtener la distribución de $d_i = \frac{4}{\alpha^2}\,\text{senh}^2\left(\frac{Y_i-\mu_i}{2}\right)$. Esta distribución es útil para pruebas de bondad de ajuste y detección de valores atípicos; ver Lange y Sinsheimer (1993). Por ejemplo, bajo la distribución SN, $d_i \sim \chi_1^2$, y la distribución Senh-St, $d_i \sim F(1, \nu)$.

## 3.3. Matriz de información observada

En cuanto a la matriz de información observada del modelo Senh-NI-ML, definida como $\mathbf{J}_o(\boldsymbol{\theta}|\mathbf{y}) = -\partial^2\ell(\boldsymbol{\theta}|\mathbf{y})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top$. Es bien conocido que bajo algunas condiciones de regularidad, la matriz de covarianza de los estimados de máxima verosimilitud $\widehat{\boldsymbol{\theta}}$ puede ser aproximada por la inversa de $\mathbf{J}_o(\boldsymbol{\theta}|\mathbf{y})$. Siguiendo Basford et al. (1997) y Lin et al. (2007), se puede estimar $\mathbf{J}_o(\boldsymbol{\theta}|\mathbf{y})$ por

$$\mathbf{J}_o(\widehat{\boldsymbol{\theta}}|\mathbf{y}) = \sum_{i=1}^{n} \widehat{\mathbf{s}}_i \widehat{\mathbf{s}}_i^\top, \tag{3.5}$$

donde $\widehat{\mathbf{s}}_i = \partial\ell(\boldsymbol{\theta}|\mathbf{y})/\partial\boldsymbol{\theta}\big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}$. Se considera el vector $\widehat{\mathbf{s}}_i$ particionado en términos de componentes de los parámetros $\boldsymbol{\theta}$ como $\widehat{\mathbf{s}}_i = (\widehat{s}_{i,\alpha}, \widehat{s}_{i,\boldsymbol{\beta}})^\top$. La demostración se encuentra en el anexo A.4., los resultados son

$$\widehat{s}_{i,\alpha} = \begin{cases} \dfrac{1}{\alpha}\left[\dfrac{\xi_{2i}^2 A_i(3/2)}{A_i(1/2)} - 1\right] & \text{si} \quad i \in D \\ \left(\dfrac{1}{\alpha\sqrt{2\pi}}\right)\dfrac{A_i(1/2)\xi_{2i}}{1-\Phi_{\text{NI}}(\xi_{2i})} & \text{si} \quad i \in C \end{cases}$$

$$\widehat{s}_{i,\boldsymbol{\beta}} = \begin{cases} \dfrac{\xi_{2i}}{4\sqrt{2\pi}}\dfrac{\partial\mu_i}{\partial\boldsymbol{\beta}}\left(\xi_{i1}^2 A_i(3/2) - A_i(1/2)\right) & \text{si} \quad i \in D \\ \dfrac{1}{2}\dfrac{\partial\mu_i}{\partial\boldsymbol{\beta}}\dfrac{\xi_{i1}\phi_{\text{NI}}(\xi_{2i})}{1-\Phi_{\text{NI}}(\xi_{2i})} & \text{si} \quad i \in C \end{cases}$$

donde $A_i(w)$ es como en (3.4). A continuación se presenta las expresiones de $A_i(w)$ para algunos casos de la familia Senh-NI:

(i) *Senh-St:* $A_i(w) = \dfrac{2^w \nu^{\nu/2}\Gamma(w+\nu/2)}{\Gamma(\nu/2)(\nu+\xi_{2i}^2)^{\nu/2+w}}$,

(ii) *Senh-SL:* $A_i(w) = \dfrac{\nu 2^{w+\nu}\Gamma(w+\nu)}{\xi_{2i}^{2(w+\nu)}} P_1\left(w+\nu, \dfrac{\xi_{2i}^2}{2}\right)$,

(iii) *Senh-NC:* $A_i(w) = \sqrt{2\pi}\,\{\nu\,\gamma^{w-1/2}\,\phi(\xi_{2i}; 0, \gamma^{-1}) + (1-\nu)\,\phi(\xi_{2i}; 0, 1)\}$.

# Capítulo 4

# Estudio de Simulación

En este capítulo se realiza un estudio de simulación a fin de evaluar el rendimiento y propiedades del modelo de regresión lineal Senh-NI-ML considerando los casos especiales Senh-NC, Senh-St y Senh-SL. Se detalla aspectos relacionados con la generación de los datos artificiales, recuperación de parámetros mediante métodos de optimización (función optim de R) aplicadas a la función de log-verosimilitud, análisis del impacto para diferentes niveles de censura y diversos tamaños de muestra.

## 4.1. Consideraciones para la simulación

Para todos los escenarios se generaron 1000 réplicas de Monte Carlo, se consideró una covariable y el vector de parámetros $\boldsymbol{\beta} = (4, -2)^\top$ se mantuvo fijo durante todas las simulaciones. También se asume $\nu = 3$ para los casos Senh-St y Senh-SL, mientras que $\boldsymbol{\nu} = (0{,}1, 0{,}1)^\top$ para el caso Senh-NC. Todos los procedimientos computacionales fueron implementados en el software R. La simulación fue generada siguiendo el modelo descrito en (4.1) y donde $Y_i$ es el logaritmo del tiempo de vida observado o censurado para el $i$-ésimo individuo.

$$Y_i = \beta_0 + \beta_1 \mathbf{x}_i + \varepsilon_i,\ i = 1, \ldots, n.$$

Además, se asume que $\varepsilon_i \sim \mathrm{Senh} - \mathrm{NI}(\alpha, 0, \sigma = 2, \boldsymbol{\nu})$. La covariable $x_i$, fue generado considerando una distribución uniforme entre 0 y 1. Para verificar si las estimaciones logran aproximarse al verdadero valor de los parámetros ($\alpha$ y $\boldsymbol{\beta}$), se generó datos simulando diferentes tamaños de muestra $n = (100, 500, 1000)$, diversos valores del parámetro $\alpha = (0{,}5, 1{,}0, 1{,}5)^\top$ , varios niveles de censura (0, 0.1, 0.2, 0.4) y casos especiales del modelo (Senh-NC, Senh-St y Senh-SL). Para recuperar los parámetros del modelo $\alpha$ y $\boldsymbol{\beta}$ se hizo uso de la función optim, del software R, aplicada a la función de log-verosimilitud y con valores iniciales descritos en la Sección 3.2. Seguidamente para obtener los errores estándar de cada parámetro se generó la matriz de información observada (según sección 3.3) y a continuación se determinó la raíz cuadrada de la diagonal de la inversa de dicha matriz para así obtener los errores estándar.

## 4.2. Resultados

Se calculó los valores promedio de las estimaciones que resultaron muy cercanos a los valores reales de los parámetros $\alpha$ y $\boldsymbol{\beta}$, así también se calculó la desviación estándar de las estimaciones (DEE) y la media de los errores estándar de las estimaciones (MEE). Los resultados se muestran en el Cuadro 4.1, donde se verifica que, para todos los parámetros, a medida que aumenta el nivel de censura se empiezan a sobrestimar los parámetros $\alpha$ y $\boldsymbol{\beta}$, incluso independiente del tamaño de muestra y caso especial del modelo. Si comparamos los resultados obtenidos con Senh-NC, Senh-St y Senh-SL, tenemos que el modelo Senh-SL produce estimaciones más robustas para el parámetro $\alpha$ frente a variaciones de nivel de censura, dado que presenta DEE e MEE más pequeños. El modelo Senh-NC presenta estimaciones más robustas para los parámetros $\beta_0$ y $\beta_1$. Sin embargo, a medida que se incrementa el valor de $\alpha$, la Senh-SL tiende a ser más robusta para todos los parámetros. También se calculó la cobertura para el intervalo definido por:

$$\mathrm{COB}(\hat{\theta}) = \frac{1}{1000}\sum_{\mathrm{j}=1}^{1000} \mathrm{I}(\theta \in [\hat{\theta}_{\mathrm{L_j}}, \hat{\theta}_{\mathrm{U_j}}]),$$

donde $I$ es una función indicadora tal que $\theta \in [\hat{\theta}_{Lj}, \hat{\theta}_{Uj}]$ con $\hat{\theta}_{L_j}$ y $\hat{\theta}_{U_j}$ que representan el límite inferior y límite superior del intervalo de confianza para $j$-ésima simulación. Para comparar el efecto de diferentes niveles de censura 0 %, 10 %, 20 % y 40 % se calculó el sesgo, el sesgo relativo (Sesgo relativo) y la raíz del error cuadrático medio (RECM) para cada uno de los estimadores y los diferentes escenarios, los cuales se definen a continuación,

$$\mathrm{Sesgo}(\hat{\theta}) = \frac{1}{1000}\sum_{\mathrm{j}=1}^{1000}(\hat{\theta}_{\mathrm{j}} - \theta), \;\; \mathrm{Sesgo-relativo}(\hat{\theta}) = \frac{\mathrm{Sesgo}(\hat{\theta})}{\theta},$$

$$\mathrm{RECM}(\hat{\theta}) = \sqrt{\frac{1}{1000}\sum_{\mathrm{j}=1}^{1000}(\hat{\theta}_{\mathrm{j}} - \theta)^2},$$

donde $\hat{\theta}_j$ es el estimado para la $j$-ésima muestra y $\theta$ es el verdadero valor del parámetro en estudio. Los resultados se muestran en el Cuadro 4.2 donde se aprecia que un incremento en el nivel de censura, genera un incremento de Sesgo, Sesgo relativo y del RECM para todos los casos (Senh-NC, Senh-St y Senh-SL) e independiente del tamaño de muestra y del valor del parámetro $\alpha$. Este comportamiento es aceptable dado que conforme se incrementa el nivel de censura se reduce la información disponible. El efecto de los diferentes tamaños de muestra en las propiedades de los estimadores (sesgo y RECM), se observa en el Cuadro 4.2, donde se observa que cuando el tamaño de muestra se incrementa, el RECM disminuye incluso independiente del nivel de censura. Se espera que el RECM tienda a cero a medida que se incremente el tamaño de muestra. Este comportamiento se da tanto para los modelos Senh-NC, Senh-St y Senh-SL.

Cuadro 4.1: Promedio de los estimadores, DEE, MEE y cobertura al 95 % para diferentes niveles de censura

| | | | | Senh-NC | | | | Senh-St | | | | Senh-SL | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\alpha$ | n | Parámetro | Medida | 0 % | 10 % | 20 % | 40 % | 0 % | 10 % | 20 % | 40 % | 0 % | 10 % | 20 % | 40 % |
| 0.50 | 100 | $\alpha = 0.5$ | Promedio | 0.4974 | 0.5056 | 0.5133 | 0.5391 | 0.4931 | 0.4999 | 0.5067 | 0.5254 | 0.4941 | 0.4955 | 0.4988 | 0.5066 |
| | | | DEE | 0.0397 | 0.0446 | 0.0505 | 0.0657 | 0.0499 | 0.0521 | 0.0555 | 0.0651 | 0.0375 | 0.0405 | 0.0437 | 0.0537 |
| | | | MEE | 0.0429 | 0.0456 | 0.0488 | 0.0584 | 0.0486 | 0.0516 | 0.0551 | 0.0645 | 0.0394 | 0.0415 | 0.0442 | 0.0517 |
| | | | 95 % COB | 94.2 | 93.6 | 91.5 | 85.0 | 93.0 | 92.9 | 93.9 | 91.7 | 94.5 | 93.4 | 92.5 | 90.7 |
| | | $\beta_0 = 4$ | Promedio | 3.9942 | 3.9948 | 4.0010 | 4.0083 | 4.0013 | 4.0026 | 4.0047 | 4.0109 | 4.0011 | 3.9998 | 3.9969 | 3.9979 |
| | | | DEE | 0.1117 | 0.1171 | 0.1258 | 0.1487 | 0.1206 | 0.1251 | 0.1336 | 0.1570 | 0.1159 | 0.1228 | 0.1303 | 0.1483 |
| | | | MEE | 0.1075 | 0.1131 | 0.1198 | 0.1391 | 0.1167 | 0.1235 | 0.1309 | 0.1510 | 0.1144 | 0.1197 | 0.1262 | 0.1447 |
| | | | 95 % COB | 91.4 | 91.8 | 91.0 | 90.7 | 92.1 | 92.6 | 92.1 | 91.2 | 93.2 | 92.1 | 91.6 | 91.7 |
| | | $\beta_1 = -2$ | Promedio | -1.9888 | -1.9750 | -1.9703 | -1.9478 | -1.9962 | -1.9842 | -1.9734 | -1.9427 | -1.9993 | -1.9818 | -1.9578 | -1.9128 |
| | | | DEE | 0.1989 | 0.2110 | 0.2243 | 0.2599 | 0.2060 | 0.2143 | 0.2278 | 0.2649 | 0.2076 | 0.2197 | 0.2313 | 0.2565 |
| | | | MEE | 0.1853 | 0.1953 | 0.2071 | 0.2414 | 0.2024 | 0.2141 | 0.2271 | 0.2620 | 0.1977 | 0.2072 | 0.2191 | 0.2516 |
| | | | 95 % COB | 92.7 | 90.4 | 91.0 | 91.0 | 92.2 | 94.3 | 93.3 | 92.0 | 92.1 | 91.8 | 92.4 | 92.0 |
| | 500 | $\alpha = 0.5$ | Promedio | 0.5030 | 0.5098 | 0.5188 | 0.5444 | 0.4995 | 0.5065 | 0.5145 | 0.5356 | 0.4991 | 0.5021 | 0.5042 | 0.5120 |
| | | | DEE | 0.0188 | 0.0215 | 0.0233 | 0.0301 | 0.0222 | 0.0237 | 0.0250 | 0.0295 | 0.0176 | 0.0194 | 0.0216 | 0.0269 |
| | | | MEE | 0.0191 | 0.0199 | 0.0211 | 0.0247 | 0.0222 | 0.0236 | 0.0251 | 0.0295 | 0.0173 | 0.0182 | 0.0193 | 0.0221 |
| | | | 95 % COB | 94.3 | 89.3 | 82.2 | 56.7 | 95.4 | 93.2 | 90.4 | 78.0 | 94.0 | 92.0 | 89.2 | 85.1 |
| | | $\beta_0 = 4$ | Promedio | 3.9940 | 3.9989 | 4.0043 | 4.0134 | 4.0007 | 4.0017 | 4.0044 | 4.0093 | 3.9994 | 3.9963 | 3.9967 | 3.9913 |
| | | | DEE | 0.0500 | 0.0534 | 0.0579 | 0.0668 | 0.0537 | 0.0554 | 0.0585 | 0.0669 | 0.0532 | 0.0575 | 0.0616 | 0.0677 |
| | | | MEE | 0.0489 | 0.0515 | 0.0546 | 0.0627 | 0.0541 | 0.0572 | 0.0608 | 0.0705 | 0.0517 | 0.0542 | 0.0570 | 0.0646 |
| | | | 95 % COB | 94.6 | 93.4 | 93.2 | 92.3 | 95.1 | 95.1 | 95.3 | 95.4 | 93.8 | 93.1 | 92.6 | 92.5 |
| | | $\beta_1 = -2$ | Promedio | -1.9919 | -1.9852 | -1.9773 | -1.9530 | -2.0029 | -1.9906 | -1.9783 | -1.9458 | -1.9997 | -1.9749 | -1.9586 | -1.9039 |
| | | | DEE | 0.0860 | 0.0898 | 0.0959 | 0.1103 | 0.0918 | 0.0952 | 0.0991 | 0.1108 | 0.0932 | 0.1019 | 0.1107 | 0.1220 |
| | | | MEE | 0.0847 | 0.0893 | 0.0947 | 0.1086 | 0.0937 | 0.0991 | 0.1054 | 0.1221 | 0.0895 | 0.0938 | 0.0987 | 0.1116 |
| | | | 95 % COB | 94.6 | 95.3 | 93.8 | 92.1 | 95.3 | 95.8 | 95.4 | 94.3 | 93.7 | 91.9 | 89.7 | 82.3 |
| | 1000 | $\alpha = 0.5$ | Promedio | 0.5028 | 0.5103 | 0.5198 | 0.5461 | 0.4996 | 0.5067 | 0.5145 | 0.5358 | 0.4996 | 0.5022 | 0.5048 | 0.5120 |
| | | | DEE | 0.0132 | 0.0149 | 0.0171 | 0.0216 | 0.0158 | 0.0165 | 0.0177 | 0.0209 | 0.0123 | 0.0142 | 0.0161 | 0.0202 |
| | | | MEE | 0.0134 | 0.0139 | 0.0146 | 0.0168 | 0.0158 | 0.0167 | 0.0178 | 0.0209 | 0.0122 | 0.0129 | 0.0136 | 0.0155 |
| | | | 95 % COB | 94.4 | 84.1 | 68.8 | 28.7 | 94.7 | 93.2 | 88.3 | 59.6 | 94.4 | 92.3 | 87.9 | 78.6 |
| | | $\beta_0 = 4$ | Promedio | 3.9984 | 4.0018 | 4.0063 | 4.0152 | 4.0036 | 4.0046 | 4.0071 | 4.0125 | 4.0020 | 3.9996 | 3.9987 | 3.9961 |
| | | | DEE | 0.0348 | 0.0359 | 0.0373 | 0.0449 | 0.0381 | 0.0401 | 0.0428 | 0.0486 | 0.0370 | 0.0396 | 0.0432 | 0.0487 |
| | | | MEE | 0.0348 | 0.0365 | 0.0386 | 0.0442 | 0.0384 | 0.0406 | 0.0432 | 0.0501 | 0.0367 | 0.0385 | 0.0405 | 0.0459 |
| | | | 95 % COB | 94.3 | 94.8 | 95.3 | 93.0 | 94.0 | 94.7 | 94.5 | 95.5 | 94.8 | 94.1 | 93.4 | 93.1 |
| | | $\beta_1 = -2$ | Promedio | -1.9982 | -1.9880 | -1.9801 | -1.9564 | -2.0063 | -1.9951 | -1.9833 | -1.9517 | -2.0033 | -1.9815 | -1.9619 | -1.9130 |
| | | | DEE | 0.0605 | 0.0621 | 0.0662 | 0.0764 | 0.0662 | 0.0684 | 0.0722 | 0.0799 | 0.0651 | 0.0714 | 0.0804 | 0.0910 |
| | | | MEE | 0.0601 | 0.0632 | 0.0669 | 0.0764 | 0.0665 | 0.0704 | 0.0749 | 0.0868 | 0.0635 | 0.0666 | 0.0700 | 0.0789 |
| | | | 95 % COB | 94.5 | 94.5 | 94.1 | 91.0 | 95.1 | 94.9 | 95.5 | 92.5 | 94.2 | 93.4 | 88.3 | 76.3 |

Cuadro 4.2: Promedio (Promedio) de los estimadores, DEE, MEE y cobertura al 95 % para diferentes niveles de censura

| | | | | Senh-NC | | | | Senh-St | | | | Senh-SL | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\alpha$ | n | Parámetro | Medida | 0 % | 10 % | 20 % | 40 % | 0 % | 10 % | 20 % | 40 % | 0 % | 10 % | 20 % | 40 % |
| 1.00 | 100 | $\alpha = 1.0$ | Promedio | 0.9853 | 0.9971 | 1.0115 | 1.0469 | 0.9859 | 0.9986 | 1.0105 | 1.0441 | 0.9875 | 0.9942 | 1.0043 | 1.0335 |
| | | | DEE | 0.0819 | 0.0898 | 0.0959 | 0.1177 | 0.0996 | 0.1035 | 0.1096 | 0.1285 | 0.0751 | 0.0810 | 0.0884 | 0.1066 |
| | | | MEE | 0.0872 | 0.0926 | 0.0991 | 0.1178 | 0.0980 | 0.1041 | 0.1112 | 0.1305 | 0.0804 | 0.0849 | 0.0913 | 0.1106 |
| | | | 95 % COB | 94.0 | 93.9 | 93.5 | 91.1 | 92.9 | 93.4 | 94.4 | 93.0 | 94.7 | 94.2 | 92.9 | 91.7 |
| | | $\beta_0 = 4$ | Promedio | 3.9945 | 3.9972 | 4.0049 | 4.0296 | 3.9980 | 4.0063 | 4.0178 | 4.0490 | 4.0012 | 4.0087 | 4.0083 | 4.0291 |
| | | | DEE | 0.2109 | 0.2210 | 0.2311 | 0.2623 | 0.2328 | 0.2391 | 0.2527 | 0.2846 | 0.2096 | 0.2187 | 0.2255 | 0.2405 |
| | | | MEE | 0.2032 | 0.2132 | 0.2251 | 0.2593 | 0.2259 | 0.2384 | 0.2522 | 0.2892 | 0.2105 | 0.2211 | 0.2345 | 0.2764 |
| | | | 95 % COB | 91.0 | 91.1 | 91.8 | 91.5 | 92.1 | 93.3 | 93.0 | 92.6 | 92.8 | 93.6 | 93.2 | 94.6 |
| | | $\beta_1 =$ -2 | Promedio | -1.9891 | -1.9593 | -1.9353 | -1.8771 | -1.9853 | -1.9629 | -1.9426 | -1.8903 | -1.9976 | -1.9462 | -1.8751 | -1.7482 |
| | | | DEE | 0.3748 | 0.3952 | 0.4115 | 0.4521 | 0.3983 | 0.4106 | 0.4307 | 0.4851 | 0.3733 | 0.3941 | 0.3994 | 0.4084 |
| | | | MEE | 0.3506 | 0.3686 | 0.3892 | 0.4507 | 0.3916 | 0.4130 | 0.4375 | 0.5013 | 0.3646 | 0.3836 | 0.4088 | 0.4860 |
| | | | 95 % COB | 91.6 | 91.0 | 91.5 | 92.7 | 93.1 | 93.8 | 94.6 | 92.5 | 92.3 | 92.8 | 92.7 | 93.9 |
| | 500 | $\alpha = 1.0$ | Promedio | 0.9969 | 1.0085 | 1.0220 | 1.0576 | 0.9991 | 1.0119 | 1.0263 | 1.0635 | 0.9981 | 1.0057 | 1.0160 | 1.0436 |
| | | | DEE | 0.0388 | 0.0412 | 0.0435 | 0.0521 | 0.0445 | 0.0471 | 0.0494 | 0.0576 | 0.0352 | 0.0389 | 0.0428 | 0.0525 |
| | | | MEE | 0.0385 | 0.0407 | 0.0433 | 0.0506 | 0.0445 | 0.0472 | 0.0503 | 0.0589 | 0.0348 | 0.0367 | 0.0387 | 0.0450 |
| | | | 95 % COB | 94.5 | 93.4 | 91.5 | 77.9 | 95.4 | 94.0 | 91.4 | 82.0 | 94.1 | 91.7 | 88.4 | 79.1 |
| | | $\beta_0 = 4$ | Promedio | 3.9951 | 4.0012 | 4.0109 | 4.0308 | 4.0018 | 4.0091 | 4.0206 | 4.0483 | 3.9989 | 4.0032 | 4.0093 | 4.0145 |
| | | | DEE | 0.0953 | 0.0989 | 0.1048 | 0.1169 | 0.1034 | 0.1058 | 0.1100 | 0.1210 | 0.0955 | 0.1022 | 0.1048 | 0.1107 |
| | | | MEE | 0.0923 | 0.0969 | 0.1022 | 0.1161 | 0.1045 | 0.1100 | 0.1165 | 0.1336 | 0.0933 | 0.0974 | 0.1021 | 0.1148 |
| | | | 95 % COB | 93.9 | 94.1 | 93.7 | 93.2 | 95.0 | 95.2 | 95.5 | 94.4 | 93.7 | 92.3 | 93.1 | 95.2 |
| | | $\beta_1 =$ -2 | Promedio | -1.9909 | -1.9681 | -1.9454 | -1.8799 | -2.0063 | -1.9823 | -1.9592 | -1.9027 | -1.9991 | -1.9397 | -1.8769 | -1.7267 |
| | | | DEE | 0.1608 | 0.1652 | 0.1730 | 0.1929 | 0.1762 | 0.1812 | 0.1860 | 0.2012 | 0.1671 | 0.1820 | 0.1915 | 0.1927 |
| | | | MEE | 0.1595 | 0.1674 | 0.1767 | 0.2003 | 0.1810 | 0.1906 | 0.2019 | 0.2312 | 0.1616 | 0.1686 | 0.1758 | 0.1969 |
| | | | 95 % COB | 94.8 | 94.8 | 94.5 | 91.3 | 95.6 | 95.6 | 96.4 | 95.2 | 94.6 | 91.6 | 86.3 | 71.1 |
| | 1000 | $\alpha = 1.0$ | Promedio | 0.9985 | 1.0102 | 1.0242 | 1.0595 | 0.9991 | 1.0120 | 1.0259 | 1.0634 | 0.9991 | 1.0065 | 1.0176 | 1.0455 |
| | | | DEE | 0.0274 | 0.0294 | 0.0317 | 0.0383 | 0.0315 | 0.0327 | 0.0348 | 0.0409 | 0.0246 | 0.0281 | 0.0328 | 0.0386 |
| | | | MEE | 0.0272 | 0.0287 | 0.0305 | 0.0354 | 0.0316 | 0.0335 | 0.0357 | 0.0417 | 0.0245 | 0.0257 | 0.0272 | 0.0312 |
| | | | 95 % COB | 94.7 | 91.7 | 84.7 | 60.6 | 94.6 | 93.9 | 89.6 | 66.4 | 94.4 | 92.1 | 84.2 | 64.7 |
| | | $\beta_0 = 4$ | Promedio | 3.9988 | 4.0048 | 4.0137 | 4.0345 | 4.0073 | 4.0148 | 4.0262 | 4.0537 | 4.0035 | 4.0083 | 4.0106 | 4.0184 |
| | | | DEE | 0.0645 | 0.0662 | 0.0679 | 0.0773 | 0.0733 | 0.0759 | 0.0804 | 0.0882 | 0.0662 | 0.0700 | 0.0772 | 0.0825 |
| | | | MEE | 0.0653 | 0.0686 | 0.0723 | 0.0820 | 0.0742 | 0.0782 | 0.0828 | 0.0949 | 0.0662 | 0.0691 | 0.0722 | 0.0805 |
| | | | 95 % COB | 94.3 | 95.4 | 95.4 | 93.9 | 94.2 | 94.5 | 94.1 | 92.6 | 94.7 | 94.7 | 92.3 | 92.6 |
| | | $\beta_1 =$ -2 | Promedio | -1.9969 | -1.9731 | -1.9494 | -1.8879 | -2.0128 | -1.9908 | -1.9687 | -1.9120 | -2.0054 | -1.9491 | -1.8769 | -1.7304 |
| | | | DEE | 0.1131 | 0.1165 | 0.1199 | 0.1321 | 0.1273 | 0.1298 | 0.1359 | 0.1462 | 0.1171 | 0.1297 | 0.1402 | 0.1457 |
| | | | MEE | 0.1129 | 0.1184 | 0.1247 | 0.1410 | 0.1285 | 0.1354 | 0.1434 | 0.1641 | 0.1145 | 0.1192 | 0.1238 | 0.1366 |
| | | | 95 % COB | 93.8 | 94.5 | 93.6 | 88.7 | 94.8 | 95.5 | 95.7 | 94.2 | 94.6 | 90.9 | 78.8 | 48.9 |

Cuadro 4.3: Promedio (Promedio) de los estimadores, DEE, MEE y cobertura al 95 % para diferentes niveles de censura

| | | | | Senh-NC | | | | Senh-St | | | | Senh-SL | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\alpha$ | n | Parámetro | Medida | 0 % | 10 % | 20 % | 40 % | 0 % | 10 % | 20 % | 40 % | 0 % | 10 % | 20 % | 40 % |
| 1.50 | 100 | $\alpha = 1.5$ | Promedio | 1.4782 | 1.4940 | 1.5131 | 1.5583 | 1.4789 | 1.4978 | 1.5148 | 1.5653 | 1.4809 | 1.5007 | 1.5253 | 1.5831 |
| | | | DEE | 0.1239 | 0.1357 | 0.1448 | 0.1749 | 0.1506 | 0.1562 | 0.1661 | 0.1946 | 0.1130 | 0.1234 | 0.1351 | 0.1670 |
| | | | MEE | 0.1327 | 0.1408 | 0.1508 | 0.1793 | 0.1490 | 0.1584 | 0.1695 | 0.2000 | 0.1228 | 0.1306 | 0.1417 | 0.1769 |
| | | | 95 % COB | 94.1 | 94.1 | 93.4 | 92.1 | 92.7 | 93.6 | 94.4 | 92.9 | 95.2 | 94.0 | 93.1 | 89.8 |
| | | $\beta_0 = 4$ | Promedio | 3.9934 | 4.0071 | 4.0257 | 4.0776 | 3.9903 | 4.0105 | 4.0366 | 4.1013 | 4.0011 | 4.0298 | 4.0567 | 4.1332 |
| | | | DEE | 0.2875 | 0.2988 | 0.3086 | 0.3443 | 0.3227 | 0.3314 | 0.3459 | 0.3817 | 0.2719 | 0.2794 | 0.2831 | 0.2865 |
| | | | MEE | 0.2776 | 0.2910 | 0.3063 | 0.3534 | 0.3155 | 0.3321 | 0.3512 | 0.4021 | 0.2782 | 0.2933 | 0.3127 | 0.3820 |
| | | | 95 % COB | 91.2 | 91.1 | 91.6 | 91.4 | 92.5 | 92.9 | 92.5 | 92.6 | 93.1 | 93.7 | 93.1 | 94.6 |
| | | $\beta_1 = -2$ | Promedio | -1.9876 | -1.9642 | -1.9424 | -1.8848 | -1.9691 | -1.9502 | -1.9374 | -1.8936 | -1.9965 | -1.9432 | -1.8814 | -1.7842 |
| | | | DEE | 0.5062 | 0.5364 | 0.5560 | 0.6155 | 0.5541 | 0.5715 | 0.5980 | 0.6654 | 0.4809 | 0.5028 | 0.5077 | 0.5136 |
| | | | MEE | 0.4796 | 0.5032 | 0.5300 | 0.6149 | 0.5468 | 0.5750 | 0.6087 | 0.6959 | 0.4834 | 0.5112 | 0.5481 | 0.6770 |
| | | | 95 % COB | 92.3 | 90.7 | 91.1 | 92.5 | 93.0 | 93.6 | 93.9 | 92.9 | 93.0 | 93.4 | 94.2 | 96.3 |
| | 500 | $\alpha = 1.5$ | Promedio | 1.4954 | 1.5107 | 1.5283 | 1.5741 | 1.4987 | 1.5175 | 1.5388 | 1.5936 | 1.4972 | 1.5174 | 1.5417 | 1.6013 |
| | | | DEE | 0.0583 | 0.0614 | 0.0644 | 0.0770 | 0.0668 | 0.0706 | 0.0740 | 0.0865 | 0.0528 | 0.0600 | 0.0665 | 0.0814 |
| | | | MEE | 0.0580 | 0.0614 | 0.0654 | 0.0764 | 0.0670 | 0.0710 | 0.0758 | 0.0888 | 0.0526 | 0.0553 | 0.0588 | 0.0695 |
| | | | 95 % COB | 94.6 | 94.3 | 92.7 | 83.2 | 95.5 | 94.0 | 91.9 | 82.7 | 94.2 | 90.6 | 84.1 | 67.5 |
| | | $\beta_0 = 4$ | Promedio | 3.9932 | 4.0102 | 4.0312 | 4.0769 | 4.0028 | 4.0210 | 4.0456 | 4.1049 | 3.9986 | 4.0246 | 4.0533 | 4.1228 |
| | | | DEE | 0.1283 | 0.1315 | 0.1373 | 0.1483 | 0.1427 | 0.1452 | 0.1496 | 0.1615 | 0.1222 | 0.1282 | 0.1284 | 0.1345 |
| | | | MEE | 0.1245 | 0.1304 | 0.1372 | 0.1547 | 0.1447 | 0.1518 | 0.1600 | 0.1815 | 0.1200 | 0.1244 | 0.1292 | 0.1441 |
| | | | 95 % COB | 93.0 | 93.7 | 93.2 | 90.8 | 94.8 | 95.3 | 94.8 | 91.9 | 93.2 | 92.1 | 91.9 | 86.2 |
| | | $\beta_1 = -2$ | Promedio | -1.9875 | -1.9689 | -1.9491 | -1.8836 | -2.0092 | -1.9879 | -1.9683 | -1.9195 | -1.9984 | -1.9420 | -1.8819 | -1.7678 |
| | | | DEE | 0.2166 | 0.2232 | 0.2320 | 0.2545 | 0.2430 | 0.2500 | 0.2568 | 0.2755 | 0.2138 | 0.2301 | 0.2317 | 0.2472 |
| | | | MEE | 0.2151 | 0.2251 | 0.2370 | 0.2665 | 0.2506 | 0.2628 | 0.2772 | 0.3139 | 0.2079 | 0.2151 | 0.2226 | 0.2478 |
| | | | 95 % COB | 94.6 | 95.0 | 95.0 | 93.5 | 95.4 | 95.2 | 96.3 | 96.5 | 93.6 | 92.6 | 89.6 | 83.0 |
| | 1000 | $\alpha = 1.5$ | Promedio | 1.4977 | 1.5131 | 1.5315 | 1.5773 | 1.4988 | 1.5178 | 1.5382 | 1.5929 | 1.4987 | 1.5194 | 1.5454 | 1.6065 |
| | | | DEE | 0.0412 | 0.0439 | 0.0472 | 0.0565 | 0.0473 | 0.0491 | 0.0520 | 0.0613 | 0.0369 | 0.0431 | 0.0488 | 0.0603 |
| | | | MEE | 0.0409 | 0.0431 | 0.0459 | 0.0534 | 0.0474 | 0.0503 | 0.0536 | 0.0627 | 0.0369 | 0.0387 | 0.0410 | 0.0479 |
| | | | 95 % COB | 94.8 | 92.5 | 87.4 | 67.6 | 94.7 | 93.8 | 90.9 | 67.9 | 94.5 | 88.9 | 75.6 | 43.8 |
| | | $\beta_0 = 4$ | Promedio | 3.9984 | 4.0147 | 4.0350 | 4.0824 | 4.0105 | 4.0291 | 4.0539 | 4.1120 | 4.0041 | 4.0289 | 4.0534 | 4.1293 |
| | | | DEE | 0.0868 | 0.0885 | 0.0897 | 0.0980 | 0.1008 | 0.1034 | 0.1086 | 0.1166 | 0.0845 | 0.0908 | 0.0916 | 0.0924 |
| | | | MEE | 0.0879 | 0.0920 | 0.0966 | 0.1087 | 0.1027 | 0.1078 | 0.1137 | 0.1286 | 0.0850 | 0.0880 | 0.0908 | 0.0988 |
| | | | 95 % COB | 94.4 | 93.8 | 93.3 | 89.3 | 94.8 | 94.5 | 93.6 | 87.9 | 95.0 | 93.2 | 89.3 | 73.4 |
| | | $\beta_1 = -2$ | Promedio | -1.9961 | -1.9759 | -1.9551 | -1.8960 | -2.0184 | -2.0000 | -1.9819 | -1.9313 | -2.0060 | -1.9482 | -1.8769 | -1.7741 |
| | | | DEE | 0.1526 | 0.1571 | 0.1618 | 0.1770 | 0.1751 | 0.1780 | 0.1863 | 0.1983 | 0.1500 | 0.1669 | 0.1708 | 0.1759 |
| | | | MEE | 0.1518 | 0.1588 | 0.1665 | 0.1867 | 0.1778 | 0.1866 | 0.1967 | 0.2223 | 0.1471 | 0.1515 | 0.1554 | 0.1672 |
| | | | 95 % COB | 94.3 | 94.5 | 94.5 | 92.5 | 94.7 | 95.2 | 96.0 | 96.3 | 94.8 | 91.0 | 84.9 | 72.8 |

Cuadro 4.4: Sesgo, Sesgo relativo y RECM para diferentes niveles de censura

| | | | | Senh-NC | | | | Senh-St | | | | Senh-SL | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\alpha$ | n | Parámetro | Valores | 0 % | 10 % | 20 % | 40 % | 0 % | 10 % | 20 % | 40 % | 0 % | 10 % | 20 % | 40 % |
| 1.00 | 100 | $\alpha = 1.0$ | Sesgo | -0.0147 | -0.0029 | 0.0115 | 0.0469 | -0.0141 | -0.0014 | 0.0105 | 0.0441 | -0.0125 | -0.0058 | 0.0043 | 0.0335 |
| | | | RECM | 0.0831 | 0.0898 | 0.0965 | 0.1267 | 0.1006 | 0.1035 | 0.1101 | 0.1358 | 0.0761 | 0.0812 | 0.0884 | 0.1116 |
| | | | Sesgo relat. % | -1.5 | -0.3 | 1.2 | 4.7 | -1.4 | -0.1 | 1.1 | 4.4 | -1.3 | -0.6 | 0.4 | 3.4 |
| | | $\beta_0 = 4$ | Sesgo | -0.0055 | -0.0028 | 0.0049 | 0.0296 | -0.0020 | 0.0063 | 0.0178 | 0.0490 | 0.0012 | 0.0087 | 0.0083 | 0.0291 |
| | | | RECM | 0.2109 | 0.2209 | 0.2311 | 0.2638 | 0.2326 | 0.2391 | 0.2533 | 0.2887 | 0.2095 | 0.2188 | 0.2256 | 0.2422 |
| | | | Sesgo relat. % | -0.1 | -0.1 | 0.1 | 0.7 | -0.1 | 0.2 | 0.4 | 1.2 | 0.0 | 0.2 | 0.2 | 0.7 |
| | | $\beta_1 = -2$ | Sesgo | 0.0109 | 0.0407 | 0.0647 | 0.1229 | 0.0147 | 0.0371 | 0.0574 | 0.1097 | 0.0024 | 0.0538 | 0.1249 | 0.2518 |
| | | | RECM | 0.3748 | 0.3971 | 0.4163 | 0.4683 | 0.3984 | 0.4120 | 0.4343 | 0.4971 | 0.3731 | 0.3976 | 0.4182 | 0.4796 |
| | | | Sesgo relat. % | -0.5 | -2.0 | -3.2 | -6.1 | -0.7 | -1.9 | -2.9 | -5.5 | -0.1 | -2.7 | -6.2 | -12.6 |
| | 500 | $\alpha = 1.0$ | Sesgo | -0.0031 | 0.0085 | 0.0220 | 0.0576 | -0.0009 | 0.0119 | 0.0263 | 0.0635 | -0.0019 | 0.0057 | 0.0160 | 0.0436 |
| | | | RECM | 0.0389 | 0.0420 | 0.0487 | 0.0776 | 0.0445 | 0.0486 | 0.0559 | 0.0857 | 0.0352 | 0.0393 | 0.0456 | 0.0682 |
| | | | Sesgo relat. % | -0.3 | 0.9 | 2.2 | 5.8 | -0.1 | 1.2 | 2.6 | 6.4 | -0.2 | 0.6 | 1.6 | 4.4 |
| | | $\beta_0 = 4$ | Sesgo | -0.0049 | 0.0012 | 0.0109 | 0.0308 | 0.0018 | 0.0091 | 0.0206 | 0.0483 | -0.0011 | 0.0032 | 0.0093 | 0.0145 |
| | | | RECM | 0.0954 | 0.0988 | 0.1053 | 0.1208 | 0.1034 | 0.1062 | 0.1118 | 0.1302 | 0.0954 | 0.1022 | 0.1052 | 0.1116 |
| | | | Sesgo relat. % | -0.1 | 0.0 | 0.3 | 0.8 | 0.0 | 0.2 | 0.5 | 1.2 | 0.0 | 0.1 | 0.2 | 0.4 |
| | | $\beta_1 = -2$ | Sesgo | 0.0091 | 0.0319 | 0.0546 | 0.1201 | -0.0063 | 0.0177 | 0.0408 | 0.0973 | 0.0009 | 0.0603 | 0.1231 | 0.2733 |
| | | | RECM | 0.1610 | 0.1682 | 0.1814 | 0.2272 | 0.1762 | 0.1820 | 0.1903 | 0.2234 | 0.1670 | 0.1916 | 0.2276 | 0.3344 |
| | | | Sesgo relat. % | -0.5 | -1.6 | -2.7 | -6.0 | 0.3 | -0.9 | -2.0 | -4.9 | 0.0 | -3.0 | -6.2 | -13.7 |
| | 1000 | $\alpha = 1.0$ | Sesgo | -0.0015 | 0.0102 | 0.0242 | 0.0595 | -0.0009 | 0.0120 | 0.0259 | 0.0634 | -0.0009 | 0.0065 | 0.0176 | 0.0455 |
| | | | RECM | 0.0275 | 0.0311 | 0.0399 | 0.0708 | 0.0315 | 0.0348 | 0.0434 | 0.0754 | 0.0246 | 0.0289 | 0.0372 | 0.0596 |
| | | | Sesgo relat. % | -0.2 | 1.0 | 2.4 | 6.0 | -0.1 | 1.2 | 2.6 | 6.3 | -0.1 | 0.7 | 1.8 | 4.6 |
| | | $\beta_0 = 4$ | Sesgo | -0.0012 | 0.0048 | 0.0137 | 0.0345 | 0.0073 | 0.0148 | 0.0262 | 0.0537 | 0.0035 | 0.0083 | 0.0106 | 0.0184 |
| | | | RECM | 0.0645 | 0.0664 | 0.0692 | 0.0846 | 0.0736 | 0.0773 | 0.0845 | 0.1032 | 0.0662 | 0.0705 | 0.0779 | 0.0845 |
| | | | Sesgo relat. % | 0.0 | 0.1 | 0.3 | 0.9 | 0.2 | 0.4 | 0.7 | 1.3 | 0.1 | 0.2 | 0.3 | 0.5 |
| | | $\beta_1 = -2$ | Sesgo | 0.0031 | 0.0269 | 0.0506 | 0.1121 | -0.0128 | 0.0092 | 0.0313 | 0.0880 | -0.0054 | 0.0509 | 0.1231 | 0.2696 |
| | | | RECM | 0.1131 | 0.1195 | 0.1301 | 0.1732 | 0.1278 | 0.1300 | 0.1394 | 0.1706 | 0.1172 | 0.1393 | 0.1865 | 0.3064 |
| | | | Sesgo relat. % | -0.2 | -1.3 | -2.5 | -5.6 | 0.6 | -0.5 | -1.6 | -4.4 | 0.3 | -2.5 | -6.2 | -13.5 |

Cuadro 4.5: Sesgo, Sesgo relativo y RECM para diferentes niveles de censura

| | | | | Senh-NC | | | | Senh-St | | | | Senh-SL | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\alpha$ | n | Parámetro | Valores | 0 % | 10 % | 20 % | 40 % | 0 % | 10 % | 20 % | 40 % | 0 % | 10 % | 20 % | 40 % |
| 1.50 | 100 | $\alpha = 1.5$ | Sesgo | -0.0218 | -0.0060 | 0.0131 | 0.0583 | -0.0211 | -0.0022 | 0.0148 | 0.0653 | -0.0191 | 0.0007 | 0.0253 | 0.0831 |
| | | | RECM | 0.1257 | 0.1357 | 0.1453 | 0.1843 | 0.1520 | 0.1561 | 0.1667 | 0.2052 | 0.1146 | 0.1233 | 0.1374 | 0.1865 |
| | | | Sesgo relat. % | -1.5 | -0.4 | 0.9 | 3.9 | -1.4 | -0.1 | 1.0 | 4.4 | -1.3 | 0.0 | 1.7 | 5.5 |
| | | $\beta_0 = 4$ | Sesgo | -0.0066 | 0.0071 | 0.0257 | 0.0776 | -0.0097 | 0.0105 | 0.0366 | 0.1013 | 0.0011 | 0.0298 | 0.0567 | 0.1332 |
| | | | RECM | 0.2874 | 0.2987 | 0.3095 | 0.3527 | 0.3227 | 0.3314 | 0.3476 | 0.3948 | 0.2718 | 0.2809 | 0.2886 | 0.3158 |
| | | | Sesgo relat. % | -0.2 | 0.2 | 0.6 | 1.9 | -0.2 | 0.3 | 0.9 | 2.5 | 0.0 | 0.7 | 1.4 | 3.3 |
| | | $\beta_1 = -2$ | Sesgo | 0.0124 | 0.0358 | 0.0576 | 0.1152 | 0.0309 | 0.0498 | 0.0626 | 0.1064 | 0.0035 | 0.0568 | 0.1186 | 0.2158 |
| | | | RECM | 0.5061 | 0.5373 | 0.5587 | 0.6259 | 0.5547 | 0.5734 | 0.6010 | 0.6736 | 0.4807 | 0.5058 | 0.5211 | 0.5569 |
| | | | Sesgo relat. % | -0.6 | -1.8 | -2.9 | -5.8 | -1.5 | -2.5 | -3.1 | -5.3 | -0.2 | -2.8 | -5.9 | -10.8 |
| | 500 | $\alpha = 1.5$ | Sesgo | -0.0046 | 0.0107 | 0.0283 | 0.0741 | -0.0013 | 0.0175 | 0.0388 | 0.0936 | -0.0028 | 0.0174 | 0.0417 | 0.1013 |
| | | | RECM | 0.0584 | 0.0623 | 0.0703 | 0.1068 | 0.0668 | 0.0727 | 0.0835 | 0.1274 | 0.0529 | 0.0624 | 0.0784 | 0.1299 |
| | | | Sesgo relat. % | -0.3 | 0.7 | 1.9 | 4.9 | -0.1 | 1.2 | 2.6 | 6.2 | -0.2 | 1.2 | 2.8 | 6.8 |
| | | $\beta_0 = 4$ | Sesgo | -0.0068 | 0.0102 | 0.0312 | 0.0769 | 0.0028 | 0.0210 | 0.0456 | 0.1049 | -0.0014 | 0.0246 | 0.0533 | 0.1228 |
| | | | RECM | 0.1284 | 0.1319 | 0.1408 | 0.1670 | 0.1426 | 0.1466 | 0.1563 | 0.1926 | 0.1222 | 0.1305 | 0.1389 | 0.1821 |
| | | | Sesgo relat. % | -0.2 | 0.3 | 0.8 | 1.9 | 0.1 | 0.5 | 1.1 | 2.6 | 0.0 | 0.6 | 1.3 | 3.1 |
| | | $\beta_1 = -2$ | Sesgo | 0.0125 | 0.0311 | 0.0509 | 0.1164 | -0.0092 | 0.0121 | 0.0317 | 0.0805 | 0.0016 | 0.0580 | 0.1181 | 0.2322 |
| | | | RECM | 0.2169 | 0.2253 | 0.2374 | 0.2797 | 0.2431 | 0.2502 | 0.2586 | 0.2869 | 0.2137 | 0.2372 | 0.2600 | 0.3391 |
| | | | Sesgo relat. % | -0.6 | -1.6 | -2.5 | -5.8 | 0.5 | -0.6 | -1.6 | -4.0 | -0.1 | -2.9 | -5.9 | -11.6 |
| | 1000 | $\alpha = 1.5$ | Sesgo | -0.0023 | 0.0131 | 0.0315 | 0.0773 | -0.0012 | 0.0178 | 0.0382 | 0.0929 | -0.0013 | 0.0194 | 0.0454 | 0.1065 |
| | | | RECM | 0.0412 | 0.0458 | 0.0567 | 0.0957 | 0.0473 | 0.0522 | 0.0645 | 0.1113 | 0.0369 | 0.0472 | 0.0666 | 0.1224 |
| | | | Sesgo relat. % | -0.2 | 0.9 | 2.1 | 5.2 | -0.1 | 1.2 | 2.5 | 6.2 | -0.1 | 1.3 | 3.0 | 7.1 |
| | | $\beta_0 = 4$ | Sesgo | -0.0016 | 0.0147 | 0.0350 | 0.0824 | 0.0105 | 0.0291 | 0.0539 | 0.1120 | 0.0041 | 0.0289 | 0.0534 | 0.1293 |
| | | | RECM | 0.0868 | 0.0896 | 0.0962 | 0.1280 | 0.1013 | 0.1073 | 0.1212 | 0.1616 | 0.0845 | 0.0953 | 0.1060 | 0.1589 |
| | | | Sesgo relat. % | 0.0 | 0.4 | 0.9 | 2.1 | 0.3 | 0.7 | 1.3 | 2.8 | 0.1 | 0.7 | 1.3 | 3.2 |
| | | $\beta_1 = -2$ | Sesgo | 0.0039 | 0.0241 | 0.0449 | 0.1040 | -0.0184 | 0.0000 | 0.0181 | 0.0687 | -0.0060 | 0.0518 | 0.1231 | 0.2259 |
| | | | RECM | 0.1526 | 0.1588 | 0.1678 | 0.2052 | 0.1760 | 0.1780 | 0.1871 | 0.2098 | 0.1501 | 0.1747 | 0.2104 | 0.2863 |
| | | | Sesgo relat. % | -0.2 | -1.2 | -2.2 | -5.2 | 0.9 | 0.0 | -0.9 | -3.4 | 0.3 | -2.6 | -6.2 | -11.3 |

Cuadro 4.6: Sesgo, Sesgo relativo y RECM para diferentes niveles de censura

| | | | | Senh-NC | | | | Senh-St | | | | Senh-SL | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\alpha$ | n | Parámetro | Valores | 0 % | 10 % | 20 % | 40 % | 0 % | 10 % | 20 % | 40 % | 0 % | 10 % | 20 % | 40 % |
| 0.50 | 100 | $\alpha = 0.5$ | Sesgo | -0.0026 | 0.0056 | 0.0133 | 0.0391 | -0.0069 | -0.0001 | 0.0067 | 0.0254 | -0.0059 | -0.0045 | -0.0012 | 0.0066 |
| | | | RECM | 0.0398 | 0.0449 | 0.0522 | 0.0764 | 0.0504 | 0.0520 | 0.0558 | 0.0699 | 0.0380 | 0.0407 | 0.0437 | 0.0540 |
| | | | Sesgo relat. % | -0.5 | 1.1 | 2.7 | 7.8 | -1.4 | 0.0 | 1.3 | 5.1 | -1.2 | -0.9 | -0.2 | 1.3 |
| | | $\beta_0 = 4$ | Sesgo | -0.0058 | -0.0052 | 0.0010 | 0.0083 | 0.0013 | 0.0026 | 0.0047 | 0.0109 | 0.0011 | -0.0002 | -0.0031 | -0.0021 |
| | | | RECM | 0.1118 | 0.1172 | 0.1258 | 0.1489 | 0.1206 | 0.1251 | 0.1336 | 0.1573 | 0.1159 | 0.1228 | 0.1302 | 0.1482 |
| | | | Sesgo relat. % | -0.1 | -0.1 | 0.0 | 0.2 | 0.0 | 0.1 | 0.1 | 0.3 | 0.0 | 0.0 | -0.1 | -0.1 |
| | | $\beta_1 =$ -2 | Sesgo | 0.0112 | 0.0250 | 0.0297 | 0.0522 | 0.0038 | 0.0158 | 0.0266 | 0.0573 | 0.0007 | 0.0182 | 0.0422 | 0.0872 |
| | | | RECM | 0.1991 | 0.2124 | 0.2261 | 0.2649 | 0.2060 | 0.2148 | 0.2292 | 0.2709 | 0.2075 | 0.2203 | 0.2350 | 0.2708 |
| | | | Sesgo relat. % | -0.6 | -1.3 | -1.5 | -2.6 | -0.2 | -0.8 | -1.3 | -2.9 | 0.0 | -0.9 | -2.1 | -4.4 |
| | 500 | $\alpha = 0.5$ | Sesgo | 0.0030 | 0.0098 | 0.0188 | 0.0444 | -0.0005 | 0.0065 | 0.0145 | 0.0356 | -0.0009 | 0.0021 | 0.0042 | 0.0120 |
| | | | RECM | 0.0190 | 0.0236 | 0.0299 | 0.0536 | 0.0222 | 0.0246 | 0.0289 | 0.0462 | 0.0176 | 0.0195 | 0.0220 | 0.0294 |
| | | | Sesgo relat. % | 0.6 | 2.0 | 3.8 | 8.9 | -0.1 | 1.3 | 2.9 | 7.1 | -0.2 | 0.4 | 0.8 | 2.4 |
| | | $\beta_0 = 4$ | Sesgo | -0.0060 | -0.0011 | 0.0043 | 0.0134 | 0.0007 | 0.0017 | 0.0044 | 0.0093 | -0.0006 | -0.0037 | -0.0033 | -0.0087 |
| | | | RECM | 0.0503 | 0.0534 | 0.0580 | 0.0681 | 0.0537 | 0.0554 | 0.0586 | 0.0675 | 0.0531 | 0.0576 | 0.0617 | 0.0682 |
| | | | Sesgo relat. % | -0.2 | 0.0 | 0.1 | 0.3 | 0.0 | 0.0 | 0.1 | 0.2 | 0.0 | -0.1 | -0.1 | -0.2 |
| | | $\beta_1 =$ -2 | Sesgo | 0.0081 | 0.0148 | 0.0227 | 0.0470 | -0.0029 | 0.0094 | 0.0217 | 0.0542 | 0.0003 | 0.0251 | 0.0414 | 0.0961 |
| | | | RECM | 0.0864 | 0.0910 | 0.0985 | 0.1198 | 0.0918 | 0.0956 | 0.1014 | 0.1233 | 0.0931 | 0.1049 | 0.1182 | 0.1552 |
| | | | Sesgo relat. % | -0.4 | -0.7 | -1.1 | -2.4 | 0.1 | -0.5 | -1.1 | -2.7 | 0.0 | -1.3 | -2.1 | -4.8 |
| | 1000 | $\alpha = 0.5$ | Sesgo | 0.0028 | 0.0103 | 0.0198 | 0.0461 | -0.0004 | 0.0067 | 0.0145 | 0.0358 | -0.0004 | 0.0022 | 0.0048 | 0.0120 |
| | | | RECM | 0.0135 | 0.0181 | 0.0262 | 0.0509 | 0.0158 | 0.0178 | 0.0228 | 0.0414 | 0.0123 | 0.0144 | 0.0167 | 0.0235 |
| | | | Sesgo relat. % | 0.6 | 2.1 | 4.0 | 9.2 | -0.1 | 1.3 | 2.9 | 7.2 | -0.1 | 0.4 | 1.0 | 2.4 |
| | | $\beta_0 = 4$ | Sesgo | -0.0016 | 0.0018 | 0.0063 | 0.0152 | 0.0036 | 0.0046 | 0.0071 | 0.0125 | 0.0020 | -0.0004 | -0.0013 | -0.0039 |
| | | | RECM | 0.0348 | 0.0359 | 0.0378 | 0.0474 | 0.0383 | 0.0404 | 0.0434 | 0.0502 | 0.0370 | 0.0396 | 0.0432 | 0.0489 |
| | | | Sesgo relat. % | 0.0 | 0.0 | 0.2 | 0.4 | 0.1 | 0.1 | 0.2 | 0.3 | 0.1 | 0.0 | 0.0 | -0.1 |
| | | $\beta_1 =$ -2 | Sesgo | 0.0018 | 0.0120 | 0.0199 | 0.0436 | -0.0063 | 0.0049 | 0.0167 | 0.0483 | -0.0033 | 0.0185 | 0.0381 | 0.0870 |
| | | | RECM | 0.0605 | 0.0632 | 0.0691 | 0.0879 | 0.0665 | 0.0685 | 0.0741 | 0.0933 | 0.0651 | 0.0738 | 0.0889 | 0.1259 |
| | | | Sesgo relat. % | -0.1 | -0.6 | -1.0 | -2.2 | 0.3 | -0.2 | -0.8 | -2.4 | 0.2 | -0.9 | -1.9 | -4.4 |

# Capítulo 5

# Aplicación

## 5.1. Exploración de los datos

El conjunto de datos analizado está relacionado con el área de salud, datos que fueron estudiados en Leiva et al. (2007) bajo el modelo SN. Considera un total de 65 pacientes diagnosticados con mieloma múltiple (un tipo de cáncer de la sangre), de los cuales 48 murieron durante el estudio y mientras que 17 sobrevivieron. La variable status indica si el paciente sobrevive o muere. Además cinco covariables (de x1 a x5) explican el comportamiento de la variable tiempo de sobrevivencia del paciente (en meses), todos medidos al momento del diagnóstico, los cuales se detallan a continuación:

- x1: Logaritmo de la medición de nitrógeno ureíco en sangre
- x2: Nivel de hemoglobina
- x3: Edad del paciente
- x4: Sexo del paciente (0:varón; 1:mujer)
- x5: Nivel de calcio en la sangre
- status: Considera 0 paciente sobrevive; 1 muere paciente

## 5.2. Estimación de parámetros del modelo

Se considera el modelo de regresión:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \varepsilon_i, \quad i = 1, .., 65, \tag{5.1}$$

donde $y_i$ es el logaritmo del tiempo de vida observado o censurado para el $i$-ésimo individuo, y donde $\varepsilon_i \sim \mathrm{Senh} - \mathrm{NI}(\alpha, 0, \sigma = 2; H)$.

## 5.3. Comparación de modelos

Para los datos analizados, mieloma múltiple, se compararon todos los casos especiales del modelo de regresión basado en la distribución Senh-NI y el modelo basado en la distribución

SN así como el caso Normal (denotado por N). El Cuadro 5.1 muestra los valores de AIC, BIC y log-verosimilitud para cada modelo y donde se aprecia que todos los casos especiales del modelo Senh-NI (para diferentes valores de $\boldsymbol{\nu}$) presentan menores AIC y BIC que las del modelo SN y el caso Normal, indicando que los casos especiales de la Senh-NI se ajustan mejor los datos comparado con la SN y el caso Normal. Específicamente, el modelo que mejor ajusta los datos es la Senh-NC dado que tiene los menores AIC y BIC de entre todos los modelos.

Cuadro 5.1: AIC, BIC y log-verosimilitud para Senh-NI y SN

| | **AIC** | Ranking | **BIC** | Ranking | **log-verosimilitud** |
|---|---|---|---|---|---|
| Senh-st $\nu = 6$ | 147.75 | 12 | 165.14 | 12 | -65.87 |
| Senh-st $\nu = 5$ | 146.01 | 10 | 163.40 | 10 | -65.00 |
| Senh-st $\nu = 4$ | 145.37 | 9 | 162.77 | 9 | -64.68 |
| Senh-st $\nu = 3$ | 145.16 | 8 | 162.55 | 8 | -64.58 |
| Senh-SL $\nu = 6$ | 144.80 | 7 | 162.20 | 7 | -64.40 |
| Senh-SL $\nu = 5$ | 144.41 | 5 | 161.81 | 5 | -64.21 |
| Senh-SL $\nu = 4$ | 144.24 | 4 | 161.64 | 4 | -64.12 |
| Senh-SL $\nu = 3$ | 144.12 | 2 | 161.51 | 2 | -64.06 |
| Senh-NC $\nu = (0{,}1, 0{,}1)$ | 146.39 | 11 | 163.78 | 11 | -65.19 |
| Senh-NC $\nu = (0{,}4, 0{,}4)$ | 144.58 | 6 | 161.97 | 6 | -64.29 |
| Senh-NC $\nu = (0{,}6, 0{,}6)$ | 144.15 | 3 | 161.55 | 3 | -64.08 |
| Senh-NC $\nu = (0{,}9, 0{,}9)$ | 143.66 | 1 | 161.05 | 1 | -63.83 |
| SN | 161.04 | 13 | 165.39 | 13 | -78.52 |
| N | 210.11 | 14 | 223.16 | 14 | -99.05 |

La estimación de los parámetros se presenta en el Cuadro 5.2 con los valores estimados de $\boldsymbol{\beta}$ y $\alpha$ para cada caso especial de la familia Senh-NI así como para el modelo SN, cabe mencionar que los errores estándar para la familia Senh-NI presentan en todos los casos un error estándar ligeramente mayor a los obtenido con el modelo SN. El Cuadro 5.2 muestra el límite superior (LS) y el límite inferior (LI) de un intervalo de confianza donde LS $= \widehat{\theta} + 1{,}96\mathrm{EE}_{\widehat{\theta}}$ y LS $= \widehat{\theta} - 1{,}96\mathrm{EE}_{\widehat{\theta}}$ con $\mathrm{EE}_{\widehat{\theta}}$ = error estándar de $\widehat{\theta}$, haremos uso de los intervalos de confianza para determinar la significancia de los parámetros estimados. Del Cuadro 5.2 notamos que los parámetros $\beta_0$, $\beta_1$ y $\alpha$ no son significativos en todos los casos mientras que $\beta_2$ es significativo para todos los casos excepto el modelo SN, $\beta_3$ y $\beta_4$ es significativo en todos los casos, mientras que $\beta_5$ es significativo para los modelos Senh-NC para $\boldsymbol{\nu} = (0{,}9, 0{,}9)$ y Senh-SL para $\nu = 3$.

## 5.4. Selección de variables

En el Cuadro 5.3 se muestra una relación de modelos en los que se ha seleccionado diferentes variables los cuales serán comparados por medio de los indicadores AIC y BIC. Podemos observar que el mejor modelo según el AIC y BIC es el Senh-NC que tiene por variables $x_1$ (Logaritmo de la medición de nitrógeno ureíco en sangre) y $x_2$ (Nivel de hemoglobina), es decir el modelo con las variables seleccionadas sería: $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$.

Cuadro 5.2: Estimación de parámetros y error estándar para casos especiales de la familia Senh-NI

| | **Parámetro** | $\beta_0$ | $\beta_1$ (Nitrógeno) | $\beta_2$ (Hemoglobina) | $\beta_3$ (Edad) | $\beta_4$ (Sexo) | $\beta_5$ (Calcio) | $\alpha$ |
|---|---|---|---|---|---|---|---|---|
| Senh-St $\nu = 3$ | Estimación | 6.6555 | -1.5541 | 0.0955 | -0.0111 | -0.0312 | -0.2058 | 0.8024 |
| | Error estándar | 1.4476 | 0.5546 | 0.0522 | 0.0146 | 0.3110 | 0.0896 | 0.1398 |
| | LI | 3.8183 | -2.6411 | -0.0068 | -0.0397 | -0.6407 | -0.3814 | 0.5284 |
| | LS | 9.4927 | -0.4671 | 0.1978 | 0.0175 | 0.5783 | -0.0302 | 1.0764 |
| Senh-NC $\nu = (0{,}9, 0{,}9)$ | Estimación | 5.9185 | -1.5979 | 0.1159 | -0.0097 | 0.0379 | -0.1585 | 0.8568 |
| | Error estándar | 1.2963 | 0.6787 | 0.0549 | 0.0160 | 0.3333 | 0.0827 | 0.1269 |
| | LI | 3.3778 | -2.9281 | 0.0083 | -0.0411 | -0.6154 | -0.3206 | 0.6081 |
| | LS | 8.4592 | -0.2677 | 0.2235 | 0.0217 | 0.6912 | 0.0036 | 1.1055 |
| Senh-SL $\nu = 3$ | Estimación | 5.2498 | -1.4465 | 0.1040 | -0.0027 | 0.0976 | -0.1378 | 0.8452 |
| | Error estándar | 1.2739 | 0.6258 | 0.0531 | 0.0154 | 0.3201 | 0.0821 | 0.1134 |
| | LI | 2.7530 | -2.6730 | -0.0001 | -0.0329 | -0.5298 | -0.2987 | 0.6229 |
| | LS | 7.7466 | -0.2200 | 0.2081 | 0.0275 | 0.7250 | 0.0231 | 1.0675 |
| SN | Estimación | 4.6744 | -1.5117 | 0.1408 | 0.0077 | -0.1964 | -0.1365 | 1.0695 |
| | Error estándar | 1.1569 | 0.4366 | 0.0491 | 0.0125 | 0.2812 | 0.0691 | 0.1108 |
| | LI | 2.4069 | -2.3674 | 0.0446 | -0.0168 | -0.7475 | -0.2719 | 0.8523 |
| | LS | 6.9419 | -0.6560 | 0.2370 | 0.0322 | 0.3547 | -0.0011 | 1.2867 |

Cuadro 5.3: AIC, BIC y log-verosimilitud para Senh-NI y SN

| Modelo | Distribución | **AIC** | Ranking | **BIC** | Ranking |
|---|---|---|---|---|---|
| $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \varepsilon_i$ | Senh-St | 147.75 | 25 | 165.14 | 34 |
| | Senh-Sl | 147.37 | 24 | 162.59 | 30 |
| | Senh-NC | 145.39 | 13 | 160.61 | 22 |
| $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_5 x_{i5} + \varepsilon_i$ | Senh-St | 146.42 | 19 | 161.64 | 24 |
| | Senh-Sl | 142.84 | 5 | 158.06 | 14 |
| | Senh-NC | 141.74 | 4 | 156.96 | 11 |
| $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_4 x_{i4} + \beta_5 x_{i5} + \varepsilon_i$ | Senh-St | 146.74 | 20 | 161.96 | 26 |
| | Senh-Sl | 142.84 | 6 | 158.06 | 15 |
| | Senh-NC | 141.69 | 2 | 156.91 | 10 |
| $y_i = \beta_0 + \beta_1 x_{i1} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \varepsilon_i$ | Senh-St | 148.25 | 27 | 163.47 | 31 |
| | Senh-Sl | 147.30 | 23 | 162.52 | 27 |
| | Senh-NC | 145.49 | 14 | 160.71 | 23 |
| $y_i = \beta_0 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \varepsilon_i$ | Senh-St | 158.07 | 43 | 173.29 | 45 |
| | Senh-Sl | 154.62 | 38 | 169.84 | 40 |
| | Senh-NC | 152.77 | 35 | 167.99 | 37 |
| $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i$ | Senh-St | 149.48 | 31 | 162.53 | 28 |
| | Senh-Sl | 145.79 | 17 | 158.84 | 18 |
| | Senh-NC | 143.66 | 9 | 156.71 | 9 |
| $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_4 x_{i4} + \varepsilon_i$ | Senh-St | 149.49 | 32 | 162.55 | 29 |
| | Senh-Sl | 145.70 | 16 | 158.74 | 17 |
| | Senh-NC | 143.30 | 7 | 156.34 | 7 |
| $y_i = \beta_0 + \beta_1 x_{i1} + \beta_3 x_{i3} + \beta_4 x_{i4} + \varepsilon_i$ | Senh-St | 150.91 | 33 | 163.96 | 32 |
| | Senh-Sl | 148.90 | 28 | 161.95 | 25 |
| | Senh-NC | 145.94 | 18 | 158.99 | 19 |
| $y_i = \beta_0 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_5 x_{i5} + \varepsilon_i$ | Senh-St | 156.15 | 39 | 169.20 | 38 |
| | Senh-Sl | 152.81 | 36 | 165.85 | 35 |
| | Senh-NC | 151.08 | 34 | 164.12 | 33 |
| $y_i = \beta_0 + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \varepsilon_i$ | Senh-St | 158.86 | 44 | 171.90 | 43 |
| | Senh-Sl | 157.90 | 42 | 170.94 | 42 |
| | Senh-NC | 154.38 | 37 | 167.42 | 36 |
| $y_i = \beta_0 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \varepsilon_i$ | Senh-St | 159.57 | 45 | 172.62 | 44 |
| | Senh-Sl | 157.73 | 41 | 170.77 | 41 |
| | Senh-NC | 156.25 | 40 | 169.30 | 39 |
| $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$ | Senh-St | 147.75 | 26 | 158.62 | 16 |
| | Senh-Sl | 143.81 | 10 | 154.68 | 4 |
| | Senh-NC | 141.69 | 1 | 152.56 | 1 |
| $y_i = \beta_0 + \beta_1 x_{i1} + \beta_3 x_{i3} + \varepsilon_i$ | Senh-St | 148.92 | 29 | 159.79 | 20 |
| | Senh-Sl | 146.90 | 21 | 157.77 | 12 |
| | Senh-NC | 144.01 | 11 | 154.88 | 5 |
| $y_i = \beta_0 + \beta_1 x_{i1} + \beta_4 x_{i4} + \varepsilon_i$ | Senh-St | 149.26 | 30 | 160.13 | 21 |
| | Senh-Sl | 146.91 | 22 | 157.78 | 13 |
| | Senh-NC | 144.02 | 12 | 154.89 | 6 |
| $y_i = \beta_0 + \beta_1 x_{i1} + \beta_5 x_{i5} + \varepsilon_i$ | Senh-St | 145.65 | 15 | 156.52 | 8 |
| | Senh-Sl | 143.54 | 8 | 154.41 | 3 |
| | Senh-NC | 141.70 | 3 | 152.57 | 2 |

# Capítulo 6

# Conclusiones

## 6.1. Conclusiones

En la literatura los modelos de regresión log-BS no consideran datos con observaciones censuradas. El modelo de regresión lineal, motivo del presente estudio, asume una distribución Senh-NI, considerando censura no informativa, y posee colas más pesadas por lo que permite modelar mejor datos con observaciones atípicas.

Para el proceso de estimación de parámetros se utilizaron métodos numéricos y un método de aproximación para obtener la matriz de información observada.

Los estudios de simulación permitieron verificar que las metodologías utilizadas recuperan los parámetros del modelo, con sesgo y RMSE aceptables, además se notó que el modelo Senh-SL es el más costoso desde el punto de vista computacional. También, la simulación demostró que los métodos utilizados para la estimación de los parámetros y para aproximar la matriz de información observada, tuvieron un buen desempeño para los diferentes niveles de censura y los distintos tamaños de muestra.

Un conjunto de datos del área médica fue analizado bajo este modelo y se comparó con el modelo de regresión SN con censura no informativa, así se evidenció que todos los casos especiales de la familia Senh-NI generan AIC y BIC menores que los obtenidos por el modelo SN. También, con el modelo propuesto se obtuvo menores errores estándar para la mayoría de los parámetros.

## 6.2. Sugerencias para futuras investigaciones

Se proponen las siguientes investigaciones futuras:

- Realizar una extensión utilizando una distribución Senh-Skew Normal/Independiente con censura, con lo cual se obtendría un modelo más robusto y así extender el trabajo de Maehara et al. (2021)
- Tomando como referencia el artículo Leiva et al. (2007) se puede realizar un análisis de diagnóstico basado en el método de influencia local y también un análisis de residuales.

- Considerar el caso no lineal, también con censura para con ello generalizar diversos trabajos publicados en la literatura, como por ejemplo el de Lemonte y Cordeiro (2009).
- Considerar el parámetro $\sigma$ desconocido en un modelo de regresión con censura.

# Apéndice A

# Apéndice

## A.1. Derivadas de $\xi_{1i}$ e $\xi_{2i}$

$$
\begin{aligned}
\frac{\partial}{\partial\alpha}\xi_{2i}^2 &= \frac{\partial}{\partial\alpha}\left[\frac{4}{\alpha^2}\mathrm{senh}^2\left(\frac{\mathrm{Y_i}-\mathbf{x}_\mathrm{i}^\top\boldsymbol{\beta}}{\sigma}\right)\right] \\
&= 4\mathrm{senh}^2\left(\frac{\mathrm{Y_i}-\mathbf{x}_\mathrm{i}^\top\boldsymbol{\beta}}{\sigma}\right)\frac{\partial}{\partial\alpha}\left(\frac{1}{\alpha^2}\right) \\
&= 4\mathrm{senh}^2\left(\frac{\mathrm{Y_i}-\mathbf{x}_\mathrm{i}^\top\boldsymbol{\beta}}{\sigma}\right)\left(-2\alpha^{-3}\right)\frac{\alpha^2}{\alpha^2} \\
&= \xi_{2i}^2\left(\frac{-2}{\alpha}\right) \\
\frac{\partial}{\partial\alpha}\xi_{1i} &= \frac{\partial}{\partial\alpha}\left[\frac{2}{\alpha}\cosh\left(\frac{\mathrm{Y_i}-\mathbf{x}_\mathrm{i}^\top\boldsymbol{\beta}}{\sigma}\right)\right] \\
&= 2\cosh\left(\frac{\mathrm{Y_i}-\mathbf{x}_\mathrm{i}^\top\boldsymbol{\beta}}{\sigma}\right)\frac{\partial}{\partial\alpha}\frac{1}{\alpha} \\
&= 2\cosh\left(\frac{\mathrm{Y_i}-\mathbf{x}_\mathrm{i}^\top\boldsymbol{\beta}}{\sigma}\right)\left(\frac{-1}{\alpha^2}\right) \\
&= \xi_{1i}\left(\frac{-1}{\alpha}\right) \\
\frac{\partial}{\partial\alpha}\xi_{2i} &= \frac{\partial}{\partial\alpha}\left[\frac{2}{\alpha}\mathrm{senh}\left(\frac{\mathrm{Y_i}-\mathbf{x}_\mathrm{i}^\top\boldsymbol{\beta}}{\sigma}\right)\right] \\
&= \left(\frac{-1}{\alpha}\right)\frac{2}{\alpha}\mathrm{senh}\left(\frac{\mathrm{Y_i}-\mathbf{x}_\mathrm{i}^\top\boldsymbol{\beta}}{\sigma}\right) \\
&= \left(\frac{-1}{\alpha}\right)\xi_{2i}.
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial}{\partial\boldsymbol{\beta}}\xi_{2i}^2 &= \frac{\partial}{\partial\boldsymbol{\beta}}\left[\frac{4}{\alpha^2}\mathrm{senh}^2\left(\frac{\mathrm{Y_i}-\mathbf{x}_\mathrm{i}^\top\boldsymbol{\beta}}{\sigma}\right)\right]\\
&= -\left(\frac{2}{\sigma}\right)\frac{2}{\alpha}\mathrm{senh}\left(\frac{\mathrm{Y_i}-\mathbf{x}_\mathrm{i}^\top\boldsymbol{\beta}}{\sigma}\right)\frac{2}{\alpha}\cosh\left(\frac{\mathrm{Y_i}-\mathbf{x}_\mathrm{i}^\top\boldsymbol{\beta}}{\sigma}\right)\frac{\partial}{\partial\boldsymbol{\beta}}(\mu_\mathrm{i})\\
&= -\frac{2}{\sigma}\xi_{2i}\xi_{1i}\frac{\partial}{\partial\boldsymbol{\beta}}(\mu_i)\\
\frac{\partial}{\partial\boldsymbol{\beta}}\xi_{1i} &= \frac{\partial}{\partial\boldsymbol{\beta}}\left[\frac{2}{\alpha}\cosh\left(\frac{\mathrm{Y_i}-\mathbf{x}_\mathrm{i}^\top\boldsymbol{\beta}}{\sigma}\right)\right]\\
&= \frac{2}{\alpha}\mathrm{senh}\left(\frac{\mathrm{Y_i}-\mathbf{x}_\mathrm{i}^\top\boldsymbol{\beta}}{\sigma}\right)\frac{\partial}{\partial\boldsymbol{\beta}}\left(\frac{\mathrm{Y_i}-\mathbf{x}_\mathrm{i}^\top\boldsymbol{\beta}}{\sigma}\right)\\
&= \xi_{2i}\left[\frac{-1}{\sigma}\frac{\partial}{\partial\boldsymbol{\beta}}(\mu_i)\right]\\
\frac{\partial}{\partial\boldsymbol{\beta}}\xi_{2i} &= \frac{\partial}{\partial\boldsymbol{\beta}}\left[\frac{2}{\alpha}\mathrm{senh}\left(\frac{\mathrm{Y_i}-\mathbf{x}_\mathrm{i}^\top\boldsymbol{\beta}}{\sigma}\right)\right]\\
&= \frac{2}{\alpha}\cosh\left(\frac{\mathrm{Y_i}-\mathbf{x}_\mathrm{i}^\top\boldsymbol{\beta}}{\sigma}\right)\frac{\partial}{\partial\boldsymbol{\beta}}\left(\frac{\mathrm{Y_i}-\mathbf{x}_\mathrm{i}^\top\boldsymbol{\beta}}{\sigma}\right)\\
&= \xi_{1i}\left(\frac{-1}{\sigma}\right)\frac{\partial}{\partial\boldsymbol{\beta}}(\mu_i)
\end{aligned}
$$

## A.2. Algunas derivadas y expresiones importantes

$$
\begin{aligned}
\phi_{\mathrm{NI}}(\xi_{2i}) &= \int_0^\infty \phi\big(\xi_{2i};0,u^{-1}\big)\,dH(u)\\
&= \frac{1}{\sqrt{2\pi}}\int_0^\infty u^{1/2}\exp\left(\frac{-\mathrm{u}}{2}\xi_{\mathrm{2i}}^2\right)dH(u)\\
&= \frac{1}{\sqrt{2\pi}}A_i(1/2)\\
\frac{\partial}{\partial\alpha}\phi_{\mathrm{NI}}(\xi_{2i}) &= \frac{1}{\sqrt{2\pi}}\int_0^\infty u^{1/2}\exp\left(\frac{-\mathrm{u}}{2}\xi_{\mathrm{2i}}^2\right)\left(\frac{-u}{2}\frac{\partial}{\partial\alpha}\xi_{2i}^2\right)dH(u)\\
&= \frac{1}{\sqrt{2\pi}}\int_0^\infty u^{1/2}\exp\left(\frac{-\mathrm{u}}{2}\xi_{\mathrm{2i}}^2\right)\left(\frac{-u}{2}\right)\left(\frac{-2}{\alpha}\xi_{2i}^2\right)dH(u)\\
&= \frac{\xi_{2i}^2}{\alpha\sqrt{2\pi}}\int_0^\infty u^{3/2}\exp\left(\frac{-\mathrm{u}}{2}\xi_{\mathrm{2i}}^2\right)dH(u)\\
&= \frac{\xi_{2i}^2}{\alpha\sqrt{2\pi}}A_i(3/2)\\
\frac{\partial}{\partial\alpha}\left[1-\Phi_{\mathrm{NI}}(\xi_{2i})\right] &= -\phi_{\mathrm{NI}}(\xi_{2i})\frac{\partial}{\partial\alpha}(\xi_{2i})\\
&= -\phi_{\mathrm{NI}}(\xi_{2i})\left(\frac{-1}{\alpha}\right)\xi_{2i}
\end{aligned}
$$

$$
\begin{aligned}
\phi_{\mathrm{NI}}(\xi_{2i}) &= \int_0^\infty \phi\big(\xi_{2i}; 0, u^{-1}\big)\, dH(u) \\
&= \frac{1}{\sqrt{2\pi}} \int_0^\infty u^{1/2} \exp\left(\frac{-\mathrm{u}}{2}\xi_{2\mathrm{i}}^2\right) dH(u) \\
&= \frac{1}{\sqrt{2\pi}} A_i(1/2) \\
\frac{\partial}{\partial \boldsymbol{\beta}} \phi_{\mathrm{NI}}(\xi_{2i}) &= \frac{1}{\sqrt{2\pi}} \int_0^\infty u^{1/2} \exp\left(\frac{-\mathrm{u}}{2}\xi_{2\mathrm{i}}^2\right) \left(\frac{-u}{2}\frac{\partial}{\partial \boldsymbol{\beta}}\xi_{2i}^2\right) dH(u) \\
&= \frac{1}{\sqrt{2\pi}} \int_0^\infty u^{1/2} \exp\left(\frac{-\mathrm{u}}{2}\xi_{2\mathrm{i}}^2\right) \left(\frac{-u}{2}\right) \left(\frac{-2}{\sigma}\xi_{2i}\xi_{1i}\right) \frac{\partial}{\partial \boldsymbol{\beta}}(\mu_i)\, dH(u) \\
&= \frac{2}{\sigma}\xi_{2i}\xi_{1i} \frac{1}{2\sqrt{2\pi}} \frac{\partial}{\partial \boldsymbol{\beta}}(\mu_i) \int_0^\infty u^{3/2} \exp\left(\frac{-\mathrm{u}}{2}\xi_{2\mathrm{i}}^2\right) dH(u) \\
&= \frac{\xi_{2i}\xi_{1i}}{\sigma\sqrt{2\pi}} \frac{\partial}{\partial \boldsymbol{\beta}}(\mu_i) A_i(3/2) \\
\frac{\partial}{\partial \boldsymbol{\beta}} [1 - \Phi_{\mathrm{NI}}(\xi_{2i})] &= -\phi_{\mathrm{NI}}(\xi_{2i}) \frac{\partial}{\partial \boldsymbol{\beta}}(\xi_{2i}) \\
&= -\phi_{\mathrm{NI}}(\xi_{2i})\xi_{1i} \left(\frac{-1}{\sigma}\right) \frac{\partial}{\partial \boldsymbol{\beta}}(\mu_i)
\end{aligned}
$$

## A.3. Función Score

Derivando la función de log-verosimilitud se obtiene la función score

$$
\begin{aligned}
U_\alpha(\boldsymbol{\theta}) &= \frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha} \\
&= \sum_{i \in D} \frac{\partial}{\partial \alpha} \log \phi_{\mathrm{NI}}(\xi_{2i}) + \sum_{i \in D} \frac{\partial}{\partial \alpha} \log(\xi_{1i}) + \sum_{i \in C} \frac{\partial}{\partial \alpha} \log(1 - \Phi_{\mathrm{NI}}(\xi_{2i})) \\
&= \sum_{i \in D} \frac{1}{\phi_{\mathrm{NI}}(\xi_{2i})} \frac{\partial}{\partial \alpha} \phi_{\mathrm{NI}}(\xi_{2i}) + \sum_{i \in D} \frac{1}{\xi_{1i}} \frac{\partial}{\partial \alpha}(\xi_{1i}) + \sum_{i \in C} \frac{1}{1 - \Phi_{\mathrm{NI}}(\xi_{2i})} \frac{\partial}{\partial \alpha} [1 - \Phi_{\mathrm{NI}}(\xi_{2i})] \\
&= \sum_{i \in D} \frac{\sqrt{2\pi}}{A_i(1/2)} \frac{\xi_{2i}^2 A_i(3/2)}{\alpha\sqrt{2\pi}} + \sum_{i \in D} \frac{1}{\xi_{1i}} \xi_{1i} \left(\frac{-1}{\alpha}\right) + \sum_{i \in C} \frac{1}{1 - \Phi_{\mathrm{NI}}(\xi_{2i})} (-\phi_{\mathrm{NI}}(\xi_{2i})) \left(\frac{-1}{\alpha}\xi_{2i}\right) \\
&= \sum_{i \in D} \frac{\xi_{2i}^2}{\alpha} \frac{A_i(3/2)}{A_i(1/2)} + \sum_{i \in D} \frac{-1}{\alpha} + \sum_{i \in C} \frac{1}{\alpha} \xi_{2i} h(\xi_{2i}) \\
&= \frac{1}{\alpha} \left[ \sum_{i \in D} \left( \xi_{2i}^2 \frac{A_i(3/2)}{A_i(1/2)} - 1 \right) + \sum_{i \in C} \xi_{2i} h(\xi_{2i}) \right]
\end{aligned}
$$

$$
\begin{aligned}
U_{\boldsymbol{\beta}}(\boldsymbol{\theta}) &= \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} \\
&= \sum_{i\in D} \frac{\partial}{\partial \boldsymbol{\beta}} \log \phi_{\mathrm{NI}}(\xi_{2i}) + \sum_{i\in D} \frac{\partial}{\partial \boldsymbol{\beta}} \log(\xi_{1i}) + \sum_{i\in C} \frac{\partial}{\partial \boldsymbol{\beta}} \log\left(1-\Phi_{\mathrm{NI}}(\xi_{2i})\right) \\
&= \sum_{i\in D} \frac{1}{\phi_{\mathrm{NI}}(\xi_{2i})} \frac{\partial}{\partial \boldsymbol{\beta}} \phi_{\mathrm{NI}}(\xi_{2i}) + \sum_{i\in D} \frac{1}{\xi_{1i}} \frac{\partial}{\partial \boldsymbol{\beta}} (\xi_{1i}) + \sum_{i\in C} \frac{1}{1-\Phi_{\mathrm{NI}}(\xi_{2i})} \frac{\partial}{\partial \boldsymbol{\beta}} \left[1-\Phi_{\mathrm{NI}}(\xi_{2i})\right] \\
&= \sum_{i\in D} \frac{\xi_{2i}\xi_{1i}A_i(3/2)}{\sigma A_i(1/2)} \frac{\partial}{\partial \boldsymbol{\beta}}(\mu_i) + \sum_{i\in D} \frac{\xi_{2i}}{\xi_{1i}} \left[\frac{-1}{\sigma} \frac{\partial}{\partial \boldsymbol{\beta}}(\mu_i)\right] + \sum_{i\in C} \frac{-\phi_{\mathrm{NI}}(\xi_{2i})}{1-\Phi_{\mathrm{NI}}(\xi_{2i})} \xi_{1i} \left(\frac{-1}{\sigma}\right) \frac{\partial}{\partial \boldsymbol{\beta}}(\mu_i) \\
&= \frac{1}{\sigma} \left[\sum_{i\in D} \xi_{2i} \frac{\partial}{\partial \boldsymbol{\beta}}(\mu_i) \left(\frac{\xi_{1i}A_i(3/2)}{A_i(1/2)} - \frac{1}{\xi_{1i}}\right) + \sum_{i\in C} \frac{\phi_{\mathrm{NI}}(\xi_{2i})}{1-\Phi_{\mathrm{NI}}(\xi_{2i})} \xi_{1i} \frac{\partial}{\partial \boldsymbol{\beta}}(\mu_i)\right] \\
&= \frac{1}{\sigma} \left[\sum_{i\in D} \xi_{2i} \left(\frac{\xi_{1i}A_i(3/2)}{A_i(1/2)} - \frac{1}{\xi_{1i}}\right) \frac{\partial}{\partial \boldsymbol{\beta}}(\mu_i) + \sum_{i\in C} \xi_{1i} h(\xi_{2i}) \frac{\partial}{\partial \boldsymbol{\beta}}(\mu_i)\right]
\end{aligned}
$$

donde $h(\xi_{2i})$ y $A_i(w)$ se encuentran definidos en **??** y 3.4 respectivamente.

## A.4. Matriz de información observada

Derivadas y expresiones requeridas para obtener $\widehat{s}_{i,\alpha}$

$$
\begin{aligned}
\frac{\partial}{\partial \alpha} f_Y(y_i;\boldsymbol{\theta}) &= \frac{\partial}{\partial \alpha} \left[\frac{1}{\sigma} \phi_{\mathrm{NI}}(\xi_{2i}) \xi_{1i}\right] \\
&= \frac{1}{\sigma} \left[\xi_{1i} \frac{\partial}{\partial \alpha} \phi_{\mathrm{NI}}(\xi_{2i}) + \phi_{\mathrm{NI}}(\xi_{2i}) \frac{\partial}{\partial \alpha}(\xi_{1i})\right] \\
&= \frac{1}{\sigma} \left[\xi_{1i} \frac{\xi_{2i}^2}{\alpha\sqrt{2\pi}} A_i(3/2) + \frac{1}{\sqrt{2\pi}} A_i(1/2) \xi_{1i} (\frac{-1}{\alpha})\right] \\
&= \frac{\xi_{1i}}{\sigma\sqrt{2\pi}\alpha} \left[\xi_{2i}^2 A_i(3/2) - A_i(1/2)\right] \\
f_Y(y_i;\boldsymbol{\theta}) &= \frac{1}{\sigma} \phi_{\mathrm{NI}}(\xi_{2i}) \xi_{1i} \\
&= \frac{1}{\sigma} \frac{1}{\sqrt{2\pi}} A_i(1/2) \xi_{1i} \\
&= \frac{A_i(1/2)\xi_{1i}}{\sigma\sqrt{2\pi}} \\
\frac{\frac{\partial}{\partial \alpha} f_Y(y_i;\boldsymbol{\theta})}{f_Y(y_i;\boldsymbol{\theta})} &= \frac{\xi_{1i}}{\sigma\sqrt{2\pi}\alpha} \left[\xi_{2i}^2 A_i(3/2) - A_i(1/2)\right] \frac{\sigma\sqrt{2\pi}}{A_i(1/2)\xi_{1i}} \\
&= \frac{1}{\alpha} \left[\frac{\xi_{2i}^2 A_i(3/2)}{A_i(1/2)} - 1\right].
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial}{\partial\alpha}S_Y(y_i;\boldsymbol{\theta}) &= \frac{\partial}{\partial\alpha}\left[1-\Phi_{\mathrm{NI}}(\xi_{2i})\right]\\
&= -\phi_{\mathrm{NI}}(\xi_{2i})\frac{\partial}{\partial\alpha}(\xi_{2i})\\
&= -\phi_{\mathrm{NI}}(\xi_{2i})\left(\frac{-1}{\alpha}\xi_{2i}\right)\\
&= \frac{1}{\alpha}\phi_{\mathrm{NI}}(\xi_{2i})\xi_{2i}\\
&= \left(\frac{1}{\alpha\sqrt{2\pi}}\right)A_i(1/2)\xi_{2i}
\end{aligned}
$$

Sea la función de sobrevivencia $S_Y(y_i;\boldsymbol{\theta}) = 1-\Phi_{\mathrm{NI}}(\xi_{2i})$

$$
\begin{aligned}
\frac{\frac{\partial}{\partial\alpha}S_Y(y_i;\boldsymbol{\theta})}{S_Y(y_i;\boldsymbol{\theta})} &= \left(\frac{1}{\alpha\sqrt{2\pi}}\right)A_i(1/2)\xi_{2i}\frac{1}{1-\Phi_{\mathrm{NI}}(\xi_{2i})}\\
&= \left(\frac{1}{\alpha\sqrt{2\pi}}\right)\frac{A_i(1/2)\xi_{2i}}{1-\Phi_{\mathrm{NI}}(\xi_{2i})}
\end{aligned}
$$

Resultados para $\widehat{s}_{i,\alpha}$

$$
\begin{aligned}
\widehat{s}_{i,\alpha} &= \begin{cases}\dfrac{\frac{\partial}{\partial\alpha}f_Y(y_i;\boldsymbol{\theta})}{f_Y(y_i;\boldsymbol{\theta})} & \text{si} \quad i \in D\\[2ex] \dfrac{\frac{\partial}{\partial\alpha}S_Y(y_i;\boldsymbol{\theta})}{S_Y(y_i;\boldsymbol{\theta})} & \text{si} \quad i \in C\end{cases}\\
&= \begin{cases}\dfrac{1}{\alpha}\left[\dfrac{\xi_{2i}^2A_i(3/2)}{A_i(1/2)}-1\right] & \text{si} \quad i \in D\\[2ex] \left(\dfrac{1}{\alpha\sqrt{2\pi}}\right)\dfrac{A_i(1/2)\xi_{2i}}{1-\Phi_{\mathrm{NI}}(\xi_{2i})} & \text{si} \quad i \in C\end{cases}
\end{aligned}
$$

donde $A_i(w)$ se encuentra definido en 3.4.

Derivadas y expresiones requeridas para obtener $\widehat{s}_{i,\boldsymbol{\beta}}$

$$
\begin{aligned}
\frac{\partial}{\partial\boldsymbol{\beta}}f_Y(y_i;\boldsymbol{\theta}) &= \frac{\partial}{\partial\boldsymbol{\beta}}\left[\frac{1}{\sigma}\phi_{\mathrm{NI}}(\xi_{2i})\xi_{1i}\right]\\
&= \frac{1}{\sigma}\left[\xi_{1i}\frac{\partial}{\partial\boldsymbol{\beta}}\phi_{\mathrm{NI}}(\xi_{2i})+\phi_{\mathrm{NI}}(\xi_{2i})\frac{\partial}{\partial\boldsymbol{\beta}}(\xi_{1i})\right]\\
&= \frac{1}{\sigma}\left[\frac{\xi_{1i}\xi_{2i}\xi_{1i}}{\alpha\sqrt{2\pi}}A_i(3/2)\frac{\partial}{\partial\boldsymbol{\beta}}(\mu_i)+\frac{A_i(1/2)}{\sqrt{2\pi}}\xi_{2i}(\frac{-1}{\alpha})\frac{\partial}{\partial\boldsymbol{\beta}}(\mu_i)\right]\\
&= \frac{\xi_{2i}}{\sigma^2\sqrt{2\pi}}\frac{\partial}{\partial\boldsymbol{\beta}}(\mu_i)\left[\xi_{1i}^2A_i(3/2)-A_i(1/2)\right]
\end{aligned}
$$

$$
\begin{aligned}
f_Y(y_i;\boldsymbol{\theta}) &= \frac{1}{\sigma}\phi_{\mathrm{NI}}(\xi_{2i})\xi_{1i}\\
&= \frac{1}{\sigma}\frac{1}{\sqrt{2\pi}}A_i(1/2)\xi_{1i}\\
&= \frac{A_i(1/2)\xi_{1i}}{\sigma\sqrt{2\pi}}\\
\frac{\frac{\partial}{\partial\boldsymbol{\beta}}f_Y(y_i;\boldsymbol{\theta})}{f_Y(y_i;\boldsymbol{\theta})} &= \frac{\xi_{2i}}{\sigma^2\sqrt{2\pi}}\frac{\partial}{\partial\boldsymbol{\beta}}(\mu_i)\left[\xi_{1i}^2A_i(3/2)-A_i(1/2)\right]\frac{\sigma\sqrt{2\pi}}{A_i(1/2)\xi_{1i}}\\
&= \frac{\xi_{2i}}{\sigma A_i(1/2)\xi_{1i}}\frac{\partial}{\partial\boldsymbol{\beta}}(\mu_i)\left[\xi_{1i}^2A_i(3/2)-A_i(1/2)\right]\\
\frac{\partial}{\partial\boldsymbol{\beta}}S_Y(y_i;\boldsymbol{\theta}) &= \frac{\partial}{\partial\boldsymbol{\beta}}\left[1-\Phi_{\mathrm{NI}}(\xi_{2i})\right]\\
&= -\phi_{\mathrm{NI}}(\xi_{2i})\frac{\partial}{\partial\boldsymbol{\beta}}(\xi_{2i})\\
&= -\phi_{\mathrm{NI}}(\xi_{2i})\left(\frac{-1}{\alpha}\xi_{1i}\right)\frac{\partial}{\partial\boldsymbol{\beta}}(\mu_i)\\
&= \frac{1}{\alpha}\frac{\partial}{\partial\boldsymbol{\beta}}(\mu_i)\xi_{1i}\phi_{\mathrm{NI}}(\xi_{2i})
\end{aligned}
$$

Ahora vamos a derivar la función de sobrevivencia respecto a $\boldsymbol{\beta}$

$$
\begin{aligned}
\frac{\frac{\partial}{\partial\boldsymbol{\beta}}S_Y(y_i;\boldsymbol{\theta})}{S_Y(y_i;\boldsymbol{\theta})} &= \frac{1}{\alpha}\frac{\partial}{\partial\boldsymbol{\beta}}(\mu_i)\xi_{1i}\phi_{\mathrm{NI}}(\xi_{2i})\frac{1}{1-\Phi_{\mathrm{NI}}(\xi_{2i})}\\
&= \frac{1}{\alpha}\frac{\partial}{\partial\boldsymbol{\beta}}(\mu_i)\frac{\xi_{1i}\phi_{\mathrm{NI}}(\xi_{2i})}{1-\Phi_{\mathrm{NI}}(\xi_{2i})}
\end{aligned}
$$

Resultados para $\widehat{s}_{i,\boldsymbol{\beta}}$

$$
\begin{aligned}
\widehat{s}_{i,\boldsymbol{\beta}} &= \begin{cases} \dfrac{\frac{\partial}{\partial\boldsymbol{\beta}}f_Y(y_i;\boldsymbol{\theta})}{f_Y(y_i;\boldsymbol{\theta})} & \text{si} \quad i\in D\\ \dfrac{\frac{\partial}{\partial\boldsymbol{\beta}}S_Y(y_i;\boldsymbol{\theta})}{S_Y(y_i;\boldsymbol{\theta})} & \text{si} \quad i\in C\end{cases}\\
&= \begin{cases} \dfrac{\xi_{2i}}{4\sqrt{2\pi}}\dfrac{\partial\mu_i}{\partial\boldsymbol{\beta}}\left(\xi_{i1}^2A_i(3/2)-A_i(1/2)\right) & \text{si} \quad i\in D,\\ \dfrac{1}{2}\dfrac{\partial\mu_i}{\partial\boldsymbol{\beta}}\dfrac{\xi_{i1}\phi_{\mathrm{NI}}(\xi_{2i})}{1-\Phi_{\mathrm{NI}}(\xi_{2i})} & \text{si} \quad i\in C,\end{cases}
\end{aligned}
$$

donde $A_i(w)$ se encuentra definido en 3.4.

## A.5. Propiedad de linearidad

Sea $Y \sim \mathrm{Senh-NI}(\alpha, \mu, \sigma; \mathrm{H})$ y $X = a + bY$ Entonces

Si $b > 0$

$$\begin{aligned} P(X \leq x) &= P(a + bY \leq x) \\ &= P(bY \leq x - a) \\ &= P(Y \leq \frac{x-a}{b}) \\ &= \Phi_{\mathrm{NI}}\left(\frac{2}{\alpha}\mathrm{senh}(\frac{\mathrm{x-a-b}\mu}{\mathrm{b}\sigma})\right) \end{aligned}$$

Si $b < 0$

$$\begin{aligned} P(X \leq x) &= P(a + bY \leq x) \\ &= P(bY \leq x - a) \\ &= P(Y \geq \frac{x-a}{b}) \\ &= 1 - \Phi_{\mathrm{NI}}\left(\frac{2}{\alpha}\mathrm{senh}(\frac{\mathrm{x-a-b}\mu}{\mathrm{b}\sigma})\right) \end{aligned}$$

Como senh(.) es una función impar, se cumple que $\mathrm{senh(w)} = -\mathrm{senh(-w)}$

$$P(X \leq x) = 1 - \Phi_{\mathrm{NI}}\left(-\frac{2}{\alpha}\mathrm{senh}(\frac{\mathrm{x-a-b}\mu}{-\mathrm{b}\sigma})\right)$$

Por lo tanto $a + bY \sim \mathrm{Senh-NI}(\alpha, \mathrm{a+b}\mu, |\mathrm{b}|\sigma)$

## A.6. Simulacion

Cuadro A.1: Sesgo, Sesgo relativo y RECM para diferentes tamaños de muestra n = (100, 500, 1000)

| | | | | Senh-NC | | | Senh-St | | | Senh-SL | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\alpha$ | Censura | Parámetro | Medida | 100 | 500 | 1000 | 100 | 500 | 1000 | 100 | 500 | 1000 |
| 0.50 | 0 % | $\alpha$ = 0.5 | Sesgo | -0.0026 | 0.0030 | 0.0028 | -0.0069 | -0.0005 | -0.0004 | -0.0059 | -0.0009 | -0.0004 |
| | | | RECM | 0.0398 | 0.0190 | 0.0135 | 0.0504 | 0.0222 | 0.0158 | 0.0380 | 0.0176 | 0.0123 |
| | | | Sesgo relat. % | -0.5 | 0.6 | 0.6 | -1.4 | -0.1 | -0.1 | -1.2 | -0.2 | -0.1 |
| | | $\beta_0$ = 4 | Sesgo | -0.0058 | -0.0060 | -0.0016 | 0.0013 | 0.0007 | 0.0036 | 0.0011 | -0.0006 | 0.0020 |
| | | | RECM | 0.1118 | 0.0503 | 0.0348 | 0.1206 | 0.0537 | 0.0383 | 0.1159 | 0.0531 | 0.0370 |
| | | | Sesgo relat. % | -0.1 | -0.2 | 0.0 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.1 |
| | | $\beta_1$ = -2 | Sesgo | 0.0112 | 0.0081 | 0.0018 | 0.0038 | -0.0029 | -0.0063 | 0.0007 | 0.0003 | -0.0033 |
| | | | RECM | 0.1991 | 0.0864 | 0.0605 | 0.2060 | 0.0918 | 0.0665 | 0.2075 | 0.0931 | 0.0651 |
| | | | Sesgo relat. % | -0.6 | -0.4 | -0.1 | -0.2 | 0.1 | 0.3 | 0.0 | 0.0 | 0.2 |
| | 10 % | $\alpha$ = 0.5 | Sesgo | 0.0056 | 0.0098 | 0.0103 | -0.0001 | 0.0065 | 0.0067 | -0.0045 | 0.0021 | 0.0022 |
| | | | RECM | 0.0449 | 0.0236 | 0.0181 | 0.0520 | 0.0246 | 0.0178 | 0.0407 | 0.0195 | 0.0144 |
| | | | Sesgo relat. % | 1.1 | 2.0 | 2.1 | 0.0 | 1.3 | 1.3 | -0.9 | 0.4 | 0.4 |
| | | $\beta_0$ = 4 | Sesgo | -0.0052 | -0.0011 | 0.0018 | 0.0026 | 0.0017 | 0.0046 | -0.0002 | -0.0037 | -0.0004 |
| | | | RECM | 0.1172 | 0.0534 | 0.0359 | 0.1251 | 0.0554 | 0.0404 | 0.1228 | 0.0576 | 0.0396 |
| | | | Sesgo relat. % | -0.1 | 0.0 | 0.0 | 0.1 | 0.0 | 0.1 | 0.0 | -0.1 | 0.0 |
| | | $\beta_1$ = -2 | Sesgo | 0.0250 | 0.0148 | 0.0120 | 0.0158 | 0.0094 | 0.0049 | 0.0182 | 0.0251 | 0.0185 |
| | | | RECM | 0.2124 | 0.0910 | 0.0632 | 0.2148 | 0.0956 | 0.0685 | 0.2203 | 0.1049 | 0.0738 |
| | | | Sesgo relat. % | -1.3 | -0.7 | -0.6 | -0.8 | -0.5 | -0.2 | -0.9 | -1.3 | -0.9 |
| | 20 % | $\alpha$ = 0.5 | Sesgo | 0.0133 | 0.0188 | 0.0198 | 0.0067 | 0.0145 | 0.0145 | -0.0012 | 0.0042 | 0.0048 |
| | | | RECM | 0.0522 | 0.0299 | 0.0262 | 0.0558 | 0.0289 | 0.0228 | 0.0437 | 0.0220 | 0.0167 |
| | | | Sesgo relat. % | 2.7 | 3.8 | 4.0 | 1.3 | 2.9 | 2.9 | -0.2 | 0.8 | 1.0 |
| | | $\beta_0$ = 4 | Sesgo | 0.0010 | 0.0043 | 0.0063 | 0.0047 | 0.0044 | 0.0071 | -0.0031 | -0.0033 | -0.0013 |
| | | | RECM | 0.1258 | 0.0580 | 0.0378 | 0.1336 | 0.0586 | 0.0434 | 0.1302 | 0.0617 | 0.0432 |
| | | | Sesgo relat. % | 0.0 | 0.1 | 0.2 | 0.1 | 0.1 | 0.2 | -0.1 | -0.1 | 0.0 |
| | | $\beta_1$ = -2 | Sesgo | 0.0297 | 0.0227 | 0.0199 | 0.0266 | 0.0217 | 0.0167 | 0.0422 | 0.0414 | 0.0381 |
| | | | RECM | 0.2261 | 0.0985 | 0.0691 | 0.2292 | 0.1014 | 0.0741 | 0.2350 | 0.1182 | 0.0889 |
| | | | Sesgo relat. % | -1.5 | -1.1 | -1.0 | -1.3 | -1.1 | -0.8 | -2.1 | -2.1 | -1.9 |
| | 40 % | $\alpha$ = 0.5 | Sesgo | 0.0391 | 0.0444 | 0.0461 | 0.0254 | 0.0356 | 0.0358 | 0.0066 | 0.0120 | 0.0120 |
| | | | RECM | 0.0764 | 0.0536 | 0.0509 | 0.0699 | 0.0462 | 0.0414 | 0.0540 | 0.0294 | 0.0235 |
| | | | Sesgo relat. % | 7.8 | 8.9 | 9.2 | 5.1 | 7.1 | 7.2 | 1.3 | 2.4 | 2.4 |
| | | $\beta_0$ = 4 | Sesgo | 0.0083 | 0.0134 | 0.0152 | 0.0109 | 0.0093 | 0.0125 | -0.0021 | -0.0087 | -0.0039 |
| | | | RECM | 0.1489 | 0.0681 | 0.0474 | 0.1573 | 0.0675 | 0.0502 | 0.1482 | 0.0682 | 0.0489 |
| | | | Sesgo relat. % | 0.2 | 0.3 | 0.4 | 0.3 | 0.2 | 0.3 | -0.1 | -0.2 | -0.1 |
| | | $\beta_1$ = -2 | Sesgo | 0.0522 | 0.0470 | 0.0436 | 0.0573 | 0.0542 | 0.0483 | 0.0872 | 0.0961 | 0.0870 |
| | | | RECM | 0.2649 | 0.1198 | 0.0879 | 0.2709 | 0.1233 | 0.0933 | 0.2708 | 0.1552 | 0.1259 |
| | | | Sesgo relat. % | -2.6 | -2.4 | -2.2 | -2.9 | -2.7 | -2.4 | -4.4 | -4.8 | -4.4 |

Cuadro A.2: Sesgo, Sesgo relativo y RECM para diferentes tamaños de muestra n = (100,500,1000)

| | | | | Senh-NC | | | Senh-St | | | Senh-SL | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\alpha$ | Censura | Parámetro | Medida | 100 | 500 | 1000 | 100 | 500 | 1000 | 100 | 500 | 1000 |
| 1.00 | 0 % | $\alpha$ = 1.0 | Sesgo | -0.0147 | -0.0031 | -0.0015 | -0.0141 | -0.0009 | -0.0009 | -0.0125 | -0.0019 | -0.0009 |
| | | | RECM | 0.0831 | 0.0389 | 0.0275 | 0.1006 | 0.0445 | 0.0315 | 0.0761 | 0.0352 | 0.0246 |
| | | | Sesgo relat. % | -1.5 | -0.3 | -0.2 | -1.4 | -0.1 | -0.1 | -1.3 | -0.2 | -0.1 |
| | | $\beta_0$ = 4 | Sesgo | -0.0055 | -0.0049 | -0.0012 | -0.0020 | 0.0018 | 0.0073 | 0.0012 | -0.0011 | 0.0035 |
| | | | RECM | 0.2109 | 0.0954 | 0.0645 | 0.2326 | 0.1034 | 0.0736 | 0.2095 | 0.0954 | 0.0662 |
| | | | Sesgo relat. % | -0.1 | -0.1 | 0.0 | -0.1 | 0.0 | 0.2 | 0.0 | 0.0 | 0.1 |
| | | $\beta_1$ = -2 | Sesgo | 0.0109 | 0.0091 | 0.0031 | 0.0147 | -0.0063 | -0.0128 | 0.0024 | 0.0009 | -0.0054 |
| | | | RECM | 0.3748 | 0.1610 | 0.1131 | 0.3984 | 0.1762 | 0.1278 | 0.3731 | 0.1670 | 0.1172 |
| | | | Sesgo relat. % | -0.5 | -0.5 | -0.2 | -0.7 | 0.3 | 0.6 | -0.1 | 0.0 | 0.3 |
| | 10 % | $\alpha$ = 1.0 | Sesgo | -0.0029 | 0.0085 | 0.0102 | -0.0014 | 0.0119 | 0.0120 | -0.0058 | 0.0057 | 0.0065 |
| | | | RECM | 0.0898 | 0.0420 | 0.0311 | 0.1035 | 0.0486 | 0.0348 | 0.0812 | 0.0393 | 0.0289 |
| | | | Sesgo relat. % | -0.3 | 0.9 | 1.0 | -0.1 | 1.2 | 1.2 | -0.6 | 0.6 | 0.7 |
| | | $\beta_0$ = 4 | Sesgo | -0.0028 | 0.0012 | 0.0048 | 0.0063 | 0.0091 | 0.0148 | 0.0087 | 0.0032 | 0.0083 |
| | | | RECM | 0.2209 | 0.0988 | 0.0664 | 0.2391 | 0.1062 | 0.0773 | 0.2188 | 0.1022 | 0.0705 |
| | | | Sesgo relat. % | -0.1 | 0.0 | 0.1 | 0.2 | 0.2 | 0.4 | 0.2 | 0.1 | 0.2 |
| | | $\beta_1$ = -2 | Sesgo | 0.0407 | 0.0319 | 0.0269 | 0.0371 | 0.0177 | 0.0092 | 0.0538 | 0.0603 | 0.0509 |
| | | | RECM | 0.3971 | 0.1682 | 0.1195 | 0.4120 | 0.1820 | 0.1300 | 0.3976 | 0.1916 | 0.1393 |
| | | | Sesgo relat. % | -2.0 | -1.6 | -1.3 | -1.9 | -0.9 | -0.5 | -2.7 | -3.0 | -2.5 |
| | 20 % | $\alpha$ = 1.0 | Sesgo | 0.0115 | 0.0220 | 0.0242 | 0.0105 | 0.0263 | 0.0259 | 0.0043 | 0.0160 | 0.0176 |
| | | | RECM | 0.0965 | 0.0487 | 0.0399 | 0.1101 | 0.0559 | 0.0434 | 0.0884 | 0.0456 | 0.0372 |
| | | | Sesgo relat. % | 1.2 | 2.2 | 2.4 | 1.1 | 2.6 | 2.6 | 0.4 | 1.6 | 1.8 |
| | | $\beta_0$ = 4 | Sesgo | 0.0049 | 0.0109 | 0.0137 | 0.0178 | 0.0206 | 0.0262 | 0.0083 | 0.0093 | 0.0106 |
| | | | RECM | 0.2311 | 0.1053 | 0.0692 | 0.2533 | 0.1118 | 0.0845 | 0.2256 | 0.1052 | 0.0779 |
| | | | Sesgo relat. % | 0.1 | 0.3 | 0.3 | 0.4 | 0.5 | 0.7 | 0.2 | 0.2 | 0.3 |
| | | $\beta_1$ = -2 | Sesgo | 0.0647 | 0.0546 | 0.0506 | 0.0574 | 0.0408 | 0.0313 | 0.1249 | 0.1231 | 0.1231 |
| | | | RECM | 0.4163 | 0.1814 | 0.1301 | 0.4343 | 0.1903 | 0.1394 | 0.4182 | 0.2276 | 0.1865 |
| | | | Sesgo relat. % | -3.2 | -2.7 | -2.5 | -2.9 | -2.0 | -1.6 | -6.2 | -6.2 | -6.2 |
| | 40 % | $\alpha$ = 1.0 | Sesgo | 0.0469 | 0.0576 | 0.0595 | 0.0441 | 0.0635 | 0.0634 | 0.0335 | 0.0436 | 0.0455 |
| | | | RECM | 0.1267 | 0.0776 | 0.0708 | 0.1358 | 0.0857 | 0.0754 | 0.1116 | 0.0682 | 0.0596 |
| | | | Sesgo relat. % | 4.7 | 5.8 | 6.0 | 4.4 | 6.4 | 6.3 | 3.4 | 4.4 | 4.6 |
| | | $\beta_0$ = 4 | Sesgo | 0.0296 | 0.0308 | 0.0345 | 0.0490 | 0.0483 | 0.0537 | 0.0291 | 0.0145 | 0.0184 |
| | | | RECM | 0.2638 | 0.1208 | 0.0846 | 0.2887 | 0.1302 | 0.1032 | 0.2422 | 0.1116 | 0.0845 |
| | | | Sesgo relat. % | 0.7 | 0.8 | 0.9 | 1.2 | 1.2 | 1.3 | 0.7 | 0.4 | 0.5 |
| | | $\beta_1$ = -2 | Sesgo | 0.1229 | 0.1201 | 0.1121 | 0.1097 | 0.0973 | 0.0880 | 0.2518 | 0.2733 | 0.2696 |
| | | | RECM | 0.4683 | 0.2272 | 0.1732 | 0.4971 | 0.2234 | 0.1706 | 0.4796 | 0.3344 | 0.3064 |
| | | | Sesgo relat. % | -6.1 | -6.0 | -5.6 | -5.5 | -4.9 | -4.4 | -12.6 | -13.7 | -13.5 |

Cuadro A.3: Sesgo, Sesgo relativo y RECM para diferentes tamaños de muestra n = (100,500,1000)

| | | | | Senh-NC | | | Senh-St | | | Senh-SL | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\alpha$ | Censura | Parámetro | Medida | 100 | 500 | 1000 | 100 | 500 | 1000 | 100 | 500 | 1000 |
| 1.50 | 0 % | $\alpha = 1.5$ | Sesgo | -0.0218 | -0.0046 | -0.0023 | -0.0211 | -0.0013 | -0.0012 | -0.0191 | -0.0028 | -0.0013 |
| | | | RECM | 0.1257 | 0.0584 | 0.0412 | 0.1520 | 0.0668 | 0.0473 | 0.1146 | 0.0529 | 0.0369 |
| | | | Sesgo relat. % | -1.5 | -0.3 | -0.2 | -1.4 | -0.1 | -0.1 | -1.3 | -0.2 | -0.1 |
| | | $\beta_0 = 4$ | Sesgo | -0.0066 | -0.0068 | -0.0016 | -0.0097 | 0.0028 | 0.0105 | 0.0011 | -0.0014 | 0.0041 |
| | | | RECM | 0.2874 | 0.1284 | 0.0868 | 0.3227 | 0.1426 | 0.1013 | 0.2718 | 0.1222 | 0.0845 |
| | | | Sesgo relat. % | -0.2 | -0.2 | 0.0 | -0.2 | 0.1 | 0.3 | 0.0 | 0.0 | 0.1 |
| | | $\beta_1 = -2$ | Sesgo | 0.0124 | 0.0125 | 0.0039 | 0.0309 | -0.0092 | -0.0184 | 0.0035 | 0.0016 | -0.0060 |
| | | | RECM | 0.5061 | 0.2169 | 0.1526 | 0.5547 | 0.2431 | 0.1760 | 0.4807 | 0.2137 | 0.1501 |
| | | | Sesgo relat. % | -0.6 | -0.6 | -0.2 | -1.5 | 0.5 | 0.9 | -0.2 | -0.1 | 0.3 |
| | 10 % | $\alpha = 1.5$ | Sesgo | -0.0060 | 0.0107 | 0.0131 | -0.0022 | 0.0175 | 0.0178 | 0.0007 | 0.0174 | 0.0194 |
| | | | RECM | 0.1357 | 0.0623 | 0.0458 | 0.1561 | 0.0727 | 0.0522 | 0.1233 | 0.0624 | 0.0472 |
| | | | Sesgo relat. % | -0.4 | 0.7 | 0.9 | -0.1 | 1.2 | 1.2 | 0.0 | 1.2 | 1.3 |
| | | $\beta_0 = 4$ | Sesgo | 0.0071 | 0.0102 | 0.0147 | 0.0105 | 0.0210 | 0.0291 | 0.0298 | 0.0246 | 0.0289 |
| | | | RECM | 0.2987 | 0.1319 | 0.0896 | 0.3314 | 0.1466 | 0.1073 | 0.2809 | 0.1305 | 0.0953 |
| | | | Sesgo relat. % | 0.2 | 0.3 | 0.4 | 0.3 | 0.5 | 0.7 | 0.7 | 0.6 | 0.7 |
| | | $\beta_1 = -2$ | Sesgo | 0.0358 | 0.0311 | 0.0241 | 0.0498 | 0.0121 | 0.0000 | 0.0568 | 0.0580 | 0.0518 |
| | | | RECM | 0.5373 | 0.2253 | 0.1588 | 0.5734 | 0.2502 | 0.1780 | 0.5058 | 0.2372 | 0.1747 |
| | | | Sesgo relat. % | -1.8 | -1.6 | -1.2 | -2.5 | -0.6 | 0.0 | -2.8 | -2.9 | -2.6 |
| | 20 % | $\alpha = 1.5$ | Sesgo | 0.0131 | 0.0283 | 0.0315 | 0.0148 | 0.0388 | 0.0382 | 0.0253 | 0.0417 | 0.0454 |
| | | | RECM | 0.1453 | 0.0703 | 0.0567 | 0.1667 | 0.0835 | 0.0645 | 0.1374 | 0.0784 | 0.0666 |
| | | | Sesgo relat. % | 0.9 | 1.9 | 2.1 | 1.0 | 2.6 | 2.5 | 1.7 | 2.8 | 3.0 |
| | | $\beta_0 = 4$ | Sesgo | 0.0257 | 0.0312 | 0.0350 | 0.0366 | 0.0456 | 0.0539 | 0.0567 | 0.0533 | 0.0534 |
| | | | RECM | 0.3095 | 0.1408 | 0.0962 | 0.3476 | 0.1563 | 0.1212 | 0.2886 | 0.1389 | 0.1060 |
| | | | Sesgo relat. % | 0.6 | 0.8 | 0.9 | 0.9 | 1.1 | 1.3 | 1.4 | 1.3 | 1.3 |
| | | $\beta_1 = -2$ | Sesgo | 0.0576 | 0.0509 | 0.0449 | 0.0626 | 0.0317 | 0.0181 | 0.1186 | 0.1181 | 0.1231 |
| | | | RECM | 0.5587 | 0.2374 | 0.1678 | 0.6010 | 0.2586 | 0.1871 | 0.5211 | 0.2600 | 0.2104 |
| | | | Sesgo relat. % | -2.9 | -2.5 | -2.2 | -3.1 | -1.6 | -0.9 | -5.9 | -5.9 | -6.2 |
| | 40 % | $\alpha = 1.5$ | Sesgo | 0.0583 | 0.0741 | 0.0773 | 0.0653 | 0.0936 | 0.0929 | 0.0831 | 0.1013 | 0.1065 |
| | | | RECM | 0.1843 | 0.1068 | 0.0957 | 0.2052 | 0.1274 | 0.1113 | 0.1865 | 0.1299 | 0.1224 |
| | | | Sesgo relat. % | 3.9 | 4.9 | 5.2 | 4.4 | 6.2 | 6.2 | 5.5 | 6.8 | 7.1 |
| | | $\beta_0 = 4$ | Sesgo | 0.0776 | 0.0769 | 0.0824 | 0.1013 | 0.1049 | 0.1120 | 0.1332 | 0.1228 | 0.1293 |
| | | | RECM | 0.3527 | 0.1670 | 0.1280 | 0.3948 | 0.1926 | 0.1616 | 0.3158 | 0.1821 | 0.1589 |
| | | | Sesgo relat. % | 1.9 | 1.9 | 2.1 | 2.5 | 2.6 | 2.8 | 3.3 | 3.1 | 3.2 |
| | | $\beta_1 = -2$ | Sesgo | 0.1152 | 0.1164 | 0.1040 | 0.1064 | 0.0805 | 0.0687 | 0.2158 | 0.2322 | 0.2259 |
| | | | RECM | 0.6259 | 0.2797 | 0.2052 | 0.6736 | 0.2869 | 0.2098 | 0.5569 | 0.3391 | 0.2863 |
| | | | Sesgo relat. % | -5.8 | -5.8 | -5.2 | -5.3 | -4.0 | -3.4 | -10.8 | -11.6 | -11.3 |

## A.7. Código R

```
# ESTUDIO DE SIMULACION - TESIS
# ******************************

## ================================================== FUNCIONES
##--- FUNCION DE DENSIDAD SINH NORMAL INDEPENDIENTE
logdbssmn.fdp<-function(y, x, alpha, beta, nu, mfamily = "t")
{
  if(!is.numeric(alpha)||!is.numeric(nu))
    if(alpha<=0){stop("alpha debe ser positivo")}
  psi2 = (2/alpha)*sinh((y-x%*%beta)/2)
  psi1 = (2/alpha)*cosh((y-x%*%beta)/2)

  if(mfamily == "t")
  {
    pdf = (gamma((nu+1)/2)/(sqrt(pi)*gamma(nu/2)))*(nu^(0.5*nu)/2)*(nu+psi2^2)^(-0.5*(nu+1))*psi1
  }

  if(mfamily == "sl")
  {
    val = vector(mode = "numeric", length = length(y))
    for(i in 1:length(y))
    {
      f = function(u) u^(nu-1)*dnorm(psi2[i],mean=0,sd=sqrt(1/u))
      val[i] = integrate(f,0,1)$value
    }
    pdf = 0.5*nu*val*psi1
  }

  if(mfamily == "cn")
  {
    pdf = 0.5*(nu[1]*sqrt(nu[2])*dnorm(sqrt(nu[2])*psi2)+(1-nu[1])*dnorm(psi2))*psi1
  }

  return(pdf)
}

##--- FUNCION DE DISTRIBUCION ACUMULADA SINH NORMAL INDEPENDIENTE
logbssmn.cdf<-function(q, y,x, alpha, beta, nu, mfamily="t")
{
  I = vector(mode = "numeric", length = length(q))

  if(mfamily == "t")
  {
    for(i in 1:length(q))
    {
      mu = x%*%beta
      yy = y[i]
      muu = mu[i]

      pdf = function(yy)
      {
        loga = log(gamma((nu+1)/2)/(sqrt(pi)*gamma(nu/2))) + ((0.5*nu)*log(nu)/2)
        logb = (-0.5*(nu+1))*log(nu+((2/alpha)*sinh((yy-muu)/2))^2)
        logc = log(2/alpha) + log(cosh((yy-muu)/2))
        return(ifelse(yy == 0 | is.nan(exp(loga + logb + logc)), 0, exp(loga + logb + logc))) #modif
      }

      I[i] = integrate(pdf,0,q[i],stop.on.error=FALSE)$value
    }
  }

  if(mfamily == "sl")
  {
    for(i in 1:length(q))
    {
      mu   = x%*%beta
      yy   = y[i]
      muu  = mu[i]
      g    = function(yy,u) { #modif
        tempo = nu*u^(nu-1)*dnorm((2/alpha)*sinh((yy-muu)/2),mean=0,sd=sqrt(1/u))*((2/alpha)*cosh((yy-muu)/2))
        return(ifelse(tempo == 0 | is.nan(tempo), 0, tempo))
      }
      I[i] = integrate(function(yy) {sapply(yy, function(yy) {integrate(function(u) g(yy,u), 0, 1)$value})}, 0, q[i])$value
    }
  }

```

```
  if(mfamily == "cn")
  {
    for(i in 1:length(q))
    {
      mu    = x%*%beta
      yy    = y[i]
      muu   = mu[i]
      pdf   = function(yy) { #modif
        tempo = 0.5*(nu[1]*sqrt(nu[2])*dnorm(sqrt(nu[2])*((2/alpha)*sinh((yy-muu)/2)))+(1-nu[1])*dnorm((2/alpha)*sinh((yy-muu)/2)))*((2/alpha)*cosh((yy-muu)/2))
        return(ifelse(tempo == 0 | is.nan(tempo), 0, tempo))
      }
      I[i] = integrate(pdf,0,q[i])$value
    }
  }

  return(I)
}

##-- FUNCION NI
NI.fdp<-function(y, x, alpha, beta, nu, mfamily = "t")
{
  if(!is.numeric(alpha)||!is.numeric(nu))
    if(alpha<=0){stop("alpha must be positive")}
  psi2 <- (2/alpha)*sinh((y-x%*%beta)/2)
  psi1 <- (2/alpha)*cosh((y-x%*%beta)/2)
  if(mfamily == "t")
  {
    pdf  <- (gamma((nu+1)/2)/(sqrt(pi)*gamma(nu/2)))*(nu^(0.5*nu))*(nu+psi2^2)^(-0.5*(nu+1))
  }

  if(mfamily == "cn")
  {
    pdf  <- (nu[1]*sqrt(nu[2])*dnorm(sqrt(nu[2])*psi2)+(1-nu[1])*dnorm(psi2))
  }

  if(mfamily == "sl")
  {
    val<-vector(mode = "numeric", length = length(y))
    for(i in 1:length(y))
    {
      f    <- function(u) u^(nu-1)*dnorm(psi2[i],mean=0,sd=sqrt(1/u))
      val[i]<-integrate(f,0,1)$value
    }
    pdf  <- nu*val
  }
  return(pdf)
}

##--- GENERA NUMEROS ALEATORIOS DE LA SINH NORMAL INDEPENDIENTE
r_shni = function(n, x, alpha, beta, nu, mfamily = "t")
{
  z = switch(mfamily,
             t = sqrt(1/rgamma(n,shape=nu/2,rate=nu/2))*rnorm(n),
             sl = sqrt(1/rbeta(n,shape1=nu,shape=1))*rnorm(n),
             cn = rmix(n, c(nu[1],1-nu[1]), "Skew.normal", list(c(0,1/nu[2],0), c(0,1,0))) )
  y = x%*%beta + 2*asinh(alpha*z*0.5)
  return(y)
}

##--- GENERADOR DE DATOS
gen_data = function(n, pc, alfa, beta, nu, mfamily="t"){
  # pc = porcentaje de datos censurados
  # n = tamano muestra
  # Genera X
  x = cbind(x1=1,x2=runif(n))
  if (mfamily == "t"){
    # Genera Y
    y = r_shni(n, x, alfa, beta, nu, mfamily="t")
  }

  if (mfamily == "sl"){
    y = r_shni(n, x, alfa, beta, nu, mfamily="sl")
  }

  if (mfamily  == "cn"){
    y = r_shni(n, x, alfa, beta, nu, mfamily="cn")
  }

  data = data.frame(x, y, status = c(rep(1,n*(1-pc)), rep(0, n*pc))) # los 0 son censurados
```

```
  return(data)
}

##--- FUNCION DE LOG-VEROSIMILITUD
log_lk_shni = function(param, y, x, status, nu, mfamily = "t"){
  alpha = param[1]
  beta = c(param[2],param[3])
  p = length(beta)
  mu = (x %*% beta)
  xi1 = (2 / alpha) * cosh((y - mu) / 2)
  xi2 = (2 / alpha) * sinh((y - mu) / 2)

  if(mfamily == "t"){
    lk1 = sum(status*(log((gamma((nu+1)/2)/(sqrt(pi)*gamma(nu/2)))))) + sum(status*log((nu^(0.5*nu)/2)*(nu+xi2^2)
         ^(-0.5*(nu+1))*xi1))
    lk2 = sum((1-status)*log(1-logbssmn.cdf(xi2, y, x, alpha, beta, nu, mfamily="t")))
    lk = lk1 + lk2
  }

  if(mfamily == "sl"){
    val<-vector(mode = "numeric", length = length(y))
    for(i in 1:length(y))
    {
      f = function(u) u^(nu-1)*dnorm(xi2[i],mean=0,sd=sqrt(1/u))
      val[i] = integrate(f,0,1)$value
    }
    pdf = 0.5*nu*val*xi1
    lk = sum( status*(log(pdf)) + (1-status)*log(abs(1-logbssmn.cdf(xi2, y, x, alpha, beta, nu, mfamily="sl"))))
  }

  if(mfamily == "cn"){
    lk = sum(
      status*(log(0.5*(nu[1]*sqrt(nu[2])*dnorm(sqrt(nu[2])*xi2)+(1-nu[1])*dnorm(xi2))*xi1)) +
        (1-status)*log(1-logbssmn.cdf(xi2, y,x,alpha, beta, nu, mfamily = "cn"))
      )
  }

  return(-lk)
}

##--- ESTIMA PARAMETROS DEL MODELO
lm_senhniml = function(y, x, status, nu, mfamily = "t"){
  fit    = lm.fit(x, y)
  beta   = c(fit$coef)
  p      = length(beta)
  n      = length(y)
  n1     = sum(status)
  mu     = x %*% beta
  alpha  = sqrt((4 / n1 ) * sum((sinh((y - mu)/2))^2))
  beta0  = beta
  alpha0 = alpha

  if(mfamily == "t"){
    tmp = optim(par = c(alpha0, beta0), log_lk_shni, y=y, x=x, status=status, nu=nu, mfamily="t")
  }

  if(mfamily == "sl"){
    tmp = optim(par = c(alpha0, beta0), log_lk_shni, y=y, x=x, status=status, nu=nu, mfamily="sl")
  }

  if(mfamily == "cn"){
    tmp = optim(par = c(alpha0, beta0), log_lk_shni, y=y, x=x, status=status, nu=nu, mfamily="cn")
  }

  if(p == 6)  beta = c(tmp$par[2], tmp$par[3], tmp$par[4], tmp$par[5], tmp$par[6], tmp$par[7]) #modif
  if(p == 5)  beta = c(tmp$par[2], tmp$par[3], tmp$par[4], tmp$par[5], tmp$par[6]) #modif
  if(p == 4)  beta = c(tmp$par[2], tmp$par[3], tmp$par[4], tmp$par[5]) #modif
  if(p == 3)  beta = c(tmp$par[2], tmp$par[3], tmp$par[4]) #modif
  if(p == 2)  beta = c(tmp$par[2], tmp$par[3]) #modif

  alpha = tmp$par[1]
  theta = c(beta,alpha)
  coef = beta
  alphaest = alpha

  AIC = - 2 * (-log_lk_shni(c(alpha,beta),y,x,status,nu, mfamily=mfamily)) + 2 * (length(theta) + 1)
  BIC = - 2 * (-log_lk_shni(c(alpha,beta),y,x,status,nu,mfamily=mfamily)) + (length(theta) + 1)  * log(length(y))

  lk = -log_lk_shni(c(alpha,beta),y,x,status,nu, mfamily=mfamily)

```

```
  result = list(alpha=alphaest, beta=coef, AIC=AIC, BIC=BIC, lk=lk)

  return(result)
}

##--- OBTIENE MATRIZ DE INFORMACION Y ERROR STANDARD
im_shni = function(y, x, status, theta, mfamily = "t")
{
  n     = length(y)
  p     = length(theta)-1
  mu    = (x %*% theta[2:(p+1)])
  xi1   = (2 / theta[1]) * cosh((y - mu) / 2)
  xi2   = (2 / theta[1]) * sinh((y - mu) / 2)
  dnlm = t(x)
  fy    = logdbssmn.fdp(y, x, alpha, beta, nu, mfamily)
  fy2   = NI.fdp(y, x, alpha, beta, nu, mfamily)
  cdf   = logbssmn.cdf(xi2, y, x, alpha, beta, nu, mfamily)

  if (mfamily == "t")
  {
    soma = 0
    Aiw = function(w=0, di=0, nu=0) as.numeric((2^w*nu^(nu/2)*gamma(w + nu/2))/(gamma(nu/2)*(nu + di)^(nu/2 + w)))
    for (i in 1:n)
    {
      S = c() #vector con todas las derivadas en relacion a cada parametro desconocido del modelo
      di = xi2[i]^2

      #derivadas
      dxi2alpha = -(1/theta[1])*xi2[i]
      dxi2beta  = - 0.5*dnlm[,i]*xi1[i]

      dxi1alpha = -(1/theta[1])*xi1[i]
      dxi1beta = - 0.5*dnlm[,i]*xi2[i]

      Ssialpha = status[i]*((1/fy[i])*(0.5/sqrt(2*pi))*(dxi1alpha *Aiw(1/2, di, nu) - xi1[i]*xi2[i]*dxi2alpha*Aiw(3/2, di, nu)))+
        (1-status[i])*(-((fy2[i])*dxi2alpha)/(1-cdf[i]))
      Ssibeta = status[i]*((1/fy[i])*(0.5/sqrt(2*pi))*(dxi1beta*Aiw(1/2, di, nu) - xi1[i]*xi2[i]*dxi2beta*Aiw(3/2, di, nu)))+
        (1-status[i])*(-((fy2[i])*dxi2beta)/(1-cdf[i]))

      S = c(S, Ssibeta, Ssialpha)
      soma = soma + S%*%t(S)
    }
  }

  if (mfamily == "sl")
  {
    soma       <- 0
    Aiw <- function(w=0, di=0, nu=0) as.numeric(((nu*2^(w + nu)*gamma(w + nu))/(di^(w + nu)))*pgamma(1, w + nu, di/2))
    for (i in 1:n)
    {
      S          <- c() #vector con todas las derivadas en relacion a cada parametro desconocido del modelo
      di         <- xi2[i]^2

      #derivadas
      dxi2alpha    <- -(1/theta[1])*xi2[i]
      dxi2beta     <- - 0.5*dnlm[,i]*xi1[i]

      dxi1alpha    <- -(1/theta[1])*xi1[i]
      dxi1beta     <- - 0.5*dnlm[,i]*xi2[i]

      Ssialpha     <- status[i]*((1/fy[i])*(0.5/sqrt(2*pi))*(dxi1alpha *Aiw(1/2, di, nu) - xi1[i]*xi2[i]*dxi2alpha*Aiw(3/2, di, nu)))+(1-status[i])*(-((fy2[i])*dxi2alpha)/(1-cdf[i]))
      Ssibeta      <- status[i]*((1/fy[i])*(0.5/sqrt(2*pi))*(dxi1beta*Aiw(1/2, di, nu) - xi1[i]*xi2[i]*dxi2beta*Aiw(3/2, di, nu)))+(1-status[i])*(-((fy2[i])*dxi2beta)/(1-cdf[i]))

      S            <- c(S, Ssibeta, Ssialpha)
      soma         <- soma + S%*%t(S)
    }
  }

  if (mfamily == "cn")
  {
    soma       <- 0
    Aiw <- function(w=0, di=0, nu=0) as.numeric(sqrt(2*pi)*(nu[1]*nu[2]^(w -0.5)*dnorm(sqrt(di), 0, sqrt(1/nu[2]))
      + (1 - nu[1])*(dnorm(sqrt(di), 0,1))))
    for (i in 1:n)
    {
```

```
      S          <- c() #vector con todas las derivadas en relacion a cada parametro desconocido del modelo
      di         <- xi2[i]^2

      #derivadas
      dxi2alpha    <- -(1/theta[1])*xi2[i]
      dxi2beta     <- - 0.5*dnlm[,i]*xi1[i]

      dxi1alpha    <- -(1/theta[1])*xi1[i]
      dxi1beta     <- - 0.5*dnlm[,i]*xi2[i]

      Ssialpha     <- status[i]*((1/fy[i])*(0.5/sqrt(2*pi))*(dxi1alpha *Aiw(1/2, di, nu) - xi1[i]*xi2[i]*dxi2alpha*Aiw(3/2, di, nu)))+(1-status[i])*(-((fy2[i])*dxi2alpha)/(1-cdf[i]))
      Ssibeta      <- status[i]*((1/fy[i])*(0.5/sqrt(2*pi))*(dxi1beta*Aiw(1/2, di, nu) - xi1[i]*xi2[i]*dxi2beta*Aiw(3/2, di, nu)))+(1-status[i])*(-((fy2[i])*dxi2beta)/(1-cdf[i]))

      S            <- c(S, Ssibeta, Ssialpha)
      soma         <- soma + S%*%t(S)
    }
  }

  EP           <-  sqrt(diag(solve(soma)))
  return(list(IM=soma,EP=EP))
}

## ================================================== SIMULACION

library(mixsmsn)

resumen = data.frame(matrix(ncol = 11, nrow = 0)) # Genera cabecera de tabla resumen
val_muestra = c(100, 500, 1000)
val_alpha = c(0.5, 1.0, 1.5)
val_pc = c(0, 0.1, 0.2, 0.4)

for(a in val_muestra){
for(b in val_alpha){
for(c in val_pc){

familia = "t" # valores permitidos t, sl, cn
nro_sim = 1000
n =  a # tamanno de muestra
alpha = b
beta = c(4,-2)
sigma = 2
nu = 3 # Para "cn" nu = c(0.1,0.1). Para "t" y "sl" asignar nu = 3
pc = c # nivel de censura

set.seed(12345)

par_est = array(0,dim=c(nrow=nro_sim,ncol=3))
se_est = array(0,dim=c(nrow=nro_sim, ncol=3))
i = 1

while (i <= nro_sim){
  flag = TRUE
  tryCatch(
    expr = {
      # genera data
      data = gen_data(n, pc, alpha, beta, nu, mfamily = familia)
      # recupera parametros
      recup_p = lm_senhniml(data$y, cbind(data$x1,data$x2), data$status, nu, mfamily = familia)
      # estima error standard
      estima_se = im_shni(data$y, cbind(data$x1,data$x2), data$status, theta=c(recup_p$alpha,recup_p$beta), mfamily = familia)
      # muestra y guarda parametros recuperados
      if (i == 1){
        cat("muestra ","Nro","====>", "n","pc" ,"alpha", "beta0", "beta1", "se_alpha" ,"se_beta0", "se_beta1", '\n') #cabecera
      }
      cat("muestra ", i, "====>", a, c, recup_p$alpha, recup_p$beta, estima_se$EP, '\n')
      par_est[i,] = c(recup_p$alpha, recup_p$beta)
      se_est[i,] = c(estima_se$EP)
      i = i + 1
    },
    error = function(e){
      flag = FALSE
    },
    warning = function(e){
      flag = FALSE
    }
  )
```

```r
  if(!flag) next
}

vt = c(alpha, beta)# valores teoricos
vtmatrix = matrix(vt, nc=length(vt), nr=nro_sim, byrow=TRUE) # matriz de valores teoricos

mean_param = array(0, dim = c(nrow = 1, ncol = length(vt)))
sd_param   = array(0, dim = c(nrow = 1, ncol = length(vt)))
mean_se    = array(0, dim = c(nrow = 1, ncol = length(vt)))
vies       = array(0, dim = c(nrow = 1, ncol = length(vt)))
rmse       = array(0, dim = c(nrow = 1, ncol = length(vt)))

mean_param[1,]  = round(apply(par_est,2,mean),digits=4) # Mean
sd_param[1,]    = round(apply(par_est,2,sd),digits=4) # MC Sd
mean_se[1,]     = round(apply(se_est,2,mean),digits=4) # IM SE
vies[1,]        = round((apply(par_est-vtmatrix,2,mean)),digits=4) # BIAS
rmse[1,]        = round(sqrt(apply((par_est-vtmatrix)^2,2,mean)),digits=4) # RMSE

alpha_csup = par_est[,1] + 1.96*se_est[,1];   alpha_cinf = par_est[,1] - 1.96*se_est[,1]
beta1_csup = par_est[,2] + 1.96*se_est[,2];   beta1_cinf = par_est[,2] - 1.96*se_est[,2]
beta2_csup = par_est[,3] + 1.96*se_est[,3];   beta2_cinf = par_est[,3] - 1.96*se_est[,3]

count.a = c()
count.b1 = c()
count.b2 = c()

for(i in 1:nro_sim)
{
  if(alpha_cinf[i]  <= alpha   && alpha_csup[i] >= alpha)   count.a[i]    = 1 else count.a[i]    = 0
  if(beta1_cinf[i]  <= beta[1] && beta1_csup[i] >= beta[1]) count.b1[i]   = 1 else count.b1[i]   = 0
  if(beta2_cinf[i]  <= beta[2] && beta2_csup[i] >= beta[2]) count.b2[i]   = 1 else count.b2[i]   = 0
}

COVMC_alpha  = sum(count.a)/nro_sim
COVMC_beta1  = sum(count.b1)/nro_sim
COVMC_beta2  = sum(count.b2)/nro_sim

COV = 100*c(COVMC_alpha,COVMC_beta1,COVMC_beta2) # COBERTURA


#COV

# Genera contenido de tabla resumen
nuevafila = data.frame(rep(familia,3),rep(n,3), rep(pc,3), c("alpha","beta0","beta1"), vt, c(mean_param), c(sd_param), c(mean_se), c(vies), c(rmse), COV)
colnames(nuevafila) = c("Familia","TamMuestra", "NivelCens", "NomParam", "ValorParam", "MEAN", "MCSD", "IMSE", "BIAS", "RMSE", "COV")
resumen = rbind(resumen, nuevafila)

}
}
}

resumen
write.csv(resumen, "resumen_t.csv")

## =================================================== APLICACION: MYELOMA

# LECTURA DE DATA
setwd("C:/Users/Max Alonzo/Dropbox/MF-BS-ST/Tesis_MaxAlonzo_codigoR")
myeloma = read.table("myeloma.txt")
x1      = myeloma$logbun # log de la medicion de nitrogeno ureico en sangre
x2      = myeloma$hgb # hemoglobina
x3      = myeloma$age # edad
x4      = myeloma$sex # sexo (0:varon; 1:mujer)
x5      = myeloma$serum.calcium # cantidad de calcio en la sangre
y       = myeloma$time # tiempo de supervivencia (en meses)
status  = myeloma$status # 0:sobrevive; 1:muere | (total 17 censurados = 17 sobrevivientes)

# ajuste en codigos para sexo
x4 = ifelse(x4 == 1, 0, 1)

# EXPLORACION DE DATA
library(ggplot2)
data_1 = data.frame(cbind(y, x1, x2, x3, x4, x5, status))
summary(data_1)
dim(data_1) # 65 filas (pacientes) y 7 columnas

# correlacion
ggplot(data = data_1) + aes(x = x1, y = y, color = factor(status)) + geom_point(size = 0.9)
```

```
ggplot(data = data_1) + aes(x = x2, y = y, color = factor(status)) + geom_point(size = 0.9)
ggplot(data = data_1) + aes(x = x3, y = y, color = factor(status)) + geom_point(size = 0.9)
ggplot(data = data_1) + aes(x = x5, y = y, color = factor(status)) + geom_point(size = 0.9)

# histograma y densidades de la variable dependiente
ggplot(data = data_1, aes(x= log(y))) + geom_histogram(color = "black", fill = "gray") # log(y)
ggplot(data = data_1, aes(x= y)) + geom_histogram(color = "black", fill = "gray") # y

# ESTIMACION DE PARAMETROS (REGRESION SINH-NI CON CENSURA)
model_t_5 = lm_senhniml(log(y), cbind(1, x1, x2, x3, x4, x5), status, nu=3, "t")
model_sl_5 = lm_senhniml(log(y), cbind(1, x1, x2, x3, x4, x5), status, nu=3, "sl")
model_cn_5 = lm_senhniml(log(y), cbind(1, x1, x2, x3, x4, x5), status, nu=c(0.1,0.1), "cn")

model_t_5$alpha; model_t_5$beta
model_sl_5$alpha; model_sl_5$beta
model_cn_5$alpha; model_cn_5$beta

# OBTENCION DEL ERROR STANDARD (REGRESION SINH-NI CON CENSURA)
im_shni(log(y), cbind(1, x1, x2, x3, x4, x5), status, theta=c(model_t_5$alpha, model_t_5$beta), nu=3, mfamily = "t"
    )$EP
im_shni(log(y), cbind(1, x1, x2, x3, x4, x5), status, theta=c(model_sl_5$alpha, model_sl_5$beta), nu=3, mfamily = "
    sl")$EP
im_shni(log(y), cbind(1, x1, x2, x3, x4, x5), status, theta=c(model_cn_5$alpha, model_cn_5$beta), nu=c(0.1,0.1),
    mfamily = "cn")$EP

# SELECCION DE VARIABLES

model_t_14  = lm_senhniml(log(y), cbind(1, x1, x2, x3, x4), status, nu=3, "t")
model_sl_14 = lm_senhniml(log(y), cbind(1, x1, x2, x3, x4), status, nu=3, "sl")
model_cn_14 = lm_senhniml(log(y), cbind(1, x1, x2, x3, x4), status, nu=c(0.9,0.9), "cn")
c(model_t_14$AIC,model_t_14$BIC)
c(model_sl_14$AIC,model_sl_14$BIC)
c(model_cn_14$AIC,model_cn_14$BIC)

model_t_24  = lm_senhniml(log(y), cbind(1, x1, x2, x3, x5), status, nu=3, "t")
model_sl_24 = lm_senhniml(log(y), cbind(1, x1, x2, x3, x5), status, nu=3, "sl")
model_cn_24 = lm_senhniml(log(y), cbind(1, x1, x2, x3, x5), status, nu=c(0.9,0.9), "cn")

c(model_t_24$AIC,model_t_24$BIC)
c(model_sl_24$AIC,model_sl_24$BIC)
c(model_cn_24$AIC,model_cn_24$BIC)

model_t_34  = lm_senhniml(log(y), cbind(1, x1, x2, x4, x5), status, nu=3, "t")
model_sl_34 = lm_senhniml(log(y), cbind(1, x1, x2, x4, x5), status, nu=3, "sl")
model_cn_34 = lm_senhniml(log(y), cbind(1, x1, x2, x4, x5), status, nu=c(0.9,0.9), "cn")

c(model_t_34$AIC,model_t_34$BIC)
c(model_sl_34$AIC,model_sl_34$BIC)
c(model_cn_34$AIC,model_cn_34$BIC)

model_t_44  = lm_senhniml(log(y), cbind(1, x1, x3, x4, x5), status, nu=3, "t")
model_sl_44 = lm_senhniml(log(y), cbind(1, x1, x3, x4, x5), status, nu=3, "sl")
model_cn_44 = lm_senhniml(log(y), cbind(1, x1, x3, x4, x5), status, nu=c(0.9,0.9), "cn")

c(model_t_44$AIC,model_t_44$BIC)
c(model_sl_44$AIC,model_sl_44$BIC)
c(model_cn_44$AIC,model_cn_44$BIC)

model_t_54  = lm_senhniml(log(y), cbind(1, x2, x3, x4, x5), status, nu=3, "t")
model_sl_54 = lm_senhniml(log(y), cbind(1, x2, x3, x4, x5), status, nu=3, "sl")
model_cn_54 = lm_senhniml(log(y), cbind(1, x2, x3, x4, x5), status, nu=c(0.9,0.9), "cn")

c(model_t_54$AIC,model_t_54$BIC)
c(model_sl_54$AIC,model_sl_54$BIC)
c(model_cn_54$AIC,model_cn_54$BIC)


model_t_13  = lm_senhniml(log(y), cbind(1, x1, x2, x3), status, nu=3, "t")
model_sl_13 = lm_senhniml(log(y), cbind(1, x1, x2, x3), status, nu=3, "sl")
model_cn_13 = lm_senhniml(log(y), cbind(1, x1, x2, x3), status, nu=c(0.9,0.9), "cn")

c(model_t_13$AIC,model_t_13$BIC)
c(model_sl_13$AIC,model_sl_13$BIC)
c(model_cn_13$AIC,model_cn_13$BIC)

model_t_23  = lm_senhniml(log(y), cbind(1, x1, x2, x4), status, nu=3, "t")
model_sl_23 = lm_senhniml(log(y), cbind(1, x1, x2, x4), status, nu=3, "sl")
model_cn_23 = lm_senhniml(log(y), cbind(1, x1, x2, x4), status, nu=c(0.9,0.9), "cn")

c(model_t_23$AIC,model_t_23$BIC)
```

```
c(model_sl_23$AIC,model_sl_23$BIC)
c(model_cn_23$AIC,model_cn_23$BIC)

model_t_33  = lm_senhniml(log(y), cbind(1, x1, x3, x4), status, nu=3, "t")
model_sl_33 = lm_senhniml(log(y), cbind(1, x1, x3, x4), status, nu=3, "sl")
model_cn_33 = lm_senhniml(log(y), cbind(1, x1, x3, x4), status, nu=c(0.9,0.9), "cn")

c(model_t_33$AIC,model_t_33$BIC)
c(model_sl_33$AIC,model_sl_33$BIC)
c(model_cn_33$AIC,model_cn_33$BIC)

model_t_43  = lm_senhniml(log(y), cbind(1, x1, x3, x5), status, nu=3, "t")
model_sl_43 = lm_senhniml(log(y), cbind(1, x1, x3, x5), status, nu=3, "sl")
model_cn_43 = lm_senhniml(log(y), cbind(1, x1, x3, x5), status, nu=c(0.9,0.9), "cn")

c(model_t_43$AIC,model_t_43$BIC)
c(model_sl_43$AIC,model_sl_43$BIC)
c(model_cn_43$AIC,model_cn_43$BIC)

model_t_53  = lm_senhniml(log(y), cbind(1, x2, x3, x5), status, nu=3, "t")
model_sl_53 = lm_senhniml(log(y), cbind(1, x2, x3, x5), status, nu=3, "sl")
model_cn_53 = lm_senhniml(log(y), cbind(1, x2, x3, x5), status, nu=c(0.9,0.9), "cn")

c(model_t_53$AIC,model_t_53$BIC)
c(model_sl_53$AIC,model_sl_53$BIC)
c(model_cn_53$AIC,model_cn_53$BIC)

model_t_63  = lm_senhniml(log(y), cbind(1, x3, x4, x5), status, nu=3, "t")
model_sl_63 = lm_senhniml(log(y), cbind(1, x3, x4, x5), status, nu=3, "sl")
model_cn_63 = lm_senhniml(log(y), cbind(1, x3, x4, x5), status, nu=c(0.9,0.9), "cn")

c(model_t_63$AIC,model_t_63$BIC)
c(model_sl_63$AIC,model_sl_63$BIC)
c(model_cn_63$AIC,model_cn_63$BIC)

model_t_73  = lm_senhniml(log(y), cbind(1, x2, x3, x4), status, nu=3, "t")
model_sl_73 = lm_senhniml(log(y), cbind(1, x2, x3, x4), status, nu=3, "sl")
model_cn_73 = lm_senhniml(log(y), cbind(1, x2, x3, x4), status, nu=c(0.9,0.9), "cn")

c(model_t_73$AIC,model_t_73$BIC)
c(model_sl_73$AIC,model_sl_73$BIC)
c(model_cn_73$AIC,model_cn_73$BIC)


model_t_12  = lm_senhniml(log(y), cbind(1, x1, x2), status, nu=3, "t")
model_sl_12 = lm_senhniml(log(y), cbind(1, x1, x2), status, nu=3, "sl")
model_cn_12 = lm_senhniml(log(y), cbind(1, x1, x2), status, nu=c(0.9,0.9), "cn")


c(model_t_12$AIC,model_t_12$BIC)
c(model_sl_12$AIC,model_sl_12$BIC)
c(model_cn_12$AIC,model_cn_12$BIC)


model_t_22  = lm_senhniml(log(y), cbind(1, x1, x3), status, nu=3, "t")
model_sl_22 = lm_senhniml(log(y), cbind(1, x1, x3), status, nu=3, "sl")
model_cn_22 = lm_senhniml(log(y), cbind(1, x1, x3), status, nu=c(0.9,0.9), "cn")

c(model_t_22$AIC,model_t_22$BIC)
c(model_sl_22$AIC,model_sl_22$BIC)
c(model_cn_22$AIC,model_cn_22$BIC)

model_t_32  = lm_senhniml(log(y), cbind(1, x1, x4), status, nu=3, "t")
model_sl_32 = lm_senhniml(log(y), cbind(1, x1, x4), status, nu=3, "sl")
model_cn_32 = lm_senhniml(log(y), cbind(1, x1, x4), status, nu=c(0.9,0.9), "cn")

c(model_t_32$AIC,model_t_32$BIC)
c(model_sl_32$AIC,model_sl_32$BIC)
c(model_cn_32$AIC,model_cn_32$BIC)

model_t_42  = lm_senhniml(log(y), cbind(1, x1, x5), status, nu=3, "t")
model_sl_42 = lm_senhniml(log(y), cbind(1, x1, x5), status, nu=3, "sl")
model_cn_42 = lm_senhniml(log(y), cbind(1, x1, x5), status, nu=c(0.9,0.9), "cn")

c(model_t_42$AIC,model_t_42$BIC)
c(model_sl_42$AIC,model_sl_42$BIC)
c(model_cn_42$AIC,model_cn_42$BIC)


model_t_11  = lm_senhniml(log(y), cbind(1, x1), status, nu=3, "t")
```

```
model_sl_11 = lm_senhniml(log(y), cbind(1, x1), status, nu=3, "sl")
model_cn_11 = lm_senhniml(log(y), cbind(1, x1), status, nu=c(0.9,0.9), "cn")
```

# Bibliografía


Akaike, H. (1974). A new look at the statistical model identification, *Autom Control IEEE Trans* **19**: 716–723.

Balakrishnan, N., Leiva, V., Sanhueza, A. y Vilca, F. (2009). Estimation in the birnbaum-saunders distribution based on scale-mixture of normals and the em-algorithm, *Stat. Oper. Res. Trans* **33**: 171–192.

Barros, M., Paula, G. y Leiva, V. (2008). A new class of survival regression models with heavy-tailed errors: robustness and diagnostics, *Lifetime Data Analysis* **14(3)**: 316–332.

Basford, K., Greenway, D., McLachlan, G. y Peel, D. (1997). Standard errors of fitted component means of normal mixtures, *Computational Statistics* **12**: 1–18.

Birnbaum, Z. W. y Saunders, S. C. (1969). A new family of life distributions, *J. Appl. Probab* **6**: 637–652.

Bogaerts, K., Komarek, A. y Lesaffre, E. (2017). *Survival Analysis with Interval-Censored Data*, CRC Press.

Johnson, N. L. (1949). Systems of frequency curves generated by methods of translation, *Biometrika* **36**: 146–176.

Johnson, N. L. y Kotz, S. (1970). *Distributions in Statistics Continuous Univariate Distributions, Vol. 1*, John Wiley & Sons, New York.

Klein, J. y Moeschberger, M. (2003). *Survival Analysis: Techniques for Censored and Truncated Data*, 2nd edn, Springer Science & Business Media.

Kleinbaum, D. y Klein, M. (2011). *Survival Analysis*, Springer.

Lachos, V., Dey, D. K., Cancho, V. G. y Louzada, F. (2017). Scale mixtures log-birnbaum-saunders regression models with censored data: a bayesian approach, *Journal of Statistical Computation and Simulation* **87**(10): 2002–2022.

Lange, K. y Sinsheimer, J. (1993). Normal/independent distributions and their applications in robust regression, *J Comput. Graph. Stat.* pp. 175–198.

Leiva, V., Barros, M., Paula, G. A. y Galea, M. (2007). Influence diagnostics in log-birnbaum–saunders regression models with censored data, *Computational Statistics & Data Analysis* **51**(12): 5694 – 5707.

Leiva, V., Barros, M., Paula, G. y Sanhueza, A. (2008). Generalized birnbaum-saunders distributions applied to air pollutant concentration, *Environmetrics* **19**: 235–249.

Leiva, V., Vilca, F., Balakrishnan, N. y Sanhueza, A. (2010). A skewed sinh-normal distribution and its properties and application to air pollution, *Communications in Statistics - Theory and Methods* **39**: 426–443.

Lemonte, A. J. y Cordeiro, G. M. (2009). Birnbaum-saunders nonlinear regression models, *Comput. Stat. Data Anal.* **53(12)**: 4441–4452.

Lin, T., Lee, J. y Yen, S. (2007). Finite mixture modelling using the skew normal distribution, *Stat. Sin.* pp. 909–927.

Maehara, R., Bolfarine, H., Vilca, F. y Balakrishnan, N. (2021). A robust birnbaum–saunders regression model based on asymmetric heavy-tailed distributions, *Matrika* .

Mikhail, S., Nikulin y Tran, X. Q. (2013). Chi-squared goodness of fit test for generalized birnbaum-saunders models for right censored data and its reliability applications.

Paula, G. A., Leiva, V., Barros, M. y Liu, S. (2011). Robust statistical modeling using the birnbaum-saunders-t distribution applied to insurance, *Appl. Stoc. Models Bus. Ind.* **28**: 16–34.

Press, W., Flanenry, B., Teolosky, S. y W.T.Vetterling (1992). *The Art of Scientic Computing*, Cambridge University.

Rieck, J. R. (1989). *Statistical analysis for the Birnbaum-Saunders fatigue life distribution*, PhD thesis, Clemson University.

Rieck, J. R. y Nedelman, J. R. (1991). A log-linear model for the birnbaum-saunders distribution, *Technometrics* **33**: 51–60.

Schwarz, G. (1978). Estimating the dimension of a model, *The Annals of Statistics* **6**: 461–464.

Vilca, F., Azevedo, C. L. N. y Balakrishnan, N. (2017). Bayesian inference for sinh-normal/independent nonlinear regression models, *Journal of Applied Statistics* **44**(11): 2052–2074.

Vilca, F., Zeller, C. B. y Cordeiro, G. M. (2015). The sinh-normal/independent nonlinear regression model, *J. Appl. Stat.* **42**: 1659–1676.