
Distribuciones de Probabilidad
Asimétrica con Soporte Positivo Beta
Potencia-Normal



Trabajo de grado para optar al título de
Estadístico

María Alejandra Martínez Guerra

Departamento de Matemáticas y Estadística

Facultad de Ciencias Básicas

Universidad Córdoba

29 de agosto de 2019

Distribuciones de Probabilidad
Asimétrica con Soporte Positivo Beta
Potencia-Normal

Trabajo de grado para optar al título de Estadístico

María Alejandra Martínez Guerra

Dirigido por

Doctor Guillermo Martínez Flórez

Departamento de Matemáticas y Estadística

Facultad de Ciencias Básicas

Universidad Córdoba

29 de agosto de 2019

Copyright © María Alejandra Martínez Guerra

UNIVERSIDAD DE CÓRDOBA
FACULTAD DE CIENCIAS BÁSICAS
DEPARTAMENTO DE MATEMÁTICAS Y ESTADÍSTICA

Los jurados abajo firmantes certifican que han leído y recomiendan a la **Facultad de Ciencias Básicas** aprobar el trabajo de grado titulado: “**Distribuciones de Probabilidad Asimétrica con Soporte Positivo Beta Potencia-Normal**”, el cual es presentado por **María Alejandra Martínez Guerra**, como requisito parcial para optar al título de **Estadístico**.

Fecha: Agosto de 2019

Director: _____
Guillermo Martínez Florez

Jurado: _____
Roger Tovar Falón

Jurado: _____
Marvin Jiménez Narváez

*A mi querida madre Enadis Guerra y mi padre Jacob Martínez un ángel que
está en el cielo, gracias por tu amor, esfuerzo y trabajo, nunca te olvidaré, algún
día nos volveremos a encontrar
Los amo.*

Agradecimientos

En primer lugar, quiero agradecer a Dios por todo su apoyo en cada etapa de mi vida y la de mi familia. A mi amado padre Jacob Martínez por su esfuerzo y amor incondicional, gracias a él he logrado convertirme en lo que soy. A mi madre Enadis Guerra, hermanos y demás familiares, gracias por su amor y comprensión en todos estos años de esfuerzo.

Al Dr. Guillermo Martínez por su apoyo, conocimiento y aliento para que concluyera este proyecto, no solo en estos meses de trabajo si no también en los años de carrera.

Asimismo, le agradezco a los profesores Roger Tovar y Marvin Jiménez por tomarse el trabajo de revisar este documento y por todas sus sugerencias. Al profesor Mario Morales, a la profesora Juana Robles, al profesor Javier Ramírez y demás profesores del departamento de Matemáticas y Estadística que contribuyeron en toda la carrera con sus conocimientos.

Por último, gracias a todos mis compañeros que de buena voluntad me ayudaron en este camino, por todos los buenos momentos vividos en la facultad, en especial quisiera agradecer a Fidel Causil, Manuel Hernández y Maryoris Pérez.

Índice general

Agradecimientos	VI
1. Introducción	12
2. Preliminares	17
2.1. Distribución Normal Asimétrica	17
2.2. Distribución Alfa-Potencia (AP)	18
2.2.1. Caso Localización-Escala	19
2.2.2. Ecuaciones <i>Score</i>	20
2.3. Distribución Ln-Alfa-Potencia (LAP)	21
2.4. Distribuciones de Tipo Bimodal	21
2.4.1. Coeficiente de Variación (CV)	25
2.5. Variable Aleatoria Censurada	26
2.5.1. Censura a la Izquierda o Derecha	26
2.6. Modelos Censurados	26
2.6.1. Modelo Tobit	27
3. Distribución Log-Beta-Normal Asimétrica.	30
3.1. Extensión de localización-escala	41
3.2. Inferencia para el Modelo LBSN	42

4. Distribución Log-Beta-Normal Asimétrico Alfa-Potencia	47
4.1. Extensión de localización-escala	52
4.2. Inferencia para el modelo LPBSN	53
5. Distribución Log-Beta-Normal Asimétrica con Datos Censurados	60
5.1. Inferencia para el modelo CLBSN	65
5.2. Distribución Log-Beta-Normal Asimétrica Alfa-Potencia con Datos Censurados	69
5.3. Inferencia para el Modelo CLPBSN	74
6. Ilustraciones	81
6.1. Ilustracion 1.	81
6.2. Ilustracion 2.	84
Bibliografía	87

Índice de figuras

3.1. fdp del modelo $LBSN(\beta)$	32
3.2. Función de sobrevivencia del modelo $LBSN(\beta)$	35
4.1. Función de sobrevivencia del modelo $LPBSN(\beta, \alpha)$	51
5.1. fdp del modelo $CLBSN(\beta)$	62
5.2. fdp del modelo $CLPBSN(\beta, \alpha)$	71
6.1. (a) Histograma para las 48 observaciones del estudio. Las líneas representan las distribuciones ajustadas usando las estimativas de máxima verosimilitud: LN (línea discontinua con puntos), BSB (línea de puntos), $LBSN$ (línea discontinua) y $LPBSN$ (línea continua). Tambien se gráfica de la FDA para la parte (b) donde: Empírica (línea continua), $LBSN$ (Línea de puntos) y $LPBSN$ (línea discontinua).	84

- 6.2. Histograma de la distribución del conjunto de datos para el estudio de la vacuna contra el sarampión. Las líneas simbolizan las distribuciones ajustadas usando las estimativas de máxima verosimilitud: LN (línea con puntos), $L\beta SN$ (línea discontinua) y $LP\beta SN$ (línea continua). (b) Para la FDA: Empírica (línea continua), $L\beta SN$ (Línea de puntos) y $LP\beta SN$ (línea discontinua) 86

Índice de tablas

6.1. Resumen de estadísticas descriptivas.	81
6.2. Estimativas de máxima verosimilitud para los modelos ajustados censurados	82
6.3. Parámetros estimados para los modelos ajustados con datos censu- rados	85

Capítulo 1

Introducción

La distribución logarítmica normal (LN) es una distribución de probabilidad con un logaritmo distribuido normalmente que se obtiene como una transformación de la distribución normal ordinaria y se suele utilizar a menudo en situaciones en las que los valores presentan sesgo a la derecha como, por ejemplo, para determinar precios de acciones, precios de propiedades inmobiliarias, escalas salariales, tamaños de depósitos de aceite entre otros. En muchas de estas situaciones, la asimetría de la distribución y su curtosis están por encima o por debajo de lo esperado para el modelo LN, por lo que es necesario pensar en un modelo más flexible que logre tal desviación al modelar datos positivos.

Varios autores han introducido familias de distribuciones que permiten modelar variables con soporte positivo y diferentes grados de asimetría y curtosis. Una de las más reconocidas corresponde al modelo *ln-skew-normal* (LSN), la cual fue estudiada por Mateus-Figuera (2003-2004), quien estudia las propiedades de este modelo.

Es de notar que esta distribución es una extensión de la distribución LN para modelar la estructura asimétrica presente en los datos, no obstante presenta una

dificultad ya que su matriz de información es singular cuando el parámetro de asimetría es cercano a cero, además existe dificultad para la estimación de este parámetro, Olmos *et al.* (2016) presentan el modelo bimodal Birnbaum - Saunder (1950), el cual además de ajustar datos positivos, su matriz de información es no singular.

Distribuciones para ajustar datos positivos se limitan generalmente a la gama, weibull, exponencial, Birnbaum-Saunders(1950) y LN. Sin embargo, estas distribuciones solo ajustan datos con distribución unimodal con alta o baja asimetría, es decir que estos modelos no pueden ser aplicados cuando la distribución de los datos es bimodal o multimodal. Para datos positivos con comportamiento bimodal, son pocos los autores que han propuesto distribuciones que ajusten este tipo de información; destacándose Bolfarine *et al.* (2011) con el modelo *The log-bimodal-skew-normal model* (logBSN) que en español corresponde a la función de densidad de la distribución logaritmica bimodal normal asimétrica y Venegas *et al.* (2016) con *the Log-Alpha-Skew-Normal Model* (LASN), en español es la función de densidad de la distribución logaritmica alfa normal asimétrica. Sin embargo, estas dos distribuciones presentan el problema que sus matrices de información son singulares, cuando su parámetro de asimetría se acerca a la frontera de cero, dado que estas se originan como extensión de la distribución normal asimétrica de Azzalini(1985). Otra posible solución es la mezcla de distribuciones LN, pero como es bien conocido en la literatura, este modelo presenta problemas de identificabilidad en la estimación de sus parámetros.

Una solución a los inconvenientes presentados con las distribuciones arriba mencionadas es usar otras familias de distribuciones bases que no presenten problemas en su matriz de información ni problemas de identificabilidad en la estimación de sus parámetros. Como solución a la problemática antes descrita, este trabajo presen-

ta una nueva distribución con la cual se pueden ajustar distribuciones de soporte positivo de tipo bimodal o multimodal basadas en la distribución ln-alfa-potencia (LAP) Martínez-Flórez *et al.* (2014) y el modelo beta-normal asimétrico exponeciado (PBSN), estudiado recientemente por Martínez *et al.* (2018). Esta nueva distribución incluye nuevos parámetros que la hacen más flexible en términos de forma (bimodalidad y multimodalidad), asimetría y curtosis que el modelo LN, puesto que se logra modelar datos con distribución unimodal, bimodal y multimodal para datos positivos, con asimetría y curtosis por fuera del rango permitido por la distribución LN.

Cabe resaltar que las distribuciones para datos no-negativos se utilizan principalmente en el campo de la teoría de la confiabilidad (Birnbaum-Saunders (1950)), el cual es el análisis de datos de tiempo de falla (es la metodología estadística que se usa para determinar la confiabilidad de una población de dispositivos y predecir el porcentaje de fallas potenciales). Las distribuciones más utilizadas en la modelación de datos positivos son las distribuciones exponencial, weibull, gamma, LN y Birnbaum-Saunders, entre otras. En los últimos tiempos se ha estado utilizando el modelo *half*-normal para el análisis de sobrevivencia o extensiones de este, como por ejemplo la extensión realizada por Elal-Olivero *et al.* (2009).

A pesar de que los modelos LN y LSN Mateus-Figueras *et al.* (2003-2004) tienen amplia aplicabilidad en datos positivos con fuerte asimetría positiva, el modelo Birnbaum-Saunders se muestra como una alternativa más utilizada para ajustar datos de sobrevivencia y fatiga de materiales. Sin embargo, todos estos modelos solo pueden ajustar datos de carácter unimodal, siendo modelos no flexibles para observaciones de tipo bimodal o multimodal.

Así, es poca la literatura existente sobre distribuciones que logren ajustar datos positivos que presenten comportamiento bimodal y aún más complejo si la distri-

bución presenta un comportamiento multimodal. Por otro lado, Durrans (1992) introdujo la distribución de estadísticas de orden fraccionaria, la cual genera un buen modelo base para la creación de distribuciones que modelen variables asimétricas, con matriz de información no singular para valores cercanos a uno en el parámetro de asimetría.

Una extensión de este modelo al caso de datos bimodales fue estudiada por Bolfarine *et al.* (2013), denominada distribución bimodal AP; este modelo también tiene la característica de que su matriz de información es no singular. Otra extensión del modelo AP para el caso de datos positivos fue realizada por Martínez-Flórez *et al.* (2014) quienes introducen el modelo ln-potencia normal (LPN), cuya matriz de información resultó ser no singular. Recientemente Martínez *et al.* (2018) proponen una familia de distribuciones PBSN; la cual es otra extensión del modelo AP que logra ajustar datos multimodales además, su matriz de información también resulta ser no singular.

En vista de lo anterior, aprovechando la no singularidad del modelo AP y la flexibilidad para el caso de datos bimodales de los modelos bimodal AP y multimodal PBSN, resulta interesante estudiar el comportamiento de la distribución resultante que mezcla la estructura genérica AP y las distribuciones base LN, LPN, bimodal AP y PBSN. La justificación de esta mezcla es que dadas las características del modelo AP al realizar la extensión de esta, se espera que la matriz de información del modelo resultante no presente los problemas de singularidad de los modelos ya estudiados en la literatura y se genere una distribución flexible capaz de modelar datos positivos de tipo bimodal y multimodal, si esto no sucede se llevará a cabo una reparametrización adecuada.

El modelo estudiado en el presente trabajo contiene la distribución LN como un caso particular y se estudiarán sus principales propiedades estadísticas, así como

el proceso inferencial de estimación de sus parámetros, matriz de información y distribución asintótica del vector de estimadores de los parámetros.

Capítulo 2

Preliminares

En las últimas décadas se ha observado un notorio interés creciente en la literatura estadística por familias de distribuciones flexibles que ayuden al modelamiento de variables con diferentes grados de asimetría y curtosis. Distintos trabajos se han publicado en torno a este tema siendo preciso resaltar los trabajos de Birnbaum (1950), Lehmann (1953), Roberts (1966), O'Hagan *and* Leonard (1976), Azzalini (1985), Henze (1986) y Durrans (1992) entre otros.

2.1. Distribución Normal Asimétrica

Azzalini (1985) desarrolla una estructura general para la función de densidad de probabilidad (fdp) asimétrica, *skew-symmetric* la cual está dada por:

$$\varphi_{fF}(z; \lambda) = 2f(z)F(\lambda z) \quad z, \lambda \in \mathbb{R} \quad (2.1.1)$$

donde f es una fdp la cual es simétrica alrededor de cero, F es una función de distribución acumulada (FDA) absolutamente continua y también simétrica alrededor de cero. Para el caso donde f es la densidad de la distribución normal estándar y de igual forma F es la FDA de la normal estándar, el modelo resultante

es llamado de *skew-normal*, se denota por $Z \sim SN(\lambda)$ y su fdp viene dado por:

$$\varphi_{\phi\Phi}(z, \lambda) = 2\phi(z)\Phi(\lambda z) \quad z \in \mathbb{R}, \lambda > 0 \quad (2.1.2)$$

donde λ es un número real, llamado el parámetro de asimetría, puesto que permite modelar la parte asimétrica de la distribución. Este es el modelo estudiado por Azzalini (1985) quien introdujo formalmente la distribución y estudió sus propiedades, cabe señalar como ya se había dicho que esta distribución es una extensión de la distribución normal para modelar la estructura asimétrica presente en los datos, no obstante presenta un problema ya que su matriz de información es singular cuando el parámetro de asimetría es cercano o igual a cero, también se han observado problemas para la estimación de este parámetro. Además de Azzalini otros aportes importantes a esta teoría de la distribución *skew-normal* fueron hechos por Henze (1986), Arnold *et al.* (1993), Chiogna (1997) y Pewsey (2000), entre otros.

Una extensión para el caso de datos positivos del modelo SN fue realizada por Mateus-Figueras (2003-2004) quien estudia el modelo ln-normal asimétrico (LNS), de localización (μ) y escala (σ), definido por la fdp así,

$$\varphi_Y(y; \xi, \eta, \lambda) = \frac{2}{\sigma y} \phi\left\{\frac{\ln(y) - \mu}{\sigma}\right\} \Phi\left\{\lambda \frac{(\ln(y) - \mu)}{\sigma}\right\} \quad y \in \mathbb{R}^+ \quad (2.1.3)$$

2.2. Distribución Alfa-Potencia (AP)

Lehmann (1953) propone la familia de distribuciones con FDA dada por,

$$F(z; \alpha) = \{F(z)\}^\alpha \quad z \in \mathbb{R} \quad (2.2.1)$$

donde F es alguna FDA y α es un número racional. En términos generales esta distribución es generada a partir de la distribución del máximo de la muestra. En

la literatura este modelo es conocido como modelo de alternativas de Lehmann y es ampliamente discutido por Gupta (1959) y Arnold (2004).

Continuando el trabajo hecho por Lehmann, Durrans (1992) en un contexto hidrológico extiende el anterior modelo para todo $\alpha \in \mathbb{R}^+$, refiriéndose a este resultado como la distribución estadística de orden fraccionaria, que en la literatura es más conocida como distribución AP, con fdp dada por:

$$\varphi(z; \alpha) = \alpha f(z) \{F(z)\}^{(\alpha-1)} \quad \text{con } z \in \mathbb{R}, \alpha \in \mathbb{R}^+ \quad (2.2.2)$$

donde F es una FDA absolutamente continua con fdp $f = dF$. Se denota con $Z \sim AP(\alpha)$. Este nuevo modelo dio origen a la familia de distribuciones AP la cual es una alternativa como entre las familias de distribuciones que modelan altos grados de asimetría y curtosis.

La FDA de Z en el modelo (2.2.2) está dada por:

$$F_F(z; \alpha) = \{F(z)\}^\alpha, \quad \text{con } z \in \mathbb{R}. \quad (2.2.3)$$

Propiedades del Modelo AP

- (i) Si $Z \sim AP(\alpha)$ entonces, $E(Z^n) = \alpha \int_0^1 [F^{-1}(u)]^n u^{(\alpha-1)} \cdot du$ donde F^{-1} es la inversa de F .
- (ii) Para el caso en que $f = \phi$ entonces, la distribución resultante es llamada distribución potencia-normal (PN) la cual tiene muchas aplicaciones en diversas áreas del conocimiento, como por ejemplo, las mediciones no negativas (las medidas del índice de masa corporal (IMC)).

2.2.1. Caso Localización-Escala

En la práctica, el interés principal se centra en ajustar el modelo de localización-escala de la distribución de estadística de orden fraccionaria AP con densidad

(2.2.1). Es fácil, ver que si $Z \sim AP(\alpha)$, entonces la extensión localización-escala de Z , es decir, $X = \xi + \omega Z$, donde $\xi \in \mathbb{R}$ y $\omega \in \mathbb{R}^+$, es la función de densidad dada por:

$$\varphi(x; \xi, \omega, \alpha) = \alpha f\left(\frac{x - \xi}{\omega}\right) \left\{ F\left(\frac{x - \xi}{\omega}\right) \right\}^{(\alpha-1)} \quad x \in \mathbb{R}, \alpha \in \mathbb{R}^+ \quad (2.2.4)$$

la cual se denotará $X \sim AP(x; \xi, \omega, \alpha)$. Como comúnmente ocurre esta distribución involucra parámetros desconocidos, para la estimación de estos se utiliza el método de máxima verosimilitud, la función de log-verosimilitud del modelo AP, esta dada por:

$$l(\theta; \mathbf{x}) = n \{ \ln(\alpha) - \ln(\omega) \} + \sum_{i=1}^n \ln \{ f(z_i) \} + (\alpha - 1) \sum_{i=1}^n \ln \{ F(z_i) \} \quad (2.2.5)$$

donde $z_i = \frac{x_i - \xi}{\omega}$.

2.2.2. Ecuaciones *Score*

Asumiendo que f' existe, las primeras derivadas parciales de la función de log-verosimilitud con respecto a cada uno de los parámetros son:

$$\begin{aligned} \frac{\partial l(\theta; \mathbf{x})}{\partial \xi} &= \frac{-1}{n} \left\{ \sum_{i=1}^n \frac{f'(z_i)}{f(z_i)} + (\alpha - 1) \sum_{i=1}^n \frac{f(z_i)}{F(z_i)} \right\}, \\ \frac{\partial l(\theta; \mathbf{x})}{\partial \omega} &= \frac{-1}{n} \left\{ n + \sum_{i=1}^n z_i \frac{f'(z_i)}{f(z_i)} + (\alpha - 1) \sum_{i=1}^n z_i \frac{f(z_i)}{F(z_i)} \right\}, \\ \frac{\partial l(\theta; \mathbf{x})}{\partial \alpha} &= \frac{n}{\alpha} + \sum_{i=1}^n \ln \{ F(z_i) \}. \end{aligned}$$

Las ecuaciones *score* son obtenidas igualando las derivadas parciales a cero. Procediendo como Chiogna (1997) y Pewsey(2006), siendo $w_i = \frac{f(z_i)}{F(z_i)}$, $v_i = \frac{f(z_i)}{f'(z_i)}$ y $u_i = \ln \{ F(z_i) \}$, se deduce inmediatamente que la solución a las ecuaciones

score satisfacen que $\bar{v} = (1 - \alpha)\bar{\omega}$, $1 + (\overline{zv}) = (1 - \alpha)(\overline{z\omega})$ y $\alpha = \frac{-1}{u}$ donde $\bar{v} = \sum_{i=1}^n \frac{v_i}{n}$, $(\overline{zv}) = \sum_{i=1}^n \frac{z_i v_i}{n}$, etc. Generalmente estos sistemas de ecuaciones son resueltos numéricamente, usando métodos iterativos tipo Newton-Raphson.

De igual manera siguiendo esta metodología se encontrará la matriz de información, tanto la observada como la esperada y podemos constatar que esta es no singular para valores del parámetro de asimetría cercanos a uno.

2.3. Distribución Ln-Alfa-Potencia (LAP)

Una extensión del modelo AP para el caso de datos positivos fue estudiado por Martínez-Flórez *et al.* (2014), denominado distribución LAP cuya fdp viene dada por:

$$\varphi_Y(y; \alpha) = \frac{\alpha}{y} f\{\ln(y)\} \{F(\ln(y))\}^{(\alpha-1)}, \quad y \in \mathbb{R}^+. \quad (2.3.1)$$

Para el caso en que $f = \phi$ entonces, la distribución resultante es llamada distribución ln-potencia-normal la cual tiene muchas aplicaciones en diversas áreas del conocimiento, cuya fdp viene dada por:

$$\varphi_Y(y; \alpha) = \frac{\alpha}{y} \phi(\ln(y)) \{\Phi(\ln(y))\}^{(\alpha-1)} \quad \text{con } y \in \mathbb{R}^+. \quad (2.3.2)$$

Este modelo (2.3.2) al igual que el Modelo PN tiene la característica que su matriz de información es no singular.

2.4. Distribuciones de Tipo Bimodal

Además del interés común del estudio de distribuciones que ayuden a modelar estructuras asimétricas, muchos de los autores que han estudiado estas familias

de distribuciones han propuesto extensiones bimodales, las cuales son muy útiles a la hora de estudiar ciertos fenómenos en el área de la cosmología y la genética. Una extensión de la distribución normal es el modelo bimodal-normal, que nos proporciona una gran herramienta para el estudio de variables con dos máximos en el soporte, su función de densidad está dada por:

$$f(u) = u^2 \phi(u) \quad (2.4.1)$$

Para el Modelo PN antes mencionado, diversas extensiones bimodales han sido estudiadas, Kim (2005) introduce el modelo

$$f(u, \theta) = c_\theta \phi(u) \Phi(\theta|u) \quad (2.4.2)$$

donde θ es un número real y c_θ una constante de normalización. Para $\theta > 0$ Kim demuestra que el modelo (2.4.2) es bimodal. Gómez *et al.* (2009), define el tipo de distribución *skew-flexible-normal* cuya fdp es dada por:

$$f(u, \theta) = c_\delta \phi(|u| + \delta) \Phi(\lambda u) \quad \text{con } \lambda \in \mathbb{R} \quad (2.4.3)$$

donde δ es un número real y c_δ es una constante de normalización. Gómez *et al.* (2009) demuestran que para $\delta < 0$ el modelo (2.3.1) es bimodal. Arnold *et al.* (2009) estudian el modelo bimodal simétrico al cual denominan La *skew-normal* de dos partes con fdp:

$$f(u, \theta) = 2c_\lambda \phi(u) \Phi(\lambda|u|) \Phi(\beta z) \quad (2.4.4)$$

donde β y λ son números reales y c_λ es una constante de normalización. Este es un modelo bimodal asimétrico para ciertos valores de β y λ . Arnold demuestra que la matriz de información del modelo es singular para $\beta = \lambda = 0$, es decir para el caso de la distribución normal, tal y como acontece con el modelo *skew-normal* de Azzalini.

Por otro lado Elal-Olivero (2009) introduce la versión bimodal de la *skew*-normal y realiza un análisis bayesiano. La función de densidad para este modelo está dada por:

$$f(y, \lambda, \alpha) = 2 \left(\frac{1 + \lambda y^2}{1 + \lambda} \right) \phi(y) \Phi(\alpha y) \quad \text{con } y \in \mathbb{R}, \lambda \in \mathbb{R}, \alpha \in \mathbb{R}^+. \quad (2.4.5)$$

Una extensión bimodal para el caso de datos positivos fue estudiada por Bolfarine *et al.* (2011), denominada de distribución bimodal LN asimétrica, cuya fdp está dada por:

$$f(x|\theta) = \frac{2}{\sigma x} \left(\frac{1 + \alpha y^2}{1 + \lambda} \right) \phi(y) \Phi(\lambda y) \quad \text{con } x > 0, \mu \in \mathbb{R}, \alpha > 0 \quad (2.4.6)$$

donde $y = \frac{\ln(x) - \mu}{\sigma}$. Esta distribución al igual que el modelo (2.4.5) presenta matriz de información singular para el caso donde sus dos parámetros toman el valor cero. Elal-Oliveros (2010) presenta el modelo bimodal asimétrico

$$f(x|\theta) = \left[\frac{(1 - \alpha y)^2 + 1}{(2 + \alpha^2)\sigma} \right] \phi(y) \quad \text{con } x > 0, \alpha \in \mathbb{R} \quad (2.4.7)$$

donde $y = \frac{x - \mu}{\sigma}$, al cual denominó como AP asimétrico, este modelo al igual que el modelo (2.4.5) tiene matriz de información singular.

La extensión para datos positivos del modelo dado en (2.4.7) para el caso de datos positivos fue estudiada por Venegas *et al.* (2016) y su fdp está dada por:

$$f(x|\theta) = \left[\frac{(1 - \alpha y)^2 + 1}{(2 + \alpha^2)\sigma x} \right] \phi(y) \quad \text{con } x > 0, \alpha \in \mathbb{R} \quad (2.4.8)$$

donde $y = \frac{\ln(x) - \mu}{\sigma}$, al cual denominó distribución ln-alfa-normal asimétrico, este modelo al igual que el modelo (2.4.7) tiene matriz de información singular.

Bolfarine *et al.* (2013) introduce dos familias de distribuciones las cuales están basadas en la distribución AP introducida por Durrans (1992). El primer modelo corresponde a una extensión directa del modelo de Durrans (1992), con la cual

se logra modelar datos con distribuciones simétricas y el segundo modelo incluye un parámetro que hace más flexible el primer modelo pues se logra obtener distribuciones bimodales tanto simétricas como asimétricas. Bolfarine *et al.* (2013) definen el modelo bimodal simétrico exponenciado mediante la fdp,

$$\varphi_X(x; \xi, \eta, \alpha) = \frac{\alpha}{\eta} \frac{2^{\alpha-1}}{2^\alpha - 1} f\left(\frac{x - \xi}{\eta}\right) \left\{ F\left(\left|\frac{x - \xi}{\eta}\right|\right) \right\}^{\alpha-1}, \quad x, y, \xi \in \mathbb{R}, \eta \in \mathbb{R}^+ \quad (2.4.9)$$

donde c_α es una constante de normalización. El caso $f = \phi$ queda dado por:

$$\varphi_X(x; \xi, \eta, \alpha) = \frac{\alpha}{\eta} \frac{2^{\alpha-1}}{2^\alpha - 1} \phi\left(\frac{x - \xi}{\eta}\right) \left\{ \Phi\left(\left|\frac{x - \xi}{\eta}\right|\right) \right\}^{\alpha-1}, \quad x, y, \xi \in \mathbb{R}, \eta \in \mathbb{R}^+ \quad (2.4.10)$$

El caso asimétrico del modelo (2.4.10) viene dado por la fdp:

$$\varphi_X(z; \xi, \eta, \beta, \alpha) = 2\alpha c_\alpha \phi\left(\frac{x - \xi}{\eta}\right) \left\{ \Phi\left(\left|\frac{x - \xi}{\eta}\right|\right) \right\}^{\alpha-1} \Phi\left(\beta \frac{x - \xi}{\eta}\right), \quad x, y, \xi \in \mathbb{R}, \eta \in \mathbb{R}^+ \quad (2.4.11)$$

Sobhan *et al.* (2016) proponen el modelo multimodal denominado la distribución alfa-Beta Skew-normal, cuya fdp viene dada por:

$$f(x) = \frac{(1 - \alpha x - \beta x^3)^2 + 1}{\alpha^2 + 15\beta^2 + 6\alpha\beta + 2} \phi(x) \quad (2.4.12)$$

donde $y, \alpha, \beta \in \mathbb{R}$. Estos autores demuestran que este modelo puede alcanzar hasta cuatro modas para distintos valores de α y β .

Recientemente, Martínez *et al.* (2018), estudian las distribuciones BSN(β) y PBSN(β, α), las cuales resultaron con matriz de información no singular y con una gran flexibilidad para ajustar un conjunto de observaciones con distribución, unimodal, bimodal o trimodal.

Se dice que la variable aleatoria Z tiene una distribución $BSN(\beta)$, si el modelo viene dado por :

$$f(x) = \frac{(1 - \beta x^3)^2 + 1}{15\beta^2 + 2} \phi(x), \quad (2.4.13)$$

donde $y, \alpha, \beta \in \mathbb{R}$. Estos autores demuestran que este modelo puede alcanzar hasta tres modas para distintos valores de β . Mientras que el caso exponenciado del modelo (2.4.13) viene representado por la fdp

$$f(z) = \alpha \frac{(1 - \beta z^3)^2 + 1}{2 + 15\beta^2} \phi(z) \left[\Phi(z) + \frac{4\beta - 15\beta^2 z + 2\beta z^2 - 5\beta^2 z^3 - \beta^2 z^5}{2 + 15\beta^2} \phi(z) \right]^{\alpha-1} \quad (2.4.14)$$

donde $z \in \mathbb{R}$, $\alpha \in \mathbb{R}^+$ y $\beta \in \mathbb{R}$. Entonces se dice que Z es una variable aleatoria con distribución PBSN. Se denotará por $Z \sim PBSN(\beta, \alpha)$.

Estos autores encontraron que la FDA acumulada del modelo dado en (2.4.14) se deja escribir como

$$\varphi_{PBSN}(z) = \left[\Phi(z) + \frac{4\beta - 15\beta^2 z + 2\beta z^2 - 5\beta^2 z^3 - \beta^2 z^5}{2 + 15\beta^2} \phi(z) \right]^\alpha \quad \text{con } z \in \mathbb{R} \quad (2.4.15)$$

2.4.1. Coeficiente de Variación (CV)

También denominado como coeficiente de variación de Spearman, es una medida estadística que nos informa acerca de la dispersión relativa de un conjunto de datos, se denota de la siguiente forma:

$$CV = \frac{\sqrt{Var(X)}}{E(X)}.$$

2.5. Variable Aleatoria Censurada

Los valores censurados son aquellos reportados como menos que algún valor (por ejemplo, menor que 5 ppb), mayores que algún valor (por ejemplo, mayor que 100 días), o como un intervalo (por ejemplo, un valor entre 67 y 75 grados). Los valores truncados son aquellos que no informan si el valor excede algún límite.

Ahora considere un laboratorio clínico que reporte la concentración de atrazina en una muestra de agua subterránea. Si reportan 0.02 ppb, se observa ese valor. Si reportan que es menor 0.05 ppb, el valor es censurado. Si solo informan la concentración si excede de 0.1 ppb, los valores se truncan. La diferencia práctica entre los datos censurados y truncados es que se conoce el número de valores censurados, pero no el número de valores truncados. Los datos ambientales observados y censurados son mucho más comunes que los datos truncados.

2.5.1. Censura a la Izquierda o Derecha

Un valor censurado a la izquierda es uno que solo se sabe que es menor que algún valor, por ejemplo, menor que 5 ppm. Un valor censurado a la derecha es aquel que solo se sabe que es más que algún valor.

2.6. Modelos Censurados

En términos generales una variable aleatoria Y que tiene una parte de sus probabilidades en puntos discretos y el resto extendido en algunos intervalos, tiene una distribución mixta. Sea $F(y)$ la FDA que representa una distribución mixta. Entonces se puede escribir

$$F(y) = c_1 f_1(y) + c_2 f_2(y) \quad (2.6.1)$$

donde $F_1(y)$ es una FDA escalonada, $F_2(y)$ una FDA continua, c_1 la probabilidad acumulada de todos los puntos discretos y $c_2 = 1 - c_1$ la probabilidad acumulada de las porciones continuas. Para la distribución dada en (2.6.1) se tiene que dado $g(Y)$, una función de Y , entonces:

$$E[g(Y)] = c_1 E[g(X_1)] + c_2 E[g(X_2)] \quad (2.6.2)$$

donde X_1 es una variable aleatoria discreta que tiene la FDA $F_1(y)$ y X_2 es una variable aleatoria que tiene la FDA $F_2(y)$.

En algunos casos además de la censura también las observaciones presentan altos (o bajos) grados de asimetría y/o curtosis. Por tanto, el análisis con datos asimétricos es necesario para afrontar este tipo de situaciones, así es usual utilizar la distribución SN de Azzalini(1985) ó Normal AP de Durrans(1992).

Cuando los datos son censurados, la distribución que siguen es una mixtura entre una distribución discreta y una continua.

2.6.1. Modelo Tobit

En los modelos de regresión censurados la variable dependiente subyacente es aproximadamente continua, pero está censurada inferior (censura a la izquierda) o superiormente (censura a la derecha), debido a la forma en que recopilamos los datos o a limitaciones institucionales.

El modelo de regresión censurado Tobit o Tobit tipo I se expresa como sigue:

$$Y_i = \begin{cases} c & \text{si } Y_i^* \leq c, \\ Y_i^* & \text{si } Y_i^* > c \end{cases}$$

Se ha formulado el modelo Tobit en términos de una variable latente Y^* que satisface los supuestos del modelo de regresión clásico, distribución normal, homocedástica y con media condicional lineal.

En algunos casos la variable en estudio puede presentar cierto grado de asimetría que no puede ser modelada por la distribución normal o la distribución *t-Student*. En estos casos es válido suponer que las observaciones obtenidas de los individuos puede ser modelada por la distribución *skew-normal* o APN.

Capítulo 3

Distribución Log-Beta-Normal Asimétrica.

En este capítulo se introduce el modelo log-beta-normal asimétrica (LBSN), Se define la fdp del modelo, se halla la FDA y se prueba la no existencia de la función generadora de momentos. También se estudian sus principales propiedades para el caso estándar del modelo, se dan a conocer las expresiones del momento de orden superior para los casos par e impar, así como también se determinan las medidas: esperanza, varianza, asimetría, curtosis y CV.

Posteriormente se presenta el modelo de Localización-Escala, se argumenta la estimación de sus parámetros y se determinan las matrices de información observada y esperada para luego determinar la no singularidad de esta.

Definición 3.0.1. Se dice que la variable aleatoria X tiene una distribución LBSN, si su fdp está dada por

$$f(x) = \frac{(1 - \beta \ln^3 x)^2 + 1}{2 + 15\beta^2} \frac{\phi(\ln x)}{x}, \quad \text{con } x > 0 \quad (3.0.1)$$

donde $\beta \in \mathbb{R}$. Denotaremos esta variable como $X \sim \text{LBSN}(\beta)$.

Proposición 3.0.1 La función de densidad del modelo $\text{LBSN}(\beta)$ tiene a lo más 3 modas.

Demostración. Diferenciando (3.0.1) con respecto a x se tiene lo siguiente:

$$f'(x) = \frac{\phi(\ln x)}{(2 + 15\beta^2)x^2} \left[5\beta^2 \ln^5 x + \beta^2 \ln^4 x + 2\beta \ln^3 x - 8\beta \ln^2 x - 2 \ln x + 2 \right] \quad (3.0.2)$$

□

Por lo tanto $f'(x)$ tiene a lo mas 6 raices, así $f(x)$ tiene a lo mucho 4 modas. Por el teorema fundamental del algebra un polinomio de grado n tiene n raices como es un polinomio de grado impar por lo menos tiene una raiz real por lo tanto la distribución $\text{LBSN}(\beta)$ tiene a lo mas 3 modas, como se puede ver a continuación en la figura (3.1).

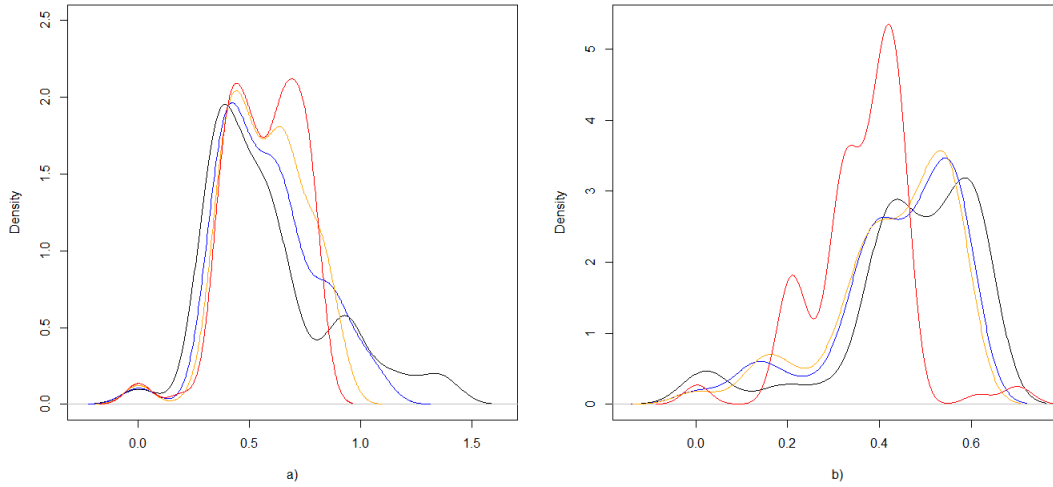


Figura 3.1: fdp del modelo LBSN (β).

Para diferentes valores de β se tienen; las líneas de color: negra, azul, naranja y roja que corresponden respectivamente, (a) $\beta = 0.2, \beta = 0.15, \beta = 0.12$ y $\beta = 0.1$; (b) $\beta = -0.05, \beta = -0.09, \beta = -0.1$ y $\beta = -0.2$.

Proposición 3.0.2 (FDA del Modelo LBSN Estándar).

Si $X \sim LBSN(\beta)$ entonces la FDA de $\Phi_{LBSN}(x)$ está dada por:

$$\Phi_{LBSN}(x) = \Phi(\ln x) + \frac{\beta(2 \ln^2 x + 4 - \beta \ln^5 x - 5\beta \ln^3 x - 15\beta \ln x)}{2 + 15\beta^2} \phi(\ln x) \quad (3.0.3)$$

Demostración.

$$\begin{aligned}
\Phi_{LBSN}(x) &= \int_{-\infty}^x \frac{(1 - \beta \ln^3 t)^2 + 1}{2 + 15\beta^2} \frac{\phi(\ln t)}{t} dt \quad \text{pero como } x > 0 \text{ se tiene} \\
&= \int_0^x \frac{2 - 2\beta \ln^3 t + \beta^2 \ln^6 t}{2 + 15\beta^2} \frac{\phi(\ln t)}{t} dt \\
&= \frac{1}{2 + 15\beta^2} \int_0^x \left[2\phi(\ln t) - 2\beta \ln^3 t \phi(\ln t) + \beta^2 \ln^6 t \phi(\ln t) \right] dt \\
&= \frac{1}{2 + 15\beta^2} \left[2 \int_0^x \phi(\ln t) dt - 2\beta \int_0^x \ln^3 t \phi(\ln t) dt + \beta^2 \int_0^x \ln^6 t \phi(\ln t) dt \right] \\
&= \frac{1}{2 + 15\beta^2} \left[(2 + 15\beta^2)\Phi(\ln x) + \beta(2 \ln^2 x + 4 - \beta \ln^5 x - 5\beta \ln^3 x - 15\beta \ln x)\phi(\ln x) \right] \\
&= \Phi(\ln x) + \frac{\beta(2 \ln^2 x + 4 - \beta \ln^5 x - 5\beta \ln^3 x - 15\beta \ln x)}{2 + 15\beta^2} \phi(\ln x) \\
&= \Phi(\ln x) + \frac{4\beta - 15\beta^2 \ln x + 2\beta \ln^2 x - 5\beta^2 \ln^3 x - \beta^2 \ln^5 x}{2 + 15\beta^2} \phi(\ln x)
\end{aligned}$$

□

Proposición 3.0.3 (Función de Supervivencia y Hazard del modelo *LBSN* Estándar).

Si $S(t)$ y $h(t)$ denotan, respectivamente, la función de supervivencia y la función hazard de la variable aleatoria $T \sim LBSN(\beta)$ en el intervalo $(0, \infty)$, entonces:

$$S(t) = 1 - \left[\Phi(\ln t) + \frac{4\beta - 15\beta^2 \ln t + 2\beta \ln^2 t - 5\beta^2 \ln^3 t - \beta^2 \ln^5 t}{2 + 15\beta^2} \phi(\ln t) \right], \quad (3.0.4)$$

$$h(t) = \frac{(1 - \beta \ln^3 t)^2 + 1}{t(2 + 15\beta^2) \left[1 - \Phi_{LBSN}(t) \right]} \phi(\ln t). \quad (3.0.5)$$

Demostración. Por definición de la función de sobrevivencia se tiene

$$\begin{aligned} S(t) = P(T > t) &= \int_t^{\infty} f(x) \cdot dx = 1 - F(t) = 1 - \Phi_{LBSN}(t) \\ &= 1 - \left[\Phi(\ln t) + \frac{4\beta - 15\beta^2 \ln t + 2\beta \ln^2 t - 5\beta^2 \ln^3 t - \beta^2 \ln^5 t}{2 + 15\beta^2} \phi(\ln t) \right] \end{aligned}$$

Ahora la función Hazard está determinada por:

$$\begin{aligned} h(t) &= \frac{f(t)}{S(t)} = \frac{\frac{(1-\beta \ln^3 t)^2 + 1}{2+15\beta^2} \frac{\phi(\ln t)}{t}}{1 - \Phi_{LBSN}(t)} \\ &= \frac{\frac{(1-\beta \ln^3 t)^2 + 1}{(2+15\beta^2)t} \phi(\ln t)}{\frac{1-\Phi_{LBSN}}{1}} \\ &= \frac{(1 - \beta \ln^3 t)^2 + 1}{t(2 + 15\beta^2)(1 - \Phi_{LBSN}(t))} \phi(\ln t) \end{aligned}$$

□

En la siguiente Figura (3.2) se observa el comportamiento de la función de sobrevivencia para distintos valores de β . Se puede apreciar que a medida que β se hace grande, la probabilidad de sobrevivencia decrece a una tasa mayor. Además, la curva se hace 0 en las proximidades de $t = 4.2$ para valores de $\beta < 0$.

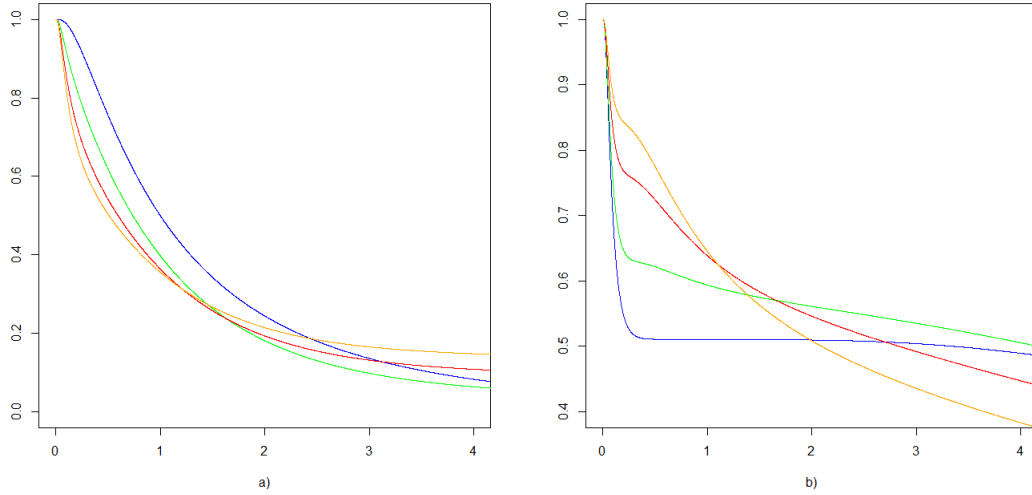


Figura 3.2: Función de sobrevivencia del modelo $LBSN(\beta)$.

Para las anteriores gráficas, con diferentes valores de β se tienen; las líneas de color: azul, verde, roja y naranja que corresponde respectivamente, (a) $\beta = 0$, $\beta = 0.15$, $\beta = 0.25$, $\beta = 0.32$ y (b) $\beta = -10$, $\beta = -1$, $\beta = -0.5$, $\beta = 0.35$.

Proposición 3.0.4 (Esperanza, varianza, asimetría y curtosis del modelo $LBSN$ estándar).

Sea $X \sim LBSN(\beta)$, γ_1 y γ_2 la asimetría y la curtosis del modelo estándar, respectivamente, entonces:

(i)

$$E(X) = \frac{2(38\beta^2 - 4\beta + 1)}{2 + 15\beta^2} e^{\frac{1}{2}} \quad (3.0.6)$$

(ii)

$$V(X) = \frac{((499\beta^2 - 28\beta + 2)(2 + 15\beta^2)e^2 - (76\beta^2 - 8\beta + 2)^2)e}{(2 + 15\beta^2)^2} \quad (3.0.7)$$

(iii)

$$\gamma_1 = \frac{2(v_1)He^{\frac{9}{2}} - 6v_2Hv_3e^{\frac{3}{2}} + 16v_2^3e^{\frac{1}{8}}}{\left(H(v_3e^2 - 2v_2e^{\frac{1}{2}})\right)^{\frac{3}{2}}} \quad (3.0.8)$$

(iv)

$$\gamma_2 = \frac{H^{\frac{3}{2}}(v_4H^3e^8 - 16v_2v_1H^2e^5 + 24v_2^2v_3He^3 - 48v_2^4e^2)}{H^4(v_3e^2 - 2v_5e^{\frac{1}{2}})^{\frac{3}{2}}} \quad (3.0.9)$$

(v)

$$CV = \frac{(v_3H - 2v_2)^{\frac{1}{2}}}{|2v_2|} \quad (3.0.10)$$

con $H = 2 + 15\beta^2$ y

$$v_1 = 1182\beta^2 - 36\beta,$$

$$v_2 = 38\beta^2 - 4\beta + 1,$$

$$v_3 = 499\beta^2 - 28\beta + 2.$$

$$v_4 = -8641\beta^2 - 152\beta + 2 \text{ y}$$

$$v_5 = 36\beta^2 - 4\beta + 1.$$

(vi) $M_X(t)$ no existe

Demostración. En general, debido a que el modelo $LBSN(\beta)$ es una transformación de la v.a normal $Y \sim BSN(\beta)$ y $X \sim LBSN(\beta)$ con $x > 0$, entonces:

Si $Y = \ln X \Rightarrow X = e^Y \Rightarrow X^n = e^{nY}$ así apoyados en esta relación y por la función generadora de momentos de el modelo $BSN(\beta)$ estudiado por Martínez *et al.* (2018), se tiene lo siguiente:

$$M_Y(t) = \left[1 + \beta t \left(\frac{\beta t^5 + 15\beta t^3 - 2t^2 + 45\beta t - 6}{2 + 15\beta^2} \right) \right] e^{\frac{t^2}{2}}$$

Por lo tanto se tiene que:

$$E(X^n) = E(e^{nY}) = M_Y(n), \text{ entonces}$$

$$M_Y(n) = \left[1 + \beta n \left(\frac{\beta n^5 + 15\beta n^3 - 2t^2 + 45\beta n - 6}{2 + 15\beta^2} \right) \right] e^{\frac{n^2}{2}}$$

Ahora para n impar con $n = 2k - 1$, se tiene:

$$E(X^{2k-1}) = E(e^{(2k-1)Y}) = M_Y(2k-1)$$

$$= \left\{ 1 + \beta(2k-1) \times \left[\frac{\beta(2k-1)^5 + 15\beta(2k-1)^3 - 2(2k-1)^2 + 45\beta(2k-1) - 6}{2 + 15\beta^2} \right] \right\} e^{\frac{(2k-1)^2}{2}}$$

(i) para $k = 1$

$$\begin{aligned} \mu_1 = E(X) &= \left[1 + \beta(1) \left(\frac{\beta(1)^5 + 15\beta(1)^3 - 2(1)^2 + 45\beta(1) - 6}{2 + 15\beta^2} \right) \right] e^{\frac{1}{2}} \\ &= \left[1 + \beta \left(\frac{\beta + 15\beta - 2 + 45\beta - 6}{2 + 15\beta^2} \right) \right] e^{\frac{1}{2}} \\ &= \left[1 + \beta \left(\frac{61\beta - 8}{2 + 15\beta^2} \right) \right] e^{\frac{1}{2}} \\ &= \frac{76\beta^2 - 8\beta + 2}{2 + 15\beta^2} e^{\frac{1}{2}} \\ &= \frac{2(38\beta^2 - 4\beta + 1)}{2 + 15\beta^2} e^{\frac{1}{2}} \\ &= \frac{2v_2}{H} e^{\frac{1}{2}} \\ &\text{con } v_2 = 38\beta^2 - 4\beta + 1 \text{ y } H = 2 + 15\beta^2. \end{aligned}$$

Para $k = 2$ de la misma forma se obtiene

$$\begin{aligned} \mu_3 = E(X^3) &= \frac{2364\beta^2 - 72\beta + 2}{2 + 15\beta^2} e^{\frac{9}{2}} \\ &= \frac{2(1182\beta^2 - 36\beta + 1)}{2 + 15\beta^2} e^{\frac{9}{2}} \\ &= \frac{2v_1}{H} e^{\frac{9}{2}} \end{aligned} \tag{3.0.11}$$

con $H = 2 + 15\beta^2$ y $v_1 = 1182\beta^2 - 36\beta + 1$. Ahora para n par, se tiene $n = 2k$, entonces,

$$\begin{aligned} E(X^{2K}) &= E(e^{(2k)Y}) = M_Y(2k) \\ &= \left[1 + \beta(2k) \left(\frac{\beta(2k)^5 + 15\beta(2k)^3 - 2(2k)^2 + 45\beta(2k) - 6}{2 + 15\beta^2} \right) \right] e^{\frac{(2k)^2}{2}} \\ &= \left[1 + 2\beta k \left(\frac{32\beta k^5 + 120\beta k^3 - 8k^2 + 90\beta k - 6}{2 + 15\beta^2} \right) \right] e^{2k^2} \end{aligned}$$

Para $k = 1$

$$\begin{aligned} \mu_2 &= E(X^2) = M_Y(2) \\ &= \left[1 + 2\beta \left(\frac{32\beta + 120\beta - 8 + 90\beta - 6}{2 + 15\beta^2} \right) \right] e^2 \\ &= \frac{499\beta^2 - 28\beta + 2}{2 + 15\beta^2} e^2 \end{aligned} \quad (3.0.12)$$

Para $k = 2$ de la misma forma se obtiene

$$\begin{aligned} \mu_4 &= E(X^4) = \left[1 + 4\beta \left(\frac{1024\beta + 960\beta - 32 + 180\beta - 6}{2 + 15\beta^2} \right) \right] e^8 \\ &= \left[1 + 4\beta \left(\frac{2164\beta - 38}{2 + 15\beta^2} \right) \right] e^8 \\ &= \left[\frac{2 + 15\beta^2 - 4\beta(2164\beta - 38)}{2 + 15\beta^2} \right] e^8 \\ &= \left[\frac{2 + 15\beta^2 - 8656\beta^2 + 152\beta}{2 + 15\beta^2} \right] e^8 \\ &= \left[\frac{-8641\beta^2 + 152\beta + 2}{2 + 15\beta^2} \right] e^8 \end{aligned} \quad (3.0.13)$$

(ii) Por definición de varianza se obtiene:

$$\begin{aligned}
V(X) &= E(X^2) - E^2(X) \\
&= \frac{499\beta^2 - 28\beta + 2}{2 + 15\beta^2} e^2 - \left(\frac{76\beta^2 - 8\beta + 2}{2 + 15\beta^2} e^{\frac{1}{2}} \right)^2 \\
&= \frac{\left[(499\beta^2 - 28\beta + 2)(2 + 15\beta^2)e^2 - (76\beta^2 - 8\beta + 2)^2 \right] e}{(2 + 15\beta^2)^2} \\
&= \frac{(v_3 H e^2 - 2v_2^2) e}{H^2} \\
&\text{con } H = 2 + 15\beta^2 \text{ y } v_2 = 38\beta^2 - 4\beta + 1, \\
&v_3 = 499\beta^2 - 28\beta + 2.
\end{aligned}$$

(iii) Usando las expresiones (3.0.6), (3.0.11) y (3.0.12), por definición de asimetría se tiene:

$$\begin{aligned}
\gamma_1 &= \frac{\mu_3 - 3\mu_1\mu_2 + 2\mu_1^3}{(\mu_2 - \mu_1^2)^{\frac{3}{2}}} \\
&= \frac{\frac{2(1182\beta^2 - 36\beta + 1)(2 + 15\beta^2)^2 e^{\frac{9}{2}}}{(2 + 15\beta^2)^3} - \frac{6(38\beta^2 - 4\beta + 1)(2 + 15\beta^2)(499\beta^2 - 28\beta + 2) e^{\frac{3}{2}}}{(2 + 15\beta^2)^3} + \frac{16(38\beta^2 - 4\beta + 1)^3 e^{\frac{1}{8}}}{(2 + 15\beta^2)^3}}{\left(\frac{(499\beta^2 - 28\beta + 2)}{2 + 15\beta^2} e^2 - \frac{2(38\beta^2 - 4\beta + 1)}{2 + 15\beta^2} e^{\frac{1}{2}} \right)^{\frac{3}{2}}} \\
&= \frac{2(v_1) H e^{\frac{9}{2}} - 6v_2 H v_3 e^{\frac{3}{2}} + 16v_2^3 e^{\frac{1}{8}}}{\left(H(v_3 e^2 - 2v_2 e^{\frac{1}{2}}) \right)^{\frac{3}{2}}} \\
&\text{con } H = 2 + 15\beta^2 \text{ y } v_1 = 1182\beta^2 - 36\beta + 1, \\
&v_2 = 38\beta^2 - 4\beta + 1, \quad v_3 = 499\beta^2 - 28\beta + 2.
\end{aligned}$$

(iv) Por definición de curtosis:

$$\begin{aligned} \gamma_2 &= \frac{\mu_4 - 4\mu_1\mu_3 + 6\mu_1^2\mu_2 - 3\mu_1^4}{(\mu_2 - \mu_1^2)^2} \\ &= \frac{(-8641\beta^2 - 152\beta + 2)(2+15\beta^2)^3 e^8 - 16(38\beta^2 - 4\beta + 1)(1182\beta^2 - 36\beta + 1)(2+15\beta^2)^2 e^5 + 24(38\beta^2 - 4\beta + 1)^2 (499\beta^2 - 28\beta + 2)(2+15\beta^2) e^3 - 48(38\beta^2 - 4\beta + 1)^3}{(2+15\beta^2)^4} \\ &= \frac{((499\beta^2 - 28\beta + 2)e^2 - 2(36\beta^2 - 4\beta + 1)e^{\frac{1}{2}})^{\frac{3}{2}}}{(2+15\beta^2)^{\frac{3}{2}}} \\ &= \frac{H^{\frac{3}{2}}(v_4 H^3 e^8 - 16v_2 v_1 H^2 e^5 + 24v_2^2 v_3 H e^3 - 48v_2^4 e^2)}{H^4(v_3 e^2 - 2v_5 e^{\frac{1}{2}})^{\frac{3}{2}}} \\ &\text{con } v_4 = -8641\beta^2 - 152\beta + 2 \text{ y } v_5 = 36\beta^2 - 4\beta + 1 \end{aligned}$$

(v) Por la definición de CV se sigue que:

$$\begin{aligned} CV &= \frac{\sqrt{Var(X)}}{E(X)} \\ &= \frac{\sqrt{\frac{[(499\beta^2 - 28\beta + 2)(2+15\beta^2)e^2 - (76\beta^2 - 8\beta + 2)^2]e}{(2+15\beta^2)^2}}}{\left| \frac{2(38\beta^2 - 4\beta + 1)e^{\frac{1}{2}}}{2+15\beta^2} \right|} \\ &= \frac{[(499\beta^2 - 28\beta + 2)(2+15\beta^2)e^2 - (76\beta^2 - 8\beta + 2)^2]^{\frac{1}{2}} e^{\frac{1}{2}}}{\left| \frac{2(38\beta^2 - 4\beta + 1)e^{\frac{1}{2}}}{2+15\beta^2} \right|} \\ &= \frac{\left\{ (499\beta^2 - 28\beta + 2)(2+15\beta^2) - [2(38\beta^2 - 4\beta + 1)] \right\}^{\frac{1}{2}}}{|2(38\beta^2 - 4\beta + 1)|} \\ &= \frac{(v_3 H - 2v_2)^{\frac{1}{2}}}{|2v_2|} \end{aligned}$$

(vi) Para la Función Generadora de momentos del modelo $LBSN(\beta)$ se tiene lo siguiente:

$$\text{Sea } u = \ln(x) \implies du = \frac{1}{x} dx \implies x du = dx$$

Ahora por definición de $M_X(t)$ se tiene

$$M_X(t) = E(e^{tu}) = \int_{-\infty}^{\infty} e^{tu} \frac{(1 - \beta u^3)^2 + 1}{(2 + 15\beta^2)x} \phi(u) x du$$

$$= \int_{-\infty}^{\infty} e^{tu} \frac{(1 - \beta u^3)^2 + 1}{(2 + 15\beta^2)} \phi(u) du$$

como $x > 0 \implies \{u > 0 \text{ si } x > 1 \text{ y } u \leq 0 \text{ si } x \leq 1\}$, entonces

$$M_X(t) = E(e^{tu}) = \int_{-\infty}^0 e^{tu} \frac{(1 - \beta u^3)^2 + 1}{(2 + 15\beta^2)} \phi(u) du + \int_0^{\infty} e^{tu} \frac{(1 - \beta u^3)^2 + 1}{(2 + 15\beta^2)} \phi(u) du$$

Resolviendo el segundo término de la integral se tiene:

$$\begin{aligned} & \int_0^{\infty} e^{tu} \frac{(1 - \beta u^3)^2 + 1}{(2 + 15\beta^2)} \phi(u) du = \frac{1}{2 + 15\beta^2} \int_0^{\infty} e^{tu} (2 - 2\beta u^3 + \beta^2 u^6) \phi(u) du \\ &= \frac{1}{2 + 15\beta^2} \left[2 \int_0^{\infty} e^{tu} \phi(u) du - 2\beta \int_0^{\infty} e^{tu} u^3 \phi(u) du + \beta^2 \int_0^{\infty} e^{tu} u^6 \phi(u) du \right] \\ &= e^{\frac{(-1)^2}{2}} \left\{ \frac{1}{\sqrt{2\pi}} \left[-(\ln x - t)^2 e^{\frac{-(\ln x - t)^2}{2}} - 2e^{\frac{-(\ln x - t)^2}{2}} \right] \Big|_0^{\infty} + \frac{3t}{\sqrt{2\pi}} \left[-(\ln x - t) e^{\frac{-(\ln x - t)^2}{2}} \right]_0^{\infty} \right. \\ &\quad \left. + 3t\Phi(\ln x - t) \Big|_0^{\infty} + \frac{3t^2}{\sqrt{2\pi}} \left(-e^{\frac{-(\ln x - t)^2}{2}} \right) \Big|_0^{\infty} + t^2\Phi(\ln x - t) \Big|_0^{\infty} \right\} \\ &= \lim_{b \rightarrow 0^+} e^{\frac{t^2}{2}} \left\{ \frac{1}{\sqrt{2\pi}} \left[-(\ln x - t)^2 e^{\frac{-(\ln x - t)^2}{2}} - 2e^{\frac{-(\ln x - t)^2}{2}} \right] \Big|_b^{\infty} + \frac{3t}{\sqrt{2\pi}} \left[-(\ln x - t) e^{\frac{-(\ln x - t)^2}{2}} \right] \Big|_b^{\infty} \right. \\ &\quad \left. + 3t\Phi(\ln x - t) \Big|_b^{\infty} + \frac{3t^2}{\sqrt{2\pi}} \left(-e^{\frac{-(\ln x - t)^2}{2}} \right) \Big|_b^{\infty} + t^2\Phi(\ln x - t) \Big|_b^{\infty} \right\} \\ &= \lim_{b \rightarrow 0^+} e^{\frac{t^2}{2}} \left\{ \frac{1}{\sqrt{2\pi}} \left[(\ln b - t)^2 e^{\frac{-(\ln b - t)^2}{2}} + 2e^{\frac{-(\ln b - t)^2}{2}} \right] + \frac{3t}{\sqrt{2\pi}} \left[-(\ln b - t) e^{\frac{-(\ln b - t)^2}{2}} \right] \right. \\ &\quad \left. + 3t \left(1 - \Phi(\ln b - t) \right) + \frac{3t^2}{\sqrt{2\pi}} \left(-e^{\frac{-(\ln b - t)^2}{2}} \right) + t^2 \left(1 - \Phi(\ln b - t) \right) \right\} \end{aligned}$$

como $\lim_{b \rightarrow 0^+} \ln b$ no existe, entonces se concluye que la función generadora de momentos del modelo $LBSN(\beta)$ no existe. \square

3.1. Extensión de localización-escala

A continuación para el caso localización escala de la variable aleatoria LBSN se hallan las funciones de densidad y de distribución acumulada, así como también

el momento r -ésimo.

Definición 3.1.1. Sea $Z \sim LBSN(\beta)$, con $\mu \in \mathbb{R}$ el parámetro de localización y $\sigma > 0$ el parámetro de escala respectivamente, entonces la función de densidad de probabilidad de Y está dada por:

$$f(y) = \frac{\left[1 - \beta \left(\frac{\ln(y) - \mu}{\sigma}\right)^2 + 1\right]}{\sigma(2 + 15\beta^2)} \phi\left(\frac{\ln(y) - \mu}{\sigma}\right) \quad (3.1.1)$$

Denotando a $Y \sim LBSN(\theta)$ con $\theta = (\mu, \sigma, \beta)'$, su respectiva función de distribución queda expresada como:

$$F(y) = \Phi(z) + \frac{4 - 15\beta^2 z + 2\beta z^2 - 5\beta^2 z^3 - \beta^2 z^5}{2 + 15\beta^2} \phi(z)$$

donde $z = \frac{\ln(y) - \mu}{\sigma}$.

El momento de orden r de Y se obtiene de la siguiente forma:

$$E(Y^r) = \sum_{k=0}^r \binom{r}{k} \mu^{r-k} \sigma^k E(Z^k)$$

donde $E(Z^k)$ es el k -ésimo momento de la variable aleatoria $Z \sim LBSN(\beta)$.

3.2. Inferencia para el Modelo LBSN

En esta sección se dan a conocer los resultados del estudio de estimación para el vector de parámetros $\theta = (\mu, \sigma, \beta)'$ del modelo localización-escala dado en (3.1.1) por el método de máxima verosimilitud, se determinan las derivadas parciales de primer orden de la función de log-verosimilitud con respecto a cada uno de los parámetros del modelo. Por último se hallan los elementos de las matrices de

información observada y esperada para luego verificar la no singularidad de la $I(\theta)$ cuando $\beta = 0$.

Por lo tanto supongase que $Y = (Y_1, Y_2, \dots, Y_n)$ es una muestra aleatoria, con $Y_i \sim LBSN(\theta)$ y vector de parámetros $\theta = (\mu, \sigma, \beta)'$, por definición la función log-verosimilitud esta dada por:

$$\mathbf{L}_{LBSN} = \mathbf{L}(\theta, \mathbf{Y}) = \sum_{i=1}^n \ln [(1 - \beta z_i^3)^2 + 1] - \frac{1}{2} \sum_{i=1}^n z_i^2 - n [\ln(\sigma) + \ln(2 + 15\beta^2) + \ln(\sqrt{2\pi})] \quad (3.2.1)$$

$$\text{con } z_i = \frac{\ln(y_i) - \mu}{\sigma}$$

Derivando la log-verosimilitud con respecto a cada parámetro se obtienen las derivadas parciales de primer orden del modelo $LBSN(\theta)$.

$$\begin{aligned} \frac{\partial l(\theta; \mathbf{z})}{\partial \mu} &= \frac{1}{\sigma} \sum_{i=1}^n \left[\frac{6\beta(1 - \beta z_i^3) z_i^2}{(1 - \beta z_i^3)^2 + 1} + z_i \right], \\ \frac{\partial l(\theta; \mathbf{z})}{\partial \sigma} &= \frac{1}{\sigma} \sum_{i=1}^n \left[\frac{6\beta(1 - \beta z_i^3) z_i^3}{(1 - \beta z_i^3)^2 + 1} - 1 + z_i^2 \right], \\ \frac{\partial l(\theta; \mathbf{z})}{\partial \beta} &= -2 \sum_{i=1}^n \left[\frac{z_i^3(1 - \beta z_i^3)}{(1 - \beta z_i^3)^2 + 1} + \frac{15\beta}{2 + 15\beta^2} \right]. \end{aligned}$$

para el modelo LBSN, A continuación usando la definición se obtienen los elementos de la matriz de información observada

$$j_{\theta_p \theta_q} = -\frac{\partial^2}{\partial \theta_p \partial \theta_q} \ell_{LBSN}, \quad \text{para } p, q = 1, 2, 3$$

para $\theta_1 = \mu, \theta_2 = \sigma, y \theta_3 = \beta$, se denotarán los elemnetos $j_{\mu\mu}, j_{\mu\sigma}, j_{\mu\beta}, j_{\sigma\sigma}, j_{\sigma\beta}$ y

$j_{\beta\beta}$, por lo tanto derivando parcialmente por segunda vez se obtiene:

$$\begin{aligned}
j_{\mu\mu} &= -\frac{1}{\sigma^2} \sum_{i=1}^n \left[\frac{6\beta z_i(5\beta z_i^3 - 2)}{(1 - \beta z_i^3)^2 + 1} - \frac{36\beta^2 z_i^4(1 - \beta z_i^3)^2}{[(1 - \beta z_i^3)^2 + 1]^2} - 1 \right], \\
j_{\mu\sigma} &= -\frac{2}{\sigma^2} \sum_{i=1}^n \left[\frac{9\beta z_i^2(2z_i^3 - 1)}{(1 - \beta z_i^3)^2 + 1} - \frac{18\beta^2 z_i^5(1 - \beta z_i^3)^2}{[(1 - \beta z_i^3)^2 + 1]^2} - z_i \right], \\
j_{\mu\beta} &= -\frac{6}{\sigma} \sum_{i=1}^n \left[\frac{z_i^2(1 - 2\beta z_i^3)}{(1 - \beta z_i^3)^2 + 1} + \frac{2\beta z_i^5(1 - \beta z_i^3)^2}{[(1 - \beta z_i^3)^2 + 1]^2} \right], \\
j_{\sigma\sigma} &= -\frac{1}{\sigma^2} \sum_{i=1}^n \left[\frac{2\beta z_i^3(21\beta z_i^3 - 12)}{(1 - \beta z_i^3)^2 + 1} - \frac{36\beta^2 z_i^6(1 - \beta z_i^3)^2}{[(1 - \beta z_i^3)^2 + 1]^2} + 1 - 3z_i^2 \right], \\
j_{\sigma\beta} &= -\frac{6}{\sigma} \sum_{i=1}^n \left[\frac{z_i^3(1 - 2\beta z_i^3)}{(1 - z_i^3)^2 + 1} + \frac{2\beta z_i^6(1 - \beta z_i^3)^2}{[(1 - \beta z_i^3)^2 + 1]^2} \right], \\
j_{\beta\beta} &= -2 \sum_{i=1}^n \left[\frac{z_i^6}{(1 - \beta z_i^3)^2 + 1} - \frac{2z_i^6(1 - \beta z_i^3)^2}{[(1 - \beta z_i^3)^2 + 1]^2} - \frac{15(2 - 15\beta^2)}{(2 + 15\beta^2)^2} \right].
\end{aligned}$$

Los elementos de la matriz de información esperada se determinan a partir de la siguiente expresión:

$$i_{\theta_p\theta_q} = \frac{1}{n} E \left[-\frac{\partial^2}{\partial\theta_p\partial\theta_q} \ell_{LBSN} \right]$$

con $\theta_1 = \mu, \theta_2 = \sigma$, y $\theta_3 = \beta$ respectivamente. Dichos elementos son denotados como $i_{\mu\mu}, i_{\mu\sigma}, i_{\mu\beta}, i_{\sigma\sigma}, i_{\sigma\beta}$ y $j_{\beta\beta}$.

Ahora definiendo a d_k , con $d_k = E \left[\frac{Z_0^k(1 - \beta Z_0^3)^2}{(1 - \beta Z_0^3)^2 + 1} \right]$, para $k = 4, 5, 6$, con $Z_0 \sim N(0, 1)$ por lo tanto la matriz de información esperada es:

$$I(\theta) = \begin{pmatrix} \frac{36\beta^2 d_4 - 75\beta^2 + 2}{\sigma^2(2 + 15\beta^2)} & \frac{36\beta^2 d_5 + 18\beta}{\sigma^2(2 + 15\beta^2)} & \frac{-12\beta d_5 - 6}{\sigma(2 + 15\beta^2)} \\ \frac{36\beta^2 d_5 + 18\beta}{\sigma^2(2 + 15\beta^2)} & \frac{-645\beta^2 + 36\beta^2 d_6 + 1}{\sigma^2(2 + 15\beta^2)} & \frac{12\beta(15 - d_6)}{\sigma(2 + 15\beta^2)} \\ \frac{-12\beta d_5 - 6}{\sigma(2 + 15\beta^2)} & \frac{12\beta(15 - d_6)}{\sigma(2 + 15\beta^2)} & \frac{60\beta^2 d_6 + 8d_6 - 900\beta^2}{(2 + 15\beta^2)^2} \end{pmatrix}$$

En particular, si $\beta = 0$ entonces la matriz de información está dada por:

$$I(\theta) = \begin{pmatrix} \frac{1}{\sigma^2} & 0 & \frac{-3}{\sigma} \\ 0 & \frac{1}{2\sigma^2} & 0 \\ \frac{-3}{\sigma} & 0 & 15 \end{pmatrix}$$

Usando la definición para hallar el determinante de una matriz 3×3 se tiene:

$$\begin{aligned} |I(\theta)| &= \left(\frac{1}{\sigma^2}\right) \begin{vmatrix} \frac{1}{2\sigma^2} & 0 \\ 0 & 15 \end{vmatrix} - (0) \begin{vmatrix} 0 & 0 \\ \frac{-3}{\sigma} & 15 \end{vmatrix} + \left(\frac{-3}{\sigma}\right) \begin{vmatrix} 0 & \frac{1}{2\sigma^2} \\ \frac{-3}{\sigma} & 0 \end{vmatrix} \\ &= \left(\frac{1}{\sigma^2}\right) \left(\frac{15}{2\sigma^2}\right) - 0 - \left(\frac{-3}{\sigma}\right) \left(\frac{-3}{2\sigma^3}\right) \\ &= \left(\frac{15}{2\sigma^4}\right) - \left(\frac{9}{2\sigma^4}\right) = \frac{6}{2\sigma^4} \neq 0. \end{aligned}$$

Así se verifica que $I(\theta)$ es no singular.

Entonces, la matriz de varianza y covarianza del vector de estimadores de los parámetros está determinada por

$$V(\hat{\theta}) = I^{-1}(\theta).$$

Para finalizar, utilizando la propiedad de convergencia asintótica para la distribución de los estimadores de máxima verosimilitud se tiene que

$$\hat{\theta} \xrightarrow{d} N_3(\theta, I^{-1}(\theta)).$$

Capítulo 4

Distribución Log-Beta-Normal Asimétrico Alfa-Potencia

En este capítulo se estudia la extensión log-alfa potencia del modelo $LBSN(\beta)$, el cual será denotado de aquí en adelante como LPBSN. Se define la fdp del modelo $LPBSN(\beta, \alpha)$, se halla la función de distribución y las expresiones de sus funciones de sobrevivencia y Hazard. También se determinan las expresiones del momento r -ésimo y de las principales características del modelo. Posteriormente se realiza el proceso de inferencia estadística para el modelo propuesto y se argumenta la estimación de sus parámetros para luego establecer las matrices de información observada y esperada y así determinar la no singularidad de esta.

Definición 4.0.1. Si la variable aleatoria Z tiene función de densidad,

$$f(z) = \alpha \frac{(1 - \beta \ln^3 z)^2 + 1}{2 + 15\beta^2} \frac{\phi(\ln z)}{z} \times \left[\Phi(\ln z) + \frac{\beta(2 \ln^2 z + 4 - \beta \ln^5 z - 5\beta \ln^3 z - 15\beta \ln z)}{2 + 15\beta^2} \phi(\ln z) \right]^{\alpha-1} \quad (4.0.1)$$

con $z > 0$, $\alpha \in \mathbb{R}^+$ y $\beta \in \mathbb{R}$, entonces se dice que Z es una variable aleatoria con distribución LPBSN estándar. Se denotará por $Z \sim LPBSN(\beta, \alpha)$.

A continuación se gráfica la distribución del modelo $LPBSN(\beta, \alpha)$. Se observa que conforme α se hace grande y β mas pequeño la función tiende a tomar un comportamiento trimodal.

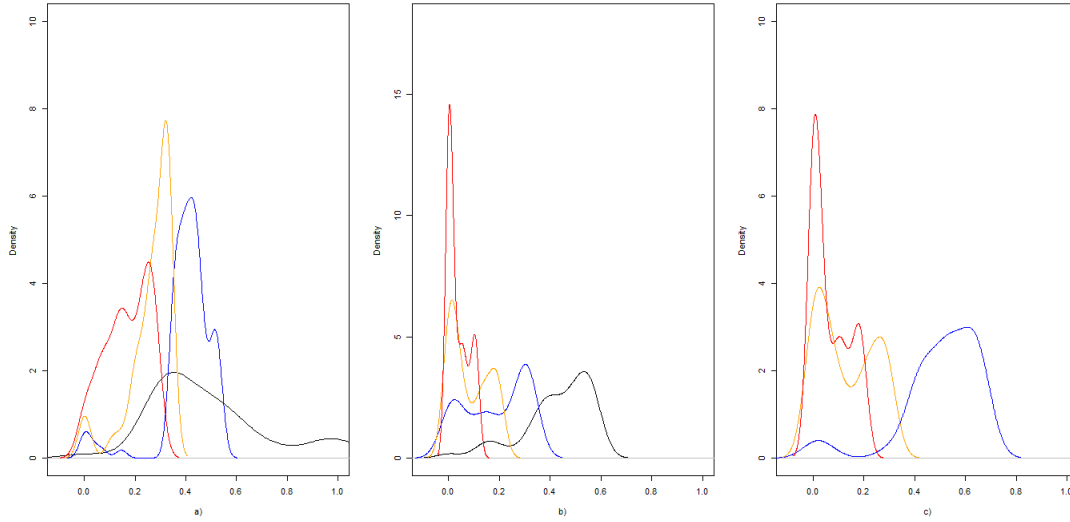


Figura (4.1): Distribución $LPBSN(\beta, \alpha)$ para valores de $\alpha = 1$ (línea negra), $\alpha = 2.0$ (línea azul), $\alpha = 3.0$ (línea naranja) y $\alpha = 4.0$ (línea roja), con (a) $\beta = 0$, $\beta = 0.25$ y (c) $\beta = -0.15$.

Proposición 4.0.1 La función de densidad del modelo $LPBSN(\beta, \alpha)$ tiene a lo más 3 modas.

Demostración. Diferenciando (4.0.1) con respecto a z se obtiene:

$$\begin{aligned} f'(z) &= \alpha \left[f_{LBSN}(z)(\alpha - 1) \left(F_{LBSN}(z) \right)^\alpha f_{LBSN}(z) + \left(F_{LBSN}(z) \right)^{\alpha-1} f'_{LBSN}(z) \right] \\ &= \alpha \left[f'_{LBSN}(z) \left(F_{LBSN}(z) \right)^{\alpha-1} + (\alpha - 1) f_{LBSN}^2(z) \left(F_{LBSN}(z) \right)^\alpha \right] \end{aligned}$$

donde $f'_{LBSN}(z)$ es la derivada de primer orden de la función de densidad de la variable aleatoria $LBSN$, como α es un parámetro de asimetría se puede tomar $\alpha = 1$, entonces se obtiene que

$$f'_{LPBSN}(z) = f_{LBSN}(z).$$

Así se concluye que la distribución LPBSN tiene a lo más tres modas. \square

Proposición 4.0.2 (FDA del modelo LPBSN(β, α) estándar).

Si $\Phi_{LPBSN}(z)$ denota la función de distribución de $Z \sim LPBSN(\beta, \alpha)$, entonces:

$$\Phi_{LPBSN}(z) = \left[\Phi(\ln z) + \frac{4\beta - 15\beta^2 \ln z + 2\beta \ln^2 z - 5\beta^2 \ln^3 z - \beta^2 \ln^5 z}{2 + 15\beta^2} \phi(\ln z) \right]^\alpha \quad (4.0.2)$$

Demostración. Por definición de la función de distribución de un modelo AP dada por (2.3.1), la FDA de $Z \sim PBSN(\beta, \alpha)$ y con $\alpha \in \mathbb{R}$, se obtiene de forma inmediata el resultado:

$$\begin{aligned} \Phi_{LPBSN}(z) &= \left[\Phi_{LBSN}(z) \right]^\alpha = \left[\Phi(\ln z) + \frac{\beta(2 \ln^2 z + 4 - \beta \ln^5 z - 5\beta \ln^3 z - 15\beta \ln z)}{2 + 15\beta^2} \phi(\ln z) \right]^\alpha \\ &= \left[\Phi(\ln z) + \frac{4\beta - 15\beta^2 \ln z + 2\beta \ln^2 z - 5\beta^2 \ln^3 z - \beta^2 \ln^5 z}{2 + 15\beta^2} \phi(\ln z) \right]^\alpha \end{aligned}$$

\square

Proposición 4.0.3 (Función de Supervivencia y Hazard del modelo LPBSN estándar).

Si $S(t)$ y $h(t)$ denotan, respectivamente, la función de supervivencia y la función hazard de la variable aleatoria $T \sim LPBSN(\beta, \alpha)$, entonces se tiene que por definición:

$$S(t) = 1 - \left[\Phi(\ln t) + \frac{4\beta - 15\beta^2 \ln t + 2\beta \ln^2 t - 5\beta^2 \ln^3 t - \beta^2 \ln^5 t}{2 + 15\beta^2} \phi(\ln t) \right]^\alpha \quad (4.0.3)$$

$$h(t) = \frac{\alpha \phi(\ln t) [(1 - \beta \ln^3 t)^2 + 1] \left[\Phi(\ln t) + \frac{4\beta - 15\beta^2 \ln t + 2\beta \ln^2 t - 5\beta^2 \ln^3 t - \beta^2 \ln^5 t}{2 + 15\beta^2} \phi(\ln t) \right]^{\alpha-1}}{t(2 + 15\beta^2) [1 - \Phi_{LPBSN}(t)]} \quad (4.0.4)$$

Demostración. Por definición de la función de supervivencia se tiene lo siguiente

$$\begin{aligned} S(t) &= P(T > t) = \int_t^\infty f(x) \cdot dx = 1 - F(t) = 1 - \Phi_{LPBSN}(t) \\ &= 1 - \left[\Phi(\ln t) + \frac{4\beta - 15\beta^2 \ln t + 2\beta \ln^2 t - 5\beta^2 \ln^3 t - \beta^2 \ln^5 t}{2 + 15\beta^2} \phi(\ln t) \right]^\alpha \end{aligned}$$

Ahora la función Hazard está determinada por:

$$\begin{aligned} h(t) &= \frac{f(t)}{S(t)} = \frac{f(t)}{1 - \Phi_{LPBSN}(t)} \\ &= \frac{\alpha \frac{[(1 - \beta \ln^3 t)^2 + 1] \phi(\ln t)}{2 + 15\beta^2} \left[\Phi(\ln t) + \frac{4\beta - 15\beta^2 \ln t + 2\beta \ln^2 t - 5\beta^2 \ln^3 t - \beta^2 \ln^5 t}{2 + 15\beta^2} \phi(\ln t) \right]^{\alpha-1}}{1 - \Phi_{LPBSN}(t)} \\ &= \frac{\alpha \phi(\ln t) (\Phi_{LPBSN})^{\alpha-1}}{t(2 + 15\beta^2) [1 - \Phi_{LPBSN}(t)]} \end{aligned}$$

□

En la siguiente gráfica (4.1) se observa el comportamiento de la función de sobrevivencia para distintos valores de β y α . En estos gráficos se aprecia que la curva se vuelve cada vez más horizontal en el intervalo 0 a 1.5 en la medida que β aumenta, y la probabilidad de sobrevivencia es mayor para valores de α más grandes, cuando β es constante.

También es importante resaltar que, sin importar los valores de β y α , la probabilidad de sobrevivencia se hace cero en la vecindad de $t = 1.5$.

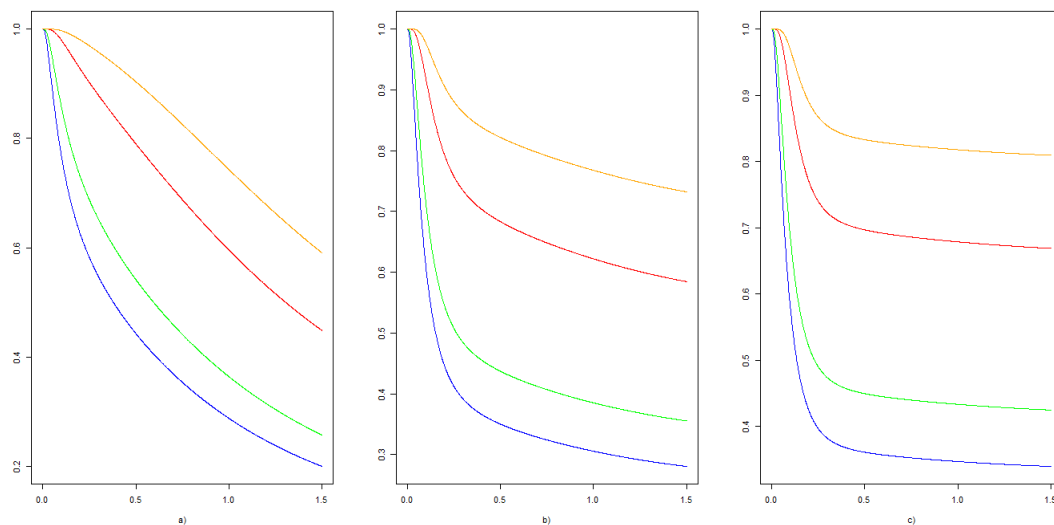


Figura 4.1: Función de sobrevivencia del modelo LPBSN (β, α) .

Nóte que para valores de $\alpha = 0.75$ (Línea Azul), $\alpha = 1.0$ (Línea Verde), $\alpha = 2.0$ (Línea Roja) y $\alpha = 3.0$ (Línea Naranja), para la figura a) con $\beta = 0.25$, b) con $\beta = 0.75$, y la figura c) con $\beta = 1.5$ respectivamente.

Proposición 4.0.4 (Momentos del modelo LPBSN estándar)

A continuación se presentan las expresiones generales del r -ésimo momento de la distribución LPBSN, así como también las siguientes medidas (la media, varianza, asimetría, curtosis y CV).

Para $Z \sim LPBSN(\beta, \alpha)$ el momento de orden r de la variable aleatoria Z viene dado por la definición general, a continuación:

$$\begin{aligned} \mu_r &= E(Z^r) \\ &= \alpha \int_0^\infty z^r \frac{[(1 - \beta \ln^3 z)^2 + 1] \phi(\ln z)}{2 + 15\beta^2} \times \\ &\quad \left[\Phi(\ln z) + \frac{4\beta - 15\beta^2 \ln z + 2\beta \ln^2 z - 5\beta^2 \ln^3 z - \beta^2 \ln^5 z}{2 + 15\beta^2} \phi(\ln z) \right]^{\alpha-1} \\ &= \alpha \int_0^\infty z^{r-1} \frac{[(1 - \beta \ln^3 z)^2 + 1] \phi(\ln z)}{2 + 15\beta^2} \times \\ &\quad \left[\Phi(\ln z) + \frac{4\beta - 15\beta^2 \ln z + 2\beta \ln^2 z - 5\beta^2 \ln^3 z - \beta^2 \ln^5 z}{2 + 15\beta^2} \phi(\ln z) \right]^{\alpha-1} \end{aligned}$$

Entonces, la media, la varianza, la asimetría, la curtosis y el CV del modelo LPBSN estándar se pueden encontrar mediante las siguientes expresiones:

- (i) $E(Z) = \mu_1$
- (ii) $V(Z) = \mu_2 - \mu_1^2$
- (iii) $\gamma_1 = \frac{\mu_3 - 3\mu_1\mu_2 + 2\mu_1^3}{(\mu_2 - \mu_1^2)^{\frac{3}{2}}}$
- (iv) $\gamma_2 = \frac{\mu_4 - 4\mu_1\mu_3 + 6\mu_1^2\mu_2 - 3\mu_1^4}{(\mu_2 - \mu_1^2)^2}$
- (v) $CV = \frac{\sqrt{Var(X)}}{E(X)}$

con $\mu_r = E[(Z - E(Z))^r]$ el momento central de orden r de la variable aleatoria $Z \sim LPBSN(\beta, \alpha)$.

4.1. Extensión de localización-escala

A continuación para la variable aleatoria $Z \sim LPBSN(\beta, \alpha)$ se halla fdp y la FDA, así como también el momento r -ésimo para el caso localización escala.

Definición 4.1.1. Sea $Z \sim LPBSN(\beta, \alpha)$, con $\mu \in \mathbb{R}$ el parámetro de localización y $\sigma > 0$ el parámetro de escala respectivamente, entonces la función de densidad de probabilidad de Y está dada por:

$$f(z) = \alpha \frac{[(1 - \beta z^3)^2 + 1]}{\sigma(2 + 15\beta^2)} \phi(z) \left[\Phi(z) + \frac{4\beta - 15\beta^2 z + 2\beta z^2 - 5\beta^2 z^3 - \beta^2 z^5}{2 + 15\beta^2} \phi(z) \right]^{\alpha-1} \quad (4.1.1)$$

donde $z = \frac{\ln(y) - \mu}{\sigma}$, con $y > 0$, $\alpha \in \mathbb{R}^+$ y $\beta \in \mathbb{R}$. Se sigue entonces que la función de distribución acumulada del modelo localización escala $LPBSN(\mu, \sigma, \beta, \alpha)$ queda expresada por:

$$F(z) = \left[\Phi(z) + \frac{4\beta - 15\beta^2 z + 2\beta z^2 - 5\beta^2 z^3 - \beta^2 z^5}{2 + 15\beta^2} \phi(z) \right]^{\alpha} \quad \text{donde } z = \frac{\ln(y) - \mu}{\sigma}$$

Proposición 4.0.5. (Momento de orden r del modelo LPBSN).

Si $Y \sim LPBSN(\mu, \sigma, \beta, \alpha)$, el momento de orden r está dado por:

$$E(Y^r) = \sum_{k=0}^r \binom{r}{k} \mu^{r-k} \sigma^k E(Z^k)$$

donde $E(Z^k)$ es el k -ésimo momento de la variable aleatoria $Z \sim LPBSN(\beta, \alpha)$.

4.2. Inferencia para el modelo LPBSN

En esta sección se dan a conocer los resultados del estudio de estimación para el vector de parámetros $\boldsymbol{\theta} = (\mu, \sigma, \beta, \alpha)'$ del modelo localización-escala dado en (4.1.1) por el método de máxima verosimilitud, se determinan las derivadas parciales de primer orden de la función de log-verosimilitud con respecto a cada uno de los parámetros del modelo. Por último se hallan los elementos de las matrices

de información observada y esperada para luego verificar la no singularidad de la $I(\theta)$ cuando $\beta = 0$.

Sea $Y = (Y_1, Y_2, \dots, Y_n)$ una muestra aleatoria, con $Y_i \sim LPBSN(\theta)$ y además $\theta = (\mu, \sigma, \beta, \alpha)'$, por definición se tiene que la función de verosimilitud es:

$$\begin{aligned} \mathbf{L}_{LPBSN} = \mathbf{L}(\theta; Y) &= \prod_{i=1}^n \alpha f_{LBSN}(y_i) \left[F_{LBSN}(y_i) \right]^{\alpha-1} \\ &= \alpha^n \prod_{i=1}^n f_{LBSN}(y_i) \left[\prod_{i=1}^n F_{LBSN}(y_i) \right]^{\alpha-1} \\ &= \alpha^n \prod_{i=1}^n \left[\frac{(1 - \beta \ln^3 y_i)^2 + 1}{2 + 15\beta^2} \frac{\phi(\ln y_i)}{y_i} \right] \times \\ &\quad \left\{ \prod_{i=1}^n \left[\Phi(\ln y_i) + \frac{4\beta - 15\beta^2 \ln y_i + 2\beta \ln^2 y_i - 5\beta^2 \ln^3 y_i - \beta^2 \ln^5 y_i}{2 + 15\beta^2} \phi(\ln y_i) \right]^{\alpha-1} \right\} \end{aligned}$$

Ahora tomando $z_i = \frac{\ln(y_i) - \mu}{\sigma}$ la función de log-verosimilitud queda expresada como:

$$\begin{aligned} \ell_{LPBSN} = \ell(\theta; Z) &= n \ln(\alpha) + \sum_{i=1}^n \ln [f_{LBSN}(z_i)] + (\alpha - 1) \sum_{i=1}^n \ln \left[\Phi_{BNS}(z_i) \right] \\ &= (\alpha - 1) \sum_{i=1}^n \ln \left[\Phi_{BNS}(z_i) \right] + \sum_{i=1}^n \ln [(1 - \beta z_i^3)^2 + 1] - \frac{1}{2} \sum_{i=1}^n z_i^2 - \\ &\quad n [\ln(\sigma) + \ln(2 + 15\beta^2) + \ln(\sqrt{2\pi}) - n \ln(\alpha)] \end{aligned}$$

Derivando la log-verosimilitud con respecto a cada parámetro se obtienen las derivadas parciales de primer orden del modelo $LPBSN(\theta)$.

$$\begin{aligned}
\frac{\partial}{\partial \mu} \ell_{LPBSN} &= -\frac{(\alpha-1)}{\sigma} \sum_{i=1}^n \frac{\phi_{BSN}(z_i)}{\Phi_{BSN}(z_i)} + \frac{1}{\sigma} \sum_{i=1}^n \left[\frac{6\beta(1-\beta z_i^3)z_i^2}{(1-\beta z_i^3)^2+1} + z_i \right] \\
\frac{\partial}{\partial \sigma} \ell_{LPBSN} &= -\frac{(\alpha-1)}{\sigma} \sum_{i=1}^n \frac{z_i \phi_{BSN}(z_i)}{\Phi_{BSN}(z_i)} + \frac{1}{\sigma} \sum_{i=1}^n \left[\frac{6\beta(1-\beta z_i^3)z_i^3}{(1-\beta z_i^3)^2+1} - 1 + z_i^2 \right] \\
\frac{\partial}{\partial \beta} \ell_{LPBSN} &= (\alpha-1) \sum_{i=1}^n \frac{W_i \phi(z_i)}{\Phi_{BSN}(z_i)} - 2 \sum_{i=1}^n \left[\frac{z_i^3(1-\beta z_i^3)}{(1-\beta z_i^3)^2+1} + \frac{15\beta}{2+15\beta^2} \right] \\
\frac{\partial}{\partial \alpha} \ell_{LPBSN} &= \sum_{i=1}^n \ln \left[\Phi_{BSN}(z_i) \right] + \frac{n}{\alpha}
\end{aligned}$$

Así, los elementos de la matriz de información observada por definición son:

$$j_{\theta_p \theta_q} = -\frac{\partial^2}{\partial \theta_p \partial \theta_q} \ell_{LPBSN}, \quad \text{para } p, q = 1, 2, 3, 4 \quad (4.2.1)$$

con $\theta_1 = \mu$, $\theta_2 = \sigma$, $\theta_3 = \beta$ y $\theta_4 = \alpha$. Estos elementos se denotarán como $j_{\mu\mu}, j_{\mu\sigma}, \dots, j_{\alpha\alpha}$ y quedan establecidos por:

$$\begin{aligned}
j_{\mu\mu} &= \frac{(\alpha-1)}{\sigma^2} \sum_{i=1}^n \left[\frac{6\beta z_i^2(1-\beta z_i^3)\phi(z_i)}{(2+15\beta^2)\Phi_{BSN}(z_i)} + \frac{z_i \phi_{BSN}(z_i)}{\Phi_{BSN}(z_i)} + \frac{\phi_{BSN}^2(z_i)}{\Phi_{BSN}^2(z_i)} \right] \\
&\quad - \frac{1}{\sigma^2} \sum_{i=1}^n \left[\frac{6\beta z_i(5\beta z_i^3-2)}{(1-\beta z_i^3)^2+1} - \frac{36\beta^2 z_i^4(1-\beta z_i^3)^2}{[(1-\beta z_i^3)^2+1]^2} - 1 \right] \\
j_{\mu\sigma} &= \frac{(\alpha-1)}{\sigma^2} \sum_{i=1}^n \left[\frac{6\beta z_i^3(1-\beta z_i^3)\phi(z_i)}{(2+15\beta^2)\Phi_{BSN}(z_i)} + \frac{(z_i^2-1)\phi_{BSN}(z_i)}{\Phi_{BSN}(z_i)} + \frac{z_i \phi_{BSN}^2(z_i)}{\Phi_{BSN}^2(z_i)} \right] \\
&\quad - \frac{2}{\sigma^2} \sum_{i=1}^n \left[\frac{9\beta z_i^2(2\beta z_i^3-1)}{(1-\beta z_i^3)^2+1} - \frac{18\beta^2 z_i^5(1-\beta z_i^3)^2}{[(1-\beta z_i^3)^2+1]^2} - z_i \right] \\
j_{\mu\beta} &= \frac{(\alpha-1)}{\sigma} \sum_{i=1}^n \left[\frac{2z_i^3(1-\beta z_i^3)\phi(z_i)}{(2+15\beta^2)\Phi_{BSN}(z_i)} + \frac{30\beta \phi_{BSN}(z_i)}{(2+15\beta^2)\Phi_{BSN}(z_i)} + W_i \frac{\phi(z_i)\phi_{BSN}(z_i)}{\Phi_{BSN}^2(z_i)} \right] \\
&\quad - \frac{6}{\sigma^2} \sum_{i=1}^n \left[\frac{z_i^2(1-2\beta z_i^3)}{(1-\beta z_i^3)^2+1} + \frac{2\beta z_i^5(1-\beta z_i^3)^2}{[(1-\beta z_i^3)^2+1]^2} \right]
\end{aligned}$$

$$\begin{aligned}
j_{\mu\alpha} &= \frac{1}{\sigma} \sum_{i=1}^n \frac{\phi_{\beta SN}(z_i)}{\Phi_{\beta SN}(z_i)} \\
j_{\sigma\sigma} &= \frac{(\alpha-1)}{\sigma} \sum_{i=1}^n \left[\frac{6\beta z_i^4(1-\beta z_i^3)\phi(z_i)}{(2+15\beta^2)\Phi_{\beta SN}(z_i)} + \frac{(z_i^3-z_i)\phi_{\beta SN}(z_i)}{\Phi_{\beta SN}(z_i)} + \frac{z_i^2\phi_{\beta SN}^2(z_i)}{\Phi_{\beta SN}^2(z_i)} \right] \\
&\quad - \frac{1}{\sigma^2} \sum_{i=1}^n \left[\frac{2\beta z_i^3(21\beta z_i^3-12)}{(1-\beta z_i^3)^2+1} + \frac{36\beta^2 z_i^6(1-\beta z_i^3)^2}{[(1-\beta z_i^3)^2+1]^2} + 1 - 3z_i^2 \right] \\
j_{\sigma\beta} &= \frac{(\alpha-1)}{\sigma} \sum_{i=1}^n \left[\frac{2z_i^4(1-\beta z_i^3)\phi(z_i)}{(2+15\beta^2)\Phi_{\beta SN}(z_i)} + \frac{30\beta z_i\phi_{\beta SN}(z_i)}{(2+15\beta^2)\Phi_{\beta SN}(z_i)} + \frac{z_i W_i \phi(z_i)\phi_{\beta SN}^2(z_i)}{\Phi_{\beta SN}^2(z_i)} \right] \\
&\quad - \frac{6}{\sigma} \sum_{i=1}^n \left[\frac{z_i^3(1-2\beta z_i^3)}{(1-\beta z_i^3)^2+1} + \frac{2\beta z_i^6(1-\beta z_i^3)^2}{[(1-\beta z_i^3)^2+1]^2} \right] \\
j_{\sigma\alpha} &= \frac{1}{\sigma} \sum_{i=1}^n \frac{z_i\phi_{\beta SN}(z_i)}{\Phi_{\beta SN}(z_i)} \\
j_{\beta\beta} &= -(\alpha-1) \sum_{i=1}^n \phi(z_i) \left[\frac{U_i}{\Phi_{\beta SN}(z_i)} - \frac{W_i^2\phi(z_i)}{\Phi_{\beta SN}^2(z_i)} \right] \\
&\quad - 2 \sum_{i=1}^n \left[\frac{z_i^6}{(1-\beta z_i^3)^2+1} - \frac{2z_i^6(1-\beta z_i^3)^2}{[(1-\beta z_i^3)^2+1]^2} - \frac{15(2-15\beta^2)}{(2+15\beta^2)^2} \right] \\
j_{\beta\alpha} &= - \sum_{i=1}^n \frac{W_i\phi(z_i)}{\Phi_{\beta SN}(z_i)} \\
j_{\alpha\alpha} &= \frac{n}{\alpha^2}
\end{aligned}$$

con

$$\begin{aligned}
U_i &= \frac{-720\beta + 1800\beta^3 - 120z_i + 2700\beta^2 z_i - 360\beta z_i^2 + 900\beta^2 z_i^2 - 40z_i^3 + 900\beta^2 z_i^3 - 8z_i^5 + 180\beta^2 z_i^5}{(2+15\beta^2)^3} \\
W_i &= \frac{8 - 60\beta^2 - 60\beta z_i + 4z_i^2 - 30\beta^2 z_i^2 - 20\beta z_i^3 - 4\beta z_i^5}{(2+15\beta^2)^2}
\end{aligned}$$

Ahora para obtener los elementos de la matriz de información esperada se utiliza la siguiente expresión

$$i_{\theta_p\theta_q} = \frac{1}{n} E \left[-\frac{\partial^2}{\partial\theta_p\partial\theta_q} \ell_{LPBSN} \right], \quad p, q = 1, 2, 3, 4 \quad (4.2.2)$$

con $\theta_1 = \mu$, $\theta_2 = \sigma$, $\theta_3 = \beta$ y $\theta_4 = \alpha$ y seran denotados como $i_{\mu\mu}, i_{\mu\sigma}, \dots, i_{\alpha\alpha}$

Definiendo:

$$\begin{aligned} h_j &= E \left[z^j \frac{\phi(z)}{\Phi_{\beta SN}(z)} \right], & v_{jk} &= E \left[z^j \left(\frac{\phi_{\beta SN}(z)}{\Phi_{\beta SN}(z)} \right)^k \right] \\ g_{jk} &= E \left[\frac{z^j}{[(1 - \beta z^3)^2 + 1]^k} \right], & u_{jk} &= E \left[\frac{z^j W^k \phi(z) \phi_{\beta SN}(z)}{\Phi_{\beta SN}^2(z)} \right] \\ a &= E \left[\frac{U \phi(z)}{\Phi_{\beta SN}(z)} \right], & y & \quad b_k = E \left[\frac{W^k \phi^k(z)}{\Phi_{\beta SN}^k(z)} \right] \end{aligned}$$

con $Z \sim LPBSN(\beta, \alpha)$, los elementos de la matriz esperada quedan dadas por:

$$\begin{aligned} i_{\mu\mu} &= \frac{(\alpha - 1)}{\sigma^2} \left[\frac{6\beta}{2 + 15\beta^2} (h_2 - \beta h_5) + v_{1,1} + v_{0,2} \right] \\ &\quad - \frac{1}{\sigma^2} [6\beta(5\beta g_{4,1} - 2g_{1,1}) - 36\beta^2(g_{4,2} - 2\beta g_{7,2} + \beta^2 g_{10,2}) - 1] \\ i_{\mu\sigma} &= \frac{(\alpha - 1)}{\sigma^2} \left[\frac{6\beta}{2 + 15\beta^2} (h_3 - \beta h_6) + v_{2,1} - v_{0,2} + v_{1,2} \right] \\ &\quad - \frac{2}{\sigma^2} [9\beta(2\beta g_{5,1} - g_{2,1}) - 18\beta^2(g_{5,2} - 2\beta g_{8,2} + \beta^2 g_{11,2}) - g_{1,2}] \\ i_{\mu\beta} &= -\frac{(\alpha - 1)}{\sigma} \left[\frac{2}{2 + 15\beta^2} (h_3 - \beta h_6) + 30\beta v_{0,1} + u_{0,1} \right] \\ &\quad - \frac{6}{\sigma} [(g_{2,1} - 2\beta g_{3,1}) + 2\beta(g_{5,2} - 2\beta g_{8,2} + \beta^2 g_{11,2})] \\ i_{\mu,\alpha} &= \frac{1}{\sigma} v_{0,1} \\ i_{\sigma\sigma} &= \frac{(\alpha - 1)}{\alpha} \left[\frac{6\beta}{2 + 15\beta^2} (h_4 - \beta h_7) + v_{3,1} - v_{1,1} + v_{2,2} \right] \\ &\quad - \frac{1}{\sigma^2} [2\beta(21\beta g_{6,1} - 12g_{3,1}) - 36\beta^2(g_{6,2} - 2\beta g_{9,2} + \beta^2 g_{12,2}) + 1 - g_{2,0}] \\ i_{\sigma\beta} &= -\frac{(\alpha - 1)}{\sigma} \left[\frac{2}{2 + 15\beta^2} (h_4 - \beta h_7) + \frac{30\beta}{2 + 15\beta^2} v_{1,1} + u_{1,1} \right] \\ &\quad - \frac{6}{\sigma} [(g_{3,1} - 2\beta g_{6,1}) + 2\beta(g_{6,2} - 2\beta g_{9,2} + \beta^2 g_{12,2})] \end{aligned}$$

$$i_{\sigma\alpha} = \frac{1}{\sigma} v_{1,1}$$

$$i_{\beta\beta} = -(\alpha - 1)[a - b_2] - 2 \left[g_{6,1} - 2(g_{6,1} - 2\beta g_{4,2} + \beta^2 g_{12,2}) - \frac{15(2 - 15\beta^2)}{(2 + 15\beta^2)^2} \right]$$

$$i_{\beta\alpha} = -b_1$$

$$i_{\alpha,\alpha} = \frac{1}{\alpha^2}$$

Ahora tomando $\beta = 0$ y $\alpha = 1$, la matriz de información estaría dada por:

$$I(\mu, \sigma, \theta, 0, 1) = \begin{pmatrix} \frac{1}{\sigma^2} & \frac{2}{\sigma^2} g_{1,0} & -\frac{6}{\sigma} g_{2,1} & \frac{1}{\sigma} v_{0,1} \\ \frac{2}{\sigma^2} g_{1,0} & \frac{1}{\sigma^2} (3g_{2,0} - 1) & -\frac{6}{\sigma} g_{3,1} & \frac{1}{\sigma} v_{1,1} \\ -\frac{6}{\sigma} g_{2,2} & -\frac{6}{\sigma} g_{3,1} & -2g_{6,1} + 4g_{6,2} + 15 & -b_1 \\ \frac{1}{\sigma} v_{0,1} & \frac{1}{\sigma} v_{1,1} & -b_1 & \frac{1}{\sigma^2} \end{pmatrix}_{4 \times 4}$$

Como $|I_p(\boldsymbol{\theta})| \neq 0$, entonces $I_p(\boldsymbol{\theta})$ es no singular.

Así, la matriz de varianza y covarianza del vector de estimadores de los parámetros de la distribución $l\beta$ normal asimétrica alfa-potencia está dada por:

$$V(\hat{\theta}) = I^{-1}(\boldsymbol{\theta}).$$

Por último, al usar la propiedad de convergencia asintótica para la distribución de los estimadores de máxima verosimilitud se obtiene:

$$\hat{\theta} \xrightarrow{d} N_4(\theta, I_p^{-1}(\boldsymbol{\theta})).$$

Capítulo 5

Distribución Log-Beta-Normal Asimétrica con Datos Censurados

En este capítulo se estudia el modelo log-beta-normal asimétrica con datos censurados (CLBSN), como un caso especial del modelo $CLPBSN$ cuando $\alpha = 1$, se define la fdp del modelo, se halla la FDA, la expresión del momento de orden r , la esperanza y la varianza de a distribución CLBSN. Luego se estudia la estimación de sus parámetros y se hallan las matrices de información de Fisher.

Sea Y una variable aleatoria que tiene una distribución CLBSN. Considere una muestra aleatoria $(Y_1^*, Y_2^*, \dots, Y_n^*)$, donde solamente son registrados los valores de Y^* que son mayores que una constante c . Además para valores $Y^* \leq c$ solo se registra el valor de c . Por lo tanto, para $i = 1, 2, \dots, n$, los valores observados son:

$$Y_i = \begin{cases} c & \text{si } Y_i^* \leq c, \\ Y_i^* & \text{si } Y_i^* > c \end{cases} \quad (5.0.1)$$

entonces la muestra resultante es una muestra censurada a la izquierda.

Sea F_{LBSN} , denota la FDA de una variable aleatoria $Y^* \sim LBSN(\mu, \sigma, \beta)$, para el

caso localización-escala se define, $z_c = \frac{\ln c - \mu}{\sigma}$ y se tiene que:

$$\begin{aligned}
 P(Y = c) &= P(Y^* \leq c) \\
 &= F_{LBSN}(z_c) \\
 &= \Phi_{LBSN}\left(\frac{\ln c - \mu}{\sigma}\right) \\
 &= \Phi(z_c) + \frac{4\beta - 15\beta^2 z_c + 2\beta z_c^2 - 5\beta^2 z_c^3 - \beta^2 z_c^5}{2 + 15\beta^2} \phi(z_c)
 \end{aligned}$$

Definición 5.0.1. Si la variable aleatoria Y , tiene función de densidad,

$$f(y) = \begin{cases} \Phi_{LBSN}\left(\frac{\ln c - \mu}{\sigma}\right) & \text{si } y \leq c, \\ f_{LBSN}(y) & \text{si } y > c \end{cases} \quad (5.0.2)$$

donde f_{LBSN} es la función de densidad de una variable aleatoria LBSN y c es una constante, entonces se dice que Y es una variable aleatoria con distribución CLBSN (cebsurada a la izquierda), se denotará como $Y \sim CLBSN(\boldsymbol{\theta})$ con $\boldsymbol{\theta} = (\mu, \sigma, \beta)'$, su función de densidad también puede ser expresada como

$$f(y) = f_{LBSN}^i(y) \Phi_{LBSN}^{1-i}\left(\frac{\ln c - \mu}{\sigma}\right) \quad (5.0.3)$$

donde

$$I = \begin{cases} 0 & \text{si } y \leq c, \\ 1 & \text{si } y > c \end{cases}$$

En la Figura (5.1) se presenta la función de densidad para la variable aleatoria $Y \sim CLBSN(\boldsymbol{\theta})$ con un punto de censura $c = 0.4$. y $c = 1$ respectivamente.

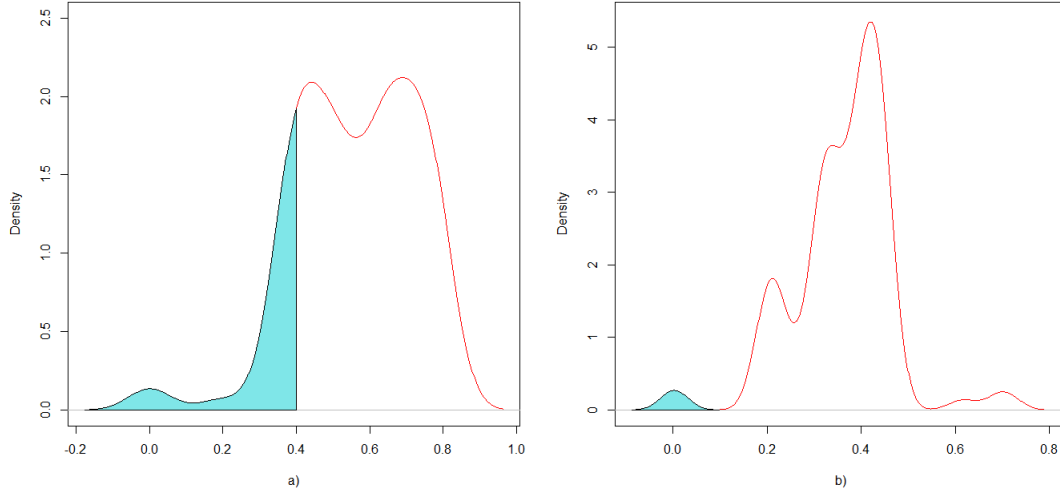


Figura 5.1: fdp del modelo $CLBSN(\beta)$.

Observe que la función está censurada a la izquierda (área de color) para $\beta = 0.1$ y $\beta = -0.2$.

Proposición 5.0.1 (Momento de orden r del modelo $CLBSN$) .

Si $Y \sim CLBSN(\mu, \sigma, \beta)$, el momento r -ésimo está dado por:

$$E(Y^r) = c^r \Phi_{BSN}(z_c) + \frac{e^{r\mu} M'_k}{2 + 15\beta^2} \quad (5.0.4)$$

$$\text{con } z_c = \frac{\ln c - \mu}{\sigma}$$

donde

$$M'_k = 2\mu'_0(z_c) - 2\beta\mu'_3(z_c) + \beta^2\mu'_6(z_c) \quad \text{y} \quad \mu'_q(z_c) = \int_{z_c}^{\infty} z^q e^{r\sigma z} \phi(z) dz \quad \text{para } q = 0, 3, 6.$$

Demostración. Por definición de valor esperado condicionado, se tiene que:

$$\begin{aligned}
E(Y^r|y > c) &= \int_c^\infty y^r f(y|y > c) dy \\
&= \int_c^\infty y^r \frac{f_{LBSN}(y)}{P(Y > c)} dy \\
&= \int_c^\infty y^r \frac{(1-\beta \ln^3 y)^2 + 1}{2+15\beta^2} \frac{\phi(\ln y)}{P(Y > c)} dy \\
&= \int_c^\infty y^r \frac{[(1-\beta \ln^3 y)^2 + 1]}{y(2+15\beta^2)P(Y > 0)} \phi(\ln y) dy
\end{aligned}$$

Ahora haciendo $z = \frac{\ln y - \mu}{\sigma}$ y redefiniendo el límite inferior de la integral como $z_c = \frac{\ln c - \mu}{\sigma}$ se tiene que:

$$\begin{aligned}
E(Y^r|y > c) &= \frac{\int_{z_c}^\infty \left[e^{(\sigma z + \mu)} \right]^r f_{BSN}}{P\left(z > \frac{\ln c - \mu}{\sigma}\right)} dz \\
&= \frac{\int_{z_c}^\infty \left[e^{r(\sigma z + \mu)} \right] \left[(1 - \beta z^3)^2 + 1 \right] \phi(z)}{\left[1 - \Phi_{BSN}(z_c) \right] (2 + 15\beta^2)} dz \\
&= \frac{e^{r\mu}}{\left[1 - \Phi_{BSN}(z_c) (2 + 15\beta^2) \right]} \times \\
&\quad \int_{z_c}^\infty e^{r\sigma z} [2 - 2\beta z^3 + \beta^2 z^6 + 1] \phi(z) dz \\
&= \frac{e^{r\mu}}{\left[1 - \Phi_{BSN}(z_c) (2 + 15\beta^2) \right]} \times \\
&\quad \left\{ 2 \int_{z_c}^\infty e^{r\sigma z} \phi(z) dz - 2\beta \int_{z_c}^\infty z^3 e^{r\sigma z} \phi(z) dz + \beta^2 \int_{z_c}^\infty z^6 e^{r\sigma z} \phi(z) dz \right\}
\end{aligned}$$

Tomando $\mu'_q(z_c) = \int_{z_c}^\infty z^q e^{r\sigma z} \phi(z) dz$ para $q = 0, 3, 6$, y

$M'_k = 2\mu'_0(z_c) - 2\beta\mu'_3(z_c) + \beta^2\mu'_6(z_c)$, entonces,

$$E(Y^r|y > c) = \frac{e^{r\mu} M'_k}{\left[1 - \Phi_{BSN}(z_c)\right] (2 + 15\beta^2)}$$

Usando la definición de la esperanza para una variable aleatoria limitada, el valor esperado de orden r para $Y \sim CLBSN(\boldsymbol{\theta})$ es:

$$\begin{aligned} E(Y^r) &= E(Y^r|y = c)P(y = c) + E(Y^r|y > c)P(y > c) \\ &= E(c^r)P(Y \leq c) + E(Y^r|y > c)P(y > c) \\ &= c^r \Phi_{BSN}\left(\frac{\ln c - \mu}{\sigma}\right) + \frac{e^{r\mu} M'_k (1 - \Phi_{SN}(z_c))}{(1 - \Phi_{BSN}(z_c))(2 + 15\beta^2)} \\ &= c^r \Phi_{BSN}(z_c) + \frac{e^{r\mu} M'_k}{(2 + 15\beta^2)} \end{aligned}$$

□

Proposición 5.0.2 (Esperanza y varianza del modelo CLBSN estándar).

Sea $Y \sim CLBSN(\boldsymbol{\theta})$ donde $\boldsymbol{\theta} = (\mu, \sigma, \beta)'$, entonces la esperanza y la varianza están dadas por:

$$(i) \quad E(Y) = c\Phi_{BSN}(z_c) + \frac{e^\mu M'_1}{2+15\beta^2}$$

$$(ii) \quad V(Y) = c^2\Phi_{BSN}(z_c) + \frac{e^{2\mu} M'_2}{2+15\beta^2} + c^2\Phi_{BSN}^2(z_c) + \frac{2ce^\mu\Phi_{BSN}(z_c)M'_1}{2+15\beta^2} + \left(\frac{e^\mu M'_1}{2+15\beta^2}\right)^2$$

Demostración. Utilizando la expresión dada en (5.8) y para $r = 1$, $k = 1$, se tiene:

i) Por definición y utilizando la expresión (5.0.4),

$$\begin{aligned}
E(Y) &= c' \Phi_{BSN}(z_c) + \frac{e^\mu M_1'}{2 + 15\beta^2} \\
E(Y^2) &= c^2 \Phi_{BSN}(z_c) + \frac{e^{2\mu} M_2'}{2 + 15\beta^2} \\
Var(Y) &= E(Y^2) - E^2(Y) \\
&= c^2 \Phi_{BSN}(z_c) + \frac{e^{2\mu} M_2'}{2 + 15\beta^2} + c^2 \Phi_{BSN}^2(z_c) \\
&\quad + \frac{2ce^\mu \Phi_{BSN}(z_c) M_1'}{2 + 15\beta^2} + \left(\frac{e^\mu M_1'}{2 + 15\beta^2} \right)^2
\end{aligned}$$

□

5.1. Inferencia para el modelo CLBSN

En esta sección se estudia por el método de máxima verosimilitud la estimación de los parámetros del modelo LBSN y se hallan las matrices de información observada y esperada.

Sea $Y = (Y_1, Y_2, \dots, Y_n)$ una muestra aleatoria, con $Y_i \sim CBSN(\boldsymbol{\theta})$ y además $\boldsymbol{\theta} = (\mu, \sigma, \beta)'$, en la cual hay n_0 observaciones censuradas y n_1 no censuradas, la función de verosimilitud está dada por:

$$\mathbf{L}_{CLBSN} = \ell(\boldsymbol{\theta}, Y) = \prod_{y_i \leq c} \Phi_{LBSN} \left(\frac{\ln c - \mu}{\sigma} \right) \prod_{y_i > c} f_{LBSN}(y_i)$$

Ahora al tomar $z_i = \frac{\ln y_i - \mu}{\sigma}$ y $z_c = \frac{\ln c - \mu}{\sigma}$ se tiene que la log-verosimilitud queda expresada como:

$$\begin{aligned}
\ell_{CLBSN} = \ell(\theta, Z) &= \sum_{z_i > z_c} \ln \left[f_{BSN}(z_i) \right] + \sum_{z_i \leq z_c} \ln \left[\Phi_{BSN}(z_c) \right] \\
&= \sum_{z_i > z_c} \ln \left[f_{BSN}(z_i) \right] + n_0 \ln \left[\Phi_{BSN}(z_c) \right]. \\
&= n_0 \ln \left[\Phi(z_c) + \frac{4\beta - 15\beta^2 z_c + 2\beta z_c^2 - 5\beta^2 z_c^3 - \beta^2 z_c^5}{2 + 15\beta^2} \phi(z_c) \right] \\
&\quad + \sum_{z_i > z_c} \ln \left[(1 - \beta z_i^3)^2 + 1 \right] - \frac{1}{2} \sum_{z_i > z_c} z_i^2 - n_1 \left[\ln(\sigma) + \ln(2 + 15\beta^2) + \ln(\sqrt{2\pi}) \right].
\end{aligned}$$

Por lo tanto, las derivadas de primer orden de la función de log-verosimilitud quedan dadas por:

$$\begin{aligned}
\frac{\partial}{\partial \mu} \ell_{CLBSN} &= -\frac{n_0}{\sigma} \frac{\phi_{BSN}(z_c)}{\Phi_{BSN}(z_c)} + \frac{1}{\sigma} \sum_{z_i > z_c} \left[\frac{6\beta(1 - \beta z_i^3) z_i^2}{(1 - \beta z_i^3)^2 + 1} + z_i \right] \\
\frac{\partial}{\partial \sigma} \ell_{CLBSN} &= -\frac{n_0}{\sigma} \frac{z_c \phi_{BSN}(z_c)}{\Phi_{BSN}(z_c)} - \frac{1}{\sigma} \sum_{z_i > z_c} \left[\frac{6\beta(1 - \beta z_i^3) z_i^3}{(1 - \beta z_i^3)^2 + 1} - 1 + z_i^2 \right] \\
\frac{\partial}{\partial \beta} \ell_{CLBSN} &= n_0 \frac{W_c \phi(z_c)}{\Phi_{BSN}(z_c)} - 2 \sum_{z_i > z_c} \left[\frac{z_i^3(1 - \beta z_i^3)}{(1 - \beta z_i^3)^2 + 1} + \frac{15\beta}{2 + 15\beta^2} \right]
\end{aligned}$$

donde $W_c = \frac{8 - 60\beta^2 - 60\beta z_c + 4z_c^2 - 30\beta^2 z_c^2 - 20\beta z_c^3 - 4\beta z_c^5}{(2 + 15\beta^2)^2}$

Los elementos de la matriz de información observada para el modelo $CLBSN$ pueden ser obtenidos, por definición, usando la siguiente expresión:

$$j_{\theta_p \theta_q} = -\frac{\partial^2}{\partial \theta_p \partial \theta_q} \ell_{CLBSN}, \quad \text{para } p, q = 1, 2, 3. \quad (5.1.1)$$

con $\theta_1 = \mu$, $\theta_2 = \sigma$, y $\theta_3 = \beta$. Estos elementos serán denotados como $j_{\mu\mu}$, $j_{\mu\sigma}$, $j_{\mu\beta}$, $j_{\sigma\sigma}$, $j_{\sigma\beta}$ y $j_{\beta\beta}$ los cuales se representan a continuación, tomando $\alpha = 1$ usado en el modelo $CLPBSN$ se tienen los siguientes resultados:

$$\begin{aligned}
j_{\mu\mu} &= \frac{n_0}{\sigma^2} \left[\frac{6\beta z_c^2(1-\beta z_c^3)\phi(z_c)}{(2+15\beta^2)\Phi_{BSN}(z_c)} + \frac{z_c\phi_{BSN}(z_c)}{\Phi_{BSN}(z_c)} + \frac{\phi_{BSN}^2(z_c)}{\Phi_{BSN}^2(z_c)} \right] \\
&\quad - \frac{1}{\sigma^2} \sum_{z_i > z_c} \left[\frac{6\beta z_i(5\beta z_i^3-2)}{(1-\beta z_i^3)^2+1} - \frac{36\beta^2 z_i^4(1-\beta z_i^3)^2}{[(1-\beta z_i^3)^2+1]^2} - 1 \right] \\
j_{\mu\sigma} &= \frac{n_0}{\sigma^2} \left[\frac{6\beta z_c^3(1-\beta z_c^3)\phi(z_c)}{(2+15\beta^2)\Phi_{BSN}(z_c)} + \frac{(z_c^2-1)\phi_{BSN}(z_c)}{\Phi_{BSN}(z_c)} + \frac{z_c\phi_{BSN}^2(z_c)}{\Phi_{BSN}^2(z_c)} \right] \\
&\quad - \frac{2}{\sigma^2} \sum_{z_i > z_c} \left[\frac{9\beta z_i^2(2\beta z_i^3-1)}{(1-\beta z_i^3)^2+1} - \frac{18\beta^2 z_i^5(1-\beta z_i^3)^2}{[(1-\beta z_i^3)^2+1]^2} - z_i \right] \\
j_{\mu\beta} &= -\frac{n_0}{\sigma} \left[\frac{2z_c^3(1-\beta z_c^3)\phi(z_c)}{(2+15\beta^2)\Phi_{BSN}(z_c)} + \frac{30\beta\phi_{BSN}(z_c)}{(2+15\beta^2)\Phi_{BSN}(z_c)} + W_i \frac{\phi(z_c)\phi_{BSN}(z_c)}{\Phi_{BSN}^2(z_c)} \right] \\
&\quad - \frac{6}{\sigma^2} \sum_{z_i > z_c} \left[\frac{z_i^2(1-2\beta z_i^3)}{(1-\beta z_i^3)^2+1} + \frac{2\beta z_i^5(1-\beta z_i^3)^2}{[(1-\beta z_i^3)^2+1]^2} \right] \\
j_{\mu\alpha} &= \frac{n_0}{\sigma} \frac{\phi_{BSN}(z_c)}{\Phi_{BSN}(z_c)} + \frac{1}{\sigma} \sum_{z_i > z_c} \frac{\phi_{BSN}(z_i)}{\Phi_{BSN}(z_i)} \\
j_{\sigma\sigma} &= \frac{n_0}{\sigma} \left[\frac{6\beta z_c^4(1-\beta z_c^3)\phi(z_c)}{(2+15\beta^2)\Phi_{BSN}(z_c)} + \frac{(z_c^3-z_c)\phi_{BSN}(z_c)}{\Phi_{BSN}(z_c)} + \frac{z_c^2\phi_{BSN}^2(z_c)}{\Phi_{BSN}^2(z_c)} \right] \\
&\quad - \frac{1}{\sigma^2} \sum_{z_i > z_c} \left[\frac{2\beta z_i^3(21\beta z_i^3-12)}{(1-\beta z_i^3)^2+1} + \frac{36\beta^2 z_i^6(1-\beta z_i^3)^2}{[(1-\beta z_i^3)^2+1]^2} + 1 - 3z_i^2 \right] \\
j_{\sigma\beta} &= -\frac{n_0}{\sigma} \left[\frac{2z_c^4(1-\beta z_c^3)\phi(z_c)}{(2+15\beta^2)\Phi_{BSN}(z_c)} + \frac{30\beta z_c\phi_{BSN}(z_c)}{(2+15\beta^2)\Phi_{BSN}(z_c)} + \frac{z_c W_c \phi(z_c)\phi_{BSN}^2(z_c)}{\Phi_{BSN}^2(z_c)} \right] \\
&\quad - \frac{6}{\sigma} \sum_{z_i > z_c} \left[\frac{z_i^3(1-2\beta z_i^3)}{(1-\beta z_i^3)^2+1} + \frac{2\beta z_i^6(1-\beta z_i^3)^2}{[(1-\beta z_i^3)^2+1]^2} \right] \\
j_{\sigma\alpha} &= \frac{n_0}{\sigma} \frac{z_c\phi_{BSN}(z_c)}{\Phi_{BSN}(z_c)} + \frac{1}{\sigma} \sum_{z_i > z_c} \frac{z_i\phi_{BSN}(z_i)}{\Phi_{BSN}(z_i)} \\
j_{\beta\beta} &= -n_0\phi(z_c) \left[\frac{U_c}{\Phi_{BSN}(z_c)} - \frac{W_c^2\phi(z_c)}{\Phi_{BSN}^2(z_c)} \right] \\
&\quad - 2 \sum_{z_i > z_c} \left[\frac{z_i^6}{(1-\beta z_i^3)^2+1} - \frac{2z_i^6(1-\beta z_i^3)^2}{[(1-\beta z_i^3)^2+1]^2} - \frac{15(2-15\beta^2)}{(2+15\beta^2)^2} \right] \\
j_{\beta\alpha} &= -n_0 \frac{W_c\phi(z_c)}{\Phi_{BSN}(z_c)} - \sum_{z_i > z_c} \frac{W_i\phi(z_i)}{\Phi_{BSN}(z_i)} \\
j_{\alpha\alpha} &= \frac{n_1}{\sigma^2}
\end{aligned}$$

donde se tiene:

$$U_c = \frac{-720\beta + 1800\beta^3 - 120z_i + 2700\beta^2 z_i - 360\beta z_i^2 + 900\beta^2 z_i^2 - 40z_i^3 + 900\beta^2 z_i^3 - 8z_i^5 + 180\beta^2 z_i^5}{(2 + 15\beta^2)^3}$$

$$W_c = \frac{8 - 60\beta^2 - 60\beta z_c + 4z_c^2 - 30\beta^2 z_c^2 - 20\beta z_c^3 - 4\beta z_c^5}{(2 + 15\beta^2)^2}$$

Los elementos de la matriz de información esperada se determinan a partir de la siguiente expresión:

$$i_{\theta_p \theta_q} = \frac{1}{n} E \left[- \frac{\partial^2}{\partial \theta_p \partial \theta_q} \ell_{LBSN} \right] \quad \text{para } p, q = 1, 2, 3. \quad (5.1.2)$$

Ahora considerando como en el Capitulo 3 a $d_k = E \left[\frac{Z_0^k (1 - \beta Z_0^3)^2}{(1 - \beta Z_0^3)^2 + 1} \right]$, para $k = 4, 5, 6$, con $Z_0 \sim N(0, 1)$ se tiene que la matriz de información esperada queda determinada por los siguientes elementos:

$$i_{\mu\mu} = \frac{1}{\sigma^2} \left[\frac{6\beta z_c^2 (1 - \beta z_c^3) \phi(z_c)}{(2 + 15\beta^2) \Phi_{BSN}(z_c)} + \frac{z_c \phi_{BSN}(z_c)}{\Phi_{BSN}(z_c)} + \frac{z_c \phi_{BSN}^2(z_c)}{\Phi_{BSN}^2(z_c)} \right] + \frac{36\beta^2 d_4 - 75\beta^2 + 2}{\sigma^2 (2 + 15\beta^2)}$$

$$i_{\mu\sigma} = \frac{1}{\sigma^2} \left[\frac{6\beta z_c^3 (1 - \beta z_c^3) \phi(z_c)}{(2 + 15\beta^2) \Phi_{BSN}(z_c)} + \frac{(z_c^2 - 1) \phi_{BSN}(z_c)}{\Phi_{BSN}(z_c)} + \frac{z_c \phi_{BSN}^2(z_c)}{\Phi_{BSN}^2(z_c)} \right] + \frac{36\beta^2 d_5 + 18\beta}{\sigma^2 (2 + 15\beta^2)}$$

$$i_{\mu\beta} = -\frac{1}{\sigma} \left[\frac{2z_c^3 (1 - \beta z_c^3) \phi(z_c)}{(2 + 15\beta^2) \Phi_{BSN}(z_c)} + \frac{30\beta \phi_{BSN}(z_c)}{(2 + 15\beta^2) \Phi_{BSN}(z_c)} + W_i \frac{\phi(z_c) \phi_{BSN}(z_c)}{\Phi_{BSN}^2(z_c)} \right] + \frac{-12\beta d_5 - 6}{\sigma (2 + 15\beta^2)}$$

$$i_{\sigma\sigma} = \frac{1}{\sigma} \left[\frac{6\beta z_c^4 (1 - \beta z_c^3) \phi(z_c)}{(2 + 15\beta^2) \Phi_{BSN}(z_i)} + \frac{(z_i^3 - z_i) \phi_{BSN}(z_c)}{\Phi_{BSN}(z_c)} + \frac{z_c^2 \phi_{BSN}^2(z_c)}{\Phi_{BSN}^2(z_c)} \right] + \frac{-645\beta^2 + 36\beta^2 d_6 + 1}{\sigma^2 (2 + 15\beta^2)}$$

$$i_{\sigma\beta} = -\frac{1}{\sigma} \left[\frac{2z_c^4 (1 - \beta z_c^3) \phi(z_c)}{(2 + 15\beta^2) \Phi_{BSN}(z_c)} + \frac{30\beta z_c \phi_{BSN}(z_c)}{(2 + 15\beta^2) \Phi_{BSN}(z_c)} + \frac{z_c W_c \phi(z_c) \phi_{BSN}^2(z_c)}{\Phi_{BSN}^2(z_c)} \right] + \frac{12\beta (15 - d_6)}{\sigma (2 + 15\beta^2)}$$

$$i_{\beta\beta} = -\phi(z_c) \left[\frac{U_c}{\Phi_{BSN}(z_c)} - \frac{W_c^2 \phi(z_c)}{\Phi_{BSN}^2(z_c)} \right] + \frac{60\beta^2 d_6 + 8d_6 - 900\beta^2}{(2 + 15\beta^2)^2}.$$

5.2. Distribución Log-Beta-Normal Asimétrica Alfa-Potencia con Datos Censurados

En esta sección se presenta el modelo log-beta-normal asimétrico alfa-potencia con datos censurados, el cual es una extensión del modelo CIBSN para $\alpha \neq 1$. Se define la fdp del modelo, se halla la FDA; también para el caso estándar del modelo, se da a conocer la expresión del momento de orden r , así como también se determinan las medidas: esperanza y varianza. Posteriormente se estudia la estimación de sus parámetros y se hallan las matrices de información.

Supongase que Y^* sigue una distribución LPBSN y que se tiene una muestra aleatoria $(Y_1^*, Y_2^*, \dots, Y_n^*)$, donde solamente son registrados los valores de Y^* que son mayores que una constante c . Además para valores $Y^* \leq c$ solo se registra el valor de c . Por lo tanto, para $i = 1, 2, \dots, n$, los valores observados son:

$$Y_i = \begin{cases} c & \text{si } Y_i^* \leq c, \\ Y_i^* & \text{si } Y_i^* > c \end{cases} \quad (5.2.1)$$

entonces la muestra resultante es una muestra censurada a la izquierda.

Sea F_{LPBSN} , la función de distribución de una variable aleatoria $Y^* \sim LPBSN(\mu, \sigma, \beta, \alpha)$, entonces:

$$\begin{aligned} F_{LPBSN}(Y^*; 0, 1, \beta, \alpha) &= \left[\Phi(\ln y^*) + \frac{\beta(2 \ln^2 y^* + 4 - \beta^5 y^* - 5\beta \ln^3 y^* - 15\beta \ln y^*)}{2 + 15\beta^2 \phi(\ln y^*)} \right]^\alpha \\ &= [\Phi_{LBSN}(y^*)]^\alpha \end{aligned}$$

con $y^* > 0$, $\alpha \in \mathbb{R}^+$ y $\beta \in \mathbb{R}$.

Ahora definiendo $z_c = \frac{\ln c - \mu}{\sigma}$, se tiene lo siguiente.

$$\begin{aligned} P(Y = c) &= P(Y^* \leq c) \\ &= F_{LPBSN}(Z_c) \\ &= \left[\Phi_{LBSN} \left(\frac{\ln c - \mu}{\sigma} \right) \right]^\alpha \\ &= \left[\Phi(z_c) + \frac{4\beta - 15\beta^2 z_c + 2\beta z_c^2 - 5\beta^2 z_c^3 - \beta^2 z_c^5}{2 + 15\beta^2} \phi(z_c) \right]^\alpha \end{aligned}$$

por lo tanto se tiene lo siguiente.

Definición 5.2.1. Si la variable aleatoria Y , tiene función de densidad,

$$f(y) = \begin{cases} \Phi_{LPBSN} \left(\frac{\ln c - \mu}{\sigma} \right) & \text{si } y \leq c, \\ f_{LPBSN}(y) & \text{si } y > c \end{cases} \quad (5.2.2)$$

donde f_{LPBSN} es la fdp de una variable aleatoria log-beta-normal asimétrica alfa-potencia para el caso localización escala y c es una constante, entonces se dice que Y es una variable aleatoria con distribución $l\beta$ -normal asimétrica alfa-potencia con datos censurados (a la izquierda). Por lo tanto denotamos $Y \sim CLPBSN(\theta)$ con $\theta = (\mu, \sigma, \beta, \alpha)'$ su función de densidad también puede ser expresada como

$$f(y) = f_{LPBSN}^i(y) \Phi_{LPBSN}^{1-i} \left(\frac{c - \mu}{\sigma} \right) \quad (5.2.3)$$

donde

$$I = \begin{cases} 0 & \text{si } y \leq c, \\ 1 & \text{si } y > c \end{cases}$$

En la Figura (5.2) se presenta la función de densidad para la variable aleatoria $Y \sim CLPBSN(\theta)$ con un punto de censura $c = 0.01$.

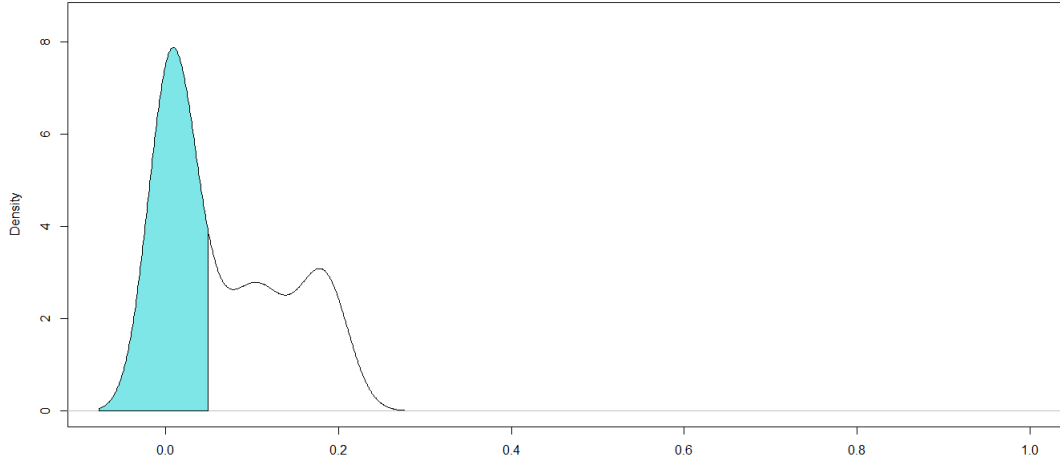


Figura 5.2: fdp del modelo $CLPBSN(\beta, \alpha)$.

Observe que la función está censurada a la izquierda (área de color) para $\beta = 0$ y $\alpha = 4$.

Proposición 5.0.1 (Momento de orden r del modelo $CLPBSN$).

Si $Y \sim CLPBSN(\mu, \sigma, \beta, \alpha)$, el momento r -ésimo está dado por:

$$E(Y^r) = c^r \Phi_{PBSN}(z_c) + \frac{\alpha e^{r\mu} M'_k}{2 + 15\beta^2} \quad (5.2.4)$$

donde

$$M'_k = 2w(z_c) - 2\beta\mu'_3(z_c) + \beta^2\mu'_6(z_c)$$

con $w = \int_{z_c}^{\infty} e^{r\sigma z} \phi(z) \left[\Phi_{BSN}(z) \right]^{\alpha-1} dz$

y $\mu'_q(z_c) = \int_{z_c}^{\infty} z^q e^{r\sigma z} \phi(z) \left[\Phi_{BSN}(z) \right]^{\alpha-1} dz$

Demostración. Por definición de valor esperado condicionado, se tiene que:

$$\begin{aligned}
E(Y^r|y > c) &= \int_c^\infty y^r f(y|y > c) dy \\
&= \int_c^\infty y^r \frac{f_{LPBSN}(y)}{P(Y > c)} dy \\
&= \int_c^\infty y^r \frac{\alpha f_{LBSN}(y) \left[\Phi_{LBSN}(y) \right]^{\alpha-1}}{P(Y > c)} dy
\end{aligned}$$

Ahora tomando $z = \frac{\ln y - \mu}{\sigma}$ y redefiniendo el límite inferior

de la integral como $z_c = \frac{\ln c - \mu}{\sigma}$ se tiene que:

$$\begin{aligned}
E(Y^r|y > c) &= \frac{\int_{z_c}^\infty \alpha \left(e^{(\sigma z + \mu)} \right)^r \left[(1 - \beta z^3)^2 + 1 \right] \phi(z) \left[\Phi_{BSN}(z) \right]^{\alpha-1}}{P\left(z > \frac{\ln c - \mu}{\sigma} \right) (2 + 15\beta^2)} dz \\
&= \frac{\int_{z_c}^\infty \alpha \left(e^{r(\sigma z + \mu)} \right)^r \left[(1 - \beta z^3)^2 + 1 \right] \phi(z) \left[\Phi_{BSN}(z) \right]^{\alpha-1}}{\left[1 - \Phi_{BSN}(z_c) \right] (2 + 15\beta^2)} dz
\end{aligned}$$

$$\text{ya que } P(Z > z) = 1 - P(Z \leq z_c) = 1 - P\left(z \leq \frac{\ln c - \mu}{\sigma} \right)$$

$$= 1 - \Phi_{BSN}(z_c), \text{ entonces,}$$

$$= \frac{\alpha \int_{z_c}^\infty e^{r\sigma z} e^{r\mu} (1 - 2\beta z^3 + \beta^2 z^6 + 1) \phi(z) \left[\Phi_{BSN}(z) \right]^{\alpha-1}}{\left[1 - \Phi_{BSN}(z_c) \right] (2 + 15\beta^2)} dz$$

$$= \frac{\alpha e^{r\mu} \int_{z_c}^\infty e^{r\sigma z} (2 - 2\beta z^3 + \beta^2 z^6) \phi(z) \left[\Phi_{BSN}(z) \right]^{\alpha-1}}{\left[1 - \Phi_{BSN}(z_c) \right] (2 + 15\beta^2)} dz$$

$$= \frac{\alpha e^{r\mu} \left\{ 2 \int_{z_c}^\infty e^{r\sigma z} \phi(z) \left[\Phi_{BSN}(z) \right]^{\alpha-1} dz - 2\beta \int_{z_c}^\infty z^3 e^{r\sigma z} \phi(z) \left[\Phi_{BSN}(z) \right]^{\alpha-1} dz + \beta^2 \int_{z_c}^\infty z^6 e^{r\sigma z} \phi(z) \left[\Phi_{BSN}(z) \right]^{\alpha-1} dz \right\}}{\left[1 - \Phi_{BSN}(z_c) \right] (2 + 15\beta^2)}$$

$$= \frac{\alpha e^{r\mu} \left[2w(z_c) - 2\beta \mu'_3(z_c) + \beta^2 \mu'_6(z_c) \right]}{\left[1 - \Phi_{BSN}(z_c) \right] (2 + 15\beta^2)}$$

$$= \frac{\alpha e^{r\mu} M'_k}{\left[1 - \Phi_{BSN}(z_c) \right] (2 + 15\beta^2)}$$

$$w = \int_{z_c}^{\infty} e^{r\sigma z} \phi(z) \left[\Phi_{BSN}(z) \right]^{\alpha-1} dz$$

$$y \quad \mu'_q(z_c) = \int_{z_c}^{\infty} z^q e^{r\sigma z} \phi(z) \left[\Phi_{BSN}(z) \right]^{\alpha-1} dz$$

Ahora usando la definición de la esperanza para una variable aleatoria limitada, el valor esperado de orden r para $Y \sim CLPBSN(\boldsymbol{\theta})$ es:

$$\begin{aligned} E(Y^r) &= E(Y^r | y = c)P(y = c) + E(Y^r | y > c)P(y > c) \\ &= c^r \Phi_{PBSN}\left(\frac{\ln c - \mu}{\sigma}\right) + \frac{\alpha e^{r\mu} M'_k}{\left[1 - \Phi_{BSN}(z_c)\right] (2 + 15\beta^2)} \left[1 - \Phi_{BSN}(z_c)\right] \\ &= c^r \Phi_{PBSN}\left(\frac{\ln c - \mu}{\sigma}\right) + \frac{\alpha e^{r\mu} M'_k}{(2 + 15\beta^2)} \\ &= c^r \Phi_{PBSN}(z_c) + \frac{\alpha e^{r\mu} M'_k}{(2 + 15\beta^2)} \end{aligned}$$

□

Proposición 5.0.2 (Esperanza y varianza del modelo CLPBSN estándar).

Sea $Y \sim CLPBSN(\theta)$ donde $\theta = (\mu, \sigma, \beta, \alpha)'$, entonces su esperanza y su varianza están dadas por:

$$(i) \quad E(Y) = c \Phi_{PBSN}(z_c) + \frac{\alpha e^{\mu} \mu'_1}{2 + 15\beta^2}$$

$$(ii) \quad V(Y) = c \Phi_{PBSN}(z_c) \left[c(1 - \Phi_{PBSN}(z_c)) - \frac{2\alpha e^{\mu} \mu'_1}{2 + 15\beta^2} \right] + \frac{\alpha e^{\mu} \mu'_1}{2 + 15\beta^2} - \left(\frac{\alpha e^{\mu} \mu'_1}{2 + 15\beta^2} \right)^2.$$

Demostración.

(i) Para $k = 1$ y $r = 1$ y reemplazando en la expresión (5.2.4), se tiene:

$$\begin{aligned} E(Y^1) &= c^1 \Phi_{PBSN}(z_c) + \frac{\alpha e^{1\mu} \mu'_1}{2 + 15\beta^2} \\ &= c \Phi_{PBSN}(z_c) + \frac{\alpha e^{\mu} \mu'_1}{2 + 15\beta^2} \end{aligned}$$

(ii) Tomando $k = 2$ y $r = 2$ en la expresión (5.2.4), se obtiene el valor esperado de orden 2

$$E(Y^2) = c^2 \Phi_{PBSN}(z_c) + \frac{\alpha e^{2\mu} \mu'_2}{2 + 15\beta^2}$$

entonces la varianza viene dada por:

$$\begin{aligned} V(Y) &= E(Y^2) - E^2(Y) \\ &= c^2 \Phi_{PBSN}(z_c) + \frac{\alpha e^{2\mu} \mu'_2}{2 + 15\beta^2} - \left[c \Phi_{PBSN}(z_c) + \frac{\alpha e^\mu \mu'_1}{2 + 15\beta^2} \right]^2 \\ &= c^2 \Phi_{PBSN}(z_c) + \frac{\alpha e^{2\mu} \mu'_2}{2 + 15\beta^2} - c^2 \Phi_{PBSN}^2(z_c) - 2 \frac{c \alpha e^\mu \Phi_{PBSN}(z_c) \mu'_1}{2 + 15\beta^2} \\ &\quad - \left(\frac{\alpha e^\mu \mu'_1}{2 + 15\beta^2} \right)^2 \\ &= c \Phi_{PBSN}(z_c) \left[c(1 - \Phi_{PBSN}(z_c)) - \frac{2\alpha e^\mu \mu'_1}{2 + 15\beta^2} \right] + \frac{\alpha e^\mu \mu'_1}{2 + 15\beta^2} \\ &\quad - \left(\frac{\alpha e^\mu \mu'_1}{2 + 15\beta^2} \right)^2. \end{aligned}$$

□

5.3. Inferencia para el Modelo CLPBSN

Sea $Y = (Y_1, Y_1, \dots, Y_n)$ una muestra aleatoria con $Y_i \sim CLPBSN(\boldsymbol{\theta})$ con n_0 observaciones censuradas y n_1 observaciones no censuradas. La función de verosimilitud, para $\boldsymbol{\theta} = (\mu, \sigma, \beta, \alpha)'$ es:

$$\begin{aligned} \mathbf{L}_{CLPBSN}(\boldsymbol{\theta}) &= \mathbf{L}(\boldsymbol{\theta}; Y) = \prod_{y_i > c} f_{LPBSN}(y_i) \prod_{y_i \leq c} \Phi_{IPBSN}(y_i) \left(\frac{\ln c - \mu}{\sigma} \right)^\alpha \\ &= \prod_{y_i > c} \alpha^{n_1} f_{LBSN}(y_i) [F_{LBSN}(y_i)]^{\alpha-1} \prod_{y_i \leq c} \left[\Phi_{LBSN}(y_i) \left(\frac{\ln c - \mu}{\sigma} \right) \right]^\alpha \end{aligned}$$

Tomando $z_i = \frac{\ln y_i - \mu}{\sigma}$ y $z_c = \frac{\ln c - \mu}{\sigma}$ la log-verosimilitud queda expresada como:

$$\begin{aligned}
\ell_{CLPBSN}(\theta) &= \ell(\theta; Z) \\
&= n_1 \ln(\alpha) + \sum_{z_i > z_c} \ln[f_{BSN}(z_i)] + (\alpha - 1) \sum_{z_i > z_c} \ln(\Phi_{BSN}(z_i)) \\
&\quad + \sum_{z_i \leq z_c} \ln \left[\Phi_{BSN} \left(\frac{\ln c - \mu}{\sigma} \right) \right]^\alpha \\
&= n_0 \alpha \ln \left[\Phi_{BSN} \left(\frac{\ln c - \mu}{\sigma} \right) \right] + (\alpha - 1) \sum_{z_i > z_c} \ln(\Phi_{BSN}(z_i)) \\
&\quad + \ell_{BSN} + n_1 \ln(\alpha) \\
&= n_0 \alpha \ln \left[\Phi(z_c) + \frac{4\beta - 15\beta^2 z_c + 2\beta z_c^2 - 5\beta^2 z_c^3 - \beta^2 z_c^5}{2 + 15\beta^2} \phi(z_c) \right] \\
&\quad + (\alpha - 1) \sum_{z_i > z_c} \left[\phi(z_i) + \frac{4\beta - 15\beta^2 z_i + 2\beta z_i^2 - 5\beta^2 z_i^3 - \beta^2 z_i^5}{2 + 15\beta^2} \phi(z_i) \right] \\
&\quad + \sum_{z_i > z_c} \ln \left[(1 - \beta z_i^3)^2 + 1 \right] - \frac{1}{2} \sum_{z_i > z_c} z_i^2 - n_1 \left[\ln(\sigma) + \ln(2 + 15\beta^2) \right] \\
&\quad + \ln \left(\sqrt{\frac{2}{\pi}} \right) - \ln(\alpha) \Big].
\end{aligned}$$

Por lo tanto, las primeras derivadas parciales de la función de log-verosimilitud están dadas por:

$$\begin{aligned}
\frac{\partial}{\partial \mu} \ell_{CLPBSN} &= -\frac{n_0 \alpha}{\sigma} \frac{\phi_{BSN}(z_c)}{\Phi_{BSN}(z_c)} - \frac{\alpha - 1}{\sigma} \sum_{z_i > z_c} \frac{\phi_{BSN}(z_i)}{\Phi_{BSN}(z_i)} + \frac{1}{\sigma} \sum_{z_i > z_c} \left[\frac{6\beta(1 - \beta z_i^3)^2}{(1 - \beta z_i^3)^2 + 1} + z_i \right] \\
\frac{\partial}{\partial \sigma} \ell_{CLPBSN} &= -\frac{n_0 \alpha}{\sigma} \frac{\phi_{BSN}(z_c)}{\Phi_{BSN}(z_c)} - \frac{\alpha - 1}{\sigma} \sum_{z_i > z_c} \frac{z_i \phi_{BSN}(z_i)}{\Phi_{BSN}(z_i)} + \frac{1}{\sigma} \sum_{z_i > z_c} \left[\frac{6\beta(1 - \beta z_i^3) z_i^3}{(1 - \beta z_i^3)^2 + 1} - 1 + z_i^2 \right] \\
\frac{\partial}{\partial \beta} \ell_{CLPBSN} &= n_0 \alpha \frac{W_c \phi(z_c)}{\Phi_{BSN}(z_c)} + (\alpha - 1) \sum_{z_i > z_c} \frac{W_i \phi(z_i)}{\Phi_{BSN}(z_i)} - 2 \sum_{z_i > z_c} \left[\frac{z_i^3(1 - \beta z_i^3)}{(1 - \beta z_i^3)^2 + 1} + \frac{15\beta}{2 + 15\beta^2} \right] \\
\frac{\partial}{\partial \alpha} \ell_{CLPBSN} &= n_0 \ln[\phi_{BSN}(z_c)] + \sum_{z_i > z_c} \ln[\phi_{BSN}(z_i)] + \frac{n_1}{\alpha}
\end{aligned}$$

Ahora los elementos de la matriz de información observada son

$$\begin{aligned}
j_{\mu\mu} &= \frac{(n_0\alpha)}{\sigma^2} \left[\frac{6\beta z_c^2(1-\beta z_c^3)\phi(z_c)}{(2+15\beta^2)\Phi_{BSN}(z_c)} + \frac{z_c\phi_{BSN}(z_c)}{\Phi_{BSN}(z_c)} + \frac{\phi_{BSN}^2(z_c)}{\Phi_{BSN}^2(z_c)} \right] \\
&\quad + \frac{(\alpha-1)}{\sigma^2} \sum_{z_i > z_c} \left[\frac{6\beta z_i^2(1-\beta z_i^3)\phi(z_i)}{(2+15\beta^2)\Phi_{BSN}(z_i)} + \frac{z_i\phi_{BSN}(z_i)}{\Phi_{BSN}(z_i)} + \frac{\phi_{BSN}^2(z_i)}{\Phi_{BSN}^2(z_i)} \right] \\
&\quad - \frac{1}{\sigma^2} \sum_{z_i > z_c} \left[\frac{6\beta z_i(5\beta z_i^3-2)}{(1-\beta z_i^3)^2+1} - \frac{36\beta^2 z_i^4(1-\beta z_i^3)^2}{[(1-\beta z_i^3)^2+1]^2} - 1 \right] \\
j_{\mu\sigma} &= \frac{(n_0\alpha)}{\sigma^2} \left[\frac{6\beta z_c^3(1-\beta z_c^3)\phi(z_c)}{(2+15\beta^2)\Phi_{BSN}(z_c)} + \frac{(z_c^2-1)\phi_{BSN}(z_c)}{\Phi_{BSN}(z_c)} + \frac{z_c\phi_{BSN}^2(z_c)}{\Phi_{BSN}^2(z_c)} \right] \\
&\quad + \frac{(\alpha-1)}{\sigma^2} \sum_{z_i > z_c} \left[\frac{6\beta z_i^3(1-\beta z_i^3)\phi(z_i)}{(2+15\beta^2)\Phi_{BSN}(z_i)} + \frac{(z_i^2-1)\phi_{BSN}(z_i)}{\Phi_{BSN}(z_i)} + \frac{z_i\phi_{BSN}^2(z_i)}{\Phi_{BSN}^2(z_i)} \right] \\
&\quad - \frac{2}{\sigma^2} \sum_{z_i > z_c} \left[\frac{9\beta z_i^2(2\beta z_i^3-1)}{(1-\beta z_i^3)^2+1} - \frac{18\beta^2 z_i^5(1-\beta z_i^3)^2}{[(1-\beta z_i^3)^2+1]^2} - z_i \right] \\
j_{\mu\beta} &= -\frac{(n_0\alpha)}{\sigma} \left[\frac{2z_c^3(1-\beta z_c^3)\phi(z_c)}{(2+15\beta^2)\Phi_{BSN}(z_c)} + \frac{30\beta\phi_{BSN}(z_c)}{(2+15\beta^2)\Phi_{BSN}(z_c)} + W_i \frac{\phi(z_c)\phi_{BSN}(z_c)}{\Phi_{BSN}^2(z_c)} \right] \\
&\quad - \frac{(\alpha-1)}{\sigma} \sum_{z_i > z_c} \left[\frac{2z_i^3(1-\beta z_i^3)\phi(z_i)}{(2+15\beta^2)\Phi_{BSN}(z_i)} + \frac{30\beta\phi_{BSN}(z_i)}{(2+15\beta^2)\Phi_{BSN}(z_i)} + W_i \frac{\phi(z_i)\phi_{BSN}(z_i)}{\Phi_{BSN}^2(z_i)} \right] \\
&\quad - \frac{6}{\sigma^2} \sum_{z_i > z_c} \left[\frac{z_i^2(1-2\beta z_i^3)}{(1-\beta z_i^3)^2+1} + \frac{2\beta z_i^5(1-\beta z_i^3)^2}{[(1-\beta z_i^3)^2+1]^2} \right] \\
j_{\mu\alpha} &= \frac{n_0}{\sigma} \frac{\phi_{BSN}(z_c)}{\Phi_{BSN}(z_c)} + \frac{1}{\sigma} \sum_{z_i > z_c} \frac{\phi_{BSN}(z_i)}{\Phi_{BSN}(z_i)} \\
j_{\sigma\sigma} &= \frac{(n_0\alpha)}{\sigma} \left[\frac{6\beta z_c^4(1-\beta z_c^3)\phi(z_c)}{(2+15\beta^2)\Phi_{BSN}(z_c)} + \frac{(z_c^3-z_c)\phi_{BSN}(z_c)}{\Phi_{BSN}(z_c)} + \frac{z_c^2\phi_{BSN}^2(z_c)}{\Phi_{BSN}^2(z_c)} \right] \\
&\quad + \frac{(\alpha-1)}{\sigma} \sum_{z_i > z_c} \left[\frac{6\beta z_i^4(1-\beta z_i^3)\phi(z_i)}{(2+15\beta^2)\Phi_{BSN}(z_i)} + \frac{(z_i^3-z_i)\phi_{BSN}(z_i)}{\Phi_{BSN}(z_i)} + \frac{z_i^2\phi_{BSN}^2(z_i)}{\Phi_{BSN}^2(z_i)} \right] \\
&\quad - \frac{1}{\sigma^2} \sum_{z_i > z_c} \left[\frac{2\beta z_i^3(21\beta z_i^3-12)}{(1-\beta z_i^3)^2+1} + \frac{36\beta^2 z_i^6(1-\beta z_i^3)^2}{[(1-\beta z_i^3)^2+1]^2} + 1 - 3z_i^2 \right] \\
j_{\sigma\beta} &= -\frac{n_0\alpha}{\sigma} \left[\frac{2z_c^4(1-\beta z_c^3)\phi(z_c)}{(2+15\beta^2)\Phi_{BSN}(z_c)} + \frac{30\beta z_c\phi_{BSN}(z_c)}{(2+15\beta^2)\Phi_{BSN}(z_c)} + \frac{z_c W_c \phi(z_c)\phi_{BSN}^2(z_c)}{\Phi_{BSN}^2(z_c)} \right] \\
&\quad + \frac{(\alpha-1)}{\sigma} \sum_{z_i > z_c} \left[\frac{2z_i^4(1-\beta z_i^3)\phi(z_i)}{(2+15\beta^2)\Phi_{BSN}(z_i)} + \frac{30\beta z_i\phi_{BSN}(z_i)}{(2+15\beta^2)\Phi_{BSN}(z_i)} + \frac{z_i W_i \phi(z_i)\phi_{BSN}^2(z_i)}{\Phi_{BSN}^2(z_i)} \right] \\
&\quad - \frac{6}{\sigma} \sum_{z_i > z_c} \left[\frac{z_i^3(1-2\beta z_i^3)}{(1-\beta z_i^3)^2+1} + \frac{2\beta z_i^6(1-\beta z_i^3)^2}{[(1-\beta z_i^3)^2+1]^2} \right] \\
j_{\sigma\alpha} &= \frac{n_0}{\sigma} \frac{z_c\phi_{BSN}(z_c)}{\Phi_{BSN}(z_c)} + \frac{1}{\sigma} \sum_{z_i > z_c} \frac{z_i\phi_{BSN}(z_i)}{\Phi_{BSN}(z_i)}
\end{aligned}$$

$$\begin{aligned}
j_{\beta\beta} &= -n_0\phi(z_c) \left[\frac{U_c}{\Phi_{BSN}(z_c)} - \frac{W_c^2\phi(z_c)}{\Phi_{BSN}^2(z_c)} \right] \\
&\quad - (\alpha - 1) \sum_{z_i > z_c} \phi(z_i) \left[\frac{U_i}{\Phi_{BSN}(z_i)} - \frac{W_i^2\phi(z_i)}{\Phi_{BSN}^2(z_i)} \right] \\
&\quad - 2 \sum_{z_i > z_c} \left[\frac{z_i^6}{(1 - \beta z_i^3)^2 + 1} - \frac{2z_i^6(1 - \beta z_i^3)^2}{[(1 - \beta z_i^3)^2 + 1]^2} - \frac{15(2 - 15\beta^2)}{(2 + 15\beta^2)^2} \right] \\
j_{\beta\alpha} &= -n_0 \frac{W_c\phi(z_c)}{\Phi_{BSN}(z_c)} - \sum_{z_i > z_c} \frac{W_i\phi(z_i)}{\Phi_{BSN}(z_i)} \\
j_{\alpha\alpha} &= \frac{n_1}{\sigma^2}
\end{aligned}$$

donde se tiene:

$$\begin{aligned}
U_c &= \frac{-720\beta + 1800\beta^3 - 120z_i + 2700\beta^2 z_i - 360\beta z_i^2 + 900\beta^2 z_i^2 - 40z_i^3 + 900\beta^2 z_i^3 - 8z_i^5}{(2 + 15\beta^2)^3} \\
W_c &= \frac{8 - 60\beta^2 - 60\beta z_c + 4z_c^2 - 30\beta^2 z_c^2 - 20\beta z_c^3 - 4\beta z_c^5}{(2 + 15\beta^2)^2}
\end{aligned}$$

Los elementos de la matriz de información esperada son:

$$i_{\theta_p\theta_q} = \frac{1}{n} E \left[-\frac{\partial^2}{\partial\theta_p\partial\theta_q} \ell_{CLPBSN} \right], \quad p, q = 1, 2, 3, 4 \quad (5.3.1)$$

con $\theta_1 = \mu$, $\theta_2 = \sigma$, $\theta_3 = \beta$ y $\theta_4 = \alpha$ y seran denotados como $i_{\mu\mu}, i_{\mu\sigma}, \dots, i_{\alpha\alpha}$

Definiendo:

$$\begin{aligned}
h_j &= E \left[z^j \frac{\phi(z)}{\Phi_{BSN}(z)} \right], & v_{jk} &= E \left[z^j \left(\frac{\phi_{BSN}(z)}{\Phi_{BSN}(z)} \right)^k \right] \\
g_{jk} &= E \left[\frac{z^j}{[(1 - \beta z^3)^2 + 1]^k} \right], & u_{jk} &= E \left[\frac{z^j W^k \phi(z) \phi_{BSN}(z)}{\Phi_{BSN}^2(z)} \right] \\
a &= E \left[\frac{U\phi(z)}{\Phi_{BSN}(z)} \right], \quad y & b_k &= E \left[\frac{W^k \phi^k(z)}{\Phi_{BSN}^k(z)} \right]
\end{aligned}$$

con $z \sim CLPBSN(\beta, \alpha)$, así los elementos de la matriz esperada quedan dadas

por:

$$\begin{aligned}
i_{\mu,\mu} &= \frac{\alpha}{\sigma^2} \left[\frac{6\beta z_c^2(1-\beta z_c^3)\phi(z_c)}{(2+15\beta^2)\Phi_{BSN}(z_c)} + \frac{z_c\phi_{BSN}(z_c)}{\Phi_{BSN}(z_c)} + \frac{z_c\phi_{BSN}^2(z_c)}{\Phi_{BSN}^2(z_c)} \right] \\
&\quad + \frac{\alpha-1}{\alpha^2} \left[\frac{6\beta}{2+15\beta^2}(h_2-\beta h_5) + v_{1,1} + v_{0,2} \right] \\
&\quad - \frac{1}{\sigma^2} [6\beta(5\beta g_{4,1} - 2g_{1,1}) - 36\beta^2(g_{4,2} - 2\beta g_{7,2} + \beta^2 g_{10,2}) - 1] \\
i_{\mu,\sigma} &= \frac{\alpha}{\sigma^2} \left[\frac{6\beta z_c^2(1-\beta z_c^3)\phi(z_c)}{(2+15\beta^2)\Phi_{BSN}(z_c)} + \frac{(z_c^2-1)\phi_{BSN}(z_c)}{\Phi_{BSN}(z_c)} + \frac{z_c\phi_{BSN}^2(z_c)}{\Phi_{BSN}^2(z_c)} \right] \\
&\quad + \frac{\alpha-1}{\alpha^2} \left[\frac{6\beta}{2+15\beta^2}(h_3-\beta h_6) + v_{2,1} - v_{0,2} + v_{1,2} \right] \\
&\quad - \frac{2}{\sigma^2} [9\beta(2\beta g_{5,1} - g_{2,1}) - 18\beta^2(g_{5,2} - 2\beta g_{8,2} + \beta^2 g_{11,2}) - g_{1,2}] \\
i_{\mu,\beta} &= -\frac{(\alpha)}{\sigma} \left[\frac{2z_c^3(1-\beta z_c^3)\phi(z_c)}{(2+15\beta^2)\Phi_{BSN}(z_c)} + \frac{30\beta\phi_{BSN}(z_c)}{(2+15\beta^2)\Phi_{BSN}(z_c)} + W_i \frac{\phi(z_c)\phi_{BSN}(z_c)}{\Phi_{BSN}^2(z_c)} \right] \\
&\quad - \frac{\alpha-1}{\alpha} \left[\frac{2}{2+15\beta^2}(h_3-\beta h_6) + 30\beta v_{0,1} + u_{0,1} \right] \\
&\quad - \frac{6}{\sigma} [(g_{2,1} - 2\beta g_{3,1}) + 2\beta(g_{5,2} - 2\beta g_{8,2} + \beta^2 g_{11,2})] \\
i_{\mu,\alpha} &= \frac{1}{\sigma} \frac{\phi_{BSN}^2(z_c)}{\Phi_{BSN}^2(z_c)} + \frac{1}{\sigma} v_{0,1} \\
i_{\sigma,\sigma} &= \frac{(\alpha)}{\sigma} \left[\frac{6\beta z_c^4(1-\beta z_c^3)\phi(z_c)}{(2+15\beta^2)\Phi_{BSN}(z_i)} + \frac{(z_i^3-z_i)\phi_{BSN}(z_c)}{\Phi_{BSN}(z_c)} + \frac{z_c^2\phi_{BSN}^2(z_c)}{\Phi_{BSN}^2(z_c)} \right] \\
&\quad + \frac{\alpha-1}{\sigma} \left[\frac{6\beta}{2+15\beta^2}(h_4-\beta h_7) + v_{3,1} - v_{1,1} + v_{2,2} \right] \\
&\quad - \frac{1}{\sigma^2} [2\beta(21\beta g_{6,1} - 12g_{3,1}) - 36\beta^2(g_{6,2} - 2\beta g_{9,2} + \beta^2 g_{12,2}) + 1 - g_{2,0}] \\
i_{\sigma,\beta} &= -\frac{\alpha}{\sigma} \left[\frac{2z_c^4(1-\beta z_c^3)\phi(z_c)}{(2+15\beta^2)\Phi_{BSN}(z_c)} + \frac{30\beta z_c\phi_{BSN}(z_c)}{(2+15\beta^2)\Phi_{BSN}(z_c)} + \frac{z_c W_c \phi(z_c)\phi_{BSN}^2(z_c)}{\Phi_{BSN}^2(z_c)} \right] \\
&\quad - \frac{\alpha-1}{\sigma} \left[\frac{2}{2+15\beta^2}(h_4-\beta h_7) + \frac{30\beta}{2+15\beta^2} v_{1,1} + u_{1,1} \right] \\
&\quad - \frac{6}{\sigma} [(g_{3,1} - 2\beta g_{6,1}) + 2\beta(g_{6,2} - 2\beta g_{9,2} + \beta^2 g_{12,2})] \\
i_{\sigma,\alpha} &= \frac{1}{\sigma} \frac{z_c\phi_{BSN}(z_c)}{\Phi_{BSN}(z_c)} + \frac{1}{\sigma} v_{1,1} \\
i_{\beta,\beta} &= -\alpha\phi(z_c) \left[\frac{U_c}{\Phi_{BSN}(z_c)} - \frac{W_c^2\phi(z_c)}{\Phi_{BSN}^2(z_c)} \right] \\
&\quad - (\alpha-1)[a-b_2] - 2 \left[g_{6,1} - 2(g_{6,1} - 2\beta g_{4,2} + \beta^2 g_{12,2}) - \frac{15(2-15\beta^2)}{(2+15\beta^2)^2} \right]
\end{aligned}$$

$$i_{\beta,\alpha} = - \frac{W_c \phi(z_c)}{\Phi_{BSN}(z_c)} - b_1$$

$$i_{\alpha,\alpha} = \frac{1}{\sigma^2}$$

Capítulo 6

Ilustraciones

En este capítulo se presentan 2 ilustraciones con datos reales de las distribuciones propuestas en los capítulos anteriores, y se comparan los ajustes con otros modelos propuestos en la literatura.

6.1. Ilustracion 1.

La erosión del acero de refuerzo es un problema serio en estructuras de concreto localizados en ambientes afectados por condiciones climáticas severas. Por esta razón los investigadores han estado estudiando el uso de barras de refuerzo hechas de un material compuesto.

En Ehsani *et al.*(1996) se muestran las 48 observaciones de un estudio donde se mide la fuerza adhesiva para adherir barras de refuerzo reforzadas con fibra de vidrio a concreto.

La Tabla 6.1 muestra algunas estadísticas descriptivas para los datos.

Tabla 6.1: Resumen de estadísticas descriptivas.

n	\bar{X}	S^2	<i>Median</i>
48	8.079	23.702	5.950

Los modelos LN, Birnbaum Saunders bimodal (BSB) (ver Olmos et al.(2016), LBSN y LPBSN son ajustados para el conjunto de datos.

Para comparar el ajuste de los modelos se utilizaron los criterios de Akaike(1974) o AIC y el criterio Bayesiano o BIC, definidos por

$$AIC = -2\hat{\ell}(\cdot) + 2k,$$

y

$$BIC = -2\hat{\ell}(\cdot) + k \ln(n),$$

donde k es el número de parámetros en cada modelo. De acuerdo con estos criterios, el modelo con mejor ajuste es aquel con menos AIC o BIC. De acuerdo a los criterios AIC y BIC el modelo con mejor ajuste es el LPBSN seguido en ese orden por el modelo BSB y el modelo LBSN. Las estimativas de los parámetros fueron calculadas maximizando numéricamente la función de log-verosimilitud, esta maximización se realiza usando la función `optim`, disponible en el software estadístico R-Project.

Tabla 6.2: Estimativas de máxima verosimilitud para los modelos ajustados censurados

estimativas	$LN(SD)$	$BSB(SD)$	$LBSN(SD)$	$LPBSN(SD)$
$\hat{\mu}$	1.940(0.076)	0.317(0.050)	2.077(0.045)	1.103(0.169)
$\hat{\sigma}$	0.528(0.053)	7.380(0.330)	0.252(0.016)	0.469(0.056)
$\hat{\beta}$		-1.307(0.372)	0.441(0.083)	0.216(0.046)
$\hat{\alpha}$				7.893(3.141)
AIC	265.3	260.0	263.6	258.2
BIC	269.0	265.6	272.2	265.6

Usando los resultados de la Tabla 6.2, podemos realizar una prueba de hipótesis

del modelo LPBSN contra el modelo LBSN(SD), es decir,

$$H_0 : \alpha = 0 \quad \text{versus} \quad H_1 : \alpha \neq 0$$

usando la estadística de razón de verosimilitud basada en

$$\Lambda = L_{LBSN}(\hat{\mu}, \hat{\sigma}, \hat{\beta}) / L_{LPBSN}(\hat{\mu}, \hat{\sigma}, \hat{\beta}, \hat{\alpha}).$$

Reemplazando los valores de las estimativas en la razón anterior, nosotros obtenemos $-2 \ln(\Lambda) = -2(125.13 - 128.83) = 7.4$, el cual es superior al valor del percentil del 95 % de la distribución χ^2 -cuadrado, dado por, $\chi_1^2 = 3.84$, llevando al rechazo de la hipótesis nula, lo cual indica claramente que el modelo LPBSN(β, α) presenta un mejor ajuste que el modelo LBSN(β).

La figura 6.1-(a), muestra que el modelo LPBSN(SD) presenta el mejor ajuste comparado con el resto de modelos ajustados mientras que el gráfico de la figura 6.1-(b) muestra la FDA de los modelos LBSN y LPBSN note que, estos presentan un buen ajuste.

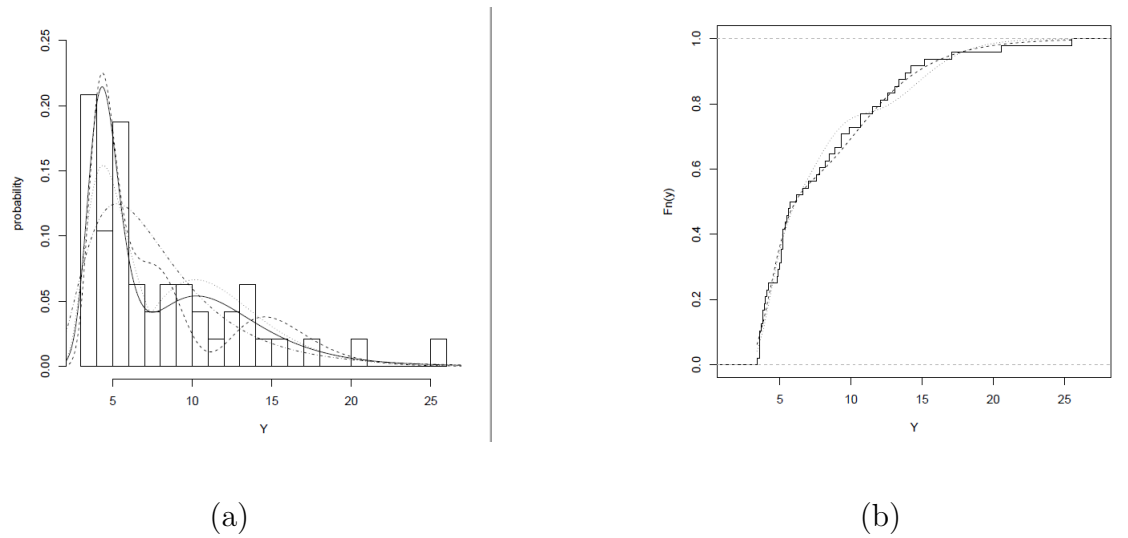


Figura 6.1: (a) Histograma para las 48 observaciones del estudio. Las líneas representan las distribuciones ajustadas usando las estimativas de máxima verosimilitud: LN (línea discontinua con puntos), BSB (línea de puntos), $LBSN$ (línea discontinua) y $LPBSN$ (línea continua). También se gráfica de la FDA para la parte (b) donde: Empírica (línea continua), $LBSN$ (línea de puntos) y $LPBSN$ (línea discontinua).

6.2. Ilustracion 2.

Una ilustración práctica del uso del modelo $CLPBSN(\beta, \alpha)$ se lleva a cabo ahora con datos de un gran estudio de seguridad e inmunogenicidad de vacunas contra el sarampión realizado en Haití durante (1987-1990), ver Moulton *et al.* (2007). El estudio tuvo como objetivo demostrar que altas dosis de la vacuna podrán inmunizar eficazmente a bebés de hasta 6 meses de edad. La inmunogenicidad estaba indicada por las respuestas de anticuerpos, mucho más altas entre los receptores con altas dosis de la vacuna. Se realizaron ensayos de anticuerpos de neutralización en sueros a 330 niños de hasta 12 meses de edad. El límite de

detección fue 0.1 unidades internacionales (UI) o $\ln(0.1) = -2.306$ en la escala del logaritmo natural. Se encontró que 86 (26.1 %) de los niños cayeron por debajo del límite de detección y todos se registraron como 0.1. En este caso se ajustaron los modelos censurados Tobit (CN), LN, LBSN y LPBSN. Las estimativas de máxima verosimilitud son dadas en la Tabla 6.3. De acuerdo a las medidas de comparación AIC y BIC el modelo que presenta el mejor ajuste es el modelo CLPBSN, seguido en su orden de los modelos CLBSN y LN.

Tabla 6.3: Parámetros estimados para los modelos ajustados con datos censurados

estimativas	<i>CN</i>	<i>CLN(SD)</i>	<i>CLBSN(SD)</i>	<i>CLPBSN(SD)</i>
$\hat{\mu}$	1.202(0.116)	-0.736(0.071)	-1.125(0.101)	-1.808(0.242)
$\hat{\sigma}$	2.105(0.082)	1.297(0.050)	1.137(0.054)	1.162(0.063)
$\hat{\beta}$			-0.100(0.022)	-0.134(0.022)
$\hat{\alpha}$				1.729(0.350)
<i>AIC</i>	1423.1	623.0	607.3	597.9
<i>BIC</i>	1430.7	630.5	618.7	613.1

La figura 6.2-(a), muestra que el modelo *LPBSN(SD)* presenta el mejor ajuste comparado con el resto de modelos ajustados mientras que el gráfico de la figura 6.1-(b) muestra la FDA de los modelos *LBSN* y *LPBSN* donde se puede observar que estos presentan un buen ajuste.

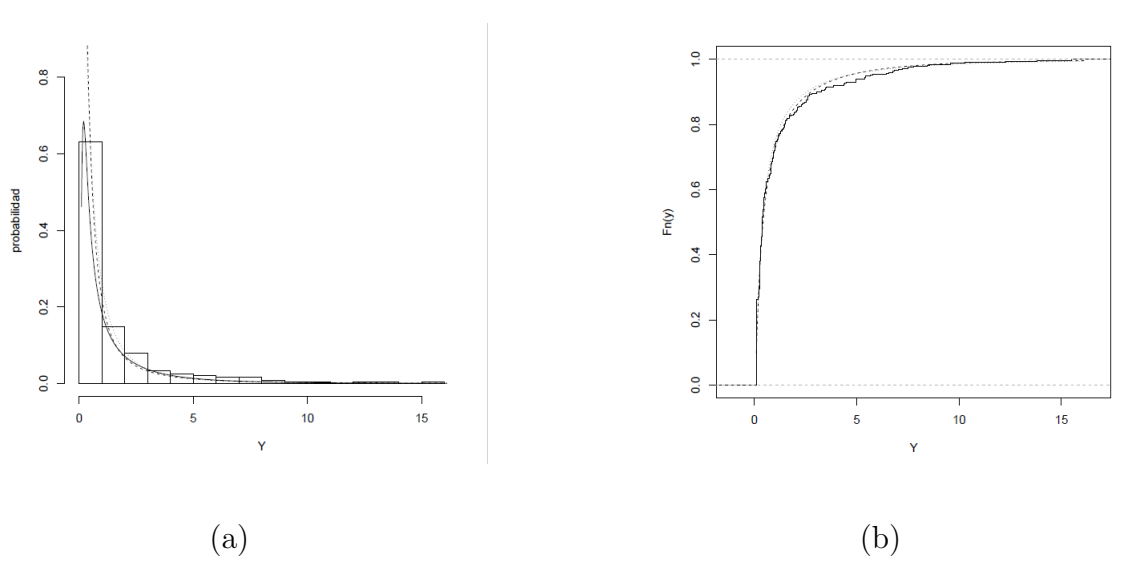


Figura 6.2: Histograma de la distribución del conjunto de datos para el estudio de la vacuna contra el sarampión. Las líneas simbolizan las distribuciones ajustadas usando las estimativas de máxima verosimilitud: LN (línea con puntos), $L\beta SN$ (línea discontinua) y $LP\beta SN$ (línea continua). (b) Para la FDA: Empírica (línea continua), $L\beta SN$ (línea de puntos) y $LP\beta SN$ (línea discontinua) .

Bibliografía

- [1] AZZALINI, A., *A class of distributions which includes the normal ones. Scand. J. Statist. (1985),*12, 171-178.
- [2] ARNOLD, B., BEAVER R., GROENEVELD R., MEEKER W., *The nontruncated marginal of a truncated bivariate normal distribution. Psychometrika. (1993) ,* 58:471-488
- [3] ARNOLD, B., GÓMEZ, H., SALINAS, H.S., *On Multiple Constraint Skewed Models. Statistic: A Journal of Theoretical and Applied Statistic, 43. (2009).,pp.* 279-293.
- [4] BIRNBAUM, Z. W AND SAUNDERS, S.C., *Effect of linear truncation on a multinormal population. Annals of Mathematical Statistics. (1950)., 21,*272-279.
- [5] BOLFARINE H., MARTÍNEZ-FLÓREZ, G.; SALINAS,H., *Bimodal Symmetric-Asymmetric Power-Normal Families. (2013).,28 : 1649?1655.*
- [6] BOLFARINE H., GÓMEZ H.W., RIVAS L., *The log-bimodal-skew-normal model. A geochemical application. Journal of Chemometrics. (2011).,25(6):* 329-332.

- [7] CHIOGNA, M., *Notes on estimation problems with scalar skew-normal distributions. Technical Report 15, University of Padua, Dept Statistical Sciences. (1997).*
- [8] DURRANS, S.R, *Distributions of fractional order statistics in hydrology. Water resources research. (1992).* 28,1649-1655.
- [9] OLIVERO E.; GÓMEZ, H. AND QUINTANA, F., *Bayesian modeling using a class of bimodal skew-elliptical distributions. Journal of Statistical Planning and Inference. (2009).*; 139, pp. 1484-1492.
- [10] OLIVERO E.D., *Alpha-skew-normal distribution. Proyecciones Journal of Mathematics. (2010).*,29, 224-240.
- [11] GÓMEZ, H.; OLIVERO E., D.; SALINAS, H.; BOLFARINE, H., *Bimodal extension based on the skew-normal distribution with application to pollendata. Environmetrics. (2009).* , Sin press.
- [12] MARTÍNEZ-FLÓREZ, G., VERGARA, S., GONZÁLEZ, L. (2014), *The family of log-skew-normal alpha-power distributions using precipitation data*, Revista Colombiana de Estadística (2013a) Volume 36, Issue 1, p. 43-57.
- [13] MARTÍNEZ-FLÓREZ, G.; BOLFARINE, H.; GÓMEZ, H. W. (2018), *Censored bimodal symmetric-asymmetric families.*
- [14] MARTÍNEZ-FLÓREZ, G., JIMÉNEZ-NARVÁEZ, M., TOVAR-FALÓN, R. (2018), *Normal Asymmetric Distribution Alpha-Power.*
- [15] HENZE, N., *A probabilistic representation of the skew-normal distribution. Scandinavian Journal of Statistics, 13 (1986),*pp. 271-275.

- [16] KIM, H.J., *On a class of two-piece skew-normal distributions. Statistics: A journal of Theoretical and Applied Statistics*, 39 (2005)pp. 537-553.
- [17] LEHMANN, E. L., *The power of rank tests. Annals of Mathematical Statistics*, (1953), 24,23-43.
- [18] MATEUS-FIGUERAS G; PAWLOSKY-GLANH V., *Una alternativa a la distribución lognormal*, Actas del XXVII congreso nacional de estadística e investigación operativa (SEIO). Sociedad de Estadística e Investigación Operativa: Barcelona. (2003) 1849-1858.
- [19] NEVEKA, O., MARTÍNEZ-FLÓREZ, G.; BOLFARINE, H., *Bimodal Birnbaum Saunders distribution with applications to non negative measurements*.
- [20] O'HAGAN A., LEONARD T., *Bayes estimation subject to uncertainty about parameter constraints. Biometrika*. (1976), vol. 63: 201-203.
- [21] PEWSEY, A., *Problems of inference for Azzalini's skew-normal distribution. Journal of Applied Statistics*, 7 (2000),pp. 859-870.
- [22] R CORE TEAM. R: A LANGUAGE AND ENVIRONMENT FOR STATISTICAL COMPUTING. R FOUNDATION FOR STATISTICAL COMPUTING, VIENNA, AUSTRIA. (2017) URL [HTTPS://WWW.R-PROJECT.ORG/](https://www.R-project.org/). 33. ROBERTS C., *A correlation model useful in the study of twins. Journal of the American Statistical Association* (1966),61: 1184- 1190.
- [23] SHAH,S.;CHAKRABORTY, S.;HAZARIKA,P., *Generalized Balakrishnan Alpha Skew Normal Distribution and Its Applications*.