

UNIVERSIDAD NACIONAL AGRARIA LA MOLINA

FACULTAD DE ECONOMÍA Y PLANIFICACIÓN



**“MIXTURA FINITA BASADA EN LA DISTRIBUCIÓN
BIRNBAUM-SAUNDERS NORMAL ASIMÉTRICA”**

Presentada por:

ROCÍO PAOLA MAEHARA ALIAGA

**TESIS PARA OPTAR EL TÍTULO DE
INGENIERO ESTADÍSTICO E INFORMÁTICO**

Lima - Perú

2018

*El presente trabajo se lo dedico a
Jesús, mi Señor y Redentor por
ser mi fortaleza, roca mía y
torre fuerte.*

AGRADECIMIENTOS

- A Dios por su presencia constante en mi vida, por confortarme y ayudarme a enfrentar todos los momentos difíciles.
- A mi esposo Luis por su amor, apoyo, ánimos y paciencia en medio de la adversidad.
- A mis padres por todo su amor, apoyo y cariño en cada etapa de mi vida, han realizado un trabajo excelente como padres.
- A mi hermano por buscar siempre ayudarme en todo lo que hago.
- A mi patrocinadora Ana Cecilia Vargas Paredes por su apoyo constante.
- A los profesores del Departamento de Estadística e Informática por sus importantes enseñanzas.

ÍNDICE

RESUMEN

| | |
|--|-----------|
| I. INTRODUCCIÓN | 1 |
| 1.1 JUSTIFICACIÓN DE LA INVESTIGACIÓN | 2 |
| 1.2 OBJETIVOS DE LA INVESTIGACIÓN | 2 |
| 1.2.1 OBJETIVOS ESPECÍFICOS | 3 |
| 1.3 ORGANIZACIÓN DEL TRABAJO | 3 |
| II. REVISIÓN DE LITERATURA | 4 |
| III. MATERIALES Y MÉTODOS | 6 |
| 3.1 LA DISTRIBUCIÓN BIRNBAUM-SAUNDERS | 6 |
| 3.2 LA DISTRIBUCIÓN NORMAL ASIMÉTRICA | 7 |
| 3.3 LA DISTRIBUCIÓN BIRNBAUM-SAUNDERS NORMAL ASIMÉTRICA | 8 |
| 3.4 LA MIXTURA FINITA BIRNBAUM-SAUNDERS NORMAL ASIMÉTRICA | 14 |
| 3.4.1 ESTIMACIÓN DE PARÁMETROS VIA EL ALGORITMO EM | 16 |
| 3.4.2 VALORES INICIALES PARA EL ALGORITMO EM | 19 |
| 3.4.3 REGLA DE PARADA | 19 |
| 3.4.4 MATRIZ DE INFORMACIÓN OBSERVADA | 20 |
| 3.4.5 SELECCIÓN DE MODELOS | 21 |
| IV. RESULTADOS Y DISCUSIÓN | 22 |
| 4.1 ESTUDIO DE SIMULACIÓN | 22 |
| 4.1.1 ESTUDIO 1: RECUPERACIÓN DE LOS PARÁMETROS Y CONSIS- TENCIA DE LOS ERRORES ESTÁNDAR DE LOS ESTIMADORES | 23 |
| 4.1.2 ESTUDIO 2: PROPIEDADES ASINTÓTICAS DE LOS ESTIMADORES | 24 |
| 4.1.3 ESTUDIO 3: PROPIEDADES DE ROBUSTEZ DE LOS ESTIMADORES | 28 |
| 4.2 APLICACIÓN 1 | 30 |
| 4.3 APLICACIÓN 2 | 34 |
| V. CONCLUSIONES | 37 |
| VI. RECOMENDACIONES | 38 |
| VII. REFERENCIAS BIBLIOGRÁFICAS | 39 |
| VIII. ANEXOS | 43 |

ÍNDICE DE CUADROS

| | |
|---|----|
| Cuadro N°1: Media estimada, errores estándar y porcentaje de la cobertura de probabilidad para el escenario 1. Los valores teóricos de los parámetros están entre paréntesis. | 25 |
| Cuadro N°2: Media estimada, errores estándar y porcentaje de la cobertura de probabilidad para el escenario 2. Los valores teóricos de los parámetros están entre paréntesis. | 26 |
| Cuadro N°3: Comparación de la máxima log-verosimilitud, AIC y BIC para el ajuste de los modelos MF-BS y MF-BS-NA usando el conjunto de datos del índice de masa corporal. El número de parámetros es denotado por m | 31 |
| Cuadro N°4: EMV con sus correspondientes EE para el modelo de MF-BS-NA aplicados al conjunto de datos IMC. | 33 |
| Cuadro N°5: Comparación de la máxima log-verosimilitud, AIC y BIC para el ajuste de los modelos MF-BS y MF-BS-NA usando el conjunto de datos de enzima. El número de parámetros es denotado por m | 34 |
| Cuadro N°6: EMV con sus correspondientes EE para el modelo de MF-BS-NA aplicados al conjunto de datos Enzima. | 36 |

ÍNDICE DE FIGURAS

| | |
|--|----|
| Figura N°1: Funciones de densidad de las distribuciones BS(0.8, 1) y BS-NA(0, 8, 1, λ). | 9 |
| Figura N°2: Funciones de densidad de la distribución BS-NA para diferentes valores de α , considerando $\beta = 1$, $\lambda = 2$ para (panel izquierdo) y $\beta = 1$, $\lambda = -2$ para (panel derecho). | 9 |
| Figura N°3: Funciones de densidad de la distribución BS-NA para diferentes valores de λ , considerando $\beta = 1$, $\alpha = 0.5$ para (a)-(b) y $\beta = 1$, $\alpha = 1$ para (c)-(d). | 10 |
| Figura N°4: Funciones de densidad de la distribución BS-NA para diferentes valores de β , considerando los mismos valores de α que en la Figura 2 y $\lambda = 2$. | 11 |
| Figura N°5: Gráficos de la función de densidad para algunos valores de p_1 ; (panel izquierdo) $(\alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.35, 0.65, 8, 5, 2, 4)$ y (panel derecho) $(\alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.35, 0.65, 8, 5, 4, -6)$. | 15 |
| Figura N°6: Gráficos de la función de razón de riesgo algunos valores de p_1 (panel izquierdo) $(\alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.35, 0.65, 8, 5, 2, 4)$ y (panel derecho) $(\alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (1.25, 1.25, 10, 4, 4, -6)$. | 16 |
| Figura N°7: Funciones de densidad a partir de las cuales se realizaron las simulaciones (panel izquierdo) $\theta = (0.4, 0.15, 0.45, 2, 3, -2, 4)$ y (panel derecho) $\theta = (0.4, 0.25, 0.35, 3, 7, -3, 3)$. | 22 |
| Figura N°8: Promedio de sesgo para todos los estimadores. | 27 |
| Figura N°9: Promedio de la raíz del error cuadrático medio para todos los estimadores. | 28 |
| Figura N°10: Cambios relativos promedio de los estimadores de α_1 , α_2 , β_1 y β_2 para diferentes perturbaciones ϑ . | 29 |
| Figura N°11: Cambios relativos promedio de los estimadores de λ_1 , λ_2 y p_1 para diferentes perturbaciones ϑ . | 30 |
| Figura N°12: (a) Histograma del conjunto de datos del IMC con ajustes de la MF-BS-NA y MF-BS con $G = 1 - 3$ componentes. (b) Gráfico de cajas del IMC. | 32 |
| Figura N°13: (a) Función de distribución acumulada y (b) función de supervivencia estimada con la función empírica de supervivencia para el ajuste de 3 modelos de la MF-BS-NA con $G = 1 - 3$ componentes. | 32 |
| Figura N°14: Histograma del IMC superpuesta con el ajuste de las densidades del modelo de MF-BS-NA (a) para $G = 1 - 3$ componentes y (b) $G = 3$ componentes. | 33 |

| | |
|---|----|
| Figura N° 15: (a) Histograma del conjunto de datos de Enzima con ajustes de la MF-BS-NA y MF-BS con $G = 1 - 3$ componentes. (b) Gráfico de cajas del conjunto de datos. | 35 |
| Figura N° 16: (a) Función de supervivencia estimada con la función empírica de supervivencia y (b) función de distribución acumulada para el ajuste de tres modelos MF-BS-NA con $G = 1 - 3$ componentes. | 35 |
| Figura N° 17: Histograma del conjunto de datos de enzima con las densidades de FM-BS-NA superpuestas (a) para $G = 2$ componentes (b) para cada componente. | 36 |

ÍNDICE DE ANEXOS

| | |
|---|----|
| Anexo N°1.: Demostración del Teorema 5 | 43 |
| Anexo N°2.: El paquete bssn | 45 |
| Anexo N°3.: Funciones para el ajuste de una BS-NA | 45 |
| Anexo N°4.: Gráficos para la BS-NA | 47 |
| Anexo N°5.: Funciones para el ajuste MF-BS-NA | 50 |
| Anexo N°6.: Gráficos para la MF-BS-NA | 52 |
| Anexo N°7.: Portafolio GitHub | 56 |

RESUMEN

Los modelos de mixtura han recibido una gran atención en el área de estadística debido a la amplia gama de aplicaciones encontradas en los últimos años. Por otro lado el modelo Birnbaum-Saunders (BS) surgió en un contexto de fatiga de materiales. Este modelo ha sido aplicado en otras áreas como por ejemplo, ciencias de la salud, ambiental, forestal demográficas, actuarial, financiera, entre otras. Teniendo en cuenta que la distribución BS se caracteriza por describir procesos de degradación acumulativa. La distribución Birnbaum-Saunders Normal Asimétrica (BS-NA) es una extensión de la distribución BS, ya que permite predecir percentiles extremos especialmente en la cola izquierda y a su vez modelar datos asimétricos.

Este trabajo discute el modelo de Mixtura Finita Birnbaum-Saunders Normal Asimétrica con G componentes, como una extensión del trabajo desarrollado por Benites *et al.* (2017), Vilca *et al.* (2011) y Balakrishnan *et al.* (2011). Esta propuesta es una clase flexible de distribuciones de probabilidad que permite modelar datos con comportamiento asimétrico, que poseen observaciones atípicas y que a su vez son provenientes de poblaciones heterogéneas. Para obtener los estimadores de máxima verosimilitud se usa el algoritmo EM con maximización condicional. Además, la matriz de información empírica se deriva analíticamente para obtener el error estándar. También se realizan estudios de simulación y analizan dos conjuntos de datos reales para ilustrar la utilidad del método propuesto. Finalmente, la propuesta del algoritmo y métodos son implementados en el programa R y posteriormente introducidos en el paquete **bssn** y en el portafolio GitHub.

ABSTRACT

Mixture models have received great attention in the area of statistics due to the wide range of applications found in recent years. On the other hand the Birnbaum-Saunders (BS) model emerged in a context of materials fatigue. This model has been applied in other areas such as, for example, health sciences, environmental, demographic forestry, actuarial, financial, among others. Taking into account that the BS distribution is characterized by describing cumulative degradation processes. The Birnbaum-Saunders Skew Normal distribution (BS-SN) The Birnbaum-Saunders Normal Asymmetric distribution is an extension of the BS distribution, since it allows predicting extreme percentiles especially in the left tail and in turn modeling asymmetric data.

This paper discusses the Finite Mixture Birnbaum-Saunders Skew Normal model with G components, as an extension of the work developed by Benites *et al.* (2017), Vilca *et al.* (2011) and Balakrishnan *et al.* (2011). This proposal is a flexible class of probability distributions that allows the modeling data with asymmetric behavior, which have atypical observations and which in turn come from heterogeneous populations. To obtain the maximum likelihood estimators, the EM algorithm with conditional maximization is used. In addition, the empirical information matrix is derived analytically to obtain the standard error. Simulation studies are also carried out and two sets of real data are analyzed to illustrate the usefulness of the proposed method. Finally, the proposal of the algorithm and methods are implemented in the program R and later introduced in the package **bssn** and in the GitHub profile.

I. INTRODUCCIÓN

La mixtura finita de distribuciones, es decir, la combinación lineal convexa de densidades (conocida como componente de la mixtura), ha sido ampliamente utilizada como una herramienta poderosa para modelar datos heterogéneos y para aproximar densidades de probabilidad complicadas, que presentan multimodalidad, asimetría y colas pesadas. Estos modelos han sido aplicados en varias áreas como genética, procesamiento de imágenes, medicina y economía. Estudios exhaustivos se encuentran disponibles en Böhning (2000), McLachlan y Peel (2000) y desde el punto de vista Bayesiano, en Frühwirth-Schnatter (2006).

Uno de los grandes problemas en el área industrial es la ruptura o falla de los materiales. Las fallas son ocasionadas por diversas causas y una de ellas es la fatiga. La fatiga es la falla estructural que resulta cuando el material es sometido repetidamente a diferentes esfuerzos dinámicos. Un modelo de fatiga de materiales que ha sido utilizado en el área industrial esta basada en la distribución para tiempos de vida propuesta por Birnbaum y Saunders (1969a), que fue inspirado en problemas de vibración encontrados en aviones comerciales. Este problema particular de fatiga permitió derivar una familia de distribuciones, la cual modela el tiempo de vida de materiales y equipos sujetos a cargas dinámicas.

Los modelos estadísticos para procesos de fatiga hacen posible describir la variación aleatoria de los tiempos de falla asociados a los materiales expuestos a la fatiga como resultado de diferentes patrones de fuerzas cíclicas. Estos materiales pueden ser caracterizados por parámetros de distribuciones de tiempos de vida. Tales caracterizaciones son importantes para predecir el desempeño de los materiales sujetos a cargas dinámicas.

Algunos modelos probabilísticos que han sido popularmente propuestos para describir el tiempo de vida por fatiga son por ejemplo las distribuciones Gamma, Gaussiana Inversa (GI), Log-normal y Weibull, que se ajustan con gran precisión en la región central de la distribución. En situaciones en que los tiempos de vida son pequeños o muy grandes, tales distribuciones no son apropiadas, pues producen un ajuste inapropiado. Por ello surge la necesidad de considerar distribuciones con colas más pesadas y/o distribuciones que permitan ajustar mejor los tiempos de vida pequeños, para obtener resultados estadísticos robustos y sobretodo que sean de fácil implementación en situaciones prácticas. Por otro lado también existen casos en que nos enfrentamos con datos provenientes de diferentes poblaciones y que además poseen asimetría surgiendo la necesidad de trabajar distribuciones que puedan ser de utilidad en este tipo de casos.

En este trabajo proponemos la distribución Mixtura Finita Birnbaum Saunders Normal Asimétrica (MF-BS-NA), la cual tiene una relación estrecha con la distribución BS, por ello se puede usar en el mismo contexto. De acuerdo con la literatura la distribución BS puede ser usada eficientemente para modelar fenómenos que se caracterizan por el desgaste o daños acumulativos. De acuerdo al origen de esta distribución se puede utilizar para analizar tiempos de vida relacionados al deterioro, desgaste o marchitamiento tanto en el área de confiabilidad como de sobrevivencia.

La distribución BS no solamente ha sido usada en las áreas mencionadas anteriormente, recientemente ha sido usada en negocios, criptografía, economía, finanzas, industria, seguros, inventario, nutrición, psicología, control de calidad y toxicología. Para mayores detalles sobre las diversas aplicaciones en una amplia variedad de contextos; véase, por ejemplo, Johnson *et al.* (1995) (pp. 651-663) y Leiva (2015) (pp. 11-15).

Se puede aplicar la MF-BS-NA cuando luego de realizar un análisis exploratorio del conjunto de datos, este presenta un comportamiento asimétrico, posee observaciones atípicas y a su vez parece ser proveniente de poblaciones heterogéneas (podría presentar multimodalidad).

1.1. JUSTIFICACIÓN DE LA INVESTIGACIÓN

Aprovechando las ventajas de la mixtura finita de distribuciones (MF) y la distribución BS basada en la distribución NA, la cual se denota por BS-NA, es posible obtener una clase general de distribuciones llamada MF-BS-NA. Los motivos principales para desarrollar esta nueva clase de distribuciones son las siguientes:

1. La mixtura finita de distribuciones permite modelar datos que pueden ser provenientes de una mixtura de observaciones procedentes de varias poblaciones.
2. La distribución BS-NA permite predecir muy bien los percentiles extremos, especialmente en la cola izquierda de la distribución, estos percentiles no pueden ser calculados fácilmente empleando modelos de uso común. Supóngase que los datos siguen una distribución BS-NA con parámetro de asimetría negativo, y en lugar de utilizar este modelo se usa la distribución BS o un algún otro modelo BS basado en una distribución simétrica, se terminará sobre-estimando los percentiles extremos inferiores.
3. La MF-BS-NA permitiría modelar datos de tiempos de vida heterogéneos, asimétricos y con colas pesadas y a su vez predecir adecuadamente los percentiles extremos.

1.2. OBJETIVOS DE LA INVESTIGACIÓN

El objetivo general es presentar un estudio de inferencia para el nuevo modelo propuesto, el cual es una extensión de los modelos mostrados en Vilca *et al.* (2011), Benites *et al.* (2017) y

Balakrishnan *et al.* (2011). El cual es obtenido al considerar una distribución BS-NA en lugar de la BS en un contexto de mixtura finita de distribuciones.

1.2.1. OBJETIVOS ESPECÍFICOS

1. Presentar un estudio de inferencia y verificar si las propiedades de la distribución MF-BS son preservadas cuando consideramos la distribución MF-BS-NA.
2. Mostrar el procedimiento de estimación para obtener los estimadores de máxima verosimilitud (EMV) usando el algoritmo EM.
3. Realizar estudios de simulación que permitan ilustrar el desempeño de la metodología utilizada.
4. Aplicar nuestro modelo en conjuntos de datos reales que permitan ilustrar la metodología utilizada.
5. Los códigos que serán incorporados al paquete estadístico R **bssn** (Maehara y Benites, 2016) para el modelo MF-BS-NA son útiles para obtener la función de densidad de probabilidad, la función de distribución acumulada, los estimadores de máxima verosimilitud usando el algoritmo EM, la función de supervivencia y la función de tasa de falla. A su vez se mostraran ejemplos de como usar dichos códigos.

1.3. ORGANIZACIÓN DEL TRABAJO

El Capítulo 2 presenta la revisión de la literatura mostrando los trabajos más representativos relacionados a la distribución BS, así como a las diferentes generalizaciones y aplicaciones desarrolladas a partir de esta distribución. También muestra los trabajos mas relevantes realizados considerando modelos de mixtura finita y su relevancia.

El Capítulo 3 describe brevemente las distribuciones BS y NA. Luego muestra las distribuciones BS-NA y MF-BS-NA describiendo sus propiedades, momentos, representación estocástica y función de supervivencia. También abordamos el proceso de estimación usando el algoritmo EM para la distribución MF-BS-NA, así como los valores iniciales y la regla de parada usando el método de Aitken. Finalmente, se muestra la metodología usada para obtener la matriz de información y los criterios utilizados para la selección de modelos.

El Capítulo 4 muestra diversos estudios de simulación y también dos aplicaciones a dos conjuntos de datos reales con la finalidad de mostrar el desempeño del modelo propuesto. Finalmente, algunas conclusiones y recomendaciones para posibles trabajos futuros son presentadas en los Capítulo 5 y 6 respectivamente.

II. REVISIÓN DE LITERATURA

Aunque la mayoría de las aplicaciones estadísticas se conciben para tratar con datos unimodales, en la investigación real los datos pueden exhibir heterogeneidad debido a la asimetría y multimodalidad. Más importante aún, aunque la asimetría y las variaciones de forma que pueden ser debidas a aspectos intrínsecos de los datos, asimetría y variaciones de forma que se asemejan a la multimodalidad pueden indicar que no todas las observaciones provienen de la misma población. En otras palabras, un conjunto de datos puede ser una mezcla de observaciones procedentes de varias poblaciones.

Los modelos de mezcla finita han sido ampliamente discutidos en la literatura y aplicados en diversas áreas, por ejemplo, biología, medicina, física, economía y marketing, entre otros. La importancia de la mezcla de distribuciones puede notarse a partir de la gran cantidad de libros disponibles sobre el tema, recientemente McNicholas (2017) muestra un enfoque en conglomerados y clasificación, Ferraro *et al.* (2017, pag. 181-186) propusieron un método restringido totalmente dirigido por datos para la estimación de mezcla de modelos de regresión lineal, Pyne *et al.* (2016, pag. 57-74) muestra la aplicación de mezcla de modelos a grandes conjuntos de datos, Tatarinova y Schumitzky (2015) desarrolla un amplio estudio sobre la mezcla de modelos no lineal desde una perspectiva bayesiana, teniendo como tema especial el *Modelado de Ecuaciones Estructurales: Una Revista Multidisciplinaria* que se centra en los avances en el modelo de mezclas finitas (Lubke y Grimm, 2017).

Birnbaum y Saunders (1969b) obtuvieron originalmente los estimadores de máxima verosimilitud de los parámetros. Mann *et al.* (1974) mostraron que la distribución BS es unimodal. Engelhardt *et al.* (1981) propusieron intervalos de confianza para los parámetros. Además, presentaron un estudio de prueba de hipótesis. Desmond (1985) derivó la distribución BS en base a un modelo biológico. El autor también extendió la justificación física para el uso de esta distribución, relajando algunas suposiciones hechas en Birnbaum y Saunders (1969a). Desmond (1986) investigó la relación entre la distribución BS y la distribución GI. Achcar (1993) presentó un método de estimación bayesiana, Lu y Chang (1997) utilizaron métodos bootstrap para construir intervalos de predicción para las observaciones futuras de una distribución BS. Dupuis y Mills (1997) utilizaron métodos robustos de estimación para los parámetros de la distribución BS en presencia de outliers.

La distribución BS surgió en el contexto de ingeniería de materiales, ya que la distribución es apropiada para describir procesos de degradación acumulativa, esto hace que la distribución BS sea un modelo que puede ser usado en otras áreas. En Leiva *et al.* (2008) y Leiva *et al.* (2010)

pueden ser encontrados algunos ejemplos de aplicaciones (o referencias bibliográficas) en ciencias de la salud, ambiental, forestal, demográficas, actuarial, financiera, entre otras. Todos estos aspectos han permitido considerar a la distribución BS como un modelo de probabilidad, en lugar de un modelo que permite apenas modelar tiempos de vida; ver Marshall y Olkin (2007).

Algunas generalizaciones y aplicaciones de la distribución BS fueron consideradas introduciendo nuevos parámetros o considerando distribuciones simétricas/asimétricas en lugar de la distribución normal, con la intención de obtener modelos más robustos o más flexibles para ser utilizados en problemas reales donde existen observaciones atípicas. Por ejemplo Owen y Padgett (1999) desarrollaron la distribución con tres parámetros. Barros *et al.* (2008) consideraron la distribución t de Student en lugar de la distribución normal estándar, donde los grados de libertad permiten controlar la curtosis de la distribución resultante, este nuevo modelo representa una extensión robusta de la distribución BS usual. Díaz-García y Leiva (2005) generalizaron la distribución BS basada en la clase de distribuciones elípticas que incluye las distribuciones Cauchy, Laplace, Logística, Normal, t de Student, entre otras.

La distribución NA ha sido estudiada ampliamente en las últimas décadas: véase, por ejemplo, Azzalini (1985) y Azzalini (2005). Cuando los datos son asimétricos, la NA podría ser utilizada en lugar de la distribución normal. En este caso, los percentiles ubicados en las colas (izquierda o derecha) deben ser ubicados de una mejor manera.

Extensiones de la distribución BS basada en la familia de distribuciones asimétricas han sido consideradas, específicamente sobre la clase de distribuciones asimétrico-elípticas. Algunos resultados pueden ser encontrados en Vilca y Leiva (2006), Leiva *et al.* (2007) y Leiva *et al.* (2010). En estos trabajos fueron obtenidos resultados teóricos, extendiendo las propiedades de las distribuciones BS y log-BS. El desarrollo de métodos de estimación e inferencia utilizando el algoritmo ECM para la distribución BS-NA pueden ser encontrados en Vilca *et al.* (2011).

El paquete de R, **bssn** (Maehara y Benites, 2016) fue creado el mes de agosto del año 2015 en su versión 0.5. En marzo del año 2016 fue actualizado y se obtuvo la versión 0.7 que posee diversos códigos útiles para el modelo BS-NA, los cuales permiten obtener la función densidad de probabilidad, la función de distribución acumulada, los cuantiles, la generación de números aleatorios, la función de supervivencia, la función de tasa de falla, la función de verosimilitud, los momentos, el algoritmo EM para obtener los estimadores de máxima verosimilitud y el gráfico cuantil-cuantil con su respectiva banda de confianza. Además, proporciona la generación de números aleatorios para el modelo MF-BS-NA.

III. MATERIALES Y MÉTODOS

3.1. LA DISTRIBUCIÓN BIRNBAUM-SAUNDERS

Birnbaum y Saunders (1969a) mostraron que sobre ciertas condiciones la función de distribución acumulada (fda) de una variable aleatoria T real no negativa que representa el tiempo total hasta que ocurra una falla es dada por

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

En la que T tiene una distribución Birnbaum-Saunders con parámetro de localización $\alpha > 0$ y parámetro de escala $\beta > 0$ denotada por $T \sim \text{BS}(\alpha, \beta)$.

Una propiedad importante en la construcción de esta distribución es que la variable aleatoria

$$Z = \frac{1}{\alpha} \left(\sqrt{\frac{T}{\beta}} - \sqrt{\frac{\beta}{T}} \right) \sim N(0, 1). \quad (2)$$

Así la distribución Birnbaum-Saunders está estrechamente relacionada con la distribución normal a través de la relación

$$T = \beta \left[\frac{\alpha Z}{2} + \sqrt{\left(\frac{\alpha Z}{2} \right)^2 + 1} \right]^2, \quad (3)$$

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

$$f_T(t) = \phi(a_t(\alpha, \beta)) A_t(\alpha, \beta), \quad (4)$$

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

$$a_t(\alpha, \beta) = \frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \quad \text{y} \quad A_t(\alpha, \beta) = \frac{d}{dt} a_t(\alpha, \beta) = \frac{t^{-3/2}(t + \beta)}{2\alpha\beta^{1/2}}. \quad (5)$$

Puede mostrarse fácilmente que si $T \sim \text{BS}(\alpha, \beta)$, entonces

- $aT \sim \text{BS}(\alpha, a\beta)$, para toda constante positiva a ,

- $T^{-1} \sim \text{BS}(\alpha, \beta^{-1})$

La media y la variancia son dadas por

$$\mathbb{E}(T) = \beta \left(1 + \frac{\alpha^2}{2}\right) \quad \text{y} \quad \text{var}(T) = (\alpha\beta)^2 \left(1 + \frac{5}{4}\alpha^2\right). \quad (6)$$

3.2. LA DISTRIBUCIÓN NORMAL ASIMÉTRICA

De acuerdo con Azzalini (1985), una variable aleatoria Z tiene una distribución normal asimétrica con parámetro de localización μ , parámetro de escala $\sigma > 0$ y parámetro de asimetría $\lambda \in \mathbb{R}$ si su fdp es dada por

$$f(z) = 2\phi(z|\mu, \sigma^2)\Phi\left(\frac{\lambda}{\sigma}(z - \mu)\right), \quad z \in \mathbb{R}, \quad (7)$$

donde $\phi(\cdot|\mu, \sigma^2)$ es la fdp de una distribución $N(\mu, \sigma^2)$ y $\Phi(\cdot)$ es la fda de una distribución normal estándar. El parámetro λ regula la asimetría de la función de densidad de Z . Particularmente, si $\lambda = 0$, $f(z)$ es la función de densidad de la distribución $N(\mu, \sigma^2)$. La distribución (7) es denotada por $Z \sim \text{NA}(\mu, \sigma^2, \lambda)$.

La media y la variancia de Z son dadas por

$$\mathbb{E}(Z) = \mu + \sqrt{\frac{2}{\pi}}\sigma\delta \quad \text{y} \quad \text{Var}(Z) = \sigma^2 - \frac{2}{\pi}\sigma^2\delta^2. \quad (8)$$

La representación estocástica de Z es dada por

$$Z \stackrel{d}{=} \mu + \sigma \left[\delta |W_0| + \sqrt{1 - \delta^2} W_1 \right], \quad \text{con} \quad \delta = \frac{\lambda}{\sqrt{1 + \lambda^2}}, \quad (9)$$

donde W_0 y W_1 son independientes y tienen una distribución normal estándar. Esta representación estocástica es importante para obtener los estimadores de máxima verosimilitud (EMV) usando el algoritmo EM, propuesto por Dempster *et al.* (1977). Si $H = |W_0|$ entonces H tiene una distribución Half Normal estándar denotada por $\text{HN}(0, 1)$ y tenemos que

$$Z|H = h \sim \text{H}(\mu + \sigma\delta h, \sigma^2(1 - \delta^2)). \quad (10)$$

3.3. LA DISTRIBUCIÓN BIRNBAUM-SAUNDERS NORMAL ASIMÉTRICA

De acuerdo con Vilca *et al.* (2011), sea T el tiempo total hasta que ocurra una falla, entonces en lugar de la suposición de normalidad, se asume que

$$Z = \frac{1}{\alpha} \left(\sqrt{\frac{T}{\beta}} - \sqrt{\frac{\beta}{T}} \right) \sim \text{NA}(0, 1, \lambda), \quad (11)$$

por lo tanto Z tiene una distribución normal asimétrica estándar con parámetro de asimetría λ . De esta forma podemos definir la distribución de la variable aleatoria T en términos de Z

$$T = \beta \left[\frac{\alpha Z}{2} + \sqrt{\left(\frac{\alpha Z}{2} \right)^2 + 1} \right]^2, \quad (12)$$

la cual depende de la distribución $\text{NA}(0, 1, \lambda)$. La variable aleatoria T basada en la distribución normal asimétrica sigue una distribución Birnbaum-Saunders Normal Asimétrica, la cual es denotada de la siguiente forma $T \sim \text{BS} - \text{NA}(\alpha, \beta, \lambda)$. Esta distribución es una generalización de la distribución Birnbaum-Saunders clásica, ya que la distribución de T para $\lambda = 0$ coincide con la distribución Birnbaum-Saunders.

Una variable aleatoria T sigue una distribución BS-NA con parámetro de forma $\alpha > 0$, parámetro de escala $\beta > 0$ y parámetro de asimetría $\lambda \in \mathbb{R}$, si su fdp es dada por

$$f_T(t) = 2\phi(a_t(\alpha, \beta))A_t(\alpha, \beta)\Phi(\lambda a_t(\alpha, \beta)), \quad (13)$$

donde $\phi(\cdot)$ y $\Phi(\cdot)$ son respectivamente la fdp y la fda de la normal estándar, $a_t(\alpha, \beta)$ y $A_t(\alpha, \beta)$ son dadas como en (5). La Figura 1 presenta la densidad de la distribución BS y la distribución BS-NA considerando $\lambda = 3, 4$ y 5 . Los valores de las densidades son transformados de manera que tengan los mismos valores en el origen, para facilitar la comparación de las densidades. Puede verse que las densidades son asimétricas a la derecha y que las densidades BS-NA(0.8, 1, λ), considerando $\lambda = 3, 4$ y 5 tiene colas más pesadas comparada con la densidad de la BS.

La Figura 2 muestra los gráficos de la densidad (13) para algunos valores de α considerando $\beta = 1$, $\lambda = 2$ y $\lambda = -2$. Se puede apreciar que a medida que el valor de α decrece, la densidad se torna más simétrica alrededor de β y la variancia también decrece con α tanto para $\lambda = 2$ como para $\lambda = -2$, lo cual es similar en la distribución BS.

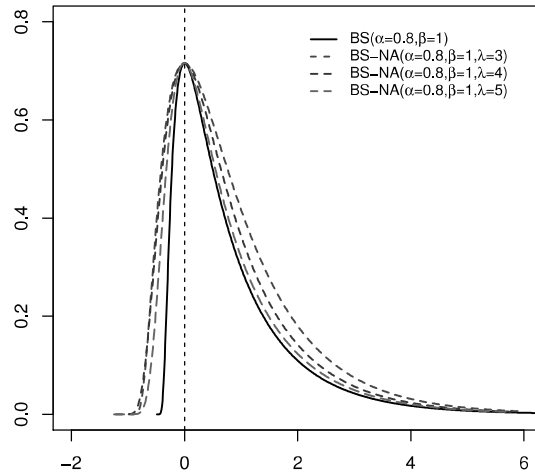


Figura 1: Funciones de densidad de las distribuciones $BS(0.8, 1)$ y $BS-NA(0.8, 1, \lambda)$.

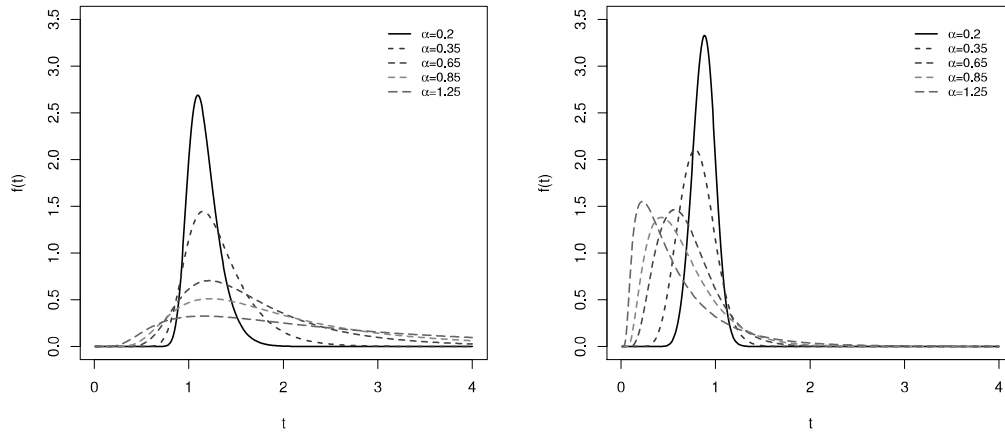


Figura 2: Funciones de densidad de la distribución $BS-NA$ para diferentes valores de α , considerando $\beta = 1$, $\lambda = 2$ para (panel izquierdo) y $\beta = 1$, $\lambda = -2$ para (panel derecho).

La Figura 3 muestra los gráficos de la densidad (13) para algunos valores de λ tanto positivos como negativos, considerando $\alpha = 0.5$ (Figuras (a) y (b)), $\alpha = 1$ (Figuras (c) y (d)) y $\beta = 1$. Puede notarse que para valores de $\lambda > 0$ las colas de la distribución $BS-NA$ son más pesadas en relación a la distribución BS , la cual se obtiene cuando $\lambda = 0$. Cuando $\lambda < 0$ la distribución $BS-NA$ modela mejor los percentiles bajos, siendo más flexible en la curtosis y asimetría que la distribución BS .

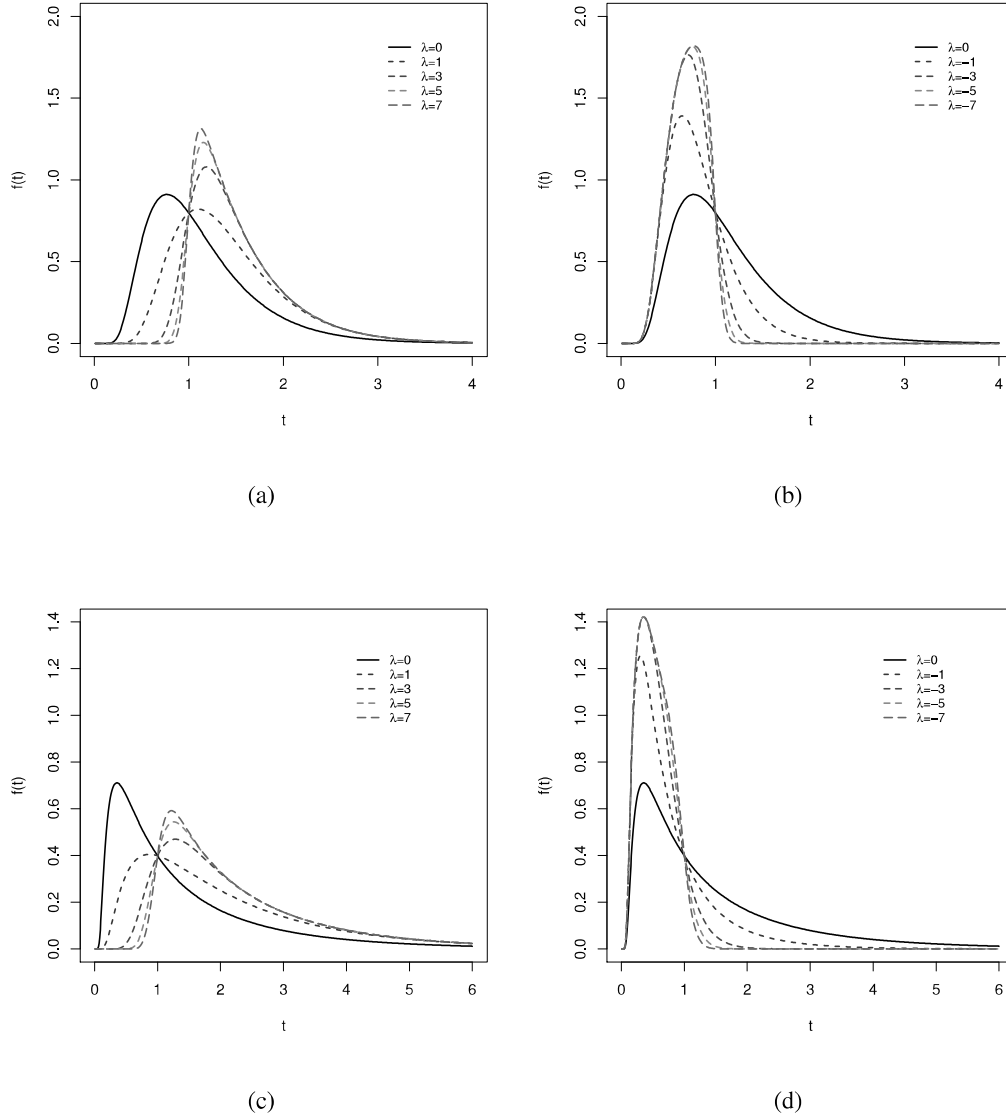


Figura 3: Funciones de densidad de la distribución BS-NA para diferentes valores de λ , considerando $\beta = 1, \alpha = 0.5$ para (a)-(b) y $\beta = 1, \alpha = 1$ para (c)-(d).

La Figura 4 presenta los gráficos para la densidad (13) para algunos valores de β , considerando los mismos valores de α dados en la Figura 2. Note que las curvas en los cuatro gráficos tienen el mismo comportamiento que la Figura 2 (a), con la diferencia que la escala en el eje x crece a medida que aumenta el valor de β . Lo que muestra que efectivamente β es el parámetro de escala.

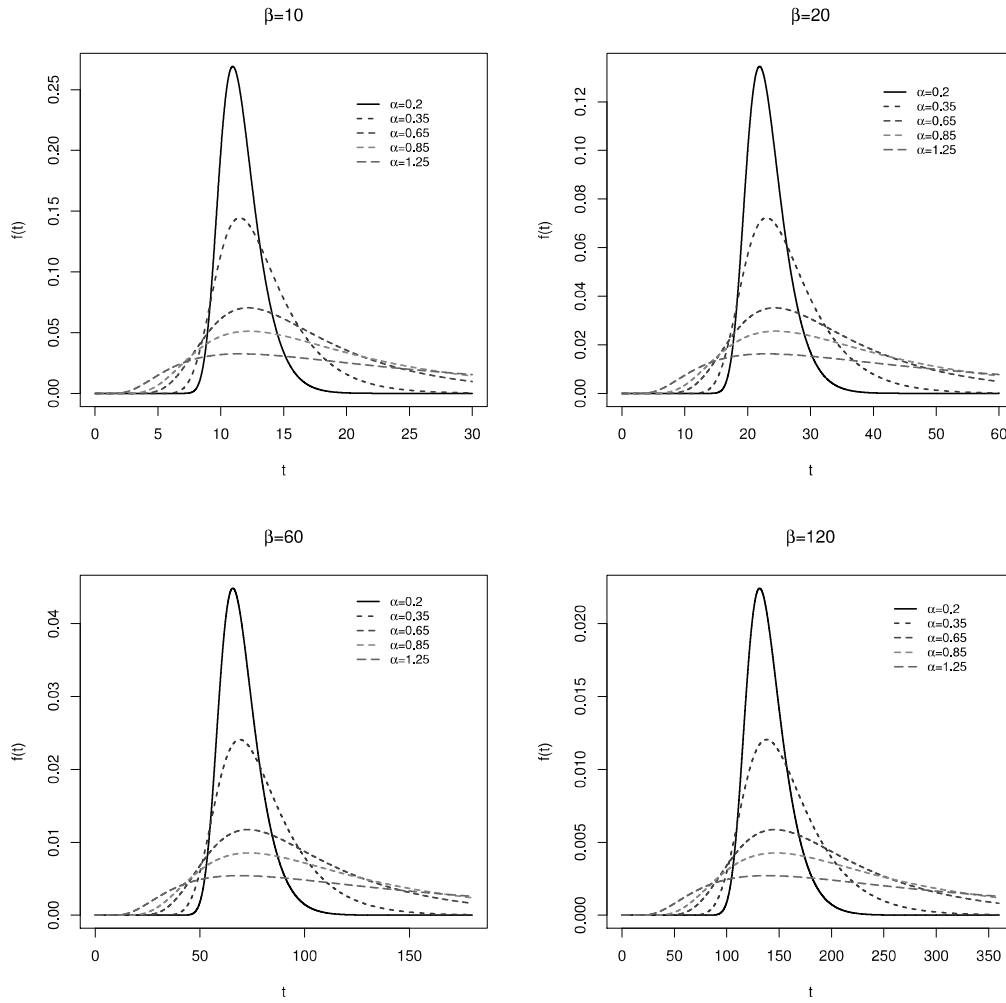


Figura 4: Funciones de densidad de la distribución BS-NA para diferentes valores de β , considerando los mismos valores de α que en la Figura 2 y $\lambda = 2$.

De acuerdo con Vilca *et al.* (2011) se presenta las propiedades más importantes de la distribución BS-NA.

Teorema 1. Sea $T \sim \text{BS-NA}(\alpha, \beta, \lambda)$. Entonces

- $aT \sim \text{BS-NA}(\alpha, a\beta, \lambda)$, para toda constante positiva a ,
- $T^{-1} \sim \text{BS-NA}(\alpha, \beta^{-1}, \lambda)$.

Sean $T \sim \text{BS-NA}(\alpha, \beta, \lambda)$ y Z dado en (11). Entonces

$$Z^2 = \frac{1}{\alpha^2} \left(\frac{T}{\beta} + \frac{\beta}{T} - 2 \right) \sim \chi_1^2, \quad (14)$$

donde χ_1^2 denota la distribución chi cuadrado con 1 grado de libertad.

Teorema 2. Sean $T_{NA} \sim \text{BS-NA}(\alpha, \beta, \lambda)$ y $T_N \sim \text{BS}(\alpha, \beta)$. Entonces la media y la variancia de estas dos variables aleatorias satisfacen las siguientes relaciones

- $\mathbb{E}(T_{NA}) = \mathbb{E}(T_N) + \frac{\alpha\beta}{2}\omega_1$,
- $\text{Var}(T_{NA}) = \text{Var}(T_N) + \left(\frac{\alpha\beta}{2}\right)^2 [2\alpha(\omega_3 - \omega_1) - \omega_1^2]$,

donde $\omega_k = \mathbb{E}\left(Z^k \sqrt{\alpha^2 Z^2 + 4}\right)$ para $k = 1, 3$, con $Z \sim \text{NA}(0, 1, \lambda)$. Para calcular los valores de ω_k , las integrales involucradas deben ser resueltas usando métodos numéricos.

Siguiendo las ideas de Santana (2009) se muestra los siguientes teoremas, los cuales serán de utilidad para la implementación del algoritmo EM

Teorema 3. Sea Z dada en (11) y $T \sim \text{BS} - \text{NA}(\alpha, \beta, \lambda)$. Considerando que Z tiene la representación estocástica dada en (9), la cual es $Z = \delta H + \sqrt{1 - \delta^2} W_1$. Entonces

$$T|H = h \sim \text{BSE}(\alpha_\delta, \beta, \sigma = 2, v_h)$$

donde $\alpha_\delta = \alpha\sqrt{1 - \delta^2}$ y $v_h = -\frac{\delta}{\sqrt{1 - \delta^2}}h$. La función de densidad condicional de T dado $H = h$ es dada por

$$f_{T|H}(t|h) = \phi(v_h + a_t(\alpha_\delta, \beta)) A_t(\alpha_\delta, \beta), \quad (15)$$

donde $a_t(\alpha_\delta, \beta)$ y $A_t(\alpha_\delta, \beta)$ se dan en (5)

Demostración: Como $Z \sim \text{NA}(0, 1, \lambda)$, entonces se usa la representación estocástica dada en (9) podemos expresar Z como

$$Z = \delta H + \sqrt{1 - \delta^2} W_1 = a_t(\alpha, \beta).$$

De esta forma la distribución condicional de Z dado $H = h$ es dada por

$$Z|(H = h) = \delta h + \sqrt{1 - \delta^2} W_1 = a_t(\alpha, \beta)|(H = h) \sim N(\delta h, 1 - \delta^2).$$

Después de algunas operaciones algebraicas tenemos que

$$\frac{a_t(\alpha, \beta)|(H = h) - \delta h}{\sqrt{1 - \delta^2}} = X_1 \sim N(0, 1)$$

Finalmente

$$v_h + a_t(\alpha_\delta, \beta) \sim N(0, 1) \quad (16)$$

Así del resultado (16), tenemos la demostración. ■

La fdp que resulta Teorema 3 es conocida como la distribución Birnbaum-Saunders extendida. Una variable aleatoria T que sigue esta distribución es denotada por $T \sim \text{BSE}(\alpha_\delta, \beta, \sigma = 2, v_h)$. Algunas propiedades de esta distribución fueron discutidas en Leiva *et al.* (2010).

Lema 1. Sean $X \sim N(\mu_x, \sigma_x^2)$ e $Y \sim N(\mu_y, \sigma_y^2)$. Entonces

$$\phi(y|\mu_y + cx, \sigma_y^2)\phi(x|\mu_x, \sigma_x^2) = \phi(y|\mu_y + c\mu_x, \sigma_y^2 + c^2\sigma_x^2)\phi\left(x|\mu_x + \Lambda\frac{c}{\sigma_y^2}(y - \mu_y - c\mu_x), \Lambda\right),$$

donde $\Lambda = \frac{\sigma_x^2\sigma_y^2}{\sigma_y^2 + c^2\sigma_x^2}$.

Lema 2. Sea $X \sim N(\mu, \sigma^2)$. Entonces para cualquier constante real c

$$\mathbb{E}(X|X > c) = \mu + \frac{\phi\left(\frac{c-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{c-\mu}{\sigma}\right)}\sigma,$$

$$\mathbb{E}(X^2|X > c) = \mu^2 + \sigma^2 + \frac{\phi\left(\frac{c-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{c-\mu}{\sigma}\right)}(\mu + c)\sigma.$$

Teorema 4. Sea $T \sim \text{BS} - \text{NA}(\alpha, \beta, \lambda)$ y $H \sim \text{HN}(0, 1)$. Entonces $H|(T = t)$ también tiene una distribución Half Normal con la siguiente función de densidad

$$f_{H|T}(h|t) = \frac{\phi(h|\delta a_t(\alpha, \beta), 1 - \delta^2)}{\Phi(\lambda a_t(\alpha, \beta))}, \quad t > 0,$$

donde $\phi(\cdot|\mu, \sigma^2)$ denota la función de densidad de una distribución normal con media μ y varianza σ^2 . Por lo tanto

$$\mathbb{E}[H|(T = t)] = \delta a_t(\alpha, \beta) + \sqrt{1 - \delta^2} W_{\Phi}\left(\frac{\delta a_t(\alpha, \beta)}{\sqrt{1 - \delta^2}}\right)$$

y

$$\mathbb{E}[H^2|(T = t)] = \delta^2 a_t^2(\alpha, \beta) + (1 - \delta^2) + \delta \sqrt{1 - \delta^2} a_t(\alpha, \beta) W_{\Phi}\left(\frac{\delta a_t(\alpha, \beta)}{\sqrt{1 - \delta^2}}\right),$$

donde $W_{\Phi}(u) = \phi(u)/\Phi(u)$, para $u \in \mathbb{R}$.

Demostración: Dado que $f_H(h) = 2\phi(h|0, 1)I_{[0, \infty)}$ y

$$\phi\left(-\frac{\delta}{\sqrt{1 - \delta^2}}h + \frac{1}{\sqrt{1 - \delta^2}}a_t(\alpha, \beta)|0, 1\right) = \sqrt{1 - \delta^2}\phi(a_t(\alpha, \beta)|\delta h, 1 - \delta^2).$$

Luego tenemos que

$$\begin{aligned} f_{H|T}(h|t) &= \frac{f_{T|H}(t|h)f_H(h)}{f_T(t)}, \\ &= \frac{2\phi(h|0, 1)\phi(a_t(\alpha, \beta)|\delta h, 1 - \delta^2)A_t(\alpha, \beta)}{2a_t(\alpha, \beta)A_t(\alpha, \beta)\Phi(\lambda a_t(\alpha, \beta))}, \end{aligned}$$

Entonces usando el Lema 1 en el numerador, obtenemos la función de densidad deseada de la siguiente forma

$$\begin{aligned} f_{H|T}(h|t) &= \frac{\phi(a_t(\alpha, \beta)|0, 1) \phi(h|\delta a_t(\alpha, \beta), 1 - \delta^2)}{a_t(\alpha, \beta) \Phi(\lambda a_t(\alpha, \beta))}, \\ &= \frac{\phi(h|\delta a_t(\alpha, \beta), 1 - \delta^2)}{\Phi(\lambda a_t(\alpha, \beta))}. \end{aligned}$$

Para demostrar las esperanzas condicionales para $k = 1, 2$ tenemos que

$$\begin{aligned} \mathbb{E}[H^k|T] &= \int_0^\infty h^k \frac{\phi(h|\delta a_t(\alpha, \beta), 1 - \delta^2)}{\Phi(\lambda a_t(\alpha, \beta))} dh, \\ &= \frac{1}{\Phi(\lambda a_t(\alpha, \beta))} \int_0^\infty h^k \phi(h|\delta a_t(\alpha, \beta), 1 - \delta^2) dh, \\ &= \mathbb{E}[X^k|X > 0] \end{aligned}$$

Luego, usando el Lema 2 se pueden demostrar ambas esperanzas condicionales. ■

3.4. LA MIXTURA FINITA BIRNBAUM-SAUDERS NORMAL ASIMÉTRICA

Considerando una variable aleatoria Y , la cual se ha considerado que proviene de G diferentes distribuciones BS-NA, con fdp como sigue

$$f_Y(y; \mathbf{p}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda}) = \sum_{j=1}^G p_j f_{T_j}(y; \alpha_j, \beta_j, \lambda_j), \quad y \in \mathbb{R}_+, \quad \sum_{j=1}^G p_j = 1, \quad (17)$$

donde $j = 1, \dots, G$, $\mathbf{p} = (p_1, \dots, p_G)^\top$, $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_G)^\top$, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_G)^\top$ y $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_G)^\top$, con $T_j \sim BS-NA(\alpha_j, \beta_j, \lambda_j)$, con fdp definida en (13), además $\alpha_j > 0$, $\beta_j > 0$, $\lambda_j \in \mathbb{R}$, p_j son las proporciones de mixtura que están restringidas a ser positivas. Esta siendo considerado que el número de componentes $G \in \mathbb{N}_+$ es conocido, fijo y positivo, mientras que \mathbf{p} es un vector desconocido de los pesos de mixtura. Note que α_j , β_j , λ_j y p_j son los parámetros para el componente j -ésimo. Una variable aleatoria con función de densidad (17), sigue una distribución de Mixtura Finita de Birnbaum-Saunders Normal Asimétrica, la cual será denotada como MF-BS-NA($\mathbf{p}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda}$). La Figura 5 muestra que la densidad de la MF-BS-NA puede tomar diferentes formas, es decir, en esta figura tenemos una mixtura de dos componentes, la cual puede ser tanto unimodal como bimodal dependiendo de los valores de los parámetros.

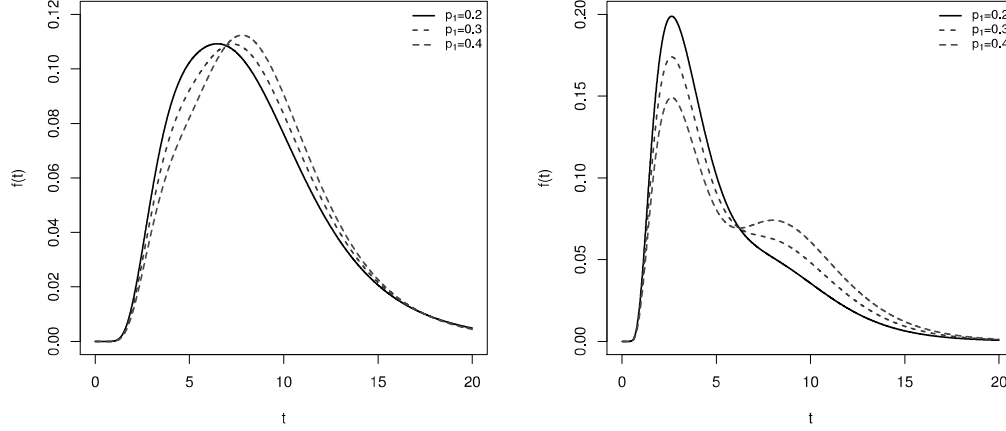


Figura 5: Gráficos de la función de densidad para algunos valores de p_1 ; (panel izquierdo) $(\alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.35, 0.65, 8, 5, 2, 4)$ y (panel derecho) $(\alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.35, 0.65, 8, 5, 4, -6)$.

Algunas propiedades de la distribución MF-BS-NA pueden derivarse usando la relación estrecha entre la distribución T_j y la distribución normal asimétrica.

Teorema 5. Si $\text{MF-BS-NA}(\mathbf{p}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda})$, Específicamente, sea $Y \sim \text{MF-BS-NA}(\mathbf{p}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda})$. Entonces,

- (i) $cY \sim \text{MF-BS-NA}(\mathbf{p}, \boldsymbol{\alpha}, c\boldsymbol{\beta}, \boldsymbol{\lambda})$, donde $c \in \mathbb{R}_+$;
- (ii) Si $Y^{-1} \sim \text{MF-BS-NA}(\mathbf{p}, \boldsymbol{\alpha}, \boldsymbol{\beta}^{-1}, \boldsymbol{\lambda})$, donde $\boldsymbol{\beta}^{-1} = (1/\beta_1, \dots, 1/\beta_G)^\top$;
- (iii) La función de distribución de Y es $F_Y(y) = \sum_{j=1}^G p_j \Phi_{\lambda_j}(a_y(\alpha_j, \beta_j))$ donde $\Phi_{\lambda_j}(\cdot)$ es la función de distribución acumulada (fda) de una variable aleatoria $\text{NA}(0, 1, \lambda_j)$;
- (iv) Si $W = \log(Y)$, entonces la función de densidad de probabilidad de W es

$$f_W(w; \mathbf{p}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda}) = \sum_{j=1}^G p_j f_{V_j}(w; \alpha_j, \gamma_j, \lambda_j), \quad w \in \mathbb{R}_+,$$

$$\text{donde } f_{V_j}(w; \alpha_j, \gamma_j, \lambda_j) = 2\phi(\xi_2(w; \alpha_j, \gamma_j))\xi_1(w; \alpha_j, \gamma_j)\Phi(\lambda_j \xi_2(w; \alpha_j, \gamma_j)), \text{ con}$$

$$\xi_2(w; \alpha_j, \gamma_j) = \frac{2}{\alpha_j} \sinh\left(\frac{w-\gamma_j}{2}\right) \text{ y } \xi_1(w; \alpha_j, \gamma_j) = \frac{2}{\alpha_j} \cosh\left(\frac{w-\gamma_j}{2}\right), \quad \gamma_j = \log(\beta_j).$$

La demostración del Teorema 5 se encuentra en el Anexo 1.

Teorema 6. Sea $Y \sim \text{MF-BS-NA}(\mathbf{p}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda})$. Entonces la moda (modas) y la mediana de la distribución MF-BS-NA se obtienen, respectivamente, resolviendo las siguientes ecuaciones no lineales con respecto a y

$$\text{Moda: } \sum_{j=1}^G p_j \phi(a_y(\alpha_j, \beta_j)) \left[A_y^2(\alpha_j, \beta_j) b_y(\alpha_j, \beta_j, \lambda_j) - \frac{y^{-5/2}(y+3\beta_j)}{4\alpha_j \beta_j^{1/2}} \right] = 0,$$

$$\text{Mediana: } \sum_{j=1}^G p_j \Phi_{\lambda_j}(a_y(\alpha_j, \beta_j)) = 0.5,$$

$$\text{donde } b_y(\alpha_j, \beta_j, \lambda_j) = \lambda_j \phi(\lambda_j a_y(\alpha_j, \beta_j)) - a_y(\alpha_j, \beta_j).$$

Teorema 7. Sea $Y \sim \text{MF-BS-NA}(\mathbf{p}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda})$. Entonces, la función de supervivencia (*fs*) y la función de razón de riesgos o tasa instantánea de fallas (*fr*) de Y son, respectivamente,

$$S_Y(y) = \sum_{j=1}^G p_j S_{T_j}(y) \quad \text{y} \quad h_Y(y) = \frac{\sum_{j=1}^G p_j f_{T_j}(y; \alpha_j, \beta_j, \lambda_j)}{S_Y(y)},$$

$$\text{donde } S_{T_j}(y) = 1 - \Phi_{\lambda_j}(a_y(\alpha_j, \beta_j)).$$

En la Figura 6 se encuentra representada la función de la razón de riesgo del modelo de MF-BS-NA para una combinación diferente de parámetros. Podemos notar que la figura del panel derecho presenta mayor asimetría hacia la derecha comparada con la del panel izquierdo.

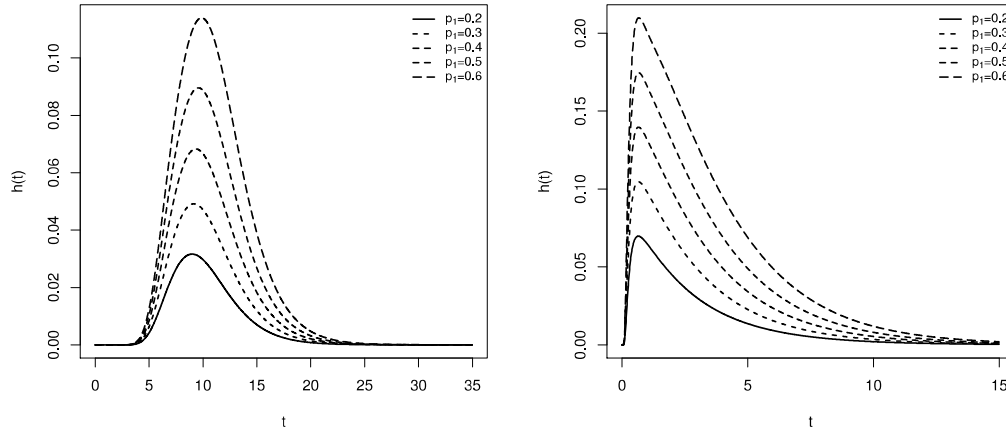


Figura 6: Gráficos de la función de razón de riesgo algunos valores de p_1 (panel izquierdo) $(\alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (0.35, 0.65, 8, 5, 2, 4)$ y (panel derecho) $(\alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2) = (1.25, 1.25, 10, 4, 4, -6)$.

3.4.1. ESTIMACIÓN DE PARÁMETROS VIA EL ALGORITMO EM

Para cada i y j , consideramos la variable indicadora latente Z_{ij} tal que:

$$P(Z_{ij} = 1) = 1 - P(Z_{ij} = 0) = p_j, \quad \sum_{j=1}^G Z_{ij} = 1, \quad y_i | Z_{ij} = 1 \sim \text{BS-NA}(\alpha_j, \beta_j, \lambda_j). \quad (18)$$

Tenga en cuenta que al integrar $\mathbf{Z}_i = (Z_{i1}, \dots, Z_{iG})^\top$ es obtenido la densidad marginal (13). $\mathbf{Z}_1, \dots, \mathbf{Z}_n$ son vectores aleatorios independientes, cada uno con una distribución multinomial con función de probabilidad

$$f(\mathbf{z}_i) = p_1^{z_{i1}} p_2^{z_{i2}} \dots (1 - p_1 - \dots - p_{G-1})^{z_{iG}},$$

que se denota por $\mathbf{Z}_i \sim \text{Multinomial}(1; p_1, \dots, p_G)$. Estos vectores latentes aparecen en la representación jerárquica dada a continuación, que se utiliza para construir el algoritmo EM con maximización condicional (ECM). Dado que de (18) tenemos que $y_i|Z_{ij} = 1 \sim \text{BS} - \text{NA}(\alpha_j, \beta_j, \lambda_j)$ y usando el Teorema 3, el modelo MF-BS-NA puede ser representado como

$$Y_i|H_i = h_i, Z_{ij} = 1 \stackrel{\text{ind}}{\sim} \text{BSE}(\alpha_j \delta_j, \beta_j, \sigma_j = 2, \mathbf{v}_{jh_i}), \quad (19)$$

$$H_i|Z_{ij} = 1 \stackrel{\text{ind}}{\sim} \text{HN}(0, 1), \quad (20)$$

$$\mathbf{Z}_i \stackrel{\text{iid}}{\sim} \text{Multinomial}(1, p_1, \dots, p_G) \quad (i = 1, \dots, n). \quad (21)$$

Para obtener los EMV del modelo MF-BS-NA vamos a utilizar el algoritmo ECM que fue propuesto por Meng y Rubin (1993). La idea básica del algoritmo ECM es que el paso de maximización (M) es reemplazado por varias etapas de maximización condicional (pasos CM). Para facilitar la notación, sea $\mathbf{y} = (y_1, \dots, y_n)$, $\mathbf{u} = (u_1, \dots, u_n)$ y $\mathbf{Z} = (Z_1, \dots, Z_n)$. De acuerdo (19), (20) y (21), la función de log-verosimilitud completa de $\boldsymbol{\theta} = (p_1, \dots, p_G, \alpha_1, \dots, \alpha_G, \beta_1, \dots, \beta_G, \lambda_1, \dots, \lambda_G)$ dado $(\mathbf{y}, \mathbf{u}, \mathbf{Z})$, sin considerar las constantes aditivas, es dada por

$$\begin{aligned} \ell_c(\boldsymbol{\theta}|\mathbf{y}, \mathbf{u}, \mathbf{Z}) &= c + \sum_{i=1}^n \sum_{j=1}^G z_{ij} \left[\log p_j - \log(\alpha_j) - \frac{1}{2} \log(\beta_j) + \log(y_i + \beta_j) - \frac{1}{2} \log(1 - \delta_j^2) \right. \\ &\quad \left. - \frac{1}{2(1 - \delta_j^2)} a_{y_i}^2(\alpha_j, \beta_j) + \frac{\delta_j}{(1 - \delta_j^2)} a_{y_i}(\alpha_j, \beta_j) h_i - \frac{\delta_j^2}{(1 - \delta_j^2)} h_i^2 \right], \end{aligned} \quad (22)$$

donde c es una constante que es independiente del vector de parámetros $\boldsymbol{\theta}$. Vamos a definir las siguiente expresiones

$$\begin{aligned} \hat{z}_{ij} &= E[Z_{ij}|\hat{\boldsymbol{\theta}}, y_i], \\ \hat{s}_{1ij} &= E[Z_{ij}H_i|\hat{\boldsymbol{\theta}}, y_i], \\ &\text{y} \\ \hat{s}_{2ij} &= E[Z_{ij}H_i^2|\hat{\boldsymbol{\theta}}, y_i], \end{aligned}$$

y usando algunas propiedades conocidas de la esperanza condicional, obtenemos lo siguiente

$$\begin{aligned} \hat{z}_{ij}^{(k)} &= \frac{\hat{p}_j^{(k)} f_{T_j}(y_i; \alpha_j^{(k)}, \beta_j^{(k)}, \lambda_j^{(k)})}{\sum_{j=1}^G \hat{p}_j^{(k)} f_{T_j}(y_i; \alpha_j^{(k)}, \beta_j^{(k)}, \lambda_j^{(k)})}, \\ \hat{s}_{1ij} &= \hat{z}_{ij} \hat{h}_{ij} \\ \text{and} \\ \hat{s}_{2ij} &= \hat{z}_{ij} \hat{h}_{ij}^2, \quad i = 1, \dots, n, \quad j = 1, \dots, G, \end{aligned}$$

donde las expresiones para \widehat{h}_{ij} y \widehat{h}_{ij}^2 son obtenidas usando el Teorema 4 y son dadas por

$$\begin{aligned}\widehat{h}_{ij} &= \delta_j a_{y_i}(\alpha_j, \beta_j) + W_\Phi \left(\frac{\delta_j a_{y_i}(\alpha_j, \beta_j)}{\sqrt{1 - \delta_j^2}} \right) \sqrt{1 - \delta_j^2}, \\ \widehat{h}_{ij}^2 &= (\delta_j a_{y_i}(\alpha_j, \beta_j))^2 + (1 - \delta_j^2) + \delta_j \sqrt{1 - \delta_j^2} a_{y_i}(\alpha_j, \beta_j) W_\Phi \left(\frac{\delta_j a_{y_i}(\alpha_j, \beta_j)}{\sqrt{1 - \delta_j^2}} \right).\end{aligned}$$

Por lo tanto, el valor esperado de la log-verosimilitud completa dada en la ecuación (22) evaluado en $\boldsymbol{\theta} = \boldsymbol{\theta}^{(k)}$, que denotaremos por la función- Q , después de algunos cálculos algebraicos, tenemos

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}) = E(\ell_c(\boldsymbol{\theta}|\mathbf{y}, \mathbf{u}, \mathbf{Z})|\mathbf{t}, \boldsymbol{\theta}^{(k)}). \quad (23)$$

Por tanto, la función- Q puede ser escrita como

$$\begin{aligned}Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}) &= c + \sum_{i=1}^n \sum_{j=1}^G \left[\widehat{z}_{ij}^{(k)} \left(\log p_j - \log \widehat{\alpha}_j - \frac{1}{2} \log \widehat{\beta}_j + \frac{1}{2} \log(y_i + \widehat{\beta}_j) - \frac{1}{2} \log(1 - \widehat{\delta}_j^2) \right. \right. \\ &\quad \left. \left. - \frac{1}{2\widehat{\alpha}_j(1 - \widehat{\delta}_j^2)} \xi^2(y_i; \widehat{\beta}_j) \right) + \frac{\widehat{\delta}_j}{\widehat{\alpha}_j(1 - \widehat{\delta}_j^2)} \xi(y_i; \widehat{\beta}_j) \widehat{s}_{1ij}^{(k)} - \frac{\widehat{\delta}_j}{2(1 - \widehat{\delta}_j^2)} \widehat{s}_{2ij}^{(k)} \right], \quad (24)\end{aligned}$$

donde $\xi(y_i; \widehat{\beta}_j) = \sqrt{y_i/\widehat{\beta}_j} - \sqrt{\widehat{\beta}_j/y_i}$.

En resumen, la implementación del algoritmo ECM para obtener los EMV de los parámetros del modelo MF-BS-NA se desarrolla de la siguiente manera:

Paso-E: Dado $\boldsymbol{\theta} = \boldsymbol{\theta}^{(k)}$, calculamos $\widehat{z}_{ij}^{(k)}$ para $i = 1, \dots, n$ y $j = 1, \dots, G$.

Paso-MC 1: Fijamos $\beta_j^{(k)}$ y actualizamos $\alpha_j^{(k)}$ y $\delta_j^{(k)}$ como

$$\begin{aligned}p_j^{(k+1)} &= \frac{\sum_{i=1}^n \widehat{z}_{ij}^{(k)}}{n}, \\ \alpha_j^{2(k)} &= \frac{\sum_{i=1}^n \widehat{z}_{ij}^{(k)} \xi(y_i; \widehat{\beta}_j^{(k)})}{\sum_{i=1}^n \widehat{z}_{ij}^{(k)}} + \left(1 - \frac{\sum_{i=1}^n \widehat{h}_{ij}^{2(k+1)}}{\sum_{i=1}^n \widehat{z}_{ij}^{(k)}} \right) \left(\frac{\sum_{i=1}^n \widehat{h}_{ij}^{(k)} \xi(y_i; \widehat{\beta}_j^{(k)})}{\sum_{i=1}^n \widehat{h}_{ij}^{(k)}} \right), \\ \text{and} \\ \widehat{\delta}_j^{(k+1)} &= \frac{\sum_{i=1}^n \widehat{h}_{ij}^{(k)} \xi(y_i; \widehat{\beta}_j^{(k)})}{\widehat{\alpha}_j^{2(k)} \sum_{i=1}^n \widehat{h}_{ij}^{2(k)}},\end{aligned}$$

Paso-MC 2: Fijamos $\widehat{\boldsymbol{\alpha}}^{(k+1)}$ y $\widehat{\boldsymbol{\delta}}^{(k+1)}$ luego actualizamos $\widehat{\boldsymbol{\beta}}^{(k+1)}$

$$\widehat{\boldsymbol{\beta}}^{(k+1)} = \arg \max_{\boldsymbol{\beta}} Q(\widehat{\boldsymbol{\alpha}}^{(k+1)}, \boldsymbol{\beta}, \widehat{\boldsymbol{\delta}}^{(k+1)} | \widehat{\boldsymbol{\theta}}^k).$$

Los EMV para λ_j son obtenidos utilizando la propiedad de invariancia de los EMV aprovechando la siguiente relación $\lambda_j = \delta_j / \sqrt{1 - \delta_j^2}$.

3.4.2. VALORES INICIALES PARA EL ALGORITMO EM

Es bien conocido que los modelos de mixtura pueden proporcionar una función de verosimilitud multimodal. En ese sentido, el método de estimación de máxima verosimilitud a través del algoritmo ECM puede no dar soluciones máximas globales si los valores iniciales están lejos de los valores reales de los parámetros. Los valores iniciales adoptados se resumen a continuación:

- Realizamos el algoritmo de K -médias (Benites *et al.*, 2016) o K -medoides (Kaufman y Rousseeuw, 1990) para inicializar con respecto a un centro de clúster elegido al azar;
- Inicializamos el indicador $\widehat{Z}_j^{(0)} = \{\widehat{z}_{ij}^{(0)}\}_{i=1}^G$ de acuerdo con el algoritmo de agrupamiento mostrado anteriormente;
- Inicializamos las proporciones de mixtura como sigue: $\widehat{p}_j^{(0)} = (1/n) \sum_{i=1}^n \widehat{z}_{ij}^{(0)}$. Esto es el valor inicial para p_j ;
- Para cada j -ésimo grupo, calculamos los valores iniciales $\alpha_j^{(0)}$ y $\beta_j^{(0)}$. Es utilizado la estimación por el método de momentos modificados propuesto por Ng *et al.* (2003) y que se encuentra implementado en la función `mmeth()`, la cual está disponible en el paquete de R llamado **bssn** (Maehara y Benites, 2016);
- Utilizando los valores iniciales obtenidos en el paso anterior calculamos $a_y(\alpha_j^{(0)}, \beta_j^{(0)})$ que de acuerdo a la teoría tiene una distribución normal asimétrica estándar y aprovechando la función `msn.mle(t)` que se encuentra implementada en el paquete de R conocido como **sn** (Azzalini, 2017) obteniendo así los valores iniciales para $\lambda_j^{(0)}$.

3.4.3. REGLA DE PARADA

Para evaluar la convergencia del algoritmo ECM para el modelo MF-BS-NA, se adopta el criterio de parada basado en el acelerador de Aitken (McLachlan y Krishnan, 2008)

$$|\ell^{(k+1)} - \ell_{\infty}^{(k+1)}| < \varepsilon,$$

donde $\ell^{(k+1)}$ es la log-verosimilitud observada en $\boldsymbol{\theta}^{(k)}$ y ε es el la tolerancia deseada. Para los análisis numéricos que serán presentados en el estudio de simulación y las aplicaciones,

será considerando un $\varepsilon = 10^{-6}$. Asumiendo convergencia de los EMV, $\widehat{\boldsymbol{\theta}}$, también $\ell_\infty^{(k+1)}$ es asintóticamente estimada de la log-verosimilitud en la iteración $k + 1$ (McLachlan y Krishnan (2008); Chap. 4.9), donde

$$\ell_\infty^{(k+1)} = \ell^{(k)} + \frac{\ell^{(k+1)} - \ell^{(k)}}{1 - c^{(k)}},$$

con $c^{(k)}$ denotando el acelerador de Aitken en la iteración k , dado por

$$c^{(k)} = \frac{\ell^{(k+1)} - \ell^{(k)}}{\ell^{(k)} - \ell^{(k-1)}}.$$

El procedimiento mostrado arriba es también aplicable para un caso más simple, es decir ($G = 1$) considerando $Z_{ij} = 1$.

3.4.4. MATRIZ DE INFORMACIÓN OBSERVADA

En esta sección se presenta el procedimiento utilizado para la obtención de los errores estándar de los EMV del modelo MF-BS-NA, los cuales son usualmente calculados derivando la función log-verosimilitud dos veces y obteniendo la inversa. Sin embargo, esto es algo complejo de llevar a cabo en un modelo de mixtura finita. Por ello, asumiendo las condiciones de regularidad usuales, las cuales garantizan que los EMV resuelven la ecuación gradiente y que existe la información de Fisher de acuerdo con Louis (1982). Así, los estimadores de la variancia se obtienen a partir de la diagonal de la inversa de la matriz de información empírica, definida como:

$$\mathbf{I}_o = \sum_{i=1}^n s(y_i|\boldsymbol{\theta})s^\top(y_i|\boldsymbol{\theta}) - n^{-1}S(\mathbf{y}|\boldsymbol{\theta})S^\top(\mathbf{y}|\boldsymbol{\theta}), \quad (25)$$

donde $S(\mathbf{y}|\boldsymbol{\theta}) = \sum_{i=1}^n s(y_i|\boldsymbol{\theta})$, con $s(y_i|\boldsymbol{\theta}) = \partial \log f(y_i|\boldsymbol{\theta})/\partial \boldsymbol{\theta}$ siendo la función score empírica para el i -ésimo individuo. Sustituyendo los EMV $\widehat{\boldsymbol{\theta}}$ por $\boldsymbol{\theta}$ en (25), \mathbf{I}_o se reduce a

$$\mathbf{I}_o = \sum_{i=1}^n \widehat{\mathbf{s}}_i \widehat{\mathbf{s}}_i^\top, \quad (26)$$

donde $\widehat{\mathbf{s}}_i$ es un vector score individual dado por $\widehat{\mathbf{s}}_i = (\widehat{s}_{i,p_1}, \dots, \widehat{s}_{i,p_{G-1}}, \widehat{s}_{i,\alpha_1}, \dots, \widehat{s}_{i,\alpha_G}, \widehat{s}_{i,\beta_1}, \dots, \widehat{s}_{i,\beta_G}, \widehat{s}_{i,\lambda_1}, \dots, \widehat{s}_{i,\lambda_G})^\top$. Las expresiones explícitas de los elementos de $\widehat{\mathbf{s}}_i$ son dadas por

$$\widehat{s}_{i,p_j} = \frac{f_{T_j}(y_i; \alpha_j, \beta_j, \lambda_j) - f_{T_G}(y_i; \alpha_G, \beta_G, \lambda_G)}{f_{Y_i}(y_i; \mathbf{p}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda})}, \quad \widehat{s}_{i,\alpha_j} = \frac{p_j D_{\alpha_j}(f_{T_j}(y_i; \alpha_j, \beta_j, \lambda_j))}{f_{Y_i}(y_i; \mathbf{p}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda})}$$

$$\widehat{s}_{i,\beta_j} = \frac{p_j D_{\beta_j}(f_{T_j}(y_i; \alpha_j, \beta_j, \lambda_j))}{f_{Y_i}(y_i; \mathbf{p}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda})}, \quad \widehat{s}_{i,\lambda_j} = \frac{p_j D_{\lambda_j}(f_{T_j}(y_i; \alpha_j, \beta_j, \lambda_j))}{f_{Y_i}(y_i; \mathbf{p}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda})},$$

donde

$$D_{\alpha_j}(f_{T_j}(y_i; \alpha_j, \beta_j, \lambda_j)) = \frac{\partial f_{T_j}(y_i; \alpha_j, \beta_j, \lambda_j)}{\partial \alpha_j}, \quad D_{\beta_j}(f_{T_j}(y_i; \alpha_j, \beta_j, \lambda_j)) = \frac{\partial f_{T_j}(y_i; \alpha_j, \beta_j, \lambda_j)}{\partial \beta_j}$$

$$y \quad D_{\lambda_j}(f_{T_j}(y_i; \alpha_j, \beta_j, \lambda_j)) = \frac{\partial f_{T_j}(y_i; \alpha_j, \beta_j, \lambda_j)}{\partial \lambda_j}.$$

Para simplificar la notación, omitimos el índice i en las expresiones con la finalidad de evitar cualquier confusión:

$$\begin{aligned} D_{\alpha_j}(f_{T_j}(y_i; \alpha_j, \beta_j, \lambda_j)) &= 2 \left\{ A_y(\alpha_j, \beta_j) \frac{\partial a_y(\alpha_j, \beta_j)}{\partial \alpha_j} \phi(a_y(\alpha_j, \beta_j)) [\lambda_j \phi(\lambda_j a_y(\alpha_j, \beta_j)) \right. \\ &\quad \left. - a_y(\alpha_j, \beta_j) \Phi(\lambda_j a_y(\alpha_j, \beta_j))] + \frac{\partial A_y(\alpha_j, \beta_j)}{\partial \alpha_j} \phi(a_y(\alpha_j, \beta_j)) \right. \\ &\quad \left. \Phi(\lambda_j a_y(\alpha_j, \beta_j)) \right\} \\ D_{\beta_j}(f_{T_j}(y_i; \alpha_j, \beta_j, \lambda_j)) &= 2 \left\{ A_y(\alpha_j, \beta_j) \frac{\partial a_y(\alpha_j, \beta_j)}{\partial \beta_j} \phi(a_y(\alpha_j, \beta_j)) [\lambda_j \phi(\lambda_j a_y(\alpha_j, \beta_j)) \right. \\ &\quad \left. - a_y(\alpha_j, \beta_j) \Phi(\lambda_j a_y(\alpha_j, \beta_j))] + \frac{\partial A_y(\alpha_j, \beta_j)}{\partial \beta_j} \phi(a_y(\alpha_j, \beta_j)) \right. \\ &\quad \left. \Phi(\lambda_j a_y(\alpha_j, \beta_j)) \right\} \\ D_{\lambda_j}(f_{T_j}(y_i; \alpha_j, \beta_j, \lambda_j)) &= 2 a_y(\alpha_j, \beta_j) A_y(\alpha_j, \beta_j) \phi(a_y(\alpha_j, \beta_j)) \phi(\lambda_j a_y(\alpha_j, \beta_j)), \end{aligned}$$

donde

$$\begin{aligned} \frac{\partial a_y(\alpha_j, \beta_j)}{\partial \alpha_j} &= -\frac{1}{\alpha_j} a_y(\alpha_j, \beta_j), \quad \frac{\partial A_y(\alpha_j, \beta_j)}{\partial \alpha_j} = -\frac{1}{\alpha_j} A_y(\alpha_j, \beta_j), \\ \frac{\partial a_y(\alpha_j, \beta_j)}{\partial \beta_j} &= -\frac{1}{2\alpha_j \beta_j} \left(\sqrt{\frac{y}{\beta_j}} + \sqrt{\frac{\beta_j}{y}} \right), \quad \frac{\partial A_y(\alpha_j, \beta_j)}{\partial \beta_j} = \frac{y^{-3/2}(\beta_j - y)}{4\alpha_j \beta_j^{1/2}}. \end{aligned}$$

Los errores estándar de $\hat{\boldsymbol{\theta}}$ se extraen de la raíz cuadrada de los elementos diagonales de la inversa de la ecuación (26). La aproximación obtenida de la matriz de información observada es asintóticamente aplicable.

3.4.5. SELECCIÓN DE MODELOS

Consideramos el problema de comparar varios modelos MF-BS-NA, con diferentes números de componentes. Para ello usaremos dos criterios de selección de modelos el Criterio de Información Akaike (AIC) (Akaike, 1974), el Criterio de Información Bayesiano (BIC) (Schwarz, 1978). Estos criterios son definidos como

$$AIC = -2\ell(\hat{\boldsymbol{\theta}}) + 2\rho \quad y \quad BIC = -2\ell(\hat{\boldsymbol{\theta}}) + \rho \log(n),$$

donde $\ell(\boldsymbol{\theta})$ es la actual log-verosimilitud, ρ es el número de parámetros libres que serán estimados en el modelo, y n es el tamaño de muestra.

IV. RESULTADOS Y DISCUSIÓN

4.1. ESTUDIO DE SIMULACIÓN

Con la finalidad de examinar el desempeño de los EMV, realizaremos tres estudios de simulación. Para realizar estos experimentos numéricos, utilizamos el entorno informático estadístico R (R Core Team , 2016). El primer estudio tiene como objetivo evaluar la exactitud de los estimadores obtenidos usando el algoritmo EM para el modelo MF-BS-NA y a su vez evaluar la consistencia de los errores estándar. El segundo estudio tiene como finalidad demostrar que el algoritmo propuesto proporciona estimadores con buenas propiedades asintóticas. En el tercer estudio se compara el desempeño de los estimadores del modelo MF-BS-NA en presencia de observaciones atípicas. En los tres casos, los datos de la simulación se generaron artificialmente a partir del siguiente modelo con dos componentes ($G = 2$):

$$f_Y(y; \mathbf{p}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda}) = p_1 f_{T_1}(y; \alpha_1, \beta_1, \lambda_1) + p_2 f_{T_2}(y; \alpha_2, \beta_2, \lambda_2). \quad (27)$$

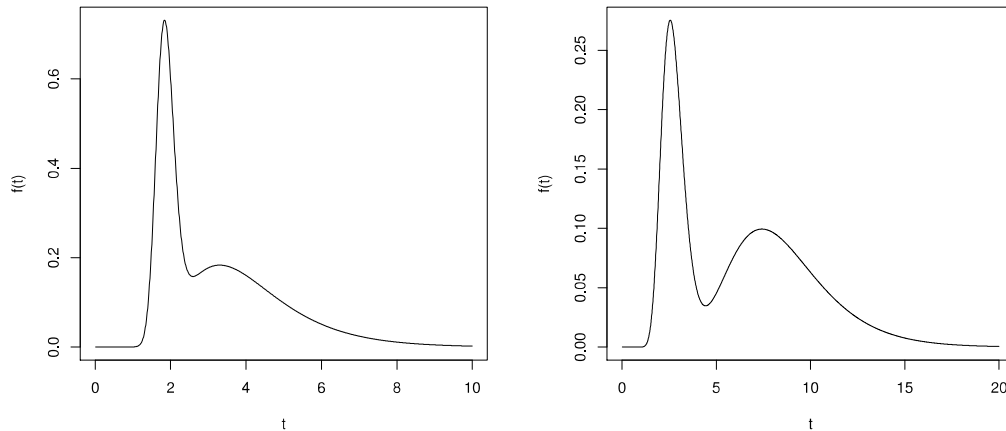


Figura 7: Funciones de densidad a partir de las cuales se realizaron las simulaciones (panel izquierdo) $\boldsymbol{\theta} = (0.4, 0.15, 0.45, 2, 3, -2, 4)$ y (panel derecho) $\boldsymbol{\theta} = (0.4, 0.25, 0.35, 3, 7, -3, 3)$.

Los estudios de simulación se realizaron utilizando un procesador Intel Core i7 de 3,40 GHz con 31,9 GB de RAM. La Figura 7 (izquierda) y (derecha) muestra las funciones de densidad generadas usando los escenarios 1 y 2, respectivamente.

4.1.1. ESTUDIO 1: RECUPERACIÓN DE LOS PARÁMETROS Y CONSISTENCIA DE LOS ERRORES ESTÁNDAR DE LOS ESTIMADORES

En este estudio de simulación se obtienen los EMV de los parámetros de cada muestra a través del algoritmo EM usando la regla de parada descrita en (3.4.3). Para ello generamos 1000 muestras aleatorias Montecarlo de diversos tamaños de muestra $n = 100, 250, 500, 1000$ y 2500 del modelo MF-BS-NA considerando dos escenarios, el primero con componentes deficientemente separados (DS) y el segundo con componentes bien separados (BS):

Escenario 1: $p_1 = 0.4, \alpha_1 = 0.15, \alpha_2 = 0.45, \beta_1 = 2, \beta_2 = 3, \lambda_1 = -2$ y $\lambda_2 = 4$ (componentes DS);

Escenario 2: $p_1 = 0.4, \alpha_1 = 0.25, \alpha_2 = 0.35, \beta_1 = 3, \beta_2 = 7, \lambda_1 = -3$ y $\lambda_2 = 3$ (componentes BS).

Puede notarse en los Cuadros 1 y 2 que las estimaciones de los parámetros se aproximan a los valores verdaderos de los parámetros a medida que aumenta el tamaño de muestra. Se puede apreciar que cuando los componentes se encuentran bien separados, el algoritmo EM recupera bien los parámetros. A su vez, examinamos la consistencia del método de aproximación sugerido para la obtención de los errores estándar (SE) de los EMV de los parámetros $\theta = (\alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2, p_1)$. También calculamos la media de los errores estándar aproximados (IM SE) obtenidos mediante la metodología descrita en (3.4.4) y el porcentaje de cobertura que resulta de los intervalos de confianza del 95 % (COV) asumiendo normalidad asintótica. La COV se define como

$$COV(\hat{\theta}) = \frac{1}{m} \sum_{j=1}^m I(\theta \in [\hat{\theta}_L, \hat{\theta}_U])$$

donde I es la función indicadora tal que θ se encuentra en el intervalo, con $\hat{\theta}_L$ y $\hat{\theta}_U$ siendo los límites inferiores y superiores estimados del IC del 95 %, respectivamente. La COV para los parámetros es bastante estable para ambos escenarios, lo que indica que la aproximación asintótica propuesta para la variancia de los EMV es confiable. En general, los resultados sugieren que el modelo MF-BS-NA propuesto produce estimaciones satisfactorias, como se esperaba. También vemos en los Cuadros 1 y 2 que el método de estimación de los errores estándar proporciona resultados relativamente cercanos (IM SE y MC Sd), lo que indica que la aproximación asintótica propuesta para las variancias de los EMV también es confiable. Esto también se puede ver en la probabilidad de cobertura de los parámetros (COV), ya que en general se mantiene un intervalo de confianza por encima del 95 % de cobertura para cada parámetro.

4.1.2. ESTUDIO 2: PROPIEDADES ASINTÓTICAS DE LOS ESTIMADORES

En esta sección el objetivo es mostrar las propiedades asintóticas de los EMV. Nuestra estrategia es generar muestras artificiales para el modelo definido en (27). Para ello hemos utilizado varios tamaños de muestra ($n = 75, 100, 500, 1000, 5000, 10000, 20000, 25000, 30000$ y 40000). Los valores verdaderos de los parámetros en este estudio son los mismos considerados en el estudio 1 (escenario 2). Para cada combinación de parámetros y tamaños de muestra, se generaron 1000 muestras aleatorias del modelo MF-BS-NA. Para evaluar las propiedades asintóticas de los estimadores obtenidos por el algoritmo propuesto, se comparó el sesgo (Bias) y el error cuadrático medio (RMSE) para cada parámetro en las 1000 repeticiones. Los cuales se definen de la siguiente forma

$$\text{Bias}(\theta_i) = \frac{1}{1000} \sum_{j=1}^{1000} (\hat{\theta}_i^{(j)} - \theta_i) \quad \text{and} \quad \text{RMSE}(\theta_i) = \sqrt{\frac{1}{1000} \sum_{j=1}^{1000} (\hat{\theta}_i^{(j)} - \theta_i)^2},$$

donde $\hat{\theta}^{(j)}$ es la estimación de θ_i para la j -ésima muestra. Los resultados para $\alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1, \lambda_2$ y p_1 , son mostrados en la Figura 8 y 9. Como regla general, podemos decir que el sesgo y el error cuadrático medio tienden a acercarse a cero cuando el tamaño de la muestra aumenta, lo que indica que los estimadores obtenidos usando el algoritmo EM bajo el modelo MF-BS-NA proporcionan buenas propiedades asintóticas.

Cuadro 1: Media estimada, errores estándar y porcentaje de la cobertura de probabilidad para el escenario 1. Los valores teóricos de los parámetros están entre paréntesis.

| Parámetro | Medidas | Tamaño de muestra | | | | |
|-------------------|---------|-------------------|-----------|-----------|------------|------------|
| | | $n = 100$ | $n = 250$ | $n = 500$ | $n = 1000$ | $n = 2500$ |
| α_1 (0.15) | Mean | 0.1404 | 0.1448 | 0.1481 | 0.1492 | 0.1495 |
| | IM SE | 0.0627 | 0.0247 | 0.0153 | 0.0102 | 0.0063 |
| | MC Sd | 0.0283 | 0.0179 | 0.0130 | 0.0094 | 0.0058 |
| | COV | 99.5 % | 98.9 % | 98.8 % | 97.6 % | 96.2 % |
| α_2 (0.45) | Mean | 0.4345 | 0.4457 | 0.4471 | 0.4484 | 0.4497 |
| | IM SE | 0.0670 | 0.0381 | 0.0266 | 0.0186 | 0.0117 |
| | MC Sd | 0.0569 | 0.0352 | 0.0244 | 0.0175 | 0.0105 |
| | COV | 97.1 % | 96.3 % | 96.5 % | 95.1 % | 96.4 % |
| β_1 (2) | Mean | 1.9694 | 1.9862 | 1.9948 | 1.9976 | 1.9981 |
| | IM SE | 0.2761 | 0.0690 | 0.0378 | 0.0243 | 0.0149 |
| | MC Sd | 0.0705 | 0.0455 | 0.0312 | 0.0222 | 0.0134 |
| | COV | 99.1 % | 99.0 % | 98.7 % | 97.6 % | 97.4 % |
| β_2 (3) | Mean | 3.0458 | 3.0127 | 3.0066 | 3.0025 | 3.0011 |
| | IM SE | 0.2070 | 0.0998 | 0.0697 | 0.0486 | 0.0306 |
| | MC Sd | 0.1723 | 0.0924 | 0.0651 | 0.0443 | 0.0288 |
| | COV | 95.1 % | 96.1 % | 96.3 % | 95.7 % | 96.3 % |
| λ_1 (-2) | Mean | -1.8080 | -1.9436 | -2.0127 | -2.0254 | -1.9928 |
| | IM SE | 3.1055 | 1.1068 | 0.6921 | 0.4629 | 0.2821 |
| | MC Sd | 0.9976 | 0.6882 | 0.5279 | 0.3994 | 0.2472 |
| | COV | 99.7 % | 99.0 % | 99.0 % | 99.0 % | 96.4 % |
| λ_2 (4) | Mean | 4.8855 | 4.4225 | 4.1420 | 4.0702 | 4.0373 |
| | IM SE | 3.9934 | 1.6994 | 1.0496 | 0.7072 | 0.4374 |
| | MC Sd | 3.1097 | 1.6522 | 1.0020 | 0.6307 | 0.3693 |
| | COV | 98.8 % | 96.7 % | 96.2 % | 97.5 % | 97.0 % |
| p_1 (0.4) | Mean | 0.3975 | 0.3994 | 0.3992 | 0.3994 | 0.4003 |
| | IM SE | 0.0491 | 0.0310 | 0.0220 | 0.0155 | 0.0098 |
| | MC Sd | 0.0462 | 0.0314 | 0.0224 | 0.0150 | 0.0100 |
| | COV | 97.1 % | 94.2 % | 94.4 % | 95.8 % | 94.2 % |

Cuadro 2: Media estimada, errores estándar y porcentaje de la cobertura de probabilidad para el escenario 2. Los valores teóricos de los parámetros están entre paréntesis.

| Parámetro | Medidas | Tamaño de muestra | | | | |
|-------------------|---------|-------------------|-----------|-----------|------------|------------|
| | | $n = 100$ | $n = 250$ | $n = 500$ | $n = 1000$ | $n = 2500$ |
| α_1 (0.25) | Mean | 0.2334 | 0.2448 | 0.2468 | 0.2493 | 0.2494 |
| | IM SE | 0.0628 | 0.0296 | 0.0200 | 0.0137 | 0.0086 |
| | MC Sd | 0.0414 | 0.0252 | 0.0178 | 0.0124 | 0.0080 |
| | COV | 98.4% | 97.0% | 97.3% | 96.3% | 96.3% |
| α_2 (0.35) | Mean | 0.3351 | 0.3451 | 0.3480 | 0.3501 | 0.3498 |
| | IM SE | 0.0604 | 0.0324 | 0.0226 | 0.0158 | 0.0099 |
| | MC Sd | 0.0482 | 0.0299 | 0.0204 | 0.0146 | 0.0095 |
| | COV | 97.4% | 96.4% | 96.7% | 95.7% | 95.9% |
| β_1 (3) | Mean | 2.9425 | 2.9835 | 2.9919 | 2.9979 | 2.9982 |
| | IM SE | 0.2408 | 0.0898 | 0.0593 | 0.0402 | 0.0251 |
| | MC Sd | 0.1316 | 0.0760 | 0.0531 | 0.0352 | 0.0234 |
| | COV | 99.9% | 98.8% | 98.2% | 97.0% | 96.5% |
| β_2 (7) | Mean | 7.1153 | 7.0345 | 7.0125 | 7.0016 | 7.0031 |
| | IM SE | 0.5123 | 0.2242 | 0.1546 | 0.1074 | 0.0675 |
| | MC Sd | 0.3795 | 0.2085 | 0.1388 | 0.0977 | 0.0636 |
| | COV | 98.0% | 96.9% | 96.8% | 96.3% | 96.7% |
| λ_1 (-3) | Mean | -2.7421 | -2.9785 | -3.0108 | -3.0537 | -3.0159 |
| | IM SE | 2.5654 | 1.2910 | 0.8574 | 0.5853 | 0.3516 |
| | MC Sd | 1.2098 | 0.8325 | 0.6641 | 0.4889 | 0.3129 |
| | COV | 99.9% | 98.3% | 96.1% | 97.6% | 96.8% |
| λ_2 (3) | Mean | 3.2349 | 3.1832 | 3.0795 | 3.0547 | 3.0181 |
| | IM SE | 2.1480 | 1.0840 | 0.6956 | 0.4709 | 0.2880 |
| | MC Sd | 1.5557 | 0.9427 | 0.6187 | 0.4010 | 0.2623 |
| | COV | 99.9% | 98.1% | 97.2% | 97.5% | 96.4% |
| p_1 (0.4) | Mean | 0.3991 | 0.4002 | 0.4000 | 0.4002 | 0.4002 |
| | IM SE | 0.0488 | 0.0309 | 0.0219 | 0.0155 | 0.0098 |
| | MC Sd | 0.0477 | 0.0312 | 0.0213 | 0.0159 | 0.0098 |
| | COV | 96.6% | 94.8% | 95.6% | 92.5% | 94.0% |

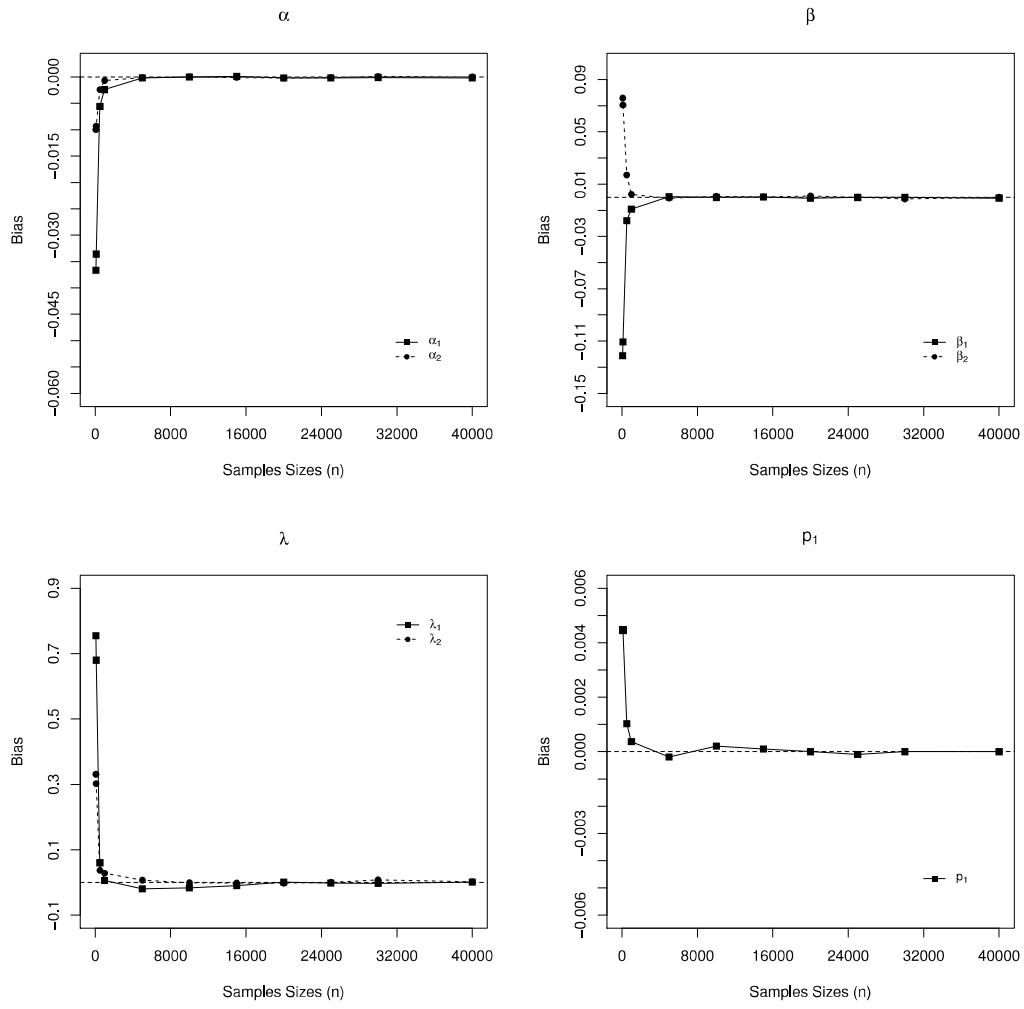


Figura 8: Promedio de sesgo para todos los estimadores.

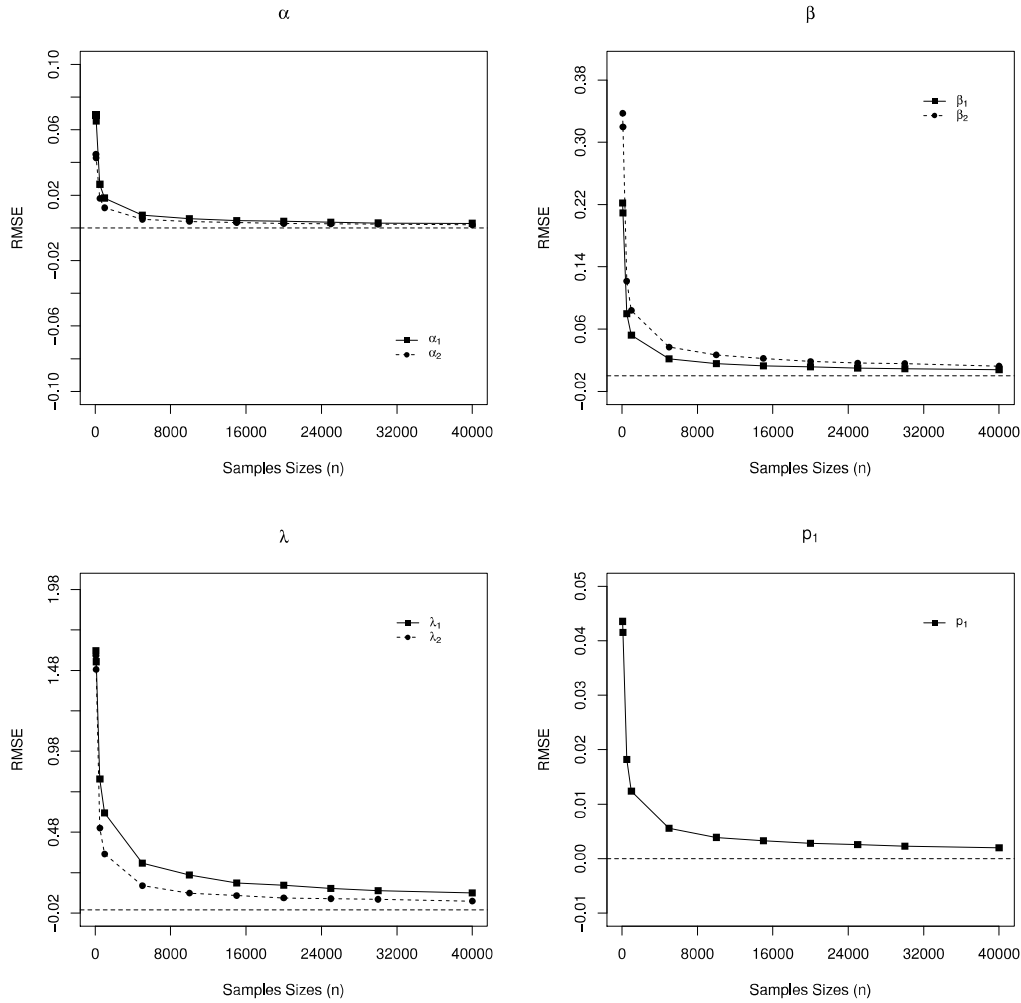


Figura 9: Promedio de la raíz del error cuadrático medio para todos los estimadores.

4.1.3. ESTUDIO 3: PROPIEDADES DE ROBUSTEZ DE LOS ESTIMADORES

El propósito de este estudio de simulación es comparar el efecto de la robustez de los estimadores del modelo MF-BS-NA, a través de la influencia de una sola observación atípica sobre los EMV de los parámetros del modelo. Para ello simulamos un conjunto de datos del modelo MF-BS-NA usando los parámetros utilizados en el escenario 2 del estudio de simulación 1 y lo comparamos con el modelo MF-BS. Para ambos casos, se generaron 1000 muestras de tamaño $n = 1000$, y evaluamos cuanto los estimadores son influenciadas por la presencia de valores atípicos, reemplazando una sola observación t_{150} por $t_{150} = t_{150} + \vartheta$, con $\vartheta = 1, \dots, 10$. Para cada replica, se obtuvieron las estimadores de los parámetros con y sin valores atípicos, en todos los modelos. Se evaluó el cambio relativo (CR) en los estimadores de $\theta = (p_1, \alpha_1, \alpha_2, \beta_1, \beta_2)^\top$ mediante

$$CR(\hat{\theta}_i(\vartheta)) = \left| \frac{\hat{\theta}_i(\vartheta) - \hat{\theta}_i}{\hat{\theta}_i} \right|,$$

donde $\hat{\theta}_i(\vartheta)$ y $\hat{\theta}_i$ denotan los EMV de θ_i con y sin perturbación, respectivamente. Las Figuras 10 y 11 muestran los valores promedio de los cambios relativos sufridos por todos los parámetros. Observamos que para todos los parámetros, los cambios relativos promedio aumentan repentinamente en el modelo MF-BS a medida que crece el valor ϑ . Por el contrario, para el modelo MF-BS-NA, las medidas varían en menor proporción, indicando que son más robustas que el modelo MF-BS en la capacidad de acomodar observaciones discrepantes.

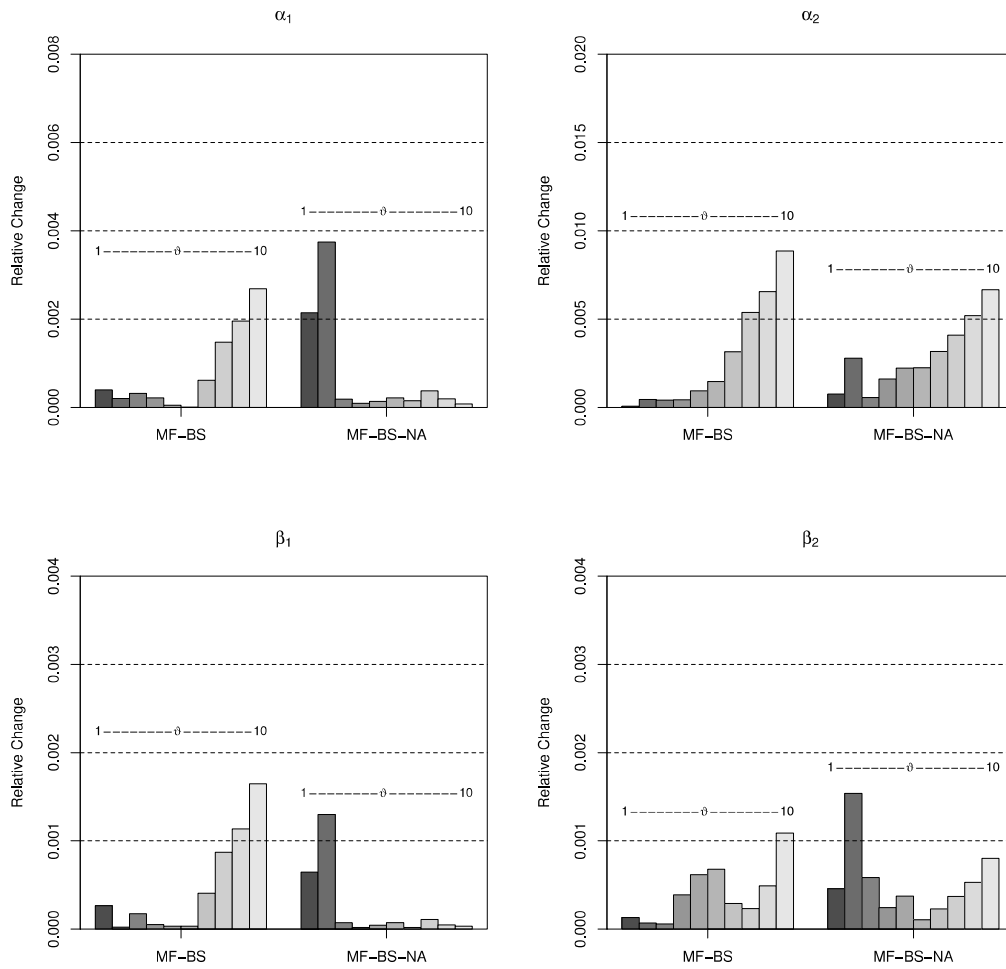


Figura 10: Cambios relativos promedio de los estimadores de α_1 , α_2 , β_1 y β_2 para diferentes perturbaciones ϑ .

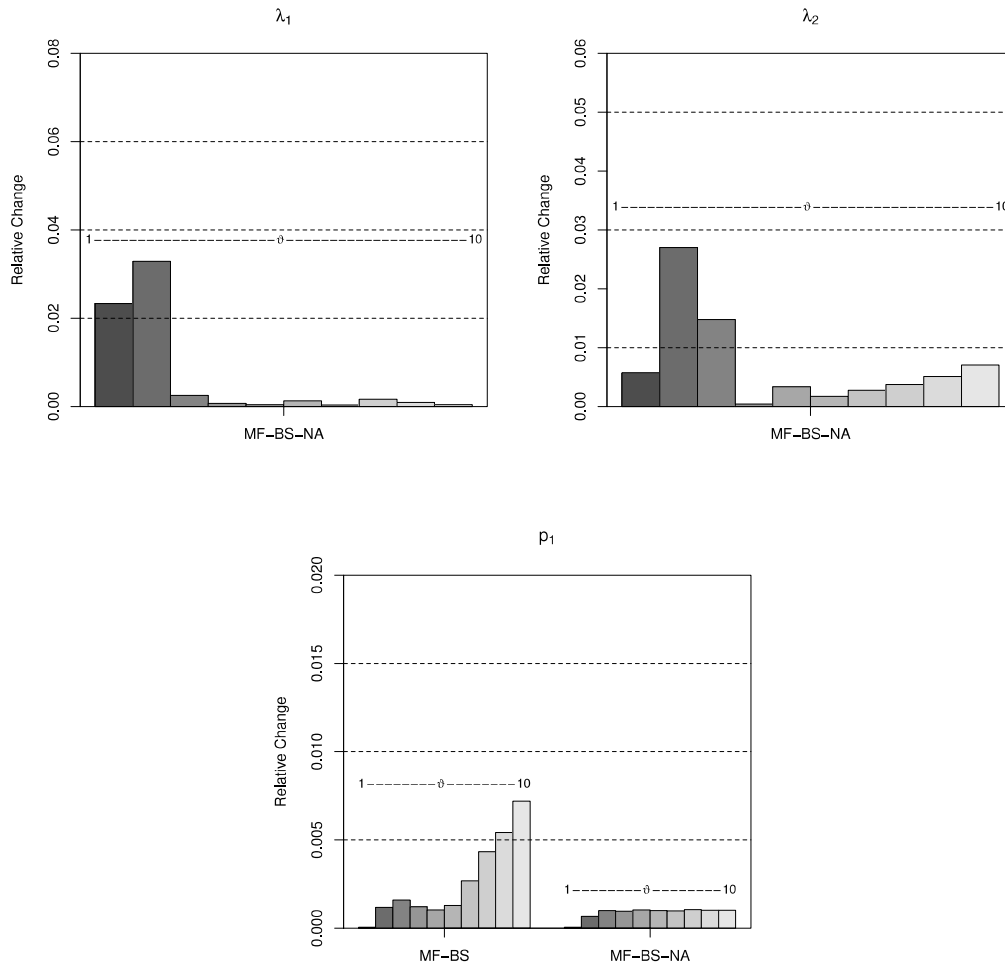


Figura 11: Cambios relativos promedio de los estimadores de λ_1 , λ_2 y p_1 para diferentes perturbaciones ϑ .

4.2. APLICACIÓN 1

En la primera aplicación consideraremos el índice masa corporal (IMC) para hombres de edades comprendidas entre 18 y 80 años. El conjunto de datos proviene de la Encuesta Nacional de Exámenes de Salud y Nutrición, realizada por el Centro Nacional de Estadísticas de Salud (NCHS) del Centro para el Control de Enfermedades (CDC) en los Estados Unidos. El problema de la obesidad ha atraído la atención en los últimos años debido a su fuerte relación con muchas enfermedades crónicas. El índice de masa corporal (IMC, Kg/m^2) se ha convertido en la medida estándar para el sobrepeso y la obesidad. El IMC se define como la relación entre el peso en kilogramos y la altura en metros al cuadrado. Originalmente el conjunto de datos tenía 4579 participantes con registros de IMC entre 1999-2000 y 2001-2002, sin embargo en el trabajo de Lin *et al.* (2007) con el objetivo de poder explorar los patrones de la mixtura, consideraron solamente pesos que estaban dentro de los intervalos de [39.50kg, 70.00kg] y [95.01kg, 196.80kg], por lo que solo quedaron 1069 participantes para el primer grupo y 1054 para segundo grupo.

Este conjunto de datos ha sido también analizado por Benites *et al.* (2017). El Cuadro 3 muestra una comparación entre el modelo de MF-BS-NA y el MF-BS según el criterio de AIC y BIC, la log-verosimilitud y las iteraciones hasta alcanzar la convergencia de los estimadores, donde consideramos un algoritmo k-means para obtener los valores iniciales. Notamos que los dos indicadores, AIC y BIC, mostraron que el modelo MF-BS-NA con $G = 3$ componentes es el que mejor se ajusta al conjunto de datos, indicando que el modelo propuesto se ajusta mejor que el modelo presentado en Benites *et al.* (2017), caso simétrico, usando el algoritmo k-médias para obtener los valores iniciales permitiendo modelar mejor los percentiles extremos especialmente en la cola izquierda de la distribución. Observando el valor de la log-verosimilitud el mayor valor es obtenido para el modelo MF-BS-NA con $G = 3$ componentes. Notamos que el número de iteraciones para el modelo simétrico (MF-BS) son bastante menores que el asimétrico (MF-BS-NA) y esto se debe a que el modelo MF-BS es una modelo más simple.

Cuadro 3: Comparación de la máxima log-verosimilitud, AIC y BIC para el ajuste de los modelos MF-BS y MF-BS-NA usando el conjunto de datos del índice de masa corporal. El número de parámetros es denotado por m .

| | G | m | log-vero | AIC | BIC | Iteraciones |
|----------|-----|-----|-----------|-----------------|-----------------|-------------|
| MF-BS | 1 | 2 | -7099.455 | 14202.91 | 14214.22 | 1 |
| MF-BS | 2 | 5 | -6953.946 | 13917.89 | 13946.16 | 9 |
| MF-BS | 3 | 8 | -7098.833 | 14213.67 | 14258.89 | 8 |
| MF-BS-EN | 1 | 3 | -7418.912 | 15099.73 | 15116.69 | 335 |
| MF-BS-EN | 2 | 7 | -7227.093 | 13828.41 | 13867.98 | 116 |
| MF-BS-EN | 3 | 11 | -7128.554 | 13745.55 | 13807.73 | 37 |

Ahora vamos a comparar el ajuste de los modelos MF-BS-NA y MF-BS al conjunto de datos. La Figura 12 (b) muestra que el conjunto de datos presenta una asimetría positiva y que a su vez presenta valores atípicos. La Figura 12 (a) muestra el histograma del conjunto de datos con las densidades ajustadas usando el algoritmo EM tanto para el modelo MF-BS como para el modelo MF-BS-NA. Puede notarse en el gráfico, que el modelo MF-BS-NA con $G = 2$ y $G = 3$ tiene un ajuste bastante razonable comparado con el modelo de MF-BS. Cabe resaltar que para este caso el modelo asimétrico con $G = 2$ y $G = 3$ se ajusta mejor al conjunto de datos tanto en la cola izquierda como en la región central.

La Figura 13 (a) muestra la fda para el modelo MF-BS-NA considerando $G = 1 - 3$. Se puede apreciar que efectivamente se obtiene un mejor ajuste al conjunto de datos cuando tomamos en cuenta $G = 2$ o $G = 3$. La Figura 13 (b) confirma nuevamente que el modelo de MF-BS-NA con $G = 2$ y $G = 3$ componentes se ajusta mejor al conjunto de datos cuando se usa la función empírica de supervivencia.

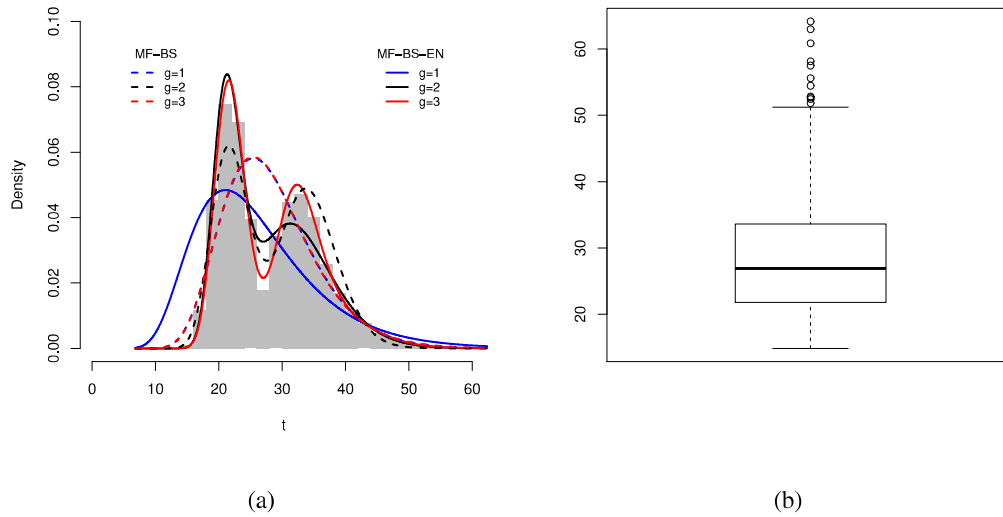


Figura 12: (a) Histograma del conjunto de datos del IMC con ajustes de la MF-BS-NA y MF-BS con $G = 1 - 3$ componentes. (b) Gráfico de cajas del IMC.

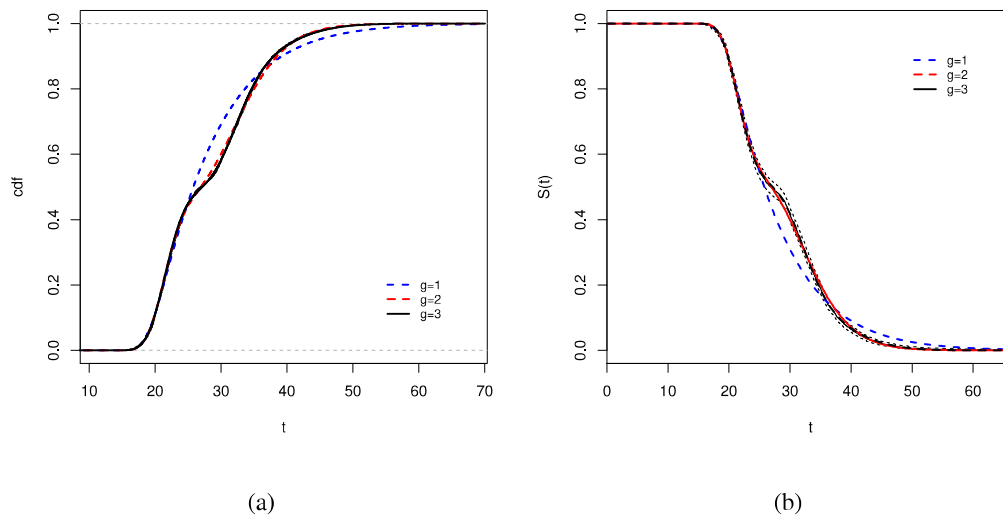


Figura 13: (a) Función de distribución acumulada y (b) función de supervivencia estimada con la función empírica de supervivencia para el ajuste de 3 modelos de la MF-BS-NA con $G = 1 - 3$ componentes.

Por lo tanto, tomando en cuenta los resultados basados en los indicadores AIC, BIC y también las figuras mostradas arriba concluimos que modelo que se ajusta mejor al conjunto de datos es la MF-BS-NA con $G = 3$ componentes. El Cuadro 4 presenta los EMV de $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, p_1$ y p_2 para el modelo de MF-BS-NA con tres componentes con sus correspondientes errores estándar (EE), calculados por medio del método de aproximación presentado en la Sección 3.4.4.

Cuadro 4: EMV con sus correspondientes EE para el modelo de MF-BS-NA aplicados al conjunto de datos IMC.

| Parámetro | EMV | EE |
|-------------|---------|--------|
| α_1 | 0.1095 | 0.0837 |
| α_2 | 0.1621 | 0.0514 |
| α_3 | 0.1192 | 0.0174 |
| β_1 | 33.7107 | 5.7852 |
| β_2 | 36.1252 | 6.1771 |
| β_3 | 22.8075 | 1.0172 |
| λ_1 | -0.6473 | 3.5941 |
| λ_2 | 0.7077 | 8.8646 |
| λ_3 | -0.6345 | 0.7586 |
| p_1 | 0.3583 | 0.7085 |
| p_2 | 0.1500 | 0.7215 |

La Figura 14 (a) muestra el histograma del conjunto de datos con las respectivas funciones de densidad de probabilidad superpuestas considerando $G = 1 - 3$ componentes, así como los EMV para cada densidad. La Figura 14 (b) muestra el mismo histograma con la fdp para tres componentes y a su vez las funciones de densidad de probabilidad para cada componente con sus respectivos EMV.

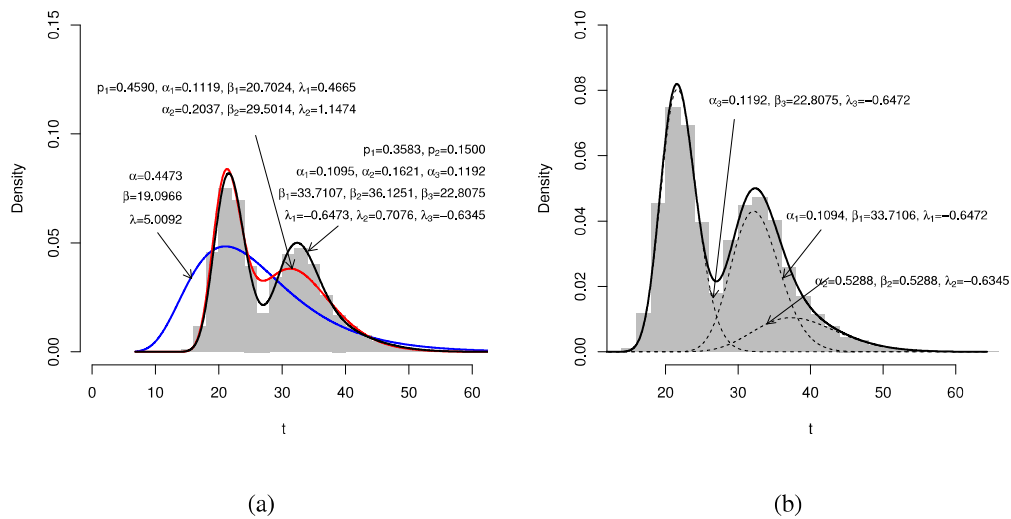


Figura 14: Histograma del IMC superpuesta con el ajuste de las densidades del modelo de MF-BS-NA (a) para $G = 1 - 3$ componentes y (b) $G = 3$ componentes.

4.3. APLICACIÓN 2

Para la segunda aplicación es aplicado el método propuesto para el conjunto de datos correspondiente a la actividad enzimática en la sangre y representando el metabolismo de sustancias carcinógenas entre 245 individuos no relacionados que fueron estudiados previamente por Bechtel *et al.* (1993), que ajusta una mixtura de dos distribuciones asimétricas, en Balakrishnan *et al.* (2011), consideraron tres modelos de mixtura diferentes basados en los modelos BS y longitud-sesgada, y en Benites *et al.* (2017) para una MF-BS. En este caso usaremos el algoritmo EM descrito en la Sección 3.4.1 para obtener los EMV para el modelo MF-BS-NA para compararlos con los obtenidos usando la MF-BS. El Cuadro 5 permite comparar el modelo de MF-BS-NA y el MF-BS usando los criterios AIC y BIC, la log-verosimilitud y el número de iteraciones necesarias para alcanzar la convergencia de los estimadores. Notamos que usando el criterio AIC el modelo que se ajusta mejor al conjunto de datos es la MF-BS con $G = 3$ componentes, sin embargo cuando tomamos en cuenta el BIC tenemos que el mejor modelo es una MF-BS-NA con $G = 2$ componentes. Podemos notar que usando una MF-BS-NA con $G = 2$ componentes tenemos menor número de parámetros, así de acuerdo al principio de parsimonia elegimos este modelo como el que mejor se ajusta al conjunto de datos. Cabe mencionar que el modelo propuesto en este trabajo resultó ser mejor, en términos del BIC, que el propuesto en Benites *et al.* (2017) que usa el algoritmo k-médias para obtener los valores iniciales, ya que el nuestro es parsimonioso y permite modelar mejor los percentiles extremos.

Cuadro 5: Comparación de la máxima log-verosimilitud, AIC y BIC para el ajuste de los modelos MF-BS y MF-BS-NA usando el conjunto de datos de enzima. El número de parámetros es denotado por m .

| | G | m | log-vero | AIC | BIC | Iteraciones |
|----------|-----|-----|-----------|-----------------|-----------------|-------------|
| MF-BS | 1 | 2 | -105.5071 | 215.0141 | 222.0167 | 2 |
| MF-BS | 2 | 5 | -91.3855 | 192.7711 | 210.2774 | 6 |
| MF-BS | 3 | 8 | -51.9166 | 119.8333 | 147.4342 | 19 |
| MF-BS-EN | 1 | 3 | -120.4081 | 215.9663 | 226.4701 | 40 |
| MF-BS-EN | 2 | 7 | -92.4439 | 121.3128 | 145.8216 | 47 |
| MF-BS-EN | 3 | 11 | -84.1544 | 127.1514 | 165.6419 | 43 |

Podemos apreciar en la Figura 15 (b) que el conjunto de datos presenta asimetría positiva y algunos valores atípicos. En la Figura 15 (a) se puede ver que el modelo MF-BS-NA con $G = 2$ componentes presenta un mejor ajuste especialmente en la cola derecha de la distribución comparada con el modelo simétrico MF-BS.

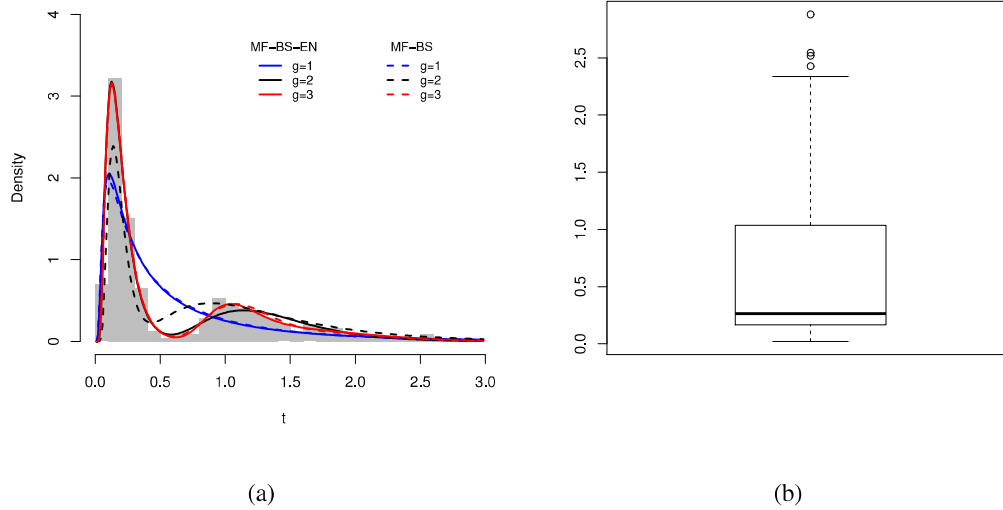


Figura 15: (a) Histograma del conjunto de datos de Enzima con ajustes de la MF-BS-NA y MF-BS con $G = 1 - 3$ componentes. (b) Gráfico de cajas del conjunto de datos.

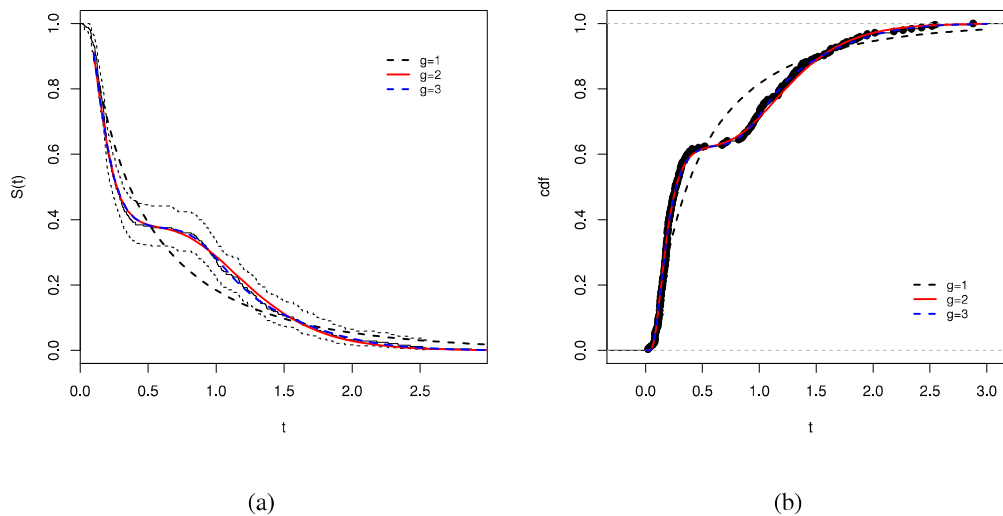


Figura 16: (a) Función de supervivencia estimada con la función empírica de supervivencia y (b) función de distribución acumulada para el ajuste de tres modelos MF-BS-NA con $G = 1 - 3$ componentes.

Notamos en la Figura 16 (a) y (b) que el comportamiento de la función de supervivencia y distribución acumulada acompaña al comportamiento teórico de dichas curvas, lo que significa que la elección MF-BS-NA con $G = 2$ componentes es adecuado.

En la Figura 17 (a) se muestra el ajuste para la MF-BS-NA con $G = 2$ componentes al histograma del conjunto de datos y en la Figura 17 (b) se muestra las curvas de cada componente.

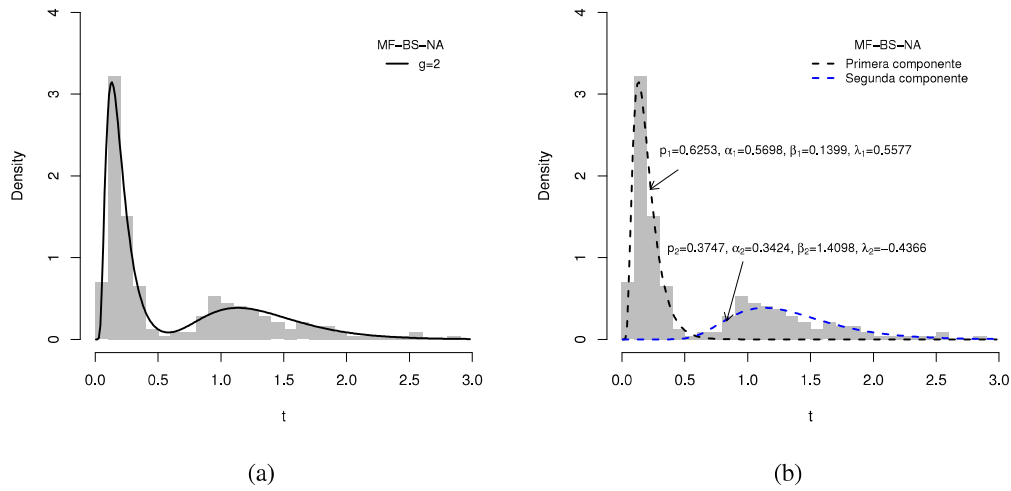


Figura 17: Histograma del conjunto de datos de enzima con las densidades de FM-BS-NA superpuestas (a) para $G = 2$ componentes (b) para cada componente.

Cuadro 6: EMV con sus correspondientes EE para el modelo de MF-BS-NA aplicados al conjunto de datos Enzima.

| Parámetro | EMV | SE |
|-------------|---------|---------|
| α_1 | 0.3426 | 0.7472 |
| α_2 | 0.5699 | 1.9731 |
| β_1 | 1.4107 | 3.0111 |
| β_2 | 0.1399 | 0.6272 |
| λ_1 | -0.4393 | 9.3041 |
| λ_2 | 0.5580 | 13.0073 |
| p_1 | 0.3747 | 0.0346 |

El Cuadro 6 presenta los EMV de α_1 , α_2 , β_1 , β_2 y p_1 para el modelo de MF-BS-NA con sus correspondientes errores estándar (EE), los cuales fueron calculados por medio del método de aproximación presentado en la Sección 3.4.4.

V. CONCLUSIONES

1. El presente trabajo propone una nueva distribución llamada Mixtura Finita Birnbaum-Saunders Normal Asimétrica como una extensión de los modelos propuestos por Vilca et al. (2011), Benites et al. (2017) y Balakrishnan *et al.* (2011).
2. Este nuevo modelo permite a los profesionales de diferentes áreas modelar datos con gran flexibilidad, acomodando simultáneamente multimodalidad, asimetría y colas pesadas.
3. Presentamos algunas características y propiedades importantes del modelo MF-BS-NA que permite obtener mejores EMV y calcularlos eficientemente utilizando el algoritmo EM.
4. Las funciones de R que fueron elaboradas para la metodología propuesta, serán incorporadas al paquete **bssn** que se encuentra disponible en el CRAN. A su vez estarán disponibles en el portafolio de GitHub donde se encuentran los códigos de la simulación 1 y 3, proporcionando a los profesionales una herramienta conveniente para otras aplicaciones en su dominio.
5. La utilidad práctica del nuevo método se ilustra con el análisis de dos conjuntos de datos reales y tres estudios de simulación.

VI. RECOMENDACIONES

El modelo MF-BS-NA puede extenderse usando la distribución mixtura de escala normal asimétrica en lugar de la distribución normal asimétrica, la cual es una clase rica de distribuciones que contienen a las distribuciones normal asimétrica, t-asimétrica, slash-asimétrica y normal contaminada-asimétrica como elementos propios que permitiría modelar datos que tienen colas más pesadas que la MF-BS-NA. Por otro lado, también es posible realizar una mixtura de regresiones con respuesta con cola pesada, siguiendo el enfoque de Zeller *et al.* (2016) o modelos de regresión cuyo término de error se distribuye como una mixtura finita de distribuciones basándose en la idea propuesta Benites *et al.* (2017). Otro enfoque que vale la pena es desarrollar una inferencia completamente bayesiana a través del método Monte Carlo de la cadena de Markov. Finalmente, nuestra propuesta puede llevarse a entornos multivariados, siguiendo la reciente propuesta de Khosravi *et al.* (2014) para mixtura de distribuciones bivariadas Birnbaum-Saunders.

VII. REFERENCIAS BIBLIOGRÁFICAS

- Achcar, J. (1993). Inference for the birnbaum-saunders fatigue life model using bayesian methods. *Computational Statistics and Data Analysis*, **15**, 367–380.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Autom. Cont.*, **19**, 716–723.
- Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, **12**, 171–178.
- Azzalini, A. (2005). The skew-normal distribution and related multivariate families. *Scandinavian Journal of Statistics*, **32**, 159–188.
- Azzalini, A. (2017). *sn: The Skew-Normal and Related Distributions, such as the Skew-t*. R package version 1.5-0.
- Balakrishnan, N., Gupta, R., Kundu, D., Leiva, V. y Sanhueza, A. (2011). On some mixture models based on the birnbaum-saunders distribution and associated inference. *Journal of Statistical Planning and Inference*, **141**, 2175–2190.
- Barros, M., Paula, G. y Leiva, V. (2008). A new class of survival regression models with heavy-tailed errors: robustness and diagnostics. *Lifetime Data Analysis*, **14**, 316–332.
- Bechtel, Y., Bonaiti-Pellie, C., Poisson, N., Maignette, J. y Bechtel, P. (1993). A population and family study of n-acetyltransferase using caffeine urinary metabolites. *Clinical Pharmacology and Therapeutics*, **54**, 134–141.
- Benites, L., Lachos, V. H. y Maehara, R. (2016). Linear regression models with finite mixtures of skew heavy-tailed errors. TR-UCONN 27, University of Connecticut - USA.
- Benites, L., Maehara, R., Vilca, F. y Marmolejo-Ramos, F. (2017). Finite Mixture of Birnbaum-Saunders distributions using the k bumps algorithm. *ArXiv e-prints*.
- Birnbaum, Z. y Saunders, S. (1969a). A new family of life distributions. *Journal of Applied Probability*, **6**, 319–327.
- Birnbaum, Z. y Saunders, S. (1969b). Estimation for a family of life distributions with applications to fatigue. *John Wiley and Sons*, **6**, 328–347.
- Böhning, D. (2000). *Computer-assisted analysis of mixtures and applications. Meta-analysis, disease mapping and others*. Chapman&Hall/CRC, Boca Raton.

- Dempster, A., Laird, N. y Rubin, D. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B*, **39**(1), 1–38.
- Desmond, A. F. (1985). Stochastic models of failure in random environments. *Canadian Journal of Statistics*, **13**, 171–183.
- Desmond, A. F. (1986). On the relationship between two fatigue-life models. *IEEE Transactions on Reliability*, **35**, 167–169.
- Díaz-García, J. y Leiva, V. (2005). A new family of life distributions based on the contoured elliptically distributions. *Journal of Statistical Planning and Inference*, **128**, 88–95.
- Dupuis, D. y Mills, J. (1997). Robust estimation of the birnbaum-saunders distribution. *IEEE Transactions on Reliability*, **47**, 88–95.
- Engelhardt, M., Bain, L. y Wright, F. (1981). Inferences on the parameters of the birnbaum-saunders fatigue life distribution based on maximum likelihood estimation. *Technometrics*, **23**, 251–256.
- Ferraro, M., Giordani, P., Vantaggi, B., Gagolewski, M., Gil, M., Grzegorzewski, P. y Hryniewicz, O. (2017). *Soft Methods for Data Science*. Advances in Intelligent Systems and Computing 456. Springer International Publishing, first edition. ISBN 978-3-319-42972-4, 978-3-319-42971-7.
- Frühwirth-Schnatter, S. (2006). *Finite mixture and markov switching models*. Springer, New York.
- Johnson, N., Kotz, S. y Balakrishnan, N. (1995). *Continuous Univariate Distributions*. Wiley, New York.
- Kaufman, L. y Rousseeuw, P. (1990). *Finding Groups in Data*. New Wiley, New York.
- Khosravi, M., Kundu, D. y Jamalizadeh, A. (2014). On bivariate and mixture of bivariate birnbaum-saunders distributions. *Statistical Methodology*.
- Leiva, V. (2015). *The Birnbaum-Saunders Distribution*. Academic Press-Elsevier.
- Leiva, V., Barros, M., Paula, G. y Sanhueza, A. (2007). Generalized birnbaum-saunders distribution applied to air pollutant concentration. *Environmetrics*, **19**, 235–249.
- Leiva, V., Barros, M., Paula, G. y Sanhueza, A. (2008). Generalized birnbaum-saunders distributions applied to air pollutant concentration. *Environmetrics*, **19**, 235–249.
- Leiva, V., Vilca, F., Balakrishnan, N. y Sanhueza, A. (2010). A skewed sinh-normal distribution and its properties and application to air pollution. *Communications in Statistics - Theory and Methods*, **39**, 426–443.

- Lin, T., Lee, J. y Hsieh, W. (2007). Robust mixture modelling using the skew t distribution. *Statistics and Computing*, **17**, 81–92.
- Louis, T. (1982). Finding the observed information matrix when using the em algorithm. *Journal of the Royal Statistical Society, Series B*, **44**, 226–233.
- Lu, M. y Chang, D. (1997). Bootstrap prediction intervals for the birnbaum-saunders distribution. *Microelectron Reliability*, **37**, 1213–1216.
- Lubke, G. y Grimm, K. J. (2017). Novel approaches in mixture modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, **24**(2), 157–158.
- Maehara, R. y Benites, L. (2016). *bssn: Birnbaum-Saunders Model Based on Skew-Normal Distribution*. R package version 0.7.
- Mann, N. R., Schafer, R. E. y Singpurwalla, N. (1974). *Methods for Statistical Analysis of Reliability and Life Data*. John Wiley and Sons, New York.
- Marshall, A. y Olkin, I. (2007). *Life Distributions-Structure of Nonparametric, Semiparametric, and Parametric Families*. Springer, New York.
- McLachlan, G. J. y Krishnan, T. (2008). *The EM algorithm and extensions*. John Wiley & Sons, New Jersey.
- McLachlan, G. J. y Peel, D. (2000). *Finite mixture models*. Wiley, New York.
- McNicholas, P. (2017). *Mixture model-based classification*. Chapman & Hall/CRC, Boca Raton.
- Meng, X. y Rubin, D. B. (1993). Maximum likelihood estimation via the ECM algorithm: a general framework. *Biometrika*, **81**, 633–648.
- Ng, H., Kundu, D. y Balakrishnan, N. (2003). Modified moment estimation for the two-parameter birnbaum-saunders distribution. *Computational Statistics and Data Analysis*, **43**, 283–298.
- Owen, W. y Padgett, W. (1999). Accelerated test models for system strength based on birnbaum-saunders distribution. *Lifetime Data Analysis*, **5**, 133–147.
- Pyne, S., Rao, B. P. y (eds.), S. R. (2016). *Big Data Analytics: Methods and Applications*. Springer India, first edition. ISBN 978-81-322-3626-9, 978-81-322-3628-3.
- R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Santana, L. (2009). *Estimação e Diagnóstico em Modelos Birnbaum-Saunders Skew-Normal*. Master's thesis, Instituto De Matemática, Estatística e Computação Científica, Campinas, Brasil.

- Schwarz, G. (1978). Estimating the dimension of a model. *Ann Stat*, **6**, 461–464.
- Tatarinova, T. y Schumitzky, A. (2015). *Nonlinear Mixture Models : A Bayesian Approach*. Imperial College Press. ISBN 1848167563, 9781848167568.
- Vilca, F. y Leiva, V. (2006). A new fatigue life model based on family of skew-elliptical distributions. *Communications in Statistics - Theory and Methods*, **35**, 229–244.
- Vilca, F., Santana, L., Leiva, V. y Balakrishnan, N. (2011). Estimation of extreme percentiles in birnbaum–saunders distributions. *Computational Statistics and Data Analysis*, **55**, 1665–1678.
- Zeller, C. B., Cabral, C. R. B. y Lachos, V. H. (2016). Robust mixture regression modeling based on scale mixtures of skew-normal distributions. *Test*, **25**, 375–396.

VIII. ANEXOS

Anexo 1.: Demostración del Teorema 5

- (i) Sea $c \in \mathbb{R}_+$ y $W = cY$ entonces $Y = W/c$, luego tenemos $dY = (1/c)dW$. Por lo tanto, el Jacobiano es $|J| = 1/c$. Luego, dado que

$$f_Y(y; \mathbf{p}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda}) = \sum_{j=1}^G p_j f_{T_j}(y; \alpha_j, \beta_j, \lambda_j),$$

encontramos que

$$\begin{aligned} f_Y(y; \mathbf{p}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda}) &= f_Y(w/c) |J| \\ &= \frac{1}{c} \sum_{j=1}^G p_j f_{T_j}(w/c; \alpha_j, \beta_j, \lambda_j) \\ &= \frac{1}{c} \sum_{j=1}^G p_j 2\phi(a_w(\alpha_j, \beta_j)) A_w(\alpha_j, \beta_j) \Phi(\lambda_j a_w(\alpha_j, \beta_j)) \\ &= \frac{1}{c} \sum_{j=1}^G p_j 2\phi\left(\frac{1}{\alpha_j} \left(\sqrt{\frac{w}{c\beta_j}} - \sqrt{\frac{c\beta_j}{w}}\right)\right) \times \\ &\quad c \frac{w^{-3/2}(w + c\beta_j)}{2\alpha_j(c\beta_j)^{1/2}} \Phi\left(\lambda_j \frac{1}{\alpha_j} \left(\sqrt{\frac{w}{c\beta_j}} - \sqrt{\frac{c\beta_j}{w}}\right)\right) \\ &= \sum_{j=1}^G p_j 2\phi(a_w(\alpha_j, c\beta_j)) A_w(\alpha_j, c\beta_j) \Phi(\lambda_j a_w(\alpha_j, c\beta_j)) \\ &= \sum_{j=1}^G p_j f_{T_j}(w; \alpha_j, c\beta_j, \lambda_j). \quad \blacksquare \end{aligned}$$

- (ii) Ahora sea $W = Y^{-1}$. Entonces $Y = W^{-1}$ y en consecuencia $dY = -W^{-2}dW$. Así el Jacobiano es $|J| = W^{-2}$ y por lo tanto

$$\begin{aligned} f_W(w; \mathbf{p}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda}) &= f_Y(w^{-1}) |J| \\ &= \sum_{j=1}^G p_j f_{T_j}(w^{-1}; \alpha_j, \beta_j, \lambda_j) w^{-2} \\ &= \sum_{j=1}^G p_j 2\phi(a_w(\alpha_j, \beta_j)) A_w(\alpha_j, \beta_j) \Phi(\lambda_j a_w(\alpha_j, \beta_j)) w^{-2} \end{aligned}$$

$$f_W(w; \mathbf{p}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda}) = \sum_{j=1}^G p_j 2\phi \left(\frac{1}{\alpha_j} \left(\sqrt{\frac{w^{-1}}{\beta_j}} - \sqrt{\frac{\beta_j}{w^{-1}}} \right) \right) \times \frac{(w^{-1})^{-3/2} (w^{-1} + \beta_j)}{2\alpha_j \beta_j^{1/2}} w^{-2} \Phi \left(\lambda_j \frac{1}{\alpha_j} \left(\sqrt{\frac{w^{-1}}{\beta_j}} - \sqrt{\frac{\beta_j}{w^{-1}}} \right) \right),$$

Dado que $\phi(x-y) = \phi(y-x)$ y que $\Phi(x-y) = \Phi(y-x)$ donde $\phi(\cdot)$ y $\Phi(\cdot)$ denotan la fdp y fda de la distribución normal estandar, respectivamente. Así tenemos que

$$\begin{aligned} f_W(w; \mathbf{p}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda}) &= \sum_{j=1}^G p_j 2\phi \left(\frac{1}{\alpha_j} \left(\sqrt{\frac{w}{\beta_j^{-1}}} - \sqrt{\frac{\beta_j^{-1}}{w}} \right) \right) \times \frac{w^{-3/2} (w + \beta_j^{-1})}{2\alpha_j (\beta_j^{-1})^{1/2}} \Phi \left(\lambda_j \frac{1}{\alpha_j} \left(\sqrt{\frac{w}{\beta_j^{-1}}} - \sqrt{\frac{\beta_j^{-1}}{w}} \right) \right) \\ &= \sum_{j=1}^G p_j 2\phi(a_w(\alpha_j, \beta_j^{-1})) A_w(\alpha_j, \beta_j^{-1}) \Phi(\lambda_j a_w(\alpha_j, \beta_j^{-1})) \\ &= \sum_{j=1}^G p_j f_{T_j}(w; \alpha_j, \beta_j^{-1}, \lambda_j). \quad \blacksquare \end{aligned}$$

(iii) De acuerdo con la definición de la función de distribución acumulada de una variable aleatoria continua tenemos lo siguiente

$$\begin{aligned} F_Y(y) &= \sum_{j=1}^G p_j \int_0^y 2\phi(a_x(\alpha_j, \beta_j)) A_x(\alpha_j, \beta_j) \Phi(\lambda_j a_x(\alpha_j, \beta_j)) dx \\ &= \sum_{j=1}^G p_j \int_{-\infty}^{a_y(\alpha_j, \beta_j)} 2\phi(x) \Phi(\lambda_j x) dx \\ &= \sum_{j=1}^G p_j \Phi_{\lambda_j}(a_y(\alpha_j, \beta_j)). \quad \blacksquare \end{aligned}$$

(iv) Si $Y \sim \text{MF-BS-NA}(\mathbf{p}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda})$ entonces su fda es como en (iii). Luego si $W = \log(Y)$ tenemos que

$$\begin{aligned} P(W \leq w) &= P(\log(Y) \leq w) \\ &= P(Y \leq \exp(w)) \\ &= \sum_{j=1}^G p_j \Phi_{\lambda_j} \left[\frac{1}{\alpha_j} \left(\sqrt{\frac{\exp(w)}{\beta_j}} - \sqrt{\frac{\beta_j}{\exp(w)}} \right) \right] \end{aligned}$$

$$\begin{aligned}
P(W \leq w) &= \sum_{j=1}^G p_j \Phi_{\lambda_j} \left\{ \frac{1}{\alpha_j} \left[\exp\left(\frac{w - \log(\beta_j)}{2}\right) - \exp\left(\frac{\log(\beta_j) - w}{2}\right) \right] \right\} \\
&= \sum_{j=1}^G p_j \Phi_{\lambda_j} \left[\frac{2}{\alpha_j} \sinh\left(\frac{w - \log(\beta_j)}{2}\right) \right] \\
&= \sum_{j=1}^G p_j \Phi_{\lambda_j} \left[\frac{2}{\alpha_j} \sinh\left(\frac{w - \gamma_j}{2}\right) \right],
\end{aligned}$$

Derivamos la fda de W con respecto a w y obtenemos

$$\begin{aligned}
P(W \leq w) &= \sum_{j=1}^G p_j 2\phi(\xi_2(w; \alpha_j, \gamma_j)) \xi_1(w; \alpha_j, \gamma_j) \Phi(\lambda_j \xi_2(w; \alpha_j, \gamma_j)) \\
&= \sum_{j=1}^G p_j f_{W_j}(w; \alpha_j, \gamma_j, \lambda_j). \quad \blacksquare
\end{aligned}$$

Anexo 2.: El paquete `bssn`

No disponemos en el entorno informático estadístico R (R Core Team, 2017) un paquete que pueda facilitar el análisis de datos provenientes de una MF-BS-NA, en la Red Completa de Archivos R (CRAN), en el paquete `bssn` (Maehara y Benites, 2016) serán incorporadas las funciones que ayudarán analizar conjuntos de datos que pueden ser modelados usando una distribución MF-BS-NA. En los siguientes anexos vamos a describir brevemente algunas funciones disponibles en el paquete `bssn`, el cual permite analizar conjuntos de datos que poseen una distribución BS-NA que esta disponible en el CRAN <https://cran.r-project.org/web/packages/bssn/index.html>. Luego mostraremos las salidas de algunas funciones que serán incorporadas posteriormente al paquete `bssn` que son útiles para analizar conjuntos de datos provenientes de una distribución MF-BS-NA.

Anexo 3.: Funciones para el ajuste de una BS-NA

Los códigos que siguen a continuación muestran como podemos obtener los estimadores de maxima verosimilitud para datos provenientes de una BS-NA. Para ello vamos considerar el conjunto de datos reales ozono.

```

# Cargando el conjunto de datos
data(ozone)
attach(ozone)
# El modelo
ti      <- dailyozonelevel
# Valores iniciales
initial <- mmmeth(ti)

```

```

alpha0    <- initial$alpha0ini
beta0     <- initial$beta0init
lambda0   <- 0
delta0    <- lambda0/sqrt(1+lambda0^2)

```

La función `EMbssn()` permite obtener el gráfico cuantil-cuantil con su respectiva banda de confianza colocando dentro de la función `show.envelope = "TRUE"`, esta herramienta gráfica permite conocer si es conveniente hacer uso de un modelo BS-NA para el conjunto de datos.

```

#Estimated parameters of the model (by default)
est_param <- EMbssn(ti,alpha0,beta0,delta0,loglik=T,
accuracy = 1e-8,show.envelope = "TRUE",
iter.max=500)

```

Se puede usar el código que se muestra a continuación para extraer cada uno de los estimadores de máxima verosimilitud

```

#ML estimates
alpha    <- est_param$res$alpha
beta     <- est_param$res$beta
lambda   <- est_param$res$lambda

```

El paquete **bssn** muestra la siguiente salida para la función `EMbssn()`

```

-----
Birnbbaum-Saunders model based on Skew-Normal distribution
-----

Observations = 116
-----
Estimates
-----

      Estimate Std. Error z value Pr(>|z|)
alpha  1.26014    0.23673  5.32311  0.00000
beta   14.65730    4.01984  3.64624  0.00027
lambda  1.06277    0.54305  1.95706  0.05034

-----
Model selection criteria
-----

      Loglik   AIC   BIC   HQC
Value -542.768 4.705 4.741 4.719

```



```
-----  
Details  
-----
```

```
Iterations = 415  
Processing time = 0.4283214 secs  
Convergence = TRUE
```

Se puede apreciar usando la función `EMbssn()` es posible saber si los estimadores de máxima verosimilitud son significativos, así como también el número de iteraciones usadas, el tiempo que demoró el proceso de estimación y finalmente podemos saber si efectivamente el algoritmo alcanzó la convergencia.

Anexo 4.: Gráficos para la BS-NA

```
#####  
## Densidad ##  
#####  
  
sseq <- seq(0,3,0.01)  
dens <- dbssn(sseq,alpha=0.2,beta=1,lambda=1.5)  
plot(sseq, dens,type="l", lwd=2,col="red", xlab="x", ylab="f(x)",  
main="BSSN Density function")  
  
#####  
## Diferentes densidades en un solo gráfico ##  
#####  
  
## para valores positivos de lambda  
y <- seq(0,3,0.01)  
f1 <- dbssn(y,0.2,1,1)  
f2 <- dbssn(y,0.2,1,2)  
f3 <- dbssn(y,0.2,1,3)  
f4 <- dbssn(y,0.2,1,4)  
den <- cbind(f1,f2,f3,f4)  
matplot(y,den,type="l", col=c("deepskyblue4", "firebrick1",  
"darkmagenta", "aquamarine4"), ylab="Density function",  
,xlab="y",lwd=2,sub="(a)")  
legend(1.5,2.8,c("BSSN(0.2,1,1)", "BSSN(0.2,1,2)", "BSSN(0.2,1,3)",  
"BSSN(0.2,1,4)"),col = c("deepskyblue4", "firebrick1",  
"darkmagenta", "aquamarine4"),lty=1:4,lwd=2,seg.len=2,cex=0.8,  
box.lty=0,bg=NULL)
```

```

## para valores negativos de lambda
y <- seq(0,3,0.01)
f1 <- dbssn(y,0.2,1,-1)
f2 <- dbssn(y,0.2,1,-2)
f3 <- dbssn(y,0.2,1,-3)
f4 <- dbssn(y,0.2,1,-4)
den <- cbind(f1,f2,f3,f4)
matplot(y,den,type="l", col=c("deepskyblue4", "firebrick1",
"darkmagenta","aquamarine4"),ylab ="Density function",
xlab="y",lwd=2,sub="(a)")
legend(1.5,2.8,c("BSSN(0.2,1,-1)", "BSSN(0.2,1,-2)",
"BSSN(0.2,1,-3)", "BSSN(0.2,1,-4)"),col=c("deepskyblue4",
"firebrick1","darkmagenta","aquamarine4"),lty=1:4,lwd=2,
seg.len=2,cex=1,box.lty=0,bg=NULL)

#####
## Función de distribución acumulada ##
#####

sseq <- seq(0.1,6,0.05)
df <- pbssn(q=sseq,alpha=0.75,beta=1,lambda=3)
plot(sseq, df, type = "l", lwd=2, col="blue", xlab="x", ylab="F(x)",
main = "BSSN Distribution function")
abline(h=1,lty=2)

#####
## Inversa de la función de distribución acumulada ##
#####

prob <- seq(0,1,length.out = 1000)
idf <- qbssn(p=prob,alpha=0.75,beta=1,lambda=3)
plot(prob, idf, type="l", lwd=2, col="gray30", xlab="x", ylab =
expression(F^{-1}~(x)), mgp=c(2.3,1,.8))
title(main="BSSN Inverse Distribution function")
abline(v=c(0,1),lty=2)

#####
## Histograma de una muestra aleatoria con linea ajustada ##
#####

```

```

sample <- rbssn(n=10000,alpha=0.75,beta=1,lambda=3)
hist(sample,breaks = 70,freq = FALSE,main="")
title(main="Histogram and True density")
sseq  <- seq(0,8,0.01)
dens  <- dbssn(sseq,alpha=0.75,beta=1,lambda=3)
lines(sseq,dens,col="red",lwd=2)

#####
## Densidades reescaladas ##
#####

alpha <- 0.8
beta  <- 1

x     <- seq(0.01,7,0.01)
f1    <- dbssn(x,alpha,beta,0)
xs1   <- x-x[f1 >= max(f1)]
f2    <- dbssn(x,alpha,beta,1)*max(f1)/max(dbssn(x,alpha,beta,1))
xs2   <- x-x[f2 >= max(f2)]
f3    <- dbssn(x,alpha,beta,2)*max(f1)/max(dbssn(x,alpha,beta,2))
xs3   <- x-x[f3 >= max(f3)]
f4    <- dbssn(x,alpha,beta,3)*max(f1)/max(dbssn(x,alpha,beta,3))
xs4   <- x-x[f4 >= max(f4)]

boundy <- c(0,0.8)
boundx <- c(-2,6)

gris1 <- rgb(54,54,54,max=255)
gris2 <- rgb(72,72,72,max=255)
gris3 <- rgb(136,136,136,max=255)
gris4 <- rgb(102,102,102,max=255)

plot(xs1,f1,xlim=boundx,ylim=boundy,xlab="",ylab="",type="l",lty=1,
lwd=1.5, col="black")
lines(xs2,f2,lty=3,lwd=1.5,col=gris2)
lines(xs3,f3,lty=4,lwd=1.5,col=gris1)
lines(xs4,f4,lty=5,lwd=1.5,col=gris4)
abline(v=0,lty=2)

```

```

legend(3,0.8,c(expression(paste("BS(*alpha,"=0.8,"*beta,"=1)")),
c(expression(paste("BS-NA(*alpha,"=0.8,"*beta,"=1,"*lambda,"=3)"))),
c(expression(paste("BS-NA(*alpha,"=0.8,"*beta,"=1,"*lambda,"=4)"))),
c(expression(paste("BS-NA(*alpha,"=0.8,"*beta,"=1,"*lambda,"=5)")))),
col = c("black", gris2,gris1,gris4),lty=1:4,lwd=1.5,seg.len=2,
cex=0.8,box.lty=0,bg=NULL)

```

Anexo 5.: Funciones para el ajuste MF-BS-NA

Las funciones que se muestra a continuación serán incorporadas posteriormente al paquete **bssn** del R . La función para obtener los estimadores de máxima verpsimilitud es `EMmixbssn()`, la cual posee los siguientes argumentos

```

EMmixbssn(ti, alpha = NULL, beta = NULL, delta = NULL, pii = NULL,
g = NULL, initial.lambda="1", get.init = TRUE, criteria = TRUE,
group = FALSE, accuracy = 10^-6, iter.max = 500,
algorithm = "Hartigan-Wong")

```

Donde `ti` es el vector de observaciones de dimensión n , `alpha`, `beta`, `delta` y `pii` son los valores iniciales para el algoritmo EM. Recordemos que `pii` es un vector de dimensión G para los pesos de cada cluster y que deben sumar uno. Tenemos que `g` es el número de componentes los cuales son escogidos a priori. El argumento `initial.lambda` puede tomar los valores “1” o “2” y representan las maneras de obtener los valores iniciales para el parámetro de asimetría λ . Si `get.init` es TRUE significa que el programa va obtener los valores iniciales en forma automática y si es FALSE es porque el usuario va indicar los valores iniciales de los parámetros. En `iter.max` el usuario indica el número máximo de iteraciones que desea que realice algoritmo EM. Para `algorithm` el usuario indica el algoritmo que será utilizado dentro del algoritmo k -médias. Mostramos a continuación un ejemplo de como usar la función `EMmixbssn()` para datos simulados de una MF-BS-NA de dos componentes y la salida obtenida. Primero vamos a comenzar cargando el paquete y luego procedemos a definir los parámetros

```

library(bssn)
n      <- 1000
pii    <- c(0.4,0.6)
g      <- 2

#Primera componente
alpha1 <- 0.25
beta1  <- 3
lambda1 <- -1

#Segunda componente

```

```
alpha2 <- 0.35
beta2 <- 7
lambda2 <- 1
```

```
alpha <- c(alpha1,alpha2)
beta <- c(beta1,beta2)
lambda <- c(lambda1,lambda2)
```

Luego de definir los parámetros, el algoritmo propuesto trabaja con una reparametrización del parámetro lambda, esto es debido a que cuando se usa $\delta_i = \lambda_i / \sqrt{1 + \lambda_i^2}$ facilita matemáticamente la obtención de los parámetros usando el algoritmo EM, es decir, el algoritmo estima el parámetro delta y luego aprovechando la propiedad de invariancia de los estimadores de máxima verosimilitud podemos obtener la estimación de lambda, a continuación se muestra la reparametrización

```
delta <- lambda/sqrt(1 + lambda^2)
```

También es posible obtener la media y la variancia de cada componente, para así conocer la variabilidad en cada una de ellas y también tener una idea de que tan separados o juntos están los componentes, esto es importante porque si los componentes están muy próximos, el algoritmo puede tener algunas dificultades para alcanzar las convergencia.

```
meanbssn(alpha1,beta1,lambda1)
varbssn(alpha1,beta1,lambda1)
```

```
meanbssn(alpha2,beta2,lambda2)
varbssn(alpha2,beta2,lambda2)
```

A continuación se muestra la función para generar una muestra aleatoria de una distribución MF-BS-NA

```
ti <- rmixbssn(n,alpha,beta,lambda,pii)$y
```

El siguiente código permite generar un histograma con la densidad superpuesta para de esta forma probar si la función usada para generar números aleatorios de la distribución MF-BS-NA es correcta.

```
hist(ti,breaks = 20, freq = FALSE, main="")
sseq <- seq(0.1,20,0.01)
dens <- d.mixed.bssn(sseq,pii,alpha,beta,lambda)
lines(sseq,dens,col="red",lwd=2)
```

Luego se muestra la función para obtener los estimadores de máxima verosimilitud.

```
fit <- EMmixbssn(ti, alpha=NULL, beta=NULL, delta=NULL,
pii=NULL, g = 2, get.init = "k-means", criteria = TRUE,
group = FALSE, accuracy = 10^-6, iter.max = 1000)
```

Luego la función EMmixbssn() muestra la siguiente salida

```
-----  
Finite Mixture Birnbaum-Saunders model based on Skew-Normal distribution  
-----
```

```
Observations = 1000
```

```
-----  
Estimates  
-----
```

| | Estimate | Std. Error |
|---------|----------|------------|
| alpha1 | 0.25512 | 0.04326 |
| alpha2 | 0.33631 | 0.04789 |
| beta1 | 2.95323 | 0.24346 |
| beta2 | 7.15316 | 0.68713 |
| lambda1 | -0.85484 | 0.73583 |
| lambda2 | 0.80110 | 0.60088 |
| pii1 | 0.37391 | 0.01691 |

```
-----  
Model selection criteria  
-----
```

| | Loglik | AIC | BIC | EDC |
|-------|-----------|----------|----------|----------|
| Value | -2415.619 | 4827.589 | 4861.944 | 4857.861 |

```
-----  
Details  
-----
```

```
Iterations = 177
```

```
Processing time = 3.180946 secs
```

```
Convergence = TRUE
```

Anexo 6.: Gráficos para la MF-BS-NA

```
#####
```

```
## Densidad ##
```

```
#####
```

```
# Primera componente
```

```
alpha1 <- 0.25
```

```
beta1 <- 3
```

```

lambda1 <- -1

# Segunda componente
alpha2 <- 0.35
beta2 <- 7
lambda2 <- 1

alpha <- c(alpha1,alpha2)
beta <- c(beta1,beta2)
lambda <- c(lambda1,lambda2)

sseq <- seq(0,20,0.01)
dens <- d.mixed.bssn2(sseq, pii, alpha, beta, lambda)
plot(sseq, dens,type="l", lwd=1,col="black", xlab="t", ylab="f(t)")

#####
## Diferentes densidades en un solo gráfico ##
#####

g <- 2

# Primera componente
alpha1 <- 0.35
beta1 <- 8
lambda1 <- 2
meanbssn(alpha1,beta1,lambda1)
varbssn(alpha1,beta1,lambda1)

# Segunda componente
alpha2 <- 0.65
beta2 <- 5
lambda2 <- 4
meanbssn(alpha2,beta2,lambda2)
varbssn(alpha2,beta2,lambda2)

alpha <- c(alpha1,alpha2)
beta <- c(beta1,beta2)
lambda <- c(lambda1,lambda2)

```

```

x <- seq(0,20,0.01)
f1 <- d.mixed.bssn2(sseq, pii=c(0.2,0.8), alpha, beta, lambda)
f2 <- d.mixed.bssn2(sseq, pii=c(0.3,0.7), alpha, beta, lambda)
f3 <- d.mixed.bssn2(sseq, pii=c(0.4,0.6), alpha, beta, lambda)
den <- cbind(f1,f2,f3)
gris1 <- rgb(54,54,54,max=255)
gris2 <- rgb(72,72,72,max=255)
matplot(x,den,type="l", col=c("black", gris1, gris2),
        ylab="f(t)",xlab="t",lty=1:3,lwd=1.5,ylim=c(0,0.12))
legend("topright",c(expression(paste(p[1],"=0.2")),
                    expression(paste(p[1],"=0.3")), expression(paste(p[1],"=0.4"))),
      col = c("black", gris1,gris2),lty=1:3,lwd=1.5,seg.len=2,cex=0.8,
      box.lty=0,bg=NULL)

```

```

#####
## Histograma de una muestra aleatoria con linea ajustada ##
#####

```

```

alpha=c(0.55,0.25);beta=c(1,1.5);lambda=c(3,2);pii=c(0.3,0.7)
sample <- rmixbssn(n=1000,alpha,beta,lambda,pii)
hist(sample$y,breaks = 70,freq = FALSE,main="")
title(main="Histogram and True density")
temp <- seq(min(sample$y), max(sample$y), length.out=1000)

lines(temp, (pii[1]*dbssn(temp, alpha[1], beta[1],lambda[1]))+
        (pii[2]*dbssn(temp, alpha[2], beta[2],lambda[2])),
      col="red",lty=3, lwd=3)

# the first component
lines(temp, pii[1]*dbssn(temp, alpha[1], beta[1],lambda[1]),
      col="blue", lty=2, lwd=3)

# the second component
lines(temp, pii[2]*dbssn(temp, alpha[2], beta[2],lambda[2]),
      col="green", lty=2, lwd=3)

```



```
#####
## Función tasa de falla ##
#####

ti          <- seq(0.001,35,0.01)
HMIXBSSN_1 <- H.mix.bssn(ti,pii=c(0.2),alpha=c(0.35,0.65),beta=c(8,5),
                        lambda=c(2,4))
HMIXBSSN_2 <- H.mix.bssn(ti,pii=c(0.3),alpha=c(0.35,0.65),beta=c(8,5),
                        lambda=c(2,4))
HMIXBSSN_3 <- H.mix.bssn(ti,pii=c(0.4),alpha=c(0.35,0.65),beta=c(8,5),
                        lambda=c(2,4))
HMIXBSSN_4 <- H.mix.bssn(ti,pii=c(0.5),alpha=c(0.35,0.65),beta=c(8,5),
                        lambda=c(2,4))
HMIXBSSN_5 <- H.mix.bssn(ti,pii=c(0.6),alpha=c(0.35,0.65),beta=c(8,5),
                        lambda=c(2,4))
den          <- cbind(HMIXBSSN_1,HMIXBSSN_2,HMIXBSSN_3,HMIXBSSN_4,
                    HMIXBSSN_5)
matplot(ti,den,type="l", col=c("black"), ylab="h(t)",xlab="t",lty=1:5,
        lwd=1.5)
legend("topright",c(expression(paste(p[1],"=0.2")),
                    expression(paste(p[1],"=0.3")),expression(paste(p[1],"=0.4")),
                    expression(paste(p[1],"=0.5")),expression(paste(p[1],"=0.6"))),
      col = c("black"), lty=1:5,lwd=1.5,seg.len=2,cex=0.8,box.lty=0,
      bg=NULL)
```

Anexo 7.: Portafolio GitHub

Los códigos de las simulaciones 1 y 3 se encuentran disponibles en el portafolio de GitHub, https://github.com/rmaeharaa/thesis_graduate.

Código en R de la simulación 1

Los códigos de la simulación 1, nos permiten verificar si el algoritmo EM usado para estimar los parámetros está implementado correctamente usando una simulación de Monte Carlo. Nuestro código hace uso de la función `try()` con la cual es posible filtrar aquellas muestras aleatorias que están dando problemas en el proceso de estimación, lo cual permite no interrumpir el proceso de simulación.

Código en R de la simulación 3

Los códigos de la simulación 3, permitirán obtener los valores del cambio relativo, guardar los resultados obtenidos en un archivo en formato csv y finalmente poder construir los gráficos, lo interesante de este último proceso es la posibilidad de mostrar al usuario como es la construcción paso a paso de una figura de este tipo.