

Distribución Weibull Unitaria Bivariada

Luis Iván Páez Martínez

Trabajo de Grado Presentado como Requisito Parcial para
Optar al Título de Estadístico



Universidad de Córdoba
Facultad Ciencias Básicas
Departamento de Matemáticas y Estadística
Programa de Estadística
Montería, Abril de 2022

Distribución Weibull Unitaria Bivariada

Luis Iván Páez Martínez

Trabajo de Grado Presentado como Requisito Parcial para
Optar al Título de Estadístico

Director:

Prof. Dr. Roger Jesús Tovar Falón

Universidad de Córdoba

Facultad Ciencias Básicas

Departamento de Matemáticas y Estadística

Programa de Estadística

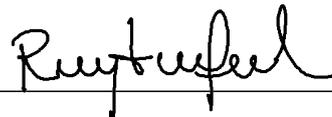
Montería, Abril de 2022

UNIVERSIDAD DE CÓRDOBA
FACULTAD DE CIENCIAS BÁSICAS
PROGRAMA DE ESTADÍSTICA

Los jurados abajo firmantes certifican que han leído y recomiendan a la **Facultad de Ciencias Básicas** aprobar el trabajo de grado titulado “**Distribución Weibull Unitaria Bivariada**” presentado por **Luis Iván Páez Martínez** como requisito parcial para optar al título de **Estadístico**.

Fecha: Abril de 2022

Director:



Prof. Dr. Roger Jesús Tovar Falón
Universidad de Córdoba
Departamento de Matemáticas y Estadística

Jurado:



Prof. Dr. Guillermo Martínez Flórez
Universidad de Córdoba
Departamento de Matemáticas y Estadística

Jurado:

Profa. Dra. Sandra Vergara Cardozo
Universidad Nacional de Colombia, Bogotá
Departamento de Matemáticas y Estadística

Agradecimientos

Quiero agradecer en primera instancia a Dios por bendecirnos con la vida, por guiarnos a lo largo de nuestra existencia, por ser el apoyo y fortaleza en aquellos momentos de dificultad y de debilidad.

Agradezco a todos los docentes del departamento de Matemáticas y Estadística de la Universidad de Córdoba que hicieron parte de mi formación y que me acompañaron durante este proceso, por todos los conocimientos brindados y en especial al director de este trabajo, Roger Jesús Tovar Falón por su ayuda, paciencia, dedicación, lo cual me permitió crecer mucho a nivel personal y académico, aprendiendo del proceso investigativo en un área tan fascinante como la Estadística.

También quiero darle las gracias a toda mi familia que han sido el principal apoyo en momentos de flaqueza y debilidad ya que con sus consejos, enseñanzas y tolerancia me han brindado ese respaldo incondicional para culminar esta etapa de mi vida, así mismo agradecer a mis amigos y compañeros que me han acompañado en este proceso.

Resumen

Los datos cuya respuesta se encuentran en el intervalo $(0, 1)$ tales como proporciones, tasas o índices surgen muchas veces en la investigaciones en diferentes áreas del conocimiento. La proporción de muertes causadas por el tabaquismo, la tasa de incidencia o prevalencia de una determinada enfermedad en una comunidad, el porcentaje de votos a favor de un candidato después de una campaña presidencial, el índice de desarrollo humano en un determinado país y la proporción de los ingresos que se gastan en educación, son algunos ejemplos de este tipo de respuestas. Para modelar este tipo de datos, existen varias propuestas, siendo la distribución beta la más conocida y aplicada en este tipo de situaciones. Entre otras propuestas, se incluyen las distribuciones Kurumaraswamy, Birnbaum-Saunders Unitaria, Weibull Unitaria, Normal-potencia unitaria y gamma unitaria, en cambio, en la teoría de distribuciones existe poca literatura estadística acerca de distribuciones para ajustar datos multivariados cuyas respuestas se encuentran en el intervalo unitario.

En este trabajo, se propone una extensión bivariada de la distribución univariada Weibull unitaria introducida por Mazucheli, Menezes y Ghitany. [*The Unit-Weibull distribution and Associated Inference*. Journal of Applied Probability and Statistics, 13 (2018), págs. 1-22], la cual ha demostrado ser una alternativa viable a las otras distribuciones utilizadas para ajustar datos en el intervalo $(0, 1)$. La cópula de Farlie-Gumbel-Morgenstern y la distribución Weibull unitaria se utilizan para producir una distribución bivariada denominada distribución Weibull unitaria bivariada. Se estudian algunas propiedades de la distribución WUB tales como las funciones de: densidad de probabilidad conjunta, distribución acumulada, supervivencia, generadora de momentos, entre otras. En la estimación de los parámetros de la distribución propuesta se considera un enfoque clásico utilizando el método de máxima verosimilitud junto con el método de estimación por inferencia de funciones marginales (Joe, 2005). Para evaluar el desempeño de los estimadores de máxima verosimilitud de los parámetros en la distribución, se realiza un estudio de simulación de Monte Carlo y se presenta una aplicación con un conjunto de datos reales para mostrar la utilidad del modelo.

Palabras claves: Distribución Weibull unitaria, cópula FGM, estimación de máxima verosimilitud, estimación por inferencia de funciones marginales, simulación de Monte Carlo.

Índice general

Índice de figuras	III
Índice de tablas	IV
1. Introducción	1
2. Conceptos preliminares	3
2.1. Distribución Weibull	3
2.1.1. Propiedades	4
2.2. Distribución Weibull Unitaria	5
2.2.1. Propiedades	6
2.2.2. Estimación por Máxima Verosimilitud	8
2.3. Función de cópulas	10
2.3.0.1. Cópula FGM	10
2.4. Distribución Weibull Bivariada FGM	11
2.4.1. Propiedades	12
2.4.2. Estimación por inferencia de funciones marginales	14
2.4.3. Estimación de máxima verosimilitud	15
2.4.4. Estimación por método semiparamétrico	17
3. Distribución Weibull Unitaria Bivariada	19
3.1. Propiedades de la distribución weibull unitaria bivariada FGM	22
3.1.1. Distribuciones marginales	22
3.1.2. Distribuciones condicionales	24
3.1.3. Generación de variables aleatorias	26
3.1.4. Momentos del producto	26
3.1.5. Función de confiabilidad	28
3.2. Estimación basada en cópulas	28

3.2.1. Estimación de máxima verosimilitud	29
3.2.2. Estimación por inferencia de funciones marginales	30
3.2.3. Estimación por método semi-paramétrico (SM)	32
3.3. Matriz de información	33
3.4. Intervalos de confianza asintóticos	36
4. Estudio de simulación	38
5. Aplicación con datos reales	45
6. Conclusiones y trabajos futuros	49
A. Tablas de simulaciones	50
Bibliografía	50

Índice de figuras

2.1. Función de densidad de probabilidad Weibull para $\alpha = 1$ y $\alpha = 2.5$ y algunos valores seleccionados del parámetro β	4
2.2. Función de densidad de probabilidad Weibull Unitaria para $\alpha = 1$ y $\alpha = 2.5$ y algunos valores seleccionados del parámetro β	6
3.1. Función de densidad conjunta de la distribución weibull unitaria para diferentes vaores de los parámetros. (a) $\varphi = (1.5, 1.5, 2.0, 2.0, 0.75)$, (b) $\varphi = (2.2, 2.5, 3.5, 2.5, 0.25)$, (c) $\varphi = (2.2, 2.5, 1.1, 1.0, -0.25)$, (d) $\varphi = (1.0, 1.2, 0.5, 2.5, -0.75)$, (e) $\varphi = (0.9, 0.9, 0.7, 0.7, 0.5)$, (f) $\varphi = (3.0, 3.0, 3.5, 3.5, -0.5)$	20
3.2. Gráficas de contorno para varios valores de los parámetros. (a) $\varphi = (1.5, 0.5, 3.5, 2.0, 0.25)$, (b) $\varphi = (2.5, 1.5, 1.5, 1.0, 0.75)$, (c) $\varphi = (2.2, 2.5, 3.5, 2.0, -0.25)$, (d) $\varphi = (1.0, 1.2, 0.5, 2.5, -0.75)$	21
4.1. Comportamiento de las estadísticas de la estimación de α_1 según el tamaño de muestra.	40
4.2. Comportamiento de las estadísticas de la estimación de β_1 según el tamaño de muestra.	41
4.3. Comportamiento de las estadísticas de la estimación de α_2 según el tamaño de muestra.	42
4.4. Comportamiento de las estadísticas de la estimación de β_2 según el tamaño de muestra.	43
4.5. Comportamiento de las estadísticas de la estimación de θ según el tamaño de muestra.	44
5.1. Gráficas de contorno para los modelos WUB, BJSB y BSB.	46
5.2. Gráficas de densidad de las variables tasa neta de escolaridad y proporción de necesidades básicas insatisfechas.	47

Índice de tablas

3.1. Valores del coeficiente de correlación.	27
5.1. Medidas descriptivas.	46
5.2. Estimaciones de los parámetros de los modelos WUB, BJSB y BSB.	47
A.1. Estimación de los parámetros de la distribución WUB: caso 1.1: $\alpha_1 = 1.5, \beta_1 = 3.5, \alpha_2 = 0.5, \beta_2 = 2.0, \theta = 0.25$	51
A.2. Estimación de los parámetros de la distribución WUB: caso 1.2: $\alpha_1 = 1.5, \beta_1 = 3.5, \alpha_2 = 0.5, \beta_2 = 2.0, \theta = 0.75$	52
A.3. Estimación de los parámetros de la distribución WUB: caso 1.3: $\alpha_1 = 1.5, \beta_1 = 3.5, \alpha_2 = 0.5, \beta_2 = 2.0, \theta = -0.25$	53
A.4. Estimación de los parámetros de la distribución WUB: caso 1.4: $\alpha_1 = 1.5, \beta_1 = 3.5, \alpha_2 = 0.5, \beta_2 = 2.0, \theta = -0.75$	54
A.5. Estimación de los parámetros de la distribución WUB: caso 2.1: $\alpha_1 = 2.5, \beta_1 = 1.5, \alpha_2 = 1.5, \beta_2 = 1.0, \theta = 0.25$	55
A.6. Estimación de los parámetros de la distribución WUB: caso 2.2: $\alpha_1 = 2.5, \beta_1 = 1.5, \alpha_2 = 1.5, \beta_2 = 1.0, \theta = 0.75$	56

Capítulo 1

Introducción

La distribución Weibull (1951) ha sido ampliamente utilizada en campos del conocimiento debido a su gran versatilidad y su flexibilidad para aproximarse a la distribución exponencial, normal y Rayleigh. Esta distribución ha recibido mucha atención en la literatura estadística y ha sido estudiada por muchos autores, especialmente en el desarrollo de la teoría de confiabilidad, el análisis de supervivencia y la teoría de colas. Con la distribución Weibull se ha evidenciado el ajuste de datos simétricos, datos que presentan asimetría a la derecha o a la izquierda, por lo tanto, tiene un rango amplio de aplicabilidad como por ejemplo, en la fiabilidad con datos de cero fallo (Zhang, 2021), el modelamiento directo del tiempo de falla de un sistema en serie compuesto por componentes dependientes (Xu et al., 2021), el análisis de la resistencia a la flexión y la fiabilidad de las cerámicas estructurales avanzadas (Liu et al., 2021), el estudio del tiempo de vida del rebaño en vacas Holstein canadienses (Dürr et al., 1999), entre otros.

En todas estas aplicaciones anteriormente mencionadas en donde se utiliza la distribución Weibull, se tiene que la variable de estudio se encuentra en el intervalo $(0, +\infty)$, sin embargo, en muchos estudios la variable de interés está restringida al intervalo $(0, 1)$, como es el caso de índices, tasas o proporciones, las cuales han recibido gran atención en las últimas décadas, ya que son muy comunes en el campo de la ingeniería, medicina, economía, entre otras. Esto llevó a Mazucheli et al. (2018b) proponer una nueva distribución denominada Weibull Unitaria con soporte en el intervalo $(0, 1)$, la cual nace de una pequeña transformación a la distribución Weibull de dos parámetros.

En la teoría de probabilidad y estadística la distribución más conocida para modelar datos en el intervalo $(0, 1)$ es la distribución beta de dos parámetros debido a su gran fle-

xibilidad, trabajos como los de Cook et al. (2008) y Gupta and Nadarajah (2004), tratan generalidades de la distribución beta y sus aplicaciones, otras propuestas de distribuciones diferentes a la distribución beta son: la distribución Kumaraswamy (1980), la distribución Birnbaum-Saunders Unitaria (UBS) de Mazucheli et al. (2018a). Para el caso bivariado, se conocen pocos trabajos que abordan este problema, ejemplos de ellos son la distribución Birnbaum Saunders Bivariada (Kundu et al., 2010) y la distribución Birnbaum Saunders Multivariada (Lemonte et al., 2015), con base en los trabajos de (Arnold et al., 2002), es por ello, que en este trabajo se propone introducir una nueva distribución capaz de modelar respuestas con soporte en el plano unitario $(0, 1) \times (0, 1)$. Esta nueva distribución extiende la distribución Weibull unitaria de Mazucheli et al. (2018b).

Por lo anterior, considerando el problema de modelar tasas y proporciones de manera conjunta el presente trabajo busca extender al caso bivariado la distribución Weibull unitaria, partiendo del métodos de las cópulas, más precisamente a partir de la cópula Farlie-Gumbel-Morgenstern (FGM), esta nueva distribución es denominada weibull unitaria bivariada. A esta nueva distribución se le estudia las características de la función de densidad conjunta, distribución acumulada conjunta, marginales, condicionales, momentos, matrices de varianza y coeficientes de correlación, entre otras. Con el modelo propuesto se llevará a cabo el proceso de inferencia estadística, partiendo de la estimación de sus parámetros desde un punto de vista clásico usando el método de máxima verosimilitud, también por el método de estimación de inferencia para marigales. Para evaluar el desempeño de los estimadores, se realiza un estudio de simulaciones de Monte Carlo para comparar las preferencias entre los métodos de estimación, se encontrarán las matrices de información (o de Fisher) del modelo y se introduce un conjunto de datos reales, se analiza para investigar el modelo y se obtienen resultados útiles con fines ilustrativos.

El resto del documento está organizado de la siguiente manera: En el Capítulo 2 se presenta la distribución Weibull y sus principales propiedades, la distribución Weibull Unitaria con sus principales propiedades, la función de cópulas (FGM) la distribución Weibull Bivariada FGM y sus principales propiedades. En la capítulo 3 se introduce la distribución Weibull Unitaria Bivariada y se estudian sus propiedades en detalle con la estimación de los parámetros. En el capítulo 4 se realizan un estudio de simulación. En el capítulo 5 se hace una aplicación con datos reales de la distribución Weibull Unitaria Bivariada. Finalmente el capítulo 6 se presentan las principales conclusiones y se dan algunas recomendaciones para trabajos futuros.

Capítulo 2

Conceptos preliminares

En este capítulo se presentan algunos elementos básicos que facilitan el desarrollo de este trabajo. Específicamente, se presenta la distribución Weibull, su función de densidad de probabilidad, función de distribución y sus propiedades más importantes, luego se define la distribución de Weibull unitaria de dos parámetros y sus propiedades más importantes. Por último definimos la familia de cópulas de Farlie Gumbel Morgenstern (FGM) la cual se utiliza para obtener la función de distribución conjunta de una variable aleatoria bidimensional con distribución de Weibull unitaria bivarida.

2.1. Distribución Weibull

La distribución Weibull fue introducida por Weibull (1951) y ha recibido mucha atención en la literatura debido a su gran flexibilidad inherente, es de gran importancia en el análisis de supervivencia al igual que las distribuciones log-normal, Gamma y exponencial. Esta distribución ha sido muy usada para resolver problemas de ingeniería, investigaciones médicas, control de calidad, finanzas y climatología, ejemplos de estas aplicaciones son: el modelamiento directo del tiempo de falla de un sistema en serie compuesto por componentes dependientes (Xu et al., 2021), el análisis de la resistencia a la flexión y la fiabilidad de las cerámicas estructurales avanzadas (Liu et al., 2021), el modelamiento de materiales frágiles (Fok et al., 2001), y el estudio del tiempo de vida del rebaño en vacas Holstein canadienses (Dürr et al., 1999).

Definición 1. Una variable aleatoria X tiene distribución Weibull con parámetros α y β si su función de densidad de probabilidad (fdp) está dada por

$$g(x) = \alpha \beta x^{\beta-1} e^{-\alpha x^\beta}, \quad x \geq 0, \quad \alpha, \beta > 0, \quad (2.1)$$

donde α y β son los parámetros de escala y forma, respectivamente.

Si una variable aleatoria X tiene distribución Weibull con parámetros α y β se denota como $X \sim W(\alpha, \beta)$. La Figura 2.1 muestra diferentes formas de la densidad Weibull para algunos valores seleccionados de los parámetros α y β . Se puede observar que la distribución exponencial es un caso particular del modelo Weibull cuando $\beta = 1$. Asimismo, para $\beta = 2$ se obtiene la distribución de Rayleigh.

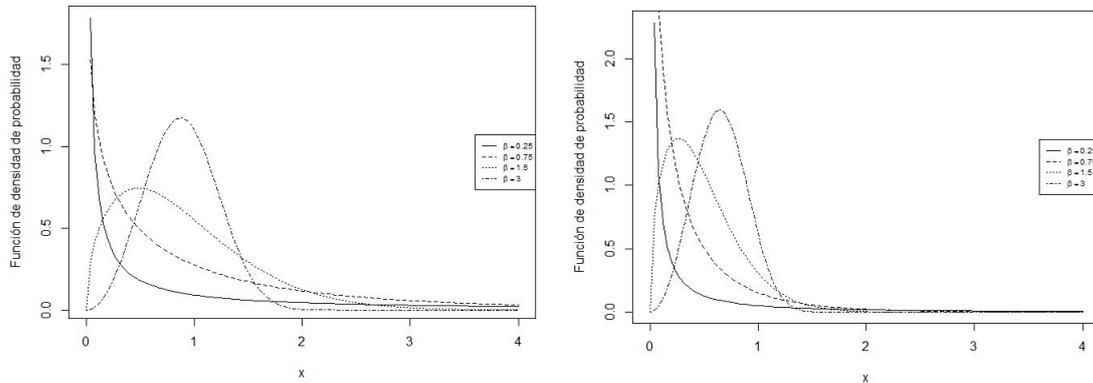


Figura 2.1: Función de densidad de probabilidad Weibull para $\alpha = 1$ y $\alpha = 2.5$ y algunos valores seleccionados del parámetro β .

2.1.1. Propiedades

A continuación se presentan algunas de las principales propiedades de la distribución Weibull (Weibull, 1951).

Proposición 1 (Función de Distribución). Sea $X \sim W(\alpha, \beta)$ entonces la fda de X es dada por:

$$F(x) = e^{-\alpha x^\beta} \quad (2.2)$$

Proposición 2 (Función de Supervivencia y Función Hazard). Sea $T \sim W(\alpha, \beta)$, entonces la función de supervivencia S , y función Hazard h , de la variable T son dadas por:

$$S(t) = 1 - F(t; \alpha, \beta) = 1 - e^{-\alpha t^\beta} \quad (2.3)$$

y

$$h(t; \alpha, \beta) = \frac{f(t; \alpha, \beta)}{1 - F(t; \alpha, \beta)} = \frac{\alpha \beta t^{\beta-1} e^{-\alpha t^\beta}}{1 - e^{-\alpha t^\beta}}, \quad t > 0 \quad (2.4)$$

Demostración. La prueba es directa a partir de la definición de las funciones S y h . \square

Proposición 3 (Esperanza y varianza). *Sea $X \sim W(\alpha, \beta)$ entonces la esperanza y la varianza de X están dadas por:*

$$E(X) = \left(\frac{1}{\alpha}\right)^{\frac{1}{\beta}} \Gamma\left(1 + \frac{1}{\beta}\right) \quad y \quad V(X) = \left(\frac{1}{\alpha}\right)^{\frac{2}{\beta}} \left(\Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma^2\left(1 + \frac{1}{\beta}\right)\right) \quad (2.5)$$

2.2. Distribución Weibull Unitaria

La distribución de Weibull unitaria (WU) fue introducida por Mazucheli et al. (2018b) y es una alternativa para modelar datos con soporte en el intervalo $(0, 1)$. Esta distribución surge a partir de la transformación $Y = e^{-X}$, donde X sigue distribución Weibull. La distribución WU es útil para modelar tasas, índices o porcentajes, como por ejemplo, la tasa de muerte por determinada enfermedad, porcentaje de ingreso dedicado a telefonía celular, porcentaje de viviendas en ciertos suburbios con antena parabólica o por satélite, entre otras.

Definición 2. *Una variable aleatoria Y tiene distribución Weibull Unitaria con parámetros α y β la cual es denotado por $Y \sim WU(\alpha, \beta)$, si su fdp está dada por:*

$$f(y) = \begin{cases} \frac{1}{y} \alpha \beta (-\ln y)^{\beta-1} \exp[-\alpha (-\ln y)^\beta], & \text{si } 0 < y < 1, \quad \alpha, \beta > 0, \\ 0, & \text{en otro caso.} \end{cases} \quad (2.6)$$

Nótese que α ya no es un parámetro de escala, ya que $f(\alpha y; \alpha, \beta) \neq \frac{1}{\alpha} f(y; 1, \beta)$. Los casos especiales de la distribución de WU incluyen: la distribución uniforme estándar sobre el intervalo $(0, 1)$ ($\alpha = \beta = 1$), la distribución función-potencia ($\beta = 1$) y la distribución unitaria de Rayleigh ($\beta = 2$). Por lo tanto, la distribución WU tiene conexión con algunas distribuciones bien conocidas y, por tanto, puede ser muy útil en muchas situaciones prácticas. La forma de la fdp de la distribución WU para algunos valores seleccionados de los parámetros α y β pueden ser vistos en la Figura 2.2.

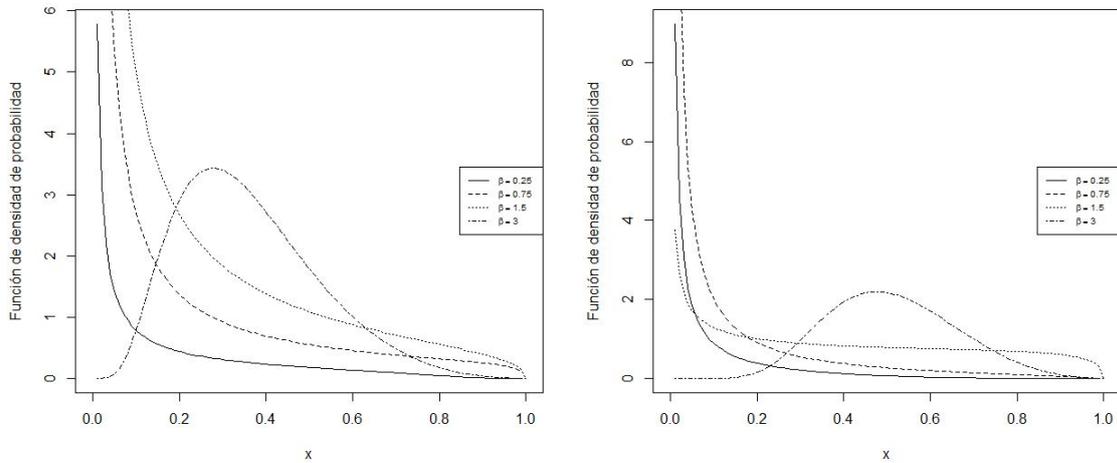


Figura 2.2: Función de densidad de probabilidad Weibull Unitaria para $\alpha = 1$ y $\alpha = 2.5$ y algunos valores seleccionados del parámetro β .

2.2.1. Propiedades

A continuación se presentan algunas de las principales propiedades de la distribución WU introducida por Mazucheli et al. (2018b).

Proposición 4 (Función de Distribución). *Sea $Y \sim WU(\alpha, \beta)$ entonces la fda de Y es dada por:*

$$F(y) = \exp\left(-\alpha(-\ln y)^\beta\right) \quad (2.7)$$

Demostración. La fda de una variable aleatoria WU puede ser obtenida usando el método de la función de distribución. Así, sean $X \sim W(\alpha, \beta)$ y $Y = \exp(X)$, entonces

$$\begin{aligned} F(y) &= P(Y \leq y) \\ &= P(e^{-X} \leq y) \\ &= 1 - P(X < -\ln y) \\ &= 1 - \int_0^{-\ln y} \alpha \beta x^{\beta-1} e^{-\alpha x^\beta} dx \\ &= 1 - \left(-e^{-\alpha(-\ln y)^\beta} + 1\right) \\ &= e^{-\alpha(-\ln y)^\beta} \end{aligned}$$

□

Proposición 5 (Función de Supervivencia y Función Hazard). *Sea $T \sim WU(\alpha, \beta)$, entonces la función de supervivencia S , y función Hazard h , de la variable T son dadas por:*

$$S(t) = 1 - F(t; \alpha, \beta) = 1 - e^{-\alpha(-\ln t)^\beta} \quad (2.8)$$

y

$$h(t; \alpha, \beta) = \frac{f(t; \alpha, \beta)}{1 - F(t; \alpha, \beta)} = \frac{\alpha \beta (-\ln t)^{\beta-1} \exp[-\alpha(-\ln t)^\beta]}{t(1 - \exp[-\alpha(-\ln t)^\beta])}, \quad 0 < t < 1 \quad (2.9)$$

Demostración. La prueba es directa a partir de la definición de las funciones S y h . □

Proposición 6 (Momentos y Medidas Asociadas). *El r -ésimo momento central de una variable aleatoria Y con distribución WU viene dado por*

$$\mu'_r = E(Y^r) = E(e^{-rY}) = M_y(-r) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n! \alpha^{n/\beta}} \Gamma\left(\frac{n}{\beta} + 1\right) \quad (2.10)$$

A partir de la Proposición 6 se pueden obtener los coeficientes de asimetría y curtosis de la distribución WU . Si γ_1 y γ_2 denotan los coeficientes de asimetría y curtosis, entonces,

$$\gamma_1 = \frac{\mu'_3 - 3\mu'_2\mu + \mu^3}{\sigma^3} \quad (2.11)$$

$$\gamma_2 = \frac{\mu'_4 - 4\mu'_3\mu + 6\mu'_2\mu^2 + -3\mu^4}{\sigma^4} \quad (2.12)$$

Para el caso especial $\beta = 1$, es decir, la distribución función potencia, se tiene,

$$\mu'_r = E(Y^r) = \frac{\alpha}{r + \alpha}, \quad r = 1, 2, \dots \quad (2.13)$$

En este caso, la media, varianza y los coeficientes de asimetría y curtosis, están dados por,

$$\mu = \frac{\alpha}{1 + \alpha}, \quad \sigma^2 = \frac{\alpha}{(1 + \alpha)^2(2 + \alpha)}, \quad (2.14)$$

$$\gamma_1 = \frac{2(1 - \alpha)}{(2 + \alpha)} \sqrt{1 + \frac{2}{\alpha}}, \quad \gamma_2 = \frac{3(2 + \alpha)(2 - \alpha + 3\alpha^2)}{\alpha(3 + \alpha)(4 + \alpha)} \quad (2.15)$$

respectivamente. Nótese que, la asimetría puede ser negativa ($\alpha < 1$), cero ($\alpha = 1$) o positiva ($\alpha > 1$). Para $\alpha = 1$, se obtiene la distribución uniforme estándar mientras que para $\beta = 2$, se

sigue la distribución de Rayleigh, y por tanto,

$$\mu'_r = E(Y^r) = 1 - \frac{\sqrt{\pi}}{2\sqrt{\alpha}} r e^{r^2/(4\alpha)} \operatorname{erfc}\left(\frac{r}{2\sqrt{\alpha}}\right), \quad r = 1, 2, \dots$$

en que

$$\operatorname{erfc}(z) = \frac{2}{\sqrt{\pi}} = \int_z^\infty e^{-x^2} dx, \quad z > 0,$$

es la función error complementaria. En este caso, la media y la varianza de la distribución WU son dadas por,

$$\mu = 1 - \frac{\sqrt{\pi}}{2\sqrt{\alpha}} e^{1/(4\alpha)} \operatorname{erfc}\left(\frac{1}{2\sqrt{\alpha}}\right)$$

$$\sigma^2 = 1 - \frac{\sqrt{\pi}}{\sqrt{\alpha}} e^{1/\alpha} \operatorname{erfc}\left(\frac{1}{\sqrt{\alpha}}\right) - \left[1 - \frac{\sqrt{\pi}}{2\sqrt{\alpha}} e^{1/(4\alpha)} \operatorname{erfc}\left(\frac{1}{2\sqrt{\alpha}}\right)\right]^2$$

Proposición 7 (Función Cuantil). *La función cuantil de la distribución WU está dada por:*

$$Q(p) = \exp\left[-\left(-\frac{\log p}{\alpha}\right)^{\frac{1}{\beta}}\right], \quad 0 < p < 1.$$

Para los casos especiales $\alpha = \beta = 1$, $\beta = 1$ y $\beta = 2$, se tiene $Q(p) = p$, $Q(p) = p^{1/\alpha}$ y $Q(p) = \exp(-\sqrt{-\log(p)/\alpha})$, respectivamente. Los cuartiles de la distribución WU, así como los casos especiales, se obtienen configurando $p = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$, respectivamente.

2.2.2. Estimación por Máxima Verosimilitud

Sea $\mathbf{x} = (x_1, \dots, x_n)^\top$ una muestra aleatoria de tamaño n de una población con distribución $WU(\alpha, \beta)$. La función de log-verosimilitud para estimar $\boldsymbol{\theta} = (\alpha, \beta)^\top$ es dada por:

$$\begin{aligned} \ell(\boldsymbol{\theta}; \mathbf{x}) &= \sum_{i=1}^n \log f(x_i; \boldsymbol{\theta}) \\ &= n(\log \alpha + \log \beta) - \sum_{i=1}^n \log x_i + (\beta - 1) \sum_{i=1}^n \log(-\log x_i) - \alpha \sum_{i=1}^n (-\log x_i)^\beta \end{aligned} \quad (2.16)$$

El estimador de máxima verosimilitud $\hat{\boldsymbol{\theta}}$ de $\boldsymbol{\theta}$ se obtiene resolviendo las ecuaciones no lineales

$$\frac{\partial \ell}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=1}^n (-\log x_i)^\beta = 0, \quad (2.17)$$

y

$$\frac{\partial \ell}{\partial \beta} = \frac{n}{\beta} + \sum_{i=1}^n \log(-\log x_i) - \alpha \sum_{i=1}^n (-\log x_i)^\beta \log(-\log x_i) = 0. \quad (2.18)$$

La ecuación (2.17) se puede resolver para α , y se obtiene $\hat{\alpha}(\beta) = \frac{n}{\alpha \sum_{i=1}^n (-\log x_i)^\beta}$. Para obtener $\hat{\beta}$, se sustituye $\hat{\alpha}(\beta)$ en (2.18) y resolviendo para β se obtiene

$$g(\beta) = \frac{n}{\beta} + \sum_{i=1}^n \log(-\log x_i) - \frac{n \sum_{i=1}^n (-\log x_i)^\beta \log(-\log x_i)}{\sum_{i=1}^n (-\log x_i)^\beta}. \quad (2.19)$$

La ecuación (2.19) no tiene solución explícita y debe ser resuelta numéricamente usando, por ejemplo, el método de Brent disponible en el software R (R Core Team, 2021) a través de la función "uniroot". Este método tiene la ventaja de que no requiere el cálculo de la derivada $g'(\beta)$ y la estimación inicial de β se puede proporcionar como un intervalo. Note que (2.7) satisface $\log[-\log F(x_i; \alpha, \beta)] = \log \alpha + \beta \log(-\log x_i)$, para $i = 1, \dots, n$. Así, una gráfico $\log[-\log \hat{F}(x_{(i)})]$ versus $\log(-\log x_{(i)})$ sería aproximadamente lineal si una distribución WU es apropiada, donde $\hat{F}(x_{(i)})$ es la función de distribución empírica en el valor observado ordenado $x_{(i)}$. Además, cuando la gráfica es aproximadamente lineal, se pueden obtener estimaciones empíricas de α y β trazando una línea recta. Esta estimación empírica de β se puede utilizar como estimación inicial para resolver la ecuación numérica (2.19).

La matriz de información de Fisher esperada de $\theta = (\alpha, \beta)^\top$ basado en una sola observación viene dado por:

$$\begin{aligned} \mathbf{I}(\theta) &= [I_{ij}] = \left[-E \left(\frac{\partial^2 \log f(x_i; \theta)}{\partial \theta_i \partial \theta_j} \right) \right], \quad i, j = 1, 2 \\ &= \begin{pmatrix} \frac{1}{\alpha} & \frac{1}{\alpha\beta} (1 - \gamma - \log \alpha) \\ \frac{1}{\alpha\beta} (1 - \gamma - \log \alpha) & \frac{1}{6\beta^2} [\pi^2 + 6(1 - \gamma - \log \alpha)^2] \end{pmatrix} \end{aligned}$$

donde $\pi \approx 3.141593$ y $\gamma \approx 0.577216$ es la constante de Euler.

Bajo condiciones de regularidad leve la distribución asintótica del estimador de máxima verosimilitud $\hat{\theta}$ de θ es tal que

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{D} N(0, \mathbf{I}(\theta)),$$

donde $\mathbf{I}^{-1}(\theta)$ es la inversa de la matriz $\mathbf{I}(\theta)$, con,

$$\mathbf{I}^{-1}(\theta) = [\sigma_{ij}] = \begin{pmatrix} \frac{\alpha^2}{\pi^2} [\pi^2 + 6(1 - \gamma - \log \alpha)^2] & -\frac{6\alpha\beta}{\pi^2} (1 - \gamma - \log \alpha) \\ -\frac{6\alpha\beta}{\pi^2} (1 - \gamma - \log \alpha) & \frac{6\beta^2}{\pi^2} \end{pmatrix}$$

Los intervalos de confianza para α y β para muestra grande del $100(1 - \delta)\%$ respectivamente, están dados por,

$$\hat{\alpha} \pm z_{\delta/2} \frac{\sqrt{\hat{\sigma}_{11}}}{\sqrt{n}} \text{ y } \hat{\beta} \pm z_{\delta/2} \frac{\sqrt{\hat{\sigma}_{22}}}{\sqrt{n}},$$

donde $\hat{\sigma}_{11}$ y $\hat{\sigma}_{22}$ respectivamente, son las varianzas asintóticas estimadas de los estimadores de máxima verosimilitud $\hat{\alpha}$, $\hat{\beta}$ y z_q es el q -ésimo cuantil superior de la distribución normal estándar.

2.3. Función de cópulas

Una cópula es un método conveniente para la descripción de una distribución multivariante con estructura de dependencia. En el trabajo (Nelsen, 2007) se introdujo la función de cópulas como; una función que une funciones de distribución multivariadas con márgenes $[0, 1]$ uniformes. La cópula n -dimensional (C) existe para todo y_1, \dots, y_n , $F(y_1, \dots, y_n) = C(F_1(y_1), \dots, F_n(y_n))$, si F es continua, entonces C se define de forma única.

Proposición 8 (Función de Distribución). *Sklar (1973) establece que, consideradas las dos variables aleatorias Y_1 y Y_2 , con funciones de distribución $F_1(Y_1)$ y $F_2(Y_2)$, la fda es dada por:*

$$F(y_1, y_2) = C(F_1(y_1), F_2(y_2)) \quad (2.20)$$

Proposición 9 (Función de densidad). *Sklar (1973) establece que, consideradas las dos variables aleatorias Y_1 y Y_2 , con funciones de distribución $F_1(Y_1)$ y $F_2(Y_2)$, la fdp es dada por:*

$$f(y_1, y_2) = f_1(y_1)f_2(y_2)c(F_1(y_1), F_2(y_2)) \quad (2.21)$$

2.3.0.1. Cópula FGM

La cópula Farlie–Gumbel–Morgenstern (FGM) discutida por Gumbel (1960) es una de las familias de cópulas paramétricas más populares.

Proposición 10 (Función de densidad y acumulada). *La conjunta fda y la conjunta fdp para cópulas FGM son respectivamente,*

$$C(y_1, y_2) = F_1(y_1)F_2(y_2)(1 + \theta(1 - F_1(y_1))(1 - F_2(y_2))); \quad -1 < \theta < 1 \quad (2.22)$$

y

$$c(y_1, y_2) = (1 + \theta(1 - 2F_1(y_1))(1 - 2F_2(y_2))) \quad (2.23)$$

Proposición 11 (Coeficiente de correlación). *El coeficiente de correlación de Spearman y Kendall se dan de la siguiente manera:*

$$\rho_{sperman} = 12 \int \int uv(1 + \theta(1 - u)(1 - v))dudv - 3 = \frac{\theta}{3} \quad (2.24)$$

y

$$\rho_{Kendall} = 1 - 4 \int \int \frac{\partial C}{\partial u} C(u, v) \frac{\partial C}{\partial v} C(u, v) dudv = \frac{2}{9}\theta \quad (2.25)$$

donde se observa que $\frac{-1}{3} \leq \rho_{sperman} \leq \frac{1}{3}$ y $\frac{-2}{9} \leq \rho_{Kendall} \leq \frac{2}{9}$.

2.4. Distribución Weibull Bivariada FGM

La distribución Weibull es una distribución versátil y muy importante en el análisis de supervivencia. Con la cópula de Farlie-Gumbel-Morgenstern (FGM) y la distribución marginal Weibull se crea una distribución bivariada que se denomina distribución Weibull bivariada FGM, propuesta por Almetwally et al. (2020). Se usa frecuentement para describir datos bivariados que tienen una correlación débil entre las variables en los datos de por vida.

Proposición 12 (Función de densidad conjunta para cualquier cópula). *Consideradas las variables aleatorias Y_1 y Y_2 , con $Y_1 \sim W(\alpha_1, \beta_1)$ y $Y_2 \sim W(\alpha_2, \beta_2)$, la fdp es dada por:*

$$f(y_1, y_2) = \alpha_1 \beta_1 y_1^{\beta_1 - 1} e^{-\alpha_1 y_1^{\beta_1}} \alpha_2 \beta_2 y_2^{\beta_2 - 1} e^{-\alpha_2 y_2^{\beta_2}} c\left(\left(1 - e^{-\alpha_1 y_1^{\beta_1}}\right), \left(1 - e^{-\alpha_2 y_2^{\beta_2}}\right)\right) \quad (2.26)$$

Proposición 13 (Función acumulada conjunta FGM). *Consideradas las variables aleatorias Y_1 y Y_2 , con $Y_1 \sim W(\alpha_1, \beta_1)$ y $Y_2 \sim W(\alpha_2, \beta_2)$, la fda es dada por:*

$$F(y_1, y_2) = \left(1 - e^{-\alpha_1 y_1^{\beta_1}}\right) \left(1 - e^{-\alpha_2 y_2^{\beta_2}}\right) \times \left[1 + \theta \left(1 - \left(1 - e^{-\alpha_1 y_1^{\beta_1}}\right)\right) \left(1 - \left(1 - e^{-\alpha_2 y_2^{\beta_2}}\right)\right)\right] \quad (2.27)$$

Proposición 14 (Función de densidad conjunta FGM). *Consideradas las variables aleatorias Y_1 y Y_2 , con $Y_1 \sim W(\alpha_1, \beta_1)$ y $Y_2 \sim W(\alpha_2, \beta_2)$, la fdp es dada por:*

$$f(y_1, y_2) = \alpha_1 \beta_1 y_1^{\beta_1-1} e^{-\alpha_1 y_1^{\beta_1}} \alpha_2 \beta_2 y_2^{\beta_2-1} e^{-\alpha_2 y_2^{\beta_2}} \times \left[1 + \theta \left(1 - 2 \left(1 - e^{-\alpha_1 y_1^{\beta_1}} \right) \right) \left(1 - 2 \left(1 - e^{-\alpha_2 y_2^{\beta_2}} \right) \right) \right] \quad (2.28)$$

2.4.1. Propiedades

A continuación se dan algunas propiedades estadísticas importantes de la distribución weibull bivariada FGM como distribuciones marginales, momentos del producto, función generadora de momentos, distribución condicional, generación de variables aleatorias y función de confiabilidad.

Proposición 15 (Distribuciones marginales). *Consideradas las variables aleatorias Y_1 y Y_2 , con $Y_1 \sim W(\alpha_1, \beta_1)$ y $Y_2 \sim W(\alpha_2, \beta_2)$, las funciones de densidad marginal para Y_1 y Y_2 son respectivamente:*

$$f(y_1; \alpha_1, \beta_1) = \alpha_1 \beta_1 y_1^{\beta_1-1} e^{-\alpha_1 y_1^{\beta_1}}; \quad y_1 > 0, \quad \alpha_1, \quad \beta_1 > 0, \quad (2.29)$$

$$f(y_2; \alpha_2, \beta_2) = \alpha_2 \beta_2 y_2^{\beta_2-1} e^{-\alpha_2 y_2^{\beta_2}}; \quad y_2 > 0, \quad \alpha_2, \quad \beta_2 > 0, \quad (2.30)$$

tienen distribución weibull, donde la distribución marginal de Y_1 y Y_2 se puede calcular directamente por

$$f(y_i) = \int_{y_j} f(y_1, y_2) dy_j; \quad i, j = 1, 2, \quad i \neq j. \quad (2.31)$$

Proposición 16 (Distribución de densidad condicionada). *La función de densidad condicionada de Y_2 dado Y_1 se da como sigue:*

$$f(y_2|y_1) = \frac{f(y_1, y_2)}{f(y_1)} = \alpha_2 \beta_2 y_2^{\beta_2-1} e^{-\alpha_2 y_2^{\beta_2}} \left[1 + \theta \left(1 - 2 \left(1 - e^{-\alpha_1 y_1^{\beta_1}} \right) \right) \left(1 - 2 \left(1 - e^{-\alpha_2 y_2^{\beta_2}} \right) \right) \right] \quad (2.32)$$

La función de densidad condicionada de Y_1 dado Y_2 se da como sigue

$$f(y_1|y_2) = \frac{f(y_1, y_2)}{f(y_2)} = \alpha_1 \beta_1 y_1^{\beta_1-1} e^{-\alpha_1 y_1^{\beta_1}} \left[1 + \theta \left(1 - 2 \left(1 - e^{-\alpha_1 y_1^{\beta_1}} \right) \right) \left(1 - 2 \left(1 - e^{-\alpha_2 y_2^{\beta_2}} \right) \right) \right] \quad (2.33)$$

Proposición 17 (Distribución acumulada condicionada). *La función de distribución acumulada condicionada de Y_2 dado Y_1 se da como sigue:*

$$F(y_2|y_1) = 1 - e^{-\alpha_2 y_2^{\beta_2}} - \theta e^{-\alpha_2 y_2^{\beta_2}} + \theta e^{-2\alpha_2 y_2^{\beta_2}} + 2\theta e^{-\alpha_1 y_1^{\beta_1} - \alpha_2 y_2^{\beta_2}} - 2\theta e^{-\alpha_1 y_1^{\beta_1} - 2\alpha_2 y_2^{\beta_2}} \quad (2.34)$$

La función de distribución acumulada condicionada de Y_1 dado Y_2 se da como sigue:

$$F(y_1|y_2) = 1 - e^{-\alpha_1 y_1^{\beta_1}} - \theta e^{-\alpha_1 y_1^{\beta_1}} + \theta e^{-2\alpha_1 y_1^{\beta_1}} + 2\theta e^{-\alpha_1 y_1^{\beta_1} - \alpha_2 y_2^{\beta_2}} - 2\theta e^{-2\alpha_2 y_2^{\beta_2} - \alpha_1 y_1^{\beta_1}} \quad (2.35)$$

Proposición 18 (Generación de variables aleatorias). *Generación de una muestra a partir de una distribución conjunta específica. Por el método de distribución condicional, la función de distribución conjunta es la siguiente*

$$f(y_1, y_2) = f(y_1) f(y_2|y_1)$$

Al usar los siguientes pasos, podemos generar una muestra bivariada usando el enfoque condicional:

1. Genere U y V independientemente de una distribución uniforme $(0, 1)$
2. Sea $Y_1 = \beta_1 [-\ln(1 - U)]^{\alpha_1}$
3. Sea $F(y_2|y_1) = V$ para encontrar Y_2 mediante simulación numérica.
4. Repita los pasos 1 – 3 (n) veces para obtener (y_{1i}, y_{2i}) , $i = 1, 2, \dots, n$.

Proposición 19 (Función de confiabilidad). *La función de confiabilidad es más conveniente para expresar una función de supervivencia conjunta como una cópula de sus funciones marginales de supervivencia, donde Y_1 y Y_2 son variables aleatorias con funciones de supervivencia $\bar{F}(y_1)$ y $\bar{F}(y_2)$ como sigue. La función de confiabilidad de las distribuciones marginales se define como*

$$R(y_j; \alpha_j, \beta_j) = 1 - F(y_j; \alpha_j, \beta_j) = e^{-\alpha_j y_j^{\beta_j}}; \quad y > 0, \quad \alpha, \beta > 0, \quad j = 1, 2. \quad (2.36)$$

La expresión de la función de supervivencia conjunta para la cópula es la siguiente:

$$\mathbf{R}(y_1, y_2) = C(\mathbf{R}(y_1), \mathbf{R}(y_2))$$

entonces la función de confiabilidad de la distribución weibull FGM es:

$$\mathbf{R}(y_1, y_2) = e^{-\alpha_1 y_1^{\beta_1}} e^{-\alpha_2 y_2^{\beta_2}} \left[1 + \theta \left(1 - e^{-\alpha_1 y_1^{\beta_1}} \right) \left(1 - e^{-\alpha_2 y_2^{\beta_2}} \right) \right] \quad (2.37)$$

Proposición 20 (Función generadora de momentos). *Sea (Y_1, Y_2) una variable aleatoria con la función de densidad de probabilidad (3.24), entonces la función generadora de momentos de (Y_1, Y_2) está dada por:*

$$M_{(y_1, y_2)}(t_1, t_2) = \sum_{n=0}^{\infty} \left(\frac{t_1^n \beta_1^n}{n!} \Gamma\left(1 + \frac{n}{\alpha_1}\right) \right) \sum_{m=0}^{\infty} \left(\frac{t_2^m \beta_2^m}{m!} \Gamma\left(1 + \frac{m}{\alpha_2}\right) \right) \times \left[1 + \theta - 2\theta \frac{1}{2^{(1+\frac{m}{\alpha_2})}} - 2\theta \frac{1}{2^{(1+\frac{n}{\alpha_1})}} + 4\theta \frac{1}{2^{(1+\frac{n}{\alpha_1})}} \frac{1}{2^{(1+\frac{m}{\alpha_2})}} \right] \quad (2.38)$$

Para probar la función generadora de momentos, se comienza con

$$M_{(y_1, y_2)}(t_1, t_2) = E(e^{t_1 y_1} e^{t_2 y_2}) = \int_0^{\infty} \int_0^{\infty} e^{t_1 y_1} e^{t_2 y_2} f(y_1, y_2) dy_1 dy_2$$

2.4.2. Estimación por inferencia de funciones marginales

El método de estimación por inferencia de funciones marginales (Joe, 2005) (IFM, por sus siglas en inglés), también es llamado estimación en dos pasos de Joe. Es un método paramétrico con estimación en dos pasos. En el primer paso, cada distribución marginal se estima por separado.

$$\ln L_1 = \sum_{j=1}^n \ln f_1(y_{1j}, \delta_1); \quad \ln L_2 = \sum_{j=1}^n \ln f_2(y_{2j}, \delta_2) \quad (2.39)$$

Luego, en el segundo paso, el parámetro de la cópula se estima maximizando la función de log-verosimilitud de la densidad de la cópula utilizando las estimaciones de máxima verosimilitud de las marginales $\hat{F}_1(y_{1j}, \delta)$ y $\hat{F}_2(y_{2j}, \delta)$. La función de log-verosimilitud de una distribución de Weibull se define como

$$\begin{aligned} \ln L_j &= \sum_{i=1}^n \ln \left(\alpha_j \beta_j (y_{ji})^{\beta_j - 1} e^{-\alpha_j (y_{ji})^{\beta_j}} \right) \\ &= \sum_{i=1}^n \left(\ln(\alpha_j) + \ln(\beta_j) + (\beta_j - 1) \ln(y_{ji}) - \alpha_j (y_{ji})^{\beta_j} \right) \\ &= \sum_{i=1}^n \ln(\alpha_j) + \sum_{i=1}^n \ln(\beta_j) + \sum_{i=1}^n (\beta_j - 1) \ln(y_{ji}) - \sum_{i=1}^n \alpha_j (y_{ji})^{\beta_j} \\ &= n (\ln(\alpha_j) + \ln(\beta_j)) + (\beta_j - 1) \sum_{i=1}^n \ln(y_{ji}) - \alpha_j \sum_{i=1}^n (y_{ji})^{\beta_j} \end{aligned}$$

Los estimadores de máxima verosimilitud ($\hat{\alpha}_1, \hat{\beta}_1, \hat{\alpha}_2, \hat{\beta}_2$) se pueden obtener resolviendo

simultáneamente las ecuaciones de verosimilitud,

$$\frac{\partial \ln L}{\partial \beta_j} = 0, \quad \frac{\partial \ln L}{\partial \alpha_j} = 0, \quad j = 1, 2$$

Esto es

$$\frac{\partial \ln L_j}{\partial \beta_j} = \frac{n}{\beta_j} + \sum_{i=1}^n (\ln y_{ji}) - \alpha_j \sum_{i=1}^n (y_{ji})^{\beta_j} \ln(y_{ji}) = 0$$

$$\frac{\partial \ln L_j}{\partial \alpha_j} = \frac{n}{\alpha_j} - \sum_{i=1}^n (y_{ji})^{\beta_j} = 0$$

entonces

$$\hat{F}_j(y_j) = 1 - e^{-\hat{\alpha}_j(y_j)^{\hat{\beta}_j}}$$

y considerando el paso anterior, la estimación FIM de una distribución weibull bivariada se define como,

$$\ln L_{FIM} = \sum_{i=1}^n \ln(1 + \theta(1 - 2\hat{F}_1(y_{1i}))(1 - 2\hat{F}_2(y_{2i}))) \quad (2.40)$$

Las estimaciones de todos los parámetros se obtienen diferenciando la función logarítmica de verosimilitud con respecto a cada parámetro por separado. Basándose en esto, la diferenciación de la función de log-verosimilitud con respecto a θ , así,

$$\frac{\partial \ln L_{FIM}}{\partial \theta} = \sum_{i=1}^n \frac{(a(y_{1i}, \hat{\alpha}_1, \hat{\beta}_1))(a(y_{2i}, \hat{\alpha}_2, \hat{\beta}_2))}{(1 + \theta(a(y_{1i}, \hat{\alpha}_1, \hat{\beta}_1))(a(y_{2i}, \hat{\alpha}_2, \hat{\beta}_2)))} \quad (2.41)$$

Las estimaciones de los parámetros se manejan numéricamente simultáneamente las ecuaciones de verosimilitud,

$$\frac{\partial \ln L_{FIM}}{\partial \theta} = 0$$

No existe una expresión de forma cerrada para el estimador de máxima verosimilitud $\hat{\theta}$ y su cálculo debe realizarse numéricamente utilizando un algoritmo de optimización no lineal.

2.4.3. Estimación de máxima verosimilitud

El estimador de máxima verosimilitud (MLE, por sus siglas en inglés) sirve para estimar todos los parámetros del modelo en forma conjunta, es un método paramétrico de un solo

paso. Por lo tanto, la función de log-verosimilitud es,

$$\ln L = \sum_{i=1}^n [\ln(f_1(y_{1i})f_2(y_{2i})c(F_1(y_{1i}, \delta_1), F_2(y_{2i}, \delta_2); \theta))] \quad (2.42)$$

Las estimaciones de los parámetros se obtienen maximizando la función logarítmica de verosimilitud con la expectativa de cada parámetro por separado. Definimos

$$\begin{aligned} a(y_j; \alpha_j, \beta_j) &= 1 - 2F(y_j; \alpha_j, \beta_j) \\ &= 1 - 2 \left(1 - e^{-\alpha_j y_j^{\beta_j}} \right), \quad j = 1, 2. \end{aligned}$$

La función de verosimilitud de una distribución weibull bivariada FGM se define como

$$\begin{aligned} L &= (\alpha_1 \beta_1 \alpha_2 \beta_2)^n \prod_{i=1}^n \left((y_{1i}^{\beta_1 - 1}) (y_{2i}^{\beta_2 - 1}) \right) e^{-\sum_{i=1}^n (\alpha_1 y_{1i}^{\beta_1}) - \sum_{i=1}^n (\alpha_2 y_{2i}^{\beta_2})} \\ &\quad \times \prod_{i=1}^n (1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2))) \end{aligned} \quad (2.43)$$

y la función de log-verosimilitud se puede escribir como

$$\begin{aligned} \ln L &= n(\ln \alpha_1 + \ln \beta_1 + \ln \alpha_2 + \ln \beta_2) + (\beta_1 - 1) \sum_{i=1}^n \ln(y_{1i}) + (\beta_2 - 1) \sum_{i=1}^n \ln(y_{2i}) \\ &\quad - \sum_{i=1}^n (\alpha_1 y_{1i}^{\beta_1}) - \sum_{i=1}^n (\alpha_2 y_{2i}^{\beta_2}) + \sum_{i=1}^n \ln(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2))) \end{aligned} \quad (2.44)$$

Las estimaciones de todos los parámetros se obtienen diferenciando la función de log-verosimilitud (2.44) con respecto a cada parámetro por separado, como sigue

$$\frac{\partial \ln L}{\partial \alpha_1} = \frac{n}{\alpha_1} - \sum_{i=1}^n y_{1i}^{\beta_1} + \sum_{i=1}^n \frac{-2\theta(a(y_{2i}; \alpha_2, \beta_2)) y_{1i}^{\beta_1} e^{-\alpha_1 y_{1i}^{\beta_1}}}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))}$$

$$\frac{\partial \ln L}{\partial \alpha_2} = \frac{n}{\alpha_2} - \sum_{i=1}^n y_{2i}^{\beta_2} + \sum_{i=1}^n \frac{-2\theta(a(y_{1i}; \alpha_1, \beta_1)) y_{2i}^{\beta_2} e^{-\alpha_2 y_{2i}^{\beta_2}}}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))}$$

$$\begin{aligned} \frac{\partial \ln L}{\partial \beta_1} &= \frac{n}{\beta_1} + \sum_{i=1}^n \ln(y_{1i}) - \alpha_1 \sum_{i=1}^n y_{1i}^{\beta_1} \ln(y_{1i}) \\ &+ \sum_{i=1}^n \frac{-2\alpha_1 \theta(a(y_{1i}; \alpha_1, \beta_1)) e^{-\alpha_1 y_{1i}^{\beta_1}} y_{1i}^{\beta_1} \ln(y_{1i})}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))} \end{aligned}$$

$$\begin{aligned} \frac{\partial \ln L}{\partial \beta_2} &= \frac{n}{\beta_2} + \sum_{i=1}^n \ln(y_{2i}) - \alpha_2 \sum_{i=1}^n y_{2i}^{\beta_2} \ln(y_{2i}) \\ &+ \sum_{i=1}^n \frac{-2\alpha_2 \theta(a(y_{2i}; \alpha_2, \beta_2)) e^{-\alpha_2 y_{2i}^{\beta_2}} y_{2i}^{\beta_2} \ln(y_{2i})}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))} \end{aligned}$$

y

$$\frac{\partial \ln L}{\partial \theta} = \sum_{i=1}^n \frac{(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2))}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))}$$

El estimador de máxima verosimilitud $\hat{\delta} = (\hat{\alpha}_1, \hat{\beta}_1, \hat{\alpha}_2, \hat{\beta}_2, \hat{\theta})$ se puede obtener resolviendo simultáneamente las ecuaciones de log-verosimilitud

$$\frac{\partial \ln L}{\partial \theta} = 0, \quad \frac{\partial \ln L}{\partial \beta_j} = 0, \quad \frac{\partial \ln L}{\partial \alpha_j} = 0, \quad j = 1, 2$$

Pero las ecuaciones deben realizarse numéricamente utilizando un algoritmo de optimización no lineal.

2.4.4. Estimación por método semiparamétrico

Kundu and Dey (2009), introdujo la Estimación que se realiza en dos etapas como en IFM, pero la diferencia es que las distribuciones marginales se estiman no paramétricamente por sus distribuciones empíricas muestrales. En este método, las observaciones se transforman en pseudo-observaciones utilizando la función de distribución empírica de cada distribución marginal. La función de distribución empírica se define como:

$$\tilde{F}_i(y_i) = \frac{\sum_{j=1}^n I(Y_{i,j} \leq y_i)}{n+1}; \quad i = 1, 2 \quad (2.45)$$

Entonces, θ es estimado por el maximizador de la pseudo verosimilitud,

$$\sum_{i=1}^n \ln c(\tilde{F}_1(Y_{1i}), \tilde{F}_2(Y_{2i}); \theta) \quad (2.46)$$

Considerando la ecuación (2.46), la función de log-verosimilitud de una distribución weibull unitaria bivariada se define,

$$\ln L_{SP} = \sum_{i=1}^n \ln(1 + \theta(1 - 2\tilde{F}_1(y_{1i}))(1 - 2\tilde{F}_2(y_{2i})))$$

No existe una expresión de forma cerrada para el estimador de máxima verosimilitud θ mediante el uso de ecuación (anterior) y su cálculo debe realizarse numéricamente utilizando un software estadístico.

Capítulo 3

Distribución Weibull Unitaria Bivariada

En este capítulo se propone una extensión bivariada de la distribución WU propuesta por Mazucheli et al. (2018b). Esta extensión obtenida a partir de la cópula de Farlie-Gumbel-Morgenstern (FGM) (Gumbel, 1960) es denominada Weibull Unitaria Bivariada (WUB) y es útil para ajustar datos bivariados en el el plano $(0, 1) \times (0, 1)$.

Para la distribución WUB, se determinan las funciones de densidad de probabilidad conjunta, distribución acumulada conjunta y momentos. Adicionalmente, se lleva a cabo la estimación de los parámetros de la distribución desde una perspectiva clásica usando el método de MV. También se lleva a cabo un estudio de simulación de Monte Carlo para estudiar las propiedades de los estimadores de MV. Finalmente, se presenta una aplicación a un conjunto de datos reales para ilustrar la utilidad de la distribución propuesta.

Definición 3. *Un vector aleatorio $\mathbf{Y} = (Y_1, Y_2)^\top$ se dice que tiene distribución WUB si su fdp conjunta está dada por,*

$$f(y_1, y_2) = \frac{1}{y_1} \alpha_1 \beta_1 (-\ln y_1)^{\beta_1 - 1} e^{-\alpha_1 (-\ln y_1)^{\beta_1}} \frac{1}{y_2} \alpha_2 \beta_2 (-\ln y_2)^{\beta_2 - 1} e^{-\alpha_2 (-\ln y_2)^{\beta_2}} \\ \times \left(1 + \theta \left(1 - 2e^{-\alpha_1 (-\ln y_1)^{\beta_1}} \right) \left(1 - 2e^{-\alpha_2 (-\ln y_2)^{\beta_2}} \right) \right), \quad 0 < y_1, y_2 < 1 \quad (3.1)$$

para $-1 < \theta < 1$ y $\alpha_1, \alpha_2, \beta_1, \beta_2 > 0$

Una vector aleatorio bidimensional \mathbf{Y} con distribución WUB se denota como $\mathbf{Y} \sim WUB(\varphi)$, donde $\varphi = (\alpha_1, \alpha_2, \beta_1, \beta_2, \theta)^\top$. La fdp conjunta en la Ecuación (3.1) es obtenida usando la cópula de FGM. En la Figura 3.1 se muestra el comportamiento de la fdp conjunta de una distribución WUB para algunos valores seleccionados de los parámetros de la distribución. Nótese que la fdp conjunta WUB presenta diferentes formas en las que se incluyen formas simétricas, asimétricas y en forma de bañera, entre otras.

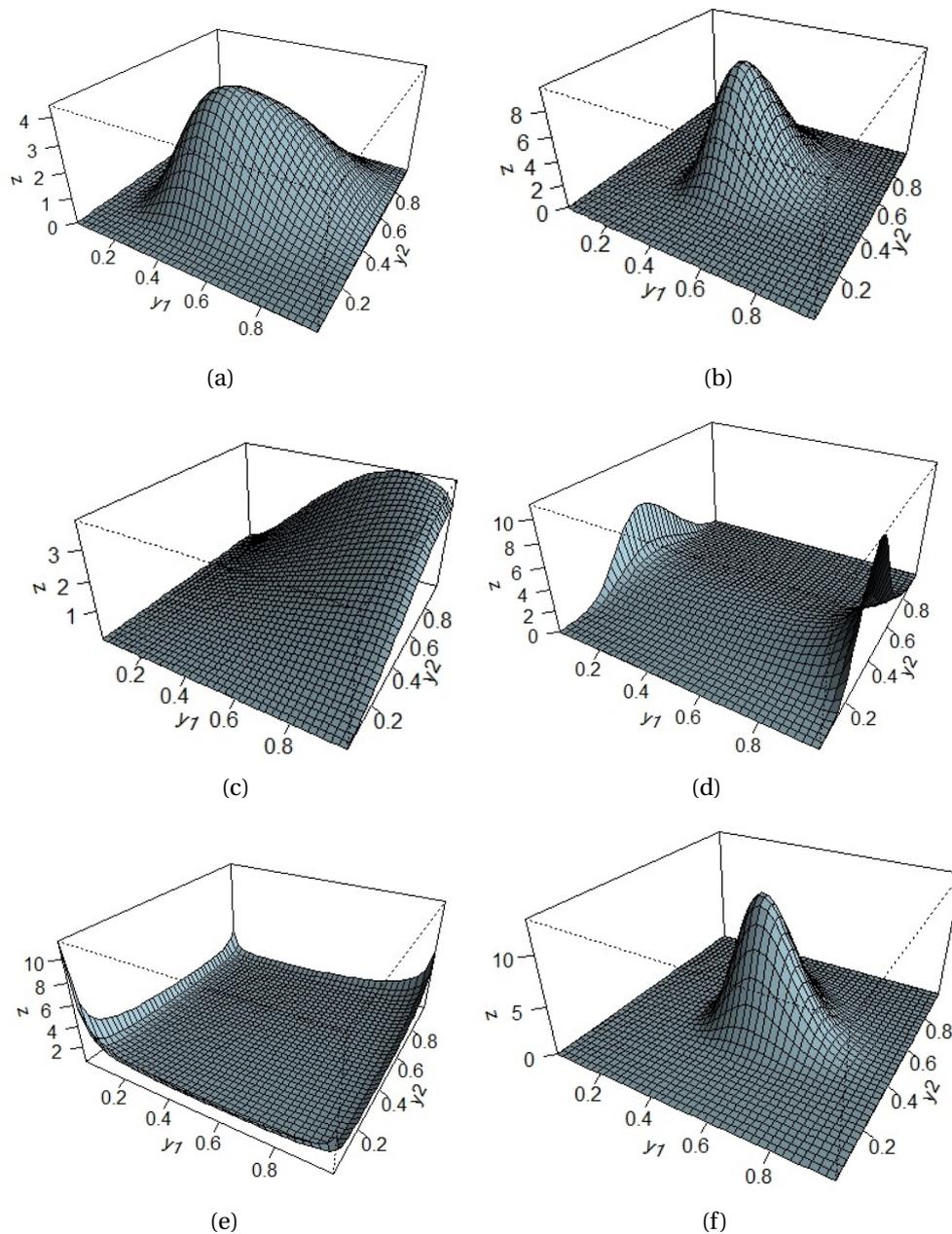


Figura 3.1: Función de densidad conjunta de la distribución weibull unitaria para diferentes valores de los parámetros. **(a)** $\varphi = (1.5, 1.5, 2.0, 2.0, 0.75)$, **(b)** $\varphi = (2.2, 2.5, 3.5, 2.5, 0.25)$, **(c)** $\varphi = (2.2, 2.5, 1.1, 1.0, -0.25)$, **(d)** $\varphi = (1.0, 1.2, 0.5, 2.5, -0.75)$, **(e)** $\varphi = (0.9, 0.9, 0.7, 0.7, 0.5)$, **(f)** $\varphi = (3.0, 3.0, 3.5, 3.5, -0.5)$.

En la Figura 3.1 se observa que la fdp conjunta WUB puede tomar muchas formas dependiendo del valor de los parámetros, en particular la gráfica **(d)** es muy similar a una silla de montar, las gráficas **(a)**, **(b)** y **(f)** tienen formas de lomas.

Proposición 21 (Función de distribución conjunta WUB). *Un vector aleatorio $\mathbf{Y} = (Y_1, Y_2)^\top$ se dice que tiene distribución WUB si su fda conjunta está dada por:*

$$F(y_1, y_2) = e^{-\alpha_1(-\ln y_1)^{\beta_1}} e^{-\alpha_2(-\ln y_2)^{\beta_2}} \left(1 + \theta \left(1 - e^{-\alpha_1(-\ln y_1)^{\beta_1}} \right) \left(1 - e^{-\alpha_2(-\ln y_2)^{\beta_2}} \right) \right) \quad (3.2)$$

Para $-1 < \theta < 1$; $0 < y_1, y_2 < 1$; $\alpha_1, \alpha_2, \beta_1, \beta_2 > 0$

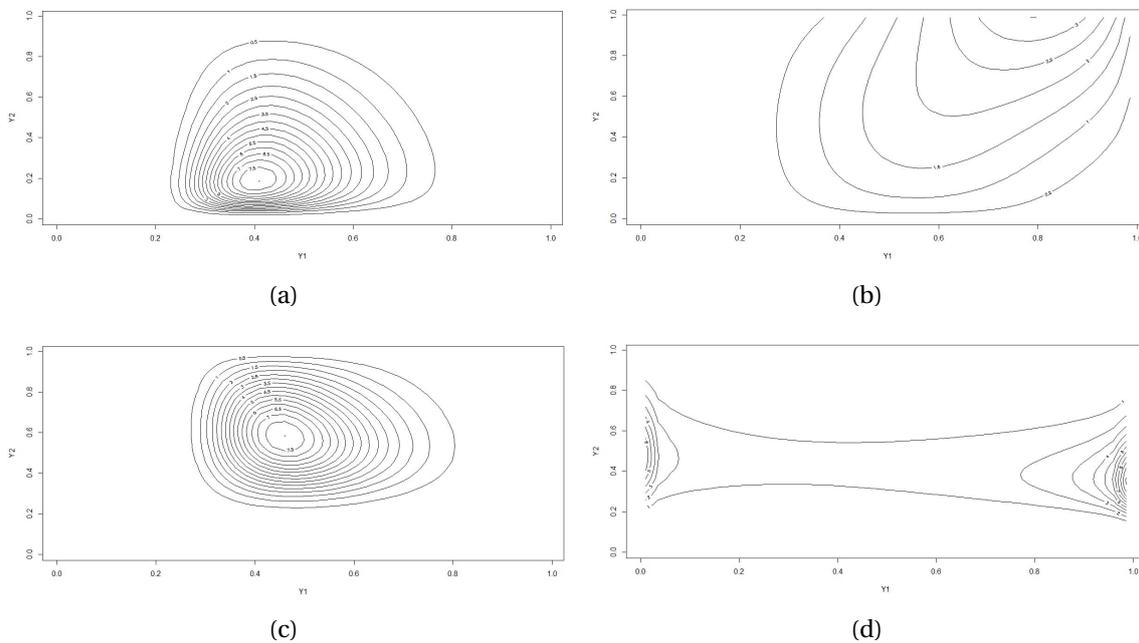


Figura 3.2: Gráficas de contorno para varios valores de los parámetros. **(a)** $\varphi = (1.5, 0.5, 3.5, 2.0, 0.25)$, **(b)** $\varphi = (2.5, 1.5, 1.5, 1.0, 0.75)$, **(c)** $\varphi = (2.2, 2.5, 3.5, 2.0, -0.25)$, **(d)** $\varphi = (1.0, 1.2, 0.5, 2.5, -0.75)$.

En la Figura 3.2 se muestra 4 gráficas de contorno para la distribución weibull unitaria bivariada FGM con diferentes valores de los parámetros. Las gráficas **(a)** y **(c)** son similares, ambas mostrando que son unimodales, centradas en $(0.4, 0.2)$ y $(0.46, 0.6)$ respectivamente, la gráfica **(b)** también muestra que es unimodal cenrada en $(0.8, 1)$, pero con mucha más variabilidad y en la gráfica **(d)** se observan pequeños indicios de bimodalidad bivariada, con posibles modas $(0, 0.44)$ y $(1, 0.4)$.

3.1. Propiedades de la distribución weibull unitaria bivariada FGM

En esta sección, damos algunas propiedades estadísticas importantes de la distribución Weibull Unitaria Bivariada FGM como distribuciones marginales, momentos, distribución condicional, función de confiabilidad, generación de variables aleatorias. Algunas expresiones algebraicas para determinar algunas propiedades estadísticas de la distribución Weibull Unitaria Bivariada FGM pueden ser más eficientes que calcularlas directamente mediante simulación numérica.

3.1.1. Distribuciones marginales

En esta sección se presentan las funciones de densidad marginales en forma general para Y_1 e Y_2 respectivamente.

Proposición 22. Sea $\mathbf{Y} = (Y_1, Y_2)^\top$ un vector aleatorio con distribución $WUB(\varphi)$, entonces,

1. La fdp marginal de Y_j es:

$$f(y_j) = \frac{1}{y_j} \alpha_j \beta_j (-\ln y_j)^{\beta_j - 1} e^{-\alpha_j (-\ln y_j)^{\beta_j}} \quad 0 < y_j < 1, \quad \alpha_j, \beta_j > 0, \quad j = 1, 2.$$

Demostración. 1. Por definición las funciones de densidad marginales se pueden hallar,

$$f(y_i) = \int_0^1 f(y_1, y_2) dy_j; \quad i, j = 1, 2, \quad i \neq j.$$

por lo tanto, la fdp marginal de Y_j se puede hallar de la siguiente manera,

$$\begin{aligned} f(y_j) = & \int_0^1 \frac{1}{y_j} \alpha_j \beta_j (-\ln y_j)^{\beta_j - 1} e^{-\alpha_j (-\ln y_j)^{\beta_j}} \frac{1}{y_i} \alpha_i \beta_i (-\ln y_i)^{\beta_i - 1} e^{-\alpha_i (-\ln y_i)^{\beta_i}} \\ & \times \left(1 + \theta \left(1 - 2e^{-\alpha_j (-\ln y_j)^{\beta_j}} \right) \left(1 - 2e^{-\alpha_i (-\ln y_i)^{\beta_i}} \right) \right) dy_i \end{aligned}$$

resolviendo primero las operaciones dentro de parentesis se tiene lo siguiente:

$$f(y_j) = \int_0^1 \frac{1}{y_j} \alpha_j \beta_j (-\ln y_j)^{\beta_j-1} e^{-\alpha_j (-\ln y_j)^{\beta_j}} \frac{1}{y_i} \alpha_i \beta_i (-\ln y_i)^{\beta_i-1} e^{-\alpha_i (-\ln y_i)^{\beta_i}} \\ \times \left[1 + \theta (1 - 2e^{-\alpha_i (-\ln y_i)^{\beta_i}} - 2e^{-\alpha_j (-\ln y_j)^{\beta_j}} + 4e^{-\alpha_i (-\ln y_i)^{\beta_i} - \alpha_j (-\ln y_j)^{\beta_j}} \right] dy_i$$

haciendo las respectivos productos se obtiene,

$$f(y_j) = \frac{1}{y_j} \alpha_j \beta_j (-\ln y_j)^{\beta_j-1} e^{-\alpha_j (-\ln y_j)^{\beta_j}} \int_0^1 \frac{1}{y_i} \alpha_i \beta_i (-\ln y_i)^{\beta_i-1} e^{-\alpha_i (-\ln y_i)^{\beta_i}} dy_i + \\ \frac{1}{y_j} \alpha_j \beta_j (-\ln y_j)^{\beta_j-1} e^{-\alpha_j (-\ln y_j)^{\beta_j}} \theta \int_0^1 \frac{1}{y_i} \alpha_i \beta_i (-\ln y_i)^{\beta_i-1} e^{-\alpha_i (-\ln y_i)^{\beta_i}} dy_i - \\ \frac{1}{y_j} \alpha_j \beta_j (-\ln y_j)^{\beta_j-1} e^{-\alpha_j (-\ln y_j)^{\beta_j}} 2\theta \int_0^1 \frac{1}{y_i} \alpha_i \beta_i (-\ln y_i)^{\beta_i-1} e^{-2\alpha_i (-\ln y_i)^{\beta_i}} dy_i - \\ \frac{1}{y_j} \alpha_j \beta_j (-\ln y_j)^{\beta_j-1} e^{-\alpha_j (-\ln y_j)^{\beta_j}} 2\theta e^{-\alpha_j (-\ln y_j)^{\beta_j}} \int_0^1 \frac{1}{y_i} \alpha_i \beta_i (-\ln y_i)^{\beta_i-1} e^{-\alpha_i (-\ln y_i)^{\beta_i}} dy_i + \\ \frac{1}{y_j} \alpha_j \beta_j (-\ln y_j)^{\beta_j-1} e^{-\alpha_j (-\ln y_j)^{\beta_j}} 4\theta e^{-\alpha_j (-\ln y_j)^{\beta_j}} \int_0^1 \frac{1}{y_i} \alpha_i \beta_i (-\ln y_i)^{\beta_i-1} e^{-2\alpha_i (-\ln y_i)^{\beta_i}} dy_i$$

Entonces,

$$f(y_j) = \frac{1}{y_j} \alpha_j \beta_j (-\ln y_j)^{\beta_j-1} e^{-\alpha_j (-\ln y_j)^{\beta_j}} \left(\int_0^1 \frac{1}{y_i} \alpha_i \beta_i (-\ln y_i)^{\beta_i-1} e^{-\alpha_i (-\ln y_i)^{\beta_i}} dy_i + \right. \\ \left. \theta \int_0^1 \frac{1}{y_i} \alpha_i \beta_i (-\ln y_i)^{\beta_i-1} e^{-\alpha_i (-\ln y_i)^{\beta_i}} dy_i - 2\theta \int_0^1 \frac{1}{y_i} \alpha_i \beta_i (-\ln y_i)^{\beta_i-1} e^{-2\alpha_i (-\ln y_i)^{\beta_i}} dy_i - \right. \\ \left. 2\theta \int_0^1 \frac{1}{y_i} \alpha_i \beta_i (-\ln y_i)^{\beta_i-1} e^{-\alpha_i (-\ln y_i)^{\beta_i} - \alpha_j (-\ln y_j)^{\beta_j}} dy_i + \right. \\ \left. 4\theta \int_0^1 \frac{1}{y_i} \alpha_i \beta_i (-\ln y_i)^{\beta_i-1} e^{-2\alpha_i (-\ln y_i)^{\beta_i} - \alpha_j (-\ln y_j)^{\beta_j}} dy_i \right)$$

Notese que,

$$\int \frac{1}{y} \alpha \beta (-\ln y)^{\beta-1} e^{-\alpha (-\ln y)^{\beta}} dy$$

$$\text{haciendo } u = -\alpha (-\ln y)^{\beta} \Rightarrow du = \frac{1}{y} \alpha \beta (-\ln y)^{\beta-1}$$

por lo tanto,

$$\int \frac{1}{y} \alpha \beta (-\ln y)^{\beta-1} e^{-\alpha(-\ln y)^\beta} dy = \int e^u du = e^u + c = e^{-\alpha(-\ln y)^\beta} + c$$

así, se tiene que,

$$\int_0^1 \frac{1}{y_i} \alpha_i \beta_i (-\ln y_i)^{\beta_i-1} e^{-\alpha_i(-\ln y_i)^{\beta_i}} dy_i = \left[e^{-\alpha_i(-\ln y_i)^{\beta_i}} \right]_0^1 = 1$$

de esta forma se tiene que,

$$\begin{aligned} f(y_j) &= \frac{1}{y_j} \alpha_j \beta_j (-\ln y_j)^{\beta_j-1} e^{-\alpha_j(-\ln y_j)^{\beta_j}} \left[1 + \theta - \theta - 2\theta e^{-\alpha_j(-\ln y_j)^{\beta_j}} + 2\theta e^{-\alpha_j(-\ln y_j)^{\beta_j}} \right] \\ &= \frac{1}{y_j} \alpha_j \beta_j (-\ln y_j)^{\beta_j-1} e^{-\alpha_j(-\ln y_j)^{\beta_j}} [1] \\ &= \frac{1}{y_j} \alpha_j \beta_j (-\ln y_j)^{\beta_j-1} e^{-\alpha_j(-\ln y_j)^{\beta_j}} \end{aligned}$$

□

3.1.2. Distribuciones condicionales

En esta sección se presenta la función de densidad condicional y la función de distribución condicional en forma general.

Proposición 23. Sea $\mathbf{Y} = (Y_1, Y_2)^\top$ un vector aleatorio con distribución WUB(φ), entonces,

1. La función de densidad condicionada de Y_i dado Y_j se da como sigue:

$$f(y_i|y_j) = \frac{1}{y_i} \alpha_i \beta_i (-\ln y_i)^{\beta_i-1} e^{-\alpha_i(-\ln y_i)^{\beta_i}} \times \left(1 + \theta \left(1 - 2e^{-\alpha_j(-\ln y_j)^{\beta_j}} \right) \left(1 - 2e^{-\alpha_i(-\ln y_i)^{\beta_i}} \right) \right)$$

para $i, j = 1, 2, i \neq j$.

Demostración. 1. Por definición de función de densidad de tiene que:

$$\begin{aligned} f(y_i|y_j) &= \frac{f(y_j, y_i)}{f(y_j)} = \frac{f(y_j) f(y_i) \left(1 + \theta \left(1 - 2e^{-\alpha_j(-\ln y_j)^{\beta_j}} \right) \left(1 - 2e^{-\alpha_i(-\ln y_i)^{\beta_i}} \right) \right)}{f(y_i)} \\ &= f(y_i) \left(1 + \theta \left(1 - 2e^{-\alpha_j(-\ln y_j)^{\beta_j}} \right) \left(1 - 2e^{-\alpha_i(-\ln y_i)^{\beta_i}} \right) \right) \\ &= \frac{1}{y_i} \alpha_i \beta_i (-\ln y_i)^{\beta_i-1} e^{-\alpha_i(-\ln y_i)^{\beta_i}} \times \left(1 + \theta \left(1 - 2e^{-\alpha_j(-\ln y_j)^{\beta_j}} \right) \left(1 - 2e^{-\alpha_i(-\ln y_i)^{\beta_i}} \right) \right) \end{aligned}$$

□

Proposición 24. Sea $\mathbf{Y} = (Y_1, Y_2)^\top$ un vector aleatorio con distribución WUB(φ), entonces,

1. La función de distribución de Y_i dado Y_j se da como sigue:

$$F(y_i|y_j) = e^{-\alpha_i(-\ln y_i)^{\beta_i}} + \theta e^{-\alpha_i(-\ln y_i)^{\beta_i}} - \theta e^{-2\alpha_i(-\ln y_i)^{\beta_i}} - 2\theta e^{-\alpha_j(-\ln y_j)^{\beta_j} - \alpha_i(-\ln y_i)^{\beta_i}} + 2\theta e^{-\alpha_j(-\ln y_j)^{\beta_j} - 2\alpha_i(-\ln y_i)^{\beta_i}}$$

Demostración. 1. La función de distribución acumulada de Y_i dado Y_j por definición se puede hallar,

$$F(y_i|y_j) = \frac{\frac{\partial F(y_j, y_i)}{\partial y_j}}{f(y_j)}$$

primero notese que,

$$\begin{aligned} F(y_j, y_i) &= F_j(y_j)F_i(y_i)(1 + \theta(1 - F_j(y_j))(1 - F_i(y_i))) \\ &= e^{-\alpha_j(-\ln y_j)^{\beta_j}} e^{-\alpha_i(-\ln y_i)^{\beta_i}} \left(1 + \theta \left(1 - e^{-\alpha_j(-\ln y_j)^{\beta_j}}\right) \left(1 - e^{-\alpha_i(-\ln y_i)^{\beta_i}}\right)\right) \\ &= e^{-\alpha_j(-\ln y_j)^{\beta_j}} e^{-\alpha_i(-\ln y_i)^{\beta_i}} \left(1 + \theta - \theta e^{-\alpha_i(-\ln y_i)^{\beta_i}} - \theta e^{-\alpha_j(-\ln y_j)^{\beta_j}} + \theta e^{-\alpha_j(-\ln y_j)^{\beta_j} - \alpha_i(-\ln y_i)^{\beta_i}}\right) \\ &= e^{-\alpha_j(-\ln y_j)^{\beta_j} - \alpha_i(-\ln y_i)^{\beta_i}} + \theta e^{-\alpha_j(-\ln y_j)^{\beta_j} - \alpha_i(-\ln y_i)^{\beta_i}} - \theta e^{-\alpha_j(-\ln y_j)^{\beta_j} - 2\alpha_i(-\ln y_i)^{\beta_i}} - \theta e^{-2\alpha_j(-\ln y_j)^{\beta_j} - \alpha_i(-\ln y_i)^{\beta_i}} + \theta e^{-2\alpha_j(-\ln y_j)^{\beta_j} - 2\alpha_i(-\ln y_i)^{\beta_i}} \end{aligned}$$

ahora se calcula la derivada con respecto a y_j ,

$$\begin{aligned} \frac{\partial F(y_j, y_i)}{\partial y_j} &= \frac{1}{y_j} \alpha_j \beta_j (-\ln y_j)^{\beta_j - 1} e^{-\alpha_j(-\ln y_j)^{\beta_j} - \alpha_i(-\ln y_i)^{\beta_i}} \\ &\quad + \theta \frac{1}{y_j} \alpha_j \beta_j (-\ln y_j)^{\beta_j - 1} e^{-\alpha_j(-\ln y_j)^{\beta_j} - \alpha_i(-\ln y_i)^{\beta_i}} \\ &\quad - \theta \frac{1}{y_j} \alpha_j \beta_j (-\ln y_j)^{\beta_j - 1} e^{-\alpha_j(-\ln y_j)^{\beta_j} - 2\alpha_i(-\ln y_i)^{\beta_i}} \\ &\quad - 2\theta \frac{1}{y_j} \alpha_j \beta_j (-\ln y_j)^{\beta_j - 1} e^{-2\alpha_j(-\ln y_j)^{\beta_j} - \alpha_i(-\ln y_i)^{\beta_i}} \\ &\quad + 2\theta \frac{1}{y_j} \alpha_j \beta_j (-\ln y_j)^{\beta_j - 1} e^{-2\alpha_j(-\ln y_j)^{\beta_j} - 2\alpha_i(-\ln y_i)^{\beta_i}} \end{aligned}$$

ahora realizando el cociente,

$$\frac{\frac{\partial F(y_j, y_i)}{\partial y_j}}{f(y_j)}$$

se obtiene que la función de distribución acumulada de Y_j dado Y_i es,

$$F(y_i|y_j) = e^{-\alpha_i(-\ln y_i)^{\beta_i}} + \theta e^{-\alpha_i(-\ln y_i)^{\beta_i}} - \theta e^{-2\alpha_i(-\ln y_i)^{\beta_i}} - 2\theta e^{-\alpha_j(-\ln y_j)^{\beta_j} - \alpha_i(-\ln y_i)^{\beta_i}} + 2\theta e^{-\alpha_j(-\ln y_j)^{\beta_j} - 2\alpha_i(-\ln y_i)^{\beta_i}}$$

□

3.1.3. Generación de variables aleatorias

Para la generación de una muestra a partir de una distribución conjunta Weibull Unitaria se utiliza el método de distribución condicional propuesto por Nelsen (2007), la función de distribución conjunta se puede escribir de la siguiente manera,

$$f(y_1, y_2) = f(y_1)f(y_2|y_1)$$

Al usar los siguientes pasos, podemos generar una muestra bivariada usando el enfoque condicional:

1. Genere U y V independientemente de una distribución uniforme $(0, 1)$
2. Sea $Y_1 = e^{-\left(-\frac{\ln U}{\alpha_1}\right)^{1/\beta_1}}$
3. Sea $F(y_2|y_1) = V$ para así encontrar a Y_2 numéricamente, por ejemplo, usando el método de Brent disponible en el software R Core Team (2021) a través de la función "uniroot"
4. Repita los pasos 1 – 3 (n) veces para obtener $(y_{1i}, y_{2i}), i = 1, 2, \dots, n$.

3.1.4. Momentos del producto

En esta sección se presentan los momentos del producto para la distribución Weibull Unitaria Bivariada (WUB).

Proposición 25. Sea $\mathbf{Y} = (Y_1, Y_2)^\top$ un vector aleatorio con distribución $WUB(\varphi)$, entonces los momentos r -ésimo y s -ésimo alrededor de cero se puede calcular de la siguiente forma:

$$\begin{aligned} \mu_{rs} = E(Y_1^r Y_2^s) &= \int_0^1 \int_0^1 (y_1^r y_2^s) \frac{1}{y_1} \alpha_1 \beta_1 (-\ln y_1)^{\beta_1 - 1} e^{-\alpha_1 (-\ln y_1)^{\beta_1}} \frac{1}{y_2} \alpha_2 \beta_2 (-\ln y_2)^{\beta_2 - 1} e^{-\alpha_2 (-\ln y_2)^{\beta_2}} \\ &\times \left(1 + \theta \left(1 - 2e^{-\alpha_1 (-\ln y_1)^{\beta_1}} \right) \left(1 - 2e^{-\alpha_2 (-\ln y_2)^{\beta_2}} \right) \right) dy_1 dy_2 \end{aligned} \quad (3.3)$$

Proposición 26 (Coeficiente de correlación). *Sea $\mathbf{Y} = (Y_1, Y_2)^\top$ un vector aleatorio con distribución $WUB(\varphi)$, entonces el coeficiente de correlación lineal entre Y_1 y Y_2 está dado por:*

$$\rho = cor(Y_1, Y_2) = \frac{E(Y_1 Y_2) - E(Y_1)E(Y_2)}{\sigma_1 \sigma_2} \quad (3.4)$$

donde,

$$\sigma_j = \sqrt{E(Y_j - E(Y_j))^2} \quad \text{para } j = 1, 2.$$

Por medio de una simulación numérica realizada en el programa R Core Team (2021), se encontró que el rango de valores posibles para el coeficiente de correlación es:

$$-\frac{1}{3} \leq \rho \leq \frac{1}{3}$$

Algunos valores del coeficiente de correlación para algunos valores de los parámetros se presentan en la Tabla 3.1.

Tabla 3.1: Valores del coeficiente de correlación.

α_1	α_2	β_1	β_2	θ	ρ
3.0	2.0	1.0	1.0	-0.75	-0.235
3.0	2.0	1.0	1.0	-0.25	-0.078
3.0	2.0	1.0	1.0	0.25	0.078
3.0	2.0	1.0	1.0	0.75	0.235
1.5	2.5	1.0	1.0	-0.99	-0.317
1.5	2.5	1.0	1.0	0.5	0.16
4.5	1.5	1.0	1.0	-0.99	-0.307
4.5	1.5	1.0	1.0	0.99	0.307
7.0	2.0	1.0	1.0	-0.9	-0.269
7.0	2.0	1.0	1.0	-0.5	-0.15
7.0	2.0	1.0	1.0	-0.1	-0.03
3.5	1.8	1.0	1.0	0.9	0.281
2.0	12.0	1.0	2.5	-0.8	-0.258
3.5	1.8	1.8	1.5	0.99	0.325
0.9	3.0	1.0	1.0	-0.99	-0.315

3.1.5. Función de confiabilidad

La función de confiabilidad es más conveniente para expresar una función de supervivencia conjunta como una cópula de sus funciones marginales de supervivencia, donde Y_1 y Y_2 son variables aleatorias con funciones de supervivencia $\bar{F}(y_1)$ y $\bar{F}(y_2)$ como sigue,

Proposición 27. *La función de confiabilidad de las distribuciones marginales Weibull Unitarias para el caso univariado se da como sigue:*

$$R(y_j) = 1 - F(y_j) = 1 - e^{-\alpha_j(-\ln y_j)^{\beta_j}}; \quad 0 < y < 1, \quad \alpha, \beta > 0, \quad j = 1, 2.$$

Proposición 28. *La función de supervivencia conjunta para la cópula FGM (Nelsen, 2007) es la siguiente*

$$\bar{H}(y_1, y_2) = 1 - F(y_1) - F(y_2) + C(F(y_1), F(y_2))$$

la función de confiabilidad de la distribución conjunta weibull unitaria FGM es

$$\begin{aligned} \bar{H}(y_1, y_2) = & 1 - e^{-\alpha_1(-\ln y_1)^{\beta_1}} - e^{-\alpha_2(-\ln y_2)^{\beta_2}} \dots \\ & + e^{-\alpha_1(-\ln y_1)^{\beta_1}} e^{-\alpha_2(-\ln y_2)^{\beta_2}} \left[1 + \theta \left(1 - e^{-\alpha_1(-\ln y_1)^{\beta_1}} \right) \left(1 - e^{-\alpha_2(-\ln y_2)^{\beta_2}} \right) \right] \end{aligned}$$

Proposición 29. *Basu (1971) define la función hazard para el caso bivariado como*

$$h(y_1, y_2) = \frac{f(y_1, y_2)}{\bar{H}(y_1, y_2)}$$

Entonces la función hazard de la distribución conjunta weibull unitaria FGM es,

$$\begin{aligned} h(y_1, y_2) = & \frac{\frac{1}{y_1} \alpha_1 \beta_1 (-\ln y_1)^{\beta_1-1} e^{-\alpha_1(-\ln y_1)^{\beta_1}} \frac{1}{y_2} \alpha_2 \beta_2 (-\ln y_2)^{\beta_2-1} e^{-\alpha_2(-\ln y_2)^{\beta_2}}}{1 - F(y_1) - F(y_2) + e^{-\alpha_1(-\ln y_1)^{\beta_1}} e^{-\alpha_2(-\ln y_2)^{\beta_2}} \left[1 + \theta \left(1 - e^{-\alpha_1(-\ln y_1)^{\beta_1}} \right) \left(1 - e^{-\alpha_2(-\ln y_2)^{\beta_2}} \right) \right]} \\ & \times \left(1 + \theta \left(1 - 2e^{-\alpha_1(-\ln y_1)^{\beta_1}} \right) \left(1 - 2e^{-\alpha_2(-\ln y_2)^{\beta_2}} \right) \right) \end{aligned}$$

3.2. Estimación basada en cópulas

En esta sección se presentan tres métodos diferentes de estimación que se utilizan para estimar los parámetros de distribución Weibull Unitaria Bivariada, tales como: estimación de máxima verosimilitud, estimación por inferencia de funciones marginales y método semiparamétrico.

3.2.1. Estimación de máxima verosimilitud

Se halla el estimador de máxima verosimilitud (MLE, por sus siglas en inglés) para estimar los parámetros del modelo en forma conjunta, el cual es un método paramétrico de un solo paso. La función de log-verosimilitud es,

$$\ln L = \sum_{i=1}^n [\ln(f_1(y_{1i})f_2(y_{2i})c(F_1(y_{1i}, \delta_1), F_2(y_{2i}, \delta_2); \theta))]]$$

Las estimaciones de los parámetros se obtienen maximizando la función de log-verosimilitud con la expectativa de cada parámetro. Definimos,

$$\begin{aligned} a(y_j; \alpha_j, \beta_j) &= 1 - 2F(y_j; \alpha_j, \beta_j) \\ &= 1 - 2e^{-\alpha_j(-\ln y_j)^{\beta_j}}, \quad j = 1, 2. \end{aligned}$$

La función de verosimilitud de una distribución Weibull Unitaria Bivariada se define como,

$$\begin{aligned} L &= (\alpha_1 \alpha_2 \beta_1 \beta_2)^n \prod_{i=1}^n \left(\left(\frac{1}{y_{1i}} \right) \left(\frac{1}{y_{2i}} \right) (-\ln(y_{1i}))^{\beta_1-1} (-\ln(y_{2i}))^{\beta_2-1} \right) e^{-\sum_{i=1}^n \alpha_1 (-\ln y_1)^{\beta_1} - \sum_{i=1}^n \alpha_2 (-\ln y_2)^{\beta_2}} \\ &\times \prod_{i=1}^n (1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2))) \end{aligned}$$

y la función de log-verosimilitud se puede escribir,

$$\begin{aligned} \ln L &= n(\ln \alpha_1 + \ln \beta_1 + \ln \alpha_2 + \ln \beta_2) + \sum_{i=1}^n \ln \left(\frac{1}{y_{1i}} \right) + \sum_{i=1}^n \ln \left(\frac{1}{y_{2i}} \right) + (\beta_1 - 1) \sum_{i=1}^n \ln(-\ln(y_{1i})) \\ &+ (\beta_2 - 1) \sum_{i=1}^n \ln(-\ln(y_{2i})) - \sum_{i=1}^n \alpha_1 (-\ln y_1)^{\beta_1} - \sum_{i=1}^n \alpha_2 (-\ln y_2)^{\beta_2} \\ &+ \sum_{i=1}^n \ln(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2))) \end{aligned}$$

Las estimaciones se consiguen igualando las derivadas a cero, es decir, encontrando los score y encontrando las ecuaciones score,

$$\frac{\partial \ln L}{\partial \alpha_1} = \frac{n}{\alpha_1} - \sum_{i=1}^n (-\ln y_{1i})^{\beta_1} + \sum_{i=1}^n \frac{2\theta(a(y_{2i}; \alpha_2, \beta_2)) (-\ln y_{1i})^{\beta_1} e^{-\alpha_1 (-\ln y_{1i})^{\beta_1}}}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))}$$

$$\frac{\partial \ln L}{\partial \alpha_2} = \frac{n}{\alpha_2} - \sum_{i=1}^n (-\ln y_{2i})^{\beta_2} + \sum_{i=1}^n \frac{2\theta(a(y_{1i}; \alpha_1, \beta_1)) (-\ln y_{2i})^{\beta_2} e^{-\alpha_2(-\ln y_{2i})^{\beta_2}}}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))}$$

$$\begin{aligned} \frac{\partial \ln L}{\partial \beta_1} &= \frac{n}{\beta_1} + \sum_{i=1}^n \ln(-\ln y_{1i}) - \sum_{i=1}^n \alpha_1 (-\ln y_{1i})^{\beta_1} \ln(-\ln y_{1i}) \\ &\quad - \sum_{i=1}^n \frac{2\alpha_1 \theta(a(y_{2i}; \alpha_2, \beta_2)) \ln(-\ln y_{1i}) (-\ln y_{1i})^{\beta_1} e^{-\alpha_1(-\ln y_{1i})^{\beta_1}}}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))} \end{aligned}$$

$$\begin{aligned} \frac{\partial \ln L}{\partial \beta_2} &= \frac{n}{\beta_2} + \sum_{i=1}^n \ln(-\ln y_{2i}) - \sum_{i=1}^n \alpha_1 (-\ln y_{2i})^{\beta_2} \ln(-\ln y_{2i}) \\ &\quad - \sum_{i=1}^n \frac{2\alpha_2 \theta(a(y_{1i}; \alpha_1, \beta_1)) \ln(-\ln y_{2i}) (-\ln y_{2i})^{\beta_2} e^{-\alpha_2(-\ln y_{2i})^{\beta_2}}}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))} \end{aligned}$$

y

$$\frac{\partial \ln L}{\partial \theta} = \sum_{i=1}^n \frac{(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2))}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))}$$

El estimador de máxima de verosimilitud es $\hat{\delta} = (\hat{\alpha}_1, \hat{\beta}_1, \hat{\alpha}_2, \hat{\beta}_2, \hat{\theta})$, se puede obtener resolviendo simultáneamente las ecuaciones de log-verosimilitud (ecuaciones score),

$$\frac{\partial \ln L}{\partial \theta} = 0, \quad \frac{\partial \ln L}{\partial \beta_j} = 0, \quad \frac{\partial \ln L}{\partial \alpha_j} = 0, \quad j = 1, 2$$

Las estimaciones se consiguen realizando métodos numéricos iterativos tipo Newton Raphson o cuasi-Newton, por ejemplo.

3.2.2. Estimación por inferencia de funciones marginales

Es un método paramétrico con estimación en dos pasos, propuesto por Joe (2005), (IFM, por sus siglas en inglés). En el primer paso se estiman los parámetros de cada distribución marginal por separado.

$$\ln L_1 = \sum_{j=1}^n \ln f_1(y_{1j}, \delta_1); \quad \ln L_2 = \sum_{j=1}^n \ln f_2(y_{2j}, \delta_2)$$

Luego, en el segundo paso, el parámetro de la cópula se estima maximizando la función

de log-verosimilitud de la densidad de la cópula utilizando las estimaciones de máxima verosimilitud de las marginales $\hat{F}_1(y_{1j}, \delta)$ y $\hat{F}_2(y_{2j}, \delta)$. La función de log-verosimilitud de una distribución marginal Weibull Unitaria se define como,

$$\begin{aligned} \ln L_j &= \sum_{i=1}^n \ln \left(\frac{1}{y_{ji}} \alpha_j \beta_j (-\ln y_{ji})^{\beta_j-1} e^{-\alpha_j (-\ln y_{ji})^{\beta_j}} \right) \\ &= \sum_{i=1}^n \left(\ln \left(\frac{1}{y_{ji}} \right) + \ln(\alpha_j) + \ln(\beta_j) + (\beta_j - 1) \ln(-\ln y_{ji}) - \alpha_j (-\ln y_{ji})^{\beta_j} \right) \\ &= \sum_{i=1}^n \ln \left(\frac{1}{y_{ji}} \right) + \sum_{i=1}^n \ln(\alpha_j) + \sum_{i=1}^n \ln(\beta_j) + \sum_{i=1}^n (\beta_j - 1) \ln(-\ln y_{ji}) - \sum_{i=1}^n \alpha_j (-\ln y_{ji})^{\beta_j} \\ &= \sum_{i=1}^n \ln \left(\frac{1}{y_{ji}} \right) + n (\ln(\alpha_j) + \ln(\beta_j)) + (\beta_j - 1) \sum_{i=1}^n \ln(-\ln y_{ji}) - \alpha_j \sum_{i=1}^n (-\ln y_{ji})^{\beta_j} \end{aligned}$$

Los estimadores de máxima verosimilitud ($\hat{\alpha}_1, \hat{\beta}_1, \hat{\alpha}_2, \hat{\beta}_2$) se pueden obtener resolviendo simultáneamente las ecuaciones de verosimilitud,

$$\frac{\partial \ln L}{\partial \beta_j} = 0, \quad \frac{\partial \ln L}{\partial \alpha_j} = 0, \quad j = 1, 2$$

Esto es,

$$\frac{\partial \ln L_j}{\partial \beta_j} = \frac{n}{\beta_j} + \sum_{i=1}^n (-\ln y_{ji}) - \alpha_j \sum_{i=1}^n (-\ln y_{ji})^{\beta_j} \ln(-\ln y_{ji}) = 0$$

$$\frac{\partial \ln L_j}{\partial \alpha_j} = \frac{n}{\alpha_j} - \sum_{i=1}^n (-\ln y_{ji})^{\beta_j} = 0$$

entonces,

$$\hat{F}_j(y_j) = e^{-\alpha_j (-\ln y_j)^{\beta_j}}$$

y considerando el paso anterior, la estimación IFM de una distribución weibull unitaria bivariada se define como,

$$\ln L_{IFM} = \sum_{i=1}^n \ln (1 + \theta (1 - 2\hat{F}_1(y_{1i})) (1 - 2\hat{F}_2(y_{2i})))$$

Basándose en esto, la diferenciación de la función logarítmica de verosimilitud con respecto a θ , así,

$$\frac{\partial \ln L_{IFM}}{\partial \theta} = \sum_{i=1}^n \frac{(a(y_{1i}, \hat{\alpha}_1, \hat{\beta}_1))(a(y_{2i}, \hat{\alpha}_2, \hat{\beta}_2))}{(1 + \theta(a(y_{1i}, \hat{\alpha}_1, \hat{\beta}_1))(a(y_{2i}, \hat{\alpha}_2, \hat{\beta}_2)))}$$

La estimación del parámetro θ se calcula numéricamente,

$$\frac{\partial \ln L_{IFM}}{\partial \theta} = 0$$

No existe una expresión de forma cerrada para el estimador de máxima verosimilitud $\hat{\theta}$ y su cálculo debe realizarse numéricamente utilizando un algoritmo de optimización no lineal como optimize o uniroot.

3.2.3. Estimación por método semi-paramétrico (SM)

Este método se realiza en dos etapas como en IFM, pero la diferencia es que las distribuciones marginales se estiman de forma no paramétrica por sus distribuciones empíricas muestrales. En este método, las observaciones se transforman en pseudo-observaciones utilizando la función de distribución empírica de cada distribución marginal. La función de distribución empírica se define como,

$$\tilde{F}_i(y_i) = \frac{\sum_{j=1}^n I(Y_{i,j} \leq y_i)}{n+1}; \quad i = 1, 2$$

Entonces, θ es estimado maximizando la pseudo verosimilitud,

$$\sum_{i=1}^n \ln c(\tilde{F}_1(Y_{1i}), \tilde{F}_2(Y_{2i}); \theta)$$

la función de log-verosimilitud de una distribución weibull unitaria bivariada se da como sigue,

$$\ln L_{SP} = \sum_{i=1}^n \ln(1 + \theta(1 - 2\tilde{F}_1(y_{1i}))(1 - 2\tilde{F}_2(y_{2i}))) \quad (3.5)$$

No existe una expresión de forma cerrada para el estimador de máxima verosimilitud θ mediante el uso de ecuación (3.5), por lo tanto su cálculo debe realizarse numéricamente utilizando un software estadístico como R Core Team (2021).

3.3. Matriz de información

En esta sección se presentan los elementos de las matrices de información observada y esperada para el modelo Weibull Unitario Bivariado. Por definición, los elementos de la matriz de información observada son hallados usando la expresión,

$$j_{\Theta_p, \Theta_q} = -\frac{\partial^2 \ln L(\Theta)}{\partial \Theta_p \partial \Theta_q}$$

donde $L(\Theta)$ es la función de log-verosimilitud

Estos elementos será denotados como j_{α_1, α_1} , j_{α_1, α_2} , j_{α_1, β_1} , j_{α_1, β_2} , $j_{\alpha_1, \theta}$, j_{α_2, α_2} , j_{α_2, β_1} , j_{α_2, β_2} , $j_{\alpha_2, \theta}$, j_{β_1, β_1} , j_{β_1, β_2} , $j_{\beta_1, \theta}$, j_{β_2, β_2} , $j_{\beta_2, \theta}$, y $j_{\theta, \theta}$ y son presentados a continuación:

$$j_{\alpha_1, \alpha_1} = \frac{n}{\alpha_1^2} + \sum_{i=1}^n \frac{2\theta(a(y_2; \alpha_2, \beta_2)) (-\ln y_{1i})^{2\beta_1} e^{-\alpha_1(-\ln y_{1i})^{\beta_1}} ((1 + \theta(a(y_1; \alpha_1, \beta_1)))(a(y_2; \alpha_2, \beta_2)))}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1)))(a(y_{2i}; \alpha_2, \beta_2))^2} \\ + \sum_{i=1}^n \frac{2\theta(a(y_2; \alpha_2, \beta_2)) (-\ln y_{1i})^{2\beta_1} e^{-\alpha_1(-\ln y_{1i})^{\beta_1}} (2\theta(a(y_2; \alpha_2, \beta_2)) e^{-\alpha_1(-\ln y_{1i})^{\beta_1}})}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1)))(a(y_{2i}; \alpha_2, \beta_2))^2}$$

$$j_{\alpha_1, \alpha_2} = \sum_{i=1}^n \frac{-4\theta(-\ln y_{1i})^{\beta_1} e^{-\alpha_1(-\ln y_{1i})^{\beta_1}} (-\ln y_{2i})^{\beta_2} e^{-\alpha_2(-\ln y_{2i})^{\beta_2}}}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1)))(a(y_{2i}; \alpha_2, \beta_2))^2}$$

$$j_{\alpha_1, \beta_1} = \frac{-2\theta(a(y_2; \alpha_2, \beta_2)) (-\ln y_{1i})^{\beta_1} \ln(-\ln y_{1i}) e^{-\alpha_1(-\ln y_{1i})^{\beta_1}}}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1)))(a(y_{2i}; \alpha_2, \beta_2))} \\ \times \frac{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1)))(a(y_{2i}; \alpha_2, \beta_2)) (1 - \alpha_1 (-\ln y_{1i})^{\beta_1})}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1)))(a(y_{2i}; \alpha_2, \beta_2))} \\ + \frac{2\theta(a(y_2; \alpha_2, \beta_2)) (-\ln y_{1i})^{\beta_1} \ln(-\ln y_{1i}) e^{-\alpha_1(-\ln y_{1i})^{\beta_1}}}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1)))(a(y_{2i}; \alpha_2, \beta_2))} \\ \times \frac{(2\theta\alpha_1(a(y_2; \alpha_2, \beta_2)) (-\ln y_{1i})^{\beta_1} e^{-\alpha_1(-\ln y_{1i})^{\beta_1}})}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1)))(a(y_{2i}; \alpha_2, \beta_2))} \\ + \sum_{i=1}^n (-\ln y_{1i})^{\beta_1} \ln(-\ln y_{1i})$$

$$j_{\alpha_1, \beta_2} = \sum_{i=1}^n \frac{-4\theta\alpha_2(-\ln y_{1i})^{\beta_1} e^{-\alpha_1(-\ln y_{1i})^{\beta_1}} \ln(-\ln y_{2i}) (-\ln y_{2i})^{\beta_2} e^{-\alpha_2(-\ln y_{2i})^{\beta_2}}}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1)))(a(y_{2i}; \alpha_2, \beta_2))^2}$$

$$j_{\alpha_1, \theta} = \sum_{i=1}^n \frac{-2(a(y_{2i}; \alpha_2, \beta_2)) (-\ln y_{1i})^{\beta_1} e^{-\alpha_1 (-\ln y_{1i})^{\beta_1}}}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))^2}$$

$$j_{\alpha_2, \alpha_2} = \frac{n}{\alpha_2^2} + \sum_{i=1}^n \frac{2\theta(a(y_{1i}; \alpha_1, \beta_1)) (-\ln y_{2i})^{2\beta_2} e^{-\alpha_2 (-\ln y_{2i})^{\beta_2}} ((1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))^2} \\ + \sum_{i=1}^n \frac{2\theta(a(y_{1i}; \alpha_1, \beta_1)) (-\ln y_{2i})^{2\beta_2} e^{-\alpha_2 (-\ln y_{2i})^{\beta_2}} (2\theta(a(y_{2i}; \alpha_2, \beta_2)) e^{-\alpha_1 (-\ln y_{1i})^{\beta_1}})}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))^2}$$

$$j_{\alpha_2, \beta_1} = \sum_{i=1}^n \frac{-4\theta\alpha_1 (-\ln y_{2i})^{\beta_2} e^{-\alpha_2 (-\ln y_{2i})^{\beta_2}} \ln(-\ln y_{1i}) (-\ln y_{1i})^{\beta_1} e^{-\alpha_1 (-\ln y_{1i})^{\beta_1}}}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))^2}$$

$$j_{\alpha_2, \beta_2} = \sum_{i=1}^n \frac{-2\theta(a(y_{1i}; \alpha_1, \beta_1)) (-\ln y_{2i})^{\beta_2} \ln(-\ln y_{2i}) e^{-\alpha_2 (-\ln y_{2i})^{\beta_2}}}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))} \\ \times \frac{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2))) (1 - \alpha_2 (-\ln y_{2i})^{\beta_2})}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))} \\ + \sum_{i=1}^n \frac{2\theta(a(y_{1i}; \alpha_1, \beta_1)) (-\ln y_{2i})^{\beta_2} \ln(-\ln y_{2i}) e^{-\alpha_2 (-\ln y_{2i})^{\beta_2}}}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))} \\ \times \frac{(2\theta\alpha_2(a(y_{1i}; \alpha_1, \beta_1)) (-\ln y_{2i})^{\beta_2} e^{-\alpha_2 (-\ln y_{2i})^{\beta_2}})}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))} \\ + \sum_{i=1}^n (-\ln y_{2i})^{\beta_2} \ln(-\ln y_{2i})$$

$$j_{\alpha_2, \theta} = \sum_{i=1}^n \frac{-2(a(y_{1i}; \alpha_1, \beta_1)) (-\ln y_{2i})^{\beta_2} e^{-\alpha_2 (-\ln y_{2i})^{\beta_2}}}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))^2}$$

$$\begin{aligned}
j_{\beta_1, \beta_1} &= \sum_{i=1}^n \frac{-2\alpha_1 \theta(a(y_{2i}; \alpha_2, \beta_2)) \ln(-\ln y_{1i})^2 (-\ln y_{1i})^{\beta_1} e^{-\alpha_1 (-\ln y_{1i})^{\beta_1}}}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))} \\
&\quad \times \frac{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2))) (1 - \alpha_1 (-\ln y_{1i})^{\beta_1})}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))} \\
&\quad + \sum_{i=1}^n \frac{4\alpha_1^2 \theta^2(a(y_{2i}; \alpha_2, \beta_2)^2) \ln(-\ln y_{1i})^2 (-\ln y_{1i})^{2\beta_1} e^{-2\alpha_1 (-\ln y_{1i})^{\beta_1}}}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))^2} \\
&\quad + \sum_{i=1}^n \alpha_1 \ln(-\ln y_{1i})^2 (-\ln y_{1i})^{\beta_1} + \frac{n}{\beta_1^2}
\end{aligned}$$

$$j_{\beta_1, \beta_2} = \sum_{i=1}^n \frac{-2\alpha_1 \theta \ln(-\ln y_{1i}) (-\ln y_{1i})^{\beta_1} e^{-\alpha_1 (-\ln y_{1i})^{\beta_1}}}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))^2}$$

$$j_{\beta_1, \theta} = \sum_{i=1}^n \frac{-2\alpha_1 (a(y_{2i}; \alpha_2, \beta_2)) \ln(-\ln y_{1i}) (-\ln y_{1i})^{\beta_1} e^{-\alpha_1 (-\ln y_{1i})^{\beta_1}}}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))^2}$$

$$\begin{aligned}
j_{\beta_2, \beta_2} &= \sum_{i=1}^n \frac{-2\alpha_2 \theta(a(y_{1i}; \alpha_1, \beta_1)) \ln(-\ln y_{2i})^2 (-\ln y_{2i})^{\beta_2} e^{-\alpha_2 (-\ln y_{2i})^{\beta_2}}}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))} \\
&\quad \times \frac{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2))) (1 - \alpha_2 (-\ln y_{2i})^{\beta_2})}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))} \\
&\quad + \sum_{i=1}^n \frac{4\alpha_2^2 \theta^2(a(y_{1i}; \alpha_1, \beta_1)^2) \ln(-\ln y_{2i})^2 (-\ln y_{2i})^{2\beta_2} e^{-2\alpha_2 (-\ln y_{2i})^{\beta_2}}}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))^2} \\
&\quad + \sum_{i=1}^n \alpha_2 \ln(-\ln y_{2i})^2 (-\ln y_{2i})^{\beta_2} + \frac{n}{\beta_2^2}
\end{aligned}$$

$$j_{\beta_2, \theta} = \sum_{i=1}^n \frac{-2\alpha_2 (a(y_{1i}; \alpha_1, \beta_1)) \ln(-\ln y_{2i}) (-\ln y_{2i})^{\beta_2} e^{-\alpha_2 (-\ln y_{2i})^{\beta_2}}}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))^2}$$

$$j_{\theta, \theta} = \sum_{i=1}^n \frac{(a(y_{1i}; \alpha_1, \beta_1)^2)(a(y_{2i}; \alpha_2, \beta_2)^2)}{(1 + \theta(a(y_{1i}; \alpha_1, \beta_1))(a(y_{2i}; \alpha_2, \beta_2)))^2}$$

Los elementos de la matriz de información esperada se determinana partir de la expresión

$$i_{\Theta_p, \Theta_q} = \frac{1}{n} E \left(- \frac{\partial^2 \ln L(\Theta)}{\partial \Theta_p \partial \Theta_q} \right)$$

Dichos elementos son denotados como $i_{\alpha_1, \alpha_1}, i_{\alpha_1, \alpha_2}, i_{\alpha_1, \beta_1}, i_{\alpha_1, \beta_2}, i_{\alpha_1, \theta}, i_{\alpha_2, \alpha_2}, i_{\alpha_2, \beta_1}, i_{\alpha_2, \beta_2}, i_{\alpha_2, \theta}, i_{\beta_1, \beta_1}, i_{\beta_1, \beta_2}, i_{\beta_1, \theta}, i_{\beta_2, \beta_2}, i_{\beta_2, \theta}$, y $i_{\theta, \theta}$

Para el método IFM, la matriz de varianzas asintótica no tienen forma cerrada. Joe (2005) define las funciones de inferencia como,

$$\sum_{i=1}^n g(\mathbf{Y}; \eta)$$

donde,

$\mathbf{g}^T = (g_1^T, \dots, g_m^T, g_d^T)$, $g_j = \frac{\partial l_j}{\partial \alpha_j}$ para $j = 1, \dots, m$; y $d = m + 1$, corresponde a la ecuación,

$$g_d = \frac{\partial l}{\partial \delta}$$

así la matriz de varianzas asintótica viene dada por,

$$V = (-D_g^{-1}) M_g (-D_g^{-1}), \quad (3.6)$$

donde, $M_g = \text{Cov}(g(\mathbf{Y}; \eta)) = E(gg^T)$ y $D_g = E\left(\frac{\partial g(\mathbf{Y}; \eta)}{\partial \eta^T}\right)$

3.4. Intervalos de confianza asintóticos

En esta sección, proponemos los intervalos de confianza asintóticos utilizando métodos de estimación. Teniendo esto en cuenta, se proponen los intervalos de confianza asintóticos utilizando los métodos MLE, IFM y SM se pueden utilizar para construir los intervalos de confianza para los parámetros. Primero obtenemos $\hat{\sigma}(\hat{\alpha}_1, \hat{\beta}_1, \hat{\alpha}_2, \hat{\beta}_2, \hat{\theta})$ que es la matriz de covarianza estimada de los parámetros y se define como,

$$\hat{\sigma}(\hat{\alpha}_1, \hat{\beta}_1, \hat{\alpha}_2, \hat{\beta}_2, \hat{\theta}) = \begin{bmatrix} \hat{\sigma}_{\hat{\alpha}_1 \hat{\alpha}_1} & \hat{\sigma}_{\hat{\alpha}_1 \hat{\beta}_1} & \hat{\sigma}_{\hat{\alpha}_1 \hat{\alpha}_2} & \hat{\sigma}_{\hat{\alpha}_1 \hat{\beta}_2} & \hat{\sigma}_{\hat{\alpha}_1 \hat{\theta}} \\ \hat{\sigma}_{\hat{\beta}_1 \hat{\alpha}_1} & \hat{\sigma}_{\hat{\beta}_1 \hat{\beta}_1} & \hat{\sigma}_{\hat{\beta}_1 \hat{\alpha}_2} & \hat{\sigma}_{\hat{\beta}_1 \hat{\beta}_2} & \hat{\sigma}_{\hat{\beta}_1 \hat{\theta}} \\ \hat{\sigma}_{\hat{\alpha}_2 \hat{\alpha}_1} & \hat{\sigma}_{\hat{\alpha}_2 \hat{\beta}_1} & \hat{\sigma}_{\hat{\alpha}_2 \hat{\alpha}_2} & \hat{\sigma}_{\hat{\alpha}_2 \hat{\beta}_2} & \hat{\sigma}_{\hat{\alpha}_2 \hat{\theta}} \\ \hat{\sigma}_{\hat{\beta}_2 \hat{\alpha}_1} & \hat{\sigma}_{\hat{\beta}_2 \hat{\beta}_1} & \hat{\sigma}_{\hat{\beta}_2 \hat{\alpha}_2} & \hat{\sigma}_{\hat{\beta}_2 \hat{\beta}_2} & \hat{\sigma}_{\hat{\beta}_2 \hat{\theta}} \\ \hat{\sigma}_{\hat{\theta} \hat{\alpha}_1} & \hat{\sigma}_{\hat{\theta} \hat{\beta}_1} & \hat{\sigma}_{\hat{\theta} \hat{\alpha}_2} & \hat{\sigma}_{\hat{\theta} \hat{\beta}_2} & \hat{\sigma}_{\hat{\theta} \hat{\theta}} \end{bmatrix}$$

Un intervalo de confianza del 95% de dos lados aproximado para $(\alpha_1, \beta_1, \alpha_2, \beta_2, \theta)$ respectivamente,

$$\hat{\alpha}_i \pm Z_{0.025} \sqrt{\hat{\sigma}_{\hat{\alpha}_i \hat{\alpha}_i}}, \quad \hat{\beta}_i \pm Z_{0.025} \sqrt{\hat{\sigma}_{\hat{\beta}_i \hat{\beta}_i}}, \quad i = 1, 2, \quad \text{y} \quad \hat{\theta} \pm Z_{0.025} \sqrt{\hat{\sigma}_{\hat{\theta} \hat{\theta}}}$$

Capítulo 4

Estudio de simulación

En esta sección; presentaremos la simulación de Monte Carlo realizada para la comparación entre métodos de estimación basados en cópula como: máxima verosimilitud y estimación por inferencia de funciones marginales para estimar los parámetros de distribución Weibull Unitaria Bivariada por medio del paquete estadístico R Core Team (2021).

El algoritmo de simulación se llevó a cabo generando los datos de la distribución de Weibull Unitaria Bivariada, considerando los siguientes valores de los parámetros α_1 , β_1 , α_2 , β_2 y θ ,

$$\text{Caso 1} = \begin{cases} \alpha_1 = 1.5, \beta_1 = 3.5, \alpha_2 = 0.5, \beta_2 = 2.0, \theta = 0.25 \\ \alpha_1 = 1.5, \beta_1 = 3.5, \alpha_2 = 0.5, \beta_2 = 2.0, \theta = 0.75 \\ \alpha_1 = 1.5, \beta_1 = 3.5, \alpha_2 = 0.5, \beta_2 = 2.0, \theta = -0.25 \\ \alpha_1 = 1.5, \beta_1 = 3.5, \alpha_2 = 0.5, \beta_2 = 2.0, \theta = -0.75 \end{cases}$$
$$\text{Caso 2} = \begin{cases} \alpha_1 = 2.5, \beta_1 = 1.5, \alpha_2 = 1.5, \beta_2 = 1, \theta = 0.25 \\ \alpha_1 = 2.5, \beta_1 = 1.5, \alpha_2 = 1.5, \beta_2 = 1, \theta = 0.75 \end{cases}$$

Para diferentes tamaños de muestra $n = 30, 50, 75, 100, 150, 200$ y 500 . Todos los cálculos se obtienen en base al lenguaje R. Los métodos de simulación se comparan utilizando los criterios de estimación de parámetros, la comparación se realiza calculando el sesgo (Bias), el error cuadrático medio (MSE) y la longitud del intervalo de confianza (L.CI) para cada método de la siguiente manera

$$Bias(\hat{\delta}) = \frac{1}{1500} \sum_{i=1}^{1500} (\hat{\delta}_i - \delta)$$

donde $\hat{\delta}_i$ es la i -ésima estimación de δ

$$MSE(\hat{\delta}) = \frac{1}{1500} \sum_{i=1}^{1500} (\hat{\delta}_i - \delta)^2$$

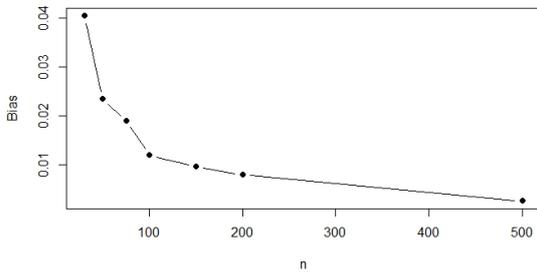
y

$$L.CI = \frac{1}{1500} \sum_{i=1}^{1500} (\hat{\delta} + z_{(0.975)} \hat{\sigma}(\hat{\delta}) - [\hat{\delta} - z_{(0.975)} \hat{\sigma}(\hat{\delta})])$$

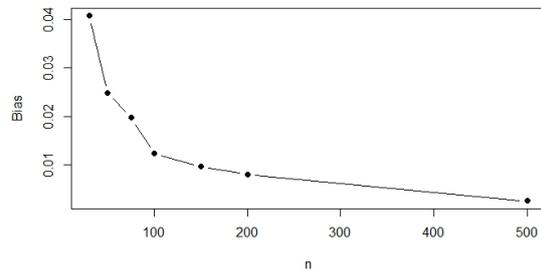
Esta simulación se lleva a cabo con 1500 muestras en cada caso.

Sobre los resultados resumidos en las tablas (A.1,A.2,A.3,A.4,A.5,A.6), se pueden extraer algunas conclusiones que se expresan de la siguiente manera: se observa que a medida que aumenta el tamaño de la muestra y el valor de parámetros permanece fijo, Bias, MSE, L.CI disminuye en los dos métodos considerados, mientras que PCI varía en cada tamaño de n. En un tamaño de muestra grande, el método de máxima verosimilitud y el método de estimación por inferencia de funciones marginales son casi equivalentes, donde la diferencia es menor y no hay diferencias significativas en los valores de Bias, MSE y L.CI. Ambos métodos tienen mejores estimaciones cuando el parámetro de cópula no es alto, esto es, que se encuentre en el intervalo $(-0.6, 0.6)$. El método IFM es mejor a la hora de estimar el parámetro de cópula θ , esto se debe a que el método IFM trata de dos pasos de estimación, primero, se estiman los parámetros de distribución marginal y segundo se estima el parámetro de cópula, teniendo en cuenta las estimaciones de parámetros previos de distribución marginal. Eso consigue más eficiencia.

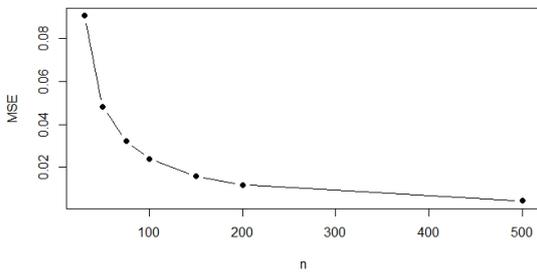
A manera de ejemplo se hacen las gráficas del tamaño de muestra contra las estadísticas calculadas para el caso 1.1, para así ver sus comportamientos a medida que el n se incrementa. Los resultados de todas las simulaciones se presentan en detalle en el Apéndice A



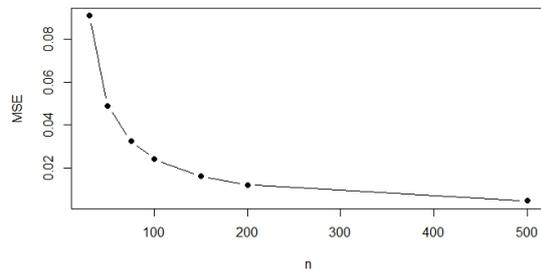
(a) Bias(MLE)



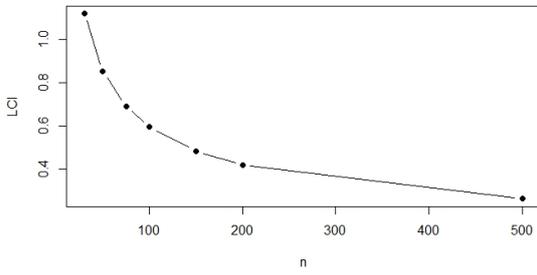
(b) Bias(IFM)



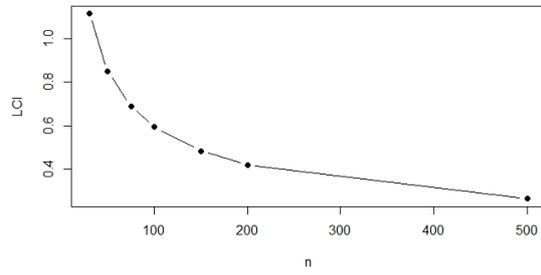
(c) MSE(MLE)



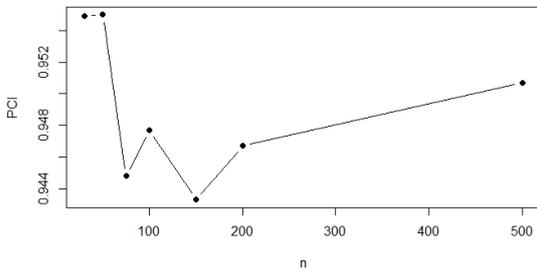
(d) MSE(IFM)



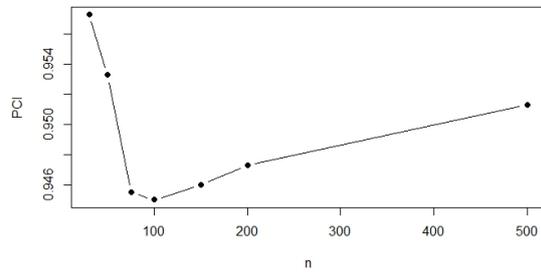
(e) L.CI(MLE)



(f) L.CI(IFM)

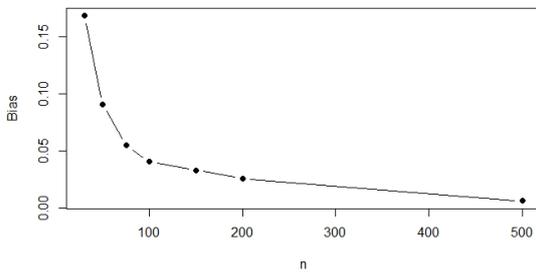


(g) PCI(MLE)

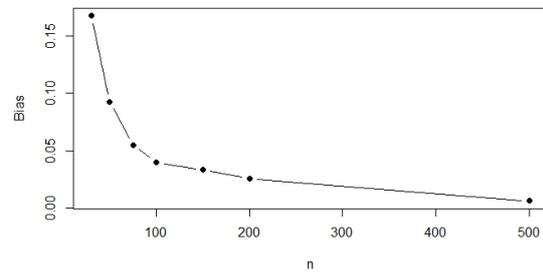


(h) PCI(IFM)

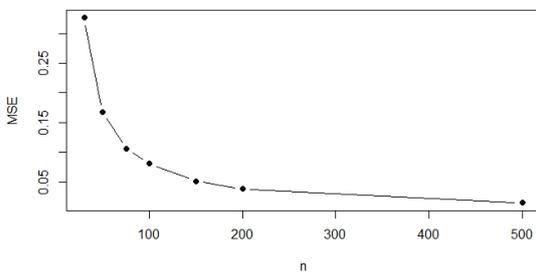
Figura 4.1: Comportamiento de las estadísticas de la estimación de α_1 según el tamaño de muestra.



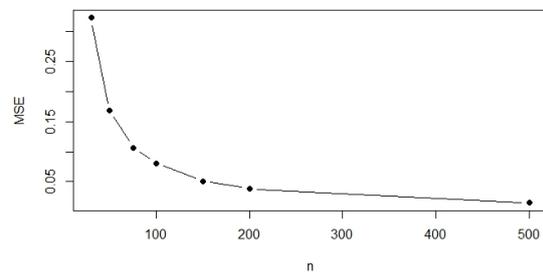
(a) Bias(MLE)



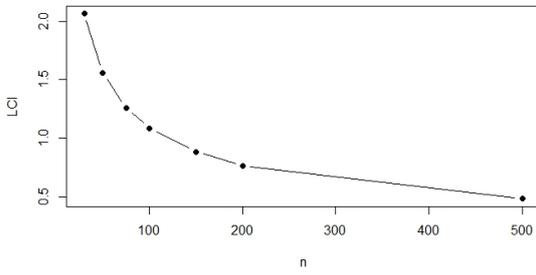
(b) Bias(IFM)



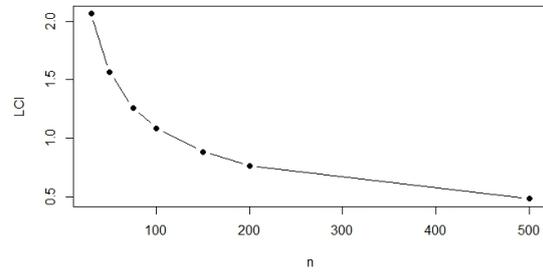
(c) MSE(MLE)



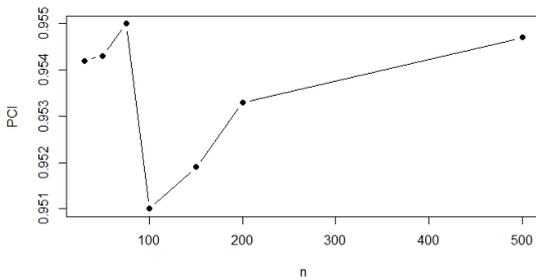
(d) MSE(IFM)



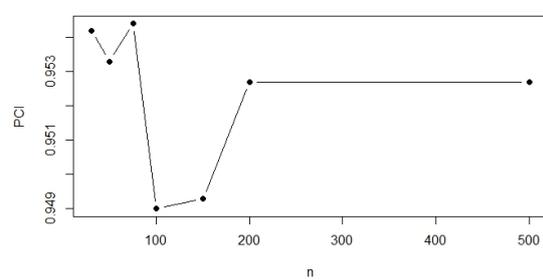
(e) L.CI(MLE)



(f) L.CI(IFM)

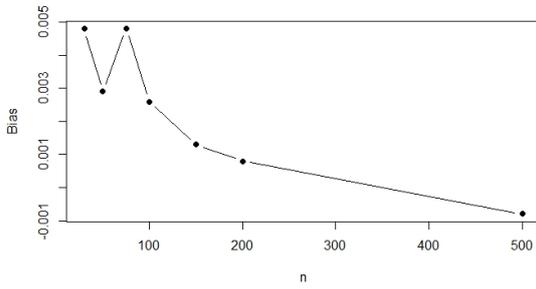


(g) PCI(MLE)

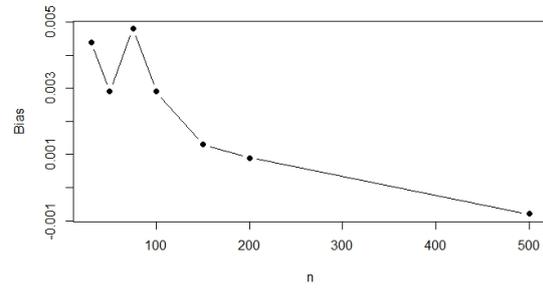


(h) PCI(IFM)

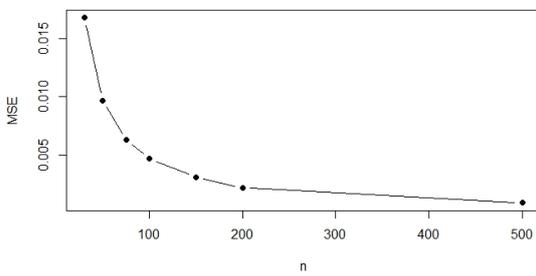
Figura 4.2: Comportamiento de las estadísticas de la estimación de β_1 según el tamaño de muestra.



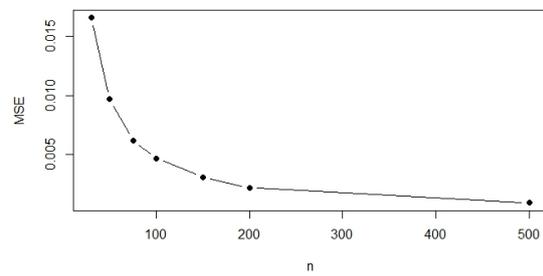
(a) Bias(MLE)



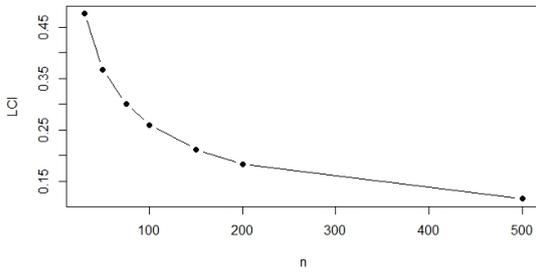
(b) Bias(IFM)



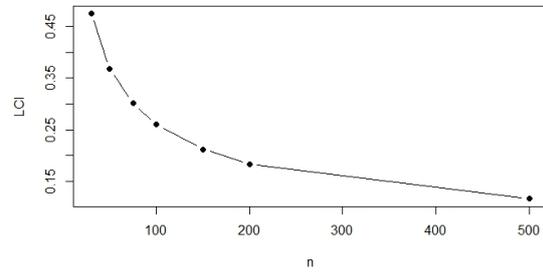
(c) MSE(MLE)



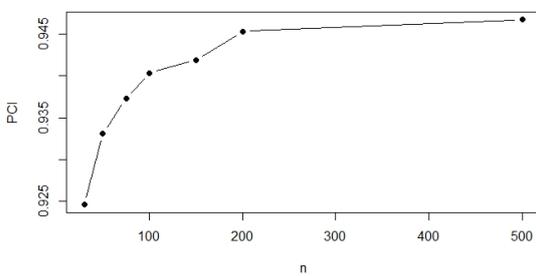
(d) MSE(IFM)



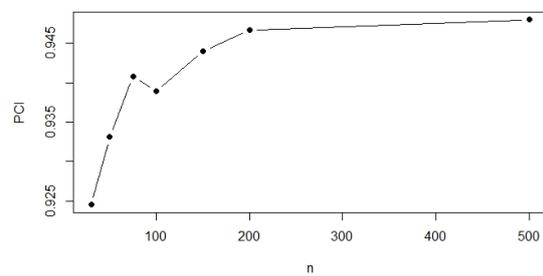
(e) L.CI(MLE)



(f) L.CI(IFM)

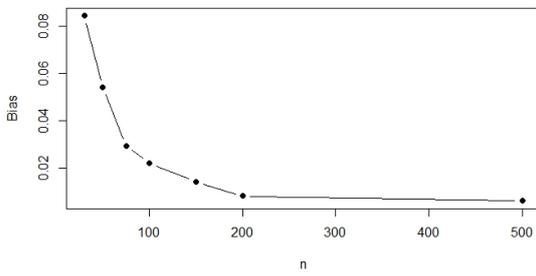


(g) PCI(MLE)

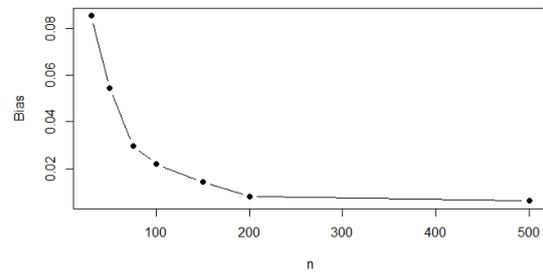


(h) PCI(IFM)

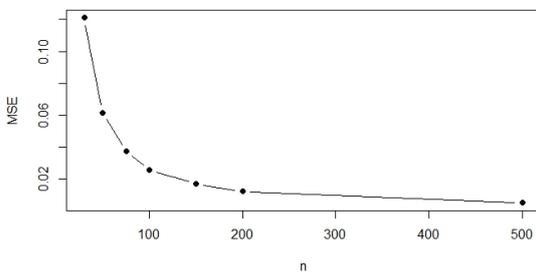
Figura 4.3: Comportamiento de las estadísticas de la estimación de α_2 según el tamaño de muestra.



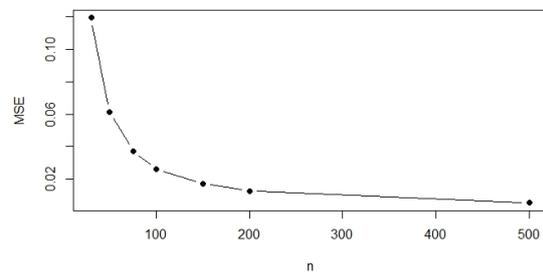
(a) Bias(MLE)



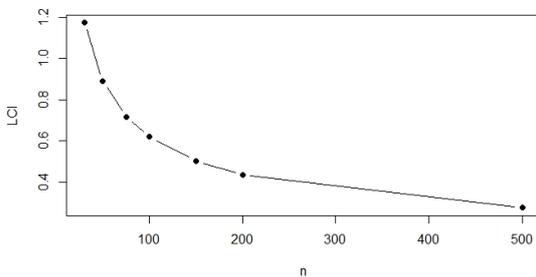
(b) Bias(IFM)



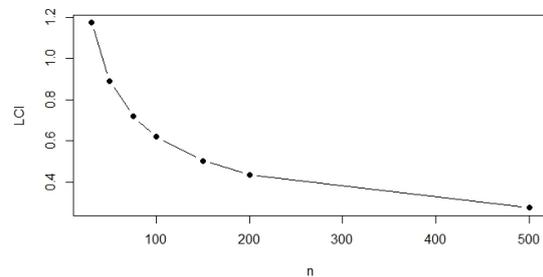
(c) MSE(MLE)



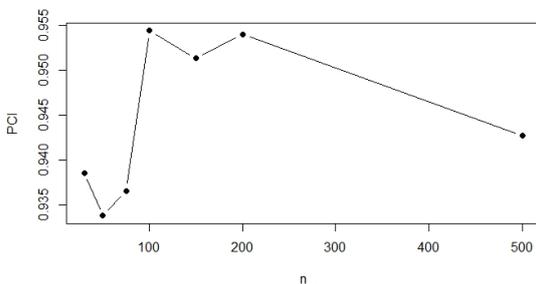
(d) MSE(IFM)



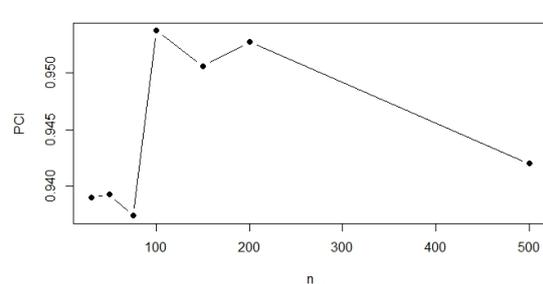
(e) L.CI(MLE)



(f) L.CI(IFM)

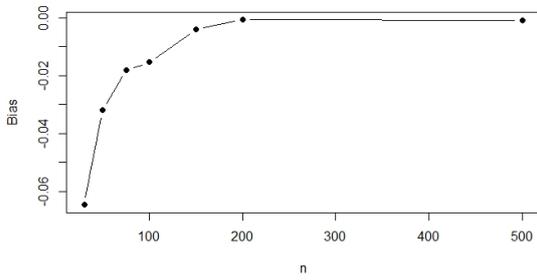


(g) PCI(MLE)

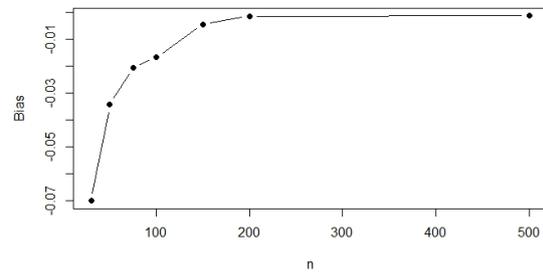


(h) PCI(IFM)

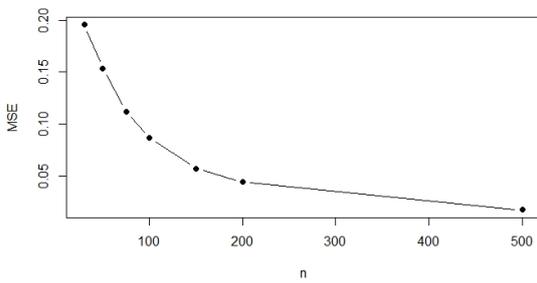
Figura 4.4: Comportamiento de las estadísticas de la estimación de β_2 según el tamaño de muestra.



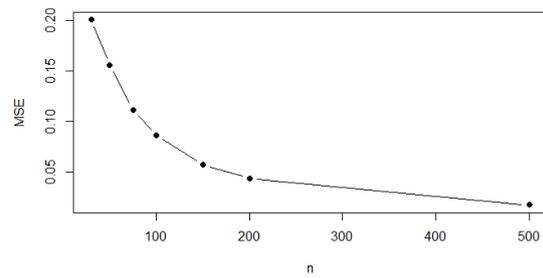
(a) Bias(MLE)



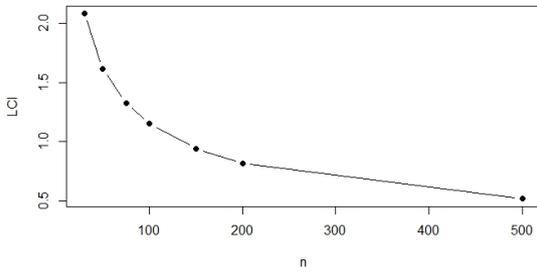
(b) Bias(IFM)



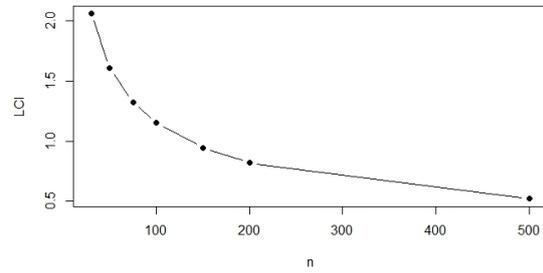
(c) MSE(MLE)



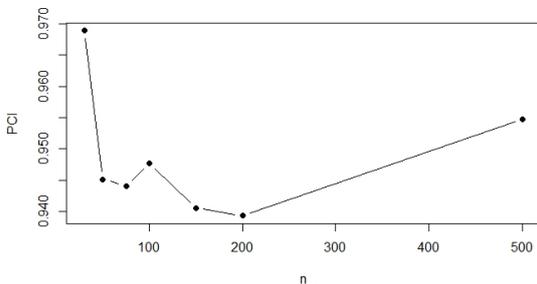
(d) MSE(IFM)



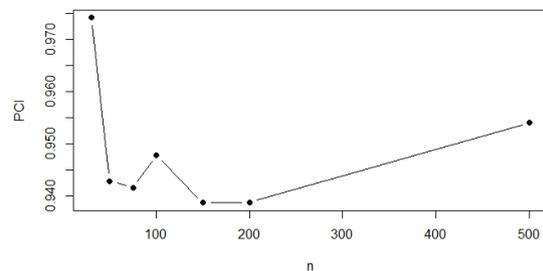
(e) L.CI(MLE)



(f) L.CI(IFM)



(g) PCI(MLE)



(h) PCI(IFM)

Figura 4.5: Comportamiento de las estadísticas de la estimación de θ según el tamaño de muestra.

Capítulo 5

Aplicación con datos reales

En este capítulo se presenta una aplicación del modelo propuesto a un conjunto de datos reales con fines ilustrativos. Los datos reales corresponden a mediciones realizadas en 117 municipios del Departamento de Antioquia, Colombia. El objetivo es modelar de forma conjunta la tasa neta de escolaridad (TNE), y la proporción de personas pobres y en mísera en los municipios de Antioquia en el año 2018. La TNE es calculada como el cociente entre el número de personas matriculadas en los niveles preescolar, primaria, secundaria y educación media y el total de personas en el municipio. En lo que respecta a la proporción de personas pobres y en mísera, hacemos referencia a la proporción de personas con necesidades básicas insatisfechas (NBI). Los datos fueron tomados del anuario estadístico de Antioquia y se encuentran disponibles para descarga en el sitio web: <https://www.antioquiadatos.gov.co/index.php/anuario-estadistico-home>.

En esta aplicación, se ajustó el modelo Weibull Unitario Bivariado (WUB) introducido en el Capítulo 3, y para efectos de comparación con otras distribuciones, se ajustaron los modelos bivariados Johnson S_B , (BJSB) (Lemonte and Moreno-Arenas, 2019) y Birnbaum-Saunders, (BSB) (Kundu et al., 2010). Todos los cálculos se realizaron con el paquete R Core Team (2021) versión 4.1.1 disponible de forma gratuita en <https://cran.r-project.org>. La estimación de los parámetros se realizó usando el algoritmo de optimización de Newton-Raphson (NR) a través de la función "maxLike" de R y el método de máxima verosimilitud (MV). Todos los gráficos se construyeron utilizando el programa R.

Las medidas descriptivas de las variables anteriores se presentan en la Tabla 5.1

Tabla 5.1: Medidas descriptivas.

Medidas	Y_1 (TNE)	Y_2 (NBI)
Mínimo	0.3420	0.0206
Q_1	0.5792	0.0980
Mediana	0.6702	0.1371
Media	0.6665	0.1839
Q_3	0.7552	0.2231
Máximo	0.9900	0.8169

Para ajustar el modelo WUB, se tomaron como valores iniciales para α_i y β_i , las desviaciones estándar de Y_i , para $i = 1, 2$. Para el parámetro θ se tomó el valor del coeficiente de correlación muestral entre Y_1 y Y_2 como valor inicial dentro del proceso iterativo.

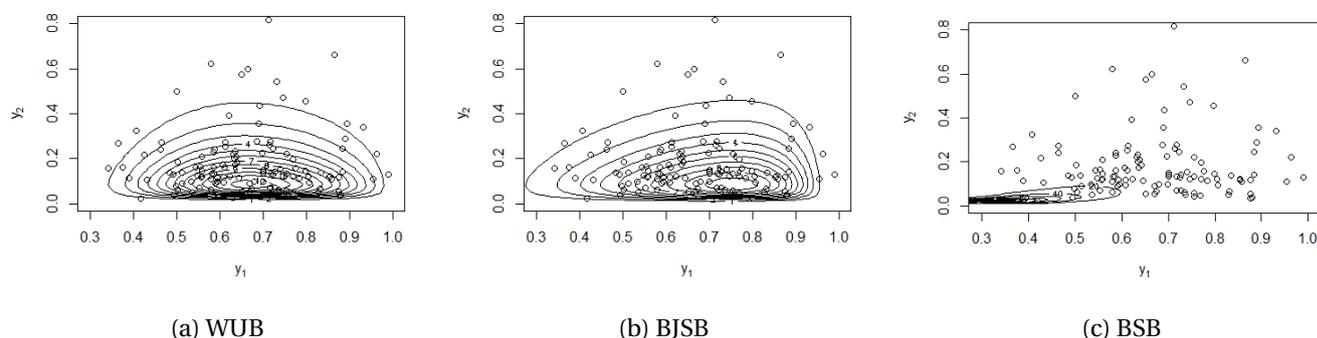


Figura 5.1: Gráficas de contorno para los modelos WUB, BJSB y BSB.

La Figura 5.1 muestra los gráficos de contorno para los modelos WUB, BJSB y BSB ajustados. En particular, el gráfico de contornos del modelo BSB muestra que el ajuste de este modelo a los datos no es muy adecuado.

La Tabla 5.2 presenta las estimaciones de máxima verosimilitud (MV) junto con sus correspondientes errores estándar asintóticos (EE) de los modelos WUB, BJSB y (BSB). Para comparar estos modelos se proporcionan los criterios AIC (Criterio de información de Akaike), BIC (Criterio de información bayesiano) y CAIC (AIC Consistente). También se proporcionan los IC asintóticos del 95% de los parámetros del modelo WUB.

Tabla 5.2: Estimaciones de los parámetros de los modelos WUB, BJSB y BSB.

Parámetro	WUB	BJSB	BSB	IC 95 %
α_1	4.2145(0.5027)		0.2239(0.0146)	(3.2292 , 5.1998)
β_1	1.9791(0.1443)		0.6231(0.0934)	(1.6963, 2.2619)
α_2	0.0950(0.0211)		0.7373(0.0497)	(0.0536, 0.1364)
β_2	3.0501(0.2230)		0.1252(0.0584)	(2.6130 , 3.4872)
θ	-0.0655(0.2856)			(-0.6253 , 0.4943)
ρ			0.0000(0.0925)	
ξ_1		0.6906(0.0160)		
ξ_2		0.1530(0.0108)		
δ_1		2.6756(1.4596)		
δ_2		3.0437(1.1375)		
σ_{11}		4.7009(5.1157)		
σ_{22}		7.5117(5.5310)		
σ_{12}		0.0305(0.5433)		
Log-verosimilitud	166.430	152.394	163.364	
AIC	-322.859	-290.788	-316.728	
BIC	-309.048	-271.453	-302.917	
CAIC	-322.131	-289.492	-316.001	

Nótese que los valores obtenidos para los criterios AIC, BIC y CAIC son menores para el modelo WUB que los obtenidos con los modelos BJSB y BSB, por lo tanto, el modelo WUB ajusta mejor al conjunto de datos que dichos modelos.

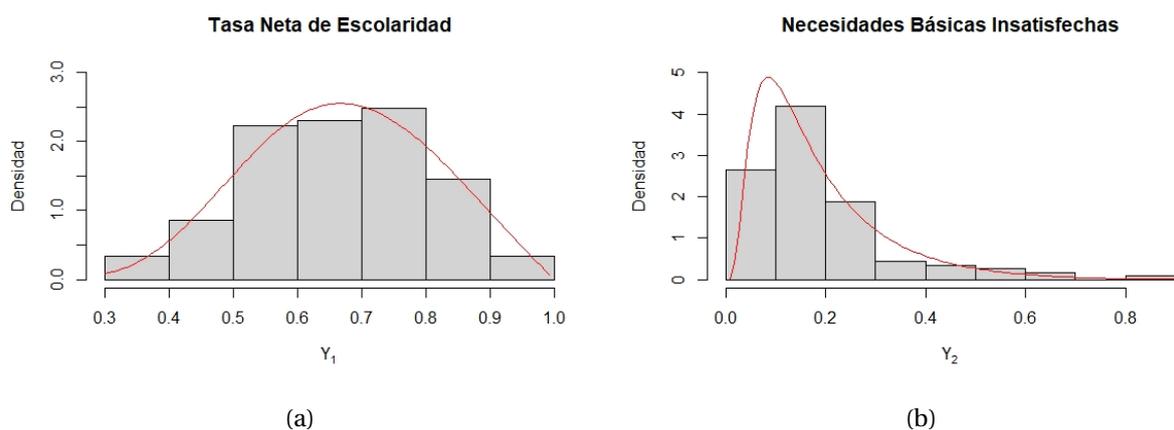


Figura 5.2: Gráficas de densidad de las variables tasa neta de escolaridad y proporción de necesidades básicas insatisfechas.

La Figura 5.2 muestra los histogramas para las variables TNE (a) y NBI (b) junto con

las densidades marginales Weibull unitarias, donde se puede apreciar el buen ajuste de las distribuciones marginales WU.

La prueba de bondad de ajuste multivariada de Kolmogorov-Smirnov propuesta por Justel et al. (1997), en especial para el caso de una distribución bivariada, que denotamos por BKS (bivariada Kolmogorov-Smirnov), el estadístico está dado por,

$$d_n = \sup_{(y_1, y_2) \in \mathbb{R}^2} |F_n(y_1, y_2) - F(y_1, y_2)|$$

donde F_n es la función de distribución empírica de la muestra y F es alguna función de distribución especificada. Cuando se desconoce la distribución de F , la estadística de Kolmogorov-Smirnov se define,

$$dn(F) = \max\{D^1, D^2\},$$

así,

$$D^1 = \sup_{(y_1, y_2) \in \mathbb{R}^2} |\mathcal{G}_n(x_1, x_2) - x_1 \times x_2|$$

usando la transformación $x_1 = F_{Y_1}(Y_1)$ y $x_2 = F_{Y_2|Y_1}(Y_2|Y_1)$, y

$$D^2 = \sup_{(y_1, y_2) \in \mathbb{R}^2} |\mathcal{G}_n(x_2, x_1) - x_2 \times x_1|$$

usando la transformación $x_2 = F_{Y_2}(Y_2)$ y $x_1 = F_{Y_1|Y_2}(Y_1|Y_2)$, donde \mathcal{G} es la función de distribución empírica de la muestra, para este caso el modelo WUB

$$dn(WUB) = \max\{0.07538, 0.10563\} = 0.10563,$$

que es inferior a 0,1685, que es el valor crítico de la Tabla 1 dado por Justel et al. (1997), al nivel del 5%. Por lo tanto, se concluye que el modelo WUB se ajusta bien al conjunto de datos de los municipios de Antioquia en el año 2018.

Capítulo 6

Conclusiones y trabajos futuros

- La distribución Weibull Unitaria Biavariada (WUB) propuesta en el presente trabajo permite ajustar proporciones o tasas bivariadas. Se hallaron las funciones: marginales, condicionales, de supervivencia, hazard, entre otras. Se introdujo un método para la generación de variables aleatorias, por medio de simulación numérica se halló el rango de valores que puede tomar el coeficiente de correlación y se estimaron los parámetros con tres diferentes métodos de estimación.
- En el estudio de simulación realizado, se observó que los resultados obtenidos con el método de máxima verosimilitud y el método por inferencia de funciones marginales son muy similares.
- En la aplicación con datos reales el modelo WUB se compara con otros dos modelos: el modelo bivariado Johnson S_B , (BJSB) (Lemonte and Moreno-Arenas, 2019) y Birnbaum-Saunders, (BSB) (Kundu et al., 2010), nuestra propuesta se ajusta mejor al conjunto de datos. Esto nos permite concluir que la distribución WUB es una buena alternativa para modelar tasas o proporciones de manera conjunta.
- Como posible trabajo futuro se considera trabajar la extensión del modelo WUB al caso multivariado y estudiar sus propiedades más importantes, también explorar la extensión del modelo propuesto al caso de modelos de regresión.

Apéndice A

Tablas de simulaciones

Tabla A.1: Estimación de los parámetros de la distribución WUB: caso 1.1: $\alpha_1 = 1.5, \beta_1 = 3.5, \alpha_2 = 0.5, \beta_2 = 2.0, \theta = 0.25$

n	Método	Estadísticas	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\alpha}_2$	$\hat{\beta}_2$	$\hat{\theta}$
30	MLE	Bias	0.0405	0.1680	0.0048	0.0845	-0.0647
		MSE	0.0907	0.3265	0.0168	0.1211	0.1956
		L.CI	1.1173	2.0657	0.4760	1.1742	2.0777
		PCI	0.9549	0.9542	0.9246	0.9386	0.9689
	IFM	Bias	0.0408	0.1676	0.0044	0.0854	-0.0701
		MSE	0.0910	0.3229	0.0166	0.1195	0.2005
		L.CI	1.1157	2.0645	0.4749	1.1736	2.0586
		PCI	0.9573	0.9542	0.9245	0.9390	0.9741
50	MLE	Bias	0.0235	0.0906	0.0029	0.0543	-0.0320
		MSE	0.0482	0.1672	0.0097	0.0617	0.1535
		L.CI	0.8501	1.5614	0.3678	0.8916	1.6133
		PCI	0.9550	0.9543	0.9331	0.9338	0.9451
	IFM	Bias	0.0248	0.0926	0.0029	0.0543	-0.0342
		MSE	0.0489	0.1682	0.0097	0.0614	0.1558
		L.CI	0.8508	1.5624	0.3678	0.8917	1.6055
		PCI	0.9533	0.9533	0.9331	0.9393	0.9428
75	MLE	Bias	0.0189	0.0551	0.0048	0.0296	-0.0181
		MSE	0.0323	0.1062	0.0063	0.0375	0.1121
		L.CI	0.6904	1.2591	0.3011	0.7176	1.3245
		PCI	0.9448	0.9550	0.9373	0.9366	0.9441
	IFM	Bias	0.0197	0.0552	0.0048	0.0296	-0.0206
		MSE	0.0326	0.1067	0.0062	0.0373	0.1113
		L.CI	0.6909	1.2596	0.3013	0.7177	1.3186
		PCI	0.9455	0.9544	0.9408	0.9374	0.9415
100	MLE	Bias	0.0119	0.0405	0.0026	0.0222	-0.0154
		MSE	0.0239	0.0808	0.0047	0.0260	0.0869
		L.CI	0.5944	1.0851	0.2601	0.6189	1.1532
		PCI	0.9477	0.9510	0.9403	0.9544	0.9477
	IFM	Bias	0.0123	0.0400	0.0029	0.0219	-0.0166
		MSE	0.0239	0.0808	0.0047	0.0259	0.0859
		L.CI	0.5948	1.0853	0.2604	0.6190	1.1485
		PCI	0.9450	0.9490	0.9389	0.9537	0.9477
200	MLE	Bias	0.0079	0.0258	0.0008	0.0083	-0.0007
		MSE	0.0119	0.0383	0.0022	0.0125	0.0439
		L.CI	0.4184	0.7630	0.1836	0.4342	0.8182
		PCI	0.9467	0.9533	0.9453	0.9540	0.9393
	IFM	Bias	0.0080	0.0258	0.0009	0.0081	-0.0013
		MSE	0.0119	0.0383	0.0022	0.0125	0.0437
		L.CI	0.4187	0.7633	0.1838	0.4343	0.8160
		PCI	0.9473	0.9527	0.9467	0.9527	0.9387
500	MLE	Bias	0.0025	0.0063	-0.0008	0.0062	-0.0010
		MSE	0.0044	0.0146	0.0009	0.0052	0.0172
		L.CI	0.2634	0.4794	0.1160	0.2743	0.5192
		PCI	0.9507	0.9547	0.9467	0.9427	0.9547
	IFM	Bias	0.0025	0.0062	-0.0008	0.0061	-0.0012
		MSE	0.0044	0.0146	0.0009	0.0052	0.0172
		L.CI	0.2636	0.4796	0.1160	0.2744	0.5183
		PCI	0.9513	0.9527	0.9480	0.9420	0.9540

Tabla A.2: Estimación de los parámetros de la distribución WUB: caso 1.2: $\alpha_1 = 1.5, \beta_1 = 3.5, \alpha_2 = 0.5, \beta_2 = 2.0, \theta = 0.75$

n	Método	Estadísticas	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\alpha}_2$	$\hat{\beta}_2$	$\hat{\theta}$
30	MLE	Bias	0.0593	0.1801	0.0029	0.0908	-0.2892
		MSE	0.0997	0.3186	0.0159	0.1133	0.2248
		L.CI	1.1301	2.0703	0.4740	1.1771	1.9724
		PCI	0.9568	0.9568	0.9290	0.9424	0.9302
	IFM	Bias	0.0625	0.1920	0.0034	0.0916	-0.2993
		MSE	0.1049	0.3375	0.0165	0.1185	0.2341
		L.CI	1.1329	2.0781	0.4744	1.1778	2.0189
		PCI	0.9557	0.9547	0.9241	0.9389	0.9505
50	MLE	Bias	0.0291	0.1007	0.0012	0.0637	-0.1823
		MSE	0.0495	0.1678	0.0093	0.0644	0.1246
		L.CI	0.8510	1.5651	0.3658	0.8948	1.5255
		PCI	0.9545	0.9516	0.9370	0.9302	0.9486
	IFM	Bias	0.0263	0.1021	0.0006	0.0637	-0.1826
		MSE	0.0495	0.1703	0.0094	0.0651	0.1251
		L.CI	0.8517	1.5682	0.3669	0.8962	1.5299
		PCI	0.9536	0.9536	0.9366	0.9290	0.9612
75	MLE	Bias	0.0229	0.0679	0.0026	0.0417	-0.1263
		MSE	0.0321	0.1070	0.0060	0.0378	0.0777
		L.CI	0.6900	1.2622	0.2990	0.7203	1.2294
		PCI	0.9486	0.9593	0.9398	0.9486	0.9575
	IFM	Bias	0.0238	0.0673	0.0030	0.0404	-0.1264
		MSE	0.0326	0.1077	0.0061	0.0381	0.0783
		L.CI	0.6927	1.2644	0.3006	0.7215	1.2356
		PCI	0.9477	0.9582	0.9425	0.9442	0.9747
100	MLE	Bias	0.0172	0.0522	0.0015	0.0302	-0.1007
		MSE	0.0240	0.0827	0.0047	0.0266	0.0576
		L.CI	0.5946	1.0875	0.2586	0.6200	1.0706
		PCI	0.9452	0.9486	0.9351	0.9503	0.9646
	IFM	Bias	0.0172	0.0523	0.0015	0.0297	-0.1016
		MSE	0.0243	0.0830	0.0047	0.0266	0.0575
		L.CI	0.5967	1.0897	0.2599	0.6211	1.0661
		PCI	0.9432	0.9491	0.9307	0.9524	0.9766
200	MLE	Bias	0.0095	0.0325	0.0004	0.0101	-0.0368
		MSE	0.0120	0.0391	0.0022	0.0126	0.0277
		L.CI	0.4169	0.7626	0.1823	0.4334	0.7398
		PCI	0.9463	0.9553	0.9433	0.9530	0.9605
	IFM	Bias	0.0102	0.0313	0.0005	0.0099	-0.0363
		MSE	0.0122	0.0390	0.0022	0.0126	0.0278
		L.CI	0.4193	0.7647	0.1837	0.4347	0.7382
		PCI	0.9453	0.9542	0.9424	0.9549	0.9749
500	MLE	Bias	0.0023	0.0078	-0.0009	0.0063	-0.0068
		MSE	0.0044	0.0145	0.0009	0.0052	0.0125
		L.CI	0.2619	0.4782	0.1151	0.2735	0.4667
		PCI	0.9563	0.9549	0.9494	0.9460	0.9706
	IFM	Bias	0.0026	0.0070	-0.0008	0.0061	-0.0052
		MSE	0.0044	0.0147	0.0009	0.0052	0.0129
		L.CI	0.2636	0.4797	0.1160	0.2744	0.4621
		PCI	0.9506	0.9526	0.9438	0.9445	0.9675

Tabla A.3: Estimación de los parámetros de la distribución WUB: caso 1.3: $\alpha_1 = 1.5, \beta_1 = 3.5, \alpha_2 = 0.5, \beta_2 = 2.0, \theta = -0.25$

n	Método	Estadísticas	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\alpha}_2$	$\hat{\beta}_2$	$\hat{\theta}$
30	MLE	Bias	0.0374	0.1699	0.0068	0.0826	0.0878
		MSE	0.0868	0.3288	0.0167	0.1187	0.2061
		L.CI	1.1148	2.0670	0.4777	1.1732	2.0815
		PCI	0.9527	0.9566	0.9310	0.9380	0.9744
	IFM	Bias	0.0429	0.1724	0.0074	0.0821	0.0927
		MSE	0.0940	0.3299	0.0168	0.1182	0.2156
		L.CI	1.1176	2.0667	0.4767	1.1715	2.0661
		PCI	0.9533	0.9563	0.9337	0.9352	0.9721
50	MLE	Bias	0.0225	0.0900	0.0043	0.0517	0.0400
		MSE	0.0487	0.1688	0.0099	0.0602	0.1509
		L.CI	0.8496	1.5605	0.3685	0.8906	1.6170
		PCI	0.9532	0.9546	0.9362	0.9376	0.9489
	IFM	Bias	0.0214	0.0903	0.0047	0.0511	0.0450
		MSE	0.0486	0.1682	0.0098	0.0598	0.1514
		L.CI	0.8489	1.5606	0.3686	0.8904	1.6097
		PCI	0.9519	0.9534	0.9385	0.9350	0.9519
75	MLE	Bias	0.0175	0.0519	0.0045	0.0300	0.0158
		MSE	0.0328	0.1063	0.0063	0.0370	0.1133
		L.CI	0.6896	1.2578	0.3009	0.7181	1.3217
		PCI	0.9420	0.9556	0.9413	0.9365	0.9427
	IFM	Bias	0.0176	0.0518	0.0044	0.0303	0.0138
		MSE	0.0326	0.1062	0.0062	0.0370	0.1136
		L.CI	0.6899	1.2582	0.3011	0.7184	1.3177
		PCI	0.9450	0.9552	0.9409	0.9382	0.9402
100	MLE	Bias	0.0114	0.0393	0.0031	0.0214	-0.0051
		MSE	0.0239	0.0807	0.0047	0.0257	0.0850
		L.CI	0.5940	1.0847	0.2603	0.6189	1.1487
		PCI	0.9489	0.9482	0.9408	0.9509	0.9509
	IFM	Bias	0.0113	0.0392	0.0031	0.0212	-0.0044
		MSE	0.0238	0.0809	0.0047	0.0258	0.0844
		L.CI	0.5943	1.0851	0.2605	0.6191	1.1441
		PCI	0.9456	0.9489	0.9415	0.9489	0.9469
200	MLE	Bias	0.0079	0.0258	0.0010	0.0082	-0.0005
		MSE	0.0119	0.0383	0.0022	0.0124	0.0446
		L.CI	0.4184	0.7630	0.1837	0.4342	0.8182
		PCI	0.9493	0.9513	0.9507	0.9507	0.9407
	IFM	Bias	0.0080	0.0258	0.0009	0.0083	0.0001
		MSE	0.0119	0.0383	0.0022	0.0124	0.0444
		L.CI	0.4187	0.7633	0.1838	0.4344	0.8158
		PCI	0.9473	0.9527	0.9540	0.9527	0.9400
500	MLE	Bias	0.0025	0.0062	-0.0008	0.0060	-0.0013
		MSE	0.0044	0.0146	0.0009	0.0051	0.0173
		L.CI	0.2634	0.4794	0.1160	0.2743	0.5193
		PCI	0.9513	0.9540	0.9507	0.9447	0.9573
	IFM	Bias	0.0025	0.0062	-0.0008	0.0061	-0.0011
		MSE	0.0044	0.0146	0.0009	0.0051	0.0172
		L.CI	0.2636	0.4796	0.1160	0.2744	0.5184
		PCI	0.9513	0.9527	0.9513	0.9447	0.9573

Tabla A.4: Estimación de los parámetros de la distribución WUB: caso 1.4: $\alpha_1 = 1.5, \beta_1 = 3.5, \alpha_2 = 0.5, \beta_2 = 2.0, \theta = -0.75$

n	Método	Estadísticas	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\alpha}_2$	$\hat{\beta}_2$	$\hat{\theta}$
30	MLE	Bias	0.0402	0.1742	0.0083	0.0899	0.2911
		MSE	0.0844	0.3241	0.0166	0.1144	0.2160
		L.CI	1.1148	2.0707	0.4777	1.1762	1.9978
		PCI	0.9588	0.9631	0.9262	0.9360	0.9393
	IFM	Bias	0.0463	0.1864	0.0093	0.0926	0.2995
		MSE	0.0988	0.3481	0.0172	0.1169	0.2258
		L.CI	1.1207	2.0754	0.4778	1.1774	2.0209
		PCI	0.9508	0.9559	0.9231	0.9364	0.9579
50	MLE	Bias	0.0235	0.1047	0.0044	0.0544	0.1941
		MSE	0.0496	0.1712	0.0099	0.0591	0.1292
		L.CI	0.8488	1.5659	0.3677	0.8918	1.5209
		PCI	0.9553	0.9553	0.9329	0.9377	0.9436
	IFM	Bias	0.0271	0.1017	0.0049	0.0550	0.1966
		MSE	0.0515	0.1708	0.0103	0.0603	0.1295
		L.CI	0.8523	1.5661	0.3687	0.8926	1.5376
		PCI	0.9478	0.9573	0.9307	0.9402	0.9658
75	MLE	Bias	0.0167	0.0658	0.0027	0.0365	0.1315
		MSE	0.0325	0.1136	0.0060	0.0379	0.0813
		L.CI	0.6872	1.2606	0.2992	0.7198	1.2148
		PCI	0.9400	0.9472	0.9463	0.9418	0.9346
	IFM	Bias	0.0198	0.0598	0.0021	0.0373	0.1324
		MSE	0.0329	0.1130	0.0061	0.0381	0.0809
		L.CI	0.6910	1.2610	0.3003	0.7215	1.2337
		PCI	0.9435	0.9496	0.9426	0.9417	0.9664
100	MLE	Bias	0.0114	0.0495	0.0025	0.0245	0.0905
		MSE	0.0238	0.0842	0.0045	0.0250	0.0544
		L.CI	0.5918	1.0856	0.2588	0.6192	1.0481
		PCI	0.9503	0.9435	0.9409	0.9572	0.9563
	IFM	Bias	0.0104	0.0495	0.0016	0.0261	0.0895
		MSE	0.0238	0.0850	0.0045	0.0253	0.0543
		L.CI	0.5940	1.0885	0.2601	0.6213	1.0537
		PCI	0.9452	0.9427	0.9435	0.9536	0.9755
200	MLE	Bias	0.0084	0.0297	0.0013	0.0100	0.0387
		MSE	0.0120	0.0386	0.0022	0.0125	0.0285
		L.CI	0.4166	0.7617	0.1825	0.4337	0.7375
		PCI	0.9482	0.9512	0.9497	0.9474	0.9684
	IFM	Bias	0.0089	0.0287	0.0011	0.0100	0.0399
		MSE	0.0120	0.0383	0.0022	0.0125	0.0285
		L.CI	0.4189	0.7638	0.1839	0.4350	0.7376
		PCI	0.9461	0.9528	0.9506	0.9513	0.9813
500	MLE	Bias	0.0024	0.0071	-0.0008	0.0063	0.0046
		MSE	0.0044	0.0147	0.0009	0.0051	0.0127
		L.CI	0.2619	0.4781	0.1151	0.2735	0.4650
		PCI	0.9526	0.9560	0.9464	0.9444	0.9547
	IFM	Bias	0.0023	0.0069	-0.0007	0.0061	0.0036
		MSE	0.0044	0.0148	0.0009	0.0051	0.0129
		L.CI	0.2635	0.4797	0.1161	0.2744	0.4621
		PCI	0.9517	0.9523	0.9510	0.9462	0.9598

Tabla A.5: Estimación de los parámetros de la distribución WUB: caso 2.1: $\alpha_1 = 2.5, \beta_1 = 1.5, \alpha_2 = 1.5, \beta_2 = 1.0, \theta = 0.25$

n	Método	Estadísticas	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\alpha}_2$	$\hat{\beta}_2$	$\hat{\theta}$
30	MLE	Bias	0.1408	0.0687	0.0665	0.0435	-0.0683
		MSE	0.3577	0.0572	0.1152	0.0302	0.1945
		L.CI	2.0970	0.8835	1.1390	0.5880	2.0834
		PCI	0.9664	0.9578	0.9445	0.9391	0.9766
	IFM	Bias	0.1419	0.0716	0.0653	0.0424	-0.0720
		MSE	0.3623	0.0592	0.1125	0.0299	0.1997
		L.CI	2.0962	0.8848	1.1362	0.5866	2.0627
		PCI	0.9633	0.9541	0.9442	0.9388	0.9755
50	MLE	Bias	0.0780	0.0391	0.0420	0.0274	-0.0310
		MSE	0.1716	0.0309	0.0613	0.0155	0.1540
		L.CI	1.5612	0.6693	0.8616	0.4460	1.6160
		PCI	0.9592	0.9529	0.9424	0.9339	0.9487
	IFM	Bias	0.0804	0.0397	0.0429	0.0271	-0.0353
		MSE	0.1740	0.0309	0.0611	0.0154	0.1553
		L.CI	1.5625	0.6696	0.8620	0.4458	1.6061
		PCI	0.9588	0.9532	0.9455	0.9392	0.9441
75	MLE	Bias	0.0552	0.0239	0.0313	0.0149	-0.0176
		MSE	0.1094	0.0197	0.0370	0.0094	0.1123
		L.CI	1.2555	0.5397	0.6963	0.3588	1.3253
		PCI	0.9632	0.9537	0.9469	0.9367	0.9442
	IFM	Bias	0.0559	0.0237	0.0319	0.0148	-0.0206
		MSE	0.1101	0.0196	0.0370	0.0093	0.1113
		L.CI	1.2558	0.5398	0.6967	0.3589	1.3186
		PCI	0.9592	0.9544	0.9462	0.9374	0.9415
100	MLE	Bias	0.0367	0.0174	0.0206	0.0110	-0.0149
		MSE	0.0800	0.0148	0.0268	0.0065	0.0872
		L.CI	1.0753	0.4651	0.5980	0.3094	1.1532
		PCI	0.9631	0.9510	0.9370	0.9544	0.9484
	IFM	Bias	0.0374	0.0172	0.0212	0.0109	-0.0166
		MSE	0.0803	0.0148	0.0267	0.0065	0.0859
		L.CI	1.0757	0.4651	0.5984	0.3095	1.1484
		PCI	0.9624	0.9490	0.9362	0.9537	0.9477
200	MLE	Bias	0.0237	0.0111	0.0065	0.0041	-0.0012
		MSE	0.0393	0.0070	0.0118	0.0031	0.0436
		L.CI	0.7532	0.3270	0.4180	0.2171	0.8183
		PCI	0.9453	0.9533	0.9460	0.9540	0.9400
	IFM	Bias	0.0238	0.0110	0.0065	0.0041	-0.0013
		MSE	0.0393	0.0070	0.0118	0.0031	0.0437
		L.CI	0.7535	0.3271	0.4183	0.2172	0.8160
		PCI	0.9480	0.9527	0.9487	0.9527	0.9387
500	MLE	Bias	0.0069	0.0027	0.0015	0.0031	-0.0010
		MSE	0.0147	0.0027	0.0046	0.0013	0.0172
		L.CI	0.4716	0.2055	0.2633	0.1372	0.5192
		PCI	0.9533	0.9547	0.9507	0.9427	0.9547
	IFM	Bias	0.0069	0.0027	0.0015	0.0031	-0.0012
		MSE	0.0147	0.0027	0.0046	0.0013	0.0172
		L.CI	0.4717	0.2055	0.2634	0.1372	0.5183
		PCI	0.9533	0.9527	0.9493	0.9420	0.9533

Tabla A.6: Estimación de los parámetros de la distribución WUB: caso 2.2: $\alpha_1 = 2.5, \beta_1 = 1.5, \alpha_2 = 1.5, \beta_2 = 1.0, \theta = 0.75$

n	Método	Estadísticas	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\alpha}_2$	$\hat{\beta}_2$	$\hat{\theta}$
30	MLE	Bias	0.1775	0.0747	0.0646	0.0428	-0.2893
		MSE	0.3854	0.0579	0.1121	0.0281	0.2221
		L.CI	2.1327	0.8858	1.1364	0.5872	2.0001
		PCI	0.9701	0.9590	0.9369	0.9435	0.9391
	IFM	Bias	0.1901	0.0834	0.0664	0.0470	-0.2994
		MSE	0.4210	0.0615	0.1114	0.0298	0.2341
		L.CI	2.1476	0.8912	1.1369	0.5895	2.0130
		PCI	0.9684	0.9568	0.9399	0.9389	0.9515
50	MLE	Bias	0.0854	0.0433	0.0442	0.0308	-0.1791
		MSE	0.1687	0.0306	0.0552	0.0157	0.1200
		L.CI	1.5643	0.6708	0.8600	0.4471	1.5351
		PCI	0.9639	0.9513	0.9561	0.9318	0.9581
	IFM	Bias	0.0858	0.0435	0.0437	0.0320	-0.1838
		MSE	0.1726	0.0312	0.0559	0.0162	0.1253
		L.CI	1.5669	0.6721	0.8619	0.4482	1.5338
		PCI	0.9630	0.9526	0.9554	0.9298	0.9649
75	MLE	Bias	0.0669	0.0297	0.0356	0.0200	-0.1269
		MSE	0.1082	0.0197	0.0365	0.0095	0.0776
		L.CI	1.2607	0.5412	0.6959	0.3600	1.2439
		PCI	0.9654	0.9592	0.9504	0.9459	0.9690
	IFM	Bias	0.0672	0.0288	0.0359	0.0202	-0.1264
		MSE	0.1091	0.0198	0.0366	0.0095	0.0783
		L.CI	1.2628	0.5419	0.6984	0.3607	1.2354
		PCI	0.9616	0.9582	0.9538	0.9442	0.9738
100	MLE	Bias	0.0486	0.0223	0.0244	0.0150	-0.0983
		MSE	0.0807	0.0153	0.0266	0.0067	0.0576
		L.CI	1.0801	0.4660	0.5972	0.3100	1.0676
		PCI	0.9582	0.9465	0.9406	0.9498	0.9632
	IFM	Bias	0.0500	0.0224	0.0234	0.0148	-0.1014
		MSE	0.0823	0.0152	0.0268	0.0066	0.0575
		L.CI	1.0827	0.4670	0.5992	0.3106	1.0655
		PCI	0.9608	0.9491	0.9391	0.9525	0.9750
200	MLE	Bias	0.0307	0.0139	0.0068	0.0049	-0.0361
		MSE	0.0402	0.0072	0.0118	0.0032	0.0277
		L.CI	0.7541	0.3268	0.4160	0.2167	0.7413
		PCI	0.9480	0.9554	0.9546	0.9524	0.9673
	IFM	Bias	0.0294	0.0134	0.0071	0.0048	-0.0359
		MSE	0.0402	0.0072	0.0119	0.0031	0.0278
		L.CI	0.7557	0.3277	0.4184	0.2173	0.7375
		PCI	0.9454	0.9543	0.9550	0.9550	0.9742
500	MLE	Bias	0.0070	0.0033	0.0011	0.0030	-0.0059
		MSE	0.0147	0.0027	0.0045	0.0013	0.0126
		L.CI	0.4702	0.2049	0.2617	0.1367	0.4666
		PCI	0.9544	0.9558	0.9483	0.9455	0.9741
	IFM	Bias	0.0073	0.0029	0.0014	0.0030	-0.0051
		MSE	0.0147	0.0027	0.0046	0.0013	0.0129
		L.CI	0.4718	0.2056	0.2634	0.1372	0.4620
		PCI	0.9527	0.9520	0.9466	0.9446	0.9655

Bibliografía

- Almetwally, E. M., Muhammed, H. Z., and El-Sherpieny, E.-S. A. (2020). Bivariate weibull distribution: properties and different methods of estimation. *Annals of Data Science*, 7(1):163–193.
- Arnold, B. C., Castillo, E., and Sarabia, J. M. (2002). Conditionally specified multivariate skewed distributions. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 206–226.
- Basu, A. (1971). Bivariate failure rate. *Journal of the American Statistical Association*, 66(333):103–104.
- Cook, D. O., Kieschnick, R., and McCullough, B. D. (2008). Regression analysis of proportions in finance with self selection. *Journal of empirical finance*, 15(5):860–867.
- Dürr, J., Monardes, E., and Cue, R. (1999). Genetic analysis of herd life in quebec holsteins using weibull models. *Journal of Dairy Science*, 82(11):2503–2513.
- Fok, S., Mitchell, B., Smart, J., and Marsden, B. (2001). A numerical study on the application of the weibull theory to brittle materials. *Engineering Fracture Mechanics*, 68(10):1171–1179.
- Gumbel, E. J. (1960). Bivariate exponential distributions. *Journal of the American Statistical Association*, 55(292):698–707.
- Gupta, A. K. and Nadarajah, S. (2004). *Handbook of beta distribution and its applications*. CRC press.
- Joe, H. (2005). Asymptotic efficiency of the two-stage estimation method for copula-based models. *Journal of multivariate Analysis*, 94(2):401–419.
- Justel, A., Peña, D., and Zamar, R. (1997). A multivariate kolmogorov-smirnov test of goodness of fit. *Statistics & Probability Letters*, 35(3):251–259.

- Kumaraswamy, P. (1980). A generalized probability density function for double-bounded random processes. *Journal of hydrology*, 46(1-2):79–88.
- Kundu, D., Balakrishnan, N., and Jamalizadeh, A. (2010). Bivariate birnbaum–saunders distribution and associated inference. *Journal of multivariate Analysis*, 101(1):113–125.
- Kundu, D. and Dey, A. K. (2009). Estimating the parameters of the marshall–olkin bivariate weibull distribution by em algorithm. *Computational Statistics & Data Analysis*, 53(4):956–965.
- Lemonte, A. J., Martínez-Florez, G., and Moreno-Arenas, G. (2015). Multivariate birnbaum–saunders distribution: Properties and associated inference. *Journal of Statistical Computation and Simulation*, 85(2):374–392.
- Lemonte, A. J. and Moreno-Arenas, G. (2019). On a multivariate regression model for rates and proportions. *Journal of Applied Statistics*, 46(6):1084–1106.
- Liu, C., Aengenheister, S., Herzog, S., Deng, Y., Kaletsch, A., Schmitz, K., and Broeckmann, C. (2021). Application of weibull theory to laser surface textured al₂o₃. *Journal of the European Ceramic Society*, 41(2):1415–1426.
- Mazucheli, J., Menezes, A., and Dey, S. (2018a). The unit-birnbaum-saunders distribution with applications. *Chilean Journal of Statistics*, 9(1):47–57.
- Mazucheli, J., Menezes, A., and Ghitany, M. (2018b). The unit-weibull distribution and associated inference. *J. Appl. Probab. Stat*, 13:1–22.
- Nelsen, R. B. (2007). *An introduction to copulas*. Springer Science & Business Media.
- R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Sklar, A. (1973). Random variables, joint distribution functions, and copulas. *Kybernetika*, 9(6):449–460.
- Weibull, W. (1951). A statistical distribution function of wide applicability. *Journal of Applied Mechanics*, 18(3):293–297.
- Xu, M., Herrmann, J. W., and Droguett, E. L. (2021). Modeling dependent series systems with q-weibull distribution and clayton copula. *Applied Mathematical Modelling*, 94:117–138.
- Zhang, C. W. (2021). Weibull parameter estimation and reliability analysis with zero-failure data from high-quality products. *Reliability Engineering & System Safety*, 207:107321.