Simulación Monte Carlo para la estimación de parámetros de la distribución Weibull Inversa *

Johan David Marin Benjumea *Universidad de Antioquia*
Freddy Hernández Barajas *Universidad Nacional de Colombia - sede Medellín-*
Olga Usuga Manco *Universidad de Antioquia*
Carmen Elena Patiño Rodriguez *Universidad de Antioquia*

Resumen: La distribución Weibull Inversa (IW) ha sido ampliamente utilizada en estudios estadísticos de datos con distribución sesgada decreciente, en diversas áreas del conocimiento, incluida la confiabilidad. En este trabajo se realiza un estudio del comportamiento asintótico de los estimadores de los parámetros de la distribución IW del paquete `RelDists` para diferentes tamaños de muestra. Para esto se realizó un estudio de simulación mediante el método Monte Carlo, concluyendo que el error de estimación es inversamente proporcional al tamaño de muestra.

Keywords: Inverse Weibull, Simulación Monte Carlo, RelDists, gamlss

Introducción

Debido al uso de distribuciones asimétricas en el estudio de variables respuesta, con distribución diferente a la normal, los autores del presente tabajo crearon el paquete `RelDists`, que recopila varias distribuciones de probabilidad, permitiendo realizar estimación de parámetros y ajustar modelos de regresión, para variables no normales, haciendo uso del paquete `gamlss` (Rigby and Stasinopoulos, 2005).

El objetivo del trabajo es analizar el comportamiento asintótico de los estimadores de los parámetros de la distribución IW por medio de un estudio de simulación Monte Carlo, considerando dos casos: estimación sin covariables y con covariables. El estudio se realizó con la distribución Weibull Inversa (IW), una de las distribuciones implementadas en `RelDists`, que por su flexibilidad ha sido analizada, modificada y utilizada en varios estudios.

El documento está organizado como se describe a continuación. Se hace una introducción a la distribución Weibull Inversa, se presenta la metodología usada en el estudio de simulación seguido de los resultados obtenidos y se finaliza con las conclusiones.

Distribución Weibull inversa

La distribución Weibull Inversa, presentada por Mudholkar and Kollia (1994), es una extensión de la distribución Weibull con dos parámetros (μ y σ), usada para modelar variables cuya inversa tenga una distribución Weibull.

Sea X una variable aleatoria que sigue una distribución Weibull inversa, $X \sim IW(\mu, \sigma)$, cuya función de densidad de probabilidad (PDF), función densidad de probabilidad acumulada (CDF)

* **Autor de contacto**: johand.marin@udea.edu.co.

y función de riesgo (HF) están dadas por las siguientes ecuaciones:

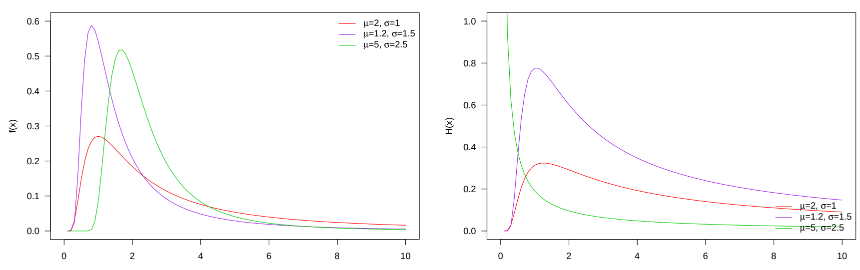
$$f(x, \mu, \sigma) = \mu\sigma x^{-\sigma-1} \exp(\mu x^{-\sigma}), \quad x > 0,\ \mu > 0,\ \sigma > 0 \quad (1)$$

$$F(x, \mu, \sigma) = \exp(-\mu x^{-\sigma}), \quad x > 0,\ \mu > 0,\ \sigma > 0 \quad (2)$$

$$h(x, \mu, \sigma) = \mu\sigma x^{-\sigma-1}, \quad x > 0,\ \mu > 0,\ \sigma > 0, \quad (3)$$

donde μ es el parámetro de escala y σ el parámetro de forma.

La Figura 1 ilustra algunos comportamientos de las funciones PDF y HF para diferentes valores de μ y σ. En la misma Figura 1 puede observarse que tanto la PDF como la HF son unimodales y decrecientes, dependiendo del parámetro de forma μ el comportamiento de la distribución IW es bastante similar a la distribución Log-normal (Kundu and Howlader, 2010) y cuando el parámetro de escala $\sigma = 1$ se comporta como la distribución exponencial (Khan, Pasha and Pasha, 2008).



(a) PDF para la distribución IW (b) HF para la distribución IW

Figure 1: PDF y HF de la distribución IW para distintos valores de μ y σ.

El uso de la distribución IW es muy diverso. Una búsqueda en Scopus, utilizando el algoritmo **TITLE-ABS-KEY ("Inverse weibull" OR "IW distribution" AND NOT(exponential OR exponentiated OR generalized))**, arrojó un resultado de 104 documentos, clasificados como se muestra en la Figura 2. De la Figura 2 (a) se observó un incremento en su uso durante los últimos 4 años comparado con años anteriores y de la Figura 2 (b) se observó un alto porcentaje de documentos de Artículos y Documentos de conferencias. Por su parte, en la Figura 2 (c) se puede ver que la distribución IW se ha utilizado en múltiples áreas del conocimiento, desde áreas como las matemáticas y la ingeniería, hasta ciencias sociales, ciencias planetarias y Veterinaria, mostrando ser adecuada para modelar fenómenos como la degradación de componentes mecánicos, los tiempos de descomposición de un fluido aislante (Khan, Pasha and Pasha, 2008), la velocidad del viento (Simiu and Heckert, 1996), el tiempo de supervivencia de cobayas (Mudholkar and Kollia, 1994), y muchos otros usos. Por último, la Figura 2 (d) muestra su uso en diversos países, en especial en India y Estados Unidos.

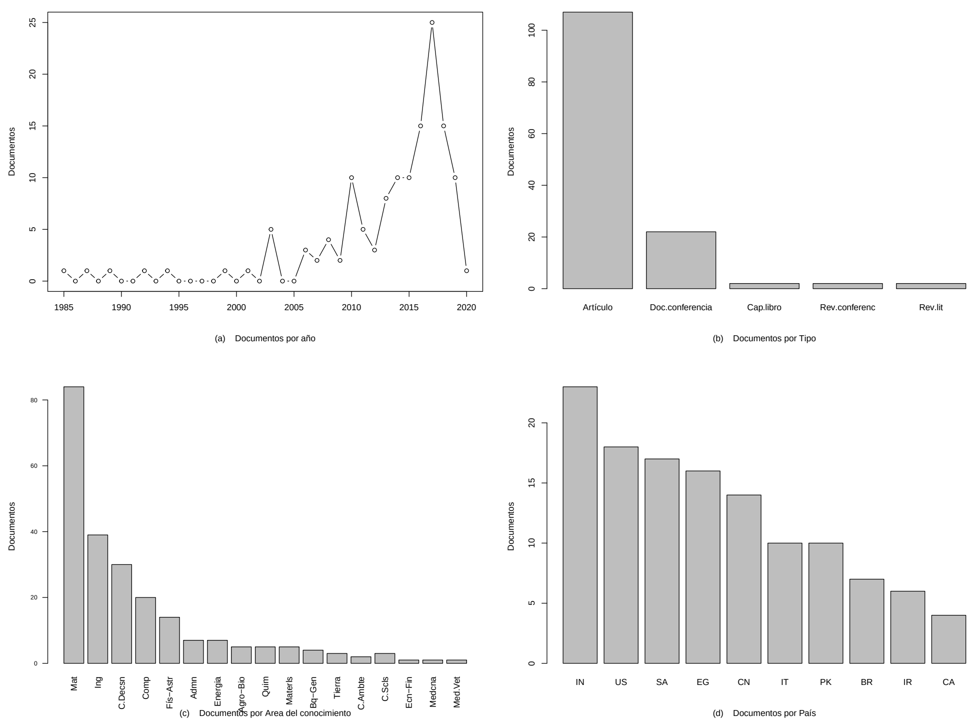


Figure 2: Clasificación de documentos relacionados con la distribución IW encontrados en Scopus y clasificados por, (a) Año, (b) Tipo de documento, (c) Área de conocimiento y (d) País de origen .

Estudio de Simulación

Para el proceso de simulación se utilizó el lenguaje de programación R (R Core Team, 2019), el paquete `gamlss` (Rigby and Stasinopoulos, 2005) y las funciones densidad de probabilidad `dIW`, densidad de probabilidad acumulada `pIW`, la función quantil `qIW`, la función generadora de números aleatorios `rIW`, la función de riesgo `hW` y la familia `IW` para la función Weibull Inversa, implementadas en el paquete `RelDists`.

Utilizando estas funciones se realizó un estudio de simulación Monte Carlo, sin y con covariables. En el primer caso se consideraron los valores de $\mu = 2$ y $\sigma = 1.5$, mientras que en el segundo se generaron aleatoriamente valores para dos covariables (X_1 y X_2) para luego obtener los valores para μ y σ.

Para la estimación de parámetros sin covariables se consideraron tamaños de muestra de $n = 20, 40, ..., 240$ y para la estimación de parámetros con covariables se usaron tamaños de muestra de $n = 50, 70, ..., 790$. Los parámetros se estimaron por medio de la función `gamlss` (Rigby and Stasinopoulos, 2005) y la familia `IW` como se muestra en el segundo bloque de código, de ello se obtienen las estimaciones de los coeficientes μ y σ para la estimación sin covariables y para la estimación con covariables los coeficientes β_0 y β_1 para μ y los coeficientes γ_0 y γ_1 para σ. Para evaluar el desempeño del proceso de estimación se uso el MSE.

El código para estimar los parámetros de $Y \sim IW(\mu, \sigma)$ fue el siguiente:

```
y <- rIW(n=n, mu, sigma)
mod <- gamlss(y~1, sigma.fo=~1, family=IW,
              control=gamlss.control(n.cyc=5000, trace=FALSE))
```

El código para estimar los parámetros de regresión de $Y \sim IW(\mu, \sigma)$ en función de X_1 y X_2 fue el siguiente:

```
x1 <- rpois(n=size, lambda=2)
x2 <- runif(n=size)
mu <- exp(3 + -1 * x1)
sigma <- exp(2 - 2 * x2)
y <- rIW(n=n, mu, sigma)
mod <- gamlss(y~x1, sigma.fo=~x2, family=IW,
              control=gamlss.control(n.cyc=5000, trace=FALSE))
```

Resultados

En la Figura 3 se puede observar una mejora en la estimación de los parámetros a medida que el tamaño de la muestra aumenta, es decir, el valor promedio de las estimaciones, $\hat{\mu}$ y $\hat{\sigma}$, es más cercano al valor utilizado para generar los datos a medida que el tamaño de la muestra aumenta. En la Figura 4 se ve como el MSE tiende a cero a medida que aumenta el tamaño de muestra n, lo que indica que tanto para modelar la distribución de una variable $Y \sim IW(\mu, \sigma)$, como para realizar un modelo de regresión el tener un tamaño de muestra grande es importante para una buena estimación de los parámetros.

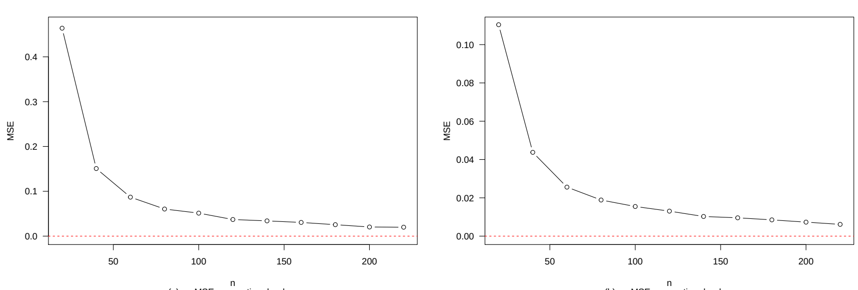


Figure 3: MSE para los estimadores de (a) $\mu = 2$ y (b) $\sigma = 1.5$, con distintos n.

Conclusión

Los Resultados de la simulación mostraron que la implementación de la distribución Weibull inversa en el paquete `RelDists` es adecuada, ya que cuando el tamaño de muestra es grande la estimación de parámetros converge al valor real de estos, convirtiendo a los paquetes `RelDists` y `gamlss` en un conjunto de herramientas adecuadas para ajustar parámetros de distribución y modelos de regresión con la función `IW`.

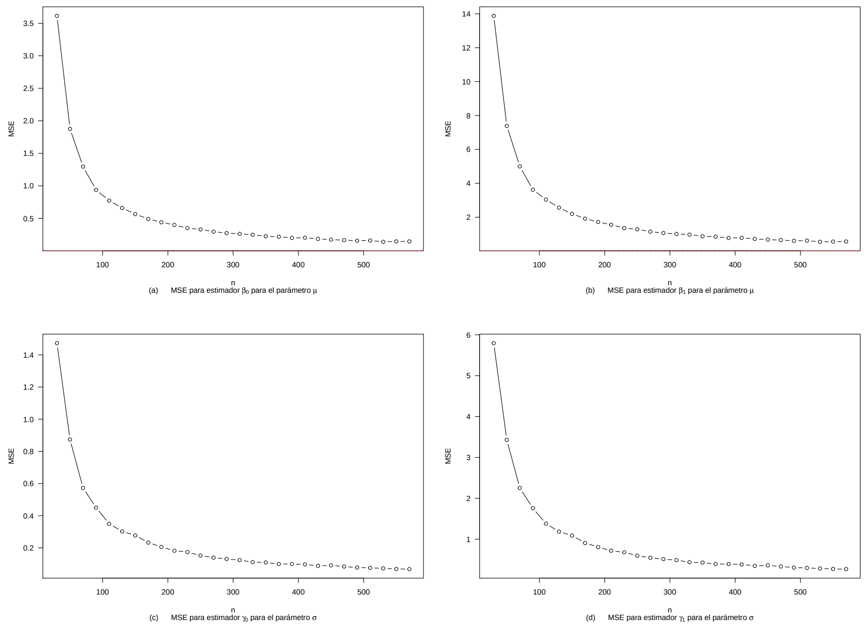


Figure 4: MSE de los estimadores con covariables de los parámetros de la distribución IW, $\hat{\mu}$: (a) β_0 y (b) β_1, y $\hat{\sigma}$: (c) γ_0 y (d) γ_1, con distintos n.

References

Khan, M Shuaib, GR Pasha and Ahmed Hesham Pasha. 2008. "Theoretical analysis of inverse Weibull distribution." *WSEAS Transactions on Mathematics* 7(2):30–38.

Kundu, Debasis and Hatem Howlader. 2010. "Bayesian inference and prediction of the inverse Weibull distribution for Type-II censored data." *Computational Statistics & Data Analysis* 54(6):1547–1558.

Mudholkar, Govind S and Georgia D Kollia. 1994. "Generalized Weibull family: a structural analysis." *Communications in statistics-theory and methods* 23(4):1149–1171.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
URL: *https://www.R-project.org/*

Rigby, R. A. and D. M. Stasinopoulos. 2005. "Generalized additive models for location, scale and shape,(with discussion)." *Applied Statistics* 54:507–554.

Simiu, Emil and NA Heckert. 1996. "Extreme wind distribution tails: a "peaks over threshold" approach." *Journal of Structural Engineering* 122(5):539–547.