Weighted Least Square Estimation : Monte Carlo Simulation and Application

Zahia Khemissi[1], **Brahim Brahimi**[2]

[1 2] Laboratory of Applied Mathematics, Mohamed Khider University, Biskra, Algeria

zahia.khemissi@univ-biskra.dz, brah.brahim@gmail.com

Least squares (LS) regression methods based on the relationship between the empirical cumulative distribution function (cdf) and the order statistics are frequently used to estimate parameters of distributions. In this work we estimate the two-parameter Fréchet distribution using the weighted least-squares method (WLS), based on an different expressions of weight. In the estimation procedure, the weights are expressed as a function of the rank of the data point in the sample.

As a motivating example for our methodology, consider two-parameter Fréchet distribution with shape parameter α, scale parameter β. The probability density function and the cdf are respectively given in the following equations:

$$f(x) = \frac{\alpha}{\beta}\left(\frac{\beta}{x}\right)^{\alpha+1} \exp\left(-\left(\frac{\beta}{x}\right)^{\alpha}\right),$$

$$F(x) = \exp\left(-\left(\frac{\beta}{x}\right)^{\alpha}\right) \quad \text{for} \quad x > 0, \alpha, \beta > 0 \tag{1}$$

After some algebraic manipulation, and for a sample size of n and $x_{(1)} \leq x_{(2)} \leq ... \leq x_{(n)}$ equation (1) can be expressed as follows:

$$\ln x_{(i)} = -\ln\beta - \frac{1}{\alpha}\ln\left(-\ln\left(F(x_{(i)})\right)\right) \tag{2}$$

where $\ln x_{(i)}$ is the i^{th} order statistics of the logarithm of the sample from the Fréchet distribution. $\frac{i-a}{(n+b)}$, $(0 \leq a \leq 0.5, 0 \leq b \leq 1)$ is used as estimate of $F(x_{(i)})$ where i is the rank of the data point in the sample in ascending order.

If we replace $\ln x_{(i)}$ with Y_i, $-\ln\beta$ with a_0, $-\frac{1}{\alpha}$ with a_1 and $\ln\left(-\ln\left(F(x_{(i)})\right)\right)$ with X_i,the regression model (2) occurs as:

$$Y_i = a_0 + a_1 X_i.$$

The WLS procedure for the Frechet distribution can be carried out by minimizing the weighted sum of squares with respect to the unknown shape and scale parameters, thus:

$$\min_{a_0,a_1} \sum_{i=1}^{n} w_i \left[Y_i - a_0 - a_0 X_i)\right]^2,$$

where w_i is the weight factor $i = 1, \ldots, n$.

In this work, we present several expressions for weights, and we show by Monte Carlo simulation that the weighted regression outperforms the in several methods such as maximum likelihood estimation, LS (X on Y, Y on X) estimation methods, in terms of bias and root mean square error for most of the considered sample sizes. In addition, a real example from Danish data is provided to demonstrate the performance of the considered method.