Maximum Likelihood Estimation: Numerical Solution for Weibull Distribution[1]

The Weibull distribution has the probability density function:

$$f\left(y|\theta,\lambda\right) \;=\; \frac{\lambda y^{\lambda-1}}{\theta^{\lambda}}\exp\left\{-\left(\frac{y}{\theta}\right)^{\lambda}\right\} \tag{1}$$

where: $y > 0$, θ is the scale parameter and λ is the shape parameter (nuisance parameter).

Maximum Likelihood Estimation

Let $y_1, \ldots, y_n$, denote the data. Assume, $\forall i : y_i$ independente random variables, share the same parameters from a Weibull distribution described in (1).

$$f\left(y_i|\theta,\lambda\right) \;=\; \frac{\lambda {y_i}^{\lambda-1}}{\theta^{\lambda}}\exp\left\{-\left(\frac{y_i}{\theta}\right)^{\lambda}\right\} \tag{2}$$

then, their join probability distribution is:

$$\begin{aligned} f\left(y_1,\ldots,y_n|\theta,\lambda\right) &= \prod_{i=1}^{n} f\left(y_i|\theta,\lambda\right) \\ &= \prod_{i=1}^{n} \frac{\lambda {y_i}^{\lambda-1}}{\theta^{\lambda}}\exp\left\{-\left(\frac{y_i}{\theta}\right)^{\lambda}\right\} \end{aligned} \tag{3}$$

The likelihood function is:

$$\mathscr{L}\left(\theta|y_1,\ldots,y_n,\lambda\right) \;=\; f\left(y_1,\ldots,y_n|\theta,\lambda\right) \tag{4}$$

The log-likelihood function is:

$$\begin{aligned} \ell\left(\theta|y_1,\ldots,y_n,\lambda\right) &= \log\left(\mathscr{L}\left(\theta|y_1,\ldots,y_n,\lambda\right)\right) \\ &= \log\prod_{i=1}^{n} \frac{\lambda {y_i}^{\lambda-1}}{\theta^{\lambda}}\exp\left\{-\left(\frac{y_i}{\theta}\right)^{\lambda}\right\} \\ &= \sum_{i=1}^{n}\log\left(\frac{\lambda {y_i}^{\lambda-1}}{\theta^{\lambda}}\exp\left\{-\left(\frac{y_i}{\theta}\right)^{\lambda}\right\}\right) \\ \ell\left(\theta|y_1,\ldots,y_n,\lambda\right) &= \sum_{i=1}^{n}\left(\log\,\lambda + (\lambda-1)\log\,y_i - \lambda\,\log\,\theta - \left(\frac{y_i}{\theta}\right)^{\lambda}\right) \end{aligned} \tag{5}$$

The maximum likelihood estimator (MLE), denoted by $\hat{\theta}$, is such that:

$$\hat{\theta} \;=\; \operatorname{argmax}\left\{\ell\left(\theta|y_1,\ldots,y_n,\lambda\right)\right\} \tag{6}$$

[1] *To refer this document and the implemented code, please cite as*: Alvarado, M. (2020, August 8). Maximum Likelihood Estimation: Numerical Solution for Weibull Distribution (Version v1.0.0). Zenodo. http://doi.org/10.5281/zenodo.3981644. Also at GitHub: https://github.com/miguel-alvarado-stats/MLE_Weibull.

To maximize the log-likelihood function (5), requires the derivative with respect to θ. The resulted function is called the Score function, denoted by $U(\theta|y_1,\ldots,y_n,\lambda)$.

$$\begin{aligned} U(\theta|y_1,\ldots,y_n,\lambda) &= \frac{\partial \ell(\theta|y_1,\ldots,y_n,\lambda)}{\partial\theta} \\ &= \sum_{i=1}^{n}\left(-\frac{\lambda}{\theta}+\frac{\lambda {y_i}^{\lambda}}{\theta^{\lambda+1}}\right) \\ &= -\frac{\lambda n}{\theta}+\frac{\lambda}{\theta^{\lambda+1}}\sum_{i=1}^{n}{y_i}^{\lambda} \end{aligned} \tag{7}$$

Then, the MLE $\hat{\theta}$, is the solution of:

$$U\left(\theta=\hat{\theta}|y_1,\ldots,y_n,\lambda\right) = 0 \tag{8}$$

Maximum Likelihood Estimation: Newton-Raphson Method

Just for notation, let write equation (8) as:

$$U(\theta^*) = 0 \tag{9}$$

The equation (9), generally, is a nonlinear equation, that can be aproximate by Taylor Series:

$$U(\theta^*) \approx U\left(\theta^{(t)}\right)+U'\left(\theta^{(t)}\right)\left(\theta^*-\theta^{(t)}\right) \tag{10}$$

Then, using (10) into (9), and solving for θ^*:

$$\begin{aligned} U\left(\theta^{(t)}\right)+U'\left(\theta^{(t)}\right)\left(\theta^*-\theta^{(t)}\right) &= 0 \\ \theta^* &= \theta^{(t)}-\frac{U\left(\theta^{(t)}\right)}{U'\left(\theta^{(t)}\right)} \end{aligned} \tag{11}$$

where U' is the derivative of the Score function (7) respect of θ.

$$\begin{aligned} U'(\theta|y_1,\ldots,y_n,\lambda) &= \frac{\partial U(\theta|y_1,\ldots,y_n,\lambda)}{\partial\theta} \\ &= \frac{\lambda n}{\theta^2}-\frac{\lambda(\lambda+1)}{\theta^{\lambda+2}}\sum_{i=1}^{n}{y_i}^{\lambda} \end{aligned} \tag{12}$$

Then, with the Newton-Raphson method: starting with an initial guess $\theta^{(1)}$ successive approximations are obtained using (13), until the iterative process converges.

$$\theta^{(t+1)} = \theta^{(t)}-\frac{U\left(\theta^{(t)}\right)}{U'\left(\theta^{(t)}\right)} \tag{13}$$

In order to example the use of Newton-Rapshon method, we use the data of lifetimes (times to failure in hours) of Kevlar epoxy strand pressure vessels at 70% stress level[2].

[2]This data was taken from Dobson, A. Barnett, A. (2018) An Introduction to Generalized Linear Models. Texts in Statistical Science. Chapman Hall/CRC.

```r
# load data
Y <- as.matrix(read.delim(file = "data.txt", header = FALSE))
```

We load the code developed into our R function `MLE_NR_Weibull`, stored in the R object with the same name.

```r
# load the function to solve by Newton-Raphson
load("MLE_NR_Weibull.RData")
```

The function `MLE_NR_Weibull` takes the sample mean $\theta = \bar{y}$ as a first guess for the iterative process and, besides some other default parameters that can be modified such as $\lambda = 2$, only needs the data vector Y.

```r
# MLE by Newton-Raphson (NR) for Weibull distribution
MLE_NR_Weibull(Y)

##      ML Estimator  Likelihood              Log-Likelihood
## [1,] "8805.693878" "3.51556268514669e-210" "-482.285669935176"
## [2,] "9633.777408" "1.37775513478331e-209" "-480.919828975047"
## [3,] "9875.898292" "1.47709267871969e-209" "-480.850208686227"
## [4,] "9892.110004" "1.47748579671734e-209" "-480.849942578555"
## [5,] "9892.176818" "1.47748580332318e-209" "-480.849942574083"
## [6,] "9892.176819" "1.47748580332318e-209" "-480.849942574084"
```

Then, the MLE by Newton-Raphson method: $\hat{\theta} =$ 9892.176819.

Maximum Likelihood Estimation: Fisher-Scoring Method

A distribution belongs to the exponential family if it can be written in the form:

$$f\left(y|\theta\right) \quad = \quad \text{exp}\left\{\frac{a\left(y\right)b\left(\theta\right)-c\left(\theta\right)}{\phi}+d\left(y,\phi\right)\right\} \tag{14}$$

Since (1) can be written as a member of exponential family as in (14):

$$\begin{aligned} f\left(y|\theta,\lambda\right) &= \text{exp}\left\{\text{log}\left(\frac{\lambda y^{\lambda-1}}{\theta^{\lambda}}\text{exp}\left\{-\left(\frac{y}{\theta}\right)^{\lambda}\right\}\right)\right\} \\ &= \text{exp}\left\{y^{\lambda}\left(-\theta^{-\lambda}\right)-\left(\lambda\ \text{log}\ \theta-\ \text{log}\ \lambda\right)+\left(\lambda-1\right)\text{log}\ y\right\} \end{aligned} \tag{15}$$

where, $a\left(y\right)=y^{\lambda}$, $b\left(\theta\right)=-\theta^{-\lambda}$, $c\left(\theta\right)=\lambda\ \text{log}\ \theta-\ \text{log}\ \lambda$, $\phi=1$, and $d\left(y,\phi\right)=\left(\lambda-1\right)\text{log}\ y$.

Then, since the Weibull distribution belongs to the exponential family, it can be show that the variance of U, denoted by $\mathcal{J}$, is:

$$\mathcal{J} \quad = \quad \text{Var}\left\{U\right\}=-\text{E}\left\{U'\right\} \tag{16}$$

where:

$$\text{E}\left\{U'\right\} \quad = \quad -\frac{1}{\phi}\left(b''\left(\theta\right)\frac{c'\left(\theta\right)}{b'\left(\theta\right)}-c''\left(\theta\right)\right) \tag{17}$$

For MLE, it is common to approximate U' by its expected value $\mathsf{E}\{U'\}$. In this case:

$$\begin{aligned}
\mathcal{J} &= -\mathsf{E}\{U'\} \\
&= \mathsf{E}\{-U'\} \\
&= \mathsf{E}\left\{-\sum_{i=1}^{n} U_i'\right\} \\
&= \sum_{i=1}^{n} -\mathsf{E}\{U_i'\} \\
&= \sum_{i=1}^{n} -\frac{1}{\phi}\left(b''(\theta)\frac{c'(\theta)}{b'(\theta)} - c''(\theta)\right) \qquad (18)
\end{aligned}$$

where, using (1), the previous derivaties:

$$\begin{aligned}
b'(\theta) &= \lambda\theta^{-(\lambda+1)} \\
b''(\theta) &= -\lambda(\lambda+1)\theta^{-(\lambda+2)} \\
c'(\theta) &= \lambda\theta^{-1} \\
c''(\theta) &= -\lambda\theta^{-2} \\
\frac{1}{\phi} &= 1
\end{aligned}$$

Then, replacing them into (18):

$$\begin{aligned}
\mathcal{J} &= \sum_{i=1}^{n} -\left(-\lambda(\lambda+1)\theta^{-(\lambda+2)}\frac{\lambda\theta^{-1}}{\lambda\theta^{-(\lambda+1)}} + \lambda\theta^{-2}\right) \\
&= \sum_{i=1}^{n} \frac{\lambda^2}{\theta^2} \\
&= n\left(\frac{\lambda}{\theta}\right)^2 \qquad (19)
\end{aligned}$$

Finally:

$$\begin{aligned}
\mathcal{J} &= -\mathsf{E}\{U'\} = n\left(\frac{\lambda}{\theta}\right)^2 \\
-\mathcal{J} &= \mathsf{E}\{U'\} = n\left(\frac{\lambda}{\theta}\right)^2 \qquad (20)
\end{aligned}$$

Then, approximating U' by its expected value $\mathsf{E}\{U'\}$, the equation (13) results into:

$$\theta^{(t+1)} = \theta^{(t)} + \frac{U\left(\theta^{(t)}\right)}{\mathcal{J}\left(\theta^{(t)}\right)} \qquad (21)$$

In order to example the use of Fisher-Scoring method, we use the same data used in the Newton-Rapshon method. We load the code developed into our R function `MLE_FS_Weibull`, stored in the R object with the same name.

```r
# load the function to solve by Fisher-Scoring
load("MLE_FS_Weibull.RData")
```

The function `MLE_FS_Weibull` takes the sample mean $\theta = \bar{y}$ as a first guess for the iterative process and besides some other default parameters that can be modified such as $\lambda = 2$, only needs the data vector Y.

```r
# MLE by Fisher-Scoring (FS) for Weibull distribution
MLE_FS_Weibull(Y)

##      ML Estimator  Likelihood              Log-Likelihood
## [1,] "8805.693878" "3.51556268514669e-210" "-482.285669935176"
## [2,] "9959.204199" "1.47092711455647e-209" "-480.854391543592"
## [3,] "9892.402373" "1.47748572804852e-209" "-480.849942625031"
## [4,] "9892.176822" "1.47748580332318e-209" "-480.849942574083"
## [5,] "9892.176819" "1.47748580332319e-209" "-480.849942574084"
```

Then, the MLE by Fisher-Scoring method: $\hat{\theta} = 9892.176819$.