CHAPTER-3

LEAST SQUARE METHOD

The Least squares technique was developed by Gauss (1795), Legendre (1805) and Adrain (1808) independently. It is published in the first decade of the nineteenth century. It is the most popular parameter estimation technique in statistics. This is due to several reasons. Because it is most common estimators can be casted within this framework. For example, the mean of a distribution is the value that minimizes the sum of squared deviations of the scores.

The method of least squares is a standard approach in regression analysis to the approximate solution of over determined systems, i.e., sets of equations in which there are more equations than unknowns. Least square means that the overall solution minimizes the sum of the squares of the errors made in the results of every single equation.

The most important application is in data fitting. The best fit in the least-squares sense minimizes the sum of squared residuals, a residual being the difference between an observed value and the fitted value provided by a model. The method of ordinary least squares assumes that there is constant variance in the errors (which is called homoscedasticity).

It is most popular method of estimation in Statistics. This technique is widely used in geophysical data analysis. It is not though for estimation of parameter of distribution. Hence it is commonly used as compared to other methods of estimation.

The least square estimation in fitting straight line, suppose a pair of observations \((x_i, y_i)\) where \(i = 1, 2, \ldots, n\). The model for straight line is,

\[
Y_i = \beta_1 + \beta_2 X_i
\]

In model (3.1) the \(X_i\) is independent variable and \(Y_i\) is dependent variable. The least square estimates for \(\alpha\) and \(\beta\) are those for which the predicted values of the curve minimize the sum of the squared deviations from the observations. By using least square method we find the values of \(\alpha\) and \(\beta\) that minimize the residual sum of squares.

\[
\varepsilon = \sum_{i=1}^{n} (Y_i - \beta_1 - \beta_2 X_i)^2
\]

Differentiating equation (3.2) w. r. t. \(\beta_1\) and \(\beta_2\) respectively, we get
\[ \frac{\partial \varepsilon}{\partial \beta_1} = 0 \]

\[ 2 \sum_{i=1}^{n} (Y_i - \beta_1 - \beta_2 X_i)(-1) = 0 \]

\[ -2 \sum_{i=1}^{n} (Y_i - \beta_1 - \beta_2 X_i) = 0 \]

\[ \sum_{i=1}^{n} (Y_i - \beta_1 - \beta_2 X_i) = 0 \]

\[ \sum_{i=1}^{n} Y_i - n \beta_1 - \beta \sum_{i=1}^{n} X_i = 0 \]

\[ -n \beta_1 = \beta_2 \sum_{i=1}^{n} X_i \cdot \sum_{i=1}^{n} Y_i \]

\[ n \beta_1 = \sum_{i=1}^{n} Y_i - \beta_2 \sum_{i=1}^{n} X_i \]

\[ \beta_1 = \frac{\sum_{i=1}^{n} Y_i - \beta_2 \sum_{i=1}^{n} X_i}{n} \]

\[ \beta_1 = \bar{y} - \beta_2 \bar{x} \quad (3.3) \]

\[ \frac{\partial \varepsilon}{\partial \beta_2} = 0 \]

\[ 2 \sum_{i=1}^{n} (Y_i - \beta_1 - \beta_2 X_i)(-X_i) = 0 \]

\[ \sum_{i=1}^{n} (X_i Y_i - X_i \beta_1 - \beta_2 X_i^2) = 0 \]

\[ \sum_{i=1}^{n} X_i Y_i - \beta_1 \sum_{i=1}^{n} X_i - \beta_2 \sum_{i=1}^{n} X_i^2 = 0 \]

\[ \beta_2 \sum_{i=1}^{n} X_i^2 = \sum_{i=1}^{n} X_i Y_i - \beta_1 \sum_{i=1}^{n} X_i \]

\[ \beta_2 = \frac{\sum_{i=1}^{n} X_i Y_i - \beta_1 \sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} X_i^2} \quad (3.4) \]
3.1 Least square method for Weibull distribution

Let \( x_1, x_2, \ldots, x_n \) be a random sample of size \( n \) from the Weibull distribution \( W(\beta, \alpha) \). The probability density function of Weibull distribution is given by

\[
f(x) = \frac{\beta}{\alpha} x^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^\beta}, \quad x \geq 0, \alpha, \beta > 0 \tag{3.5}\]

Where \( \beta \) is shape parameter and \( \alpha \) is scale parameter which indicates spread of Weibull distribution.

The cumulative distribution function of Weibull distribution is,

\[
F(x_i) = 1 - e^{-\left(\frac{x_i}{\alpha}\right)^\beta}, \quad x_i > 0 \tag{3.6}\]

The cumulative distribution function (3.6) will be transformed to a linear function.

From equation (3.6) we can write,

\[
1 - F(x_i) = e^{-\left(\frac{x_i}{\alpha}\right)^\beta} \quad (3.6)
\]

\[
ln(1 - F(x_i)) = -\left(\frac{x_i}{\alpha}\right)^\beta \quad (3.6)
\]

\[
ln(1 - F(x_i)) = -\left(\frac{x_i}{\alpha}\right)^\beta \quad (3.6)
\]

\[
ln(-ln(1 - F(x_i))) = ln\left(\frac{x_i}{\alpha}\right)^\beta \quad (3.6)
\]

\[
ln(-ln(1 - F(x_i))) = \beta ln\left(\frac{x_i}{\alpha}\right) \quad (3.7)
\]

Comparing equation (3.1) and (3.7), we get

\[
Y_i = ln(-ln(1 - F(x_i))) \quad (3.7)
\]

\[
\beta_i = -\beta ln \alpha \quad (3.7)
\]

\[
X_i = ln x_i
\]
and $\beta_2 = \beta$

Now consider $x_1 < x_2 < \ldots < x_n$ be the order statistics of $x_1, x_2, \ldots, x_n$. The mean rank function is used to estimate the value of C.D.F. $F(x)$.

$$\hat{F}(x_{(i)}) = \frac{i}{n+1}$$

Where $i$ denotes the $i^{th}$ smallest value of $x_1, x_2, \ldots, x_n$.

The regression parameter $\hat{\beta}_1$ and $\hat{\beta}_2$ are estimated from the function

$$\varepsilon = \sum_{i=1}^{n} (Y_i - \beta_1 - \beta_2 \ln x_{(i)})^2$$ \hspace{1cm} (3.8)

Differentiating equation (3.8) partially w.r.t. $\beta_1$ and equating to zero, we get

$$\frac{\partial \varepsilon}{\partial \beta_1} = 2 \sum_{i=1}^{n} (Y_i - \beta_1 - \beta_2 \ln x_{(i)}) (-1) = 0$$

$$-2 \sum_{i=1}^{n} (Y_i - \beta_1 - \beta_2 \ln x_{(i)}) = 0$$

$$\sum_{i=1}^{n} (Y_i - \beta_1 - \beta_2 \ln x_{(i)}) = 0$$

$$\sum_{i=1}^{n} Y_i - n\beta_1 - \beta_2 \sum_{i=1}^{n} \ln x_{(i)} = 0$$

$$n\beta_1 = \sum_{i=1}^{n} Y_i - \beta_2 \sum_{i=1}^{n} \ln x_{(i)}$$

$$\beta_1 = \frac{\sum_{i=1}^{n} Y_i - \beta_2 \sum_{i=1}^{n} \ln x_{(i)}}{n}$$ \hspace{1cm} (3.9)

$$\hat{\beta}_2 = \frac{\sum_{i=1}^{n} \ln (-\ln (1 - F(x_{(i)}))) - \beta \sum_{i=1}^{n} \ln x_{(i)}}{n}$$ \hspace{1cm} (3.10)

Differentiating equation (3.8) partially w.r.t. $\beta_2$ and equating to zero, we get

$$\frac{\partial \varepsilon}{\partial \beta_2} = 2 \sum_{i=1}^{n} (Y_i - \beta_1 - \beta_2 \ln x_{(i)}) (-\ln x_{(i)}) = 0$$

$$-2 \sum_{i=1}^{n} (Y_i - \beta_1 - \beta_2 \ln x_{(i)}) (\ln x_{(i)}) = 0$$
\[
\sum_{i=1}^{n}(Y_i - \beta_1 - \beta_2 \ln x(i)) \ln x(i) = 0
\]

\[
\sum_{i=1}^{n} (Y_i \ln x(i) - \beta_1 \ln x(i) - \beta_2 (\ln x(i))^2) = 0
\]

\[
\sum_{i=1}^{n} Y_i \ln x(i) - \beta_1 \sum_{i=1}^{n} \ln x(i) - \beta_2 \sum_{i=1}^{n} (\ln x(i))^2 = 0
\]

\[
\beta_2 \sum_{i=1}^{n} (\ln x(i))^2 = \sum_{i=1}^{n} Y_i \ln x(i) - \beta_1 \sum_{i=1}^{n} \ln x(i)
\]

\[
\beta_2 = \frac{\sum_{i=1}^{n} Y_i \ln x(i) - \beta_1 \sum_{i=1}^{n} \ln x(i)}{\sum_{i=1}^{n} (\ln x(i))^2}
\]

From the equation (3.9) we can write

\[
\beta_2 = \frac{n \sum_{i=1}^{n} Y_i \ln x(i) - \left( \frac{\sum_{i=1}^{n} Y_i - \beta_2 \sum_{i=1}^{n} \ln x(i)}{n} \right) \sum_{i=1}^{n} \ln x(i)}{n \sum_{i=1}^{n} (\ln x(i))^2}
\]

\[
\beta_2 = \frac{n \sum_{i=1}^{n} Y_i \ln x(i) - \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} \ln x(i) + \beta_2 \sum_{i=1}^{n} \ln x(i) \sum_{i=1}^{n} \ln x(i)}{n \sum_{i=1}^{n} (\ln x(i))^2}
\]

\[
\beta_2 = \frac{n \sum_{i=1}^{n} Y_i \ln x(i) - \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} \ln x(i)}{n \sum_{i=1}^{n} (\ln x(i))^2} + \frac{\beta_2 (\sum_{i=1}^{n} \ln x(i))^2}{n \sum_{i=1}^{n} (\ln x(i))^2}
\]

\[
\beta_2 = \frac{\beta_2 (\sum_{i=1}^{n} \ln x(i))^2}{n \sum_{i=1}^{n} (\ln x(i))^2} = \frac{n \sum_{i=1}^{n} Y_i \ln x(i) - \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} \ln x(i)}{n \sum_{i=1}^{n} (\ln x(i))^2}
\]

\[
\beta_2 (1 - \frac{(\sum_{i=1}^{n} \ln x(i))^2}{n \sum_{i=1}^{n} (\ln x(i))^2}) = \frac{n \sum_{i=1}^{n} Y_i \ln x(i) - \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} \ln x(i)}{n \sum_{i=1}^{n} (\ln x(i))^2}
\]
\[ \beta_2 \left( \frac{n \sum_{i=1}^{n} (\ln x(i))^2 - (\sum_{i=1}^{n} \ln x(i))^2}{n \sum_{i=1}^{n} (\ln x(i))^2} \right) = \frac{n \sum_{i=1}^{n} Y_i \ln x(i) - \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} \ln x(i)}{n \sum_{i=1}^{n} (\ln x(i))^2} \]

\[ \beta_2 \left( n \sum_{i=1}^{n} (\ln x(i))^2 - (\sum_{i=1}^{n} \ln x(i))^2 \right) = n \sum_{i=1}^{n} Y_i \ln x(i) - \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} \ln x(i) \]

\[ \beta_2 = \frac{n \sum_{i=1}^{n} Y_i \ln x(i) - \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} \ln x(i)}{\left( n \sum_{i=1}^{n} (\ln x(i))^2 - (\sum_{i=1}^{n} \ln x(i))^2 \right)} \] (3.11)

Now substituting the value of \( Y_i \) in equation (3.11), we get

\[ \hat{\beta}_2 = \frac{n \sum_{i=1}^{n} \ln(-\ln(1 - F(x(i)))) \ln x(i) - \sum_{i=1}^{n} \ln(-\ln(1 - F(x(i)))) \sum_{i=1}^{n} \ln x(i)}{\left( n \sum_{i=1}^{n} (\ln x(i))^2 - (\sum_{i=1}^{n} \ln x(i))^2 \right)} \] (3.12)

Thus, the estimates \( \hat{\beta}_1 \) and \( \hat{\beta}_2 \) of the parameters \( \beta_1 \) and \( \beta_2 \) are given by equations (3.10) and (3.12).

Therefore, estimates \( \hat{\beta} \) and \( \hat{\alpha} \) of the parameter \( \beta \) and \( \alpha \) are given as

\[ \hat{\beta} = \hat{\beta}_2 \]

and

\[ \hat{\beta}_1 = -\beta \ln \alpha \]

\[ -\frac{\hat{\beta}_1}{\beta} = \ln \alpha \]

\[ \hat{\alpha} = \exp(-\frac{\sum_{i=1}^{n} \ln(-\ln(1 - F(x(i)))) - \hat{\beta} \sum_{i=1}^{n} \ln x(i)}{n\hat{\beta}}) \]

Thus, the estimators given by equation (3.13) are the least square estimators of \( \beta \) and \( \alpha \) for Weibull distribution.

**3.2 Least square method for Frechet distribution**
Let $x_1, x_2, \ldots, x_n$ be a random sample of size $n$ from the Frechet distribution. The probability density function of Frechet distribution is given by,

$$f(x) = \frac{\alpha}{\beta} x^{\alpha+1} \exp\left(-\frac{x}{\beta}\right), \ x > 0, \ \alpha, \beta > 0$$  \hspace{1cm} (3.14)

The cumulative distribution function of Frechet distribution is,

$$F(x) = \exp\left(-\frac{x}{\beta}\right)$$  \hspace{1cm} (3.15)

The cumulative distribution function (3.15) will be transformed to a linear function. From equation (3.15) we can write,

$$F(x_i) = \exp\left(-\frac{x_i}{\beta}\right)$$

$$\ln(F(x_i)) = -\left(\frac{x_i}{\beta}\right)^\alpha$$

$$-\ln(F(x_i)) = \left(\frac{x_i}{\beta}\right)^\alpha$$

$$\ln(-\ln(F(x_i))) = \ln \left(\frac{x_i}{\beta}\right)^\alpha$$

$$\ln(-\ln(F(x_i))) = \alpha \ln \left(\frac{x_i}{\beta}\right)$$

$$\ln(-\ln(F(x_i))) = \alpha \ln \beta - \alpha \ln x_i$$  \hspace{1cm} (3.16)

Comparing equation (3.1) and (3.16), we get

$$Y_i = \ln(-\ln(F(x_i)))$$

$$\beta_1 = \alpha \ln \beta$$

$$X_i = \ln x_i$$

and $\beta_2 = -\alpha$

Now consider $x_{(1)} < x_{(2)} < \ldots < x_{(n)}$ be the order statistics of $x_1, x_2, \ldots, x_n$. The mean rank function is used to estimate the value of C.D.F. $F(x)$.  


\[
\hat{F}(x_{(i)}) = \frac{i}{n+1}
\]

Where \( i \) denotes the \( i^{th} \) smallest value of \( x_{(1)}, x_{(2)}, \ldots, x_{(n)} \).

The regression parameter \( \hat{\beta}_1 \) and \( \hat{\beta}_2 \) are estimated from the function

\[
\varepsilon = \sum_{i=1}^{n} (Y_i - \beta_1 - \beta_2 \ln x_{(i)})^2
\]  

(3.17)

Differentiating equation (3.17) partially w. r. t. \( \beta_1 \) and equating to zero, we get

\[
\frac{\partial \varepsilon}{\partial \beta_1} = 2 \sum_{i=1}^{n} (Y_i - \beta_1 - \beta_2 \ln x_{(i)}) (-1) = 0
\]

\[-2 \sum_{i=1}^{n} (Y_i - \beta_1 - \beta_2 \ln x_{(i)}) = 0\]

\[\sum_{i=1}^{n} (Y_i - \beta_1 - \beta_2 \ln x_{(i)}) = 0\]

\[n \beta_1 = \sum_{i=1}^{n} Y_i - \beta_2 \sum_{i=1}^{n} \ln x_{(i)}\]

\[
\hat{\beta}_1 = \frac{\sum_{i=1}^{n} Y_i - \beta_2 \sum_{i=1}^{n} \ln x_{(i)}}{n}
\]  

(3.18)

Differentiating equation (3.17) partially w. r. t. \( \beta_2 \) and equating to zero, we get

\[
\frac{\partial \varepsilon}{\partial \beta_2} = 2 \sum_{i=1}^{n} (Y_i - \beta_1 - \beta_2 \ln x_{(i)}) (-\ln x_{(i)}) = 0
\]

\[-2 \sum_{i=1}^{n} (Y_i - \beta_1 - \beta_2 \ln x_{(i)}) (\ln x_{(i)}) = 0\]

\[\sum_{i=1}^{n} (Y_i - \beta_1 - \beta_2 \ln x_{(i)}) (\ln x_{(i)}) = 0\]

\[\sum_{i=1}^{n} (Y_i \ln x_{(i)} - \beta_1 \ln x_{(i)} - \beta_2 (\ln x_{(i)})^2) = 0\]

\[\sum_{i=1}^{n} Y_i \ln x_{(i)} - \beta_1 \sum_{i=1}^{n} \ln x_{(i)} - \beta_2 \sum_{i=1}^{n} (\ln x_{(i)})^2 = 0\]
$$\beta_2 \sum_{i=1}^{n} (\ln x_{(i)})^2 = \sum_{i=1}^{n} Y_i \ln x_{(i)} - \beta_1 \sum_{i=1}^{n} \ln x_{(i)}$$

$$\beta_2 = \frac{\sum_{i=1}^{n} Y_i \ln x_{(i)} - \beta_1 \sum_{i=1}^{n} \ln x_{(i)}}{\sum_{i=1}^{n} (\ln x_{(i)})^2}$$

From the equation (3.18) we can write

$$\beta_2 = \frac{\sum_{i=1}^{n} Y_i \ln x_{(i)} - \left(\frac{\sum_{i=1}^{n} Y_i - \beta_2 \sum_{i=1}^{n} \ln x_{(i)}}{n}\right) \sum_{i=1}^{n} \ln x_{(i)}}{\sum_{i=1}^{n} (\ln x_{(i)})^2}$$

$$\beta_2 = \frac{n \sum_{i=1}^{n} Y_i \ln x_{(i)} - \left(\sum_{i=1}^{n} Y_i - \beta_2 \sum_{i=1}^{n} \ln x_{(i)}\right) \sum_{i=1}^{n} \ln x_{(i)}}{n \sum_{i=1}^{n} (\ln x_{(i)})^2}$$

$$\beta_2 = \frac{n \sum_{i=1}^{n} Y_i \ln x_{(i)} - \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} \ln x_{(i)} + \beta_2 \sum_{i=1}^{n} \ln x_{(i)} \sum_{i=1}^{n} \ln x_{(i)}}{n \sum_{i=1}^{n} (\ln x_{(i)})^2}$$

$$\beta_2 = \frac{n \sum_{i=1}^{n} Y_i \ln x_{(i)} - \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} \ln x_{(i)} + \beta_2 \left(\sum_{i=1}^{n} \ln x_{(i)}\right)^2}{n \sum_{i=1}^{n} (\ln x_{(i)})^2}$$

$$\beta_2 \left(1 - \frac{\left(\sum_{i=1}^{n} \ln x_{(i)}\right)^2}{n \sum_{i=1}^{n} (\ln x_{(i)})^2}\right) = \frac{n \sum_{i=1}^{n} Y_i \ln x_{(i)} - \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} \ln x_{(i)}}{n \sum_{i=1}^{n} (\ln x_{(i)})^2}$$

$$\beta_2 \left(\frac{n \sum_{i=1}^{n} (\ln x_{(i)})^2 - \left(\sum_{i=1}^{n} \ln x_{(i)}\right)^2}{n \sum_{i=1}^{n} (\ln x_{(i)})^2}\right) = \frac{n \sum_{i=1}^{n} Y_i \ln x_{(i)} - \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} \ln x_{(i)}}{n \sum_{i=1}^{n} (\ln x_{(i)})^2}$$

$$\beta_2 \left(\frac{n \sum_{i=1}^{n} (\ln x_{(i)})^2 - \left(\sum_{i=1}^{n} \ln x_{(i)}\right)^2}{n \sum_{i=1}^{n} (\ln x_{(i)})^2}\right) = \frac{n \sum_{i=1}^{n} Y_i \ln x_{(i)} - \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} \ln x_{(i)}}{n \sum_{i=1}^{n} (\ln x_{(i)})^2}$$
\[
\hat{\beta}_2 = \frac{n \sum_{i=1}^{n} Y_i \ln x(i) - \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} \ln x(i)}{\left(n \sum_{i=1}^{n} (\ln x(i))^2 - \left(\sum_{i=1}^{n} \ln x(i)\right)^{2}\right)} \tag{3.20}
\]

Now substituting the value of \(Y_i\) in equation (3.20), we get
\[
\hat{\beta}_2 = \frac{n \sum_{i=1}^{n} \ln(-\ln(F(x(i)))) \ln x(i) - \sum_{i=1}^{n} \ln(-\ln(F(x(i)))) \sum_{i=1}^{n} \ln x(i)}{\left(n \sum_{i=1}^{n} (\ln x(i))^2 - \left(\sum_{i=1}^{n} \ln x(i)\right)^{2}\right)}
\]
\[
(3.21)
\]

Thus, the estimates \(\hat{\beta}_1\) and \(\hat{\beta}_2\) of the parameters \(\beta_1\) and \(\beta_2\) are given by equations (3.19) and (3.21).

Therefore, estimates \(\hat{\beta}\) and \(\hat{\alpha}\) of the parameter \(\beta\) and \(\alpha\) are given as
\[
\hat{\alpha} = -\hat{\beta}_2
\]
\[
(3.22)
\]
and
\[
\hat{\beta}_1 = \alpha \ln \beta
\]

\[
\begin{align*}
\hat{\beta} &= \exp\left(\frac{\sum_{i=1}^{n} \ln(-\ln(F(x(i)))) + \hat{\alpha} \sum_{i=1}^{n} \ln x(i)}{n \hat{\alpha}}\right) \\
\hat{\beta} &= \exp\left(\frac{\sum_{i=1}^{n} \ln(-\ln(F(x(i)))) + \hat{\alpha} \sum_{i=1}^{n} \ln x(i)}{n \hat{\alpha}}\right)
\end{align*}
\]
\[
(3.23)
\]

Thus, the estimators given by equation (3.23) are the least square estimators of \(\alpha\) and \(\beta\) for Frechet distribution.