Methods of Finding Estimators

There are various methods of finding estimators for the parameters, some of which are listed below.

- Method of Maximum Likelihood
- Method of Moments
- Method of Least Squares
- Method of Minimum Chi square Estimation

We will discuss the method of moments and method of maximum likelihood estimation in detail.

Method of Maximum Likelihood Estimation: Let $X_1, \cdots, X_n$ be a random sample having joint probability density function $f_\theta(x_1, \ldots, x_n)$, $\theta \in \Theta$. The function $f_\theta(x_1, \ldots, x_n)$ may be regarded as a function of θ for given values $(x_1, \ldots, x_n)$. When regarded as a function of θ, the expression $f_\theta(x_1, \ldots, x_n)$ is referred to as the likelihood function of θ, $L(\theta|x_1, \ldots, x_n)$ and expresses the probability that the value of the random variable θ is θ for given values of observations $x_1, \ldots, x_n$. The maximum likelihood estimate (MLE) of θ is that value of θ, within the admissible range of values of θ, which makes the likelihood function a maximum, i.e. the MLE of θ is the number $\hat{\theta}$, if it exists, such that $L(\hat{\theta}|x_1, \ldots, x_n) > L(\theta'|x_1, \ldots, x_n)$ whatever be θ', any other value in Θ.
Ordinarily the parameter θ may be regarded as continuous and in this case the determination of MLE becomes simple. Assuming

1. the likelihood is a positive differentiable function of θ.
2. the maximum of the likelihood does not occur on the boundary of the interval in $\mathbb{R}$ of all admissible values of θ.

The stationary values of the likelihood function within the interval are given by the roots of the equation

$$\frac{\partial L(\theta|x_1, \ldots, x_n)}{\partial \theta} = 0.$$

A sufficient condition that any of these values, say, $\hat{\theta}$ be a real maximum is

$$\left.\frac{\partial^2 L(\theta|x_1, \ldots, x_n)}{\partial \theta^2}\right|_{\theta=\hat{\theta}} < 0.$$

Since $\log L$ attains its maximum value for the same value of θ as L it is usual to maximize $\log L$ in lieu of L. Therefore, we shall seek solution of

$$\frac{\partial \log L(\theta|x_1, \ldots, x_n)}{\partial \theta} = 0. \tag{1}$$

subject to the condition

$$\frac{\partial^2 \log L(\theta|x_1, \ldots, x_n)}{\partial \theta^2} < 0. \tag{2}$$

(2) is generally referred to as likelihood equation. If the observations are iid

$$f_\theta(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_\theta(x_i) \tag{3}$$

where $f_\theta(x)$ is the common pdf and here $\log L(\theta) = \sum_{i=1}^{n} \log f_\theta(x)$.

Remark 1. *1. If there are more than one solution satisfying (1) and (2), the maximum of these solutions is to be taken.*

2. We shall ignore any solution which is independent of the observations, i.e., any constant solution.

3. The method holds even if all the variables $X_1, \dots, X_n$ are discrete and in this case the density function is to be replaced by probability mass function (pmf).

4. If assumptions 1 and 2 do not hold, the MLE cannot be obtained by solving the likelihood equation.

If more than one parameters are involved, i.e., a sample has the pdf $f_{\theta}(x_1, \dots, x_n)$ where $\underline{\theta} = (\theta_1, \dots, \theta_n) \in \Theta \subset \mathbb{R}^k$. In this case, the MLEs are the numbers $\hat{\theta}_1, \dots, \hat{\theta}_k$, if such a set exists, which maximises f as a function of $\underline{\theta}$. If the likelihood function does not have a maxima on the boundary of set Θ, the maximum of the likelihood function is obtained by the solution of

$$\frac{\partial L(\theta|x_1, \dots, x_n)}{\partial \theta_i} = 0, \quad i = 1, \dots, k$$

subject to the condition that the matrix

$$\left(\frac{\partial^2 \log L(\theta|x_1, \dots, x_n)}{\partial \theta_i \theta_j} \right)_{i,j=1,\dots,r} \Bigg|_{\underline{\theta}=\hat{\underline{\theta}}} \tag{4}$$

is negative definite.

Example 2. *Let $X_1, \cdots, X_n$ follow Poisson distribution with parameter λ; $\lambda > 0$. Find the MLE for λ.*

Solution: Let $\underline{x} = (x_1, \cdots, x_n)$ be a realization of a random sample. Then the likelihood function is given by

$$L_{\lambda}(\underline{x}) = \prod_{i=1}^{n} f(x_i, \lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda} \lambda^{x_i}}{x_i!} = \frac{e^{-n\lambda} \lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}.$$

Therefore, the log likelihood function is given by

$$\log L_{\lambda}(\underline{x}) = l(\lambda) = -n\lambda + \sum_{i=1}^{n} x_i \log \lambda - \log \left(\prod_{i=1}^{n} x_i! \right).$$

The likelihood equation is

$$\frac{\partial l}{\partial \lambda} = -n + \frac{1}{\lambda} \sum_{i=1}^{n} x_i = 0.$$

Now, $\frac{\sum_{i=1}^{n} x_i - n\lambda}{\lambda} > 0$ if $\lambda < \bar{x}$ and $\frac{\sum_{i=1}^{n} x_i - n\lambda}{\lambda} < 0$ if $\lambda > \bar{x}$
Hence, the MLE for λ is $\hat{\lambda} = \bar{x}$.

Example 3. *Let X_1, X_2 be a random sample from a population*

$$f_{\theta}(x) = \frac{2}{\theta^2}, \quad 0 < x < \theta.$$

Find the MLE of θ.

Solution: The likelihood function is given by

$$L_\theta(\underline{x}) = \frac{4}{\theta^4}(\theta - x_1)(\theta - x^2)$$

The likelihood equation is

$$\frac{\partial \log L}{\partial \theta} = -\frac{4}{\theta} + \frac{1}{\theta - x_1} + \frac{1}{\theta - x_2} = 0.$$

$\Rightarrow$

$$\hat{\theta} = \frac{3(x_1 + x_2) + \sqrt{9(x_1 - x_2)^2 + 4x_1 x_2}}{4}.$$

Remark 4. *1. The MLE is unique. (Prove yourself).*

*2. **Invariance Property:** If $\hat{\theta}$ is the MLEof θ, then $g(\hat{\theta})$ is the MLE of $g(\theta)$ provided $g(\theta)$ is some single valued function of θ.*

Exercise 5. *Let $X_1, \ldots, X_n$ be random sample with following pdf/pmf. Find the MLE(s) of the parameter(s).*

1. $N(\theta, \theta^2)$, $\theta \in (0, \infty)$.
2. $f_{\alpha,\beta}(x) = \frac{\alpha\beta^\alpha}{x^{\alpha+1}}$, $\alpha > 0$, $x \geq \beta > 0$.
3. $P(X_i = 0) = 1 - p$, $P(X_i = 1) = p$ *where* $p \in \left[\frac{1}{4}, \frac{3}{4}\right]$.

Method of Moments: Let $X_1, \cdots, X_n$ be a random sample from a population with probability distribution $P_{\underline{\theta}}; \theta \in \Theta; \underline{\theta} = (\theta_1, \cdots, \theta_k)$.
Consider first k non central moments,

$$\mu_1^{'} = E(X_1) = g_1(\underline{\theta})$$

$$\mu_2^{'} = E(X_1^2) = g_2(\underline{\theta})$$

$$\vdots$$

$$\mu_k^{'} = E(X_1^k) = g_k(\underline{\theta}).$$

Assume that the above system of equations have solution as

$$\theta_1 = h_1(\mu_1^{'}, \cdots, \mu_k^{'})$$

$$\theta_2 = h_2(\mu_1^{'}, \cdots, \mu_k^{'})$$

$$\vdots$$

$$\theta_k = h_k(\mu_1^{'}, \cdots, \mu_k^{'}).$$

Now, define the first k non central sample moments as

$$\alpha_1 = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$\alpha_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2.$$

$$\vdots$$

$$\alpha_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k.$$

In the method of moments, we estimate k^{th} population moment by k^{th} sample moment, i.e.,

$$\hat{\mu}_{j'} = \alpha_j \ ; \quad j = 1, \cdots, k.$$

Thus, the method of moments estimators of $\theta_1, \cdots, \theta_k$ are defined as

$$\hat{\theta}_1 = h_1(\alpha_1, \cdots, \alpha_k)$$

$$\hat{\theta}_2 = h_2(\alpha_1, \cdots, \alpha_k)$$

$$\vdots$$

$$\hat{\theta}_k = h_k(\alpha_1, \cdots, \alpha_k).$$

Example 6. *Let $X_1, \cdots, X_n$ follow $N(\mu, \sigma^2)$; μ and σ^2 are unknown. Find the method of moments estimators μ and σ^2 .*

Solution: We know, for normal distribution, $\mu_1^{'} = \mu$ and $\mu_2{'} = \mu^2 + \sigma^2$. Therefore, we have

$$\mu = \mu_1^{'}$$

and

$$\sigma^2 = \mu_2^{'} - \mu_1^{'2}.$$

Now, equating the population moments to sample moments, we get

$$\hat{\mu}_{MME} = \bar{X},$$

and

$$\hat{\sigma}^2_{MME} = \alpha_2 - \alpha_1^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2.$$

Exercise 7. *1. Let $X_1, \cdots, X_k$ follow binomial distribution with parameters n and p. Find the moment estimators of p, when n in known.*

2. Let $X_1, \cdots, X_n$ be a random sample from Poisson distribution with parameter λ, find the moment estimator of λ.

Remark 8. *1. The method moment estimators need not be unbiased always.*

2. If the functions $g_i^{\prime}s$ are continuous and one-one then the functions $h_i^{\prime}s$ are also continuous and then the method of moment estimators will be consistent.