Chapter- 6.4

Maximum Likelihood Estimation

Let, X be $b(1, p)$, so that the pmf of X is,

$$f(x; p) = p^x (1-p)^{1-x}, \quad x = 0, 1, \quad 0 \leq p \leq 1.$$

where, $p \in \Omega = \{p : 0 \leq p \leq 1\}$

Given a random sample $X_1, X_2, \ldots, X_n$, the problem is to find an estimator $u(X_1, \ldots, X_n)$ such that $u(x_1, \ldots, x_n)$ is a good point estimate of p, where $x_1, \ldots, x_n$ are the observed values of the random sample.

Now,

$$P(X_1 = x_1, \ldots, X_n = x_n) = f(x_1; p) \ldots f(x_n; p)$$

$$= \prod_{i=1}^{n} p^{x_i} (1-p)^{1-x_i}$$

$$= p^{\Sigma x_i} (1-p)^{n - \Sigma x_i}$$

which is the joint pmf of $X_1, \ldots, X_n$ evaluated at the observed values.

Now, we have to find the value of p that maximizes the pmf.

Here, the likelihood function is,

$$L(p) = L(p;\, x_1, \ldots, x_n)$$
$$= f(x_1;p) \cdots f(x_n;p)$$
$$= p^{\sum x_i} (1-p)^{n-\sum x_i} \;;\; 0 \le p \le 1. \quad \text{— (1)}$$

If $\sum x_i = 0$, $L(p) = (1-p)^n$, which is maximized over $p \in [0,1]$ by taking $\hat{p} = 0$.

On the other hand, if $\sum x_i = n$, $L(p) = p^n$, which is maximized over $p \in [0,1]$ by taking $\hat{p} = 1$.

Again, if $\sum x_i$ is neither 0 nor n, $L(0) = L(1) = 0$ while $L(p) > 0$ for all $p \in (0,1)$;

To maximize $L(p)$ for $0 < p < 1$, taking "ln" on both sides of (1),

$$\ln[L(p)] = \sum x_i \ln p + (n - \sum x_i) \ln(1-p)$$

$$\therefore \frac{d}{dp}[\ln(L(p))] = \frac{\sum x_i}{p} + \frac{n - \sum x_i}{1-p} \cdot (-1) \quad \text{— (2)}$$

At, maximum value,

$$\frac{d}{dp}\left[\ln(L(p))\right] = 0$$

$$\Rightarrow \frac{\Sigma x_i}{p} - \frac{n - \Sigma x_i}{1-p} = 0$$

$$\Rightarrow (1-p)\Sigma x_i - p(n - \Sigma x_i) = 0$$

$$\Rightarrow \Sigma x_i - np = 0$$

$$\Rightarrow p = \frac{\Sigma x_i}{n} = \bar{x}$$

$\therefore$ (2) $\Rightarrow$

$$\frac{d^2}{dp^2}\left[\ln(L(p))\right] = -\frac{\Sigma x_i}{p^2} + \frac{n - \Sigma x_i}{(1-p)^2}(-1)$$

$$= -\frac{\Sigma x_i}{p^2} - \frac{n - \Sigma x_i}{(1-p)^2}$$

$\bar{x} = \frac{\Sigma x_i}{n}$ $\Rightarrow \Sigma x_i = n\bar{x}$

At, $p = \bar{x}$,

$$\frac{d^2}{dp^2}\left[\ln(L(p))\right] = -\frac{n\bar{x}}{\bar{x}^2} - \frac{n - n\bar{x}}{(1-\bar{x})^2}$$

$$= -\frac{n}{\bar{x}} - \frac{n(1-\bar{x})}{(1-\bar{x})^2}$$

$$= -\frac{n}{\bar{x}} - \frac{n}{1-\bar{x}}$$

$$= -\frac{n - n\bar{x} + n\bar{x}}{\bar{x}(1-\bar{x})}$$

$$= -\frac{n}{\bar{x}(1-\bar{x})} < 0$$

[since, $p = \bar{x}$ $\therefore 0 \le \bar{x} \le 1$]

$\therefore$ The maximum likelihood estimator for p is,

$$\hat{p} = \bar{X}.$$

Let, $X_1, X_2, \ldots, X_n$ be a random sample from a distribution that depends on one or more unknown parameters $\theta_1, \ldots, \theta_m$ with pmf or pdf that is denoted by $f(x; \theta_1, \ldots, \theta_m)$

Here, the likelihood function is,

$$L(\theta_1, \ldots, \theta_m) = f(x_1; \theta_1, \ldots \theta_m) \cdots \cdots$$
$$f(x_n; \theta_1, \ldots, \theta_m)$$

where, $\theta_1, \ldots, \theta_m \in \Omega$

Here, $[u_1(x_1, \ldots, x_n), \ldots, u_m(x_1, \ldots, x_n)]$ is that m-tuple in Ω that maximizes $L(\theta_1, \ldots, \theta_m)$.

Then, $\hat{\theta}_1 = u_1(X_1, \ldots, X_n)$

$\vdots$

$\hat{\theta}_m = u_m(X_1, \ldots, X_n)$

are maximum likelihood estimators of $\theta_1, \ldots, \theta_m$ respectively and the corresponding observed values of these statistics, namely $u_1(x_1, \ldots, x_n), \ldots, u_m(x_1, \ldots, x_n)$ are called maximum likelihood estimates.

For one unknown parameter, $L(\theta) = \prod_{i=1}^{n} f(x_i; \theta)$.

Example- 6.4.1 :- The likelihood function is,

$$L(\theta) = L(\theta; x_1, \cdots, x_n)$$
$$= f(x_1;\theta) \cdots f(x_n;\theta)$$
$$= \prod_{i=1}^{n} \frac{1}{\theta} e^{-\frac{x_i}{\theta}} = \frac{1}{\theta^n} e^{-\frac{\Sigma x_i}{\theta}} \quad \text{--- (1)}$$

By Taking logarithm on both sides of (1),

$$\ln(L(\theta)) = \ln\left(\frac{1}{\theta^n}\right) + \ln e^{-\frac{\Sigma x_i}{\theta}}$$
$$= -n \ln\theta - \frac{\Sigma x_i}{\theta}$$

$$\therefore \frac{d}{d\theta}[\ln(L(\theta))] = -\frac{n}{\theta} + \frac{\Sigma x_i}{\theta^2} \quad \text{--- (2)}$$

At maximum value, $\frac{d}{d\theta}[\ln(L(\theta))] = 0$

$$\Rightarrow -\frac{n}{\theta} + \frac{\Sigma x_i}{\theta^2} = 0$$
$$\Rightarrow -n\theta + \Sigma x_i = 0$$
$$\Rightarrow \theta = \frac{\Sigma x_i}{n} = \bar{x}$$

$$\therefore (2) \Rightarrow \frac{d^2}{d\theta^2}[\ln(L(\theta))] = \frac{n}{\theta^2} - 2\frac{\Sigma x_i}{\theta^3}$$
$$= \frac{n}{\theta^2} - 2 \cdot \frac{n\bar{x}}{\theta^3} \quad \Big| \; \bar{x} = \frac{\Sigma x_i}{n}$$
$$= \frac{n}{\theta^2} - 2 \cdot \frac{n\theta}{\theta^3}$$
$$= -\frac{n}{\theta^2} < 0$$

$\therefore$ (1) has a maximum value at $\theta = \bar{x}$,
so, the maximum likelihood estimator for θ is,

$$\hat{\theta} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Example: 6.4-2:- The likelihood function is,

$$L(P) = L(P;\, x_1, \dots, x_n)$$

$$= f(x_1; P) \cdots f(x_n; P)$$

$$= \prod_{i=1}^{n} (1-P)^{x_i - 1} \cdot P = (1-P)^{\sum x_i - n} \cdot P^n \quad \text{—(1)}, \; 0 \le P \le 1$$

By taking logarithm on both sides of (1),

$$\ln(L(P)) = (\sum x_i - n) \ln(1-P) + n \ln P$$

$$\therefore \frac{d}{dP}\left(\ln(L(P))\right) = \frac{\sum x_i - n}{1-P}(-1) + \frac{n}{P}$$

$$= \frac{n}{P} - \frac{\sum x_i - n}{1-P} \quad \text{—(2)}$$

At maximum value,

$$\frac{d}{dP}\left[\ln(L(P))\right] = 0$$

$$\Rightarrow \frac{n}{P} - \frac{\sum x_i - n}{1-P} = 0$$

$$\Rightarrow (1-P)n - P(\sum x_i - n) = 0$$

$$\Rightarrow n - P \sum x_i = 0$$

$$\Rightarrow P = \frac{n}{\sum x_i} = \frac{1}{\frac{\sum x_i}{n}} = \frac{1}{\bar{x}}$$

$\therefore$ ② $\Rightarrow$ $$\frac{d^2}{dp^2}\left[\ln(L(p))\right] = -\frac{n}{p^2} + \frac{\sum x_i - n}{(1-p)^2} \cdot (-1)$$

$$= -\frac{n}{p^2} \cancel{+} + \frac{n - \sum x_i}{(1-p)^2}$$

$$= -\frac{n}{\left(\frac{1}{\bar{x}}\right)^2} + \frac{n - n\bar{x}}{\left(1-\frac{1}{\bar{x}}\right)^2}$$

$$= -n\bar{x}^2 + \frac{n(1-\bar{x})}{\frac{-(1+\bar{x})(1-\bar{x})}{\bar{x}^2}}$$

$$= -n\bar{x}^2 - \frac{n\bar{x}^2}{1+\bar{x}} < 0, \text{ for all } \bar{x}$$

$\therefore$ ① has a maximum value at $p = \frac{1}{\bar{x}}$, ($\because 0 < p < 1 \Rightarrow 0 < \frac{1}{\bar{x}} < 1 \Rightarrow \bar{x} > 1$)

So, the maximum likelihood estimator for p is, $$\hat{p} = \frac{n}{\sum X_i} = \frac{1}{\bar{X}} = \frac{n}{\sum X_i}$$

Example:- 6.4-3:-

Here, $\theta_1 = \mu$ and $\theta_2 = \sigma^2$

The likelihood function is,

$$L(\theta_1, \theta_2) = f(x_1; \theta_1, \theta_2) \cdots f(x_n; \theta_1, \theta_2)$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\theta_2}} \exp\left[-\frac{(x_i - \theta_1)^2}{2\theta_2}\right]$$

$$= \left(\frac{1}{\sqrt{2\pi\theta_2}}\right)^n \exp\left[\frac{-\sum(x_i-\theta_1)^2}{2\theta_2}\right]$$

—— (1)

By taking logarithm on both sides of (1),

$$\ln[L(\theta_1,\theta_2)] = -\frac{n}{2}\ln(2\pi\theta_2) - \frac{\sum(x_i-\theta_1)^2}{2\theta_2}$$

—— (2)

(2) $\Rightarrow$ $$\frac{\partial}{\partial\theta_1}[\ln(L)] = -0 - \frac{\sum 2(x_i-\theta_1)(-1)}{2\theta_2}$$

$$= \frac{1}{\theta_2}\sum(x_i-\theta_1)$$

—— (3)

and, $$\frac{\partial}{\partial\theta_2}[\ln(L)] = -\frac{n}{2}\cdot\frac{1}{\theta_2} + \frac{1}{2\theta_2^2}\sum(x_i-\theta_1)^2$$

$$= -\frac{n}{2\theta_2} + \frac{1}{2\theta_2^2}\sum(x_i-\theta_1)^2$$

—— (4)

At maximum value,

$$\frac{\partial}{\partial\theta_1}(\ln(L)) = 0$$

$$\Rightarrow \frac{1}{\theta_2}\sum(x_i-\theta_1) = 0$$

$$\Rightarrow \sum x_i - n\theta_1 = 0$$

$$\Rightarrow \theta_1 = \frac{\sum x_i}{n} = \bar{x}$$

and $$\frac{\partial}{\partial \theta_2}\left[\ln(L)\right] = 0$$

$$\Rightarrow \frac{1}{2\theta_2}\left[-n + \frac{1}{\theta_2}\sum (x_i - \theta_1)^2\right] = 0$$

$$\Rightarrow \theta_2 = \frac{1}{n}\sum (x_i - \theta_1)^2$$

Here, $$D = \frac{\partial^2}{\partial \theta_1^2}\left[L(\theta_1, \theta_1)\right] \cdot \frac{\partial^2}{\partial \theta_2^2}\left[L(\theta_1, \theta_2)\right] - \left(\frac{\partial^2}{\partial \theta_1 \partial \theta_2}\left[L(\theta_1, \theta_2)\right]\right)^2 > 0$$

and $$\frac{\partial^2}{\partial \theta_1^2}\left[\ln(L(\theta_1, \theta_2))\right] = \frac{1}{\theta_2}\sum(-1) = -\frac{n}{\theta_2} < 0$$

$\therefore$ (1) has a maximum value at $\theta_1 = \bar{x}$ and $\theta_2 = \frac{1}{n}\sum (x_i - \theta_1)^2$.

Thus, the maximum likelihood estimators of $\mu = \theta_1$ and $\sigma^2 = \theta_2$ are,

$$\hat{\theta}_1 = \bar{X} \text{ and } \hat{\theta}_2 = \frac{1}{n}\sum (X_i - \bar{X})^2 = V(X).$$

Exercise :- 6.4 :-
1, 2, 3, 4, 5, 6, 7, 8 & 9.

Def^n :- 6.4.1 :-

If $E[u(X_1, \ldots, X_n)] = \theta$, then the statistic $u(X_1, \ldots, X_n)$ is called an unbiased estimator of θ. Otherwise, it is said to be biased.

Example 6.4-4 :-

The likelihood function is,

$$L(\theta) = L(\theta; x_1, x_2, x_3, x_4)$$
$$= f(x_1;\theta) f(x_2;\theta) f(x_3;\theta) f(x_4;\theta)$$
$$= \prod_{i=1}^{4} \frac{1}{\theta} = \left(\frac{1}{\theta}\right)^4 \quad ; \quad 0 < x_i \leq \theta$$

and $L(\theta) = 0$; when $x_i > \theta$, or $x_i \leq 0$.
To maximize $L(\theta)$, we must make θ as small as possible ; hence, the maximum likelihood estimator is,

$$\hat{\theta} = \max(X_i) = Y_4$$

Here, $F(x, \theta) = \int_0^x f(x;\theta)\,dx = \frac{x}{\theta}$, $0 < x \leq \theta$,

the pdf of Y_4 is,

$$g_4(y_4) = \frac{4!}{3!\,1!} \cdot \left(\frac{y_4}{\theta}\right)^3 \left(\frac{1}{\theta}\right) = 4\frac{y_4^3}{\theta^4}, \quad 0 < y_4 \leq \theta.$$

$$\therefore E(Y_4) = \int_0^\theta y_4 \cdot 4 \cdot \frac{y_4^3}{\theta^4}\,dy_4 = \frac{4}{5}\theta \neq \theta$$

$\therefore Y_4$ is a biased estimator of θ.
However, $\frac{5Y_4}{4}$ is unbiased.