PROBABILITY AND STATISTICS (UCS401)

Lecture-30

(Maximum Likelihood Estimation with illustrations)

Sampling Distributions and Theory of Estimation (Unit –V & VI)

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

Dr. Rajanish Rai

Assistant Professor

School of Mathematics

Thapar Institute of Engineering and Technology, Patiala

Rajanish Rai

Maximum Likelihood Estimation

Maximum Likelihood Estimation :

Let $X_1, X_2, X_3, \dots, X_n$ be a random (i.i.d) sample from a population with p.d.f./p.m.f. $f(x;\theta)$. The joint p.d.f/p.m.f of $X_1, X_2, X_3, \dots X_n$ is

$$f(\vec{x};\theta) = \prod_{i=1}^{n} f(x_i,\theta)$$

$$\vec{x} = (x_1, x_2, x_3, \dots, x_n)$$

The Likelihood function

$$L(\vec{x},\theta) = \prod_{i=1}^{n} f(x_i,\theta)$$

Note that : A statistics $T(X)$ is called the maximum likelihood estimator of θ if $L(\vec{x}, T(X)) \geq L(\vec{x},\theta)$,

$\forall\ \theta \in$ parameter space.

* **Procedure:** We have to maximize the Likelihood function $L(\vec{x}, \theta_i) = \prod_{i=1}^{n} f(x_i, \theta_i)$ for the value of θ_i (parameter) $(i = 1, 2, 3 \dots k)$

(i) $\frac{\delta}{\delta\theta_i}(L) = 0$ & $\frac{\delta^2}{\delta\theta_i^2}(L) < 0$

or Equivalently

$$\frac{\delta}{\delta\theta_i}(\log L) = 0 \quad \& \quad \frac{\delta^2}{\delta\theta_i^2}(\log L) < 0.$$

for $i = 1, 2, 3, \dots k$.

(i) **Estimation of Bernalli's distribution parameter:**

Let $X_1, X_2, \dots, X_n \sim B(1, p) \quad 0 < p < 1$

$$L(\vec{x}, p) = \prod_{i=1}^{n} p^{x_i} q^{1-x_i}$$

$$L = p^{\sum x_i} q^{n - \sum x_i}$$

$$\log L = \sum x_i \log p + \left(n - \sum x_i\right) \log q$$

$$\frac{\partial}{\partial p}(\log L) = \frac{\sum x_i}{p} - \left(\frac{n - \sum x_i}{1-p}\right) = 0$$

$$\Rightarrow (1-p)\sum x_i - p\left(n-\sum x_i\right) = 0$$

$$\Rightarrow (1-\cancel{p})\sum x_i - np + \cancel{p\sum x_i} = 0$$

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

$\Rightarrow \quad \bar{x}$ is MLE for p

Moreover,

$$\frac{\partial^2}{\partial p^2}(\log L) = \frac{-\sum x_i}{p^2} - \frac{\left(n-\sum x_i\right)}{(1-p)^2}$$

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

$$= -\frac{n}{p} - \frac{n}{(1-p)}$$

$$= -n\left(\frac{1-p+p}{p(1-p)}\right)$$

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

$\Rightarrow \quad \bar{X}$ is MLE for p.

(ii) Estimation of Binomial distribution parameters

Let $x_1, x_2, \dots x_N \sim B(n, p)$

find MLE of p when n is known

$\because \quad X \sim B(n, p)$

$$p(X=x) = {}^{n}C_{x}\, p^{x} q^{n-x}$$

$$L(p) = \prod_{i=1}^{N} {}^{n}C_{x_i}\, p^{x_i}\, q^{n-x_i}$$

$$= \prod_{i=1}^{N} {}^{n}C_{x_i}\; p^{\sum_{i=1}^{n} x_i}\; q^{nN - \sum x_i}$$

$$\log(L(p)) = \log\left(\prod_{i=1}^{N} {}^{n}C_{x_i}\right) + \sum x_i \log p + \left(nN - \sum x_i\right) \log q$$

$$\log(L(p)) = \log\left(\prod_{i=1}^{N} {}^{n}C_{x_i}\right) + \sum x_i \log p + \left(nN - \sum x_i\right) \log(1-p) \quad -(1)$$

For Maximum Likelihood estimation

$$\frac{\partial}{\partial p}\left(\log(L(p))\right) = 0$$

$$\Rightarrow \sum x_i \left(\frac{1}{p}\right) - \frac{(nN - \sum x_i)}{(1-p)} = 0$$

$$\Rightarrow (1-p)\sum x_i - p\left(nN - \sum x_i\right) = 0$$

$$\Rightarrow \sum x_i - \cancel{p\sum x_i} - pnN + \cancel{p\sum x_i} = 0$$

$$\Rightarrow p = \frac{\sum_{i=1}^{N} x_i}{nN} = \left(\frac{\overline{X}}{n}\right)$$

$$\Rightarrow \hat{p} = \frac{\overline{X}}{n} \text{ is MLE for } p.$$

Moreover,

$$\frac{\partial^2}{\partial p^2}(\log L)$$

$$= -\frac{\sum x_i}{p^2} - \frac{nN - \sum x_i}{(1-p)^2} \qquad \left| \begin{array}{l} \frac{\sum x_i}{N} = \overline{X} \\ \sum x_i = npN \end{array}\right.$$

$$= \frac{-nNp}{p^2} - \frac{(nN - nNp)}{(1-p)^2}$$

$$= \frac{-nN}{p} - \frac{nN}{(1-p)} = -nN\left(\frac{1-p+p}{p(1-p)}\right)$$

$$\left.\frac{\partial^2}{\partial p^2}(\log L)\right|_{p=\hat{p}} = \frac{-nN}{p(1-p)} < 0$$

$$\Rightarrow \hat{p} = \overline{X}/n \text{ is MLE for } p.$$

(iii) Estimation of Poisson distribution parameter :

Let $x_1, x_2, x_3, \ldots, x_n \sim P(\lambda)$

$\therefore$ For Poisson distribution, the p.m.f is given by

$$f(x_i, \lambda) = \frac{e^{-\lambda} \lambda^{x_i}}{x_i!} \; ; \quad x_i = 0, 1, 2, \ldots \infty$$

The Likelihood function is given by

$$L(\vec{x}, \lambda) = \prod_{i=1}^{n} f(x_i, \lambda)$$

$$L(\vec{x}, \lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda} \lambda^{x_i}}{x_i!}$$

$$L(\vec{x}, \lambda) = \frac{(e^{-\lambda})^n \lambda^{\sum x_i}}{\prod_{i=1}^{n} x_i!} \quad \text{--- (1)}$$

To find MLE, we have to find

$$\frac{\partial}{\partial \lambda}(\log L) = 0$$

$$(1) \Rightarrow \log L = -n\lambda + \sum_{i=1}^{n} x_i (\log \lambda) - \log\left(\prod_{i=1}^{n} x_i!\right)$$

$$= -n\lambda + (\log \lambda) \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \log(x_i!) \quad \text{--- (2)}$$

$$\frac{\partial}{\partial \lambda}(\log L) = -n + \frac{1}{\lambda} \sum_{i=1}^{n} x_i = 0$$

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

$$\therefore \boxed{\hat{\lambda} = \bar{x}} \quad \text{(MLE for poisson distribution)}$$

Moreover,

$$\left.\frac{\partial^2}{\partial \lambda^2}(\log L)\right|_{\lambda=\bar{x}} = -\frac{\sum_{i=1}^{n} x_i}{\lambda^2}$$

$$= \frac{-n\bar{x}}{(\bar{x})^2} = -\frac{n}{\bar{x}} < 0$$

Thus, $\hat{\lambda} = \bar{x}$ is the MLE for poisson distribution.

(iv) MLE of Geometric distribution parameter :

Let $x_1, x_2, x_3, \ldots, x_n \sim Geo(p)$

then find MLE of p.

We know that

$$X \sim G(p)$$

$$P(X=x) = \begin{cases} pq^{x-1} & x = 1, 2, 3 \ldots \infty \\ 0 & o/w \end{cases}$$

$$L(\bar{x}, p) = \prod_{i=1}^{n} pq^{x_i - 1}$$

$$L(\vec{x}, p) = p^n q^{\sum x_i - n} \quad \text{—— (i)}$$

$$\log L = n \log p + \left(\sum x_i - n\right) \log q$$

$$\log L = n \log p + \left(\sum x_i - n\right) \log(1-p) \quad \text{—— (2)}$$

For MLE of p, we have

$$\frac{\partial}{\partial p}(\log L) = 0$$

$$\Rightarrow \frac{n}{p} - \frac{\left(\sum x_i - n\right)}{(1-p)} = 0$$

$$\Rightarrow n(1-p) - \left(\sum x_i - n\right) p = 0$$

$$\Rightarrow n - \cancel{np} - p\sum x_i + \cancel{np} = 0$$

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

$$\Rightarrow \quad \hat{p} = \frac{1}{\bar{x}}$$

$\Rightarrow \frac{1}{\bar{X}}$ is MLE for p.

Moreover, $\left.\frac{\partial^2}{\partial p^2}(\log p)\right|_{p=\frac{1}{\bar{x}}}$

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

$$= \frac{-n}{p^2} - \frac{\left(\frac{n}{p} - n\right)}{(1-p)^2}$$

$$= \frac{-n}{p^2} - \frac{n\cancel{(1-p)}}{p(1-p)^{\cancel{2}}} = -\left[\frac{n}{p^2} + \frac{n}{p(1-p)}\right]$$

$$= \frac{-n}{p}\left[\frac{1-p+p}{p(1-p)}\right] = \frac{-n}{p^2(1-p)} < 0$$

$\Rightarrow \hat{p} = \frac{1}{\bar{x}}$ is MLE for p.

(5) Estimation of exponential distribution parameter:

Suppose $X_1, X_2, X_3, \dots X_n$ is a random sample from an exponential distribution with parameter α. Because of independence, the likelihood function is a product of the individual's p.d.f.

$$f(\vec{x}, \alpha) = \prod_{i=1}^{n} f(x_i, \alpha)$$

$$X \sim \exp(\alpha)$$

$$f(x, \alpha) = \alpha e^{-\alpha x} \quad ; \quad x > 0$$

$$f(\vec{x}; \alpha) = \prod_{i=1}^{n} \alpha e^{-\alpha x_i}$$

$$\therefore \quad L(\vec{x}, \alpha) = \alpha^n e^{-\alpha \sum_{i=1}^{n} x_i} \quad \text{---} (1)$$

Taking log on both side

$$\log L = n \log \alpha - \alpha \sum_{i=1}^{n} x_i \quad \text{---} (2)$$

For MLE, $\frac{\partial}{\partial \alpha}(\log L) = 0$

$$\Rightarrow \frac{n}{\alpha} - \sum_{i=1}^{n} x_i = 0$$

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

$$\Rightarrow \boxed{\hat{\alpha} = \frac{1}{\bar{x}}}$$

and $\frac{\partial^2}{\partial \alpha^2}(\log L) = \frac{-n}{\alpha^2} < 0$

Thus, $\hat{\alpha} = \frac{1}{\bar{x}}$ is MLE for α.

⑥ <u>MLE of μ and σ^2 for Normal distribution $X \sim N(\mu, \sigma^2)$</u> :

Let $X_i \sim N(\mu, \sigma^2)$ $\forall$ $i = 1, 2, 3, \cdots n$

The probability density function (p.d.f) for normal distribution is given by

$$f(x, \mu; \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad ; \quad -\infty < x < \infty$$

$\therefore$ The Likelihood function is given by

$$L = \prod_{i=1}^{n} f(x_i, \mu, \sigma^2)$$

$$= \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2}$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2}$$

$$L(\vec{x}, \mu, \sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2} \quad \text{—} (1)$$

Taking log on both side, we have

$$\log L = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\log L = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 \quad \text{—} (2)$$

For MLE of μ,

$$\frac{\partial}{\partial\mu}(\log L) = 0$$

$$\Rightarrow \quad 0 - 0 - \frac{1}{2\sigma^2}\sum_{i=1}^{n} 2(x_i-\mu)(-1) = 0$$

$$\Rightarrow \quad \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) = 0$$

$$\Rightarrow \quad \sum_{i=1}^{n} x_i - n\mu = 0$$

(i) MLE of μ, when σ^2 is known

(ii) MLE of σ^2 when μ is known.

(iii) Simultaneous estimation of μ and σ^2.

$$\Rightarrow \quad \mu = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$$

$$\Rightarrow \quad \boxed{\hat{\mu} = \bar{x}} \Rightarrow \text{MLE for } \mu.$$

Moreover, $\frac{\partial^2}{\partial \mu^2}(\log L) = \frac{-n}{\sigma^2} < 0$

Thus, $\hat{\mu} = \bar{x}$ is MLE for μ.

(ii) Now $\frac{\partial}{\partial \sigma^2}(\log L) = 0$

$$0 - \frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2}\frac{1}{(\sigma^2)^2}\sum_{i=1}^{n}(x_i-\mu)^2 = 0$$

$$\frac{1}{\cancel{2}(\sigma^2)^{\cancel{2}}}\sum_{i=1}^{n}(x_i-\mu)^2 = \frac{n}{\cancel{2}}\frac{1}{\cancel{\sigma^2}} \qquad \hat{\mu} = \bar{x}$$

$$\Rightarrow \quad \sigma^2 = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n} = \sigma^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}$$

Thus, $\hat{\sigma^2} = \sigma^2$ is MLE for σ^2.

Moreover, $$\left.\frac{\partial^2}{\partial (\sigma^2)^2}\log L\right|_{\mu=\bar{x}} = \frac{n}{2}\frac{1}{(\sigma^2)^2} - \frac{1}{2}\frac{2}{(\sigma^2)^3}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$= \frac{n}{2}\frac{1}{\sigma^4} - \frac{1}{\sigma^6}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$= \frac{n}{2}\frac{1}{\sigma^4} - \frac{1}{\sigma^6} n\sigma^2$$

$$= \frac{n}{2}\frac{1}{\sigma^4} - \frac{n}{\sigma^4} = -\frac{n}{2}\frac{1}{\sigma^4} < 0$$

Thus, $\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}$ is MLE for σ^2.

#

Question:- Let $X \sim Poiss(\alpha)$; $\alpha > 0$

In 50 observations of X, it is observed that exactly 20 out of them are zero. Find the maximum likelihood estimator of α.

Solution:- $P(X=0) = \frac{e^{-\alpha}\alpha^0}{0!} = e^{-\alpha}$

What is the probability function / Likelihood function for observing 20 zeroes.

$$L(\alpha) = {}^{50}C_{20}\,(e^{-\alpha})^{20}\,(1-e^{-\alpha})^{30}$$

find α such that $L(\alpha)$ is maximized.

$$\log(L(\alpha)) = \log {}^{50}C_{20} + 20\log(e^{-\alpha}) + 30\cdot\log(1-e^{-\alpha}) \quad \text{---(1)}$$

For MLE of α

$$\frac{\partial}{\partial\alpha}(\log L(\alpha)) = 0$$

$$\Rightarrow 0 - \frac{20}{e^{-\alpha}}(e^{-\alpha}) + \frac{30}{(1-e^{-\alpha})}(e^{-\alpha}) = 0$$

$$\Rightarrow \frac{30\,e^{-\alpha}}{1-e^{-\alpha}} = \frac{20\,e^{-\alpha}}{e^{-\alpha}}$$

put $e^{-\alpha} = p$

$$\Rightarrow \frac{30p}{1-p} = 20$$

$$\Rightarrow 30p = 20 - 20p$$

$$\Rightarrow 50p = 20$$

$$\Rightarrow p = \frac{2}{5}$$

$$\Rightarrow e^{-\alpha} = \frac{2}{5} \quad \Rightarrow \alpha = -\log\left(\frac{2}{5}\right)$$

* If $X_1, X_2, X_3, \cdots X_n$ be i.i.d. with p.d.f. $f(x, \theta)$; $\theta \in \Omega$

For a specific function g

if we know $\hat{\theta}$ is MLE of θ

$\Rightarrow g(\hat{\theta})$ is MLE of $g(\theta)$.

- If $\overline{X}$ is MLE of λ

$\Rightarrow (\overline{X})^2$ is MLE of λ^2

$\Rightarrow \dfrac{e^{-\overline{X}}(\overline{X})^2}{2!}$ is MLE of $\dfrac{e^{-\lambda}\lambda^2}{2!}$.