The Method of Maximum Likelihood Estimation

Example:- Consider a box containing black and white marbels. We know that the ratio of one colour of marbel to the other is ~~1~~ 1:3 but we ~~is~~ donot know whether black or white marbels are more numerous.

A statistician draws a random sample of size 3 from the box and decides to base his/her estimate on the reported outcome.

Let x denotes the no. of black marbels in the drawn sample.

Let $p \to$ unknown proportion of black marbels in the box.

Then $X \sim Bin(3, p)$

The problem is to estimate p.

All that we know is $p = 1/4$ or $3/4$

$$P_X(x) = \binom{3}{x} p^x (1-p)^{3-x}, \quad x = 0, 1, 2, 3$$
$$p = 1/4, 3/4.$$

x	0	1	2	3
$p = 1/4$	27/64	27/64	9/64	1/64
$p = 3/4$	1/64	9/64	27/64	27/64

$\leftarrow p(x)$

We observe that the likelihood of $p = 1/4$ is higher when $x = 0$ or 1 and that of $p = 3/4$ is higher when $x = 2$ or 3.

So,
$$\hat{p} = \begin{cases} \frac{1}{4} & \text{if } x = 0 \text{ or } 1 \\ 3/4 & \text{if } x = 2 \text{ or } 3 \end{cases}$$

Scanned by TapScanner

$p_X(x) = \binom{3}{x} p^x (1-p)^{3-x}$; $x = 0,1,2,3$

$0 \le p \le 1$

maximizing p_X & $\log p_X$ — these are same since log is a ↑ increasing fun^n

$\log p(x) = \log\binom{3}{x} + x\log p + (3-x)\cdot\log(1-p)$

$$\frac{\partial \log p(x)}{\partial p} = \frac{x}{p} - \frac{3-x}{1-p} = \frac{x-3p}{p(1-p)}$$

> 0 if $p < x/3$
< 0 if $p > x/3$

$\ln p$ is increasing for $p < x/3$
decreasing " $p > x/3$.

So, it attains a max^m at $x/3$

So, $\hat{p}_{MLE} = X/3$ → sample proportion.

To formalize the procedure we consider the likelihood fun^n of a random sample $x_1, x_2, \dots, x_n$ is observed $\underline{x} = (x_1, \dots, x_n)$

$$L(\underline{\theta}, \underline{x}) \triangleq \prod_{i=1}^{n} f_{X_i}(x_i, \underline{\theta})$$

The value of θ, say $\hat{\theta}(\underline{x})$ so that

~~$L(\hat{\underline{\theta}}(\underline{x}))$~~ $L(\hat{\underline{\theta}}, \underline{x}) \ge L(\underline{\theta}, \underline{x})$ $\forall\, \underline{\theta} \in \Theta$

is called the ML estimate of $\underline{\theta}$.

In practice we may often consider maximization of $\ln L(\underline{\theta}, \underline{x}) = l(\underline{\theta}, \underline{x})$ w.r.to $\underline{\theta}$ as ln is an increasing fun^n of θ.

A useful approach often applicable is to find solutions of the eq^n

$$\frac{\partial l}{\partial \theta} = 0, \quad \left[\frac{\partial l}{\partial \theta_1} = 0, \dots, \frac{\partial l}{\partial \theta_n} = 0\right]$$

likelihood eq^{n}s.

So, in general the method of maximum likelihood estimation can be described as below. Let $x_1, \dots, x_n$ be a random sample from a popn with pmf or pdf $f(x, \underline{\theta})$, $\underline{\theta} \in \Omega$, $\underline{\theta} = (\theta_1, \dots, \theta_k)$.

We write the likelihood function as

$$L(\underline{\theta}, \underline{x}) = \prod_{i=1}^{n} f(x_i, \underline{\theta}), \quad \underline{x} = (x_1, \dots, x_n).$$

The maxm likelihood estimator of $\underline{\theta}$ is $\hat{\underline{\theta}}(\underline{x})$ if

$$L(\hat{\underline{\theta}}(x), \underline{x}) \geq L(\underline{\theta}, \underline{x}) \quad \forall \, \underline{\theta} \in \Omega$$

Scanned by TapScanner

Finding Estimators

Maximum Likelihood Estimators

Examples 1:- Let $X \sim Bin(n, p)$, $0 \le p \le 1$

n is known.

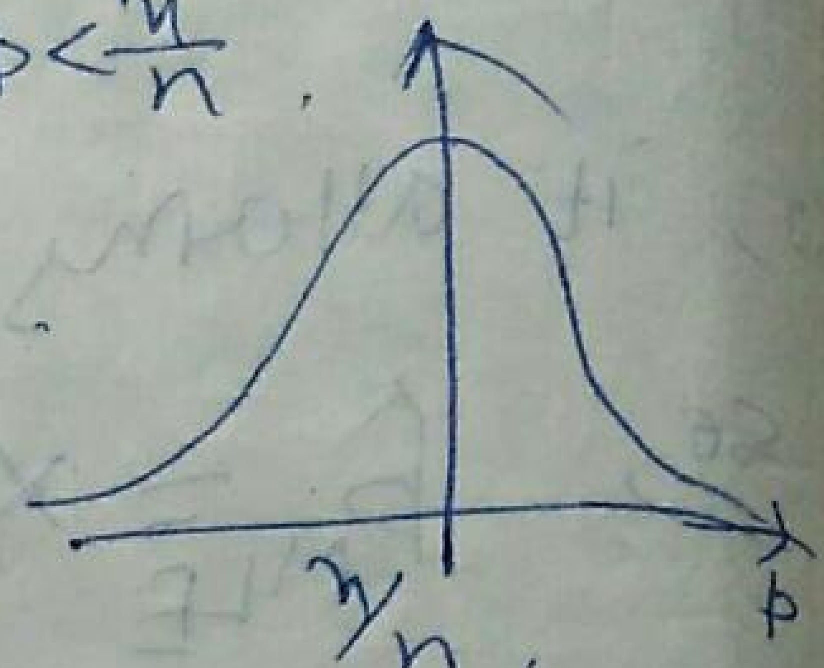
$$L(p, x) = f(x, p) = \binom{n}{x} p^{x} (1-p)^{n-x} \; ; \; x = 0, 1, 2, \dots, n \quad 0 \le p \le 1.$$

$$l(p) = \log L(p, x) = \log\binom{n}{x} + x \log p + (n-x)\log(1-p)$$

$$\frac{dl}{dp} = \frac{x}{p} - \frac{n-x}{1-p} = \frac{x - np}{p(1-p)} \begin{array}{l} < 0 \text{ if } p > x/n \\ > 0 \text{ if } p < \frac{x}{n}. \end{array}$$

$l(p)$ will be increasing if $p < x/n$.

" " decreasing if $p > x/n$



So, the maxm value is attained at x/n.

So, $\hat{p} = \frac{X}{n}$ is the maximum likelihood estimator of p (MLE).

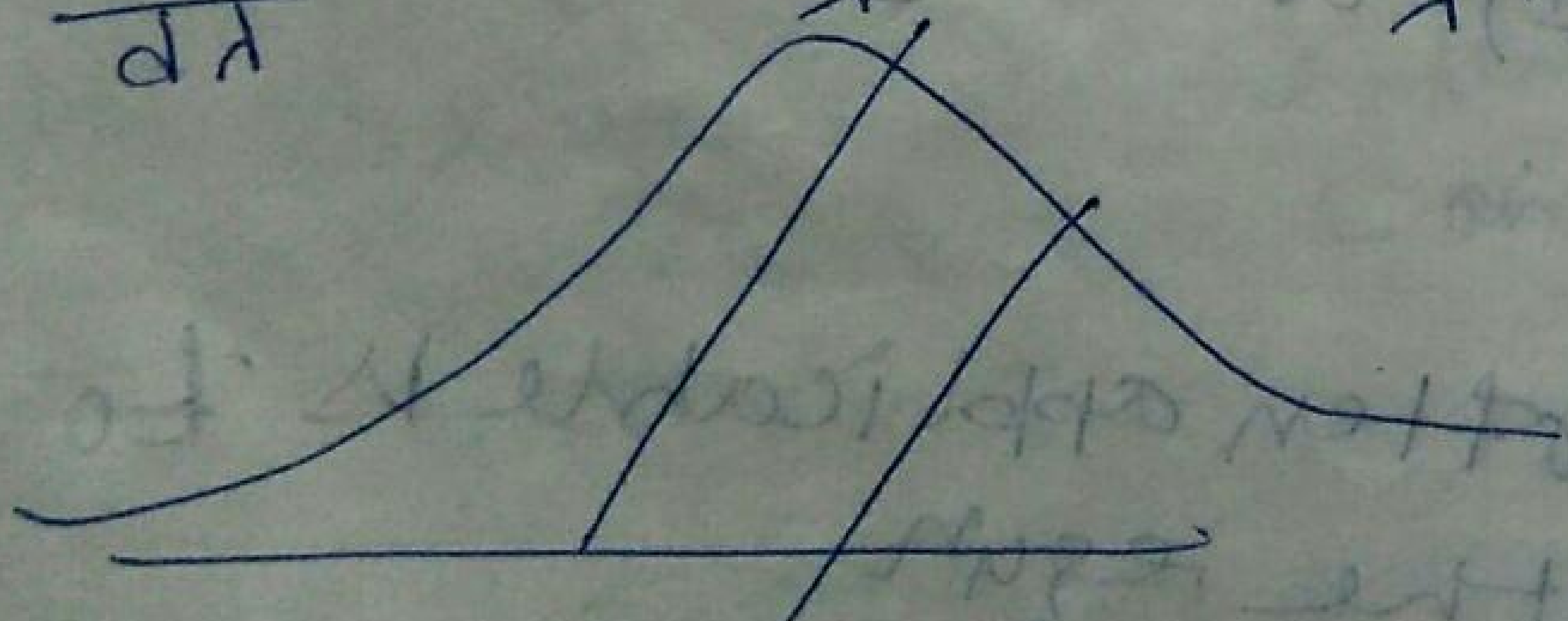
2. Ex. Let $X_1, X_2, \dots, X_n \sim P(\lambda)$, $\lambda > 0$. $\underline{x} = (x_1, \dots, x_n)$

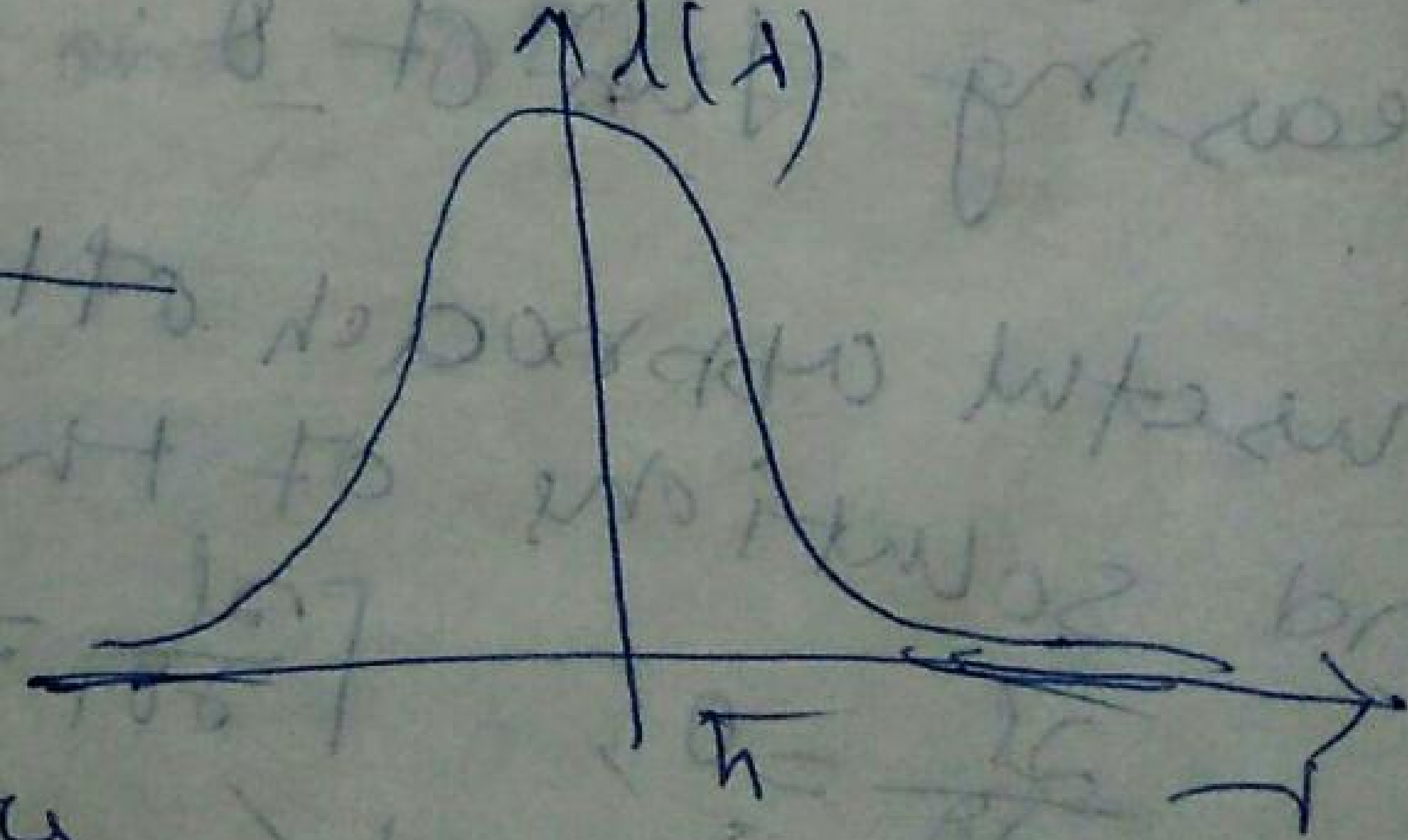
$$L(\lambda, \underline{x}) = \prod_{i=1}^{n} f(x_i, \lambda)$$

$$= \prod_{i=1}^{n} \frac{e^{-\lambda} \lambda^{x_i}}{x_i!} = \frac{e^{-n\lambda} \lambda^{\sum x_i}}{\prod_{i=1}^{n} x_i!}$$

$$l(\lambda) = \log L(\lambda, \underline{x}) = -n\lambda + \sum x_i \log\lambda - \log\left\{\prod_{i=1}^{n}(x_i!)\right\}$$

$$\frac{dl}{d\lambda} = -n + \frac{\sum x_i}{\lambda} = \frac{\sum x_i - n\lambda}{\lambda} \begin{array}{l} > 0 \text{ if } \lambda < \bar{x} = \frac{\sum x_i}{n} \\ < 0 \text{ if } \lambda > \bar{x} \end{array}$$

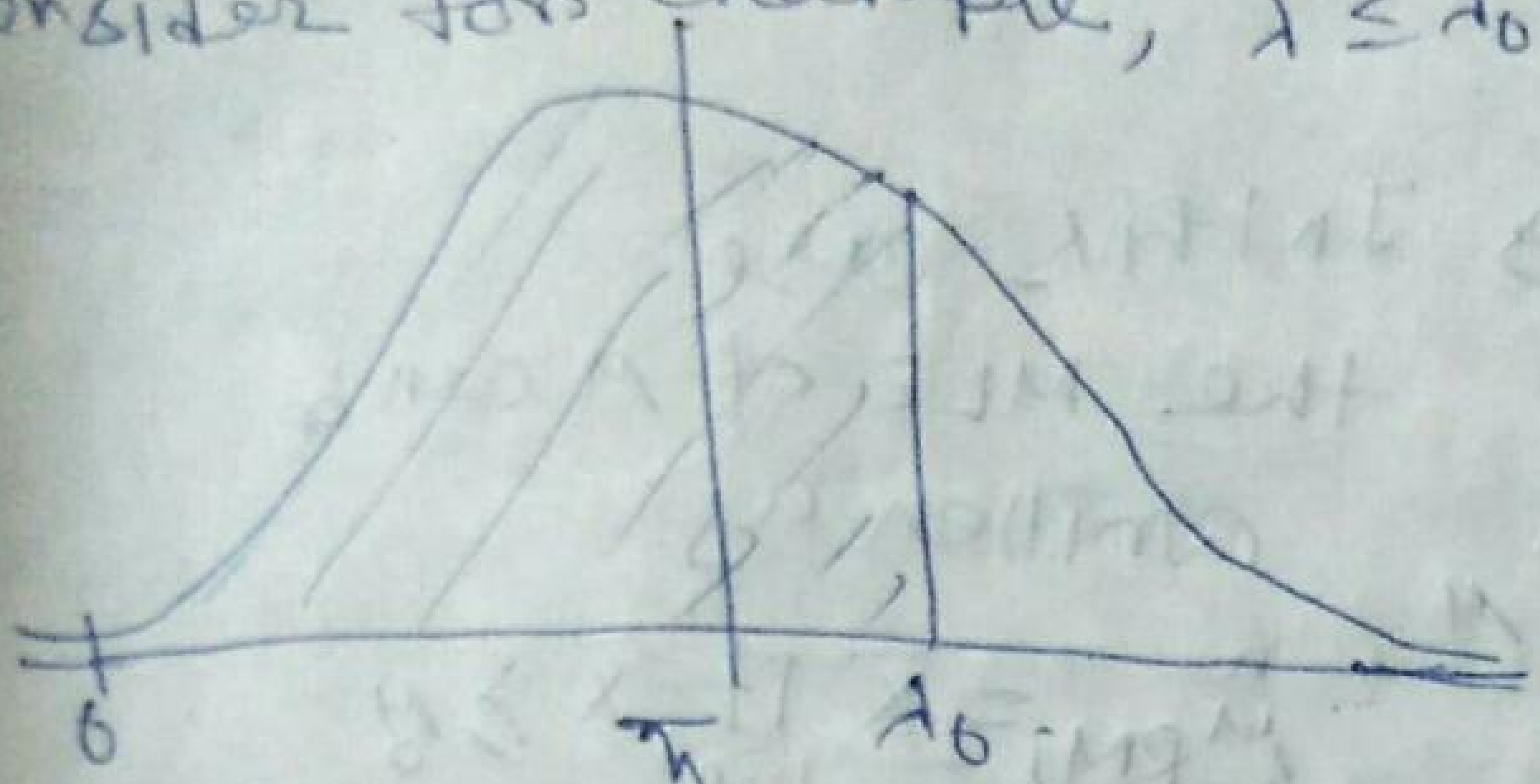


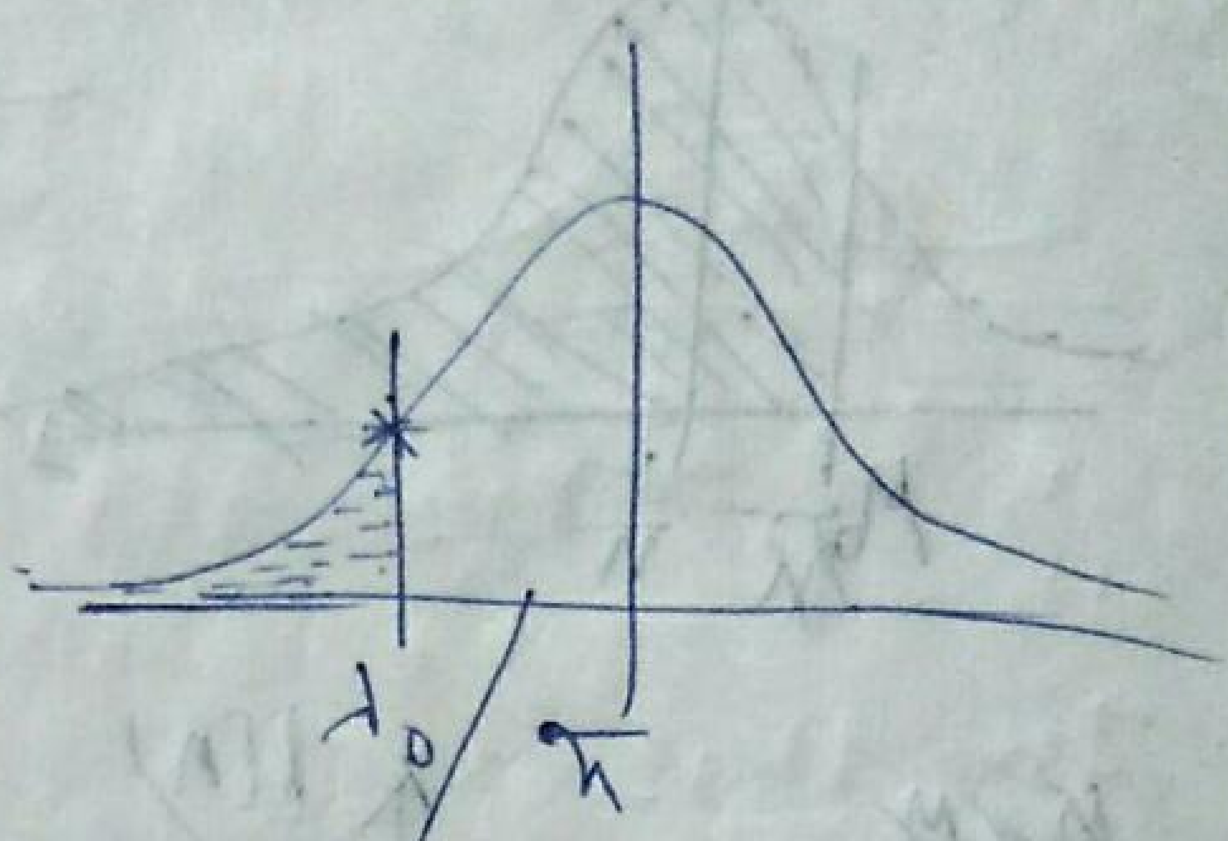


So, the maxm occurs at $\lambda = \bar{x}$.

So, $\hat{\lambda} = \overline{X}$ is the max^m likelihood esitimator of λ.

Consider for example, $\lambda \leq \lambda_0$ (λ_0 fixed unknown)





In this case MLE of λ is $\hat{\lambda}_{RML} = \begin{cases} \overline{X} & \text{if } \overline{X} \leq \lambda_0 \\ \lambda_0 & \text{if } \overline{X} > \lambda_0. \end{cases}$

restricted Maximum Likelihood.

In this MLE of λ is λ_0.

4. $X_1, X_2, \ldots, X_n \sim N(\mu, \sigma^2)$

<u>Case I:</u> σ^2 is known σ^2, say $\sigma^2 = 1$ (WLOG)

The likelihood fun^n is

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

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

$$= \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}\Sigma(x_i - \mu)^2}$$

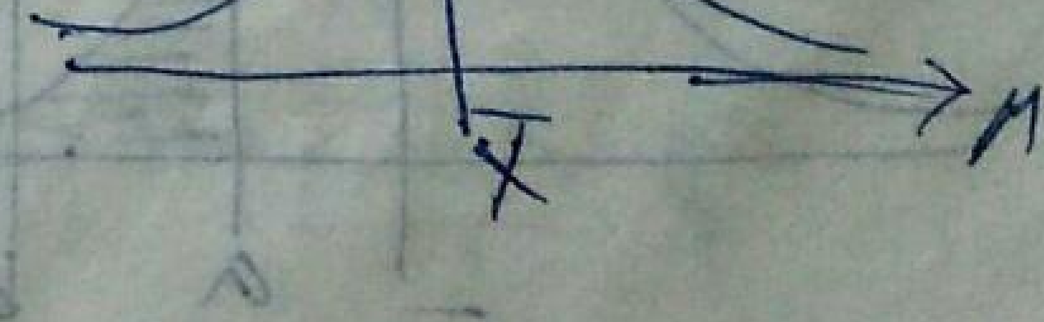
$$l(\mu) = \log L(\mu, \underline{x}) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}(x_i - \mu)^2$$

$$\frac{dl}{d\mu} = \Sigma(x_i - \mu) = 0 \Rightarrow n(\bar{x} - \mu) = 0$$

$$\Rightarrow \hat{\mu} = \bar{x}$$

$n(\bar{x} - \mu) < 0$ if $\mu > \bar{x}$

> 0 if $\mu < \bar{x}$



So, $\bar{x}$ is the MLE of μ.

Suppose (i) $\mu \geq \mu_0$.

Case II :- μ is known $\mu = 0$ (WLOG)

The likelihood fun^n is

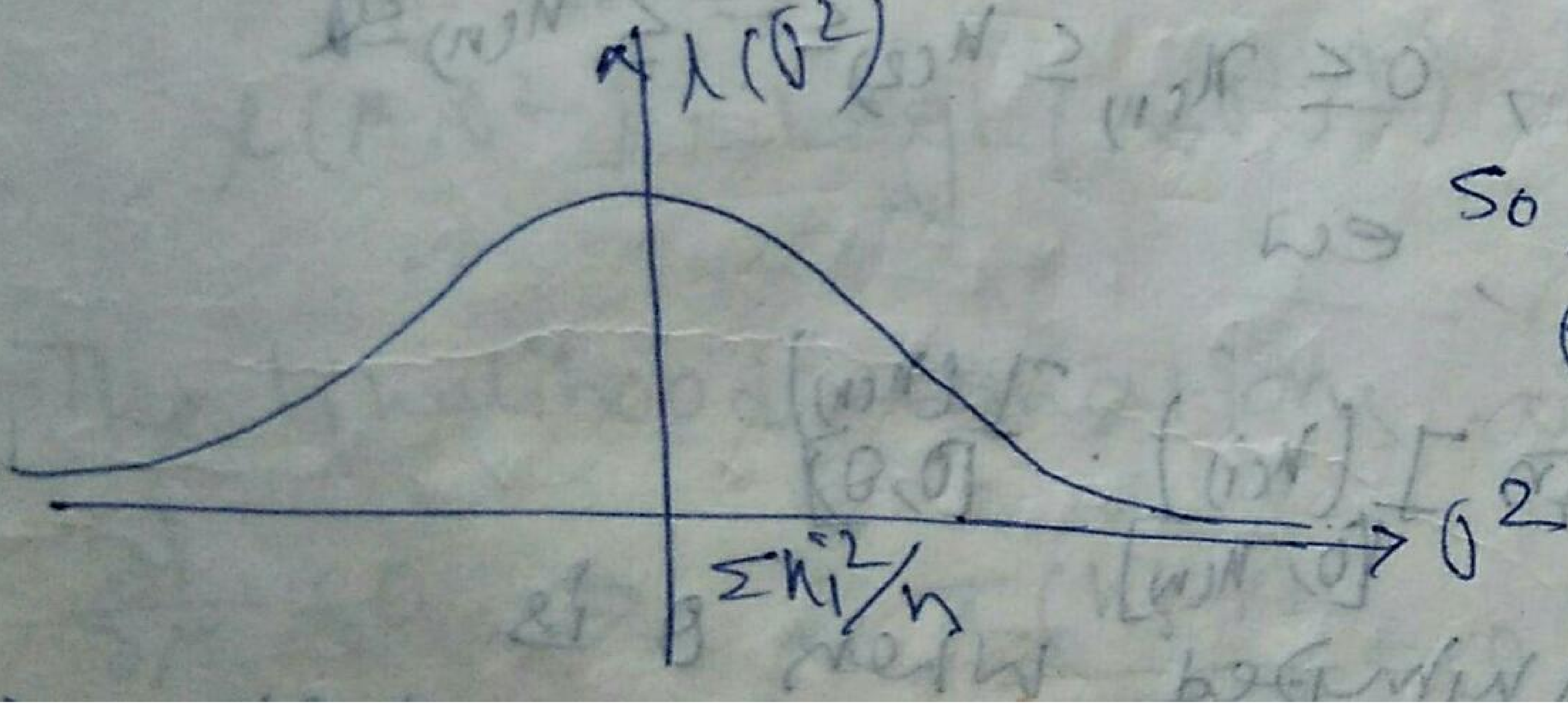
$$L(\sigma^2, \underline{x}) = \prod_{i=1}^{n} f(x_i, \sigma^2)$$

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

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

$$\lambda(\sigma^2) = \log L(\sigma^2, \underline{x}) = -\frac{n}{2} \log \sigma^2 - \frac{n}{2} \log 2\pi - \frac{\sum x_i^2}{2\sigma^2}$$

$$\frac{d\lambda}{d\sigma^2} = 0 \Rightarrow -\frac{n}{2\sigma^2} + \frac{\sum x_i^2}{2\sigma^4} = \frac{\sum x_i^2 - n\sigma^2}{2\sigma^4} \quad \begin{matrix} < 0 \text{ if } \sigma^2 > \frac{\sum x_i^2}{n} \\ > 0 \text{ if } \sigma^2 < \frac{\sum x_i^2}{n} \end{matrix}$$



So, the ~~max^m~~ MLE of σ^2 is

$$\boxed{\hat{\sigma}^2_{ML} = \frac{1}{n} \sum_{i=1}^{n} x_i^2}$$

$$\hat{\sigma}^2_{RML} = \min\left(\frac{1}{n}\sum x_i^2, \sigma_0^2\right)$$

5. $X_1, X_2, \ldots, X_n$ is a random sample from $U[0, \theta]$, $\theta > 0$.

The likelihood fun^n is

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

$$= \begin{cases} \frac{1}{\theta^n}, & 0 \le x_i \le \theta, \ i = 1, 2, \ldots, n \\ 0, & \text{otherwise.} \end{cases}$$

We may write the likelihood fun^n as

$$L(\theta, \underline{x}) = \frac{1}{\theta^n}, \quad 0 \le x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)} \le \theta$$

$$= 0, \quad \text{ow}$$

or, $$L(\theta, \underline{x}) = \frac{1}{\theta^n} I_{[0, x_{(n)}]}(x_{(1)}) \cdot I_{[0, \theta]}(x_{(n)})$$

So, L is maximized when θ is minimized which is possible when

$\theta = x_{(n)}$.

So, $\hat{\theta}_{MLE} = X_{(n)}$

$\hat{\theta}_{1ML} = \ldots(1), \hat{\theta}_{2ML} = \ldots(n).$

(7) Let $x_1, x_2, \ldots, x_n \sim N(\mu, \sigma^2)$ both μ & σ^2 are unknown.

$$\Rightarrow f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \begin{matrix} -\infty < \mu < \infty \\ \sigma^2 > 0 \end{matrix}$$

The likelihood funⁿ is

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

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

$$= \frac{1}{\sqrt{2\pi}\,(2\pi\sigma^2)^{n/2}} e^{-\frac{\Sigma(x_i-\mu)^2}{2\sigma^2}}$$

The log likelihood funⁿ is

$$l(\mu, \sigma^2) = \log L(\mu, \sigma^2, \underline{x}) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{\Sigma(x_i-\mu)^2}{2\sigma^2}$$

The likelihood equations are

$$\frac{\partial l}{\partial \mu} = 0 \Rightarrow \frac{\Sigma(x_i-\mu)}{\sigma^2} = 0 \Rightarrow \boxed{\hat{\mu} = \bar{X}}$$

$$\frac{\partial l}{\partial \sigma^2} = 0 \Rightarrow -\frac{n}{2\sigma^2} + \frac{\Sigma(x_i-\mu)^2}{2\sigma^4} = 0$$

$$\Rightarrow \boxed{\hat{\sigma}^2 = \frac{1}{n}\Sigma(x_i-\mu)^2}$$

So, the ~~ma~~ MLEs of μ & σ^2 are

$$\hat{\mu}_{ML} = \bar{X} \quad \& \quad \hat{\sigma}^2_{ML} = \frac{1}{n}\Sigma(x_i - \bar{X})^2$$

same as MMEs of μ & σ^2.

Scanned by TapScanner