



PRINCIPLES OF POINT ESTIMATION

D. Uma

Computer Science and Engineering

STATISTICS FOR DATA SCIENCE

Maximum Likelihood Estimation

D. Uma

Computer Science and Engineering

STATISTICS FOR DATA SCIENCE

Topics to be covered...



PES
UNIVERSITY
ONLINE

Maximum Likelihood Estimate (MLE)

method
Best Point Estimator

$X \sim \text{Bin}(n, p)$
 p is unknown
 $X : 0, 1, 2, \dots, n$
proportion of success in population

Maximum Likelihood for Discrete Distributions

$X = 4$
 X : No. of successes
 n : no. of trials
 $x_1, x_2, x_3, \dots, x_{10}$
 x 0 1

$$f(x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}$$

Using Sample Estimate \hat{p}

$P(X=x) :$

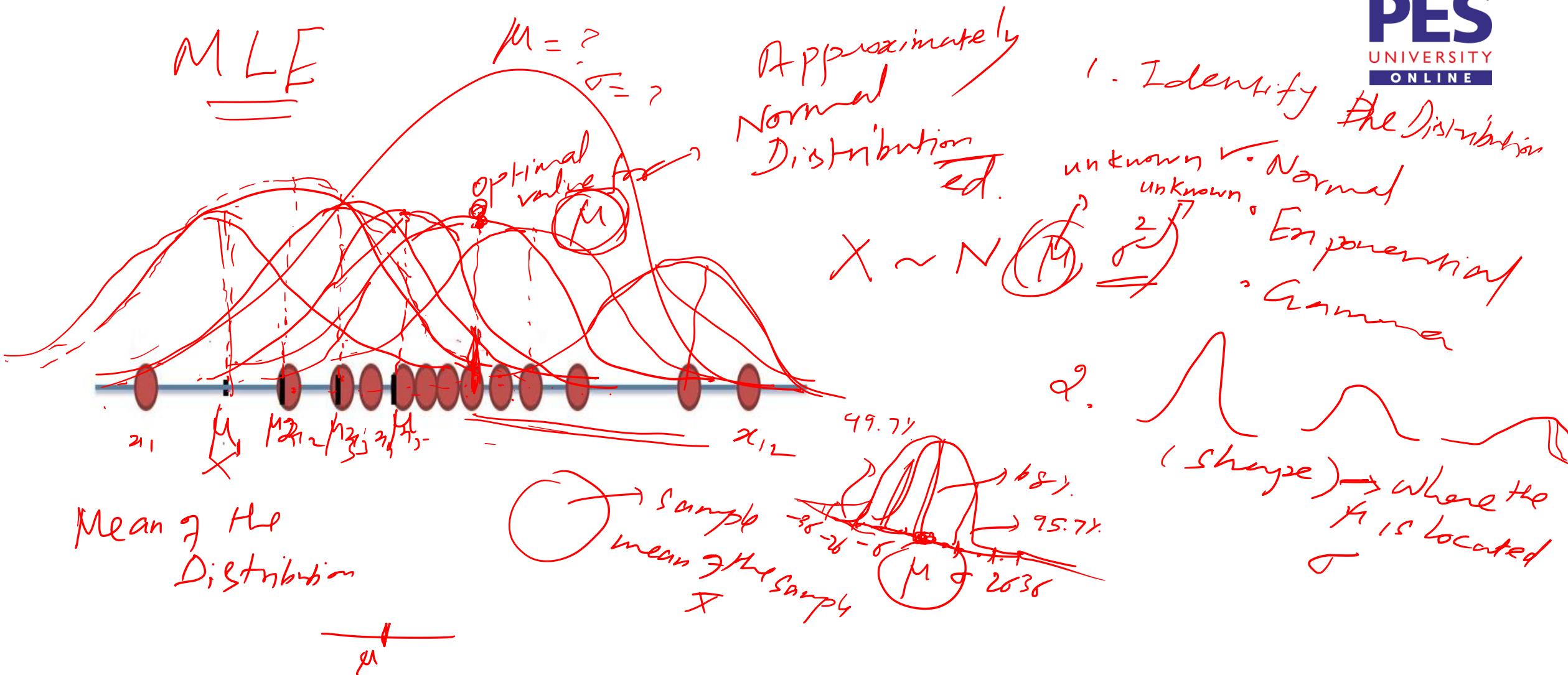
$$\hat{p} = ?$$

$$\hat{p} = \frac{x}{n} = \frac{4}{10}$$

$$\boxed{\hat{p} = \frac{x}{n}}$$

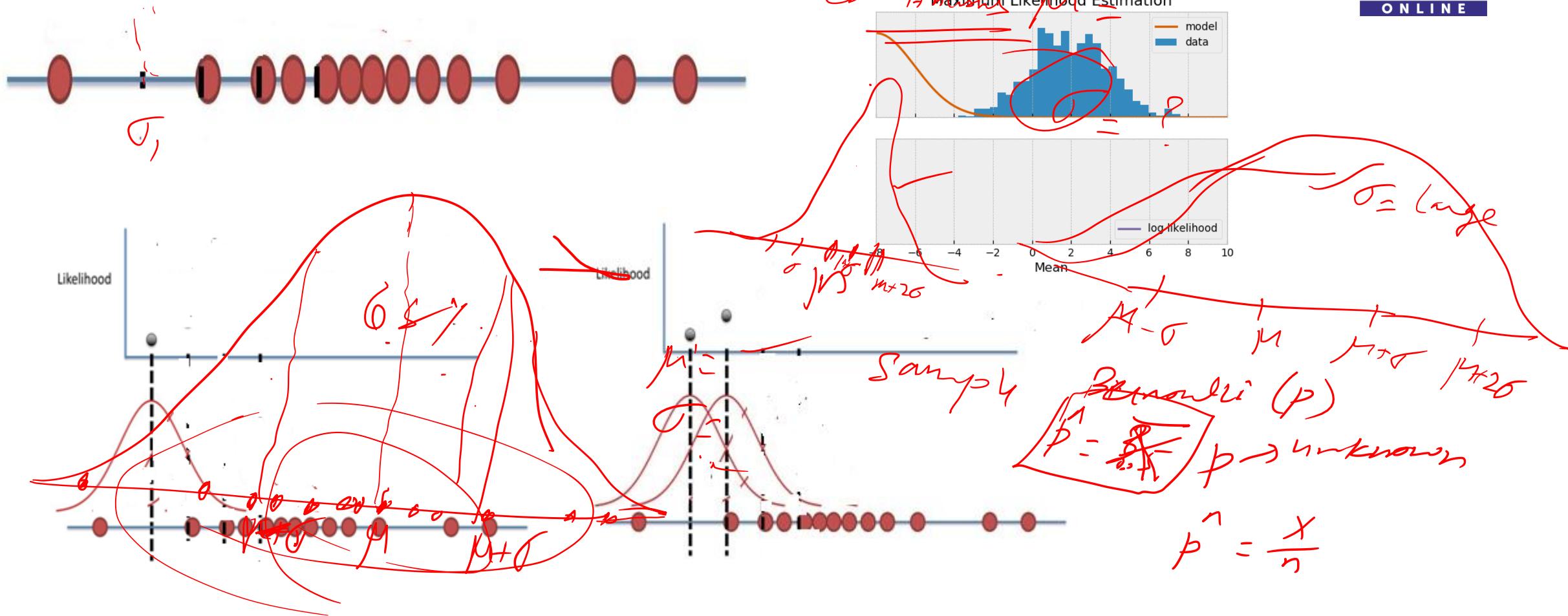
$\hat{p} = \frac{4}{10} = 0.4$ Point estimator

Maximum Likelihood Estimate (MLE)

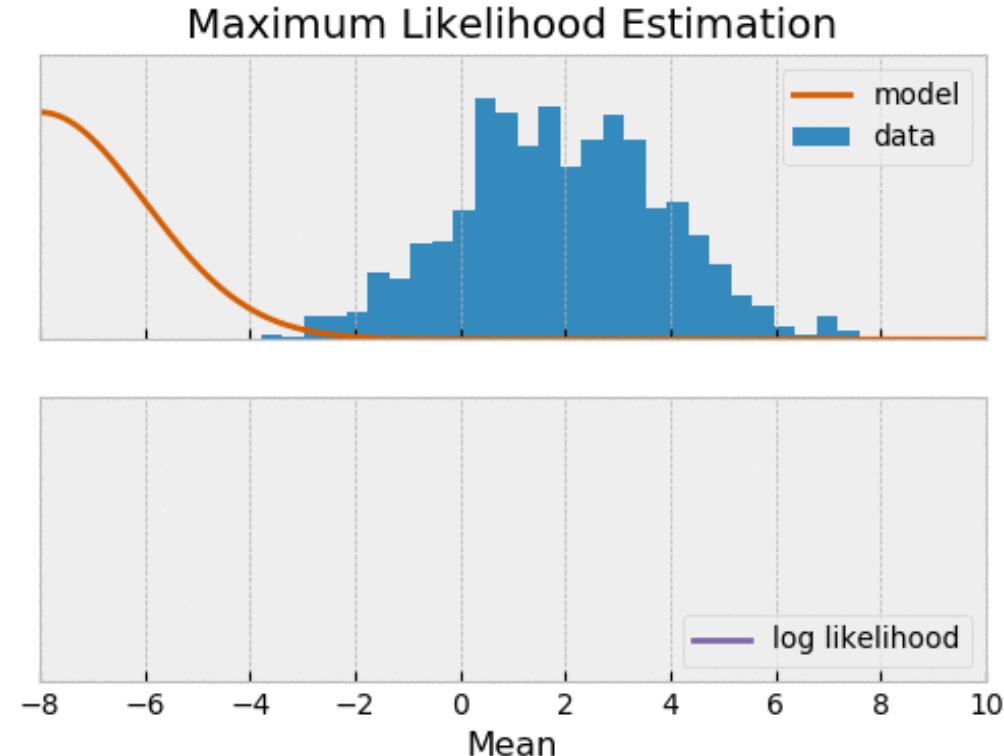


STATISTICS FOR DATA SCIENCE

Maximum Likelihood Estimate (MLE)



Maximum Likelihood Estimate(MLE)



To estimate the parameter with the value that makes the observed data most likely.

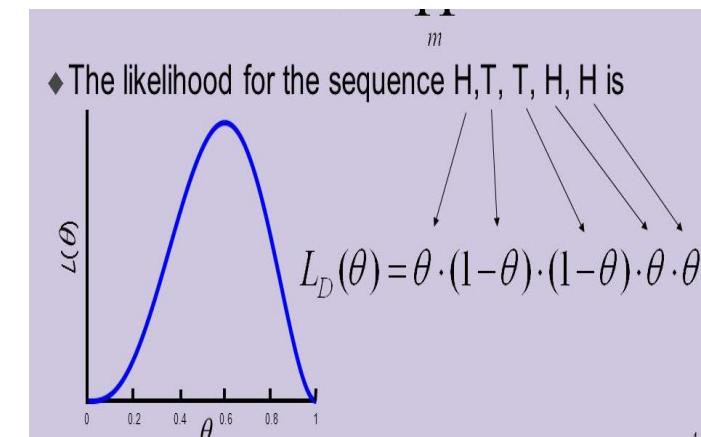
Method to Construct Good Estimator

Method to Construct Good Estimator

MLE (maximum likelihood estimator) is regarded as the **best method** to point estimation.

Likelihood function is the probability of obtaining observed value.

$$L_D(\theta) = P(D|\theta) = \prod_m P(x[m]|\theta)$$



Note: If single observation is made, Likelihood function is just the probability of obtaining that value. (There will be no product involved)

Maximum Likelihood Estimator

MLE is that value of the estimator which when substituted in place of the parameter, maximizes the Likelihood function.

$$L_D(\theta) = P(D | \theta) = \prod_m P(x[m] | \theta)$$

Method to Construct Good Estimator

Method of Maximum Likelihood(MLE) is regarded as the best method to point estimation.

Likelihood function is the probability of obtaining observed value.

If single observation is made, Likelihood function is just the probability of obtaining that value.

MLE is that value of the estimator which when substituted in place of the parameter, maximizes the Likelihood function.

$$L(P)_{x_1, x_2, \dots, x_n} = \text{sample}$$

$P = ?$

Joint Prob. Mass

$$x = 4 \quad \text{Prob. } f_{x=4}$$
$$L = P(x=4)$$

General Steps to proceed with MLE.

Step 1: Write down the likelihood function.

Step 2: Take natural log of likelihood function.

(Reason: the quantity that maximizes log of a function is always the same quantity that maximizes the function itself)

Step 3: Differentiate log-likelihood function with respect to the parameter being estimated.

Step 4: Set the derivative equal to 0 to get MLE.

General Steps to proceed with MLE.

Step 1: Write down the likelihood function.

$$\ln L(\theta | x_1, x_2, \dots, x_n)$$

Step 2: Take natural log of likelihood function.

(Reason: the quantity that maximizes log of a function is always the same quantity that maximizes the function itself)

$$\frac{\partial}{\partial \theta} (\ln L) = 0 \Rightarrow \boxed{0' = ?}$$

Step 3: Differentiate log-likelihood function with respect to the parameter being estimated.

Step 4: Set the derivative equal to 0 to get MLE.

$$\begin{aligned} \text{MLE } \hat{\theta} &= \theta' \\ \text{MLE } \hat{\mu} &= \mu' \\ \text{MLE } \hat{\sigma}^2 &= \sigma'^2 \end{aligned}$$

Points to Remember

- The **maximum likelihood estimate** is the value of the estimators that when substituted in for the parameters maximizes the likelihood function.
- The likelihood function can be a **probability density function** or a **probability mass function**.
- It can also be a **joint probability** density or mass function, and is often the joint density or mass function of independent random variables.



Maximum Likelihood Estimator for Bernoulli Distribution

Let $x_1, x_2, x_3, \dots, x_n$ be a random sample from a population with Bernoulli(p) distribution. Find MLE of p .

$x_i \sim \text{Bernoulli}(p)$ for $i = 1$ to n

Distribution	X :	0	1
$p(x)$ or $P(X=x)$:		$1-p$	p

$$P(X=x_i) = p^{x_i} (1-p)^{1-x_i}$$

1. Write down the likelihood function $L(\theta)$
2. Take natural log of likelihood function.
3. Differentiate log-likelihood function w.r.t. to the parameter being estimated.
4. Set the derivative = 0 to get MLE.

Maximum Likelihood Estimator for Bernoulli Distribution

$$P(X=x_i) = p^{x_i} (1-p)^{1-x_i}$$

$$L = f(p | x_1, x_2, \dots, x_n)$$

$$L = P(X=x_1) * P(X=x_2) * \dots * P(X=x_n) \quad L(p) =$$

$$= p^{x_1} (1-p)^{1-x_1} * p^{x_2} (1-p)^{1-x_2} * \dots * p^{x_n} (1-p)^{1-x_n}$$

$$\textcircled{1} L = \prod_{i=1}^n p^{x_i} (1-p)^{1-x_i} \simeq p^{\sum_{i=1}^n x_i} (1-p)^{n - \sum_{i=1}^n x_i}$$

$$\textcircled{2} \ln L = \sum x_i \ln p + (n - \sum x_i) \ln (1-p)$$

H T T H H

$$L(p) = p \cdot (1-p) \cdot (1-p) \cdot p \cdot p$$

Maximum Likelihood Estimator for Bernoulli Distribution

$$\text{3. } \frac{d \ln L}{dp} = 0$$

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

$$\Rightarrow \sum x_i \frac{d \ln p}{dp} + (n - \sum x_i) \frac{d \ln(1-p)}{dp} = 0$$

$$\Rightarrow \frac{\sum x_i}{p} + (n - \sum x_i) \frac{(-1)}{1-p} = 0 \quad , \quad \Rightarrow \sum x_i - p \sum x_i - np + p n = 0$$

$$\Rightarrow \sum x_i \quad \quad \quad) \quad \quad \quad) \quad \quad \quad -$$

$$\Rightarrow \frac{\sum x_i}{p(1-p)}$$

$$\text{MLE } \hat{p} = \bar{x}$$

Maximum Likelihood Estimator for Bernoulli Distribution

 $X \sim \text{Bern}$

$$\text{Let } L(p|x) = p^x (1-p)^{1-x} \quad p = ?$$

Let $x_1, x_2, x_3, \dots, x_n$ which,

$x_i \sim \text{Bernoulli}(p)$ unknown

Find MLE of p .

 $X \sim \text{Bernoulli}(p) \quad n = 5$
 $\begin{matrix} x_1 & x_2 & x_3 & x_4 & x_5 \\ H & H & T & H & T \end{matrix}$

$$L(p|x_1=H, x_2=H, x_3=T, x_4=H, x_5=T) = p^4 (1-p)^1$$

$$p = (\hat{p} = 1)$$

$$L(p|x_1=H, x_2=H, x_3=T, x_4=H, x_5=T) = \prod_{i=1}^n p^{x_i} (1-p)^{1-x_i}$$



Maximum Likelihood Estimator for Bernoulli Distribution

Find MLE of p .

Let x_1, x_2, \dots, x_n be a simple random sample which follows Bernoulli(p) distribution.

$$i. L(p|x_1, \dots, x_n) = \prod_{i=1}^n p^{x_i} (1-p)^{1-x_i}$$

$$x_i \sim \text{Bernoulli}(p)$$

$$P(x_i = 0) = 1-p \quad P(x_i = 1) = p$$

$$p^{x_i} (1-p)^{1-x_i}$$

$$L(p|x_1, \dots, x_n) = p^{\sum x_i} (1-p)^{n - \sum x_i}$$

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

multiply by $p(1-p)$

$$(1-p)\sum x_i \rightarrow (n - \sum x_i) = 0$$

$$\sum x_i - np = 0$$

$$\Rightarrow p = \frac{\sum x_i}{n} = \bar{x} \Rightarrow \boxed{\text{MLE of } p = \bar{x}}$$

$$2. \ln L(p| \cdot) = \sum_{i=1}^n x_i \ln p + (n - \sum_{i=1}^n x_i) \ln (1-p)$$

$$3. \frac{d}{dp} (\ln L(p)) = 0 \Rightarrow \sum x_i \frac{1}{p} + (n - \sum x_i) \frac{1+(1)}{1-p} = 0$$

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

Maximum Likelihood Estimator for Binomial Distribution

Let X be a Binomial r.v with $X \sim \text{Bin}(n, p)$ follows Binomial (n, p) distribution where p is unknown. Find MLE of p .

$$1. L(p|x) = \frac{n!}{x!(n-x)!} * p^x * (1-p)^{n-x} \quad p(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

$$d. \ln L(p|x) = \ln\left(\frac{n!}{x!(n-x)!}\right) + x \ln(p) + (n-x) \ln(1-p)$$

$$3. \frac{\partial}{\partial p} (\ln L) = 0 \Rightarrow \frac{d}{dp} \left(\ln n! - \ln x! - \ln(n-x)! + x \ln(p) + (n-x) \ln(1-p) \right) = 0$$

Maximum Likelihood Estimator for Binomial Distribution

$$\begin{aligned}
 - \frac{d}{dp} (\ln L) = 0 &\Rightarrow \frac{d}{dp} (x \ln(p) + (n-x) \ln(1-p)) = 0 \\
 &\Rightarrow \frac{x}{p} + \frac{n(-1)}{1-p} - \frac{x(-1)}{(1-p)} = 0 \\
 &\Rightarrow \frac{x}{p} - \frac{n}{1-p} + \frac{x}{1-p} = 0
 \end{aligned}$$

Multiplying by $p(1-p)$

$$\begin{aligned}
 x(1-p) - np + xp &= 0 \\
 x - xp - np + xp &= 0 \\
 x - np &\Rightarrow p = \frac{x}{n}
 \end{aligned}$$

i. MLE of p is

$$\boxed{\hat{p} = \frac{x}{n}}$$

Maximum Likelihood Estimator for Poisson Distribution

Let X_1, X_2, \dots, X_n be a simple random sample from a population which follows $X_i \sim \text{Poisson}(\lambda)$ distribution where λ is unknown. Find MLE of λ .

$$X_i \sim \text{Poisson}(\lambda)$$

$$P(X_i = x_i) = \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$$

$$P(X_1=x_1; X_2=x_2; \dots; X_n=x_n)$$

$$L(\lambda | x_1, x_2, \dots, x_n) = \prod_{t=1}^n \frac{\lambda^{x_t} e^{-\lambda}}{x_t!} = \frac{e^{-\lambda} \lambda^{x_1}}{x_1!} + \frac{e^{-\lambda} \lambda^{x_2}}{x_2!} + \dots + \frac{e^{-\lambda} \lambda^{x_n}}{x_n!}$$

$$\ln L(\lambda | x) = -n\lambda + (\sum_{t=1}^n x_t) \ln \lambda - \sum_{t=1}^n \ln x_t!$$

$$\frac{d}{d\lambda} (\ln L(\lambda)) = 0 \Rightarrow -n + (\sum_{t=1}^n x_t) \frac{1}{\lambda} = 0 \Rightarrow -n + \frac{\sum x_t}{\lambda} = 0$$

i.e. MLE of λ is $\boxed{\lambda = \bar{x}}$

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

Maximum Likelihood Estimator for Normal Distribution

Sample Set
 $n < 10$ $X \sim N(\mu, \sigma^2)$

$$L(\mu, \sigma^2 | x_1, x_2, \dots, x_n) = \prod f(x_i; \mu, \sigma^2)$$

Normal

$$P(a \leq X \leq b) = \int_a^b f(x) dx$$

prob. density $f_h.$

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

~~pdf~~

$$P(X_1=1; X_2=5; X_3=7, \dots, X_{10}=13)$$

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

$$= \left(\frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} \right)^n \cdot \left(\frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x_2-\mu)^2}{2\sigma^2}} \right) \cdot \dots \cdot \left(\frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x_{10}-\mu)^2}{2\sigma^2}} \right)$$

Joint Prob

$$P(X=x_1; Y=y_1)$$

$$= P(X=x_1) * P(Y=y_1)$$

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

Maximum Likelihood Estimator for Normal Distribution

Let x_1, x_2, \dots, x_n be a random sample from a $N(\mu, \sigma^2)$ distribution. Find MLE of μ and σ .

$$\text{If } x \sim N(\mu, \sigma^2) \quad f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$x_i \sim N(\mu, \sigma^2)$$

$$L(\mu, \sigma | x_1, \dots, x_n) = \prod_{i=1}^n \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

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

$$\ln L(\mu, \sigma) = -n \ln \sigma - \frac{\sum_{i=1}^n (x_i-\mu)^2}{2\sigma^2} \rightarrow 0$$

$$\begin{matrix} x_1 \\ \vdots \\ x_n \end{matrix}$$

Maximum Likelihood Estimator for Normal Distribution

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

Goal: To find MLE of μ

$$\frac{\partial}{\partial \mu} (\ln L(\mu, \sigma)) = 0 \Rightarrow 0 - \frac{1}{2\sigma^2} + 2 \sum_{i=1}^n (x_i - \mu) \cdot (-1) = 0$$

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

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

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

$$\therefore \text{MLE of } \mu \text{ is } \boxed{\mu = \bar{x}}$$

Case ii) Maximum Likelihood Estimator for σ^2

$$\ln L(\mu, \sigma | x_1, \dots, x_n) = -n\ln\sigma - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

$$\frac{\partial}{\partial \sigma} (\ln L(\mu, \sigma | x)) = 0 \Rightarrow -n + \frac{1}{\sigma} - \frac{\sum (x_i - \mu)^2}{\sigma^3} * (-2) = 0$$

MLE of σ^2

$$\hat{\sigma}^2 = \frac{\sum (x_i - \mu)^2}{n} \Rightarrow -\frac{n}{\sigma^2} + \frac{\sum (x_i - \mu)^2}{\sigma^3} = 0$$

$$\Rightarrow -n\sigma^2 + \sum (x_i - \mu)^2 = 0$$

$$\Rightarrow n\sigma^2 = \sum (x_i - \mu)^2$$

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

Maximum Likelihood Estimator: Problems

Let $X \sim \text{Bin}(20, p)$. Suppose we observe the value of $x = 7$. Find MLE of p .

X : No. of successes

$$P(x) = nC_x p^x (1-p)^{n-x}$$

$$\begin{aligned} n &= 20 \\ x &= 7 \end{aligned}$$

$$\hat{p} = \frac{x}{n} = \frac{7}{20}$$

~~$$\hat{p}(7) L(p|x=7) = 20C_7 p^7 (1-p)^{20-7} \approx p^7 (1-p)^{13}$$~~

Estimator for p is

$$\hat{p} = \frac{x}{n} = \frac{7}{20}$$

~~$$\textcircled{1} \quad L(p|x=7) = p^7 (1-p)^{13}$$~~

$$\ln(p^7 (1-p)^{13})$$

~~$$\textcircled{2} \quad \ln L(p|x=7) = 7 \ln p + 13 \ln(1-p)$$~~

~~$$\textcircled{3} \quad \frac{d \ln L}{dp} = 0 \Rightarrow 7 + \frac{1}{p} + 13 \frac{1}{1-p} (-1) = 0$$~~

$$\Rightarrow \frac{7}{p} - \frac{13}{1-p} = 0 \Rightarrow 7(1-p) - 13p = 0 \Rightarrow 7 - 7p - 13p = 0$$

$$\therefore \hat{p} = \frac{7}{20}$$

Maximum Likelihood Estimator: Problems

Suppose number of claims in a month from a certain portfolio of policies has a Poisson distribution with mean λ .

- a) If no. of claims observed in one particular month is 8.
 Find MLE of λ . $X \sim \text{Poisson}(\lambda)$

X : No. of claims $x \sim \text{Poisson}(\lambda)$

$$P(x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$$

$$L(\lambda | x=8) = \frac{e^{-\lambda} \cdot \lambda^8}{8!} \approx e^{-\lambda} \cdot \lambda^8$$

$$\ln L(\lambda | x=8) = -\lambda + 8 \ln \lambda$$

$$\frac{d}{d\lambda} (\ln L(\lambda)) = 0 \Rightarrow -1 + \frac{8}{\lambda} = 0 \Rightarrow \lambda = 8$$

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

$$\hat{\lambda} = \bar{x}$$

$$\hat{\lambda} = \bar{x}$$

$$\hat{\lambda} = 8$$

$$827537$$

$$86 \hat{\lambda} = 7$$

$$\therefore \hat{\lambda} = 8$$

Sample

Suppose number of claims in a month from a certain portfolio of policies has a Poisson distribution with mean λ .

b) In one month 8 claims were incurred and in next month 10 claims were incurred.

Find MLE of λ .

$$\begin{aligned}
 L(\lambda | x_1 = 8; x_2 = 10) &= \prod_{i=1}^2 \frac{e^{-\lambda} \cdot \lambda^{x_i}}{x_i!} \quad \boxed{\lambda = \bar{x}} \\
 L(\lambda | x_1, x_2) &= P(X=8) \cdot P(X=10) \\
 L(\lambda | x_1, x_2) &= \frac{e^{-\lambda} \cdot \lambda^8}{8!} * \frac{e^{-\lambda} \cdot \lambda^{10}}{10!} \\
 L(\lambda | x_1, x_2) &= e^{-2\lambda} \cdot \lambda^{18}
 \end{aligned}$$

$$\begin{aligned}
 \ln L(\lambda | x_1, x_2) &= -2\lambda + 18 \ln \lambda \\
 \frac{d \ln L}{d \lambda} &= 0 \Rightarrow -2 + \frac{18}{\lambda} = 0 \\
 \Rightarrow -2\lambda + 18 &= 0 \\
 \Rightarrow 2\lambda &= 18 \\
 \Rightarrow \lambda &= 9 \\
 \therefore \text{MLE } \lambda &\rightarrow \lambda' = 9 //
 \end{aligned}$$

Maximum Likelihood Estimator: Problems

Let X_1, X_2, \dots, X_n be a random sample from $N(\mu, 1)$ population. Find MLE of μ .

$$X_i \sim N(\mu, \sigma^2) \quad \text{pdf } f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\text{Given } \sigma^2 = 1 \quad X_i \sim N(\mu, 1) \quad f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2}} \quad \therefore \text{MLE of } \mu \text{ is}$$

$$L(\mu, 1 | x_1, \dots, x_n) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(x_i - \mu)^2}{2}}$$

$$L(\mu, 1 | x_1, \dots, x_n) = \prod_{i=1}^n \frac{1}{2} e^{-\frac{(x_i - \mu)^2}{2}}$$

$$\ln L(\mu, 1 | x_1, \dots, x_n) = -\frac{1}{2} \sum_{i=1}^n (x_i - \mu)^2$$

$$\left. \begin{aligned} \frac{d \ln L}{d \mu} &= 0 \Rightarrow \\ &\Rightarrow -\frac{1}{2} + 2 \sum_{i=1}^n (x_i - \mu) = 0 \\ &\Rightarrow \sum_{i=1}^n (x_i - \mu) = 0 \\ &\Rightarrow \sum_{i=1}^n x_i - n\mu = 0 \\ &\Rightarrow n\mu = \sum x_i \Rightarrow \mu = \frac{\sum x_i}{n} \end{aligned} \right\}$$

Maximum Likelihood Estimator: Problems

Let x_1, x_2, \dots, x_n be a random sample from a $N(\mu, \sigma^2)$ population. Find MLE of σ .

Example - Poisson Distribution

Problem:

The following data are the observed frequencies of occurrence of domestic accidents: we have $n = 647$ data as follows

Number of Accidents	Frequency
0	447
1	132
2	42
3	21
4	3
5	2

Example - Poisson Distribution

Solution:

$$\lambda = \frac{1}{n} \sum_{i=1}^n x_i = \bar{X}$$

The MLE of λ is $\hat{\lambda} = \bar{X}$

$$\begin{aligned}\hat{\lambda} &= \frac{1}{n} \sum_{i=1}^n x_i = \bar{X} \\ &= \frac{(447 * 0) + (132 * 1) + (42 * 2) + (21 * 3) + (3 * 4) + (2 * 5)}{674} \\ &= 0.465\end{aligned}$$

Example – Maximum Likelihood Estimate for Mean (μ)**Problem:**

A random sample of 10 weights (in pounds) of Annie's classmates are given as 115 122 130 127 149 160 152 138 149 180.

Solution:

We know that, the MLE of μ is $\hat{\mu} = \bar{X}$

$$\hat{\mu} = \sum_{i=1}^n \frac{x_i}{n} = \bar{X}$$

$$\begin{aligned}\hat{\mu} &= \frac{1}{10} (115 + 122 + 130 + 127 + 149 + 160 + 152 + 138 + 149 + 180) \\ &= 142.2\end{aligned}$$



THANK YOU

D. Uma

Computer Science and Engineering

umaprabha@pes.edu

+91 99 7251 5335