# Statistical Inference and Multivariate Analysis (MA324)

## Lecture Slides
## Lecture 13

Maximum Likelihood Estimator

Indian Institute of Technology Guwahati

Jan-May 2023

# Maximum Likelihood Estimator (MLE)

The MLE was first proposed by R. A. Fisher in 1912. This is one of the most popular methods of estimation. Let us start with an example.

Example 1: Let a box has some red balls and some black balls. It is known that number of black balls to red balls is in 1:1 or 1:2 ratio. Let two balls are drawn randomly and with replacement from the box. We want to find whether it is 1:1 or 1:2 based on that data (drawn two balls).

# Maximum Likelihood Estimator

Motivated by the previous example, we have following definitions.

Def: [**Likelihood Function**] Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be a RS from a population with PMF/PDF $f(x; \boldsymbol{\theta})$, where $\boldsymbol{\theta} \in \Theta$. The function

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

considered as a function of $\boldsymbol{\theta} \in \Theta$ for any fixed $\boldsymbol{x} \in \mathcal{X}$ ($\mathcal{X}$ is support of the RS, which is also called sample space of the RS), is called the likelihood function.

Def: [**Maximum Likelihood Estimator**] For a sample point $\boldsymbol{x} \in \mathcal{X}$, let $\widehat{\boldsymbol{\theta}}(\boldsymbol{x})$ be a value in $\Theta$ at which $L(\boldsymbol{\theta}, \boldsymbol{x})$ attains its maximum as a function of $\boldsymbol{\theta}$, with $\boldsymbol{x}$ held fixed. Then MLE of the parameter $\boldsymbol{\theta}$ based on a RS $\boldsymbol{X}$ is $\widehat{\boldsymbol{\theta}}(\boldsymbol{X})$.

# Maximum Likelihood Estimator

- Unlike MME, by definition, MLE **always lies in the parametric space**.
- Moreover, the problem of finding MLE boils down to **finding maxima** of likelihood function.
- For finding maxima, we can **use any method** that is applicable for a particular problem.
- For example, if $L\left(\boldsymbol{\theta},\, \boldsymbol{x}\right)$ is **twice differentiable**, then one can find $\widehat{\boldsymbol{\theta}}$ using simple calculus. In regular cases, we can equivalently **maximize the log-likelihood** function $l\left(\boldsymbol{\theta},\, \boldsymbol{x}\right) = \log L\left(\boldsymbol{\theta},\, \boldsymbol{x}\right)$, as $\log(\cdot)$ is an **strictly increasing** function.
- In such cases, we can **find MLE by solving**

  $$\frac{\partial}{\partial \boldsymbol{\theta}} l\left(\boldsymbol{\theta},\, \boldsymbol{x}\right) = \frac{\partial}{\partial \boldsymbol{\theta}} \log L\left(\boldsymbol{\theta},\, \boldsymbol{x}\right) = 0$$

  simultaneously. The above equation is called **likelihood equation**.
- Now onwards, for brevity, we will write $L(\boldsymbol{\theta})$ instead of $L(\boldsymbol{\theta},\, \boldsymbol{x})$ if not special emphasis is needed on $\boldsymbol{x}$. Similarly, we will use $l(\boldsymbol{\theta})$.

# Examples:

Example 2: Let $X_1, X_2, \ldots, X_n \overset{i.i.d.}{\sim} P(\lambda)$, where $\lambda > 0$. Find the MLE of $\lambda$.

Example 3: Let $X_1, X_2, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu, 1)$, $\mu \in \mathbb{R}$. Find the MLE of $\mu$.

Example 4: Let $X_1, X_2, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$, where $\mu \in \mathbb{R}$ and $\sigma > 0$. Find the MLE of $\mu$ and $\sigma^2$.

Example 5: Let $X_1, X_2, \ldots, X_n \overset{i.i.d.}{\sim} N(0, \sigma^2)$, where $\sigma > 0$. Find the MLE of $\sigma^2$.

# Remark:

Remark: Note that the estimator of $\sigma^2$ are different in the last two examples. It shows that the MLE may update itself based on any available information on the parameters.