CHAPTER 10-Part 1

ESTIMATION THEORY

Two types of approaches: Point estimation where a unique value is being assigned to estimate the parameter, and, interval estimation in which a bound on the value of the parameter is assigned with some precision.

Point Estimation
Let \( X_1, \ldots, X_n \) random variables and let the numerical representations of these random variables \( X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n \) be the experimental values.

Given \( X \) having a known probability density function which depends on unknown parameter \( \theta \) have any value in set \( \Omega \), that is,

\[
f(x; \theta), \theta \in \Omega
\]

and a family of distributions varying with respect to the value of \( \theta \)

\[
\{ f(x; \theta), \theta \in \Omega \}
\]

Select precisely one member of the family

We define a statistic \( Y_1 = u_1(X_1, \ldots, X_n) \) such that \( y_1 = u_1(x_1, \ldots, x_n) \) is a good point estimate of \( \theta \)

**DEFINITION** Let \( X_1, \ldots, X_n \) be a random sample of size \( n \) form a population with p.d.f \( f(x; \theta) \) where \( \theta \) is unknown. The joint p.d.f of \( X_1, \ldots, X_n \) regarded as a function of \( \theta \) is called the **likelihood function** of the random sample.

\[
L(\theta; X_1, \ldots, X_n) = f(X_1, X_2, \ldots, X_n; \theta) = \prod_{i=1}^{n} f(X_i; \theta) \quad \theta \in \Omega
\]
1. Method of Moments Estimation (MME)

**DEFINITION** If the population has r parameters, the method of moments consists of the solution of the system of equations

\[ \mu_k = \frac{\sum_{i=1}^{n} X_i^k}{n} ; k = 1, 2, \ldots, r , \quad \text{where} \quad \mu_k = E[X^k] \quad k = 1, 2, \ldots, r . \]

**Example:**
Let \( X_1, X_2, \ldots, X_n \) be a random sample having \( \text{Poi}(\theta) \), where \( \theta \) is unknown. Find the MME of \( \theta \).

Here r=1, \( \mu_1 = \frac{\sum_{i=1}^{n} X_i}{n} = \bar{X} \).

Equating \( \mu_1 = E[X] \Rightarrow \bar{X} = \theta \Rightarrow \hat{\theta}_{\text{MME}} = \bar{X} \)

**Example:**
Let \( X_1, X_2, \ldots, X_n \) be a random sample having \( \text{Bernoulli}(n, \theta) \), where \( \theta \) is unknown. Find the MME of \( \theta \).

Here r=1, \( \mu_1 = \frac{\sum_{i=1}^{n} X_i}{n} = \bar{X} \).

Equating \( \mu_1 = E[X] \Rightarrow \bar{X} = \theta \Rightarrow \hat{\theta}_{\text{MME}} = \bar{X} \)

**Example:**
Let \( X_1, X_2, \ldots, X_n \) be a random sample having \( \text{Ga}(\alpha, \beta) \), where \( \alpha, \beta \) are unknown. Find the MME of \( \theta = (\alpha, \beta) \).

Here r=2, \( \mu_1 = \frac{\sum_{i=1}^{n} X_i}{n} = \bar{X} \) and \( \mu_2 = \frac{\sum_{i=1}^{n} X_i^2}{n} \).

\[ \mu_1 = E[X] \Rightarrow \bar{X} = \alpha \beta \]

\[ \mu_2 = E[X^2] \Rightarrow \frac{\sum_{i=1}^{n} X_i^2}{n} = \alpha \beta ^2 \]

\[ \hat{\alpha}_{\text{MME}} = \frac{n\bar{X}^2}{\sum X_i^2 - n\bar{X}^2} ; \quad \hat{\beta}_{\text{MME}} = \frac{\sum X_i^2 - n\bar{X}^2}{n\bar{X}} \]
2. Maximum Likelihood Estimator (MLE)

**DEFINITION** Let \( U(X_1, X_2, \ldots, X_n) \) be a statistic, if the likelihood of the statistic equals or exceeds the likelihood of the random sample, i.e.

\[
L(U(X_1, X_2, \ldots, X_n); X_1 \ldots X_n) \geq L(\theta; X_1, X_2 \ldots X_n)
\]

then \( U(X_1, X_2, \ldots, X_n) \) is called the MLE of \( \theta \) Estimator.

In notation: \( \hat{\theta}_{MLE} = \text{the MLE of } \theta \)

Check the conditions for minimum:

\[
\frac{\partial L(\theta; x)}{\partial \theta} = 0 \quad \text{and} \quad \frac{\partial^2 L(\theta; x)}{\partial \theta^2} < 0
\]

Example:
Let \( X_1, X_2, \ldots, X_n \) be a random sample having \( \text{Poi}(\theta) \), where \( \theta \) is unknown.
Find the MLE of \( \theta \).

\[
f(x; \theta) = \frac{e^{-\theta} \theta^x}{x!}
\]

\[
L(\theta; x) = f(x_1; \theta) f(x_2; \theta) \ldots f(x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta).
\]

\[
L(\theta; x) = \frac{e^{-\theta} \theta^x}{x!} \cdot \frac{e^{-\theta} \theta^{x-1}}{x!} \cdots \frac{e^{-\theta} \theta^{x-n+1}}{x!} = \frac{e^{-n\theta} \theta^{\sum x_i}}{\prod x!}
\]

Taking the logarithm of likelihood function simplifies the function to take derivative.

\[
l(\theta) = \log(L(\theta; x)) = \log(e^{-n\theta} \theta^{\sum x_i} / \prod x!) = -n\theta + \theta \sum x_i - \log(\prod x!)
\]

\[
\frac{\partial}{\partial \theta} l(\theta) = -n + \sum \frac{x_i}{\theta} - \log(\prod x!) \Rightarrow \hat{\theta}_{MLE} = \frac{\sum x_i}{n} = \bar{X}
\]

Checking if it is the minimum or not

\[
\frac{\partial^2}{\partial \theta^2} l(\theta) = - \sum \frac{x_i}{\theta^2} < 0
\]
Example:
Let \( X_1, X_2, \ldots, X_n \) be a random sample having \( \text{Bernoulli}(\theta) \), where \( \theta \) is unknown. Find the MLE of \( \theta \).

\[
f(x; \theta) = \theta^x (1-\theta)^{1-x}; \quad x = 0, 1
\]

\[
L(\theta; x) = f(x_1; \theta) f(x_2; \theta) \cdots f(x_n; \theta) = \prod_{i=1}^n f(x; \theta)
\]

\[
L(\theta; x) = \theta^x (1-\theta)^{1-x} \cdots \theta^x (1-\theta)^{1-x} = \theta \sum_{i} x_i (1-\theta)^{n-\sum_{i} x_i}
\]

Taking the logarithm of likelihood function simplifies the function to take derivative.

\[
l(\theta) = \log(L(\theta; x)) = \log(\theta) \sum x_i + n \log(1-\theta) - \log(1-\theta) \sum x_i
\]

\[
\frac{\partial}{\partial \theta} l(\theta) = \sum \frac{x_i}{\theta} - n - \sum \frac{x_i}{1-\theta} = 0 \Rightarrow \hat{\theta}_{\text{MLE}} = \frac{\sum x_i}{n} = \bar{x}
\]

Checking if it is the minimum or not

\[
\frac{\partial^2}{\partial \theta^2} l(\theta) = -\sum \frac{x_i}{\theta^2} - n - \sum \frac{x_i}{(1-\theta)^2} < 0
\]

Example:
Let \( X_1, X_2, \ldots, X_n \) be a random sample having \( \text{N}(\theta_1, \theta_2) \), where \( \theta_1, \theta_2 \) are unknown. Find the MLE of \( \theta = (\theta_1, \theta_2) \).

\[
f(x; \theta) = \frac{1}{\sqrt{2\pi\theta_2}} e^{\frac{(x-\theta_1)^2}{2\theta_2}}
\]

\[
L(\theta; x) = \prod_{i=1}^n f(x_i; \theta)
\]

\[
L(\theta; x) = \frac{1}{\sqrt{2\pi\theta_2}} e^{\frac{(x_1-\theta_1)^2}{2\theta_2}} \cdots \frac{1}{\sqrt{2\pi\theta_2}} e^{\frac{(x_n-\theta_1)^2}{2\theta_2}} = \left( \frac{1}{2\pi\theta_2} \right)^{n/2} e^{-\sum \frac{(x_i-\theta_1)^2}{2\theta_2}}
\]

Taking the logarithm of likelihood function simplifies the function to take derivative.

\[
l(\theta) = \log(L(\theta; x)) = \log \left( \frac{1}{(2\pi\theta_2)^{n/2}} e^{-\sum \frac{(x_i-\theta_1)^2}{2\theta_2}} \right) = -\frac{n}{2} \log(2\pi\theta_2) - \sum \frac{(x_i-\theta_1)^2}{2\theta_2}
\]

\[
l(\theta) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\theta_2) - \sum \frac{(x_i-\theta_1)^2}{2\theta_2}
\]
\[
\frac{\partial}{\partial \theta_1} l(\theta) = \frac{2}{2\theta_2} \sum (x_i - \theta_1) = 0 \Rightarrow \hat{\theta}_{1\text{MLE}} = \frac{\sum x_i}{n} = \bar{X}
\]
\[
\frac{\partial}{\partial \theta_2} l(\theta) = -\frac{n}{2\theta_2} + \frac{\sum (x_i - \theta_2)^2}{2\theta_2^2} = 0 \Rightarrow n\theta_2 = \sum (x_i - \theta_2)^2 \Rightarrow \hat{\theta}_{2\text{MLE}} = \frac{\sum (x_i - \bar{X})^2}{n} = \frac{\sum (x_i - \bar{X})^2}{n}
\]

Checking if it is the minimum or not
\[
\frac{\partial^2}{\partial \theta_1^2} l(\theta) = -\frac{n}{\theta_2^3} < 0
\]
\[
\frac{\partial^2}{\partial \theta_2^2} l(\theta) = \frac{n}{2\theta_2^3} - \frac{2\sum (x_i - \theta_2)^2}{\theta_2^3} = n\theta_2 < 2\sum (x_i - \theta_2)^2 \Rightarrow \frac{\partial^2}{\partial \theta_2^2} l(\theta) < 0
\]
\[
\frac{\partial^2}{\partial \theta_1 \theta_2} l(\theta) = -\frac{\sum (x_i - \theta_1)}{\theta_2^3} < 0
\]

Invariance Property of MLE
**Theorem:** Given \( \hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_k) \) MLE of \( \theta = (\theta_1, \ldots, \theta_k) \), let \( Z_1(\theta) \ldots Z_k(\theta) \) be the one-to-one functions of \( \hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_k) \). The MLE of functions of \( \theta = (\theta_1, \ldots, \theta_k) \) \( Z = (Z_1(\theta) \ldots Z_k(\theta)) \) is \( \hat{Z} = (\hat{Z}_1(\theta) \ldots \hat{Z}_k(\theta)) \) by invariance property of MLE.

**Example:**
Let \( X_1, X_2 \ldots X_n \) be a random sample having \( N(\theta_1, \theta_2) \), where \( \theta_1, \theta_2 \) are unknown. Find the MLE of \( \tau(\theta_2) = \sqrt{\theta_2} \).

Given the MLE’s of \( \theta_1, \theta_2 \) and the theorem above, the MLE of \( \tau(\theta_2) = \sqrt{\theta_2} \) is
\[
\sqrt{\theta_{2\text{MLE}}} = \sqrt{\frac{\sum (x_i - \hat{\theta}_1)^2}{n}} = \sqrt{\frac{\sum (x_i - \bar{X})^2}{n}}
\]
Example:
Let \( X_1, X_2 \ldots X_n \) be a random sample having \( U(\theta_1, \theta_2) \), where \( \theta_1, \theta_2 \) are unknown. Find the MLE of \( \theta_1, \theta_2 \).

\[
f(x; \theta) = \frac{1}{\theta_2 - \theta_1}; \quad \theta_1 < x < \theta_2
\]

\[
L(\theta; x) = \prod_{i=1}^{n} f(x; \theta) = \frac{1}{(\theta_2 - \theta_1)^n}
\]

\[
l(\theta) = \log(L(\theta; x)) = \log\left(\frac{1}{(\theta_2 - \theta_1)^n}\right) = -n \log(\theta_2 - \theta_1)
\]

\[
\frac{\partial}{\partial \theta_1} l(\theta) = -\frac{n}{\theta_2 - \theta_1} = 0 \Rightarrow \text{no solution for } \hat{\theta}_{1\text{MLE}}
\]

\[
\frac{\partial}{\partial \theta_2} l(\theta) = -\frac{n}{\theta_2 - \theta_1} = 0 \Rightarrow \text{no solution for } \hat{\theta}_{2\text{MLE}}
\]

The supremum (maximum) of the likelihood function appears for \( \theta_1 \) when \( \theta_1 = \text{Min}(X_1, X_2, \ldots X_n) \)

The supremum (maximum) of the likelihood function appears for \( \theta_1 \) when \( \theta_1 = \text{Max}(X_1, X_2, \ldots X_n) \).

Therefore, the maximum likelihood estimators are \( \hat{\theta}_{1\text{MLE}} = Y_1, \quad \hat{\theta}_{2\text{MLE}} = Y_n \).