Chapters 9. Properties of Point Estimators
**Recap**

- Population parameter $\theta$. Population distribution $f(x; \theta)$. The form of $f(x; \theta)$ is known except the value of $\theta$.
- Sample: $\{X_1, X_2, \ldots, X_n\}$ iid with distribution $f(x, \theta)$.
- Estimator $\hat{\theta}$: a function of samples $\{X_1, X_2, \ldots, X_n\}$:
  $$\hat{\theta} = \hat{\theta}(X_1, X_2, \ldots, X_n).$$
- MSE, unbiased, confidence interval.
**Relative efficiency**

Two estimators for $\theta$: $\hat{\theta}_1$ and $\hat{\theta}_2$. The relative efficiency of $\hat{\theta}_1$ with respect to $\hat{\theta}_2$ is defined as

$$\text{eff}(\hat{\theta}_1, \hat{\theta}_2) = \frac{\text{MSE}(\hat{\theta}_2)}{\text{MSE}(\hat{\theta}_1)}$$

**Remark:** When $\hat{\theta}_1$ and $\hat{\theta}_2$ are both unbiased, their relative efficiency reduces to

$$\text{eff}(\hat{\theta}_1, \hat{\theta}_2) = \frac{\text{Var}(\hat{\theta}_2)}{\text{Var}(\hat{\theta}_1)}$$

**Remark:** When $\text{eff}(\hat{\theta}_1, \hat{\theta}_2) > (<)1$, $\hat{\theta}_1$ is more (less) efficient than $\hat{\theta}_2$. 
**Minimal Variance Unbiased Estimator (MVUE)**

**Goal:** Among all the unbiased estimators, find the one with the minimal variance (most efficient unbiased estimator).

**Keywords:**

1. Estimator: function of samples \( \{X_1, X_2, \ldots, X_n\} \)
2. Unbiased.
3. Minimal variance.
**MVUE: Sufficient Statistics**

**Definition:** A *Statistics* is a function of samples \( \{X_1, X_2, \ldots, X_n\} \).

**Definition:** A statistics \( t = t(X_1, X_2, \ldots, X_n) \) is said to be *sufficient* if the *likelihood* of samples \( \{X_1, X_2, \ldots, X_n\} \)

\[
L(x_1, x_2, \ldots, x_n; \theta) = f(x_1; \theta) \times f(x_2; \theta) \times \cdots \times f(x_n; \theta)
\]
can be written as

\[
L(x_1, x_2, \ldots, x_n; \theta) = g_\theta(t) \times h(x_1, x_2, \ldots, x_n).
\]
Examples of sufficient statistics

1. Bernoulli Distribution. \( \{X_1, X_2, \ldots, X_n\} \) iid Bernoulli with parameter \( p \) (target parameter). Then

\[
\sum_{i=1}^{n} X_i
\]

is sufficient.
2. Poisson Distribution. \( \{X_1, X_2, \ldots, X_n\} \) iid Poisson with parameter \( \lambda \) (target parameter). Then

\[
\sum_{i=1}^{n} X_i
\]

is sufficient.
3. **Uniform Distribution.** \( \{X_1, X_2, \ldots, X_n\} \) iid uniform on interval \([0, \theta]\) (target parameter \(\theta\)). Then

\[
X_{(n)} = \max\{X_1, X_2, \ldots, X_n\}
\]

is sufficient.
4. Normal Distribution. \( \{X_1, X_2, \ldots, X_n\} \) iid \( N(\mu, \sigma^2) \).

(a) Suppose \( \sigma \) is known, and \( \mu \) is the target parameter. Then
\[
\sum_{i=1}^{n} X_i
\]

is sufficient.

(b) Suppose \( \mu \) and \( \sigma \) are both unknown (target parameters). Then
\[
\left( \sum_{i=1}^{n} X_n, \sum_{i=1}^{n} X_i^2 \right)
\]

are (jointly) sufficient.
Remark: Sufficient statistics are not unique. Many of them.

Remark: What is the meaning of “sufficiency” — A sufficient statistics contains all the information about $\theta$ from the samples $\{X_1, X_2, \ldots, X_n\}$.

The conditional distribution of $\{X_1, X_2, \ldots, X_n\}$ given a sufficient statistics $t = t(X_1, X_2, \ldots, X_n)$ does NOT depend on $\theta$.

Verify the discrete case ....
**MVUE: Rao-Blackwell Theorem**

**Theorem:** Let \( \hat{\theta} = \hat{\theta}(X_1, X_2, \ldots, X_n) \) be an unbiased estimator for \( \theta \), and \( t \) any sufficient statistics. Define

\[
\hat{\theta}^* = E[\hat{\theta}(X_1, X_2, \ldots, X_n)|t].
\]

Then \( \hat{\theta}^* \) is an unbiased estimator for \( \theta \) and

\[
\text{Var}[\hat{\theta}^*] \leq \text{Var}[\hat{\theta}]
\]

**Remark:** \( \hat{\theta}^* \) is a function of \( t \) only.

**Observation:** If there is only one function of \( t \), say \( h(t) \), such that \( h(t) \) is an unbiased estimator for \( \theta \), that is

\[
E[h(t)] = \theta,
\]

then \( h(t) \) is the MVUE.
Definition: We say a statistics $t$ is complete if

$$E[g(t)] = 0$$

for every $\theta$ implies $g \equiv 0$.

Remark: Suppose $t$ is sufficient and complete, then there will be at most one function of $t$, say $h(t)$, that is an unbiased estimator for $\theta$. 
**MVUE: A USEFUL APPROACH**

To identify an MVUE,

1. Find a sufficient statistics, say $t$.
2. Argue this statistics is complete.
3. Find an unbiased estimator $h(t)$ for $\theta$. (One can use any unbiased estimator, say $\hat{\theta}$, and then let $h(t) = E[\hat{\theta}|t]$)
4. This estimator $h(t)$ is MVUE.
**MVUE: Examples**

1. A coin with $P(H) = p$ (target parameter). Toss coin $n$ times,

   $$X_i = \begin{cases} 
   1 & \text{if i-th toss is heads} \\
   0 & \text{if i-th toss is tails} 
   \end{cases}$$

   Identify the MVUE for $p$. 
2. Suppose \( \{X_1, X_2, \ldots, X_n\} \) are iid \( N(\mu, \sigma^2) \).

(a) If \( \sigma^2 \) is known, what is the MVUE for \( \mu \)?
(b) If \( \mu \) and \( \sigma^2 \) are both unknown, what is the MVUE for \( \mu \) for \( \sigma^2 \)?
3. Suppose \( \{X_1, X_2, \ldots, X_n\} \) are iid samples from uniform distribution on \([0, \theta]\). Find an MVUE for \( \theta \).
4. Suppose \( \{X_1, X_2, \ldots, X_n\} \) are iid samples from Poisson distribution with parameter \( \lambda \). Find an MVUE for \( \theta \). What about an MVUE for \( e^{-\theta} \)?
Maximum Likelihood Estimate (MLE)

MLE: Find $\theta$ to maximize $L(x_1, x_2, \ldots, x_n; \theta)$.

[In this maximization problem, \{x_1, x_2, \ldots, x_n\} are regarded as fixed]
1. Suppose \( \{X_1, X_2, \ldots, X_n\} \) are iid samples from Poisson distribution with parameter \( \theta \). Find the MLE for \( \theta \).
2. Suppose \( \{X_1, X_2, \ldots, X_n\} \) are iid samples from uniform distribution \([0, \theta]\). Find the MLE for \( \theta \).
3. Suppose \( \{X_1, X_2, \ldots, X_n\} \) are iid samples from \( N(\mu, \sigma^2) \).

(a) Find the MLE for \( \mu \) when \( \sigma^2 \) is known.
(b) Find the MLE for \( \mu \) and \( \sigma^2 \) when they are both unknown.
Properties of MLE

MLE has the following nice properties under mild regularity conditions.

1. MLE is a function of sufficient statistics.
2. **Consistency**: An estimator \( \hat{\theta} = \hat{\theta}(X_1, X_2, \ldots, X_n) \) is said to be *consistent* if
   \[
   \left| \hat{\theta}(X_1, X_2, \ldots, X_n) - \theta \right| \to 0
   \]
as \( n \to \infty \).
3. **Asymptotic optimality**: MLE is asymptotically normal and asymptotically most efficient.
4. **Invariance Property**: Suppose \( \hat{\theta} \) is the MLE for \( \theta \), then \( h(\hat{\theta}) \) is MLE for \( h(\theta) \) when \( h \) is a one-to-one function.