(Statistical Estimation & Sampling distribution)

- Difference between Statistics and Parameter:

BASIS FOR COMPARISON	STATISTIC	PARAMETER
Meaning	Statistic is a measure which describes a fraction of population.	Parameter refers to a measure which describes population.
Numerical value	Variable and Known	Fixed and Unknown
Statistical Notation	$\bar{x}$ = Sample Mean	μ = Population Mean
	s = Sample Standard Deviation	σ = Population Standard Deviation

- Point estimate

A **point estimate** of an unknown parameter θ is a statistic $\hat{\theta}$ that represents a "best guess" at the value of θ. There may be more than one sensible point estimate of a parameter.

- Properties of a good point estimate:
  1) Unbiased Estimates:

A point estimate $\hat{\theta}$ for a parameter θ is said to be **unbiased** if

$$E(\hat{\theta}) = \theta$$

Unbiasedness is a good property for a point estimate to possess. If a point estimate is not unbiased, then its **bias** can be defined to be

$$\text{bias} = E(\hat{\theta}) - \theta$$

2) Consistency: $lim_{n \to \infty} V(\hat{\boldsymbol{\theta}}) = 0$
3) Efficiency: $Var(\hat{\boldsymbol{\theta}}) \leq Var(\theta^*)$
4) Sufficiency: For making inference about an unknown θ, the statistician makes a reduction of the data by using a statistic, if no information about θ is lost then this statistic is called sufficient estimator.

- **Factorization theorem of Fisher and Neyman:**
  Instead of checking sufficiency via the definition it is more convenient to make use of the following theorem which gives a necessary and sufficient condition for a statistic to be sufficient.

$$f(x_1;\theta)f(x_2;\theta)\dots f(x_n;\theta) = g(t(x_1,\dots,x_n);\theta).h(x_1,\dots,x_n)$$

where g is a nonnegative function depending on θ and x_1, . . . , x_n and h is a nonnegative function, not depending on θ.

Classwork

$X_1, ..., X_m$ **be a random sample from a binomial distribution with parameter θ. Find a sufficient statistic for the parameter θ.**

Homework:

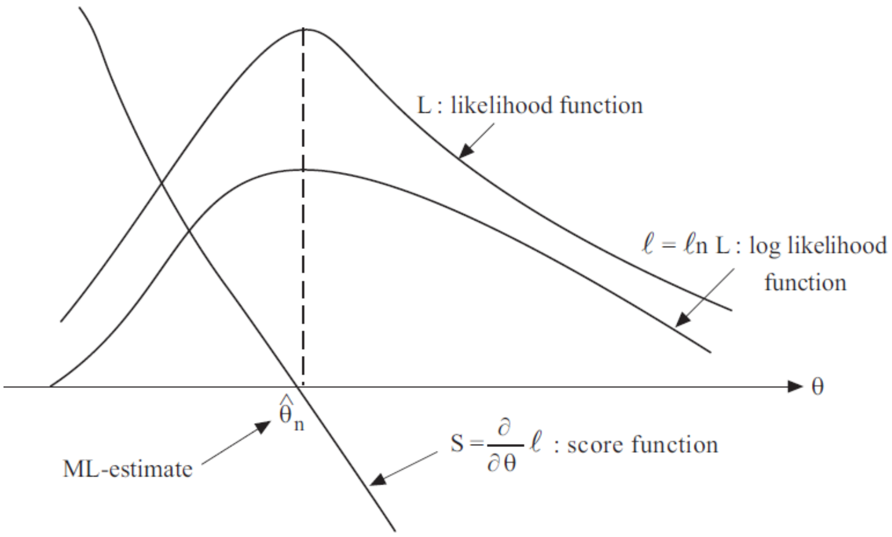
1) $X_1, ..., X_n$ be a random sample from a bernoulli distribution with parameter θ. Find a sufficient statistic for the parameter θ.
2) $X_1, ..., X_n$ be a random sample from an exponential distribution with parameter θ. Find a sufficient statistic for the parameter θ.
3) $X_1, ..., X_n$ be a random sample from a Poisson distribution with parameter θ. Find a sufficient statistic for the parameter θ.

- **Methods of point estimate**
  1) Maximum Likelihood Method (MLE)
  2) Method of moments (MM)
  3) Least Squares Method (LSM)

1) Maximum Likelihood Method (MLE) Let $(X_1, X_2, ..., X_n)$ be a random sample from a population distribution f(x; θ). The likelihood function is:
L(θ; $X_1, ..., X_n$) = f (X_1;θ)f (X_2;θ)...f (X_n;θ) = $\prod_{i=1}^{\infty} f(x_i;\theta)$

The basic idea behind this form of the method is to:

1. **Take logarithm of the likelihood function.**
2. **Derivate the log likelihood function with respect to parameters, and setting to 0.**
3. **Solve the equations to estimate the parameters.**
4. **To verify that we indeed obtain a maximum value of the parameters, we need to take the second order derivative of the log likelihood with respect to parameters.**



The maximum likelihood principle says that, out of all the possible values for θ, the value that maximizes the likelihood of the observed data should be chosen.

Advantage: The advantages of this method are:

1. consistent.
2. asymptotically unbiased and asymptotically the most efficient.

Disadvantage: The disadvantages of this method are:

1. The likelihood equations need to be specifically worked out for a given distribution.
2. Estimates can be heavily biased for small samples.
3. Can be sensitive to the choice of starting values.

- Classwork
  $X_1, ..., X_m$ be a random sample from a binomial distribution with parameter θ. Find MLE for the parameter θ. Show that this estimator is an unbiased and consistent estimator.

  Homework:

  1) $X_1, ..., X_n$ be a random sample from a bernoulli distribution with parameter θ. Find MLE for the parameter θ. Show that this estimator is an unbiased and consistent estimator.
  2) $X_1, ..., X_n$ be a random sample from an exponential distribution with parameter θ. Find MLE for the parameter θ. Show that this estimator is an unbiased and consistent estimator.

3) $X_1, ..., X_n$ **be a random sample from a Poisson distribution with parameter** θ**. Find MLE for the parameter** θ**. Show that this estimator is an unbiased and consistent estimator.**