# Statistical Inference.

- Drawing conclusions from data that has "random" variations.

## Frequentist Approach.

- $\theta$, unknown quantity
  , just a number.

Goal is to estimate this number, by collecting statistics on data.

Frequentist is all we will consider in this class.

## Bayesian Approach

$\Theta$ - unknown RV.

Goal is to update RV after making some initial guess, using Bayes rule and the data collected.

---

## Random Sampling and Point Estimators.

Def: A random sample is

$X_1, X_2, \dots, X_n$ - i.i.d.

$\mathbb{E}X_i = \mu < \infty$ , $Var(X_i) = \sigma^2 < \infty$.

**Note:** In many cases $\mu$, $\sigma^2$ are unknown.
and distribution is unknown.

**Goal** - Estimate various properties of a distribution using the "data" from a random sample

$X_1, X_2, \ldots, X_n$.

$\theta$ - parameter or quantity to be estimated.

**Ex** - $\theta = \mu$, "population mean"

$N(\theta, \sigma^2)$, $\text{Exponential}(1/\theta)$, $\text{Poisson}(\theta)$.

- $\theta = \sigma^2$ - variance.

- $\theta$ - range of a distribution.

$\text{Uniform}(0, \theta)$ $\rightsquigarrow$ $\theta$ is the upper bound.

# Point Estimators.

Given a random sample

$$X_1, X_2, X_3, \ldots, X_n$$

**Def**: An point estimator of $\Theta \in \mathbb{R}$ is a function of the data.

$$\hat{\Theta} = h(X_1, X_2, \ldots, X_n).$$

**Note** $\hat{\Theta}$ is a Random Variable

**Ex** Sample mean.

Suppose $X$ is an RV with $\mathbb{E}X = \mu$, $Var(X) = \sigma^2$

We want to estimate $\Theta = \mu$.

Let $X_1, X_2, \ldots, X_n$ be a random sample

$$\hat{\Theta} = \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

We know

① $\mathbb{E}\hat{\Theta} = \mathbb{E}\bar{X} = \mu = \Theta$

② By WLLN $\hat{\Theta} \xrightarrow{P} \Theta$ $n \to \infty$.

How do we decide how good an estimator is?

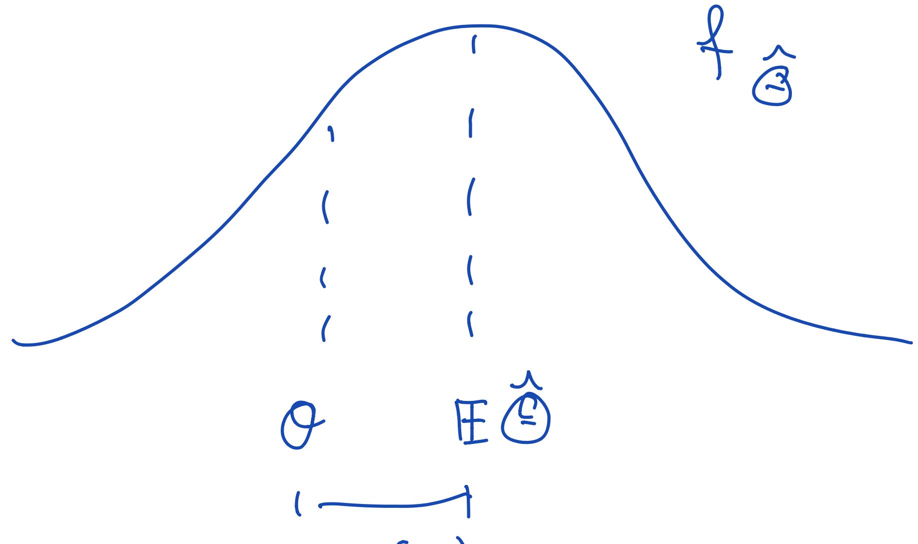
Bias, Mean square error.

Def The bias $B(\hat{\Theta})$ is

$$B(\hat{\Theta}) = \mathbb{E}(\hat{\Theta} - \Theta) = \mathbb{E}\hat{\Theta} - \Theta.$$

We say $\hat{\Theta}$ is unbiased if $B(\hat{\Theta}) = 0$.

$$B(\hat{\Theta}) = 0 \iff \mathbb{E}\hat{\Theta} = \Theta.$$



$B(\hat{\Theta})$.

Ex $X_1, X_2, \ldots X_n$ random sample.

$$\mathbb{E}X_i = \Theta$$

$$\hat{\Theta}_1 = X_1 \quad , \quad \hat{\Theta}_2 = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Both are unbiased.

## Mean square error

$$MSE(\hat{\Theta}) = \mathbb{E}\underbrace{(\hat{\Theta} - \Theta)}_{\text{Estimation error}}{}^2$$

Ex $\hat{\Theta}_1 = X_1$, $\hat{\Theta}_2 = \bar{X}$

$$MSE(\hat{\Theta}_1) = \mathbb{E}(\hat{\Theta}_1 - \underset{\mathbb{E}\hat{\Theta}_1}{\underset{\parallel}{\Theta}})^2 = Var(\hat{\Theta}_1) = Var(X_1) = \sigma^2$$

$$MSE(\hat{\Theta}_2) = Var(\hat{\Theta}_2) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{\sigma^2}{n}$$

$$MSE(\hat{\Theta}_2) < MSE(\hat{\Theta}_1) \quad n \geq 1.$$

$\hat{\Theta}_2$ is the better estimator.

## What if $\hat{\Theta}$ is not unbiased.

### Useful formula:

$$MSE(\hat{\Theta}) = Var(\hat{\Theta}) + (B(\hat{\Theta}))^2$$

Proof: $X = \hat{\Theta} - \Theta$

$$\mathbb{E}X = B(\hat{\Theta}) \quad , \quad Var(X) = Var(\hat{\Theta} - \theta)$$
$$= Var(\hat{\Theta}).$$

$$MSE(\hat{\Theta}) = \mathbb{E}X^2 = Var(X) + (\mathbb{E}X)^2$$
$$= Var(\hat{\Theta}) + (B(\hat{\Theta}))^2.$$

## Consistent Estimator

Def A sequence of estimators

$\hat{\Theta}_1, \hat{\Theta}_2, \hat{\Theta}_3, \ldots$

is consistent if.

$$\hat{\Theta}_n \xrightarrow{P} \Theta \quad n \to \infty$$

or

"$\lim_{n\to\infty} P(|\hat{\Theta}_n - \Theta| > \varepsilon) = 0 \quad \forall \varepsilon > 0.$"

<u>Ex</u> $\hat{\Theta}_n = \bar{X}_n = \frac{1}{n}\sum_{i=1}^{\infty} X_i.$

$\hat{\Theta}_n$ is consistent by WLLN.

<u>Consistency by MSE → 0.</u>

by Chebyshev.

$$P(|\hat{\Theta}_n - \Theta| > \varepsilon) \le \frac{MSE(\hat{\Theta}_n)}{\varepsilon}.$$

If.

$MSE(\hat{\Theta}_n) \to 0$ then $\hat{\Theta}_n$ is consistent.

# Practical Examples of estimators

① $\theta = \mu$ of some value in a population.

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

$B(\hat{\theta}) = 0$ , $MSE(\hat{\theta}) = \frac{\sigma^2}{n}$.

② $\theta = p$ the fraction of people in a population with a certain property (survey).

$\hat{\theta} = \frac{X}{n}$ , $X$ = # of people with that property

$\sim \text{Binomial}(n, p)$.

$B(\hat{\theta}) = 0$ , $MSE(\hat{\theta}) = \frac{p(1-p)}{n} \leq \frac{1/4}{n}$

$\hat{\theta}$ is also consistent.

③ $\Theta$ is the upper bound in Uniform $(0, \theta)$.

$$\hat{\Theta} = \max\{X_1, X_2, X_3, \ldots, X_n\}.$$

order statistic.

$$\hat{\Theta} \leq \Theta. \implies \text{Biased}.$$

<u>Estimating</u> <u>Variance</u>

$$\sigma^2 = Var(X) = \mathbb{E}(X-\mu)^2$$

$$= \mathbb{E}Y \quad , \quad Y = (X-\mu)^2$$

An estimator of $\sigma^2$ is:

$$\hat{\Theta} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2 \qquad B(\hat{\Theta}) = 0$$

$$\mathbb{E}\hat{\Theta} = \frac{1}{n}\sum_{i=1}^{n}\underbrace{\mathbb{E}(X_i - \mu)^2}_{=\sigma^2} = \sigma^2$$

we don't know this!!

Replace $\mu$ with $\bar{X}$.

$$\bar{S}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

sample mean

However, we can show,

$$\mathbb{E}\,\bar{S}^2 = \sigma^2 \left(1 - 1/n\right)$$

$$B(\bar{S}^2) = -\frac{\sigma^2}{n},$$

"Undershoots the value"

$$\boxed{S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 \longrightarrow \text{unbiased}}$$