

# Chapter 7. Point Estimation.

$$X_1, \dots, X_n \stackrel{iid}{\sim} f_X(x|\theta) \quad \theta \in \mathbb{R}^k$$

$$\underline{X} = (X_1, \dots, X_n) \sim \underbrace{\prod_{i=1}^n f_X(x_i|\theta)}_{f_{\underline{X}}(\underline{x}|\theta)} = f_{\underline{X}}(\underline{x}|\theta)$$

where  $\underline{x} = (x_1, \dots, x_n)$

Goal:

- Estimate  $\theta$
- Estimate  $\tau(\theta)$  where  $\tau: \mathbb{R}^k \rightarrow \mathbb{R}^q$  ( $q \leq k$ )

Definition: A Point Estimator of  $\theta$  for  $\tau(\theta)$   
 $\tau(\hat{\theta})$   
denoted by  $\hat{\theta}$ , which is  $W(\underline{X}) = W(X_1, \dots, X_n)$   $\hat{\tau(\theta)}$   
a function of the sample  $\underline{X}$ .  $\nwarrow$  Estimator

How to find  $\hat{\theta}$ ?

Estimate:  $W(X_1, \dots, X_n)$   
observed!

- Method of Moments (MoM)
- Maximum Likelihood Estimation (MLE)
- Bayesian Estimation

• MoM,

WLLN: iid r.v.  $Y_i, i=1, \dots, n \sim Y$  [Var(Y) < ∞]

$$\frac{1}{n} \sum Y_i \xrightarrow{P} E[Y]$$

$X_1, \dots, X_n \stackrel{iid}{\sim} f_X(x|\theta) \quad \theta = (\theta_1, \dots, \theta_k) \in \mathbb{R}^k$

$$E[X] = g_1(\theta_1, \dots, \theta_k)$$

$$E[X^2] = g_2(\theta_1, \dots, \theta_k)$$

⋮

$$E[X^k] = g_k(\theta_1, \dots, \theta_k)$$

$$N(\mu, \sigma^2) \quad \theta = \begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}$$

$$E[X] = \mu = g_1(\theta) = g_1(\mu, \sigma^2) \xleftarrow{P} \frac{1}{n} \sum X_i$$

$$E[X^2] = \sigma^2 + \mu^2 = g_2(\mu, \sigma^2) \xleftarrow{P} \frac{1}{n} \sum X_i^2$$

$$\text{Set: } \begin{cases} \frac{1}{n} \sum X_i = g_1(\mu, \sigma^2) = \mu \\ \frac{1}{n} \sum X_i^2 = g_2(\mu, \sigma^2) = \sigma^2 + \mu^2 \end{cases}$$

Solve it for  $\mu$  and  $\sigma^2$

$$\hat{\mu} = \bar{X}_n$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum X_i^2 - \bar{X}_n^2$$

} MoM Estimator of  $\mu$  and  $\sigma^2$

not unique

$$E[X^2] = \sigma^2 + \mu^2 \leftarrow \frac{1}{n} \sum X_i^2$$

$$E[X^3] = \dots \leftarrow \frac{1}{n} \sum X_i^3$$

---

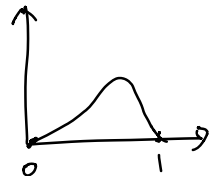
Maximum Likelihood Estimation!

10 flip      8 heads      2 tails

$p$ : Prob of a head

$P(\text{observing the data}) = \binom{10}{8} p^8 (1-p)^2$

$L(p)$



which  $\hat{p}$  is a better guess of  $p$

$$\hat{p} = 0.1 \quad \hat{p} = 0.5 \quad \hat{p} = 0.8 \quad \hat{p} = 0.9$$

---

Likelihood function.  $X_1, \dots, X_n \sim f_X(x|\underline{\theta}) \quad \underline{\theta} \in \mathcal{R}^k$

$$L(\underline{\theta}) = \prod_{i=1}^n f_X(x_i|\underline{\theta}) \quad \underline{\theta} \in \mathcal{H}$$

Maximum Likelihood Estimator.

is the  $\hat{\underline{\theta}}$  that maximizes  $L(\underline{\theta})$  in  $\mathcal{H}$