

## Summary of Lecture 2

- Not covering model misspecification.
- So we assume data  $X$  comes from  $F_{\theta_0}$  or  $f_{\theta_0}$  for some *true* value  $\theta_0 \in \Theta$ .
- Possible goals of inference: point estimation, interval estimation, hypothesis testing, prediction.
- Schools of inference: broadly, Bayes, Neyman-Pearson, Likelihood (Fisher).
- First inference example: estimation of a distribution function.
- Data  $X_1, X_2, \dots$  independent and identically distributed real random variables.
- Model: the common distribution function  $F$  can be any cdf.  
*Nonparametric.*
- $F_0$  is true cdf.



## Summary of Lecture 2

- An estimator  $\hat{\theta}$  is a function of the data  $\hat{\theta}(X)$  taking values in  $\Theta$ .
- We judge the quality of  $\hat{\theta}$  by studying the distribution of the estimation error  $\hat{\theta} - \theta_0$ .

- The *mean* of  $\hat{\theta}$  is

$$E_{\theta_0}(\hat{\theta}) = \int \hat{\theta}(x) f_{\theta_0}(x) dx.$$

- The *bias* of  $\hat{\theta}$  is

$$\text{bias}_{\theta_0}(\hat{\theta}) = E_{\theta_0}(\hat{\theta}) - \theta_0.$$

- The *variance* of  $\hat{\theta}$  is

$$\text{Var}_{\theta_0}(\hat{\theta}) = E_{\theta_0}[(\hat{\theta} - E_{\theta_0}(\hat{\theta}))^2] = E_{\theta_0}[\hat{\theta}^2] - (E_{\theta_0}(\hat{\theta}))^2.$$

- The *mean squared error* of  $\hat{\theta}$  is

$$E_{\theta_0}[(\hat{\theta} - \theta_0)^2] = \text{Var}_{\theta_0}(\hat{\theta}) + [\text{bias}_{\theta_0}(\hat{\theta})]^2.$$



## Summary of Lecture 2

- Basic properties of mean, variance, standard deviation.
- The mean of a random variable  $X$  is its expected value  $\mu \equiv E(X)$ .
- The variance of  $X$  is

$$\begin{aligned}\text{Var}(X) &= E[(X - \mu)^2] \\ &= E[X^2 - 2\mu X + \mu^2] \\ &= E(X^2) - 2\mu E(X) + \mu^2 \\ &= E(X^2) - \mu^2.\end{aligned}$$

- The standard deviation of  $X$  is  $\sqrt{\text{Var}(X)}$ .
- Units matter: the standard deviation and the mean are measured in the same units as  $X$ .



- In the non parametric model – estimate probabilities by sample proportions, means by sample averages.
- So estimate  $F(x)$  by proportion of sample less than or equal to  $x$ .
- Estimate  $F$  by

$$\hat{F}_n(x) = \frac{\# X_i \leq x}{\text{sample size}} = \frac{1}{n} \sum_{i=1}^n 1(X_i \leq x)$$

- Note that  $1(A)$  is equal to 1 if  $A$  is true and 0 otherwise.
- Called an *indicator* function by statisticians and probabilists.



## Course coverage in text

- Chapter 6, section 3.3.1.
- Review of Chapter 3.3.
- Chapter 7.1.

