STAT 5430
Lecture 02, F, Jan 24

- No new homework this week
  (assigned on Monday)
- Office hours to be announced
  Mine: FM, 12-1 PM & by appointment
  TA (Min-Yi): WR ??

Introduction to Statistical Inference

Problem Statement

- Statistical inference is about *making statements about population distributions based on samples.*

- For a collection $\mathcal{F}$ of cdf's, let $F(x) \in \mathcal{F}$ be the underlying population cdf.

  Given $X_1, \ldots, X_n$, our objective is to draw inferences about $F(x)$.

- *Definition*: If $\mathcal{F} \equiv \{F(x|\theta) : \theta \in \Theta\}$, $\Theta \subset \mathbb{R}^k$, $1 \leq k < \infty$, then the inference problem is called **parametric**; otherwise, it is nonparametric.

- Above θ is called the **parameter** and Θ is the **parameter space**.

Examples:

normal

$\mathcal{F} = \{ N(\mu, \sigma^2) : -\infty < \mu < \infty, \sigma > 0 \}$

→ parametric inference problem

$\theta = (\mu, \sigma)$ ← parameters

$\Theta = \mathbb{R} \times (0, \infty)$ ← parameter space

$\mathcal{F} = \{ F(x) :$ $F(x)$ is continuous & symmetric around zero $\}$

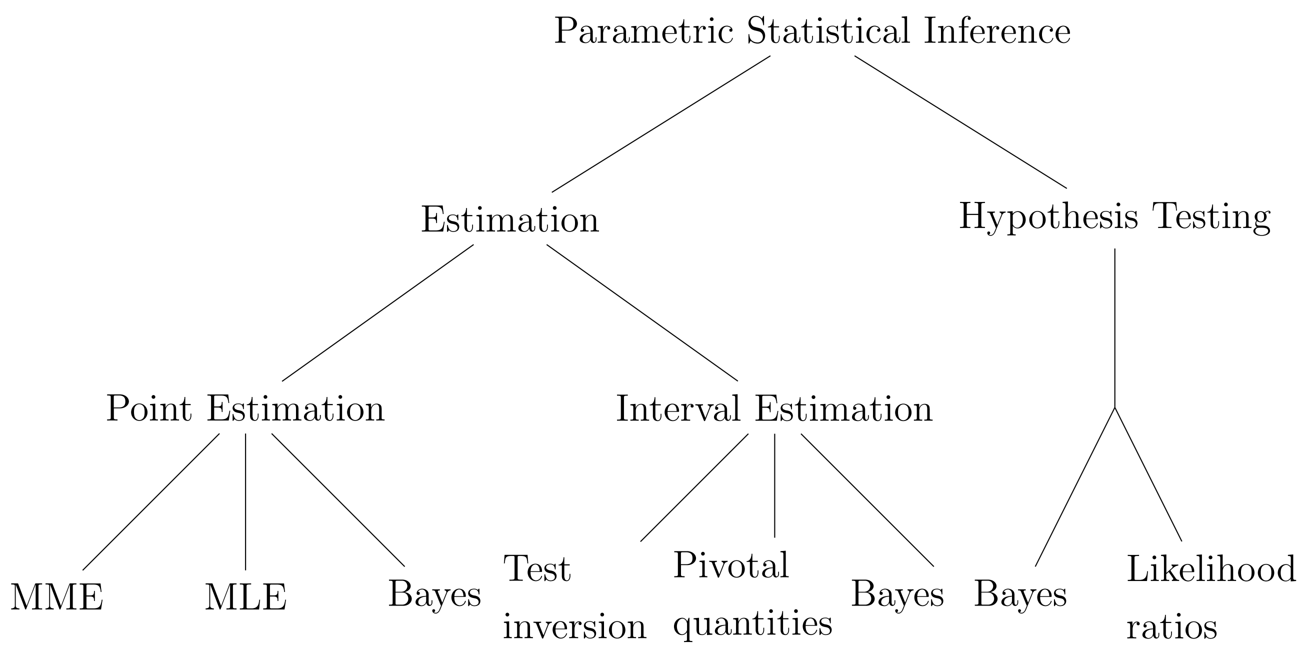
→ nonparametric inference problem

($F(x)$ is symmetric around 0 means $F(x) = 1 - F(-x)$)

Introduction to Statistical Inference

High-level Overview of STAT 5430

- We focus on *parametric* statistical inference and develop the following inference topics:



- We will answer the following types of questions:

  1. What are some strategies for finding estimators or tests?

  2. What are "good" properties of an estimator or a test?

  3. What general statistical principles exist, if any, to guarantee that we can actually find estimators/tests with good properties?

STAT 5430: Summary to date
Where we have been
& where we are headed

- Completed: Introduction to Statistical Inference
  - definitions/notation
  - random samples for inference about parametric population distributions

- Next: Point Estimation
  - Defining statistics & point estimators
  - Some strategies for point estimation
    * Method of Moments Estimation (MME)
    * Maximum Likelihood Estimation (MLE)

Point Estimation

iid from some distribution $F(x)$

Background

Definition: Let $X_1, \ldots, X_n$ be a random sample. A (Borel measurable) function of the random sample, say $T = h(X_1, \ldots, X_n)$, is called a **statistic** or an **estimator**.

(computable from data $X_1, \ldots, X_n$)

Examples:

1. $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ ← sample mean
2. $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X}_n)^2$ ← sample variance
3. $T = \begin{cases} 1 & \text{if } X_1 \le 0 \\ 0 & \text{if } X_1 > 0 \end{cases}$ statistic
4. $T_0 = (\bar{X}_n, S^2, T, X_1 + X_2^2)$ statistic
5. $T_1 = \bar{X}_n - E(X_1)$ ← not necessarily statistic (if $E(X_1)$ is unknown, T_1 is NOT statistic)

Definition: The probability distribution of a statistic T is called the **sampling distribution** of T.

("what values are possible for T & how likely those are")

Example: Suppose $X_1, \ldots, X_n$ is a r.s. from $N(\mu, \sigma^2)$.

Then, $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

$T = \sum_{i=1}^n (X_i - \bar{X}_n)^2 = (n-1)S^2 \sim \sigma^2 \chi^2_{n-1}$

since $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

Point Estimation

Background, continued

Definitions:

1. A (Borel measurable) function $\gamma : \Theta \to \mathbb{R}^d$, some $1 \le d < \infty$, is called a **parametric function**. $\gamma(\theta)$
   $\uparrow$ θ is a parameter

2. If a statistic $T = h(X_1, \ldots, X_n)$ is used to estimate $\gamma(\theta)$, then T is called an **estimator of** $\gamma(\theta)$; and the observed value $t = h(x_1, \ldots, x_n)$ is called an **estimate of** $\gamma(\theta)$.
   $\uparrow$ realized value of estimator T

Example: X_1, X_2, X_3 iid $N(\mu, \sigma^2)$

$\gamma(\mu) = \mu^2$ $\leftarrow$ parametric function

$T = h(X_1, X_2, X_3) = (\bar{X}_3)^2$ $\leftarrow$ estimator of $\gamma(\mu)$

Suppose $x_1 = 1, x_2 = 2, x_3 = 3$ are observed, then

$t = \left(\frac{1+2+3}{3}\right)^2 = 2^2 = 4$ is an estimate of $\gamma(\mu)$ (observed value of $T = (\bar{X}_3)^2$)

Some General Approaches to Point Estimation (how to get statistics or estimators)

I. Method of Moments

II. Maximum Likelihood (popular)

III. Bayes Estimators (popular)

We'll next discuss I. & II., and return to Bayes estimators at a later point.

Point Estimation

Method of Moments Estimation (MOMestimation)

Definition: Let $X_1, \ldots, X_n$ be a r.s. from pdf/pmf $f(x|\theta_1, \ldots, \theta_k)$. Then,

pop. distribution k parameters

(a) $\mathrm{E}\{(X_1)^j\} \equiv \mu_j(\theta_1, \ldots, \theta_k)$ is the j**th population moment**, $j = 1, 2, \ldots$

parametric functions

e.g. $X_1 \sim N(\mu, \sigma^2)$, $E(X_1) = \mu$

$E(X_1^2) = Var(X_1) + (EX_1)^2 = \sigma^2 + \mu^2$

(b) $\mu'_j \equiv \frac{1}{n}\sum_{i=1}^{n}(X_i)^j$ is the j**th sample moment**, $j = 1, 2, \ldots$

statistic
$j = 1, 2, 3, \ldots$

estimators based on $X_1, \ldots, X_n$

(c) The method of moments estimators (MMEs), say $\tilde{\theta}_1, \ldots, \tilde{\theta}_k$, of $\theta_1, \ldots, \theta_k$ are defined as the solution to

k parameters $\Rightarrow$ k equations

$$\left.\begin{array}{rcl} \mu_1(\tilde{\theta}_1, \ldots, \tilde{\theta}_k) & = & \mu'_1 \\ \vdots & \vdots & \vdots \\ \mu_k(\tilde{\theta}_1, \ldots, \tilde{\theta}_k) & = & \mu'_k \end{array}\right\} (*)$$

pick $\tilde{\theta}_1, \ldots, \tilde{\theta}_k$ so that pop./model moments match the sample moments

(d) The system of equations $(*)$ is called the method of moments equations (MMEquations).

Point Estimation

Method of Moments Estimation, cont'd

Example: Let $X_1, \ldots, X_n$ be a random sample from a Beta(α, β) distribution, $\alpha > 0, \beta > 0$. Find the MMEs of α, β.

Solution: $\theta_1 = \alpha$, $\theta_2 = \beta$

Then, $\mu_1(\theta_1, \theta_2) = EX_1 = \dfrac{\theta_1}{\theta_1 + \theta_2}$

& $\mu_2(\theta_1, \theta_2) = E(X_1^2) = \dfrac{(\theta_1 + 1)\theta_1}{(\theta_1 + \theta_2 + 1)(\theta_1 + \theta_2)}$

Hence, MME equations

$\mu_1(\tilde{\theta}_1, \tilde{\theta}_2) = \dfrac{\tilde{\theta}_1}{\tilde{\theta}_1 + \tilde{\theta}_2} = \mu_1' = \overline{X}_1$

$\mu_2(\tilde{\theta}_1, \tilde{\theta}_2) = \dfrac{(\tilde{\theta}_1 + 1)\tilde{\theta}_1}{(\tilde{\theta}_1 + \tilde{\theta}_2 + 1)(\tilde{\theta}_1 + \tilde{\theta}_2)} = \mu_2' = \frac{1}{n}\sum_{i=1}^{n} X_i^2$