6.434/16.391 Statistics for Engineers and Scientists

Lecture 1   09/05/2012

Laboratory for Information and Decision Systems
Massachusetts Institute of Technology

Lecture 1 09/05/2012

BACKGROUND ON STATISTICS

Motivation

- Engineers and scientists are constantly exposed to collection of facts or data both in their professional capacities as well as activities
- Population: A well defined collection of objects in the study
- Census: A complete enumeration of a population
- Sample: A subset of the population selected in a prescribed manner

Branches of statistics

- Descriptive statistics: Summarize and describe the important features of the data

- Inferential statistics: Techniques for generalizing from a sample to a population
  – We will mainly focus on inferential statistics that are useful for engineers and scientists

- Examples of inferential statistics
  – Point estimation (derivation and performance evaluation)
  – hypothesis testing
  – Confidence intervals

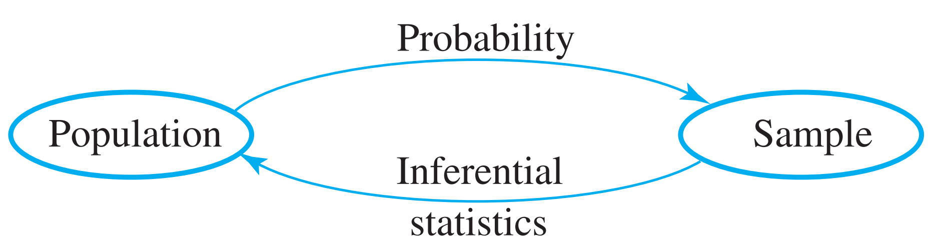
Probability

- Probability forms a bridge between the descriptive statistics and inferential techniques

- Mastery of probability leads to
  - how inferential procedures are developed and used
  - how certain statistical conclusions can be interpreted

Probability & statistics

- Probability and statistics both deal with questions involving population and samples
  - Probability problems: Properties of population is known, and questions regarding a sample taken from the population are investigated
  - Statistics problems: Characteristics of a sample are known from experiments, and conclusions regarding the population are made



The relationship between probability and inferential statistics

Example of data

- Vector
  - Experiments consisting of n individuals
  - Example: signals from output of n antennas

- Matrices or arrays
  - Digitized pictures
  - MIMO systems

Models

- "Models of course are never true. But fortunately, it is only necessary that they are useful" (George Box)

- Models give insights to engineers and scientists. Insights are very important

- All great systems are based on engineering insights provided by analysis, and we understand how and why it works. To gain insights, we have to step back from reality and build models.
  - Shannon's mathematical theory of communication, binary symmetric channel (BSC)

Our focus (1)

- Derive methods of extracting useful information from data, and give methods that infer the generality of experimental results
  - We observe a particular behavior. What extent can we expect the same effect more generally? (Estimation, testing, confidential intervals)

- Evaluate the effectiveness of our methods
  - Optimality principles
  - Compare the estimators

Our focus (2)

- Experimental design

- Algorithm issues
  - Estimators may not be derived in explicit closed-forms, but only in terms of implicit solution
  - e.g., method of bisection, steepest descent, EM algorithm

- Rather than treating as a black box and push your magic buttons, our philosophy here is to try to gain *insights*

Data & model: Example 1

- We observe $x_1,\ x_2,\ \ldots,\ x_n$

- We model the observations as independent identically distributed (i.i.d.) random variables (r.v's)
$$X_1,\ X_2,\ \ldots,\ X_n$$
with common unknown distribution F

- This model is fully specified by the set $\mathcal{F}$ of distributions which we specify

Data & model: Example 2

- Random variables $X_i = \mu + \xi_i,\ 1 \leq i \leq N$

  where $\boldsymbol{\xi} = \left[\xi_1\ \xi_2\ \cdots\ \xi_N\right]^{\mathrm{T}}$ is the vector of random errors

- Question: What should be assumed about the distribution of $\boldsymbol{\xi}$ together with μ which completely specify the joint probability density function (pdf) of $X_1,\ X_2,\ \ldots,\ X_N$?

- Answer: Of course this depends on how the experiment is carried out

i.i.d. model

- The value of error (or noise) at one instant does not affect the value of the error at other instants, i.e., $\xi_1, \xi_2, \cdots, \xi_N$ are *independent*

- The distribution of the error at one instant is the same as other instants, i.e., $\xi_1, \xi_2, \cdots, \xi_N$ are *identically* distributed

- The distribution of $\boldsymbol{\xi}$ does not depend on μ