Applied Statistics I
Math 3070-1

Spring, 2011
Jingyi Zhu,
LCB 335
zhu@math.utah.edu
Contents of the Course

• Jay L. Devore, Probability and Statistics for Engineering and the Sciences, 7th Edition
• Chapters 1-9
Subjects

- Introduction: what is statistics?
- Probability: random variables, distributions
- Point estimate
- Statistical inference: confidence interval and test of hypotheses
Course Work

- weekly homework (30%)
- 4 midterms (drop one lowest) (30%)
- comprehensive final (must take, 30%)
- SAS lab (must pass, 10%)
Chapter One: Overview and Descriptive Statistics

- What is statistics? concepts and methods
- Why do we need to study statistics?
- What do we get out of statistics?
- What kind of approaches?
What is Statistics?

- a mathematical tool that is about collecting, analyzing, interpreting/explaining, and presenting data
Why do we need statistics?

• make intelligent judgements and informed decisions

• In the presence of uncertainty and variation

• Lower the cost of repeated work
What do we get?

- Description of the data
- Nature of the phenomenon
- Effectiveness of treatments
- Draw some conclusions
Approaches

• Point estimate: simple way to present the data
• Using samples
• Confidence interval - recognizing the limitation of samples
• Hypothesis testing - about drawing conclusions with care
• Design experiments - performing experiments - use the result to help making decisions
• Making good suggestions
Examples of Experiments

- Testing a drug
- Assess the effectiveness of a treatment
- Election polls
Section 1.1

- Populations
- Samples
- Processes
Populations and Samples

- We are presented with data, only part of what we want to study: the population
- Population: a well-defined collection of objects for which we are interested in gaining some knowledge
- Sample: a subset of the population that we have access
- The whole population is usually not available for us to examine
Variables

• What aspect of the population we want to study?

• A variable is a characteristic whose value may change from one object to another in the population

• Can be quantitative or qualitative (use lowercase)

• Examples:
  • gas mileage of a particular maker
  • gender of engineering graduate students
Univariate vs. Multivariate

- **Univariate data set**: observations on a single variable
- **Multivariate data set**: multiple variables
- **Bivariate**
- **Example**: height and gender in a group
Branches of Statistics

• 1st/primitive approach: descriptive statistics - we learn how to present the data

• displays like stem-and-leaf, histogram, charts, etc.

• More sophisticated approach: inferential statistics

• use sample information to draw some conclusions
Descriptive Statistics Example

- $x = \text{temperature of the O-rings}$

```
84 49 61 40 83 67 45 66 70 69 80 58 68 60
67 72 73 70 57 70 78 52 67 53 67 75 61
70 81 76 79 75 76 58 31
```
Inferential Statistics Example

• Flexural Properties of Concrete

• data set

5.9 7.2 7.3 6.3 8.1 6.8 7.0 7.6 6.8 6.5 7.0 6.3 7.9 9.0 8.2 8.7 7.8 9.7 7.4 7.7 9.7 7.8
7.7 11.6 11.3 11.8 10.7

• Inferred:
  • with a high degree of confidence
  • the population mean strength is between 7.48 MPa and 8.80 MPa
Probability and Statistics

- Probability: study of the likelihood of certain events
- How do we use probability to ask about the sample?
- Probability: estimate how likely certain sample turns up
- Statistics: given the information in a sample, what can we say about the population?
Enumerative vs. Analytic Studies

- Enumerative studies: concrete population, samples can be made available
- Analytic studies: otherwise
- Such as taking action on a process, moving target
Collecting Data

- garbage in, garbage out
- more subtle than just good or bad
- make sure that the population is sampled objectively (unbiased):
  - simple random sampling - equally likely for each sampled
  - stratified sampling - sample from well divided groups
Collecting Data

- target population may be different from the population from which samples are drawn
- simple random sample
- stratified sampling