INTRODUCTION

Statistics: collection, processing, analysis, interpretation of numerical data

1.1 Why Statistics

1. clearly defined goals for investigation
2. what data to collect & how to collect it
3. apply statistical methods to efficiently extract information
4. interpret information & draw conclusions

1.2 Modern Statistics

- games of chance
- political science

descriptive statistics: numerical descriptions of political units

- e.g. mean, standard deviation, charts

inferential statistics: generalizations based on sample data

frequency/classical approach: statistical inferences concern fixed but unknown quantities

Data

- Discrete VS Continuous
- Quantitative VS Qualitative

Collecting Engineering Data

- **retrospective** study: using all or a sample of *historical process data*
- **observational** study: observing/recording quantities of interest of a process/population with minimal disturbance
- **designed** experiment: deliberate/purposeful change in controllable variables

1.3 Statistics & Engineering

engineer: someone who solves problems of interest to society by efficient application of scientific principles

- refining existing products
- designing new products

mechanistic model: built from knowledge of physical mechanisms that relate several variables

empirical model: built from engineering/scientific knowledge of phenomenon, but NOT directly developed from theoretical or first-principles understanding of mechanism

Notes

- **statistical techniques** are useful for describing/understanding variability (successive observations of a system with different results)
- **statistics** provides a framework for portraying variability & learning about its sources
- **variability**: encountered when a process is too complex to model/control
    - cost/time/resource constraints
    - incomplete knowledge

1.4 Role of Scientist & Engineer in QI

probability

- % of population meeting a criteria
- chance of any one observation from population meeting some criteria
- **probability models**: helps quantify risks involved in statistical inference

quality improvement: make it right the first time; do not stay content with and process/product, continuous improvement.

- **reliability**

1.5 Case Study

X-bar chart: plotting sample average VS time

1.6 Basic Concepts

unit: single entity whose characteristics are of interest

population of units: complete collection of units about which information is sought

population: set of all measurements corresponding to each unit in the entire population of units about which information is sought

- **parameter**: characteristic of population
- **variation**: changes/uncertainty in population

POPULATION	UNIT	VARIABLES/CHARACTERICS
All students in a school	student	GPA, credits, majors, etc.

sample: subset of measurements that are actually collected in an investigation

- should randomized to avoid bias
    - random number table
    - random digit dialing
- **statistic**: characteristic of sample

-