

# Statistics 571

## Statistical Methods for Bioscience I

Section I: Bret Larget  
Section II: Jun Zhu  
Department of Statistics  
University of Wisconsin–Madison

Fall 2005

### Course Information

- Read the entire syllabus carefully. Complete the survey sheet.
- Before the lectures, download the lecture notes and handouts from the course website. You are advised to attend the lecture section you are enrolled in, even though the two lecture sections are similar.
- Late homework without Instructor's prior permission will be penalized.
- You are advised to attend the discussion section that you are enrolled in. If you wish to attend a different section that is closed, you need to obtain TA's permission to attend that section.
- Block the dates and time for the exams **NOW**.
- Read "Introduction to R" on the course website.
- Policy on Auditors: Attend at least 80% of the lectures.
- Where to get help beyond the lectures?
  - Reading materials.
  - Course website, discussion sections, office hours, etc.

## Introduction to Statistics

### What is statistics?

- Statistics is a branch of scientific inquiry that helps to determine cause and association, and to make predictions.
- It provides methods to organize and summarize data from sample (i.e. any subset of a population).
- It also provides methods to use information in the data to draw various conclusions about a population (i.e. all individuals or objects of a particular type).

### Population vs. sample

- A plate vs. a spoonful of chef's special tonight.
- A book vs. a few pages of the book.
- All vs. 100 plants in a field.

## Introduction to Statistics

### Probability vs. statistics

- Probability is the mathematics of chance and randomness. In probability, properties of a population are assumed known and questions about a sample are posed and answered. Thus the approach here is deductive.
- In statistics, properties of a sample are available and conclusions about a population is drawn based on the sample. Thus the approach here is inductive.

### Three main topics

- Descriptive statistics: display and summarize data in a sample.
- Probability: Given a population, study the uncertainty associated with a sample taken from the population.
- Statistics: Given a sample, learn methods to draw conclusions about a population, while taking into account of uncertainties in the sample.

## Descriptive Statistics

### Example: height of seedlings

Thirteen (13) red pine seedlings were sampled from a nursery in Wisconsin. The heights of these seedlings were (in cm):

42 23 43 34 49 56 31 47 61 54 46 34 26

- Graphical methods describe data by visual/graphical techniques.
  - Stem-and-leaf plot, dot plot
  - Histogram
- Numerical methods extract summarizing numbers that characterize the data set and reveal main features.
  - Measures of location/center:
    - \* Sample mean
    - \* Sample median
    - \* Sample quantiles, box plot
  - Measures of spread:
    - \* Sample range
    - \* Interquartile range (IQR)
    - \* Sample variance, standard deviation

## Descriptive Statistics

### Stem-and-leaf plot

- Select leading digits for stem values and trailing digits for leaves.

- List all possible stems:

For the height of seedlings, 42 23 43 34 49 56 31 47 61 54 46 34 26

2 |

3 |

4 |

5 |

6 |

- Record leaf for each observation (or obs), besides the corresponding stem value:

2 | 36

3 | 144

4 | 23679

5 | 46

6 | 1

- Indicate the units for stems and leaves.

stem = 10 cm; leaf = 1 cm