Stat 571B Experimental Design

Topic 1: Introduction to Experimental Design

Prof. Lingling An
University of Arizona

Outline

- Syllabus
  - Goals
  - Class policies
  - Software & D2L
- Introduction to experimental design
  - Why do experiments?
  - Strategy
  - Issues in experimental design

Goals

- Learn how to plan, design and conduct experiments efficiently and effectively, and analyze the resulting data to obtain objective conclusions.
- Both design and statistical analysis issues will be discussed.
- Students will be expected to utilize standard statistical software packages for computational purposes.

Class policies

- Homework Assignments
  - generally one every two weeks
  - statistics is something you *do*
  - this should be individual work
  - late work policy
  - online students: dropbox in D2L
  - On-campus student: hardcopy
- Exams: one in-class midterm exam and one final exam.
  - online students: find a proctor at your place

Class policies

- One project:
  - each team consists of 2 students.
  - the report will be due the last day of class
  - presentation (15~20 min) will be scheduled the last week of class
  - online students: recording w/ follow-up calls or skype presentation individually

Class policies

- Grades
  - 25% Midterm Exam
  - 25% Final Exam
  - 35% Homework
  - 15% Project

A: 90 - 100

B: 80 -89

C: 70 -79

D: 60 -69

E: 0 - 59

Software

- SAS -what course is designed around
  - CDs for PC available from UA Bookstore
  - Free version:
    http://www.sas.com/en_us/software/university-edition.html
  - Computer labs:
    - ECE229, Shantz 338
    - http://www.oscr.arizona.edu/locations/computing_labs

D2L class site

- Syllabus
  - Contact information
  - Useful websites for the software
  - Textbook: *Design and Analysis of Experiment* (8th edition, by Montgomery)
- Lecture notes
- Homework assignments & Exam materials
- Data sets & SAS files
- Announcement, e.g., lab time

Are you ready for this class?

- Background: a good introductory course on Statistics that covers
  - probability distributions
  - sampling distributions
  - hypothesis testing/confidence interval
  - ANOVA
  - linear regression

Topics to cover

- Overview and Basic Principles (Chaps 1-2)
- Simple Designs and Analysis of Variance (Chap 3)
- Block Designs, Latin Squares and Related Designs (Chap 4)
- Full Factorial Designs (Chap 5)
- 2-level Full Factorial and Fractional Factorial Designs (Chaps 6-8)
- Response surface methods and designs (an overview, part of Chap 11)
- Designs with Random Factors, Nested Designs, and split-plot Designs (Chaps 13-14)

Experimental Design

- Experimental design is the part of statistics that happens *before* you carry out an experiment
- Proper planning can save many headaches
- You should design your experiments with a particular statistical test in mind

Why do experiments?

- Contrast: observational study vs. experiments
- Example:
  - Observational studies show a positive association between ice cream sales and levels of violent crime
  - What does this mean?

Alternative explanation

Ice cream

Correlation is not causation

Hot weather

Violent crime

Why do experiments? -2

- Observational studies are prone to *confounding variables:*
  - Variables that mask or distort the association between measured variables in a study
  - Example: hot weather

- In an experiment, you can use random assignments of treatments to individuals to avoid confounding variables

Why do experiments? -3

- To determine the *cause(s) of variation in the response*
- To find conditions under which the *optimal* (maximum or minimum) response is achieved
- To *compare responses at different levels of* controllable variables
- To develop a model for *predicting responses*

History of Design of Experiments

- **The agricultural origins, 1908 – 1940s**
  - W.S. Gossett and the *t*-test (1908)
  - R. A. Fisher & his co-workers
  - Profound impact on agricultural science
  - Factorial designs, ANOVA
- The **first industrial** era, 1951 – late 1970s
  - Box & Wilson, response surfaces
  - Process modeling and optimization
  - Applications in the chemical & process industries
- The **second industrial** era, late 1970s – 1990
  - Quality improvement and variation deduction initiatives in many companies
  - Taguchi and robust parameter design, process robustness

R. A. Fisher (1890 – 1962)

George E. P. Box

History -2

- The **modern** era, beginning circa 1990
  - Popular outside statistics, and an indispensable tool in many scientific/engineering endeavors
  - New challenges:
    - Large and complex experiments, e.g. screening design in pharmaceutical industry, experimental design in biotechnology
    - Computer experiments: efficient ways to model complex systems based on computer simulation
    - ..

A Systematic Approach to Experimentation

1. State objectives.
2. Choose responses.
   - What to measure? How to measure? How good is the measurement system?
3. Choose factors and levels
   - Flow chart and cause-and-effect diagram.
   - Factor experimental range is crucial for success
4. Choose experimental plan.

A Systematic Approach to Experimentation -2

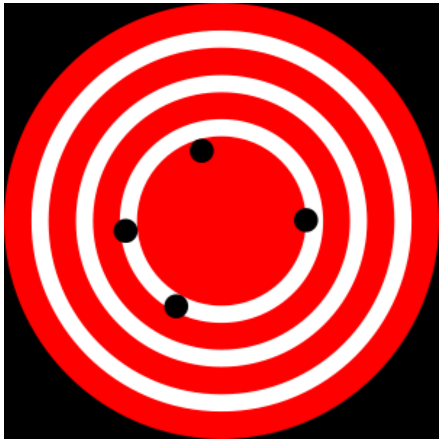
5. Conduct the experiment.
6. Analyze the data
7. Conclusion and recommendation.
    - iterative procedure
    - confirmation experiments/follow-up experiments

Issues in Experimental Design

- Eliminate bias
  1. Use a simultaneous control group
  2. Randomization
  3. Blinding

Three principles

- Reduce sampling error
  1. Replication
  2. Balance
  3. Blocking



- Calculate sample size

Issues in Experimental Design

- Eliminate bias
  1. Use a simultaneous control group
  2. Randomization
  3. Blinding
- Reduce sampling error
  1. Replication
  2. Balance
  3. Blocking
- Calculate sample size

Control Group

- A control group is a group of subjects left untreated for the treatment of interest but otherwise experiencing the same conditions as the treated subjects

- Example: one group of patients is given an placebo

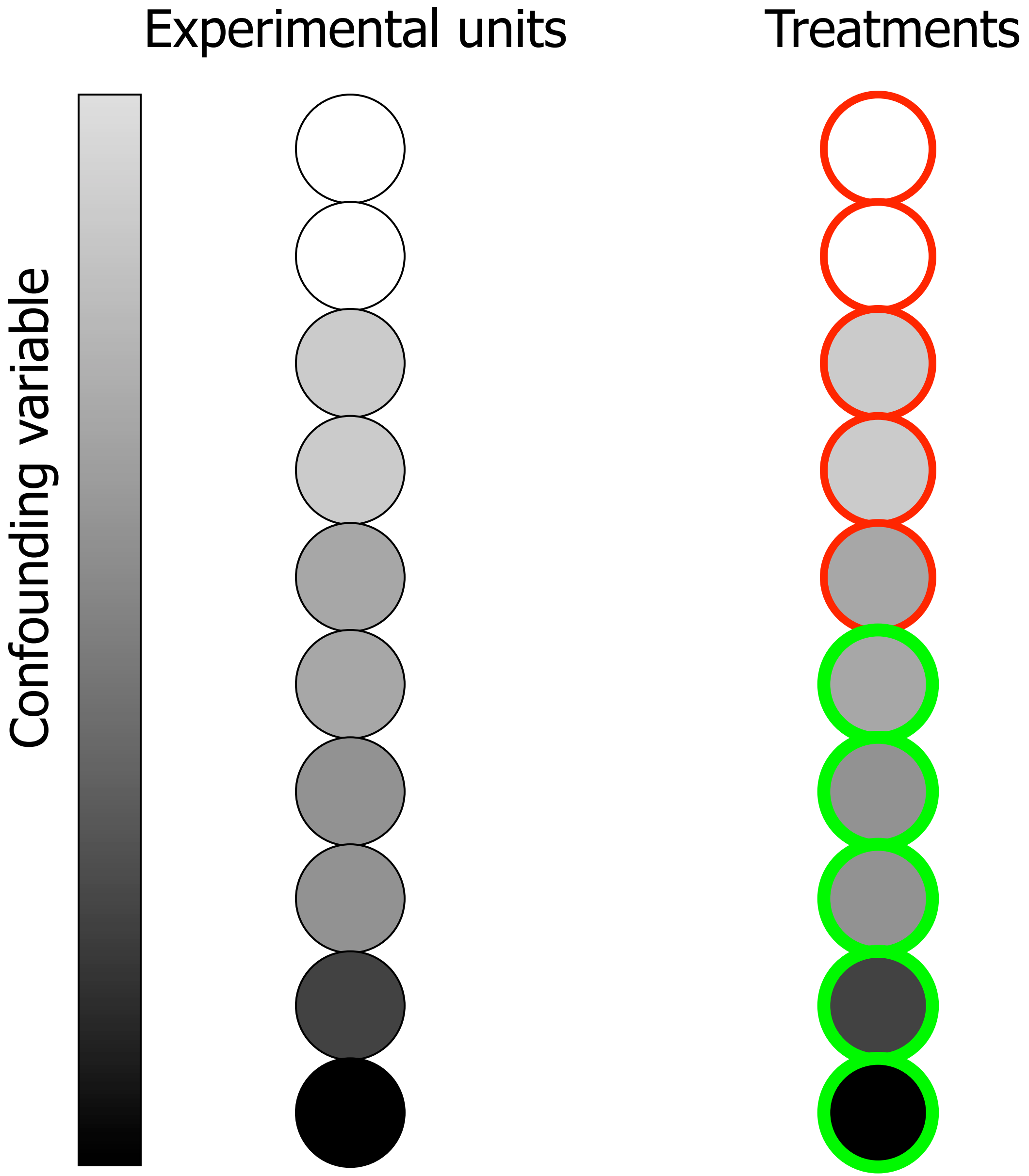
The Placebo Effect

- Patients treated with placebos, including sugar pills, often report improvement
  - Example: up to 40% of patients with chronic back pain report improvement when treated with a placebo
  - Even “sham surgeries” can have a positive effect
- This is why you need a control group!

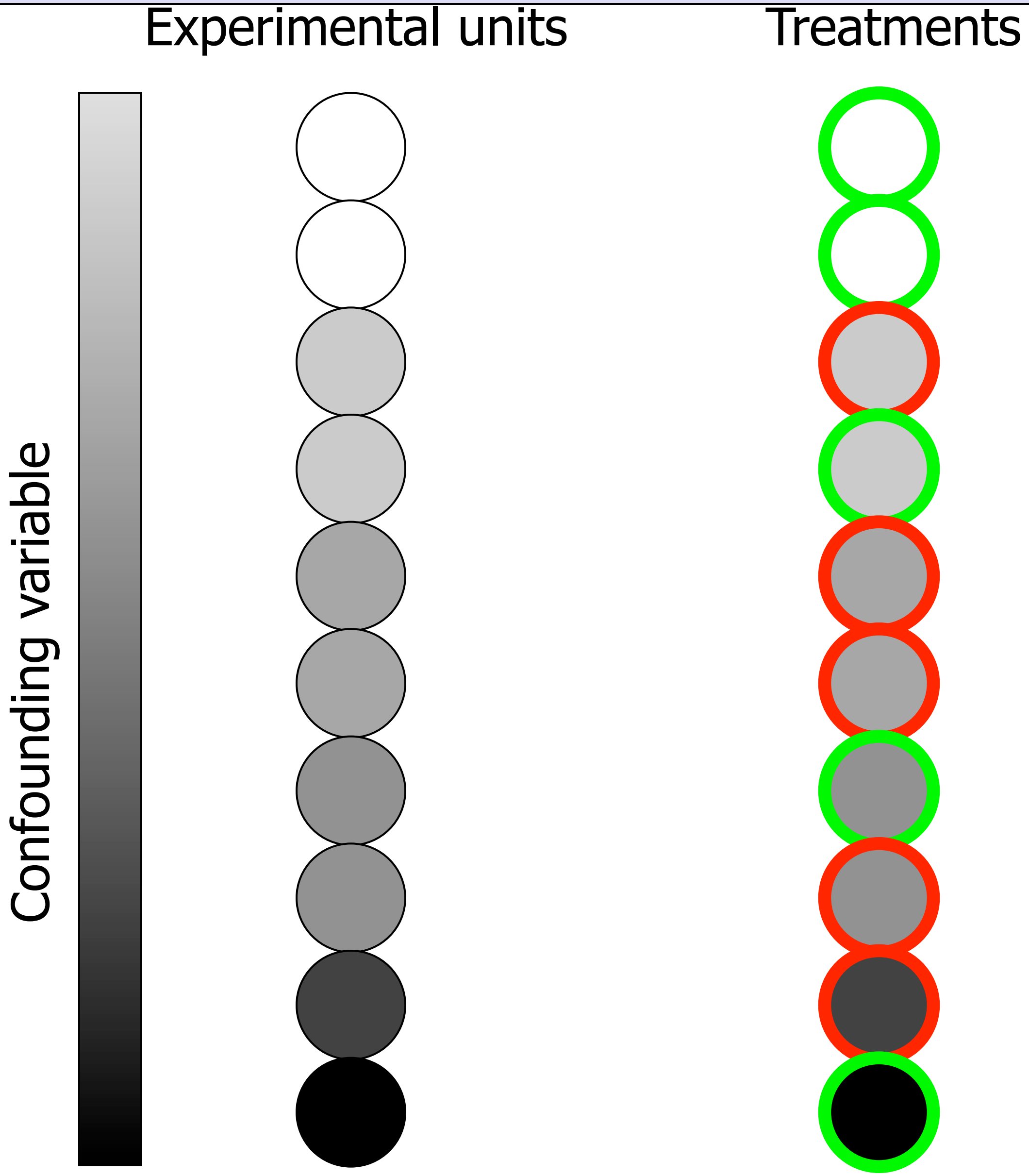
Randomization

- Randomization is the random assignment of treatments to units in an experimental study

- Breaks the association between potential confounding variables and the explanatory variables



Without randomization, the confounding variable **differs** among treatments



With randomization, the confounding variable **does not differ** among treatments

Blinding

- Blinding is the concealment of information from the participants and/or researchers about which subjects are receiving which treatments
  - Single blind: subjects are unaware of treatments
  - Double blind: subjects and researchers are unaware of treatments

Blinding -2

- Example: testing heart medication
  - Two treatments: drug and placebo
- Single blind: the patients don’t know which group they are in, but the doctors do
- Double blind: neither the patients nor the doctors administering the drug know which group the patients are in
  - The key that identifies the subjects and which group they belonged to is kept by a third party and not given to the doctors until the study is over.

Issues in Experimental Design

- Eliminate bias
  1. Use a simultaneous control group
  2. Randomization
  3. Blinding
- Reduce sampling error
  1. Replication
  2. Balance
  3. Blocking
- Calculate sample size

Replication

- Experimental unit: the individual unit to which treatments are assigned

2 Exprt. Units

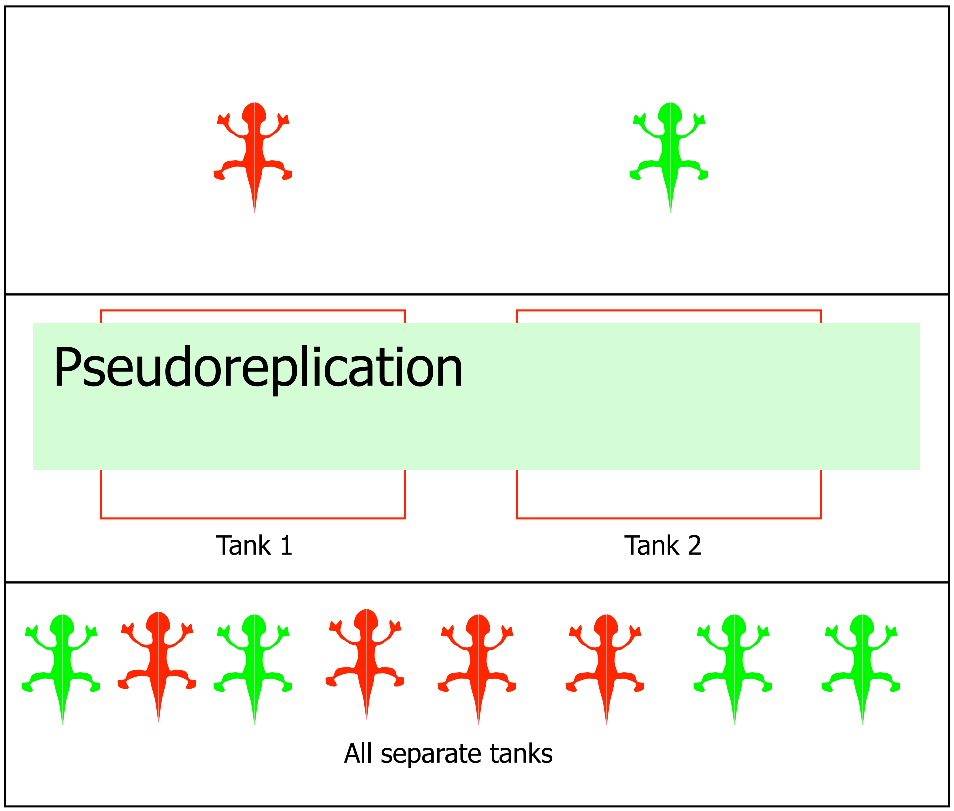
Experiment 1

2 Exprt. Units

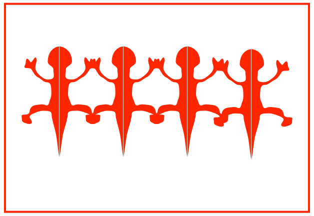
Experiment 2

8 Exprt. Units

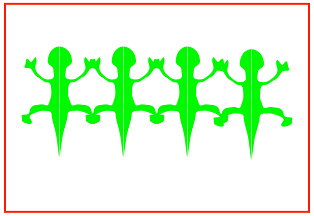
Experiment 3



Why is pseudoreplication bad?



Tank 1



Tank 2

Experiment 2

- problem with confounding and replication!
- Imagine that something strange happened, by chance, to tank 2 but not to tank 1
- Example: light burns out
- All four lizards in tank 2 would be smaller
- You might then think that the difference was due to the treatment, but it's actually just random chance

Replication vs. repeated measurements

- **Replication** is the repetition of an experimental condition so that the variability associated with the phenomenon can be estimated
  - Imagine you flip a coin and it comes up heads.
  - You ask a colleague to look at it and call out the result. Then another colleague is asked to observe the coin and state which side came up. This is a repeated measure.
  - A true replication is accomplished only by re-flipping the coin.

Why is replication good?

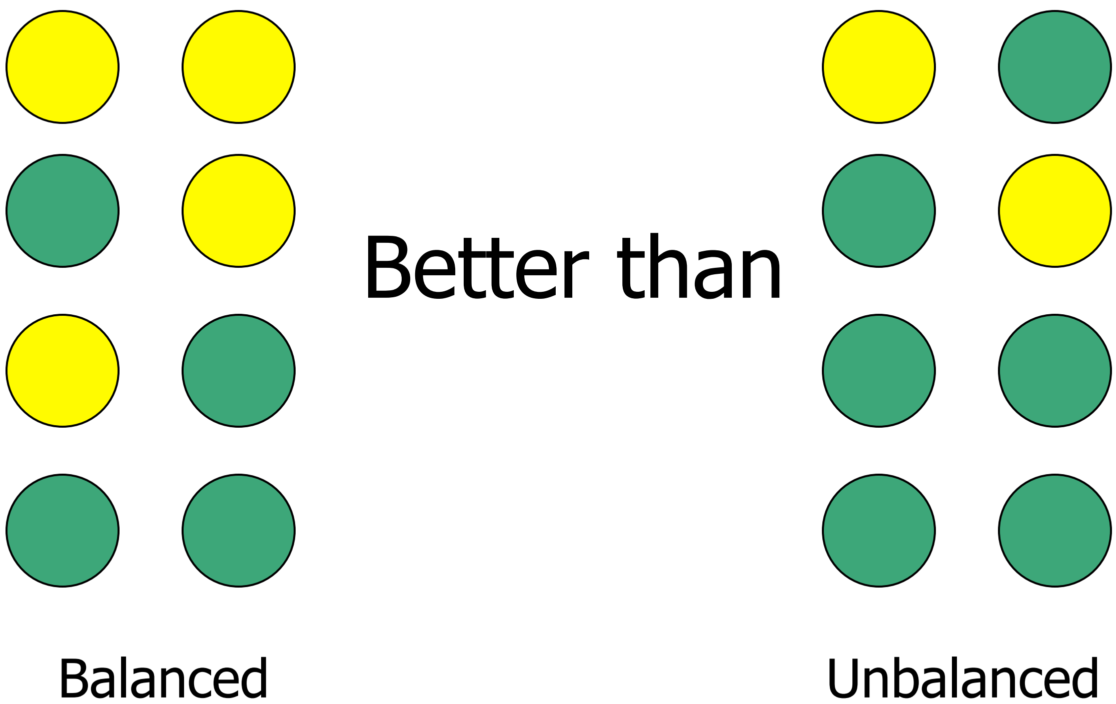
- Consider the formula for standard error of the mean:

$$SE_{\bar{Y}} = \frac{s}{\sqrt{n}}$$

Larger *n* $\longrightarrow$ Smaller SE

Balance

- In a balanced experimental design, all treatments have equal sample size



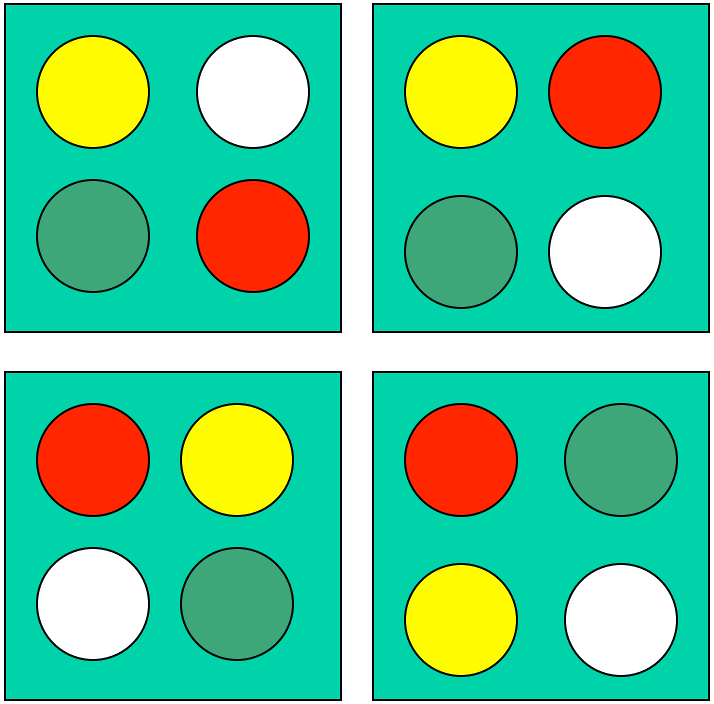
Balance -2

- In a balanced experimental design, all treatments have equal sample size
- This maximizes power
- Also makes tests more robust to violating assumptions

Blocking

- Blocking is the grouping of experimental units that have similar properties
- **Within** each block, treatments are randomly assigned to experimental units
- Randomized block design

Randomized Block Design



Randomized Block Design

- Example: cattle tanks in a field


Block 4
Block 3
Block 2
Block 1

What good is blocking?

- Blocking allows you to remove extraneous variation from the data
- Like replicating the whole experiment multiple times, once in each block

When need consider blocking, when need consider randomization?

- Nuisance factors are not our interest but they do affect the response
- For the nuisance factors
  - Block what you can control
  - Randomize what you cannot control

Issues in Experimental Design

- Eliminate bias
  1. Use a simultaneous control group
  2. Randomization
  3. Blinding
- Reduce sampling error
  1. Replication
  2. Balance
  3. Blocking
- Calculate sample size

Sample size calculation

- Before carrying out an experiment you must choose a sample size
- Too small: no chance to detect treatment effect
- Too large: too expensive

Sample size calculation – plan for precision

Example:

- Assume that the standard deviation of exam scores for a class is 10. We want to compare scores between two lab sections.
- How many exams do we need to mark to obtain a confidence limit for the difference in mean exam scores between two sections that has a width (precision) of 5?

Using confidence interval approach

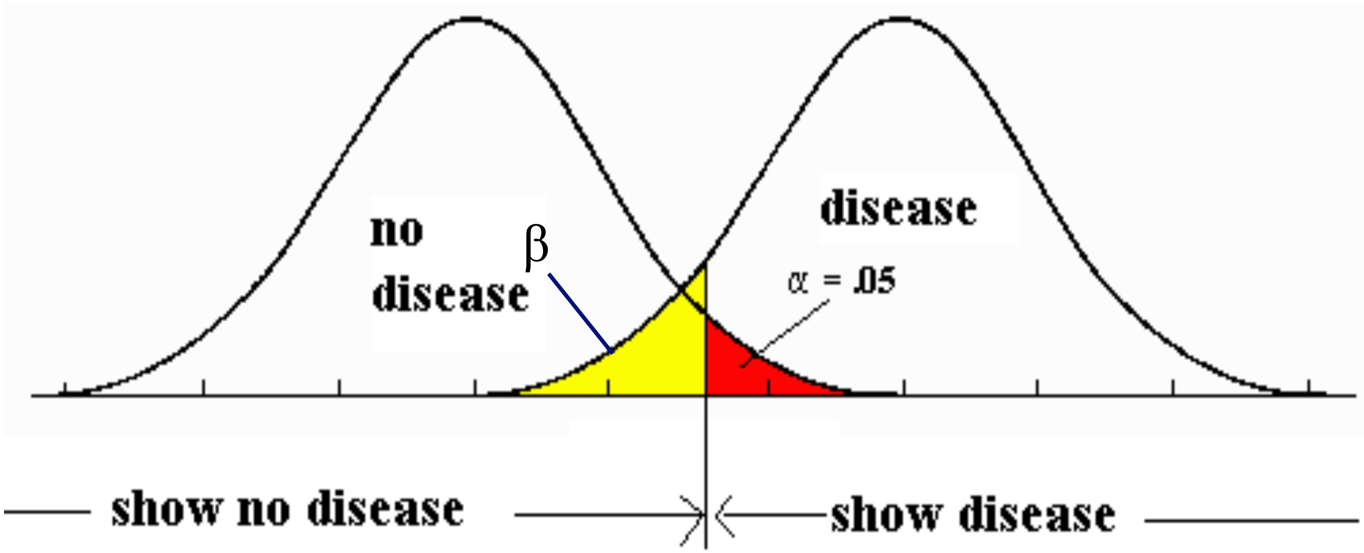
Sample size calculation
- plan for power

Example:

- Assume that the standard deviation of exam scores for a class is 10. We want to compare scores between two lab sections.
- How many exams do we need to mark to have sufficient power (80%) to detect a mean difference of 10 points between the sections?

Using power approach => type II error

Illustration of Types of error



Last slide

- Read Sections 1.1 - 1.6