STATISTICS 13V
Fall 2006
Professor Jessica Utts

COURSE WEBSITE:

http://anson.ucdavis.edu/~utts/st13v

UCDAVIS

How the Course Works

- Weekly Meetings, Mon, 5:10-6:30pm
  - Turn in Homework (*before* class starts)
  - Discussion/Question&Answer, up to 50 min.
  - Quiz, 10 questions, 30 minutes
- Between Meetings
  - Weekly Notes, posted on Mondays, directed self-study using textbook and CyberStats
  - Help available through office hours, email, optional discussion on Mon, 12:10-1:00, with TA
  - Sample quizzes are provided – good way to study!
  - See Course web site for all details!

How the Course Works, Continued

- Weekly Notes consist of one page with:
  - Short summary of the topics for the week
  - Reading assignments (text and CyberStats)
  - Where to focus (too much material!)
  - Interactivities to play with in CyberStats
    - Focus on understanding *why* you're doing it
  - Homework assignment

Some Details

- To enroll in CyberStats go to Course website, click on *How to Register for CyberStats* and follow directions. You will need:
  - Access Code (packaged with book) or buy online
  - Course Key (E-YNDJYQ78DQ6KJ), on Course website

More Details – Getting Help

- Teaching Assistant: Clayton Schupp
- Office hours and optional discussion:

  Monday
  - Optional discussion, 12:10-1, Schupp, 235 Wellman
  - Office hours: 3-4:30, Utts, 4214 MSB (except 11/6, 12/4; 1-2:30)

  Wednesday
  - Office hours: 9:30-11, Utts, 4214 MSB
  - Office hours: 12-1, Schupp, 1214 MSB

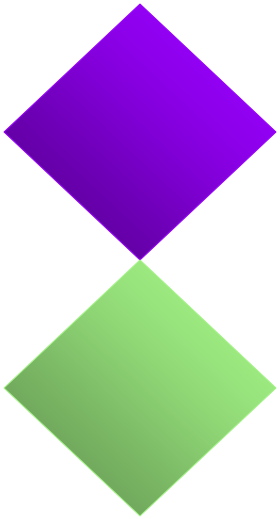
  Thursday
  - Office hours: 1-2, Schupp, 1214 MSB
- Available by email most of the time.

Homework, Tests and Grading

- **56%** Nine weekly quizzes (20 to 30 minutes, most are closed book) count 7% each; *the lowest one is dropped*
- **9%** Weekly homework – each assignment counts 1% of your grade
- **35%** Final exam - comprehensive, 2 hours, open book and notes, Fri December 15th, 10:30-12:30
- Because solutions are provided on the web right after class, *no late homework or quizzes*

Group Exercise

- Stand up and stretch.
- Divide into groups of 3 to 4 people.
- Introduce yourselves:
  - Name
  - Class (freshman, etc)
  - Home town
  - Why you are taking Statistics
- Discuss this question:
  - What is "Statistics" and why is it useful for everyday life to know something about Statistics?



Chapter 1

Statistics Success Stories and Cautionary Tales

1.1 What is Statistics?

Statistics is a collection of procedures and principles

- for gathering data and
- analyzing information
- in order to help people make decisions
- when faced with uncertainty.

1.2 Seven Statistical Stories With Morals

- **Case Study 1.1: Who Are Those Speedy Drivers?**
- **Case Study 1.2: Safety in the Skies**
- **Case Study 1.3: Did Anyone Ask Whom You've Been Dating?**
- **Case Study 1.4: Who Are Those Angry Women?**
- **Case Study 1.5: Does Prayer Lower Blood Pressure?**
- **Case Study 1.6: Does Aspirin Reduce Heart Attack Rates?**
- **Case Study 1.7: Does the Internet Increase Loneliness and Depression?**

Case Study 1.5 *Does Prayer Lower Blood Pressure?*

"Attending religious services lowers blood pressure more than tuning into religious TV or radio, a new study says"

***USA Today* headline read:**
"Prayer can lower blood pressure." (Davis, 1998)

Based on ***observational study***, followed 2391 people 6 years.

"People who attended a religious service once a week and prayed or studied the Bible once a day were 40% less likely to have high blood pressure than those who don't go to church every week and prayed and studied the Bible less."

Researchers ***did observe a relationship***, but it's a **mistake to conclude** prayer actually ***causes*** lower blood pressure.

Case Study 1.5 *Does Prayer Lower Blood Pressure?*

In observational studies, **groups can differ** by important ways that may contribute to the observed relationship.

People who attended church regularly may have …

- been less likely to smoke or drink alcohol;
- had a better social network;
- been somewhat healthier and able to go to church.

These are called possible ***confounding variables***.

Moral: *Cause-and-effect conclusions cannot generally be made based on an observational study.*

Case Study 1.6

Does Aspirin Reduce Heart Attack Rates?

Physician's Health Study (1988)

- 5-year ***randomized experiment*** …
- 22,071 male physicians of age 40 - 84
- ***randomly assigned*** to one of two ***treatment*** groups:
  - Group 1 = *aspirin* every other day
  - Group 2 = ***placebo*** every other day
- Physicians ***blinded*** as to which group they were in

Case Study 1.6

Does Aspirin Reduce Heart Attack Rates?

TABLE 1.1 The Effect of Aspirin on Heart Attacks

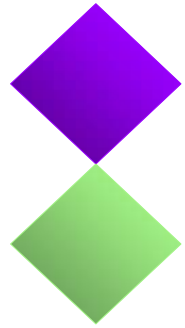
Treatment	Heart Attacks	Doctors in Group	Attacks Per 1000 Doctors
Aspirin	104	11,037	9.42
Placebo	189	11,034	17.13

Aspirin group: 9.42 heart attacks per 1000 participants
Placebo group: 17.13 heart attacks per 1000 participants

Randomization => other important factors (age, diet, etc) should have been similar for both groups. Only important difference should be whether they took aspirin or placebo.

Moral: *Unlike with observational studies, cause-and-effect conclusions can generally be made on the basis of randomized experiments.*

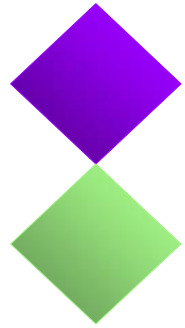
1.3 The Common Elements in the Seven Stories



In every story, ***data are used to make a judgment about a situation.***

This is what ***statistics*** is all about.

The Discovery of Knowledge



1. *Asking the right question(s).*
2. *Collecting useful data*, which includes deciding how much is needed.
3. *Summarizing and analyzing data*, with the goal of answering the questions.
4. *Making decisions and generalizations* based on the observed data.
5. *Turning the data and subsequent decisions into new knowledge.*

Step 1: Asking the right question(s)

Are differences in an **Explanatory variable** (or independent variable) related to differences in a **response variable** (or outcome or dependent variable)?

- Do people who pray and attend religious services have lower blood pressure than those who don't pray and attend religious services?
- Does taking aspirin reduce the risk of heart attack when compared to taking a placebo?

Step 2: Collecting useful data

Observational Study:

Researchers *observe* or *question* participants about opinions, behaviors, and outcomes. Participants not asked to do anything differently. (*Special case is a "Sample Survey" – more about those in reading this week.)*

Experiment:

Researchers *manipulate* something and measure the effect of the manipulation on some outcome of interest.

Randomized experiments: participants are randomly assigned to participate in one condition or another (called ***treatments***).

Some Tradeoffs (Obs. Study vs Experiment)

Advantages of Observational Studies:

Sometimes cannot conduct experiment due to practical/ethical issues.

More realistic to *observe* than *manipulate* behavior.

Advantages of Randomized Experiments:

Reduce the impact of ***confounding variables***

Usually *can* say change in explanatory variable *causes* change in response variable, if there is a change.

Problem with Observational Studies

- A **confounding variable** is a variable that
  - *affects the response variable* and
  - is *related to the explanatory variable*.
- The effect of the explanatory variable on the response variable cannot be separated from the effect of confounding variables.
- A **lurking variable** is a confounding variable that was not measured or considered in the study.

Example 4.1 *What Confounding Variables Lurk behind Lower Blood Pressure?*

Recall Case Study 1.5: people who attended church regularly had lower blood pressure than those who stayed home.

Possible confounding variables:

- Amount of social support
- Health status
- Age
- Attitude toward life

Why aren't confounding variables a concern in Randomized Experiments?

Randomized experiments: Usually allow us to determine **cause and effect**.

Random assignment makes the groups approximately equal in all respects except for the explanatory variable.

Case Study 1.6: Physicians randomly assigned to take aspirin or placebo, over 11,000 treatment. Groups should be similar in all other respects, such as age, exercise level, etc. because of random assignment.

Remaining Steps – Rest of the Course!

3. *Summarizing and analyzing data*, with the goal of answering the questions.
4. *Making decisions and generalizations* based on the observed data.
5. *Turning the data and subsequent decisions into new knowledge*.

Case Study 4.3
Quitting Smoking with Nicotine Patches

- 240 smokers randomly assigned to wear nicotine patch or placebo patch for 8 weeks
- 46% of nicotine patch wearers quit smoking
- 20% of placebo patch wearers quit smoking

Discuss:

Observational study or randomized experiment?

Explanatory and response variables?

Can “cause and effect” be concluded?

Case Study 4.3, continued

Explanatory variable: *Type of patch*

Response variable: *Quit smoking or not*

Randomized experiment, so we *can* conclude that wearing nicotine patches *causes* people to be more likely to quit smoking.

NOTE: The difference in proportion who quit is called ***statistically significant*** because it is unlikely to have happened just by chance in this sample, if there really were no difference in the population.

Part of Weekly Notes for this week

Reading and Study Assignment for this week:

Book Chapter or CyberStats Unit to Read	**Focus on:**
Chapter 1: Statistics Success Stories	All (it's very short!)
Chapter 3: Sampling and Surveys	All
Chapter 4: Gathering Useful Data	All
Unit A2: Overview of Sampling	Unit Review
Unit A3: Overview of Experiments and Observational Studies	Self assess quiz and Unit Review
Unit A4: Population vs. Sample	Basics 1 and Self assess quiz

Interactivities to play with *and understand*: Unit A3, Uses 1; Unit A4, Uses 3 (*this is also on the CD in the back of the book, and described in Chapter 3, pgs 103-106)*

Exercises to hand in (due Oct 9 at the start of class):
Chapter 1: 9, 20 **Chapter 3**: 15cd, 23, 63, 98 **Chapter 4**: 1ab, 2ab, 5a, 13, 18ab, 20, 54