Math 140
Introductory Statistics

Professor Silvia Fernández

Chapter 4
Based on the book *Statistics in Action*
by A. Watkins, R. Scheaffer, and G. Cobb.

Sample surveys and experiments

- Most of what we've done so far is data exploration—ways to uncover, display, and describe patterns in data. Unfortunately, these patterns can't take you beyond the data in hand. *With exploration, what you see is all you get.* Often, that's not enough.

Sample surveys and experiments

- **Pollster:** I asked a hundred likely voters who they planned to vote for, and fifty-two of them said they'd vote for you.
- **Politician:** Does that mean I'll win the election?
- **Pollster:** Sorry, I can't tell you. My stat course hasn't gotten to inference yet.
- **Politician:** What's inference?
- **Pollster:** Drawing conclusions based on your data. I can tell you about the hundred people I actually talked to, but I don't yet know how to use that information to tell you about *all* the likely voters.

Sample surveys and experiments

- Methods of ***inference*** can take you beyond the data you actually have, but only if your numbers come from the right kind of process.
- If you want to use 100 likely voters to tell you about ***all*** likely voters, how you choose those 100 voters is crucial.
- The quality of your inference depends on the quality of your data; in other words, bad data lead to bad conclusions.

4.1 Why Take samples, and How Not To.

- Population: Set of people or things we want to study.
- Unit: Individual element of the population.
- Population Size: Number of units.

- Sample: set of units that you **do** get to study.
- Census: collecting data on the entire population.

Census Versus Sample Discussion. D1

- In which of these situations do you think a census is used to collect data, and in which do you think sampling is used? Explain your reasoning.
  - a. An automobile manufacturer inspects its new models.
  - b. A cookie producer checks the number of chocolate chips per cookie.
  - c. The U.S. president is determined by an election.
  - d. Weekly movie attendance figures are released each Sunday.
  - e. A Los Angeles study does in-depth interviews with teachers in order to find connections between nutrition and health.

Discussion

- You want to estimate the average number of TV sets per household in your community.
  - a. What is the population? What are the units?
  - b. Explain the advantages of sampling over conducting a census.
  - c. What problems do you see in carrying out this sample survey?

Bias: A problem with survey data.

- A sampling method is biased if it tends to give samples in which some characteristic of the population is overrepresented or under-represented.
- An Unbiased Sample Method requires that **all** units in the population have a chance of being in the sample.
- A sampling frame is the list of units you use to create the sample. “bad frame, bad sample”.

Bias: (dialogue page 222)

Investigator: What makes a good sample?

Statistician: A good sample is representative; that is, it looks like a small version of the population. Proportions you compute from the sample are close to the corresponding proportions you would get if you used the whole population. The same is true for other numerical summaries, like averages and standard deviations or medians and *IQR*s.

Investigator: How can you tell if your sample is representative?

Statistician: There's the rub: in practice, you can't. You can tell only by comparing your sample with the population, and if you know that much about the population, why bother to take a sample?

Investigator: Great! First you tell me my sample should be representative, and then you tell me there's no way to know whether it is. Is that the best statisticians can do?

Bias: Sample vs. method used for choosing the sample (dialogue p. 222)

Statistician: Nope. Although you can't tell about any particular sample, it *is* possible to tell whether a sampling *method* is good or not. That's where bias comes in.

Investigator: I thought "biased" was just a fancy word for "nonrepresentative." Not true?

Statistician: Now we're getting to the point. Bias refers to the method, not the samples you get from it. A method is biased if it tends to give nonrepresentative samples.

Investigator: Now I get it. I may not be able to tell whether my sample is representative, but if I use an unbiased method, then I can be confident that my sample is likely to be representative. Right?

Statistician: Now you're thinking like a statistician. There's more detail to come, but you've got the big picture in focus.

Sampling Bias

- Size bias: Larger units are more likely to be included.
- Voluntary Response Bias: Those who care about the issue respond.
- Convenience Sample Bias: Units are chosen because of convenience.
- Judgment Sample Bias: Units are chosen according to the judgment of someone (expert)

Sample Bias Discussion (D5)

You want to know the percentage of voters who favor state funding for bilingual education. Your population of interest is the set of people likely to vote in the next election. You use as your frame the phone book listing of residential telephone numbers.

- How well do you think the frame represents the population?
- Are there important groups of individuals who belong to the population but not to the frame? To the frame but not to the population?
- If you think bias is likely, identify what kind of bias and how it might arise.

Response Bias

- Non-Response Bias: You get no data or not enough data. e.g. 80% of people contacted refuse to answer a Survey
- Questionnaire Bias: Arises from the way the questions are asked.
- Bias from incorrect responses: Might be the result of intentional lying (often, the people being interviewed want to be agreeable and tend to respond in the way they think the interviewer wants them to respond), but it is more likely to come from inaccurate measuring devices, including inaccurate memories of people being interviewed in self-reported data.

Response Bias

"One final question: Do you now own or have you ever owned a fur coat?"

Response Bias

- *Reader's Digest* commissioned a poll to determine how the wording of questions affected people's opinions. The same 1031 people were asked to respond to these two statements:

1. I would be disappointed if Congress cut its funding for public television.
2. Cuts in funding for public television are justifi ed as part of an overall effort to reduce federal spending.

- Note that agreeing with the first statement is pretty much the same as disagreeing with the second.

	Agreed	Disagreed	Didn't know
Statement 1	54%	40%	6%
Statement 2	52%	37%	10%

[*Source:* Fred Barnes, "Can You Trust Those Polls?" *Reader's Digest,* July 1995, pp. 49–54.]

Response Bias

4.2 Randomizing: Playing It Safe by Taking Chances

Randomize: Choose a sample by chance.
This is the only method guaranteed to be unbiased.

- Simple random sample (SRS)
- Stratified random samples
- Cluster samples
- Two (or more) stage samples
- Systematic samples with a random start.

Simple random sample (SRS)

- In a SRS all possible samples of a given fixed size are equally likely. That is all units have the same chance of being in the sample, all possible triples of units have the same chance, and so on.
- Steps in choosing a SRS
  - 1. Start with a list of **all** units in the population. (a frame)
  - 2. Number the units in the list.
  - 3. Use a random number table or generator to choose units from the numbered list, one at a time, until you have as many as you need.

Stratified random samples

- 1. Divide the units of the sample into non-overlapping subgroups (strata)
- 2. Choose a SRS from each subgroup (stratum)

Choose the relative sample sizes proportional to the stratum sizes.

Why stratify

- Convenience. It is easier to sample in smaller more compact groups.
- Coverage. Each stratum is assured to be covered. (this may not happen with a SRS)
- Precision. The results may be more precise if the measurement we are interested varies a lot from stratum to stratum.

Cluster samples

- 1. Create a numbered list of all the clusters in the population.
- 2. Choose a SRS of clusters
- 3. Obtain data on **each** unit in each chosen cluster.

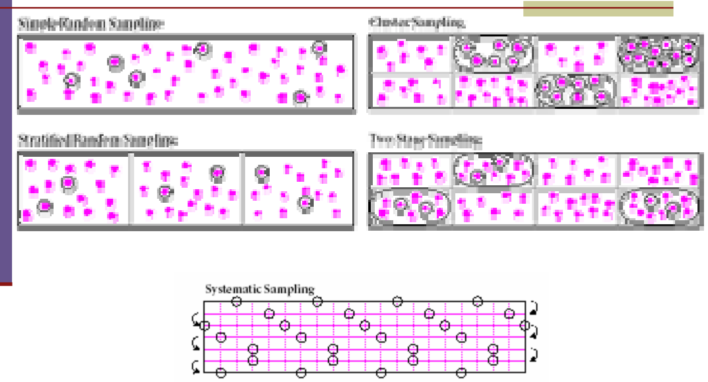
Two (or more) stage samples

- 1. Create a numbered list of clusters.
- 2. Choose a SRS of clusters.
- 3. From each selected cluster, create a list of individuals and choose a SRS from each (selected) cluster.

Systematic samples with a random start

- 1. By a method, such as counting off, divide your population into groups of the size you want for your sample.
- 2. Use a chance method to choose one of the groups for your sample.

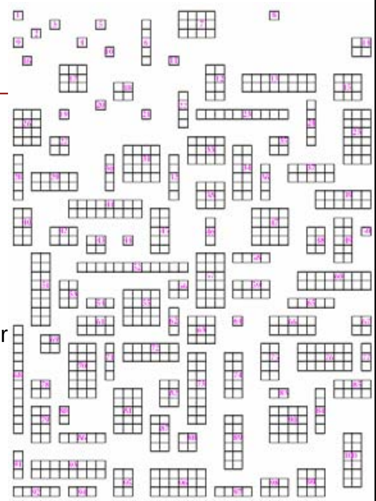
Summary of sampling methods



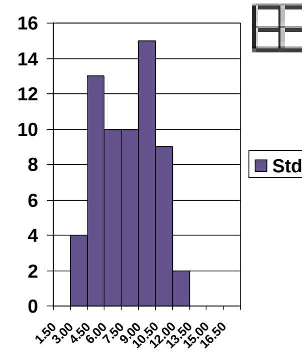
Activity 4.2
Part 1. (page 225)

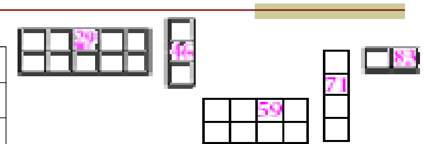
- Quickly choose 5 rectangles.
- Calculate the areas of each of your 5 rectangles
- Calculate the mean (average) of these areas.

Keep your sample data for future reference.



Results





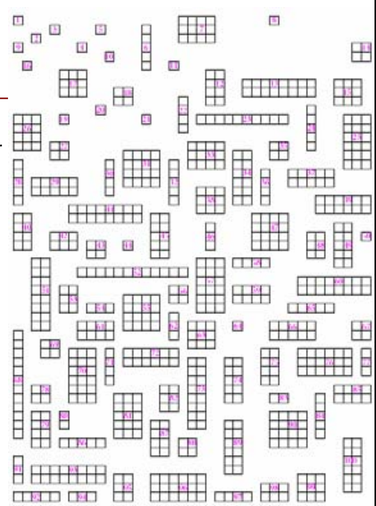
- Rectangles: 29, 46, 59, 71, 83
- Areas: 10, 3, 8, 4, 2
- Mean: (10+3+8+4+2)/5 = 5.4

Activity 4.2
Part 2.

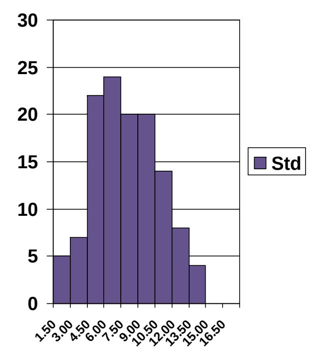
- Choose 5 random numbers between 1 and 100. Look for the rectangles associated to these numbers.

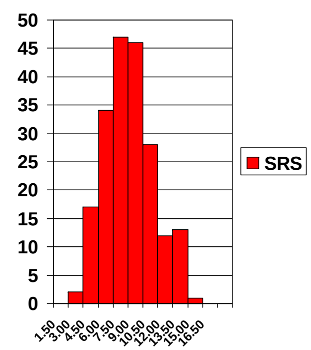
  Use $\text{randInt}(1{,}100)$
- Calculate the areas of each of these 5 rectangles
- Calculate the mean (average) of these areas.

Keep your sample data for future reference.

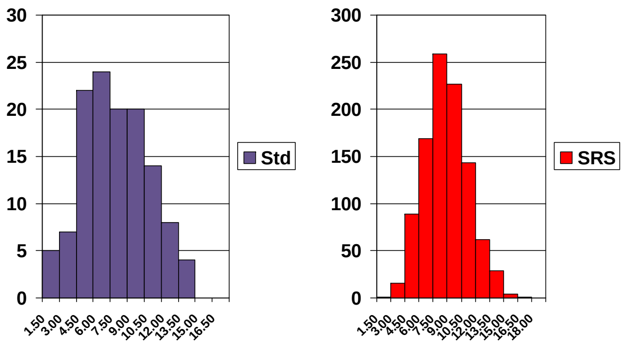


Results 1
(Computer Simulated $n = 200$)



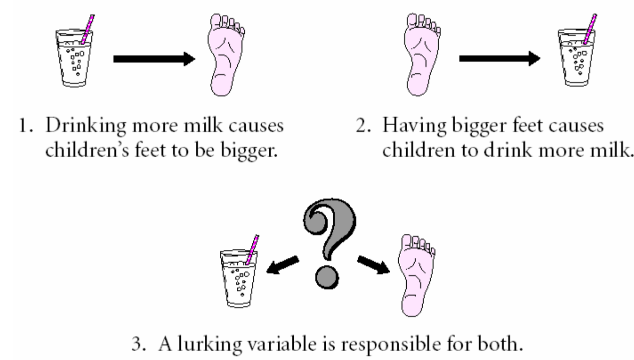


Results 2
(Computer Simulated $n = 1000$)



4.3 Experiments and Inference about Cause

- Cause and Effect



Experiments and Inference about Cause

- Lurking Variable:

  A variable in the background that could explain a pattern between the variables investigated.
- How to establish cause and effect?

Answer: Conduct an experiment.

Experiments

- Goal: To establish cause and effect by comparing two or more ***conditions*** (called treatments) using an ***outcome variable*** (called the response).
- To be a real experiment, the subjects must be randomly assigned to their treatments. To make this distinction sometimes we call these Randomized Experiments.

Example: Kelly's Hamsters

- Assumptions
  - Golden Hamsters hibernate.
  - Hamsters rely on the amount of daylight to trigger hibernation.
  - An animal's capacity to transmit nerve impulses depends in part on an enzyme called Na^+K^+ ATP-ase.
- Question: If you reduce the amount of light a hamster gets, from 16 hours to 8 hours per day, what happens to the concentration of Na^+K^+ ATP-ase.

Example: Kelly's Hamsters

- Subjects: Eight golden hamsters.
- Treatments: Raised in long days (16 hours) or short days (8 hours) of daylight.
- Random Assignment of Treatments: Kelly randomly assigns four of the hamsters to short days, and four to long days.
- Replication: Each treatment was given to four hamsters.
- Response Variable: Enzyme concentration.

Kelly's Hamsters (Results)

- Results

Enzyme concentrations in milligrams per 100 milliliters.

Short Days	12.500	11.625	18.275	13.225
Long Days	6.625	10.375	9.900	8.800

Kelly's defense of her design

Kelly: I claim that the observed difference in enzyme concentrations between the two groups of hamsters is due to the difference in daylight.

Skeptic: Wait a minute. As you can see, the concentration varies from one hamster to another. Some just naturally have higher concentrations. If you happened to assign all the high-enzyme hamsters to the group that got short days, you'd get results like the ones you got.

Kelly: I agree, and I was concerned about that possibility. In fact, that's precisely why I assigned day lengths to hamsters by using random numbers. The random assignment makes it **extremely unlikely** that all the high-enzyme hamsters would get assigned to the same group. If you have the time, I can show you how to compute the probability.

Skeptic: (*Hastily*) That's OK for now. I'll take your word for it. But maybe you can catch me in Chapter 6.

Discussion D21 (page 245)

- **D21**. Kelly has shown that hamsters raised in less daylight have higher hormone concentration than hamsters raised with more daylight. In order for Kelly to show that less daylight *causes* an increase in the hormone concentration, she must convince us that there is no other explanation. Has she done that?

Confounding in Observational Studies

- ***Confounded***: mixed-up, confused, at a dead end.
- Two possible influences on an observed outcome are said to be confounded if they are mixed together in a way that makes it impossible to separate their effects.

Confounding in Observational Studies

- Studies that claim to show that review courses increase SAT scores often ignore the important concept of confounding.
- In one study, students at a large high school were offered an SAT preparation course, and SAT scores of students who completed the course were higher than scores of students who chose not to take the course.
- The positive effect of the review course was confounded with the fact that the course was taken only by volunteers, who would tend to be more motivated to do well on the SAT.
- Consequently, you can't tell if the higher scores of those who took the course were due to the course itself or to the higher motivation of the volunteers.

Confounding in Observational Studies

Imagine yourself in this situation:

- You know that many infants are dying of what seem to be respiratory obstructions.
- You begin to do autopsies on infants who die with respiratory symptoms.
- The infants all have thymus glands that look too big in comparison to body size. Aha! That must be it:
- The respiratory problems are caused by an enlarged thymus.
- It became quite common in the early 1900s for surgeons to treat respiratory problems in children by removing the thymus. Even though a third of the children who were operated on died.

Confounding in Observational Studies

Thymus Size	Age: Child	Age: Adult
Large	Problems	No evidence
Small	No evidence	No problems

- The doctors couldn't know whether children with a large thymus tend to have more respiratory problems, because they have no evidence about children with a smaller thymus. Age and size of thymus were confounded.

Experiments vs. Observational Studies

- The best solution to guard against confounding: To randomize.
- Observational Study: No treatment gets assigned to the subjects by the experimenter.
- (Randomized) Experiment: Comparing results of treatments assigned to subjects at random.
- Clinical Trial: Randomized experiment comparing medical treatments.
- For observational studies the conditions are called factors, (not treatments)

Factors and Levels

- The term factor is also used for experiments when there are many characteristics that want to be compared.
- The different values that a factor may take are called levels.
- Example. If Kelly added the type of diet to her experiment.

Factor 1
Type of Diet

Factor 2
Length of Day

	Light	Heavy
Short	Light-Short	Heavy-Short
Long	Light-Long	Heavy-Long

Why randomization makes inference possible?

- By assigning treatments to units at random, there are only **two possible causes** for a difference in the responses to the treatments: **chance** or **the treatments**.
- If the probability is small that chance alone will give such a difference in the responses, then we can **infer** that the cause of the difference was **the treatment**.

Control or Comparison Group

- Anecdotal evidence is not proof. Why?
- Placebo Effect: When people believe they are getting special treatment they tend to improve.
- Control Group: A group of people given a placebo.
- Comparison Group: A group of people given the standard treatment (when comparing against a new treatment).
- Blind Experiment: People do not know which treatment they are given.
- Double Blind Experiment: patients and doctors do not know which treatment they are assigned.