Stat 13, Intro. to Statistical Methods for the Life and Health Sciences.

1. Syllabus, etc.
2. Textbook.
3. Example with organ donations.
4. Rough interpretation of distribution and standard deviation.
5. Sample size.
6. Statistic and parameter.
7. Categorical and quantitative variables.
8. Statistical significance and testing.
9. Null and alternative hypotheses.
10. Z statistic.
11. Simulating null distributions.
12. p-values.

Read preliminaries, chapter 1, and p592, the first page of Appendix A.
Hw1 is due Tue 10/4. 1.3.16 and 1.4.26. Also, on the bottom of your hw, print the names and emails of two other students in the class.

The course website is http://www.stat.ucla.edu/~frederic/13/F16

1. Syllabus, etc.

Read the syllabus, especially the hw policy, the gradegrubbing policy, and the 1 question not to ask me.
Here are things not on it but worth mentioning.

The CCLE website for this course is not maintained.
The only course website is http://www.stat.ucla.edu/~frederic/13/F16 .

I do not give hw hints in office hours. Conceptual questions only.

Attendance is not mandatory in lecture nor in section and lab.

You can only switch sections if we find someone to switch with you. If you want to switch sections, email me, frederic@stat.ucla.edu. Indicate your current section and preference.

2. Textbook.

Tintle N, Chance BL, Cobb GW, Rossman AJ, Roy S, Swanson T, and Vanderstoep J. (2016). Introduction to Statistical Investigations, Wiley, NY.

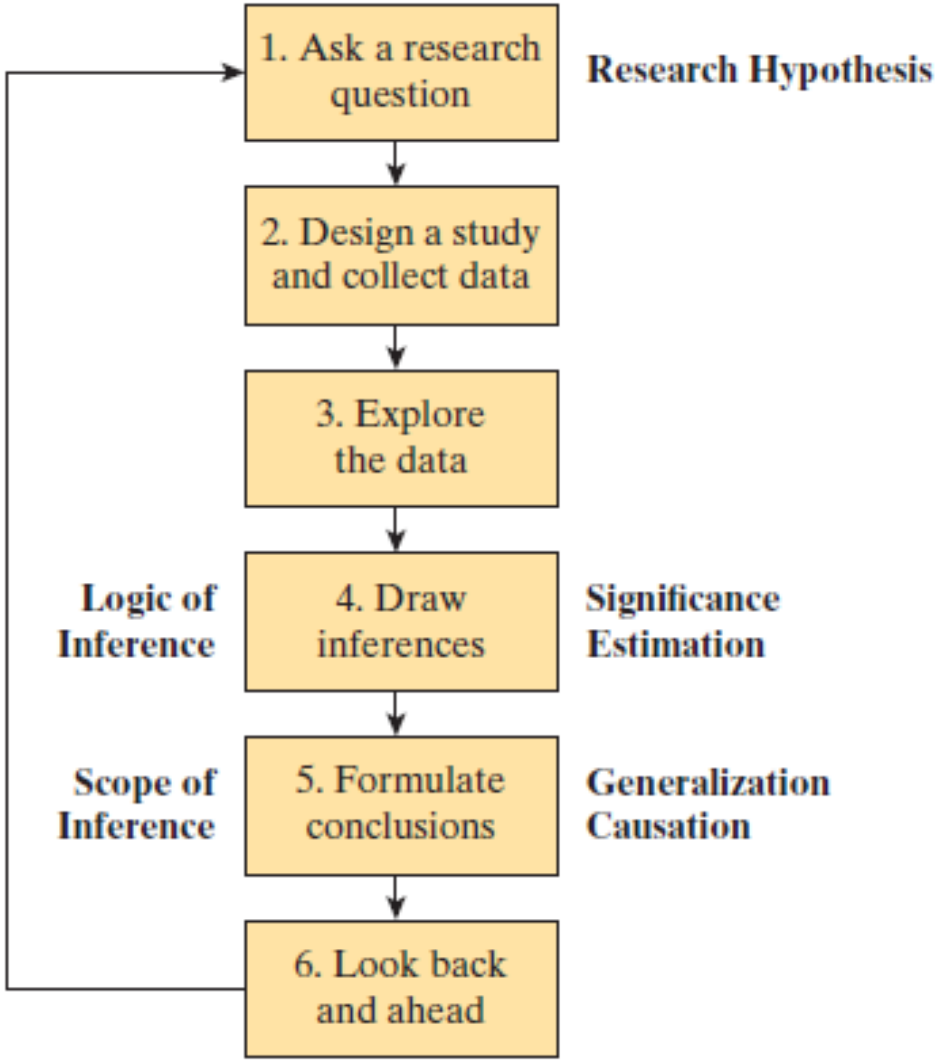
- Emphasizes concepts, not formulas.
- Emphasizes randomization tests and other nonparametric methods.
- Verbose, and some examples are phony or unimportant.

Optional reading, "Statistics for the Life Sciences", by Samuels and Witmer.

3. Example P.1: Organ Donations

- While a majority of people approve of organ donation in principle, far less than that actually sign up when getting a driver's license.
- Different states (and different countries) have different recruiting methods.
- Do these different methods result in different sign-up rates?

Six-Step Statistical Investigation Method



Recruiting Organ Donors

Step 1. Ask a Research Question

- Does the default option presented to driver's license applicants influence the likelihood of someone becoming an organ donor?

Recruiting Organ Donors

Step 2: Design a study and collect data

- The researchers decided to recruit various participants and ask them to pretend to apply for a new driver's license.
- The participants did not know in advance that different options were given for the donor question, or even that this issue was the main focus of the study.

Recruiting Organ Donors

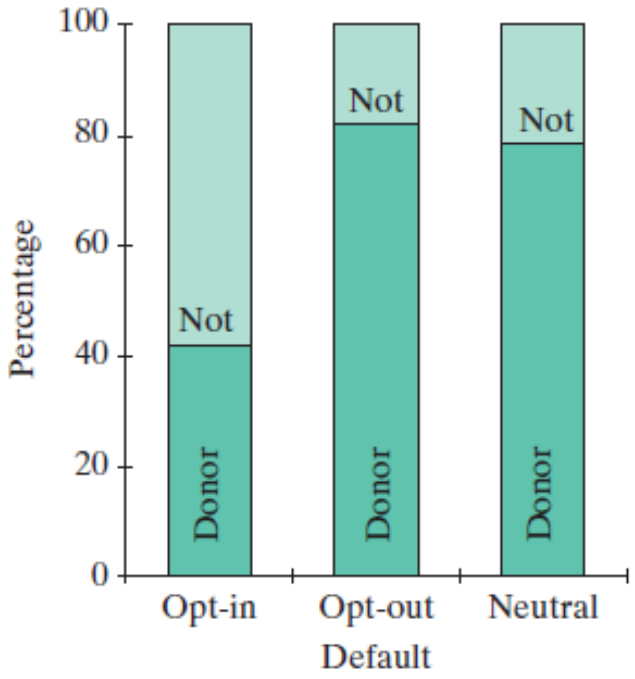
Step 2: Design a study and collect data

- Some of the participants were forced to make a choice of becoming a donor or not, without being given a default option (the "neutral" group, Michigan's current practice).
- Other participants were told that the default option was *not* to be a donor but that they could choose *to* become a donor if they wished (the "opt-in" group, Michigan's past practice).
- The remaining participants were told that the default option was to *be* a donor but that they could choose *not* to become a donor if they wished (the "opt-out" group, some countries use this practice).

Recruiting Organ Donors

Step 3: Explore the data.

- 23 of 55 (41.8%) participants in the opt-in group agreed to become organ donors
- 41 of 50 (82.0%) participants in the opt-out group agreed to become organ donors
- 44 of the 56 (78.6%) participants in the neutral group agreed to become organ donors



Recruiting Organ Donors

Step 4: Draw inferences beyond the data.

- Using methods that you will learn in this course, the researchers analyzed whether the observed differences between the groups was large enough to indicate that the default option had a genuine effect.
- In particular, they reported strong evidence that the neutral and opt-out versions do lead to a higher chance of agreeing to become a donor, as compared to the opt-in version currently used in many states.
- In fact, they could be quite confident that the neutral version increases the chances that a person agrees to become a donor by between 20 and 54 percentage points, a difference large enough to save thousands of lives per year in the United States.

Recruiting Organ Donors

Step 5: Formulate conclusions.

- Based on the analysis of the data and the design of the study, the researchers concluded that the neutral version *causes* an increase in the proportion who agree to become donors over the opt-in.
- But because the participants in the study were volunteers recruited from various general interest Internet bulletin boards, generalizing conclusions beyond these participants is only legitimate if they are representative of a larger group of people. (The authors believed their sample included a “broad range of demographics.”)

Recruiting Organ Donors

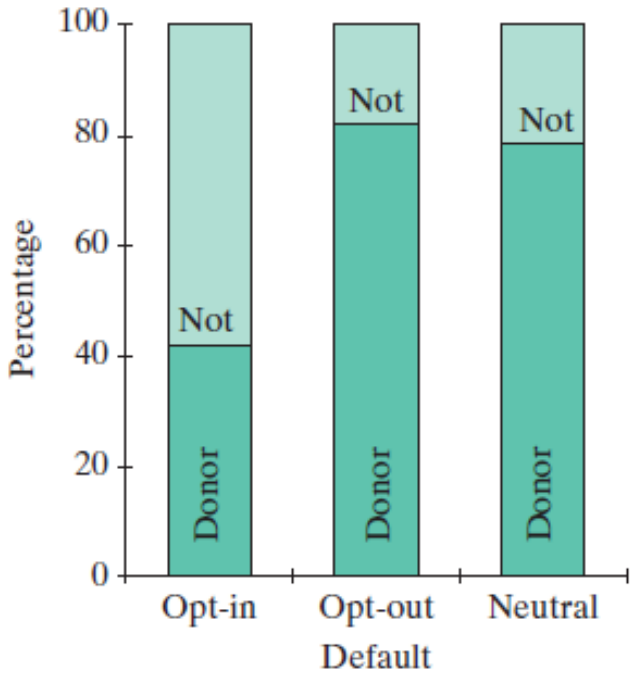
Step 6: Look back and ahead.

- One limitation of the study is that participants were asked to imagine how they would respond, which might not mirror how people would actually respond in such a situation.
- A new study might look at people's actual responses to questions about organ donation or could monitor donor rates for states that adopt a new policy.

- The individual entities on which data are recorded are called **o*bservational units***.
- The recorded characteristics of the observational units are the ***variables*** of interest.
- What are the observational units and variables in the Organ Donation Study?

4. Distribution and SD (rough definitions)

- The ***distribution*** of variable describes the pattern of value/category outcomes.
- For the organ donation study the bar chart shown displays the distribution of responses.



- One way to measure the center of a distribution is with the average, also called the mean. Sample mean $\overline{x} = \sum x_i/n$.
- One way to measure variability is with the ***standard deviation***, which is roughly the average distance between a data value in the distribution and the mean of the distribution. The sample std deviation, $s = \sqrt{[\sum (x_i - \overline{x})^2/(n-1)]}$
  - What is the standard deviation of the data set {7,7,7,7,7}?
  - Which data set has the largest standard deviation?
    - A {1, 3, 3, 3, 3, 3, 7}
    - B {1, 2, 3, 4, 5, 6, 7}
    - C {1, 1, 1, 4, 7, 7, 7}

5. Sample size.

Each record typically corresponds to an *observational unit*, and the number of observed units in the analysis is called the sample size, *n*.

In some situations, the population size might be known and you might have a Simple Random Sample (SRS) from the population. The sample size then is the number of people in your sample.
For instance, there are 4 million births every year in the United States.
Suppose we sample 1,000 of them at random from this population, and record for each pregnancy, the number of weeks of pregnancy, and the height, weight and gender of the baby at birth.
Here n = 1,000. Each baby is an observational unit.

6. Statistic and parameter.

A statistic is a numerical description of your sample. Another word for statistic is *random variable*. The sample is typically considered random, and if a different sample were obtained, then the statistic might be different.

A parameter, however, is a property of the whole population. If a different sample were obtained, the parameter would not change.

Parameters are properties of the population. Typically unknown. Represented by Greek letters (like μ or σ).
Statistics are properties of the sample.
Represented by Roman letters (like $\bar{x}$ or s).
Typically, you're interested in a value of a parameter. But you can't know it. So you *estimate* it with a statistic, based on the sample.

There are two means and two standard deviations.
The sample mean $\bar{x}$ and sample std deviation s are statistics.
Define the population average μ as the sum of all values in the population ÷ the number of subjects in the population. (parameter).
It turns out $\bar{x}$ is an unbiased estimate of μ.
That is, $\bar{x}$ is neither higher nor lower, on average, than μ, if we sampled repeatedly.

7. Categorical and quantitative variables.
For a quantitative variable, the responses are all numbers and the difference between two observations has a natural interpretable meaning. For categorical variables there is no such meaning to the difference between two observations.
The line between the two terms can sometimes be a bit blurry.
e.g. gender of baby would be categorical.
height, weight, and number of weeks would be quantitative.
eye color, birth type, or pain medication used might be examples of categorical variables here with multiple possibilities.

8. Statistical significance and Testing.

According to the CDC, 4 million babies were born in the U.S. in 2014 and 10% were born preterm (< 37 weeks). Suppose you take a simple random sample (SRS) of women with HG and you want to test whether the low birthweight proportion among women with HG might really be different from 10%.

Suppose in the sample of n=254 mothers with HG, p = 39/254 (15.35%) are preterm.

You want to test whether something like this could reasonably have happened just by chance alone, if the populations were actually identical with respect to delivery time.

Otherwise we conclude that the two population proportions are probably not equal, i.e. the difference observed is *statistically significant*.

There are different tests, but we'll just talk about the Z-test (or normal test) for now.

Assumptions:

SRS (or obs are known to be independent)

AND n is large (or pop is known to be normally distributed).

For testing proportions, there should be ≥ 10 of each type of response in the sample.

Here we have 39 preterm and 215 not preterm.

We will talk later in the course about these assumptions and also about the t test. If n is small, pop. is normal, and σ is unknown, then use t instead of Z.

After checking assumptions, the remaining steps in testing are

* stating the hypotheses,
* computing the test statistic (Z in this case),
* computing the p-value, and
* concluding.

9. Null and alternative hypotheses.

Let π be the proportion preterm in the population from which the sample was drawn.

Null hypothesis (Ho): $\pi = 10\%$.

This means that any observed difference between the sample proportion, p, and 10%, was due to chance alone. Usually we specify these hypotheses numerically.

Alternative hypothesis (Ha): $\pi \neq 10\%$. Difference is not due to chance alone. (2-sided test.)

Or Ha: $\pi > 10\%$. Or Ha: $\pi < 10\%$. (1-sided tests). We will talk about this next lecture.

When in doubt, do a two-sided test, unless there is a specific reason to do a 1-sided test.

10. Z-statistic.

A test statistic is a summary of the strength of the evidence in your data.

Z-statistic here = (p – 10%) ÷ SE.

SE means Standard Error. We will talk about ways to get the SE either analytically or via simulations in a bit. For proportion problems like this, $SE = \sqrt{[\pi(1-\pi)/n]}$.

Here, analytically, the SE would be $\sqrt{[.10 \times .90 / 254]} \sim 1.88\%$.

$Z = (15.34\% - 10\%)/1.88\% = 2.84$.

The book calls Z a *standardized* test statistic.

It indicates how many SDs the observed statistic is above its hypothesized value under Ho.

The book also calls the SE the "standard deviation of the null distribution" but it is usually called the standard error or SE.

A value of Z far from 0 (more than 2 or less than -2) indicates strong evidence against the null hypothesis. A value of Z between -2 and 2 indicates weak evidence against the null.

$|Z| > 3$ indicates very strong evidence against the null.

11. Simulating null distributions and Standard Errors.

We observe p = 15.34% in our sample, and under Ho, the population percentage π = 10%. So we see a difference of 5.34%. This is our quantity of interest, and it is usually a difference like this. We want to see if that quantity of interest, 5.34%, is bigger than what we'd expect by chance under the null hypothesis.

The Standard Error (SE) is the standard deviation of the quantity of interest under the null hypothesis.

Many stat books just tell you the formulas to get the SE. Your book is different. They want to emphasize that in many cases you can estimate the SE by simulations.

In this example, under Ho, women with HG are just like the rest in terms of probability of delivering preterm. We have a SRS of size 254 from a population with π = 10% having preterm delivery. We can simulate 254 draws on the computer, where each draw is independent of the others and has a 10% chance of being preterm, and then see what results we get. In R, I did

```
x = runif(254)
y = (x<0.1)
p = mean(y)
```

The first time, I got p = 0.1259843. 12.60%.

I tried it many times, and here is what I got.

```
a = rep(0,10000)
for(i in 1:10000){ x = runif(254); a[i] = mean(x<.1)}
hist(a*100,main="simulated preterm percentages", nclass=100,
      xlab="percentage preterm in sample")
abline(v=15.34)l
mean(abs(a-.1)>.0534)       ## 0.0051
sd(a-.10)                   ## 0.01885409
sqrt(.10 * .90 / 254) ## 0.01882367
```