DEPARTMENT OF POLITICAL SCIENCE
AND
INTERNATIONAL RELATIONS
Research Methods Posc 302

SAMPLING AND STATISTICAL INFERENCE

I. TODAY'S SESSION:
- A. Samples
  - 1. Basic concepts
- B. Statistical inference
  - 1. Purpose and overview.
- C. Some writing tips:
  - 1. "The results show there is a relationship...," not "The results show there is **such** a relationship..."
    - i. Usually "such" isn't needed. Try to economize. The second draft (what you turn in) should have about half the number of words the first one.
  - 2. "Those **who** voted for Clinton...," not "Those **that** voted for Clinton..."
  - 3. "Researchers conducted a nation-wide survey after the 1996 national election in which subjects were asked ..." not "In a study conducted following the 1996 presidential election, a massive survey was conducted. Surveyors asked their subjects several questions about..."
  - 4. "In a recent survey President Bill Clinton's character ...," not "In a recent survey in regards to the 1996 presidential election, President Bill Clinton's character..."

II. SAMPLES AND POPULATIONS:
- A. From Class 13 notes
- B. Overview:
  - 1. What does an observed relationship between, say, party identification and attitude toward capital punishment say about politics in America>
    - i. A small, perhaps a very small, part of the answer depends on whether or not we can **generalize** the observed relationship to the public as a whole.
    - ii. After all, in most instances our knowledge of political attitudes and behavior rests on **samples** of Americans.
    - iii. So besides worrying about measurement we need to think about samples and the sampling process.
  - 2. The analyses we've been considering are based on samples of 1,000 to 1,500 respondents drawn from the public as a whole.
  - 3. Are the individuals and the relationships they exhibit representative of citizens in general?
- C. Some key terms:

1. **Population**: a well-defined collection of units of analysis such as the American states, the people living in Delaware, the countries in Asia, the students taking this course.
2. **Sample**: a subset of the population drawn in some fashion. We can have, for example, a sample of the states, a sample of Delawareans, or a sample of students taking this class.
3. **Estimation**: guessing at the value of a population characteristics using a sample statistic.
4. **Random sample**: members of the population have some known probability or chance of being included in the sample.
5. **Simple Random Sample** (SRS): members of the population are selected in such a way that each member has an <u>equal</u> chance of being included.
    i. Imagine everyone one in the population being assigned a unique number and then some totally random process draws a few numbers from the list. The people corresponding to those numbers are included in the sample.
    ii. Presumably, if everyone has one and only one unique number and the numbers all have the same chance of being "chosen," then every person's chance of ending up in the sample is as good as everyone else's.
    iii. But in practice it is difficult to draw simple random samples from large populations.
        1) See the discussion in Johnson and Joslyn, ***Research Methods***, pages 174 to 184 for a discussion and examples.
6. **Sample sizes, N**: the number of units or elements in the sample.
    i. Frequently, N will refer to the number of respondents in a cross-tabulation table.
    ii. We have briefly discussed samples sizes and will do so in more detail when covering statistical inference.
    iii. For now note that large samples are not necessarily essential for producing "valid" results. In fact, N probably has more to do with "reliability."
7. **Parameter**: a statistical characteristic (e.g., mean, standard deviation) of a population.
    i. Two major goals of statistics are to estimate the magnitude of population parameters and to test hypotheses about them.
    ii. Parameters are usually denoted with Greek letters.
    iii. Example:
        1) Suppose we are examining the relationship between X and Y in cross-tabulation.
        2) It is possible to apply numerical indices to measure the strength and nature of the observed relationships.

3) These indices are parameters.
4) Suppose is ψ a measure of the association or connection between one variable and another.

8. **Sample statistic**: a statistical characteristic of a sample or batch of values.
   i. Sample statistics are generally used to make inferences about the corresponding population parameter.
   ii. Example: in the preceding paragraph ψ stands for a numerical indicator of the strength of a relationship between a X and Y in a population. Then, the same character with a "hat" means a sample **estimator** of that characteristic. That is, $\hat{\psi}$ is an estimator of ψ.

III. MORE ON POPULATION PARAMETERS AND SAMPLE STATISTICS:

A. Case 1: what was the average or **mean** family income in the United States in 1998?
   1. Although we may not know the exact amount, we can denote this quantity by the Greek letter μ.
   2. Now suppose we take a random sample of 500 Americans and ask them to report their annual family income in 1998.
   3. Suppose in addition that each person in sample responds and we find the average of these reported incomes.
   4. This number is the sample income, which we can call for short $\hat{\mu}$.
   5. To be even more concrete suppose the average family income of the 500 families is $37,500. Then, $\hat{\mu} = 37{,}500$
   6. This number, $37,500 is our best guess as to the corresponding value for the population. Or, stated a bit more formally, the estimate of family income based on our sample of N = 500 respondents is $37,500.

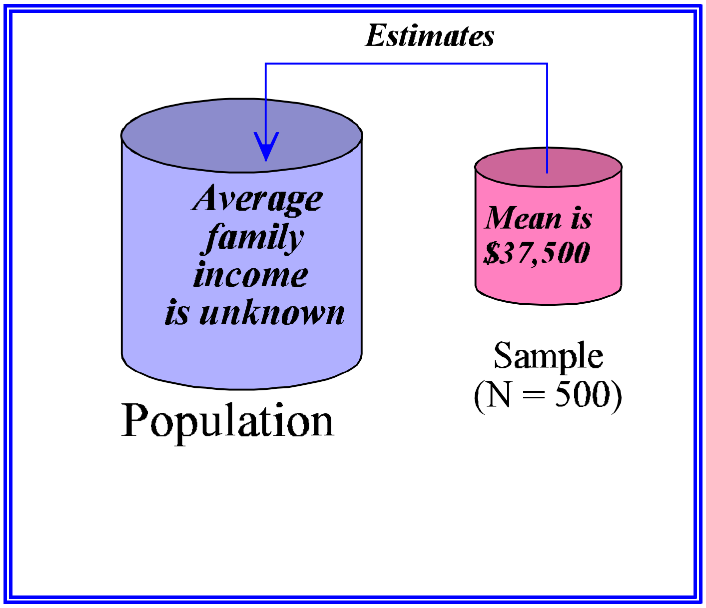

Figure 1: Sample Estimate of Population Mean Family Income

B. Case 2: Was there a racial gap in turnout in the 1996 election United States? If so, how big was it?

1. Suppose we have a cross-tabulation between sex and vote in 1996.

Race/Turnout	White	African-American	Totals
Voted	70.0 140	70.0 49	189
Did not Vote	30.0 60	30.0 21	81
Totals	100.0 200	100.0 70	200

2. This table is in a sense a sample statistic. But since it contains so much information we might want to combine or summarize the information in a single number or index, which we'll call $\hat{\theta}$.
3. Suppose $\hat{\theta} = 1.0$.
4. This number is an estimator of the population value of θ.
   i. The population value, which is usually unknown, could be any number such as 0, .23, 1.9, 22.7, or anything. But our estimate of its value based on the sample reported in the table is 1.0.

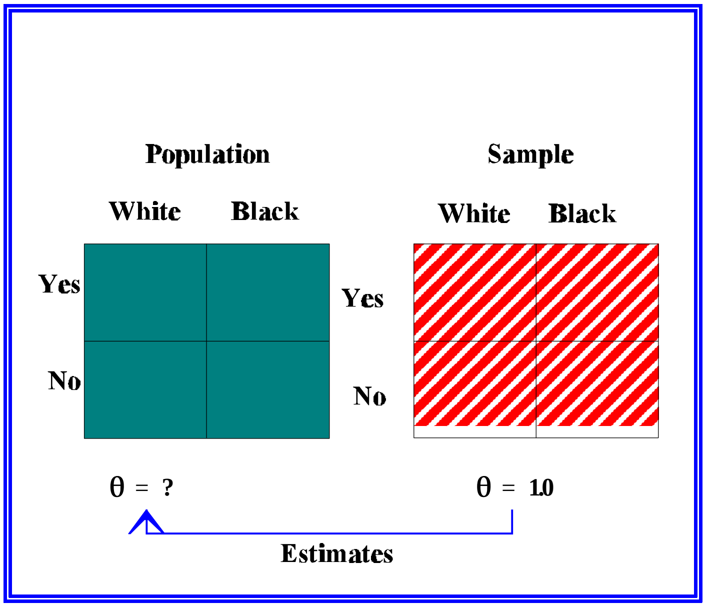

Figure 2: Sample Cross-tabulation Estimates Corresponding Population Relationship

C. Statistical inference:

1. As noted above, statistical inference has two goals.

i. Test hypothesis about population parameters on the basis of sample statistics.

1) Example: test the hypothesis that there is **no** relationship between race and voting in the population, which if true implies that any sample association we happen to observe might be due purely to chance.

ii. Estimate the value of population parameters.

1) Attempt to judge or guess the numerical value of a parameter.

a) It is 1.0 or 2.0 or 0 or what?

IV. PROBABILITY:

A. Background of hypothesis testing.

1. Remember: the goal is to test a statement or hypothesis about an unknown parameter using only sample data and mathematical theory.

B. Consider a couple of problems:

1. From "Rosencrantz and Guildenstern Are Dead" a play by Tom Stoppard. Rosencrantz (Ros) and Guildenstern (Guil) are flipping coins and keeping track of heads and tails:

- Ros: Eighty-five in a row--beaten the record
- Guild: Don't be absurd.
- Ros: Easily
- Guil (angry): Is that it, then? Is that all?
- Ros: What?
- Guil: A new record? Is that as far as you are prepared to go?
- Ros: Well...
- Guil: No questions? Not even a pause?
- Ros: You spun them yourself.
- Guil: Not even a flicker of doubt?
- Ros (aggrieved, aggressive): Well, I won--didn't I?
- Guil:...And if you'd lost? If they had come down against you, eighty-five times, one after another, just like that?
- Ros (dumbly): Eighty-five in a row? Tails?
- Guil: Yes! What would you think?
- Ros:...Well....(Jocularly.) Well, I'd have a good look at your coins for a start!

i. The issue here is that one of players, Rosencrantz, has tossed a coin 85 times and come up with heads 85 times!

ii. More concisely, the number of heads in 85 tosses is 85.

iii. The other player wonders how likely that result is **if the coin or the game is fair**.

iv. How might statisticians translate this situation into terms that lend themselves to systematic analysis?

v. To say that the coin is “fair” is to claim that the probability of getting a heads on a single toss is ½ = .5, which is the probability of getting a tails on a single toss.

1) That is, a fair coin is one such that P = .5, where P is the probability of getting heads on a toss.

vi. Now, suppose we want to state the hypothesis “the coin is fair” as precisely as possible. Such a hypothesis might be symbolized as

$$\boxed{H_0\colon P[H] = .5}$$

1) In words this “equation” simply says that the probability of obtain “heads” on a single toss is .5 or ½.

vii. The play’s protagonists wonder if the hypothesis is **true**.

1) That is, they ask (in effect) does the assertion that P[H} = .5 really apply to this coin?

2) If it does, we can say that the hypothesis is **accepted**.

3) If, on the other hand, we conclude that the coin is biased, we might **reject** the hypothesis.

viii. This hypothesis refers to the coin, of course, but what's really at stake is the long-run probability of getting heads in many, many tosses.
   1) In fact, when one speaks about the "fairness" of coin one refers to the chance of getting heads or tails on any particular toss in an infinitely long series of tosses of the coin.
   2) Thus, one way to make sure that the coin is in fact honest (i.e., P[H] = .5) would be to toss it an infinite number of times and record the number of heads and tails.

ix. But of course such an "experiment" is not possible.
   1) Instead one has to flip the coin a finite number of times, say 85 times, and count heads or tails.

x. We can consider this finite number of tosses a **sample**.
   1) In the play the sample size is N = 85, since Rosencrantz tossed the coin 85 times.
   2) The result of these 85 flips was HHHHHH...HH.

xi. The problem:
   1) Can one say anything about the hypothesis likelihood of being true based on just a sample.

xii. In short, we use sample results (the number of heads in N tosses in this case) as evidence for or against the hypothesis, which in this case is P[H] = .5.

2. Take something more commonplace. A lawyer has just called you because she has been troubled by recent trends in the state supreme court where capital punishment cases are adjudicated. Although polls report that 25 percent of the people in the state oppose the death penalty, she thinks that opponents are under-represented on the juries dealing with homicide cases. Why would this be a problem. She surmises that people who favor the death penalty are more likely to return guilty verdicts than people who oppose capital punishment, even when everything else is equal. So it's important that people who voice opposition to capital punishment be included on juries involving murders and other capital crimes. But in a recent case, only one out of 12 jurors were expressed any doubts about capital punishment. The other 11 favored it. She feels there is evidence of bias that works against defendants because this case suggests that the juries are being stacked by people who are likely to convict. Is there any evidence, she wants to know, of discrimination?

   i. Can you imagine how a precise hypothesis for this situation could be formulated?

   ii. How about this: if jury selection is totally unbiased the proportion of jury members who oppose the death penalty should be .25 (if the polls are accurate).

1) We have a panel of prospective jurors. We draw one at random. What is the probability that this person is actually selected?
2) A way of approaching the question is to formulate the hypothesis the proportion of jurors in capital punishment cases will be .25 or, more precisely, P[~CP] = .25, **if the selection process is unbiased**. (The symbol “~CP” simply means “against capital punishment.”)
3) Seen from this point of view, jury selection might be analogous to coin tosses, only the probability of the outcome of interest is .25.
4) And our sample now consists of 12 tries (“tosses”) of which only 1 is a “success.” That is, only 1 person out of 12 had doubts about the death penalty.
5) Is this sample result–1 out of 12–consistent with the hypothesis that opponents have a .25 chance of being selected.
    a) If we answer, “yes” the result is consistent, then we accept the hypothesis and conclude that the jury selection system is not biased.
    b) But, if we answer “no”–the data are not consistent–the we reject the hypothesis and claim discrimination.

iii. In short, we could use this simple sample to “test” the hypothesis that the probability of an opponent being selected is .25.

V. TESTING HYPOTHESES:

A. Let’s modify the coining tossing example:

1. Question (hypothesis): Suppose you are flipping coins with someone. A reasonable question is: is the person honest?
2. How should one "operationalize" the notion of honesty?
3. The question really asks: is the probability of getting a "head" on a single flip of a coin 1/2 = .5? The hypothesis suggested by the question, in other words, is that P[heads] = .5 More formally,

$$\boxed{H_0: P = .5}$$

4. H_0 is called a **null** hypothesis
5. In a series of , say, 10 flips of a coin how likely is it that a person will get 9

heads?

6. In order to answer this question you have to decide two things:
    - i. What is "likely"? That is, what is your definition of likely or probable and conversely what is your definition of improbable?
    - ii. Given this definition, you then have to decide what outcomes--what number of heads in 10 tosses--that you will consider probable and what number you will believe to be improbable.
7. Stated another way you have to form a mental picture like this:

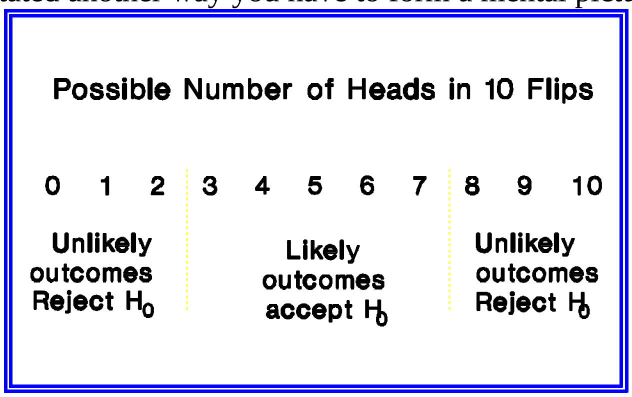

Figure 3: **Areas of "Acceptance" and "Rejection"**

8. Next, observe an <u>actual</u> set of ten tosses and count the number of heads.
9. Finally, make a decision using the guidelines presented above: if the <u>observed</u> number of heads fall in the region of rejection because this result is improbable, then reject the hypothesis (P[heads] = .5) and conclude that the coin is biased, the person is crooked, etc.
10. If, for example, someone tosses a coin 10 times and obtain 9 heads then you might suspect that the hypothesis of honesty is untenable.

VI. NEXT TIME:
- A. Sampling and statistical inference.
    - 1. How to test for the “statistical significance” of an observed relationship.
- B. Reading:
    - 1. Johnson and Joslyn, ***Research Methods***, Chapter 7, especially pages 343 to 346.