# Math 161

Glen Pugh

September 12, 2007



1 Our Focus So Far

2 Chapter 4: Sample Surveys in the Real World

3 Chapter 5: Experiments, Good and Bad



# Our Focus So Far

## Central Idea

- We wish to determine what proportion $p$ of a population of individuals possesses a particular characteristic.
- To do so, we select a Simple Random Sample (SRS), thus avoiding bias, of a size $n$ large enough to keep the variability small.
- From our SRS of $n$ individuals we calculate $\hat{p}$, the proportion of the sample which possesses the characteristic:

$$\hat{p} = \frac{\text{(\# of individuals in the sample possessing the characteristic)}}{n}$$

- $\hat{p}$ is an *estimate* for the unknown value of $p$.



## The Confidence Statement

- Once we have our estimate $\hat{p}$, we use a confidence statement to say how close $\hat{p}$ is to $p$ and how confident we are of our result:

  "Our result is accurate to within $E$ percentage points nineteen times out of twenty",
- Meaning:

  "If we repeatedly took SRSs in the same way and calculated the resulting $\hat{p}$ values for each SRS, 95% of those $\hat{p}$ would be within $E$ percentage points of the true value of $p$.
- Here $E$ is called the margin of error. As long as the population is large (at least 100 times larger than $n$), the quick method can be used for estimating the margin of error:

$$\text{margin of error} \approx \frac{1}{\sqrt{n}}$$



## Example

In a random sample of 1002 adult Canadians of voting age, 701 said that they voted in a recent election, although voting records show that only 61% of eligible voters actually voted.

1. What is the population?
2. What are the individuals?
3. What steps were taken to avoid bias in the survey?
4. Let $p$ be the proportion of eligible voters who actually voted. Calculate $\hat{p}$ (as a percentage, rounded to 1 decimal).
5. Use the "quick method" to determine the margin of error in the survey (assuming a 95% confidence level).
6. State your results, including a confidence statement.
7. Explain your confidence statement to your friend using plain English.
8. If we wanted a margin of error of 2% (at the 95% confidence level), approximately how large a sample would be required?

## Clarification from Last Day

A more precise definition of margin of error: the margin of error in a sample statistic $\hat{p}$ (at a 95% confidence level) is the maximum difference between the population parameter $p$ and the sample statistic $\hat{p}$. However, this maximum is valid for only 95% of samples taken; the difference between $p$ and $\hat{p}$ will exceed the margin of error in 5% of samples.



## Clarification from Last Day

Question: How can we say that we are within the margin of error in 95% of samples if we take only one sample?

Answer: Taking a sample of size $n$ from a large population is similar to flipping a coin $n$ times: we don't know exactly how many heads we'll get, but we can give a range for the number of heads which will come up 95% of the time. The possible outcomes for coin flipping is entirely predictable and follows the bell (normal) curve. We'll see more on this later.



# Chapter 4: Sample Surveys in the Real World

## Errors in Sampling

We saw that SRSs reduce bias and give control of variability, but they don't solve all of our problems. Other sources of error can cause inaccuracy in our statistical studies.



## Sampling Errors

- Sampling Errors are errors resulting from the act of sampling itself. These include
  - Random Sampling Error: Difference between sample statistic and population parameter; strictly due to chance, and accounted for by the confidence statement. **The margin of error captures only Random Sampling Error**
  - Undercoverage: Sample not drawn from the entire population. The sampling frame, the list from which samples are drawn, does not include the entire population

    Example: Recall example from week 1: A 1996 Ontario telephone poll of people who had gone off welfare found that 62% of these people listed finding a job as their reason for no longer collecting welfare. (Sampling frame neglects all former welfare recipients not reachable by phone.)



## Nonsampling Errors

- Nonsampling Errors: Errors unrelated to taking samples: Data entry, calculation error, response error (intentional or not), non-response.

  Example: 2001 Canadian Census: 179,788 households classified as non-response.
- Question Wording: Avoid leading/vague/open-ended questions:

  Example: BC Treaty Referendum of 2002: Do you think . . .
  1. Private property should not be expropriated for treaty settlements. (Yes/No)
  2. The terms and conditions of leases and licenses should be respected; fair compensation for unavoidable disruption of commercial interests should be ensured. (Yes/No)
  3. Hunting, fishing and recreational opportunities on Crown land should be ensured for all British Columbians. (Yes/No)

## Handling Problems

Dealing With Non-Sampling Errors  Survey results generally adjusted to account for non-sampling errors:

- Census: impute (substitute) data for missing households using similar households.
- Weight responses of individuals in your sample differently so that sample makeup matches that of population as much as possible.



## Real World Samples: Stratified Samples

We may want to ensure that subgroups of the population sharing common characteristics are represented in our survey. To do so, break the sampling frame up into strata, the distinct groups you want represented, and take an SRS from each stratum.



## Stratified Samples: Example

Example: Women make up 23.5% of the first year engineering class at the University of Alberta. Suppose the university wants to survey 200 first year engineering students to determine their level of satisfaction with the program, and the administration would like males and females represented equally in the results. The first year engineering class will be split into two strata: males and females, and an SRS of size 100 will be drawn from each.

If the first year engineering class has 450 students, what is the chance (i.e. the probability) that a female in the first year engineering class will be selected for the survey?

## Question Everything!

Don't believe everything you hear or read: survey results can be inaccurate and misleading if the methods used are unreliable. Ask questions!

- Who conducted the survey?
- What sampling method was used? (SRS?)
- How large was the sample?
- See p.65 of text.



# Chapter 5: Experiments, Good and Bad



## Experiments *vs* Observational Studies

What's the difference? So far our discussion has focused on taking SRSs from populations and gathering response variable data from those samples. These are purely observational studies: no attempt was made to influence the responses of the individuals.

An experiment, on the other hand, imposes a treatment on the individuals, called the subjects, in an effort to determine if explanatory variables cause change in the response variable.

## Experimenting Badly

- A good experiment allows one to impose a treatment on the subjects and determine with some certainty whether the explanatory variables cause changes in response variables.
- Problems occur when it is unclear what caused changes in response variables.



## Experimenting Badly: Example

- Example: Studies have found that more education leads to longer life (reduced mortality). Here

$$\begin{aligned} \text{subjects} &= \text{people} \\ \text{explanatory variable} &= \text{number of years of education} \\ \text{response variable} &= \text{age at death} \end{aligned}$$

- Could there be other explanatory variables affecting the response variable? Say, economic status, health of the subjects?
- The other (unaccounted for) explanatory variables are called lurking variables.
- Two variables are said to be confounded when their effect on the response variable cannot be distinguished from each other.



## Good Experiments

Good experiments should isolate the effect of explanatory variables on the response variable, not allowing lurking variables to confound the explanatory variables.

## Randomized Comparative Experiments

- Randomized Comparative Experiments: Divide test subjects into two groups using chance.
- Give one group the treatment, the other no treatment. Compare response variables of each group.
- Logic: By dividing the subjects into groups using chance, the two groups are similar in all respects except for the treatment received. Therefore, any differences in the response variable must be due to the treatment.
- The group receiving no treatment (given the placebo in this case) is called the control group since it controls the effect of lurking variables.



## Randomized Comparative Experiments: Example

Example:

- Elevated levels of the amino acid homocysteine make arteries stiffen and clog. It is thought that elevated homocysteine levels combined with low vitamin B12 and folic acid levels (which help brain function) lead to decreased cognitive ability in advanced age.
- In a study published last year, 818 adults aged 50 to 70 with high homocysteine levels were randomly divided into two groups of equal size. One group was given a daily 0.8-milligram folic acid supplement, the other a placebo (a fake drug which looks the same but has no active ingredients). After three years, memory was significantly better in the folic acid supplement users.
- Sketch an experiment design diagram for this experiment.



## Statistical Significance

- In an experiment, how big a difference is enough to reach a solid conclusion?
- Statistical Significance: an observed result that would rarely occur by chance is called statistically significant.
- One should expect some difference between the control groups and treatment groups because subjects were randomly assigned to each. Assuming those groups are large enough to keep variability small, large differences between the control groups and treatment groups are most likely due to the treatment.
- More to come on this later (Chapter 22 on Significance Testing).

# Observational Studies: a Weak Second Best

- Not always possible to carry out experiments, like in the case of the study of whether more education leads to longer life.
- Such observational studies still useful if we can compare the treatment group with a control group.
- For example, establish a control group similar in all respects to the treatment group except for level of education.