MATH 243 PARTIAL LECTURE NOTES (31 MARCH 2008)

N. CHRISTOPHER PHILLIPS

1. Introductory Examples

What is this course about?

Example: Do test preparation courses improve SAT scores?

Basic idea: Compare the SAT scores of some people who took test preparation courses and some who didn't.

More specifically, compare the mean (average) SAT scores in the two groups.

Need care:

- Choose groups well. (If people in one group are more motivated, can't trust results.)
- Choose enough people. (One person in each group won't do.)
- "Statistical significance": If one mean is bigger, could it be due to chance?
- Importance: If we are convinced there is a difference, but it is only one point on average, who cares?

We will learn:

- How to look at data.
- How to try to address questions like the SAT score problem. Other examples: public opinion polls, determining effectiveness of medical treatments, ...)
- Some of the hazards in using statistics to try to answer questions like these. (The SAT score problem presents major difficulties, but some others are easier.)

What a statistical conclusion really means:

The statistician said, "We conclude, at significance level $\alpha = 0.05$, that students taking Jennifer's SAT preparation course get higher SAT scores than students taking no preparation course."

The statistician meant, "We did an experiment .... If it were **not** true that students taking Jennifer's SAT preparation course get higher SAT scores than students

Date: 31 March 2008.

taking no preparation course, then the observed outcome of the experiment would be very unlikely. In fact, it would have a probability of less than $\frac{1}{20}$ of occurring."

An example and a warning:

I take a coin, presumed to be fair, and flip it 10 times. Suppose it comes up tails every time.

The statistician can say, "If it were **not** true that the coin has tails on both sides, then the observed outcome would be very unlikely. In fact, it would have a probability of less than $\frac{1}{1000}$ of occurring."

(The precise probability is $\frac{1}{1024}$.)

Do you really think the coin has tails on both sides?

We have to put up with conclusions like this because they are the best we know how to produce.

"There are three kinds of lies: lies, damned lies, and statistics."

(Mark Twain attributes this to Benjamin Disraeli.)

"Statistics is never having to say you are certain."

(Outside a UO statistician's office door, Fall 2003.)

Example 1: Misleading statistics.

One year, at every airport, Alaska Air had a *greater* proportion of on time flights than America West. But overall, Alaska Air had a *smaller* proportion of on time flights.

How it happened: Alaska Air has many more flights at airports with bad weather.

(From the previous edition of the textbook.)

Example 2: More misleading statistics—ask the right question.

(Old) ads for Tylenol: "Recommended by more doctors than the leading brand of aspirin . . . ."

For the leading brand of aspirin: "Preferred by doctors 2–1 over Tylenol, 3–1 over . . . ."

(When asked what to do about a headache, most doctors simply say to take an aspirin.)

Example 3: Yet more misleading statistics—beware of the "lurking variable".

At the (imaginary) Wang's Department Store, the advertising agency wants to take credit for a 100% increase in bathing suit sales over the previous six month period.

The previous six month period was October through March, the current one April through September.

Example 4: Misleading *non*statistics—beware of anecdotes and the placebo effect.

Nonscientific health products, such as homeopathic remedies, are often advertised via anecdotes ("testimonials"): "Mr. X used our preparation and was cured of . . . ."

Properly controlled double blind experiments usually show that homeopathic remedies are no better than dummy pills containing no active ingredients. (But *both* are better than taking nothing at all!)