Economics 120A

1. Introduction

Graham Elliott

Plan for Today

1. Give overview of ideas the course and course sequence

2. Go through syllabus and materials available in Canvas.

    (a) Syllabus
    (b) Reading file

What we expect to get out of the course

Statistical Literacy

"Statistical Literacy is the ability to understand and critically evaluate statistical results that permeate our daily lives -- coupled with the ability to appreciate the contributions that statistical thinking can make in public and private, professional and personal decisions."
(Wallman, former ASA president).

There will be plenty of math, but this is the goal of the course.

The Big Picture

What are big 'themes' or areas of statistics?

1. Describe a dataset. What does it tell us?

2. Forecasting/Prediction of some outcome given observable features. How does the data help with decision making.

3. Causal inference – if I change one variable what happens to other variables?

4. Does the data agree with how I think things work?

The Big Picture

Consider how we learn about the world.

1. We first observe phenomena

2. We construct conjectures to explain such phenomena (hypotheses or theories).

If they are consistent with each other, this is evidence in favor of the hypothesis.

3. We then examine other implications of the theory.

4. We test the predictions in experiment or any way we can.

If the predictions of the theory are consistent with the hypothesis this is even stronger support for the hypothesis. This is essentially what we mean by the `scientific method'.

Examples

(a) Einstein's theory of relativity.

General theory in 1915. Has the implication that light should bend near large mass objects. Eddington realized a test of the theory - data that could be used to examine the implications of the theory - was available through a solar eclipse in 1919. He measured the light and found what the theory suggested.

(b) New England Fishing Industry.

Catches were falling every year on the NW banks. Fisherman had a new theory each time (cold water, more storms ...), but in reality it was overfishing. Cod stocks collapsed killing the industry.

It is really a formalization of the `common sense' that we apply to everyday situations. Each step of the process is important, but notice that two of the steps involve collecting information on and examining `observations'.

Basic Quest : Understanding what observation (data) tells us about the world.
Statistics : This is the formalization of the quest.

Why Formalize?

To avoid mistakes, and to be able to be more precise in our understanding.

(a) Some problems are too complicated, e.g. models of the economy have a lot of interrelated parts.

(b) Some results too subtle to do otherwise. The advantage of one investment strategy over another might be small and hard to see, even though it adds up to much greater profits.

(c) We make mistakes, for example confirmation bias.

"When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind" -Lord Kelvin (Victorian Physicist - Kelvin temperature scale named after him).

Some examples to think about

Some Examples.

1. SAT averages - Average Math scores for 2001. California average is 510, ranked equal 35th with Alaska and Maryland. The top average was 603, recorded by both Iowa and Utah. South Carolina came in with an average of 488 to claim last place. What do these averages mean for education levels?

2. Average life expectancy has been rising dramatically over the last few hundred years. What does this mean. What does a life expectancy of 30 mean for a population?

3. Tests indicate that face recognition technology has a very high probability of picking out a bad guy in say an airport check point, and a high probability of passing a good guy. Should we use it?

4. A small scale clinical trial for a drug cures slightly more patients than the leading alternative. Should it be adopted?

Some Examples

5. Many sociologists have found high correlations between a penchant for watching violent movies/playing violent video games and being violent themselves. Does it make sense to ban these types of activities to control violence in society?

6. Mutual funds often point to past returns as a reason to invest (they did better than the market). Theory suggests that they do not do better than the market, past returns are no guarantee of future performance. Who is right?

7. Guiliani claims that crime rates in New York show that he has what it takes to take on terror. Do the numbers back him up?

8. Kidney cancer rates are lowest for rural counties in the midwest, west and south. Evidence of clean country living as a factor in cancer?

9. Zillow has tons of data and forecasted prices for homes. Why not automate home buying and flipping?

Overview of Statistics using Cards

Theory

Model/World

Data

Deck of cards

P[Red] : P[Black] = $\frac{1}{2}$.

1st Black

2nd Black

3rd Black

4th Black

Major Themes

1. Informally analyzing data (particularly economic data) to get an idea of how to present data in a clear way (and not get fooled by those trying to pull the wool over our eyes) and also understand some of the complications of informally analyzing data. (looking at data plays a large role in learning, also a large role in `popular' analysis. Also nice to see why we are going to put in so much work into doing it formally).

2. Probability. Probability plays a dual role in statistics a) We write down our models of the world in terms of probability models, so we have to be able to do probability theory. Such is life. A good example is the medical experiments b) We use probability to say how sure or unsure we are of our conclusions in formal statistics. We say things like "I am 90% sure that the theory is untrue".

3. Sampling. We will examine how data is collected, and why data is collected in the way it is. We will examine good and bad methods for data collection, and what they do for or against our ability to learn from the data.

Census or Sample?

Before we get going, however, it is of interest to examine why, if we want to learn about the world, we do not just go out and observe everything.

Such a situation where every possible piece of data is collected (or at least the attempt is made to collect every piece of data) is called a census.

In most cases, we do not conduct a census as it is not feasible for one reason or another.

a) Resources are limited. E.g. Consider the quarterly GNP numbers.

b) Scarcity of observations. E.g. Suppose we want to determine the effect of monetary policy on recessions.

c) Destructive testing. E.g. Match Factory

d) Sampling may be more accurate. E.g. US Census.

A Caveat

The motto of the Royal Society is 'Nullias in Verba', roughly 'Take nobodies word for it'.

Studies can be exceedingly poor and misrepresent reality. We will see how. A sophisticated user will understand when there are likely to be problems, what to look for. We will see this in this course.

We will go beyond the policy of rejecting studies that do not fit with what we believe, and move to a deeper understanding.