FMS1204S: Fraud, deception and data

Week 3

Surveys and opinion polls

Issues in Surveys

- In a survey or opinion poll we'd like to know a characteristic of some population of interest but it isn't feasible to measure that characteristic for every individual in the population.
- A survey, in an abroad sense, is a random sampling. In stead of measure every individual in the population, a survey takes a random sample from the population and measure the characteristic of the individuals in the sample.
- One of the most difficult parts of sampling is to obtain a truly random sample representative of the population of interest.
- Modern survey methods are very sophisticated, we will not discuss the survey methods. However we will learn something about how to spot a bad poll, and about how data from bad surveys and opinion polls can be used to mislead.

An Example of biased sampling

In a survey a random sample of people is taken and asked for a list of friends, the average number of friends of sample is computed.

Another random sample is taken from the friend lists of the first sample and also asked for a list of friends and the average number of friends in this sample is also computed.

The people sampled at the second stage have, on average, many more friends than do the people in the original sample.

This suggests that, on average, your friends are more popular than you are. Is this correct?

How can surveys be misleading?

- Some organizations will run polls (often these days they are online polls) where there is an obvious selection bias (that is, the respondents aren't randomly chosen from the population of interest).
- Respondents might be people who are motivated enough to visit a website, or to call in for example.
- In some cases there is a deliberate attempt to use such bias to generate results in a certain direction, with the intent to publicize the findings of the survey as representative of public opinion.

Example

- The newspaper *USA Today* ran a call-in poll in which 81 percent of the more than 6,000 respondents agreed that "Donald Trump symbolizes what made the U.S.A. a great country."
- However, 72 percent of the favorable calls came from two telephones in one insurance company office.
- Since the people who call in are not typical of the general population (and since the same person can vote more than once) a poll like this cannot be considered a serious effort to measure public opinion.

How can surveys be misleading? (cont.)

- Another way to bias a poll is through the wording of the question that is asked.
- Example: the Fox News network in the US asked some of the following questions on a recent survey. The wording of the questions is hardly politically neutral:
  - Do you think Barack Obama's travel and speaking schedule makes him look more like he is a candidate on the campaign trail or more like he is the president of the United States?
  - Do you think President Obama apologizes too much to the rest of the world for past U.S. policies?
  - Do you think the Obama administration is proposing more government spending than American taxpayers can afford, or not?

How can surveys be misleading? (cont.)

- There are also question order effects. The questions on the Fox News survey we just mentioned were followed by a question on health care reform: “Based on what you know about the health care reform legislation being considered right now, do you favor or oppose the plan?" The results of the Fox survey showed unusually low support for health care reform (compared to other polls).
- The pollster Nate Silver writes on his blog: (http://www.fivethirtyeight.com/2009/10/question-order-may-bias-fox-news-health.html) "A respondent who hears these questions, particularly the series of questions on the national debt, is going to be primed to react somewhat unfavorably to the mention of another big Democratic spending program like health care. And evidently, an unusually high number of them do."

How can surveys be misleading? (cont.)

- Other problems that can occur with opinion polls (not an exhaustive list):
  - Non-response bias
  - People don't always tell the truth!
  - Fatigue effect
- Problems in the way results are reported:
  - Overinterpreting trends
  - Misrepresenting the question that was asked
  - Multiple comparisons

  Some of your readings for next week go more deeply into these issues.

Groups and readings for next week

Group 1

- Goldacre, Ben (2007). When the facts get in the way of a story. *Significance*, 4 (2), pp. 84–85.
- A front page story appeared in the UK newspaper *The Times* about a doubling of cocaine use in children. This article discusses the survey data on which the story was based. Questions to address specifically from the reading:
  - Why is the story misleading? Explain what is meant by a p-value, clustering and multiple comparisons.
  - Do you think the journalist who wrote the story was misleading readers deliberately or merely through ignorance?

Group 2

- Parris, M. (2007) The truth about those little red lights: a tale of power and poppycock. Significance 4 (3), 128–129.
- This reading discusses exactly how much power is wasted by electrical devices on standby. An often quoted figure is 10%. The article discusses the very limited survey data behind the figure.
  - What data have been used to justify the figure of 10% energy wasted by appliances on standby?
  - What are the limitations of this data?

Group 3

- Spagat, M. (2010). Truth and death in Iraq under sanctions. *Significance*, 7 (3) pp. 116–120.
- This article discusses the claim that economic sanctions applied to Saddam Hussein's Iraq "caused 500,000 children to die". The claim was based on survey data.
  - Describe the surveys which have been done which might inform us about child mortality in Iraq and the conflicts between them.
  - What kinds of problems are encountered in performing surveys of this kind?

Group 4

- Discuss the article “How to read a poll" on the Fallacy Files website at `http://www.fallacyfiles.org/readpoll.html`
  - Use the information in the article to explain the correct answers to the questions in the “Fallacy Files Online Poll" at the top of the article.
  - Discuss the story behind the photograph at the top of the article.

Group 5

- Moon, N. (2010). Curtains at Number Ten: predicting the general election Significance 7 (1), 24–26.
- This article is a discussion by a pollster in the UK about the likely outcome of the UK general election (the article was written before the election took place). As you probably know the result was a hung parliament with no majority for either major party, with the conservatives finally forming a coalition with the social democrats. Questions to address:
  - Why does Moon think that reporting of polls in the newspapers is often misleading?
  - Did Moon correctly predict the hung parliament as a possibility?
  - Why is prediction of the result from the national poll difficult?