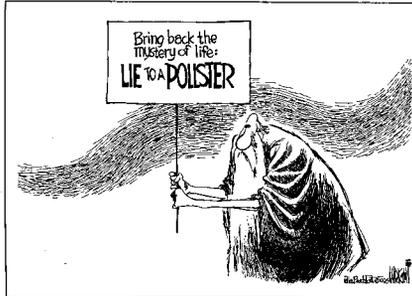


## Sample Surveys: Introduction



1



## Where are we going?

- Terminology and examples of uses of sampling
- Advantages of sampling
- Methods of sampling
- Sources of bias
- Random sampling: advantages and designs

2

## Terminology

- **Population:** a well-defined group of people or objects
- **Sample:** a part of a population
- **Parameter:** a numerical value associated with the population
- **Statistic:** a numerical value computed from a sample.

3

## Gallup Poll

Bush Approval Ratings

4

## Market Research



6

## Acceptance Sampling

Careful inspection of a sample of a large lot received from a supplier. On the basis of the sample, a decision is made whether to select or reject the whole lot.

6

## Accounting



Expensive to verify large numbers of invoices and other commercial transactions. Accountants thus routinely use sampling.

The **IRS** randomly samples a fraction of tax returns.

7

## Why Sample?

- Less expensive and time consuming
- May be more accurate than complete enumeration.

8

## Methods of Creating Samples

- “Convenience” sampling
- Quota sampling
- Random sampling. A *Simple Random Sample* is a draw from the population without replacement. Most real surveys use more elaborate kinds of random sampling.

9

## Sources of Bias

**Non-Response** People who do not respond may be different than those who do respond in ways that are related to the questions asked by the survey. For example, young people are less likely to be at home than married people.

10

### *Non-response bias: example*

Ann Landers asked her readers: “If you had it to do over again, would you have children?” Responses mailed back:

70% NO!

Newsday randomly polled 1373 parents

91% YES

11

### *Non-response bias: The Hite Reports*

Books on female sexuality based on essays of many women. The respondents were self selected, but in many ways they matched overall population demographics.

	study	US population
urban	60%	62%
rural	27%	26%
small town	13%	12%

Read more about it on the class website (under Additional Material)

12

**Timing as a source of bias:** A poll conducted by the NFL showed that football was America's favorite spectator sport. It was conducted one week before the Super Bowl.

**Wording of questions as a source of bias:**

"Do you favor adding to the Constitution a one-term limit for the President?" 50% No.

"Do you favor changing the Constitution to include a one-term limit for the President?" 65% No

Example of wording bias:

**WHY CLINTON'S RATINGS 'PLUNGED'**

A small change in the wording of a poll question by the Gallup Organization led to a significant contraction of the drop in President Clinton's personal favorability rating after his speech Monday.

Instead of the 20 point drop reported by Gallup, and picked up by several news organizations, including The Chronicle, the actual drop now appears to be considerably less.

In Monday night's survey, respondents were asked: "Now thinking about Bill Clinton as a person, do you have a favorable or unfavorable opinion of him?" That elicited responses that were 40 percent favorable and 48 percent unfavorable. In previous Gallup polls, the wording was: "Now I'd like to get your opinion about some people in the news. All read the name, please say if you have a favorable or unfavorable opinion of that person."

Last week, that traditional wording resulted in responses of Clinton that were 50 percent favorable and 38 percent unfavorable. Thus, it appeared on Monday that there had been a 20 percentage point drop.

In a poll on Tuesday, however, Gallup used the traditional wording and measured 55 percent favorable views of Clinton and 42 percent unfavorable.

"Now thinking of Bill Clinton as a person do you have a favorable or unfavorable opinion of him?"  
40% favorable

"Now I'd like to get your opinion about some people in the news. As I read the name, please say if you have a favorable or unfavorable opinion of this person."  
55% favorable

**The World Values Survey**

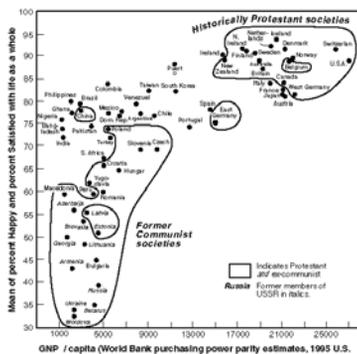
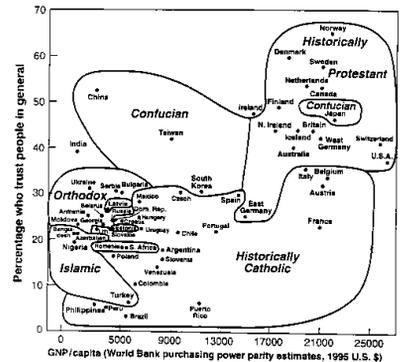


Figure 3. Subjective well-being by level of economic development and historical heritage of citizen societies.



## The Questionnaire

17

### **Attributes of interviewer as a source of bias:**

After major racial riots in 1968 in Detroit a sample of black residents were asked:

“Do you personally feel that you can trust most white people, some white people, or none at all?”

White interviewer: 35% “most”

Black interviewer” 7% “most”

18

## Some designs using random sampling

A **Simple Random Sample (SRS)** is drawing without replacement from the entire population.

**Multistage Sampling:** Take a SRS of counties, within each take a SRS of townships, and within each of those a SRS of blocks. Sample the entire block.

**Stratified Sampling:** Divide population into groups (“strata”) and take SRS from each group.

19

## Advantages of Random Sampling

- Eliminates bias
- Gives a basis for calculating probable errors. Since SRS is drawing without replacement we can use what we know about the chance variation in drawing tickets from box to say something about chance variation in SRS

20

? Which of the following are simple random samples:

- Write out the names of everyone in this room on cards; make 10 draws at random without replacement.
- As above, but draw out the 10 cards all at once.
- Take first 10 people who pass through Sproul Plaza.
- Take 5 classes at random and then take all the students in those classes.

21

? Education Testing Service got a sample of students in the following way: First they divided the population of all colleges and universities into roughly homogeneous groups. Within each group they used their judgement to choose a representative school. That school was then asked to choose a sample of students.

Is this a good way to get a representative sample of students?

22

? An accounting firm audits the transactions of a company in the following way: The transactions are listed in chronological order, a number between 1 and 100 is drawn at random. The first transaction is that one on the list. The second is 100 transactions later, etc.

Is this a probability sample?

Is it a simple random sample?

Is there selection bias in this method?

23