The Values of Polls

1. Polls and politicians’s behavior
   - Help politicians monitor their electoral standing against current or future opponents
   - Important aspect of presidential power

2. Polls and election results
   - Create expectations about who is the likely election winner
   - Influence campaign activities

3. Polls and the public
   - Poor understanding of how opinion polls work
   - In 2001, 66% said “Polls are useful for elected officials in Washington to understand how the public feels about important issues”
Public opinion polls are based on samples because it is impossible to ask the nearly 300 million people living in the United States about their political opinion.

When 500 out of 1000 survey respondents approve the president’s job performance, we can confidently say that 105 million out of 210 million American citizens approve his job performance.

What are samples and how are they collected?

How much confidence can we place in statements about a population given observations derived from a very few of its members?
“Do you favor keeping a large number of US troops in Iraq until there is a stable government there OR bringing most of the troops home?”

1. Favor keeping a larger number of US troops there until stable government
2. Bring most home in next year
3. Decline to answer / Don’t know
Respondents’ reactions to the survey question:

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Favor keeping them in Iraq</td>
<td>446</td>
<td>44.9%</td>
<td>47.9%</td>
</tr>
<tr>
<td>Bring them home</td>
<td>484</td>
<td>48.8%</td>
<td>52.1%</td>
</tr>
<tr>
<td>Subtotal</td>
<td>930</td>
<td>48.8%</td>
<td>52.1%</td>
</tr>
<tr>
<td>DK</td>
<td>63</td>
<td>6.3%</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>993</td>
<td>100%</td>
<td></td>
</tr>
</tbody>
</table>
“What about sexual relations between two adults of the same sex. Do you think it is always wrong, almost always wrong, wrong only sometimes, or not wrong at all?”

1. Always wrong
2. Almost always wrong
3. Sometimes wrong
4. Not wrong at all
5. Decline to answer / Don’t know
Respondents’ reactions to the survey question:

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Always wrong</td>
<td>1483</td>
<td>28.6%</td>
<td>57.5%</td>
</tr>
<tr>
<td>Almost always wrong</td>
<td>119</td>
<td>2.1%</td>
<td>4.6%</td>
</tr>
<tr>
<td>Sometimes wrong</td>
<td>198</td>
<td>3.5%</td>
<td>7.7%</td>
</tr>
<tr>
<td>Not wrong at all</td>
<td>781</td>
<td>14.0%</td>
<td>30.3%</td>
</tr>
<tr>
<td>DK</td>
<td>3001</td>
<td>46.2%</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>5582</td>
<td>100%</td>
<td></td>
</tr>
</tbody>
</table>

How can we use the information on these tables?
Sampling

- **Population**: A unit about which we want information → How one defines the population has an impact on poll results.
  1. Those eighteen and older or
  2. registered voters or
  3. those who will (or do) vote in the next election

- **Sample**: a part of the population selected for analysis

- If the sample is selected properly, the information it yields may be used to **make inferences** about the whole population

- **Population parameter** – Certain characteristics of population
The purpose of sampling is to collect data or samples that provide an accurate estimate of a population parameter.

Sample statistic – the estimated value of a population characteristic that is calculated from sample data.

Representativeness – samples are representative when they accurately mirror the population.

Randomness – the only method by which a representative sample can be scientifically drawn.
Example: Partisanship

- Distribution of partisanship in the U.S.
- We want to know unknown proportions $D$ (Democrats), $R$ (Republicans), and $I$ (Independents) in the population. (Note that $D + R + I = 1$.)
- A statistic based on a properly collected sample provides an estimator, $\hat{D}$, $\hat{R}$, $\hat{I}$.
- Our goal is to make a good unbiased guess about the true value of $D$, $R$, $I$. 
Types of Samples

- **Probability Sample**: A sample for which each element in the total population has a known probability of being included in the sample = high representativeness
  - Simple random samples
  - Systematic samples
  - Stratified samples
  - Cluster samples

- **Nonprobability Sample**: A sample for which each element in the total population has an unknown probability of being selected → This sampling framework rules out the use of sampling theory to make inferences
Simple random sample

- Each element has an *equal* chance of being selected
- A list of all the elements in the population must be available
- (1) Assign a number to each element in the sample frame, and then use a random numbers table (list of random numbers)
- (2) By-the-lot method
Cluster samples

- The sampling frame initially consists of clusters of elements.
- Cluster sampling is used to address the problem of having no list of the elements in the target population.
- Example of national opinion poll:
  1. List and sample States
  2. List and sample Counties within each state
  3. List and sample municipalities within counties
  4. List and sample census tracts within municipalities
  5. List and sample blocks within census tracts
  6. List and sample households within each block
Suppose we want to know the proportion of people who approve job performance of President Obama.

The population parameter is $p$.

Suppose that a recent survey asked a nationwide random sample of 1000 adults if they approve Presidents’ job performance.

Of the respondents, 600 or 60% said they approve President’s job performance.

Using the sample statistic $\hat{p} = 0.6$, we infer that 60% of the American electorate approves President’s performance.

How can we justify the use of sample statistic to make an inference of population parameter?
A second random sample taken at the same time would choose different people and might produce a different value of $\hat{p}$.

A third random sample would produce a different value of $\hat{p}$.

Every time you take another random sample, you find a different sample statistic!

The sample statistic depends on the composition of people in the sample.

**Sampling variability**: the value of a statistic varies in (hypothetical) repeated random sampling.
How do you trust $\hat{p}$ from a single random sample?

We cannot be totally confident in the sample statistic from a single random sample.

Samples provide only estimates or approximations of population attributes.

When we report a sample statistic, we always assume there will be a difference between the reported and actual values.
Example

- You are asked to conduct a poll to estimate the proportion of the electorate who registered as independents in your district.
- Suppose $P = 0.25$. (Suppose the Bureau of Elections reports 25% of the registered voters in the district is independent.) But, your client doesn’t believe that.
- So she asks you interview 10 individuals using a simple random sample method.
- You found $\hat{P} = 0.20$ from a single random sample.
- The difference between the true value and the sample statistics is called **sampling error**.
- Imagine that you repeat the sampling procedures 1000 times.
Example

Distribution of 1000 Sample Proportions (Sample Size=10)

Sample Proportions

Frequency

0.0 0.2 0.4 0.6 0.8 1.0

0 50 100 150 200 250
Example

- We see a sampling variability.
- Most sample statistics are around the population parameter ($=0.25$).
- The sample statistics from repeated samples have a distribution.
Suppose that the same sampling procedure were to be repeated an infinite number of times, and a sample statistic calculated each time.

The sampling distribution of a statistic is the distribution of the values taken by the statistic in all possible samples of the same size from the same population.
If the samples are collected independently and randomly, the average (mean, or the **expected value**) of many sample statistics will equal the true, population parameter, no matter what the sample size is: \( E(\hat{P}) = P \)

“**The expected value** of the sample statistics of \( \theta \) equals the corresponding value for the population.”

So a sample statistic is **unbiased**.

Knowing that your statistic is unbiased does not help with the immediate problem because you have, in fact, only one sample and one estimated proportion!
Standard Errors

- You need to know how far from the true value your estimate can be.
- **Standard error**: A number that measures the variability or dispersion of the sample statistics within the sampling distribution.
- If the true proportion in a population is $P$, and one takes repeated samples of size $N$, the resulting sample proportions ($\hat{P}$) will have a normal distribution, with a mean of $P$ and a standard error of $\sigma_{\hat{P}}$.
- The standard error is an indicator of how much uncertainty there is in a statistic.
Simulation Example

Sample Size=10

Sample Proportions

Sample Size=100

Sample Proportions

Sample Size=5000

Sampling Proportions

Sample Size=50

Sample Proportions

Sample Size=1000

Sampling Proportions

Sample Size=10000

Sampling Proportions
• As sample size gets larger, the standard error of the statistic – a measure of the variability of our statistic – gets smaller.

• The larger the sample size, the greater the certainty of the results.

• **Confidence level**: A statement of our belief that an estimated range of values includes the population parameter.

• 95% confidence interval \(= \hat{\theta} \pm 1.96\hat{\sigma}_\theta\)

• If we drew 100 independent random samples from a population having a parameter \(\theta\), we believe that about 95 out of the 100 estimated intervals would include this value \(\theta\).
Example

- “Obama leads McCain, 55% to 45%”
- “Obama leads McCain, 55% to 45%, with a margin of error of ±3%”
- Margin of error should be small
- The larger the sample, the smaller the sampling error
## Sample Size and Sampling Error

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Confidence Interval (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4000</td>
<td>±2</td>
</tr>
<tr>
<td>1500</td>
<td>±3</td>
</tr>
<tr>
<td>1000</td>
<td>±4</td>
</tr>
<tr>
<td>600</td>
<td>±5</td>
</tr>
<tr>
<td>400</td>
<td>±6</td>
</tr>
<tr>
<td>200</td>
<td>±8</td>
</tr>
<tr>
<td>100</td>
<td>±11</td>
</tr>
</tbody>
</table>
Bad Sampling

- Four Defects in *Literary Digest Poll* of 1936 on Alf Landon VS FDR
  - A large sample (2.2 million respondents) is no guarantee for accuracy
  - No probability sample (a sample drawn from telephone directories and automobile registration lists)
  - Timing (early September)
  - The emergence of class cleavage and New Deal coalition
  - Low response rate (2.2 million out of 10 million) and self-selection

- Scientific polling with a probability sample surged in popularity
How Accurate are the Contemporary Polls?

[Graph showing the relationship between Democratic vote percent in polls and on election day over the years 1960 to 2004.]
## Accuracy of Final Polls in the 2004 Presidential Election

<table>
<thead>
<tr>
<th>Poll Source</th>
<th>Bush</th>
<th>Kerry</th>
<th>Bush Margin</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>News Week</em></td>
<td>50</td>
<td>44</td>
<td>+6</td>
</tr>
<tr>
<td>GWU Battleground</td>
<td>50</td>
<td>46</td>
<td>+4</td>
</tr>
<tr>
<td>TIPP</td>
<td>48.6</td>
<td>45.3</td>
<td>−3.3</td>
</tr>
<tr>
<td>Pew Research Center</td>
<td>48</td>
<td>45</td>
<td>+3</td>
</tr>
<tr>
<td><strong>Election Outcome (Popular Vote)</strong></td>
<td><strong>48.9</strong></td>
<td><strong>46.5</strong></td>
<td><strong>+2.4</strong></td>
</tr>
<tr>
<td>CBS News</td>
<td>49</td>
<td>47</td>
<td>+2</td>
</tr>
<tr>
<td>CNN/<em>USA Today</em>/Gallup</td>
<td>49</td>
<td>47</td>
<td>+2</td>
</tr>
<tr>
<td>ABC News</td>
<td>49</td>
<td>48</td>
<td>+1</td>
</tr>
<tr>
<td>NBC/<em>Wall St. Journal</em></td>
<td>48</td>
<td>47</td>
<td>+1</td>
</tr>
<tr>
<td>Zogby</td>
<td>48</td>
<td>47</td>
<td>+1</td>
</tr>
<tr>
<td><em>Washington Post</em></td>
<td>49</td>
<td>48</td>
<td>+1</td>
</tr>
<tr>
<td>American Research Group</td>
<td>48</td>
<td>48</td>
<td>0</td>
</tr>
<tr>
<td>Marist College</td>
<td>49</td>
<td>50</td>
<td>−1</td>
</tr>
<tr>
<td>Democracy Corps</td>
<td>47</td>
<td>48</td>
<td>−1</td>
</tr>
<tr>
<td>Fox News</td>
<td>46</td>
<td>48</td>
<td>−2</td>
</tr>
</tbody>
</table>
The quality of data often depends on the types of surveys.

- Face-to-face
- Telephone
- Mail (and Email)
- Internet
Characteristics of Surveys

1. **Response Rate**
   - The proportion of persons selected for participation in a survey who actually participate
   - Face-to-face > telephone > mail and internet

2. **Representativeness**
   - Sample-population congruence
   - Determined by initial selection of respondents or incomplete participation of those selected
   - Face-to-face > telephone > mail and internet

3. **Response quality**
   - Respondents provide accurate and complete information
   - Face-to-face > telephone > mail and internet

4. **Overall Cost**
   - Money and time
   - Face-to-face > telephone > mail and internet
People are sensitive to the way survey questions are phrased

Measure accurately people’s beliefs, attitudes, opinions, and behavior

Example: the responses of white Americans to different phrasings of a question about the controversial policy of minority set-asides

Form A: “Sometimes you hear it said there should be a law to ensure that a certain number of federal contracts go to minority contractors. Do you favor or oppose such a law?” Favor 43% and Oppose 57%

Form B: The Congress of the United States – both the House of Representatives and the Senate – has passed a law to ensure that a certain number of federal contracts go to minority contractors. Do you favor or oppose such a law? Favor 57% and Oppose 43%
Bad Types of Questions

- **Double-barreled question**
  - “Do you agree with the statement that the situation in Iraq is deteriorating and that the US should increase the number of troops in Iraq?”

- **Ambiguous question**
  - “What is your income?”
  - “Rate the two presidential candidates’ performance in a TV debate as good, bad, or indifferent”
  - “How would you rate President Bush’s performance so far?” (A) Great (B) Somewhere between great and terrible (C) Terrible

- **Leading question**
  - “Don’t you think that global warming is a serious environmental problem?”
  - “Recently the Soviet armed forces openly invaded the independent county of Afghanistan. Do you think the U.S. should supply military equipment to the rebel freedom fighters?”
Close-ended question

“Do you agree that Mr. Kerry cares for people like me?” (A) strongly agree (B) agree (C) neither agree nor disagree (D) disagree (E) strongly disagree

Open-ended question

“Is there anything in particular about Mr. Kerry that might make you want to vote for him?”

Single-sided question

“Do you agree or disagree with the idea that the government should see to it that every person has a job and a good standard of living?” (A) agree (B) disagree (C) don’t know

Two-sided question

“Do you think that the government should see to it that every person has a job and a good standard of living, or should it let each person get ahead on his or her own?” (A) provide job (B) let people get ahead on their own
The previous content of the interview might have an impact on a response to a specific question

- “The United States should let Communist newspaper reporters from other countries come here and send back to their papers the news as they see it.” 55% said yes
- “A Communist country like Russia should allow American newspaper reporters to come in and send back the news as they see it.” If this question is asked before the above one, 75% said yes to the above question
- A consistency effect: an answer to a question may be constrained by an answer given earlier.
Today, most surveys are conducted by telephone rather than in person.

Random-digit dialing

Nonresponse and representativeness

Nonresponse led to an overrepresentation of the elderly, women, the poor, and the less educated.

The pollsters’ major telephone-related problem is the growing reliance on cell phones.
The Misuse of Survey

- Mails sent to politicians
- Biased, non-representative samples
- Biased questions
  - “Republican leaders are advocating a welfare deform plan that would end welfare benefits to single parents who cannot find work and spend the money on orphanages to house their children. Do you support or oppose this approach?”