

## **“Pretend” Clicker Quiz**

**What do you do if a company calls you and asks you to take an opinion poll that will last several minutes?**

- A. Immediately hang up**
- B. Answer the questions if you have time, otherwise decline to answer**
- C. Answer the questions if the topic interests you, otherwise decline to answer**
- D. Always answer the questions**

## Sample Surveys

- **Pew Research Center did a survey of randomly selected households.**
- **Needed a sample size of 1000 responses.**
- **Had to call 2879 households to get their 1000 responses.**
- **About 1600 of these either didn't answer the phone, or refused to do the survey questions.**
- **Nonresponse rate was  $\frac{1879}{2879} = 0.653$  (65.3%)**
- **What is the effect of this high nonresponse rate on how trustworthy the survey results are?**

## Ways Sample Surveys Can Go Wrong

- ***Sampling errors:*** Often caused by poorly designed surveys
- ***Random sampling error:*** Nothing can be done to eliminate this
- ***Nonsampling errors:*** Mistakes unrelated to sampling issues
- ***Important:*** The announced margin of error of a survey result only accounts for *random sampling error*.

## Sampling Errors

- The *sampling frame* is a list of all the individuals from which we will draw our sample.
- This is often only hypothetical – we cannot actually list all individuals in the population of interest.
- What if we used a phone directory as the basis for our sampling frame?
- What if we used a list of household addresses? (from tax records?)
- When we leave certain groups in the population out of our sampling process, the result is *undercoverage*.
- Hard to eliminate these sampling errors, but their effect is often(?) minimal.

## Nonsampling Errors

- A simple nonsampling error is a *processing error* – a typo in the data entry or calculation.
- Automated data entry systems reduce the occurrence of processing errors.
- *Response error*: When a respondent gives an incorrect or false answer to a survey question.
- Often occurs when questions are personally embarrassing or when correct answer is difficult to remember.

## Nonsampling Errors (Continued)

- **Most difficult to handle – *Nonresponse error*: When an individual chosen for a sample cannot be contacted or refuses to fully participate.**
- **Nonrespondents can form 75% or more of a poll's original sample!**

## Clicker Quiz 1

**Joe Smith is selected at random to be called for a sample survey.**

**When he is called, he listens to one question and hangs up. What type of error is this?**

- A. Random sampling error**
- B. Response error**
- C. Processing error**
- D. Nonresponse error**

## Clicker Quiz 2

***The Bradley Effect:*** In the 1982 California governor's race, exit polls predicted black candidate Tom Bradley would win. Instead, his white opponent won. Pollsters suspected respondents lied about voting against Bradley in order to not appear prejudiced to the interviewer. What type of error is this?

- A. Random sampling error
- B. Response error
- C. Processing error
- D. Nonresponse error



## Wording Issues

- **The way a survey question is worded can greatly influence the responses.**
- ***Example 1:* “Should laws be passed to eliminate all possibilities of special interests giving huge sums of money to candidates?”**
- ***Example 2:* “Should laws be passed to prohibit interest groups from contributing to campaigns, or do groups have a right to contribute to the candidate they support?”**
- **Which question is more likely to receive a response of “Yes”?**

## Wording Issues

- **The way a survey question is worded can greatly influence the responses.**
- ***Example 1:* “Should laws be passed to eliminate all possibilities of special interests giving huge sums of money to candidates?”**
- ***Example 2:* “Should laws be passed to prohibit interest groups from contributing to campaigns, or do groups have a right to contribute to the candidate they support?”**
- **For two national random samples, 80% said “Yes” to question 1, and 40% said “Yes” to question 2.**
- **Be careful of surveys whose questions are slanted to favor one response over others – this is a serious *nonsampling error*.**

## Clicker Quiz 3

Which question is more likely to result in an answer of “yes”?

- A. “Does the government spend *too much* on assistance to the poor?”
- B. “Does the government spend *too much* on welfare?”
- C. Both are equally likely

## Dealing with Nonsampling Errors

- **Nonresponse is a fact of life for almost all surveys.**
- **One solution: Substitute (similar) participants for nonrespondents.**
- **If the survey is finished, we might *weight the responses* differently to correct for bias.**
- **Example: What if there was a high nonresponse rate among Hispanics?**

## Dealing with Nonsampling Errors

- **Example: What if there was a high nonresponse rate among Hispanics?**
- **Could weight the answers of Hispanics who *did* respond more heavily.**
- **Weighting involves complicated formulas usually implemented by expert statisticians.**
- **Weighting can reduce bias, but will increase variability, so the announced margin of error must be adjusted properly.**

## Advanced Sample Survey Designs

- **With a *cluster sample*, we randomly select *clusters* of individuals, then survey everyone in each selected cluster.**
- ***Example*: Selecting neighborhoods from a city at random, then interviewing every household in each selected neighborhood.**
- **Advantage: May be easier to carry out logistically.**
- **With a *Stratified Random Sample*, first we divide the population into groups (called *strata*) so that members of each stratum are similar in some way.**
- **Then take a simple random sample *within each stratum*, which are combined to create the overall sample.**

## More about Stratified Samples

- **With a Stratified Random Sample, we can make separate conclusions about each stratum.**
- **With a Stratified Random Sample, the margin of error may be different for each stratum.**
- **The proportions of sampled individuals from each stratum may match the proportions from the population . . . but they don't have to.**
- **Sometimes we may want to deliberately overrepresent a certain group in our sample.**

## Clicker Quiz 4

**What type of sampling scheme is the following? We take a simple random sample of 80 USC freshmen, a simple random sample of 90 USC sophomores, a simple random sample of 70 USC juniors, and a simple random sample of 60 USC seniors – resulting in 300 students total.**

- A. Cluster Sample**
- B. Simple Random Sample**
- C. Stratified Random Sample**
- D. Convenience sample**



## Clicker Quiz 5

**What type of sampling scheme is the following? Out of all residence hall floors, we randomly select 10 floors across campus. Then we interview every person on those 10 floors.**

- A. Cluster Sample**
- B. Simple Random Sample**
- C. Stratified Random Sample**
- D. Convenience sample**

- **A *probability sample* uses random chance to select the individuals for the sample.**
- **All good samples are *probability samples*.**
- ***Nonprobability samples* allow human beings to make the decisions about which individuals are selected for the sample.**
- ***Probability samples* are the only kind of samples for which we can state the *margin of error* or *confidence level* of our results.**
- **With *nonprobability samples*, we have no idea how *trustworthy* our results are.**

## **Clicker Quiz 6**

**Which type of sample is a probability sample?**

- A. Convenience sample**
- B. Census**
- C. Voluntary response sample**
- D. Stratified Random Sample**

## **Questions to Ask: Should You Trust a Poll Result?**

- **Who carried out the survey?**
- **What was the population?**
- **How was the sample selected?**
- **How large was the sample?**
- **What was the response rate?**
- **How and when were the subjects contacted?**
- **What was the exact wording of the questions?**