

Chapter 3 – Surveys and Sampling

1. Roper.

- a. No, they were not using a random sample. As stated, the survey was designed to get 500 males and 500 females.
- b. They used a stratified random sample, with the strata being the gender of the respondents.

3. Software licenses.

- a. This sample was a voluntary response, not a random sample.
- b. There is no confidence in the estimate sampled. Voluntary response samples are almost always biased, and so conclusions drawn from them are almost always wrong.

5. Gallup.

- a. The population of interest is all adults in the United States aged 18 and older.
- b. The sampling frame, a list of individuals from which the sample will be drawn, consists of adults with landline telephones, which are the only numbers available for a study like this.
- c. An increasing number within the population (e.g., many college students and others with mobile only service) don't have landline phones, which could create a bias.

7. HR directors.

- a. Population–Human resources directors of Fortune 500 companies.
- b. Parameter–Proportion of those surveyed who don't feel that surveys intruded on their work day.
- c. Sampling Frame–List of all Human Resource directors at Fortune 500 companies.
- d. Sample–Those who responded (23% of all HR directors).
- e. Sampling method–Questionnaire mailed to all HR directors (not random sample).
- f. Bias–Voluntary response sample. Those who responded to the questionnaire that was mailed could be predisposed to a particular answer. Because they are responding to the survey, they may be more inclined to believe that surveys like this do not intrude on their workday.

9. Alternative medicine.

- a. Population–All Consumer Union subscribers.
- b. Parameter–Proportion of Consumer Union subscribers who have used and benefited from alternative medicine.
- c. Sampling Frame–All Consumer Union subscribers.
- d. Sample–Subscribers who responded.
- e. Sampling method–Questionnaire to all subscribers.
- f. Bias–Nonresponse. Those who respond could have strong feelings about the topic and affect the results.

11. At the bar.

- a. Population–Adult bar patrons.
- b. Parameter–Proportion of sample who thought drinking and driving was a serious problem.
- c. Sampling Frame–All chosen bar patrons.
- d. Sample–Every 10th person leaving the bar.
- e. Sampling method–Systematic sampling (every 10th person).
- f. Bias–Probably biased toward thinking drinking and driving is not a serious problem. The sample consisted of bar patrons leaving the bar. A large percentage of them had something to drink, most likely leading to a biased viewpoint. In addition, bar patrons don't reflect what all adults think about drinking and driving.

13. Toxic waste.

- a. Population–Soil located near former waste dumps.
- b. Parameter–Concentrations of toxic chemicals.
- c. Sampling Frame–Any accessible soil surrounding a former waste dump.
- d. Sample–Soil samples taken from 16 locations near a former waste dump.
- e. Sampling method–Not specified how the sample locations were chosen.
- f. Bias–Not specified how soil sample locations were chosen and therefore cannot assume they were chosen randomly, perhaps accessibility or some other factors. Unless this is known, it is possible that bias can affect the results if soil taken is more or less polluted than a random selection would produce.

15. Quality control.

- a. Population–Snack food packages.
- b. Parameter–Proportion of snack food packages passing inspection, weight of bags.
- c. Sampling Frame–All snack food packages produced in a day.
- d. Sample–Packages in 10 randomly selected cases, 1 bag from each case for inspection.
- e. Sampling method–Multistage sampling due to a combination of methods. The selection of the 10 cases is a cluster and the sampling selection of an individual bag from each case is probably a random sample, although this is not specified.
- f. Bias–Should be unbiased as long as the individual bag chosen is random. There could be differences in the first bag of a case versus the last bag.

17. Pulse poll. The station's faulty prediction is most likely the result of bias. Only people watching the local TV station news have the opportunity to respond. The responders who volunteered to participate may have different viewpoints than those of other voters, who either chose not to respond or didn't have the opportunity to participate (didn't see the news program).

19. Cable company market research.

- a. Sampling strategy is volunteer response. Bias is introduced because only those individuals who see the ad and feel strongly about the issue will respond. The opinions may not be representative of the rest of the public.
- b. Sampling strategy is a cluster of one town selected to be sampled. Bias is introduced because one town may not be representative of all towns.
- c. Sampling strategy is an attempted census, accessing all customers. Bias is introduced because of nonresponse to the mailing survey.
- d. Sampling strategy is stratified by town, selecting 20 customers at random from each town to be surveyed, including follow up. This strategy should be unbiased and representative of the public opinion about the cable issue.

21. Churches.

- a. This is a multistage design because of the combination of methods. A cluster sample consists of the 3 churches chosen at random. One hundred church members are randomly selected from each church to be surveyed.
- b. Since only 3 churches are chosen at random, if any one of the churches chosen is not representative of the entire 17 churches, bias is introduced in the form of sampling error.

23. Amusement park riders.

- a. This is a systematic sample (every 10th person in line).
- b. It is likely to be representative of all of those waiting in line to go on the roller coaster. It would be useful to compare those who have waited and are now at the front with those who are in the back of the line. Otherwise, survey every 10th person about to board the roller coaster for a more consistent response.
- c. The sampling frame consists of persons willing to wait in line for the roller coaster on a particular day within a given time frame.

25. Survey wording.

- a. Answers will vary. Question 1 is a straightforward question about the issue and certainly appropriate for the survey. Question 2 is biased in its wording and could in some way offend those surveyed because it connects the cost to a daily cost of a cappuccino. Many people don't have coffee drinks and so if they don't spend that money, why would they want to spend it on cable? Those people who drink coffee most likely would not give up a coffee drink in order to have cable. The statement would be more neutral by just stating how much the cable would cost per day.
- b. Question 1 is neutrally worded because it is a simple, straightforward statement asking for the required response.

27. Another ride.

Biases exist because it could be that only those who think it is worth waiting for the roller coaster ride are likely to still be in line. Those who don't like roller coasters or don't want to stay in lines are not part of the sampling frame. Therefore, the poll won't get a fair picture of whether park patrons overall would favour more roller coasters.

29. (Possibly) Biased questions.

- a. This statement is biased because it leads the responder toward yes because of the word "pollute". The word "pollute" conjures up a negative image leading the responder to agree that companies should pay for

this behaviour. Another way to phrase it would be “Should companies be responsible for costs of environmental cleanup?”

b. This statement is biased because it leads the responder to no because of the words “enforce” and “strict” that conjure up images that could lead a responder to having negative reaction. Another way to phrase it would be “Should companies have dress codes?”

31. Phone surveys.

a. It would be difficult to achieve a random sample in this case because not everyone in the sampling frame has an equal chance of being chosen. People with unlisted phone numbers, without phones, and those at work or away from the home at the designated calling time cannot be contacted.

b. Another strategy would be to generate random numbers and call at random times or select random numbers from the phonebook and call at random times (this doesn’t solve the unlisted phone number issue).

c. In the original plan, families that have one person at home are more likely to be included in the study. Using the second plan, more people are potentially included although people without phones or those not home when called are still not included.

d. This change does improve the chance of selected households being included in the study.

e. The random digit dialing does address all existing phone numbers, including unlisted numbers. However, there is still the issue of residents not being home at the time of the call. In addition, people without phones are still left out of the study.

33. Change.

a. Answers will vary

b. The parameter being estimated is the true mean amount of change that you carry daily just before lunch.

c. Population is now the amount of change carried by your friends. The average parameter estimates the mean of these amounts.

d. The 10 measurements in c) are more likely to be representative of your class (peer group with similar needs) but unlikely for larger groups outside of your circle of friends.

35. Accounting.

a. Assign numbers 001 to 120 (3 digits required because the maximum number is 120) representing each order in a day. Use random numbers to select 10 transactions to check for accuracy.

b. Separate the transactions and sample each type (wholesale and retail) proportionately. This would be a stratified random sample.

37. Quality control.

a. Randomly select 3 cases and then randomly select one jar from each case.

b. Assign numbers 01 to 20 to cases 07N61 to 07N80 respectively. Then generate three random numbers between 01 and 20 and select the appropriate case. Then assign random numbers 01 to 12 to each of the 12 jars within each case. For each case selected, generate a random number between 01 and 12 and select the corresponding jar within each case.

c. The method described involves two separate sampling methods and, therefore, it is multistage sampling.

39. Sampling methods.

a. Yellow pages may not include all doctor listings. If regular line listings are used, the list may include all doctors. If ads are used, not all doctors would be included and the ones with ad would not be typical of all doctors.

b. This sampling method is not appropriate. The cluster sample chosen (the randomly selected page) will only contain a handful of businesses and maybe only one or two business types.