AP Statistics

2019-01-08 Noah Overcash

Assignment 4.1

Pg. 226-229: 1,5,9,17,19,25,27,29,31,35

Question 1

Population: The local businesses, as found in the telephone book

Sample: The 73 businesses which returned the questionnaire

Question 5

This poll is almost certainly biased because it requires volunteering. Volunteering requires effort, which would only be invoked by those with very strong views. Additionally, these volunteers must forgo 50¢ to make the call.

Question 9

Part A

A convenience sample, as it was not a representative sample of the entire population, only those which were easiest to access.

Part B

This sampling method is biased as only the early-risers would make it to school earlier than the others, making it likely that they got less sleep the previous night than those who slept later, causing them to arrive at school later, excluding them from the study.

Question 17

Part A

It would be difficult to conduct a simple random sample as, although random numbers could be easily produced, it would be difficult to extract an individual device from the stock created that day (i.e., it would be difficult to count and find the 473rd item that had already been packaged into stacks, boxes, palettes, etc.

Part B

This would not be representative as it only covers a convenient sample. The displays could reflect employees eager to be done for the day, machines being prepared to be shut down, or any other such variable/ inconsistency which would only be present at the end of the day.

Part C

This is not a simple random sample as it does not include any entropy in how devices are selected.

Question 19

I will label the two strata using numbers, so that the students get alphabetically allocated numbers 00-29 and the faculty get numbers 30-39.

Then, I will clump line 123 into two-digit clumps and iterate over them.

Line 123: 54 58 08 15 07 27 10 25 60 27 55 89 23 30 63 41 84 28 18 68

Now, I will iterate over these numbers in the following manner:

If it is above 39, it will be excluded due to it being out of range

If it is within 00-29

If four students have not been chosen, continue;

If this number has not already been chosen, continue;

Take the number and select the associated student

If it is within 30-39

If two faculty have not been chosen, continue;

If this number has not already been chosen, continue;

Take the number and select the associated faculty member

If four students and two faculty now chosen, stop iterating

54, 58: out of range

08: First student (**Griswold**)

15: Second student (**Kim**)

07: Third student (**Ghosh**)

27: Fourth student (**Thompson**) — all four students have now been selected

10, 25, 60, 27, 55, 89, 23: ignored, four students chosen or number out of range

30: First faculty (**Andrews**)

63, 41, 84, 28, 18, 68: ignored, four students chosen or number out of range

Line 123 has been exhausted, pairing line 124 and continuing iteration.

Line 124: 71 03 50 90 01 43 36 74 94 97 72 71 99 67 58 27 61 19 15 96

71, 03, 50, 90, 01, 43: ignored, four students chosen or number out of range

36: Second faculty (**Moore**) - all two faculty have been selected.

The sample consists of:

Students: Griswold, Kim, Ghosh, Thompson

Faculty: Andrews, Moore

Question 25

Part A

This is a cluster sample, as the subdivision was split into blocks (clusters). Each of these clusters was inputted into a SRS, and the sample blocks (clusters) were surveyed in their entirety.

Part B

I loaded the number of each block (1-65) into a program. From there, I instructed the program to shuffle the list of numbers and select the first five within the shuffled array. This caused the program to effectively choose five unique numbers, which corresponded to five unique subdivision blocks.

Below is the demonstration of the program:

```
php > $blocks = range(1, 65);
php > shuffle($blocks);
php > echo json_encode(array_slice($blocks, 0, 5));
[36,11,7,45,31]
```

Question 27

This sampling method could cause for households without a telephone number (or that have requested not to be listed) to be excluded from the study, and households with multiple lines, depending on the aggregation method, to have increased chances of being selected. This is problematic as it favors households with more income (and thus able to afford telephone service(s)). It would also remove potential respondents who do not agree with telephones or the policies of telephone companies. Finally, those who have telephone service may be more biased based on personal experience. For example, a survey asking how fair a telephone provider's rates seemed might be responded to positively to those who do not have experience with providers, however, those that do have experience with providers may be biased against certain types of plans or be more aware of scams than the general population.

Question 29

Part A

This survey does not sample those in the richest seating, which would lead to a bias causing a lower result on overall spending (as those willing to buy very nice tickets would have more money.

Part B

Sampling error - the sample does not accurately represent the population (the "fans attending opening day for the Cleveland Indians baseball season").

Question 31

Part A

% = 100 % • (45956-5029)/45956 = 89%

Part B

Nonresponse can cause bias in this survey as those who are busier may be unable to accept telephone calls, particularly those at a landline. Those who are busier would typically spend more time commuting (and thus driving), causing the survey to be biased to lower than the actual result.

Question 35

Part A

This question is clear, however, by presenting an ambiguous "some" which may cause some to overestimate the number of cases, inciting fear. This would cause bias towards adding warning labels.

Part B

This question is clear, however, slants the response towards "yes" as it asks "do you agree" (where most people would say yes to "go with the flow") and provides a positive reason which the average person would not be against.

Part C

This question, and the context with which it should be evaluated in, are fair and clearly stated for the respondent. This allows for the sampled to respond without being directed towards a response.