

Sampling Scheme & Question Design

By: Emily Butler, Meggie Hayes, Douglas Heckmann, Christopher Peter Makris, & Thomas Todd

K. Sampling Scheme

Since we will be conducting a “man on the street” survey, it will be difficult to collect a completely random sample. Therefore, we plan to collect a pseudo-random sample by attempting to sample every third person that passes by. We will also be implementing stratification to try to gain a demographically accurate sample of our target population. This stratification will be done by soliciting survey responses in different locations where we expect certain demographic groups to be commonly located. By using this method of stratification, we hope that the demographics of our sample population will be representative of our target population.

what variable(s) will you stratify on?

this is fine. please do a little "dry run" somewhere to get an idea of what's going to be workable in practice (every 3rd person might be too frequent to

L. Proposed Questions

Before individuals take our survey, we will ask them whether or not they are an undergraduate student. If they are an undergraduate student, we will ask them to fill out our survey. We believe that by doing so our coverage error will be reduced since we completely eliminate receiving responses from non-undergraduate students who are all outside the scope of our target population. Therefore, we consider the question “Are you an undergraduate student?” an initial question that is not technically part of our survey.

0. (Verbally) Are you an undergraduate student?
 - If response is “yes,” offer the individual a survey to complete.
1. Do you consume alcoholic beverages?
 - If response is “no,” please stop this survey and return it to your surveyor.
2. What university do you attend? Please choose one:
 - Carnegie Mellon University
 - The University of Pittsburgh
 - Other
3. Are you a full-time or part-time student?
4. What is your age?
5. What is your gender?
6. Are you affiliated with the Greek life of your University?
7. How frequently do you consume at least one alcoholic beverage? Please choose one:
 - Less than 1 time per week, on average
 - From 1 to 2 times per week, on average
 - From 3 to 5 times per week, on average
 - Greater than 5 times per week, on average
8. Do you consume energy drinks in conjunction with alcohol?
9. Have you ever experienced complete or partial memory loss (i.e. a “blackout”) during an event that involved drinking?
 - If response is “no,” please skip to question #11.
10. Were you consuming an alcoholic energy drink or alcohol in conjunction with an energy drink in the events shortly prior to any blackouts?

not clear what you're using to stratify on or monitor coverage and unbiasedness here.

11. Have you ever directly or indirectly purchased alcoholic energy drinks?
-If response is "no," please skip to question #13.
12. Why did you choose to purchase alcoholic energy drinks? Please choose all that apply.
 - Price
 - Taste
 - Alcoholic content
 - Availability
 - Effects of consumption
 - Other reason(s)
13. Have you ever heard of the alcoholic energy product Four Loko?
-If response is "no," please skip to question #18.
14. How did you first hear about Four Loko? Please choose one:
 - A friend or relative
 - A print advertisement
 - News media
 - An online source
 - Other outlet
15. Have you ever consumed the product Four Loko?
-If response is "no," please skip to question #17.
16. The caffeine and energy stimulants were removed from Four Loko by January 1, 2011.
When did you consume Four Loko?
 - Before January 1, 2011
 - After January 1, 2011
 - Both before & after January 1, 2011
17. Would you consider drinking the product Four Loko since the removal of its caffeine and energy stimulants?
18. In general, would you consider drinking alcoholic energy drinks in the future?
19. Have the recent health risks associated with alcoholic energy drinks influenced your answer to the previous question?

these are related. you might consider instead a check-all-that-apply "what effect attracted you" question, along with a blank for any effects you didn't think of.

(give blank)

some "why" or "what risks" followups might be interesting here

M. Initial Sample Size Calculation

Because our survey consists of primarily "yes" or "no" answers, we consider most of our parameters to be from a Bernoulli distribution. Therefore, we will consider our standard deviation to be of the worst case scenario, by setting $p = .5$:

$$SD_{\text{Worst Case Scenario}} = \sqrt{(.5)(1 - .5)} = .5$$

It may be difficult to receive a large sample size since there are many groups conducting surveys within our class. Therefore, we will allow our margin of error to be up to 5%. We can calculate n_0 as follows:

$$n_0 = \frac{\left(\frac{z_{\alpha}}{2}\right)^2 (SD)^2}{(ME)^2}$$

90% instead of 95% CI's? OK, but you should say that.

$$n_0 = \frac{(1.645)^2 (.5)^2}{(.05)^2}$$

$$n_0 = 270.6025$$

Overall, the total number of undergraduate students attending both Carnegie Mellon University and the University of Pittsburgh is approximately 25,000. Because we are under the assumption that we are taking a random sample without replacement, we must make the following adjustment to our calculation above:

$$n \geq \frac{Nn_0}{N + n_0}$$

$$n \geq \frac{(25,000)(270.6025)}{(25,000) + (270.6025)}$$

$$n \geq 267.7048361$$

Therefore, to make inferences about our population concerning the questions we plan to ask with a margin of error of 5%, we must sample at least 268 total individuals.

Since we are conducting a “face to face” survey, with practice we may be able to get a response rate up to 70% (as noted in the lecture slides). Therefore, for the worst case scenario we will consider our response rate to be approximately 50%. If we take our response rate into account, we may need to physically ask more individuals for their participation in our survey. Specifically:

$$n_{\text{With .5 Response Rate}} \geq \frac{267.7048361}{.5}$$

this is great

$$n_{\text{With .5 Response Rate}} \geq 535.4096722$$

Thus, for a worst case scenario calculation, we need to ask approximately 536 individuals to take our survey, but only have approximately 268 of those individuals completely fill out our survey, to be able to make inferences with a 5% margin of error.