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 fifth 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. Variables we plan to stratify upon are what university students attend, their gender, and whether or not they are affiliated with Greek life.

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.

**Our revisions are in red, questions we would like to consider removing to simplify our survey are in blue. Please include your comments on their removal.**

1. (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.

1. What university do you attend? Please choose one:

-Carnegie Mellon University

-The University of Pittsburgh

-Other \_\_\_\_\_\_\_\_\_\_\_\_

1. Are you a full-time or part-time student?

-Full-Time

-Part-Time

1. What is your age? \_\_\_\_\_\_\_\_
2. What is your gender?

-Male

-Female

1. Are you affiliated with ~~the~~ Greek life ~~of~~ at your University?

-Yes

-No

1. On average, 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~~

1. ~~Do you consume energy drinks in conjunction with alcohol?~~ Have you ever experienced complete or partial memory loss (i.e. a “blackout”) while drinking?

-Yes

-No

1. ~~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.~~

Do you mix energy drinks and alcohol?

-Yes

-No

1. ~~Were you consuming an alcoholic energy drink or alcohol in conjunction with an energy drink in the events shortly prior to any blackouts?~~ Were you drinking alcoholic energy drinks when you “blacked out”?

-Yes

-No

-N/a

1. Have you ever directly or indirectly purchased alcoholic energy drinks?

-~~If response is “no,” please skip to question #13.~~

-Yes

-No

-N/a

1. ~~Why did you choose to purchase alcoholic energy drinks? Please choose all that apply.~~ Why do you drink alcoholic energy drinks? Choose all that apply:

-Price

-Taste

-Alcoholic content

-Availability

-~~Effects of consumption~~ For the caffeine

-Other ~~reason(s)~~ \_\_\_\_\_\_\_\_\_\_

1. Have you ever heard of ~~the alcoholic energy product~~ Four Loko?

~~-If response is “no,” please skip to question #18.~~

-Yes

-No

1. 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

1. Have you ever ~~consumed~~ drank ~~the product~~ Four Loko?

~~-If response is “no,” please skip to question #17.~~

-Yes

-No

1. The caffeine ~~and energy stimulants~~ ~~were~~ was removed from Four Loko by January 1, 2011. When did you ~~consume~~ drink Four Loko?

-Before January 1, 2011

 Would you consider drinking Four Loko since the removal of caffeine?

 -Yes

 -No

-After January 1, 2011

-Both before & after January 1, 2011

1. ~~Would you consider drinking the product Four Loko since the removal of its caffeine and energy stimulants?~~
2. ~~In general, would you consider drinking alcoholic energy drinks in the future?~~ Do you know the health risks associated with drinking alcoholic energy drinks?

-Yes

-No

1. ~~Have the recent health risks associated with alcoholic energy drinks influenced your answer to the previous question?~~ Will you consider drinking alcoholic energy drinks in the future?

-Yes

-No

1. Did the risks associated with drinking alcoholic energy drinks affect your answer to 19?

-Yes

-No

-N/a

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:

$$SD\_{Worst Case Scenario}= \sqrt{(.5)(1-.5)}= .5$$

 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%, allowing for a 90% confidence interval. We can calculate $n\_{0}$ as follows:

$$n\_{0}=\frac{(z\_{α/2})^{2} (SD)^{2}}{(ME)^{2}}$$

$$n\_{0}= \frac{\left(1.645\right)^{2}(.5)^{2}}{(.05)^{2}}$$

$$n\_{0}= 270.6025$$

 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:

$$n\_{With .5 Response Rate} \geq \frac{267.7048361}{.5}$$

$$n\_{With .5 Response Rate} \geq 535.4096722$$

 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.

Brief Report on Pretesting

Based on our pretesting, we made the wording of our questions more simplistic and easier to understand. We tried to make the survey less wordy so more people would be interested in taking it. We also took out questions we don’t think add anything to our research. Furthermore, we included answer choices to the questions, making them simpler to respond to.

Our concern is the length of our survey. As we indicated above, we propose to remove some of the questions we don’t think provide pertinent information. We know that the survey should

contain at least 20 questions, but since we are doing a man on the street survey while people are going to and from their obligations, we’d rather make the survey more simple and only include questions concerning the data we’re interested in collecting.

In question 2, if someone responds “other” should we tell them to stop the survey, or allow them to continue? If we allow them to continue, would we disregard their response to the survey, or would we include the information in general analysis (not where we compare the two schools)?