Stat 446 Homework 1

Leslie Gains-Germain

1. (a) Target population: For a given book, all readers who have read that book (it doesn't state that they need have bought the book from Amazon.com)

    Sampling Frame: anyone who has access to the internet and can make an account at Amazon.com (I actually checked-you don't need to have bought the book to write a review)

    Sampling Unit: an individual who has an account at Amazon.com

    Observation Unit: an individual who has an account at Amazon.com

    Two sources of selection bias or inaccuracy of responses:

    i. This is a voluntary response survey, so it will most likely over-represent readers who have strong opinions about the book, negative or positive.

    ii. They may get people posting reviews who didn't actually read the book. Anyone can make an account at Amazon.com and review a book, so there could potentially be reviews that are complete nonsense.

   (b) Target population: all homeless people

    Sampling Frame: homeless people who received medical attention from one of the clinics in the Health Care for the Homeless (HCH) project

    Sampling Unit: an individual homeless person

    Observation Unit: an individual homeless person

    Two sources of selection bias or inaccuracy of responses:

    i. This study misses all the homeless people who have never been to an HCH clinic. I would expect the estimate to be biased low because many homeless people with mental illnesses avoid health clinics.

    ii. This study also misses homeless people who live in cities without an HCH clinic. This will most likely bias the estimate high, because I would expect there to be a lower proportion of mentally ill homeless people in the cities that do not have an HCH clinic.

    No, I do not agree with their opinion that selection bias is not a problem. This is a judgement sample; they are trying to deem the sample representative because the demographics of HCH clients are similar

to the general homeless population. Demographics, however, aren't necessarily indicative of mental illness, and they will miss out on a huge portion of the homeless population by only sampling from HCH clinics.

2. (a) The major flaw is that the sample of customers flying on not very full flights is not representative of the population of all of the airline's customers. When the sample is not representative of the population, the estimates are biased.

   (b) I would expect these customer satisfaction surveys to overestimate the true customer satisfaction. The customers taking these surveys are on flights that are not very full, so they are most likely receiving a lot of attention from the flight attendants and they are very comfortable with all the extra room. I expect the estimates to be biased high.

3. (a) It might be unclear as to what is defined as a building; for example, is the football stadium considered a building? It also might be unclear as to where the campus boundaries lie. For example, are the fraternity houses considered MSU campus buildings?

   (b) It may be hard to find a conference room, if one does exist. Some buildings may have a room where they conduct meetings, but it might not be listed as an official "conference room".

4. (a) The target population is all undergraduate MSU students enrolled on the Bozeman campus.

   (b) The study population is all students who were at one of the five dining facilities for lunch on the Wednesday when sampling took place.

   (c) Undergraduates, graduate students, and guests often eat the dining facilities, so they may have sampled some people who are not in the intended target population. Also, I would be very surprised if every undergraduate student enrolled at the MSU campus was eating lunch at one of the five dining centers on the Wednesday when sampling took place. As a result, the study population and target population do not match and the estimates from this sample will not be valid for the target population.