Solutions to Chapter 11 homework problems

7. a) the population is US adults

b) the population parameter if interest is the percent of US adults that use and have benefited from alternative medicine

c) the sampling frame consisted of the subscribers in the Consumers Union

d) the sample consisted of the 12% of subscribers who responded

e) sampling method was not random, but was rather a convenience sample because though they are attempting to gain information about the entire US adult population, they were lazy and only asked their subscribers which do NOT represent the whole us adult population. This can be considered a convenience sample for that reason (easy/lazy pollsters trying to generalize a non-representative sample to the population).

f) Non-response bias since only 12% responded. Convenience sample also.

9. a) the population is adults of legal drinking age in this particular city

b) the population parameter of interest is the percent of adults of legal drinking age that believe drinking and driving is a problem

c) the sampling frame consists of the people at this particular bar that had been randomly selected

d) the sample consisted of the 4th person walking out of the bar and then every 10th person after that

e) the sampling method was a systematic sample with a random start. I would argue it was a multistage sample though because they treated the bars as clusters and then did a systematic inside the cluster.

f) They only sampled people walking out of a bar and therefore had likely been drinking. They might not be sober enough to answer the question and if they do they will probably say it's not a problem because they don't want to be pegged as a drunk driver. Additionally, since only one bar was selected only a certain part of the city would be represented which is not necessarily going to represent the city as a whole.

10. a) the population is this city's residents

b) the population parameter of interest is the top three issues resonating with voters

c) the sampling frame consists of all city residents

d) the sample consists of the residents that were found in the two blocks from each election district

e) the sample method was a multistage – first stratified by election district, then entire blocks were randomly selected and surveyed

f) Cluster samples have a large chance that they are not representative of the target population since all the people contained in the same two blocks will be similar. Additionally, not all of the residents will be registered to vote and therefore the results may not be a good estimate of the top 3 issues to *voters*.

12. a) No, this survey was not representative of the US adult population since only dorks who visit www.gamefaqs.com would have been sampled.

b) No. Since this was just a survey posted on their website for any visitor to respond to, there is voluntary response bias present. Only visitors who feel strongly about having a social life outside of the internet (or not) are likely to respond.

13. a) An SRS can't be done because cops would need a list of vehicles to choose from. OR, they can't just pull over 200 people and then randomly select a few to be sampled and checked for registration, insurance, and safety inspection stickers. Even if cops used a list of registered vehicles to choose a sample, unregistered vehicles (the problematic ones that they want to locate) will not be on the list!

b) Use a RNG (random number generator) to choose a random starting vehicle and then check every 20th car after that. Or use a 20-sided die to pick cars going through the checkpoint by rolling it for each car and if your favorite number comes up, check the car. Also, profiling will likely factor into which cars they pick, but of course that's a social issue, not a sampling design issue.

15. Stupid TV station poll was biased of course! It was a classic voluntary response poll and so only the people who cared called in their votes and the results were nowhere close to the real thing. Surprise surprise.

16. The newspaper poll sounds like a pretty legitimate attempt: they stratified by multiple characteristics in order to get as representative of a sample as possible. It is just unfortunate that the sample did not come close to the real election results. This is sampling error because it is based on natural variation that occurs when taking samples. No sample (even a great sample) can be *guaranteed* to give accurate results because of randomness. Sometimes you just get lucky or unlucky.

24. a) HA! Yeah right. If you have a question as loaded as this, simply asking more people will give you an even bigger biased sample. **You can't fix bias by getting a bigger sample. Instead you have to start over.**

b) Yes. The question is biased because the reader is given very negative information about ephedra leading up to a negative leaning question.

c) The sampling frame only consists of people visiting the site.

d) Correct. Since only those with strong opinions on the matter will respond, the results cannot be generalized to any population.

25. a) This is a biased question because of words such as "pollute" and "compelled." It makes the reader think that big bad companies are automatic polluters that owe it to society to clean up their waste. And even though no one would say, "I don't care if they pollute. We'll drink the water anyway," the question will certainly get the response the pollster is looking for. A better way to ask would be, "Should companies be held responsible for any costs associated with environmental clean-up?"

b) This is a leading question because of words such as "fair" in addition to stressing the fact that 18 year olds can serve in the military (yet not drink). A better way to ask would be, "Do you believe the drinking age should be lowered?"

26. a) Seems neutral. Though the word "required" might seem pretty strong for some people. Maybe just eliminated that word would make it better. "Should high school students wear uniforms?"

b) Biased. Prepping the question with a positive outlook on exploration will lead respondents to agree that we should continue funding space flights. Consider if it had been worded, "Given our nation's surmounting debt, do you favor continued funding of space flights?" then you will probably get more no's. A more neutral way to ask is to leave out the beginning phrase: "Do you favor continued funding of space flights?"

36. a) A petition is like a one-sided survey – lots of response bias. If you show up to someone's door saying you are gathering signatures for a new playground then sure everyone will want to sign unless they are an idiot or a grouch. This does not really give them a true neutral opportunity for dissent. Additionally, there might be many people out and about on a Saturday afternoon and not at home so there is some undercoverage involved.

b) If you like the food somewhere you will most likely eat there and the opposite is also true. If they survey people in line then those people will probably be the ones who like (or at least tolerate) the food. Timing also matters. If they sample people during the lunch rush it could just be people who only have 30 minutes before their next class and don't have a choice except to eat there. Some cafeterias serve certain things on certain days. If they survey people on a "good" food day as opposed to a day when the menu is less favorable to diners they will get different responses. It also matters if the food at this cafeteria is representative of the food served at all the cafeterias on campus. If not, then there is no way to tell in general if the students are satisfied with the food.

37. a) The question seems ok on its face, but then who is considered "family"? And what does "seek" mean? So it's a little ambiguous and might be interpreted differently by different people.

b) i) Cluster   ii) Stratified by classification   iii) Systematic

c) No. Stupid professor probably failed statistics class. Remember I stressed in class that a **SRS has a very specific definition. Each possible subset must have the same opportunity of being sampled.** Since the professor chose entire rows to be included in the sample, there is no way for people NOT on those rows to be sampled. To be a SRS the professor should have numbered the students and picked numbers out of a hat. Or she could have assigned all the students a number using a random number generator and divide them by even or odd or whatever.

d) i) This is voluntary response bias so only those with strong opinions will responds. Since the question targets students who are first in their family to go to college the estimate will likely be much higher than the real percent since those students will be the ones responding.

d) ii) This is a convenience sample. The surveyor is asking students who pass by to stop and answer. This will likely not be representative and will likely either over or underestimate the real percent of students who are first in their family to go to college.

e) The president is misusing "population parameter" and "sample statistic". The estimate from the sample survey cannot be assumed to be the actual percent of all the students who are first attenders. He should make that clear that it is only an estimate.

38. a) The company's claim should stress that it is only an estimate. The way they stated the results it sounds like they did a census, which is not true.

b) Yes. Numbers were randomly assigned in the first place. After that I could sample those divisible by 7 or 8 or evens/odds or whatever.

c) i) stratified   ii) systematic   iii) multi-stage with double SRS

d) Heavy response bias. Who is going to refuse to raise their hand so that their boss and other employees know they are unhappy at their job?

e) i) voluntary response bias: only those who feel strongly will respond. It might underestimate those who are happy since maybe only those who hate their job will respond (or visa versa).

e) ii) Non-response bias: many will not respond and therefore the estimate will be an over or underestimate.