STAT 315 Chapter 7 Review Questions

Scenario A: A <u>recent article published in Annals of Internal Medicine</u> investigated the median incubation period for the COVID19 virus, that is, the time between exposure to the virus and presenting symptoms. From 181 cases, they estimate the median exposure to be 5.1 days, with a 95% confidence interval of 4.5 to 5.8 days.

1.   What parameter is being estimated?

2.   What is the point estimate?

3.   What is the margin of error?

4.   Why didn't the researchers estimate the mean instead of the parameter they actually estimated?

5.   When the population distribution is not known, confidence intervals rely on a Central Limit Theorem to provide the sampling distribution. The CLT relies on random sampling. The article states: "Publicly reported cases may overrepresent severe cases, the incubation period for which may differ from that of mild cases." What does this statement communicate about the sampling procedure, and how should that impact your interpretation of the confidence interval?

6.   Why might researchers want to use a one sided confidence interval in this scenario?

Scenario B: In lecture, we have discussed how outliers may affect the sample mean and variance. Now let's look at how confidence intervals might be affected by outliers. Assume we are investigating a therapeutic treatment for ankle pain using class IV lasers. In the present study, patients each receive regular "doses" of the treatment until they report substantial pain relief. Our goal is to measure the average number of doses required to provide relief. 20 patients undergo the therapy and the data are provided below:

2	1	2	2	1	3	2	2	1	2	3	1	1	1	2	1	2	3	2	3

7. Create a 95% confidence interval for the mean number of doses required to provide relief.

8. Now presume that the last patient required 10 doses instead of 3, and create a new 95% confidence interval.

9. How was the point estimate affected by the introduction of an outlier?

10. How was the precision of the estimate affected by the introduction of an outlier?

Scenario C: Suppose ecologists are attempting to estimate the proportion of mule deer in rocky mountain national park with infectious keratoconjunctivitis, a disease affecting the tissue around the eye. On two separate weekends, they track herds of deer and look for visual signs of the disease. On the first weekend, they observe 56 deer, 12 of which have the disease, and on the second weekend they observe 79 deer, 23 of which have the disease.

11. Construct a 95% confidence interval for the proportion of deer which have the disease using data from the first weekend.

12. Construct a 95% confidence interval for the proportion of deer which have the disease using data from the second weekend.

13. Construct a 95% confidence interval for the proportion of deer which have the disease by combining both weekends's worth of data.

14. Which is preferred, the two confidence intervals you calculated from repeated samples in problems 11 and 12, or the single combined confidence interval from problem 13?