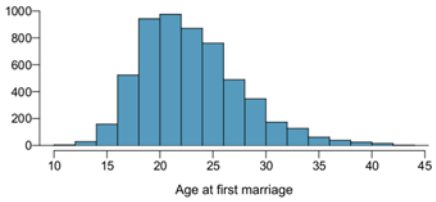
Problem 1

The National Survey of Family Growth conducted by the Centers for Disease Control gathers information on family life, marriage and divorce, pregnancy, infertility, use of contraception, and men's and women's health. One of the variables collected in this survey is the age at first marriage. The histogram shows the distribution of ages at first marriage of 5,534 randomly sampled women between 2006 and 2010. The average age at first marriage among these women is 23.44. Consider the population standard deviation of 4.72 (NSFG, 2010). (please round ages to two decimal places)



1. Estimate the average age at first marriage of women using a 95% confidence interval.

[Sample Mean ± Z * SD / $\sqrt{n}$] = [23.44 ± (1.96)$\frac{4.72}{\sqrt{5534}}$] = [23.32, 23.56]

2. Interpret the confidence interval in context.
   a) The researchers can be 95% confident that any sample of women's ages at their first marriage would be within this confidence interval
   b) The researchers can be 95% confident that the population mean age of women at their first marriage is within our confidence interval
   c) 95% of women have their first marriage when their age is within our confidence interval

Problem 2

The 2010 General Social Survey asked the question: "For how many days during the past 30 days was your mental health, which includes stress, depression, and problems with emotions, not good?" Based on responses from 1,151 US residents, the survey reported a 95% confidence interval of 3.40 to 4.24 days in 2010.

3. Interpret this interval in context of the data.
   a) 95% of surveys will report a mean number of "not good" days in the past 30 is between 3.40 and 4.24
   b) The researchers can be 95% confident that the true population mean number of "not good" days in the past 30 is between 3.40 and 4.24

[1] Math-UOttawa 2. UVermont 3 Utts [4] OpenIntro

c) The researchers can be 95% confident that the sample mean number of "not good" days in the past 30 is between 3.40 and 4.24

4. Suppose the researchers think a 99% confidence level would be more appropriate for this interval. Will this new interval be smaller or larger than the 95% confidence interval?
   a) smaller since we will be surer of our results
   b) larger since the margin for error must be larger
   c) smaller since we have less room for error
   d) larger since the standard error would be larger

5. If a new survey were to be done with 500 Americans, would the standard error of the estimate be larger, smaller, or about the same. Assume the standard deviation has remained constant since 2010.
   a) larger since we can be less sure of our estimate with a smaller sample size
   b) smaller since we can collect more accurate results from fewer individuals
   c) about the same since the standard deviation has remained constant, so we shouldn't expect different results

---

Problem 3

A hospital administrator hoping to improve wait times decides to estimate the average emergency room waiting time at her hospital. The population standard deviation is provided. She collects a simple random sample of 64 patients and determines the time (in minutes) between when they checked in to the ER until they were first seen by a doctor. A 95% confidence interval based on this sample is (128 minutes, 147 minutes), which is based on the normal model for the mean.

6. Determine whether the following statements are true or false, and explain your reasoning. 4

- This confidence interval is not valid since we do not know if the population distribution of the ER wait times is nearly Normal.
  False, because $n \geq 30$
- We are 95% confident that the average waiting time of these 64 emergency room patients is between 128 and 147 minutes.
  False, because inference is made on parameter
- We are 95% confident that the average waiting time of all patients at this hospital's emergency room is between 128 and 147 minutes.
  True
- 95% of random samples have a sample mean between 128 and 147 minutes.
  False, the CI is not about a sample mean.

[1] Math-UOttawa 2. UVermont 3 Utts [4] OpenIntro

- A 99% confidence interval would be narrower than the 95% confidence interval since we need to be "surer" of our estimate.
  False, it will be wider
- The margin of error is 9.5 and the sample mean is 137.5.
  True
- In order to decrease the margin of error of a 95% confidence interval to half of what it is now, we would need to double the sample size.
  False, To cut the SE (or margin of error) in half, we would need a sample of 256 that is four times the original sample size.

Problem 4[4]

In 2013, the Pew Research Foundation reported that 45% of U.S. adults report that they live with one or more chronic conditions". However, this value was based on a sample, so it may not be a perfect estimate for the population parameter of interest on its own. The standard deviation of about 1.2%, and a normal model may reasonably be used in this setting.

7. Which conditions were checked in order to use the normal model?
   $n\hat{p} \geq 10$ and $n(1-\hat{p}) \geq 10$
8. Create a 95% confidence interval for the proportion of U.S. adults who live with one or more chronic conditions. Also interpret the confidence interval in the context of the study.
   Recall that the general formula is: point estimate ± Z*SD

   First, identify the three different values. The point estimate is 45%, Z^* = 1.96 for a 95% confidence level, and *Sd* = 1.2%. Then, plug the values into the formula:

   45% ± 1.96 × 1.2% → (42.6%, 47.4%)

   We are 95% confident that the proportion of US adults who live with one or more chronic conditions is between 42.6% and 47.4%.

Problem 5[4]

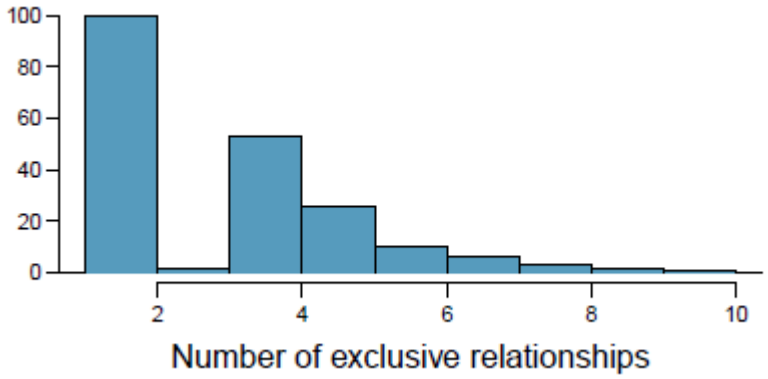
In 2013, the Pew Research Foundation reported that "45% of U.S. adults report that they live with one or more chronic conditions", and the standard deviation is 1.2%. Identify each of the following statements as true or false. Provide an explanation to justify each of your answers.

[1] Math-UOttawa 2. UVermont 3 Utts [4] OpenIntro

9. We can say with certainty that the confidence interval from Problem 1 contains the true percentage of U.S. adults who suffer from a chronic illness.
   False. Confidence intervals provide a range of plausible values, and sometimes the truth is missed. A 95% confidence interval "misses" about 5% of the time.
10. If we repeated this study 1,000 times and constructed a 95% confidence interval for each study, then approximately 950 of those confidence intervals would contain the true fraction of U.S. adults who suffer from chronic illnesses.
    True. Notice that the description focuses on the true population value.
11. Since the standard deviation is 1.2%, only 1.2% of people in the study communicated uncertainty about their answer.
    False. The standard deviation describes the uncertainty in the overall estimate from natural fluctuations due to randomness, not the uncertainty corresponding to individuals' responses.

Problem 6[4]

A survey conducted on a reasonably random sample of 203 undergraduates asked, among many other questions, about the number of exclusive relationships these students have been in. The histogram below shows the distribution of the data from this sample. The sample average is 3.2 and the sample standard deviation is 1.97.



12. Estimate the average number of exclusive relationships Duke students have been in using a 90%confidence interval and interpret this interval in context. Check any conditions required for inference, and note any assumptions you must make as you proceed with your calculations and conclusions.

Answer: Independence: random sample. We can assume that the students in this sample are independent of each other with respect to number of exclusive relationships they have been in. Notice that there are no students who have had no exclusive relationships in the sample, which suggests some student responses are likely missing (perhaps only positive values were reported). The sample size is at least 30.

[1] Math-UOttawa 2. UVermont 3 Utts [4] OpenIntro

The skew is strong, but the sample is large so this is not a concern. 90% CI: (2.97, 3.43). We are 90% confident that undergraduate students have been in 2.97 to 3.43 exclusive relationships, on average.

Problem 7[4]

The 2010 General Social Survey asked the question: "After an average work day, about how many hours do you have to relax or pursue activities that you enjoy?" to a random sample of 1,155 Americans.[41] A 95% confidence interval for the mean number of hours spent relaxing or pursuing activities they enjoy was (1.38, 1.92).

13. Interpret this interval in context of the data.

We are 95% confident that Americans spend an average of 1.38 to 1.92 hours per day relaxing or pursuing activities they enjoy.

14. Suppose another set of researchers reported a confidence interval with a larger margin of error based on the same sample of 1,155 Americans. How does their confidence level compare to the confidence level of the interval stated above?

Their confidence level must be higher as the width of the confidence interval increases as the confidence level increases.

15. Suppose next year a new survey asking the same question is conducted, and this time the sample size is 2,500. Assuming that the population characteristics, with respect to how much time people spend relaxing after work, have not changed much within a year. How will the margin of error of the 95% confidence interval constructed based on data from the new survey compare to the margin of error of the interval stated above?

The new margin of error will be smaller since as the sample size increases the standard error decreases, will decrease the margin of error.

Problem 8:

A researcher conducted an experiment on 8 randomly selected NASCAR drivers in which their reaction time was measured. The sample mean reaction time was 1.24 seconds. The sample standard deviation reaction time was 0.12 seconds. Assume that reaction time follows a normal distribution.

16. A 98% confidence interval for the population mean reaction time based on these data is given by

a) 1.24 ± 0.083
b) 1.24 ± 0.127
c) 1.24 ± 0.099
d) 1.24 ± 0.120
e) Sample size is too small to construct a confidence interval

[1] Math-UOttawa 2. UVermont 3 Utts [4] OpenIntro

Problem 9:

Laura is interested in the mean height of the girls in her 9th grade class year. Assume the population standard deviation is known to be 1.2 inches. She takes a sample of 48 students and calculates a sample mean of 64.5 inches and a sample standard deviation of 0.9.

17. What is the parameter of interest?

(a) The mean height of the 48 girls in Laura's 9th grade class year.
(b) The mean height of girls in Laura's 9th grade class year.
(c) The mean height of the girls in Laura's entire school.
(d) The mean height of the 48 girls in Laura's entire school.

18. Assume Laura wants to create a 95% confidence interval about the true parameter. What is the margin of error?

(a) 0.89
(b) 1.07
(c) 0.52
(d) 0.34

[1] Math-UOttawa 2. UVermont 3 Utts [4] OpenIntro