All questions in Homework 5 will be submitted and graded as "multiple answer" questions.
Some questions may have one correct answer; some will have multiple correct answers.
You must correctly answer ALL the parts of a question to earn ANY credit for it. (That is, every choice must be correctly chosen or not chosen.)

That is a strong requirement, but appropriate for ensuring that you fully understand the ideas. This is to emphasize the importance of reading and thinking carefully.

(2 pts) **Question 1**: In 2015, a national polling organization surveyed US adults. One of the questions was "Do you live with one or more chronic health conditions?" A 90% confidence interval for the result of this question was 0.45 to 0.49.

Based on the information provided in this course about interpreting confidence intervals, which of these is/are **incorrect** statements interpreting this result? (For your own understanding, be able to describe WHAT is incorrect about each.)

   a. I have 90% confidence that the population proportion of US adults who report living with one or more chronic health conditions is between 0.45 and 0.49.
   b. I have 90% confidence that the proportion of people who report living with one or more chronic health conditions is between 0.45 and 0.49.
   c. I have 90% confidence that the sample proportion of US adults who report living with one or more chronic health conditions is between 0.45 and 0.49.
   d. I am 90% sure that the population proportion of US adults who report living with one or more chronic health conditions is between 0.45 and 0.49.
   e. There is a 90% probability that the population proportion of US adults who report living with one or more chronic health conditions is between 0.45 and 0.49.

(2 pts) **Question 2**: It is known that the proportion of left-handed people in the overall population is 10%. A random sample of the 5th-grade students in our school district was selected and it was found that the proportion of left-handed students is 7.3%. We test the claim that the proportion of our school districts 5th grade students who are left handed is less than 10%. The computed p-value is 0.061.

Which of these is/are **incorrect** statements interpreting this p-value? (For your own understanding, be able to describe WHAT is incorrect about each.)

   a. The probability of finding a sample statistic as extreme or more extreme than the sample statistic of our data is 0.061
   b. The probability of finding a sample statistic as extreme or less extreme than the sample statistic of our data is 0.061 if the null hypothesis is true.
   c. The probability of finding a sample statistic as extreme or less extreme than the sample statistic of our data is 0.061 if the alternative hypothesis is true.
   d. The probability of finding a sample statistic as extreme or more extreme than the sample statistic of our data is 0.061 if the null hypothesis is true.
   e. The probability of finding a sample statistic as extreme or more extreme than the sample statistic of our data is 0.061 if the alternative hypothesis is true.

(2 pts) **Question 3**: Your client has done some estimates and calculations and determined that the overall cost of test-result errors will be minimized by using a significance level of 0.02. Using this, for which of these p-values would you "Reject Ho"?

   a. p-value = 0.083
   b. p-value = 7.1%
   c. p-value = 2.3%
   d. p-value = 0.009
   e. p-value = 0.0505

(1 pt) **Question 4**. You have done a hypothesis test and your *p*-value is 0.043. According to the description in our course of appropriate "strength of evidence" conclusions, which one of the following is the best summary of this result?

   a. These data provide very strong evidence for the alternative hypothesis (p = 0.043.)
   b. These data provide strong evidence for the alternative hypothesis (p = 0.043.)
   c. These data provide moderate evidence for the alternative hypothesis (p = 0.043.)
   d. These data provide some evidence for the alternative hypothesis (p = 0.043.)
   e. These data provide weak evidence for the alternative hypothesis (p = 0.043.)

(1 pt) **Question 5**: A study is done to determine which of two different drugs, Drug A or Drug B, is more effective in treating a certain medical condition. The subjects are volunteers from the patients with this medical condition who are patients of doctors from our family of hospitals. The patients are assigned randomly to Drug A or Drug B and their time for recovery is recorded.

From the design of the study, determine which one of these is the best summary of the types of conclusions that may be made.

   a. Cannot generalize to the population and cannot provide evidence for causality
   b. Cannot generalize to the population but can provide evidence for causality.
   c. Can generalize to the population but cannot provide evidence for causality.
   d. Can generalize to the population and can provide evidence for causality.

(1 pt) **Question 6**. A study is done to gather evidence on whether smoking (or equivalent) affects students' grade point average at a certain university. The researchers gather information on a random sample of the students at the university. They obtain the grade point averages of the students and information about whether they use nicotine-based products (smoking or e-cigarettes.) Assume there is a good response rate and that the information available about their habits and the GPAs is accurate.

From the design of the study, determine which one of these is the best summary of the types of conclusions that may be made.

   a. Cannot generalize to the population and cannot provide evidence for causality
   b. Cannot generalize to the population but can provide evidence for causality.
   c. Can generalize to the population but cannot provide evidence for causality.
   d. Can generalize to the population and can provide evidence for causality.

(1 pt) **Question 7**. Consider the characteristics of a "confounding variable." Which one or more of these fully describe the minimal characteristics that must be met for a variable to be a confounding variable.

   a. It is not associated with the explanatory variable or the response variable.
   b. It is not associated with the explanatory variable and it is associated with the response variable.
   c. It is associated with the explanatory variable and is not associated with the response variable.
   d. It is associated with the explanatory variable and with the response variable.