

Name: _____

AP Statistics Summer Assignment

Complete each problem to the best of your ability. If you do not know how to complete a problem, complete as much as you can and then indicate that you are unable to complete it by writing "unsure how to progress." If you do not write your answers on the page provided, please be sure to number each problem correctly.

Section 1: Variables

1. Studies have found that people who drink more red wine are less likely to develop heart disease. Give an example of a confounding variable and explain why it is confounding.

2. Researchers want to determine whether black tea or green tea is more effective at reducing stress. They split the volunteers into three groups: black tea, green tea and hot water. The subjects drink their assigned beverage each morning, and their stress levels are recorded at the end of the experiment.
 - a. What are the explanatory and response variables?

 - b. What is the purpose of the hot water group in this experiment?

3. Jimmy wants to determine whether students at his school support changing the school lunch menu. Jimmy waits by the vending machine and surveys the first 20 students that come to the vending machine each lunch period. Explain to Jimmy why his sampling method is flawed and the effect that it will have the results of his survey.

Section 2: Probability

4. A bag of marbles contains 3 green marbles, 6 yellow marbles and 5 purple marbles.
 - a. Johnny watches Clarissa pull out a yellow marble. Clarissa does not put the marble back in the bag. What is the probability that Johnny pulls a yellow marble as well?

 - b. Johnny and Clarissa put their marbles back in the bag. Then, Preston pulls a yellow marble. He puts the marble back in the bag and hands the bag to Demarco. What is the probability that Demarco also pulls a yellow marble?

5. The probability that Alfalfa does not wreck on his way to work is 0.99993. What is the probability that Alfalfa does wreck on his way to work?

6. At Chili's school, there are 92 freshmen. Every freshman can choose between Latin, Spanish, Italian or no language. They may also take more than one language class if they choose to. 42 freshmen take Spanish, 36 take Latin and 51 take Italian. 15 take all three classes, 12 take just Latin and Spanish, 3 take only Latin and 4 take only Spanish.
- Create a Venn Diagram modeling this situation.
 - How many students take only Latin?
 - How many students do not take any of the three language classes?
7. At a local shelter, they test every dog that enters the shelter for rabies before allowing it to enter the shelter. The dog must be put down if it tests positive for rabies. This test, however, like any medical test is not always completely accurate. If a dog has rabies, there is a 1% chance that the test will return negative. If a dog does not have rabies, there is still a 4% chance that the test returns positive. Based on historical data, 0.8% of all dogs that enter the shelter are actually infected with rabies.
- Create a tree diagram modeling this situation. Use your diagram to find the probability that a randomly selected dog actually has rabies if it tests positive for the disease.

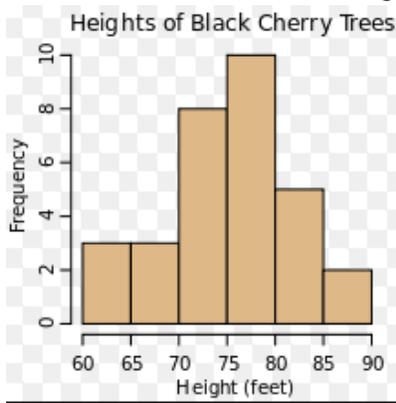
In order to ensure that they do not mistakenly put down a dog that does not have rabies, the shelter requires two consecutive positive tests before they will put the dog down.

- What is the probability that a randomly selected dog without rabies is put down due to two consecutive positive tests?
 - What is the probability that a dog with rabies will be allowed into the shelter?
8. Mars is running a promotion with M&M's at a local convenience store. Inside the package, Mars claims there is a 1 in 10 chance of receiving a voucher for a free pack of M&M's. If the customer wins the free pack, the store owner loses money. The owner is concerned when a group of 9 friends enter the store and 3 of them win a free pack of M&M's. To determine if Mars is telling the truth, he performs 100 simulations of 9 customers entering the store and buying M&M's under the assumption that the 1 in 10 claim is true, and finds that 3 of them result in 3 or more customers winning the free M&M's. Based on his simulation, does it seem likely that the 1 in 10 claim made by Mars is true?

Section 3: Creating and Interpreting Graphs

9. A data set is given as $\{0, 3, 3, 4, 5, 6, 7, 7, 7, 7, 8, 9, 9, 10, 11, 13, 15, 19, 22\}$
- Find the 5-number summary (minimum, Q1, Median, Q3, maximum) for the data set.
 - State the IQR of the data set.
 - An outlier is defined as any value more than $1.5 \cdot \text{IQR}$ above the third quartile or $1.5 \cdot \text{IQR}$ below the first quartile. Find the range of values that are NOT considered outliers. Use this determine if there are any outliers in the data set.

10. What is the median of the histogram shown below?



11. The histogram above shows the heights of black cherry trees in a local grove. Explain why a bar graph would not be an appropriate representation of this data.

Section 4: Variability

12. Justin is designing a study to estimate the proportion of registered voters who plan to vote Republican in the upcoming election. He has planned to randomly survey 100 registered voters and record their response. Paco is planning to do a similar survey, but he plans to survey 150 registered voters instead of 100. Whose survey do you expect to yield a result closer to the true proportion of registered voters who plan to vote Republican? Explain your answer.
13. Justin uses his data to create an interval that he is 99% sure contains the true proportion of registered voters who plan to vote Republican in the upcoming election. The resulting interval is $(0.4248, 0.5167)$. If Republicans need at least 50% of the vote to win the election, does Justin's interval provide clear evidence that the Republicans will win the election?

