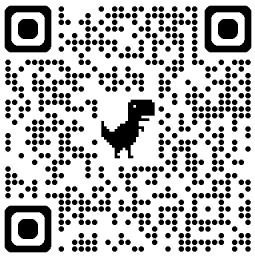
Name:

Common Core Algebra II

Unit 10

Statistics

Mr. Schlansky



2025

Lesson 1: I can identify the three statistical studies given their components.

A *survey* (sample survey) is a type of observational study that gathers data by asking people a number of questions.

A *census* (population survey) is a survey where every member of the population is being surveyed.

An *observational study* records the values of variables for members of a sample. NO TREATMENT IS ADMINISTERED.

A *controlled experiment* **randomly selects** a sample and **randomly assigns** members of the sample to a treatment and control group for the purpose of seeing what effect the treatments have on some response. THE SAMPLE RECEIVES A TREATMENT.

*They all can suggest patterns and relationships but only a CONTROLLED EXPERIMENT CAN DETERMINE A CAUSE AND EFFECT RELATIONSHIP.

Lesson 2: I can understand the components of a survey by understanding bias.

A good sample should be randomly selected where every member of the population has a chance of being chosen.

The sample should be large enough to eliminate random chance. Sample size should be a minimum of 30.

A random sample should have no bias. For example, to determine what percentage of a school's population likes soccer, ask every 10^{th} person walking into the building, not the soccer team. Asking the soccer team would be bias because you are asking people that you already know like soccer. They do not represent the population of the entire school.

Lesson 3: I can understand that as sample size increases, the mean relatively stays the same and the standard deviation decreases.

As sample size increases, the mean relatively stays the same and the standard deviation/margin of error decreases.

As sample size decreases, the mean relatively stays the same and the standard deviation/margin of error increases.

*The larger the sample size, the more accurate the results.

Lesson 4: I can calculate the margin of error $(2(S\tan dard\ \ Deviation))$ **and confidence interval** $(mean \pm 2(S\tan dard\ \ Deviation))$ **and use the confidence interval to determine if a value is expected.**

To determine if something is expected or usual, determine if it falls within the confidence interval.

$Confidence\ \ Interval = mean \pm 2(S\tan dard\ \ Deviation)$

The margin of error is how far an expected value can be from the mean.

$M\arg in\ of\ \ Error = 2(S\tan dard\ \ Deviation)$

Lesson 5: I can solve 4 point questions by determine if something is expected by finding its confidence interval using $Confidence\ Interval = mean \pm 2(S\tan dard\ \ Deviation)$

1) Find the Confidence Interval ($Confidence\ Interval = mean \pm 2(S\tan dard\ \ Deviation)$)
2) If the given value is
   a) Inside the confidence interval: Yes, it is an expected value because it's inside the CI
   b) Outside the confidence interval: No, it is not expected because it's outside the CI

Lesson 6: I can determine if an object is fair by determining whether the object's output is inside the fair object's confidence interval.

1) Calculate the proportion for the actual object. (Divide)
2) Find the confidence interval for the fair object .
$Confidence\ Interval = mean \pm 2(S\tan dard\ \ Deviation)$
3) If the actual object's proportion is inside the confidence interval, it is fair.
If the actual object's proportion is not inside the confidence interval, it is not fair.

Lesson 7: I can design a controlled experiment by following its procedure.

To design a controlled experiment:

1) Randomly select a sample.
2) Randomly assign half of the sample to a treatment group and the other half to a control group.
3) Apply the treatment to the treatment group and a placebo to the control group (if possible).
4) Analyze the data.

Lesson 8: I can determine if a treatment is effective by seeing if its mean different is in the confidence interval.

To determine if a treatment is effective:

1) Find the mean difference between the treatment and control group.
2) Rerandomize the sample many times and record the mean differences on a dot plot.

   If the mean difference falls within the confidence interval, the treatment is not effective.

   If the mean difference falls outside the confidence interval (less than 5%), the treatment is effective.

Lesson 9: I can prepare for my Probability/Statistics Test by practicing!

Name _____ Date _____
Mr. Schlansky Algebra II

Statistical Studies

1. Determine whether each scenario is a survey, census, observational study, or controlled experiment. Explain your answer.

A. Researchers want to determine if there is a relationship between whether or not a woman smoked during pregnancy and the birth weight of her baby. Researchers examined records for the past five years at a large hospital.

B. A large high school wants to know the proportion of students who would be interested in a driver's education class. Counselors asked a random sample of 200 students if they would be interested.

C. A company develops a new dog food. The company wants to know if dogs would prefer its new food over the competition's dog food. One hundred dogs, who were food deprived overnight, were given equal amounts of the two dog foods: the new food versus the competitor's food. The proportion of dogs preferring the new food versus the competitor's was recorded.

D. An elementary school Principal wants to determine if any students in the school have a peanut allergy. He has every family in the school return a questionnaire regarding peanut allergies.

2. Which scenario is best described as an observational study?

1) For a class project, students in Health class ask every tenth student entering the school if they eat breakfast in the morning.
2) A social researcher wants to learn whether or not there is a link between attendance and grades. She gathers data from 15 school districts.
3) A researcher wants to learn whether or not there is a link between children's daily amount of physical activity and their overall energy level. During lunch at the local high school, she distributed a short questionnaire to students in the cafeteria.
4) Sixty seniors taking a course in Advanced Algebra Concepts are randomly divided into two classes. One class uses a graphing calculator all the time, and the other class never uses graphing calculators. A guidance counselor wants to determine whether there is a link between graphing calculator use and students' final exam grades.

3. A doctor wants to test the effectiveness of a new drug on her patients. She separates her sample of patients into two groups and administers the drug to only one of these groups. She then compares the results. Which type of study *best* describes this situation?

1) census
2) survey
3) observation
4) controlled experiment

4. A market research firm needs to collect data on viewer preferences for local news programming in Buffalo. Which method of data collection is most appropriate?

1) census
2) survey
3) observation
4) controlled experiment

5. A school cafeteria has five different lunch periods. The cafeteria staff wants to find out which items on the menu are most popular, so they give every student in the first lunch period a list of questions to answer in order to collect data to represent the school. Which type of study does this represent?

1) observation
2) controlled experiment
3) population survey
4) sample survey

6. Determine whether each scenario is a survey, an observational study, or a controlled experiment. Explain your answer.

a) A study is done to see how high soda will erupt when mint candies are dropped into two-liter bottles of soda. You want to compare using one mint candy, five mint candies, and 10 mint candies. You design a cylindrical mechanism, which drops the desired number of mint candies all at once. You have 15 bottles of soda to use. You randomly assign five bottles into which you drop one candy, five into which you drop five candies, and five into which you drop 10 candies. For each bottle, you record the height of the eruption created after the candies are dropped into it.

b) You want to see if fifth-grade boys or fifth-grade girls are faster at solving multiplication problems. You randomly select twenty fifth-grade boys and twenty fifth-grade girls from fifth graders in your school district. You time and record how long it takes each student to solve multiplication problems.

c) You want to determine if people would be interested in watching a video of you performing Mr. Schlansky's math songs. You ask every 5^{th} student walking into Mr. Schlansky's math class if they would want to watch the video.

7. Howard collected fish eggs from a pond behind his house so he could determine whether sunlight had an effect on how many of the eggs hatched. After he collected the eggs, he divided them into two tanks. He put both tanks outside near the pond, and he covered one of the tanks with a box to block out all sunlight. State whether Howard's investigation was an example of a controlled experiment, an observation, or a survey. Justify your response.

8. Darryl conducted a study comparing the statistics of baseball players in the steroid era compared to the non steroid era. Would this investigation be an example of a controlled experiment, an observation, or a survey? Justify your response.

Name _____ Date _____
Mr. Schlansky Algebra II

Surveys

Determining an unbiased, random sample. A *random sample* is one that is selected in a way that gives every different possible sample an equal chance of being chosen. Every member of the population must have the opportunity to be chosen in the sample. The sample should be large enough (30) to represent a good portion of the population.

1. A survey team wants to determine what the favorite foods are of students in a high school. Determine whether the following would be unbiased, random samples. Assume all students have a normal schedule.
   a) Asking every fifth student entering the cafeteria

   b) Asking all fifth period English 9 students

   c) Asking 5 randomly selected students in every physical education class

   d) Asking 20 randomly selected students in all study halls

   e) Asking students whose street address name starts with a vowel

   f) Asking randomly selected students in the culinary club

   g) Asking all students in a randomly selected English 9, English 10, English 11, and English 12 class

   h) Asking every fifth student entering the building in the morning

   i) Asking every fifth student at the Varsity Basketball game

2. Determine if each of the following scenarios is an unbiased, random sample to produce accurate results. If it is not a good sample, explain why.
   a) Asking 100 randomly selected people walking into Westbury's public library what their favorite food is to determine the favorite foods of people in Westbury.

   b) Asking 85 randomly selected people walking into Westbury's public library what their favorite food is to determine the favorite foods of people in the United States.

   c) Asking 80 randomly selected people at Westbury's football game what their favorite sport is to determine the favorite sports of people in Westbury

   d) Asking 15 randomly selected teenagers at the Broadway Mall what their favorite type of music is to determine the favorite type of music for teenagers that live near the Broadway Mall.

   e) Asking 100 randomly selected teenagers at the Broadway Mall what their favorite type of music is to determine the favorite type of music for teenagers that live near the Broadway Mall.

   f) Asking 50 randomly selected people walking down Westbury's main street who their favorite athlete is to determine who Westbury's favorite athletes are.

   g) Posting a QR code at Westbury High School for to determine interest in a cooking club at Westbury High School.

h) Asking 10 randomly selected people at every library in the United States what their favorite streaming service is to determine favorite streaming services of people in the United States.

i) Asking 5 randomly selected adults at a supermarket in Westbury what type of entertainment they prefer to determine the types of entertainment adults in Westbury prefer

j) Asking 100 randomly selected people at a rap concert in New York what their favorite type of music to determine what New Yorker's favorite type of music is.

k) Asking 50 randomly selected people at a Dominican Restaurant in Westbury what their favorite type of food is to determine what people in Westbury's favorite types of food are.

l) Posting a phone number to call on the back of a truck to comment on what drivers think of the truck driver's driving

m) Asking 75 randomly selected students in the Westbury High School 9^{th} grade cafeteria what their favorite type of music is to determine the favorite type of music of Westbury High School students.

n) Asking 40 randomly selected Westbury High School Seniors where they plan to go to college to determine where Westbury High School Seniors plan to go to college.

3. Which statement(s) about statistical studies is true?

I. A survey of all English classes in a high school would be a good sample to determine the number of hours students throughout the school spend studying.

II. A survey of all ninth graders in a high school would be a good sample to determine the number of student parking spaces needed at that high school.

III. A survey of all students in one lunch period in a high school would be a good sample to determine the number of hours adults spend on social media websites.

IV. A survey of all Calculus students in a high school would be a good sample to determine the number of students throughout the school who don't like math.

1) I, only 2) II, only 3) I and III 4) III and IV

4. Which survey is *least* likely to contain bias?

1) surveying a sample of people leaving a movie theater to determine which flavor of ice cream is the most popular
2) surveying the members of a football team to determine the most watched TV sport
3) surveying a sample of people leaving a library to determine the average number of books a person reads in a year
4) surveying a sample of people leaving a gym to determine the average number of hours a person exercises per week

5. A survey is to be conducted in a small upstate village to determine whether or not local residents should fund construction of a skateboard park by raising taxes. Which segment of the population would provide the most unbiased responses?

1) a club of local skateboard enthusiasts
2) senior citizens living on fixed incomes
3) a group opposed to any increase in taxes
4) every tenth person 18 years of age or older walking down Main St.

7. A survey is being conducted about American's favorite musicians. Which of the following survey methods would most likely produce a random sample?

(1) Asking every 20^{th} person at a Green Day concert
(2) Asking every 10^{th} person at a vintage record store
(3) Asking every 10^{th} person at the Westbury Public Library
(4) Sending out surveys to random households across the country.

8. Which method of collecting data would most likely result in an unbiased random sample?
(1) selecting every third teenager leaving a movie theater to answer a survey about entertainment
(2) placing a survey in a local newspaper to determine how people voted in the 2004 presidential election
(3) selecting students by the last digit of their school ID number to participate in a survey about cafeteria food
(4) surveying honor students taking Trigonometry to determine the average amount of time students in a school spend doing homework each night

9. A survey completed at a large university asked 2,000 students to estimate the average number of hours they spend studying each week. Every tenth student entering the library was surveyed. The data showed that the mean number of hours that students spend studying was 15.7 per week. Which characteristic of the survey could create a bias in the results?
(1) the size of the sample (3) the method of analyzing the data
(2) the size of the population (4) the method of choosing the students who were surveyed

10. The yearbook staff has designed a survey to learn about student opinions on how the yearbook could be improved for this year. If they want to distribute this survey to 100 students and obtain the most reliable data, they should survey
(1) Every third student sent to the office
(2) Every third student to enter the library
(3) Every third student to enter the gym for the basketball game
(4) Every third student arriving at school in the morning

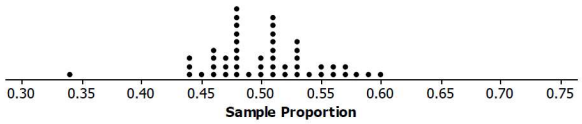
11. Chuck's Trucking Company has decided to initiate an Employee of the Month program. To determine the recipient, they put the following sign on the back of each truck.
The driver who receives the highest number of positive comments will win the recognition.
Explain *one* statistical bias in this data collection method.

Name _____ Date _____
Mr. Schlansky Algebra II

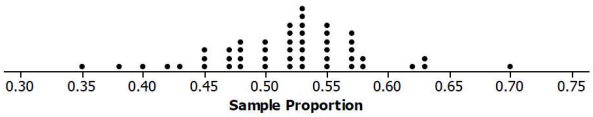
Sample Distributions and Sample Size

1. Below are three dot plots of the proportion of tails in 20, 60, or 120 simulated flips of a coin. The mean and standard deviation of the sample proportions are also shown for each of the three dot plots. Match each dot plot with the appropriate number of flips. Clearly explain how you matched the plots with the number of simulated flips.

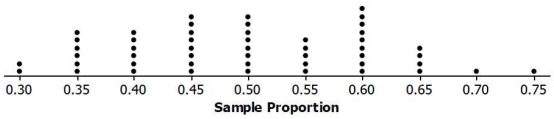
Dot Plot 1:



Dot Plot 2

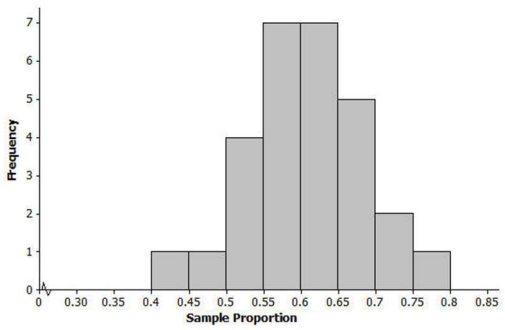


Dot Plot 3

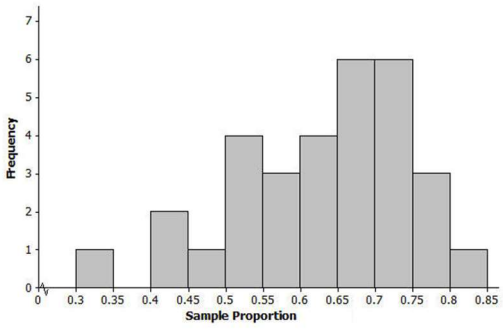


2. A group of eleventh graders wanted to estimate the proportion of all students at their high school who suffer from allergies. Each student in one group of eleventh graders took a random sample of 20 students, while each student in another group of eleventh graders each took a random sample of 40 students. Below are the two sampling distributions (shown as histograms) of the sample proportions of high school students who said that they suffer from allergies. Which histogram is based on random samples of size 40? Explain.

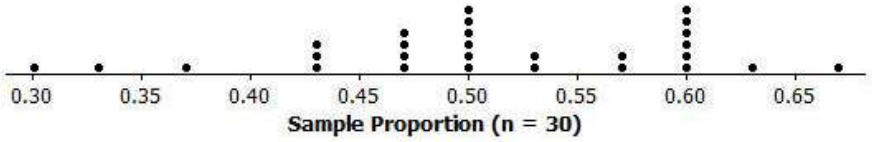
Histogram A



Histogram B

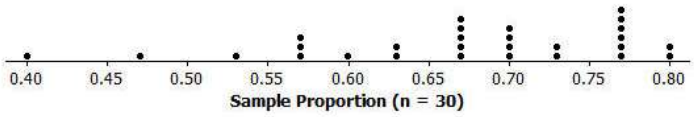


3. A group of eleventh graders wanted to estimate the population proportion of students in their high school who drink at least one soda per day. Each student selected a different random sample of 30 students and calculated the proportion that drink at least one soda per day. The dot plot below shows the sampling distribution. This distribution has a mean of 0.51 and a standard deviation of 0.09.



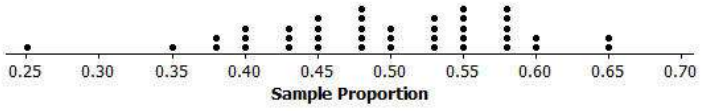
If, instead of taking random samples of 30 students in the high school, the eleventh graders randomly selected samples of size 60, describe what will happen to the mean and standard deviation of the sampling distribution of the sample proportions.

4. A class of 28 eleventh graders wanted to estimate the proportion of all juniors and seniors at their high school with part-time jobs after school. Each eleventh grader took a random sample of 30 juniors and seniors and then calculated the proportion with part-time jobs. A dot plot is created to represent the data. The mean is 0.67 and the standard deviation is 0.1.



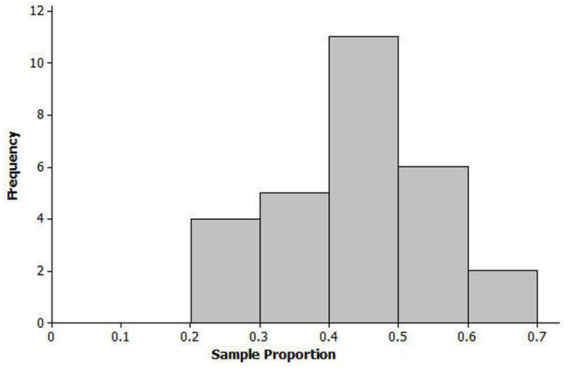
Suppose the eleventh graders had taken random samples of size 60. How would the distribution of sample proportions based on samples of size 60 compare to the distribution for samples of size 30?

5. The following is an example of a sampling distribution of sample proportions of heads in **40** flips of a coin. The mean is .4955 and the sample standard deviation is .0852.



If this experiment was performed where the coin was flipped 20 times, how would the data be affected?

6. The nurse in your school district would like to study the proportion of all high school students in the district who usually get at least eight hours of sleep on school nights. Suppose each student in your class takes a random sample of 20 high school students in the district and each calculates their sample proportion of students who said that they usually get at least eight hours of sleep on school nights. Below is a histogram of the sampling distribution.

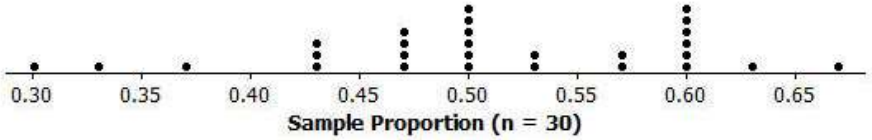


Suppose students had taken random samples of size 60. How would the distribution of sample proportions based on samples of size 60 differ from those of size 20?

Name _____ Date _____
Mr. Schlansky Algebra II

Confidence Interval and Margin of Error

1. A group of eleventh graders wanted to estimate the population proportion of students in their high school who drink at least one soda per day. Each student selected a different random sample of 30 students and calculated the proportion that drink at least one soda per day. The dot plot below shows the sampling distribution. This distribution has a mean of 0.51 and a standard deviation of 0.09.



Find the margin of error and the confidence interval.

Is it reasonable to say that 40% of students drink at least one soda per day? Explain your answer.

Is it reasonable to say that 80% of students drink at least one soda per day? Explain your answer.

2. A class of 28 eleventh graders wanted to estimate the proportion of all juniors and seniors at their high school with part-time jobs after school. Each eleventh grader took a random sample of 30 juniors and seniors and then calculated the proportion with part-time jobs. The mean is 0.67 and the standard deviation is 0.1.

Find the margin of error and the confidence interval.

Do you think that the proportion of all juniors and seniors at the school with part-time jobs could be 0.7? Do you think it could be 0.5? Justify your answers.

3. The following is an example of a sampling distribution of sample proportions of heads in **40** flips of a coin. The mean is .4955 and the sample standard deviation is .0852.

Find the margin of error and the confidence interval.

Fred flipped a coin 40 times and 65% of the flips came up heads. Is this an expected outcome? Explain your answer.

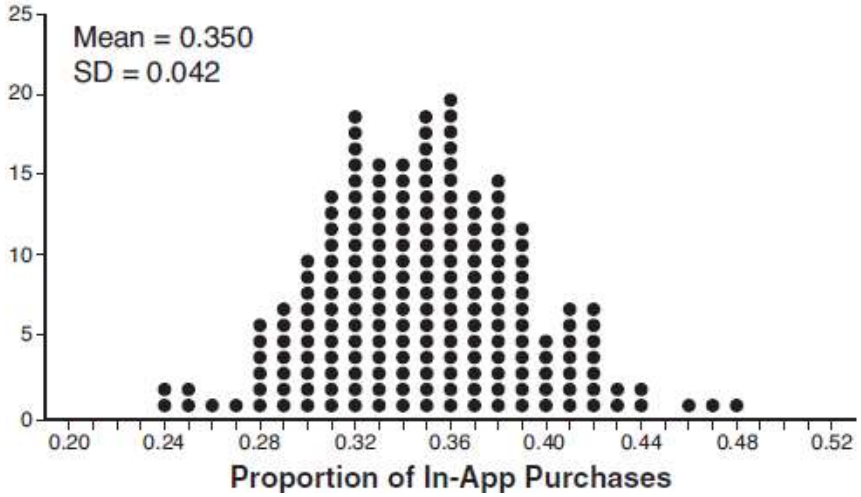
Fred flipped a coin 40 times and 75% of the flips came up heads. Is this an expected outcome? Explain your answer.

4. Elizabeth waited for 6 minutes at the drive thru at her favorite fast-food restaurant the last time she visited. She was upset about having to wait that long and notified the manager. The manager assured her that her experience was very unusual and that it would not happen again. A study of customers commissioned by this restaurant found an approximately normal distribution of results. The mean wait time was 226 seconds and the standard deviation was 38 seconds. Given these data, and using a 95% level of confidence, was Elizabeth's wait time unusual? Justify your answer.

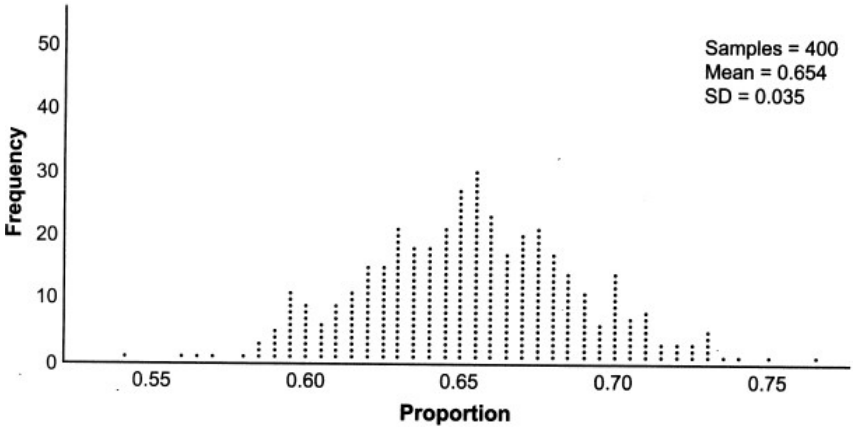
5. Jessica got 20 math problems for homework and complained to her teacher that this was an unusual amount of homework. Her teacher told her to look at the number of questions in all of her past homework assignments from the school year and find the range of the expected number of math problems. She found that the mean was 11.2 and the standard deviation was 3. Was Jessica correct that 20 math problems was unusual? Justify your answer.

6. Fatima bought a chicken burrito for dinner and was unhappy with the amount of chicken that she received. She received 4.75 ounces of chicken and believed that this was less than normal. The manager conducted a study and found that the mean amount of chicken on their burritos was 5.1 ounces with a standard deviation of .25 ounces. Did Fatima's burrito have an expected amount of chicken? Justify your answer.

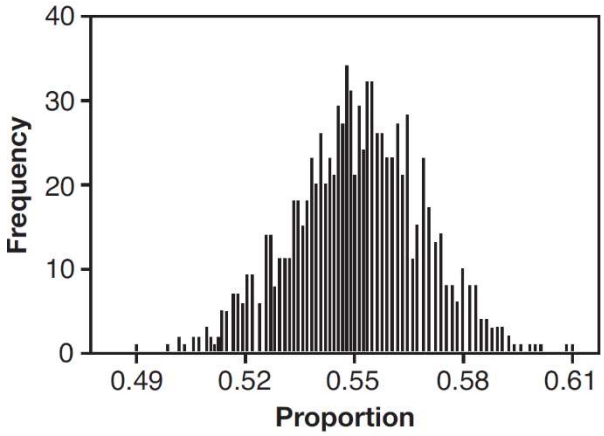
7. Some smart-phone applications contain "in-app" purchases, which allow users to purchase special content within the application. A random sample of 140 users found that 35 percent made in-app purchases. A simulation was conducted with 200 samples of 140 users assuming 35 percent of the samples make in-app purchases. The approximately normal results are shown below. Considering the middle 95% of the data, determine the margin of error, to the *nearest hundredth*, for the simulated results.



8. Betty conducted a survey of her class to see if they like pizza. She gathered 200 responses and 65% of the voters said they did like pizza. Betty then ran a simulation of 400 more surveys, each with 200 responses, assuming that 65% of the voters would like pizza. The output of the simulation is shown below. Considering the middle 95% of the data, what is the margin of error for the simulation?



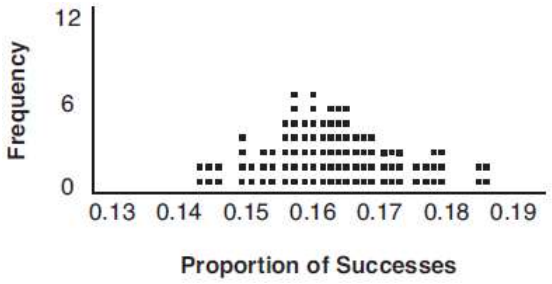
9. A candidate for political office commissioned a poll. His staff received responses from 900 likely voters and 55% of them said they would vote for the candidate. The staff then conducted a simulation of 1000 more polls of 900 voters, assuming that 55% of voters would vote for their candidate. The output of the simulation is shown in the diagram below.



Given this output, and assuming a 95% confidence level, the margin of error for the poll is closest to

1) 0.01
2) 0.03
3) 0.06
4) 0.12

10. A study conducted in 2004 in New York City found that 212 out of 1334 participants had hypertension. Kim ran a simulation of 100 studies based on these data. The output of the simulation is shown in the diagram below.



At a 95% confidence level, the proportion of New York City residents with hypertension and the margin of error are closest to

1) proportion ≈ .16; margin of error ≈ .01
2) proportion ≈ .16; margin of error ≈ .02
3) proportion ≈ .01; margin of error ≈ .16
4) proportion ≈ .02; margin of error ≈ .16

Name _____ Date _____

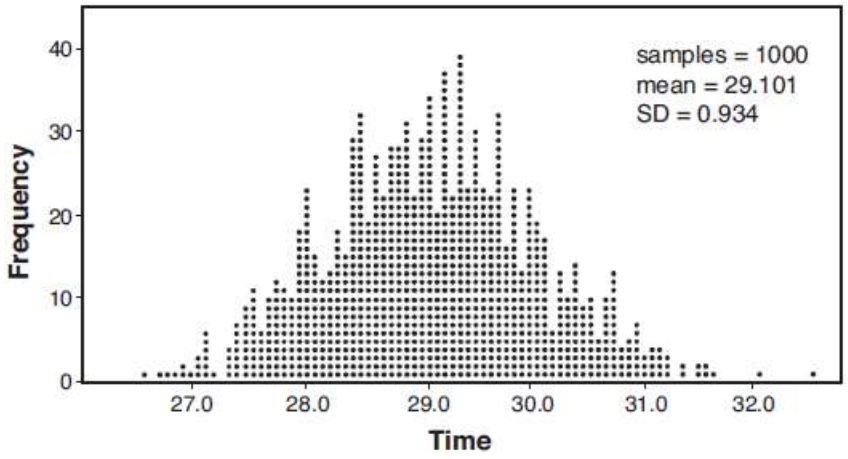
Mr. Schlansky Algebra II

Confidence Interval (Part IIIs)

1. A radio station claims to its advertisers that the mean number of minutes commuters listen to the station is 30. The station conducted a survey of 500 of their listeners who commute. The sample statistics are shown below.

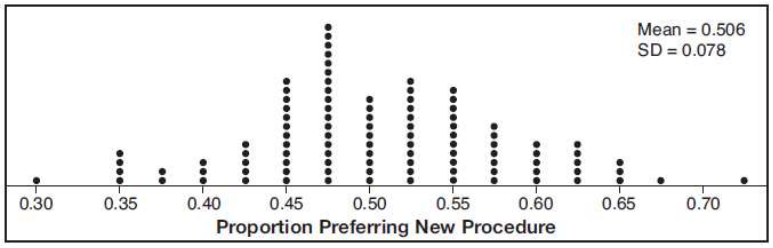
$\bar{x}$	29.11
s_x	20.718

A simulation was run 1000 times based upon the results of the survey. The results of the simulation appear below.



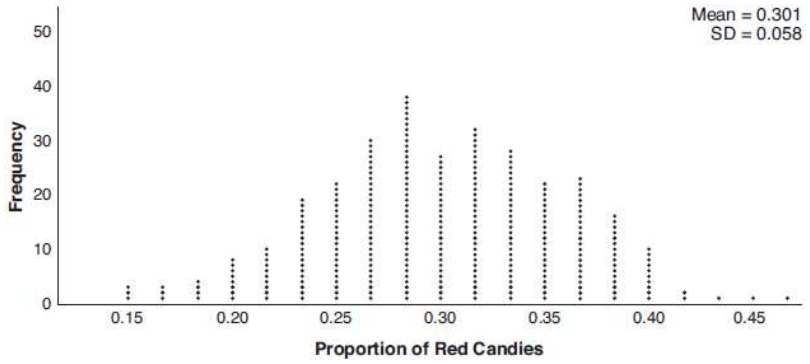
Based on the simulation results, is the claim that commuters listen to the station on average 30 minutes plausible? Explain your response including an interval containing the middle 95% of the data, rounded to the *nearest hundredth*.

2. Charlie's Automotive Dealership is considering implementing a new check-in procedure for customers who are bringing their vehicles for routine maintenance. The dealership will launch the procedure if 50% or more of the customers give the new procedure a favorable rating when compared to the current procedure. The dealership devises a simulation based on the minimal requirement that 50% of the customers prefer the new procedure. Each dot on the graph below represents the proportion of the customers who preferred the new check-in procedure, each of sample size 40, simulated 100 times.



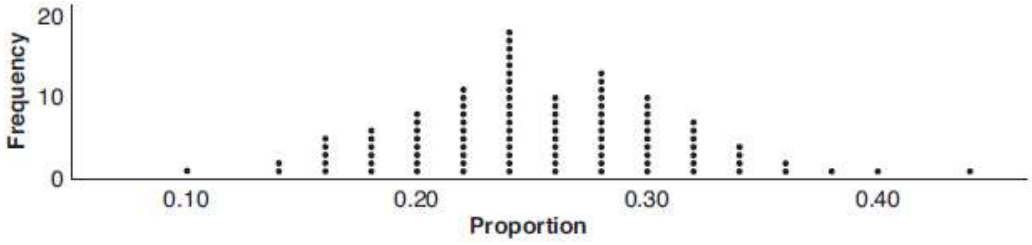
Assume the set of data is approximately normal and the dealership wants to be 95% confident of its results. Determine an interval containing the plausible sample values for which the dealership will launch the new procedure. Round your answer to the *nearest hundredth*. Forty customers are selected randomly to undergo the new check-in procedure and the proportion of customers who prefer the new procedure is 32.5%. The dealership decides *not* to implement the new check-in procedure based on the results of the study. Use statistical evidence to explain this decision.

3. Mary bought a pack of candy. The manufacturer claims that 30% of the candies manufactured are red. In her pack, 14 of the 60 candies are red. She ran a simulation of 300 samples, assuming the manufacturer is correct. The results are shown below.



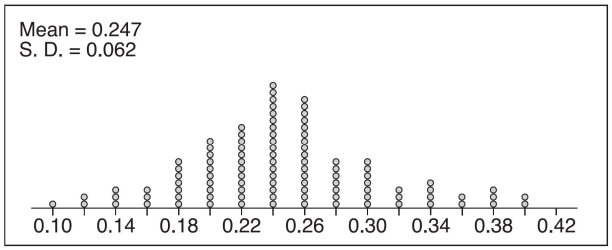
Based on the simulation, determine the middle 95% of plausible values that the proportion of red candies in a pack is within. Based on the simulation, is it unusual that Mary's pack had 14 red candies out of a total of 60? Explain.

4. A group of students was trying to determine the proportion of candies in a bag that are blue. The company claims that 24% of candies in bags are blue. A simulation was run 100 times with a sample size of 50, based on the premise that 24% of the candies are blue. The approximately normal results of the simulation are shown in the dot plot below.



The simulation results in a mean of 0.254 and a standard deviation of 0.060. Based on this simulation, what is a plausible interval containing the middle 95% of the data? A student found that 18 out of 50 of the candies were blue. Use statistical evidence to explain why this is an expected value.

5. Stephen's Beverage Company is considering whether to produce a new brand of cola. The company will launch the product if at least 25% of cola drinkers will buy the product. Fifty cola drinkers are randomly selected to take a blind taste-test of products *A*, *B*, and the new product. Nine out of fifty participants preferred Stephen's new cola to products *A* and *B*. The company then devised a simulation based on the requirement that 25% of cola drinkers will buy the product. Each dot in the graph shown below represents the proportion of people who preferred Stephen's new product, each of sample size 50, simulated 100 times.

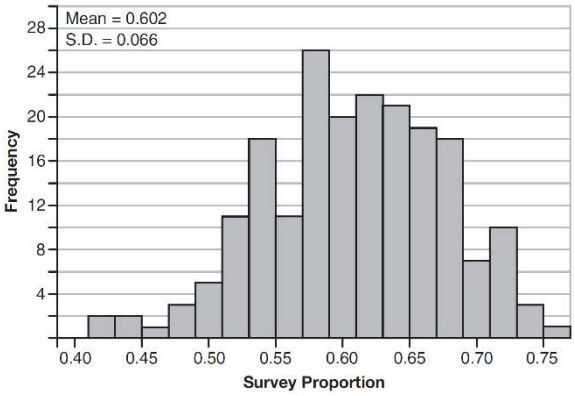


Proportion Preferring Stephen's Product

Assume the set of data is approximately normal and the company wants to be 95% confident of its results. Does the sample proportion obtained from the blind taste-test, nine out of fifty, fall within the margin of error developed from the simulation? Justify your answer. The company decides to continue developing the product even though only nine out of fifty participants preferred its brand of cola in the taste-test. Describe how the simulation data could be used to support this decision.

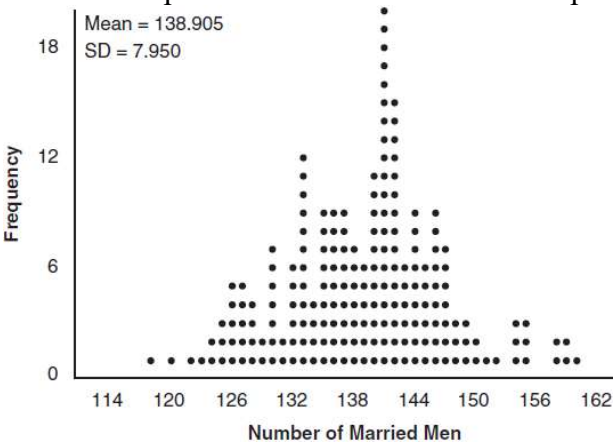
6. Fifty-five students attending the prom were randomly selected to participate in a survey about the music choice at the prom. Sixty percent responded that a DJ would be preferred over a band. Members of the prom committee thought that the vote would have 50% for the DJ and 50% for the band. A simulation was run 200 times, each of sample size 55, based on the premise that 60% of the students would prefer a DJ. The approximate normal simulation results are shown below.

Using the results of the simulation, determine a plausible interval containing the middle 95% of the data. Round all values to the *nearest hundredth*. Members of the prom committee are concerned that a vote of all students attending the prom may produce a 50%-50% split. Explain what statistical evidence supports this concern.



7. In a random sample of 250 men in the United States, age 21 or older, 139 are married. The graph below simulated samples of 250 men, 200 times, assuming that 139 of the men are married.

a) Based on the simulation, create an interval in which the middle 95% of the number of married men may fall. Round your answer to the *nearest integer*.
b) A study claims "50 percent of men 21 and older in the United States are married." Do your results from part a contradict this claim? Explain.



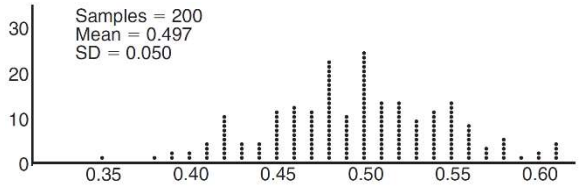
Name _____ Date _____
Mr. Schlansky Algebra II

Confidence Interval (Fair)

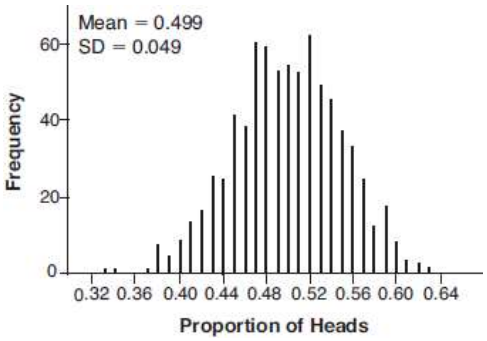
1. Anne has a coin. She does not know if it is a fair coin. She flipped the coin 100 times and obtained 73 heads and 27 tails. She ran a computer simulation of 200 samples of 100 fair coin flips. The output of the proportion of heads is shown below.

Given the results of her coin flips and of her computer simulation, which statement is most accurate?

1) 73 of the computer's next 100 coin flips will be heads.
2) 50 of her next 100 coin flips will be heads.
3) Her coin is not fair.
4) Her coin is fair.

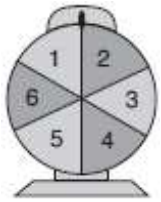


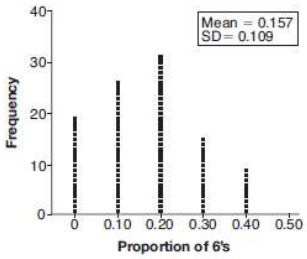
2. Robin flips a coin 100 times. It lands heads up 43 times, and she wonders if the coin is unfair. She runs a computer simulation of 750 samples of 100 fair coin flips. The output of the proportion of heads is shown below. Do the results of the simulation provide strong evidence that Robin's coin is unfair? Explain your answer.



3. Juanita rolls a 6 sided die and recorded that it landed on 6 five times out of 50. She questioned whether the die was fair so she ran a computer simulation of 1000 samples of 50 rolls of a fair die. The mean of the simulation was .159 with a standard deviation of .102. Is her die fair? Explain your answer.

4. A game spinner is divided into 6 equally sized regions, as shown in the diagram below. For Miles to win, the spinner must land on the number 6. After spinning the spinner 10 times, and losing all 10 times, Miles complained that the spinner is unfair. At home, his dad ran 100 simulations of spinning the spinner 10 times, assuming the probability of winning each spin is $\frac{1}{6}$. The output of the simulation is shown in the diagram below.





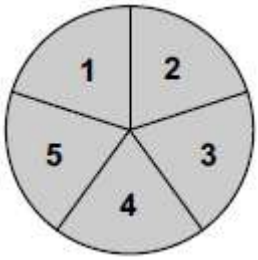
Is there strong evidence to suggest that the spinner is unfair? Explain your answer.

5. A spinner with 8 sectors labeled A, B, C, D, E, F, G, H was spun 100 times. The spinner landed on sector B 20 times out of 100. A computer simulation of 500 samples of 100 spins of a fair 8 sector spinner was run and it was found that the mean proportion of landing on sector B was .126 with a standard deviation of .027. Is the spinner fair? Explain your answer.

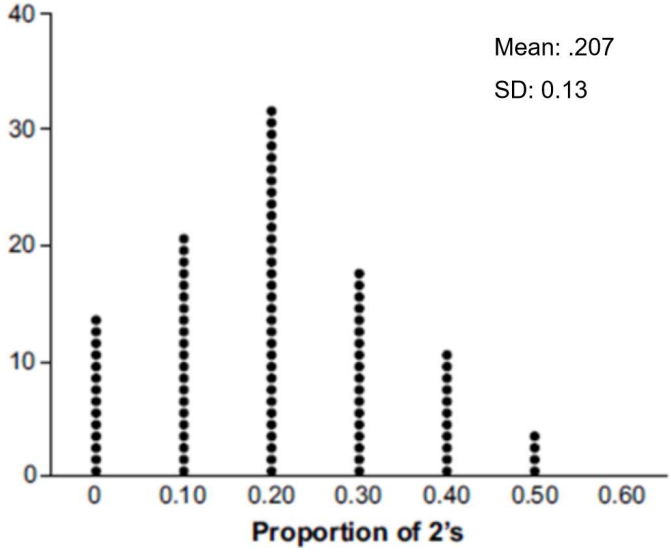
6. Ally flipped a coin 100 times a got a proportion of .41 heads. She believed this coin was unfair so she ran a computer simulation of 200 samples of 100 coin flips of a fair coin. The mean of the simulation was .502 and the standard deviation was .024. Is Ally's coin fair? Explain your answer.

7. A spinner below is spun and it landed on the number "2" 3 times out of 50. A computer simulation of 500 samples of 50 spins of a fair spinner was spun. The mean of the simulation was .128 and the standard deviation was .07. Is the spinner fair? Explain your answer.

8. Joette is playing a carnival game. To win a prize, one has to correctly guess which of five equally sized regions a spinner will land on, as shown in the diagram below.



She complains that the game is unfair because her favorite number, 2, has only been spun once in ten times she played the game. State the proportion of 2's that were spun. State the theoretical probability of spinning a 2. The simulation output below shows the results of simulating ten spins of a fair spinner, repeated 100 times.



Does the output indicate that the carnival game was unfair? Explain your answer.

Name _____ Date _____
Mr. Schlansky Algebra II

Designing Controlled Experiments

1. A high school wants to determine if taking a Regents Review course causes students to score higher on the Common Core Geometry Regents Exam. Describe a controlled experiment that can be created to test this hypothesis.

2. A company created a pill that is supposed to reduce the length of the common cold. Describe a controlled experiment to test this hypothesis.

3. A farm is experimenting with a drug to increase the size of its chickens. Describe a controlled experiment that can be performed to see if this drug is effective.

4. Describe how a controlled experiment can be created to examine the effect of ingredient X in a toothpaste.

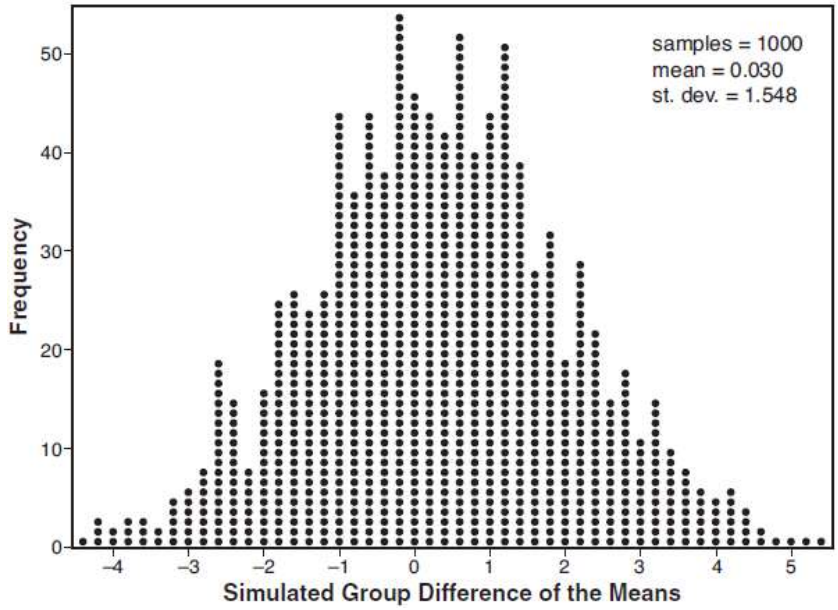
Name _____ Date _____
Mr. Schlansky Algebra II

Evaluating Effectiveness of Treatments (Mean Differences)

1. Joseph was curious to determine if scent improves memory. A test was created where better memory is indicated by higher test scores. A controlled experiment was performed where one group was given the test on scented paper and the other group was given the test on unscented paper. The summary statistics from the experiment are given below.

	Scented Paper	**Unscented Paper**
$\bar{x}$	23	18
s_x	2.898	2.408

Calculate the difference in means in the experimental test grades (scented -unscented). A simulation was conducted in which the subjects' scores were rerandomized into two groups 1000 times. The differences of the group means were calculated each time. The results are shown below.

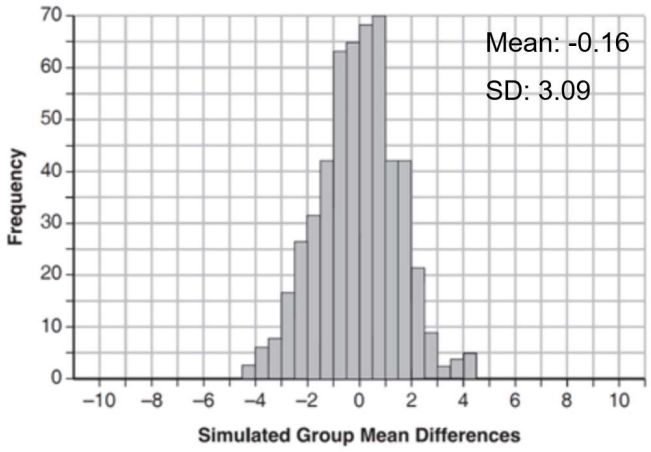


Use the simulation results to determine the interval representing the middle 95% of the difference in means, to the *nearest hundredth*. Is the difference in means in Joseph's experiment statistically significant based on the simulation? Explain.

2. Seventy-two students are randomly divided into two equally-sized study groups. Each member of the first group (group 1) is to meet with a tutor after school twice each week for one hour. The second group (group 2), is given an online subscription to a tutorial account that they can access for a maximum of two hours each week. Students in both groups are given the same tests during the year. A summary of the two groups' final grades is shown below:

	Group 1	Group 2
$\bar{x}$	80.16	83.8
S_x	6.9	5.2

Calculate the mean difference in the final grades (group 1 – group 2) and explain its meaning in the context of the problem. A simulation was conducted in which the students' final grades were rerandomized 500 times. The results are shown below.



Use the simulation to determine if there is a significant difference in the final grades. Explain your answer.

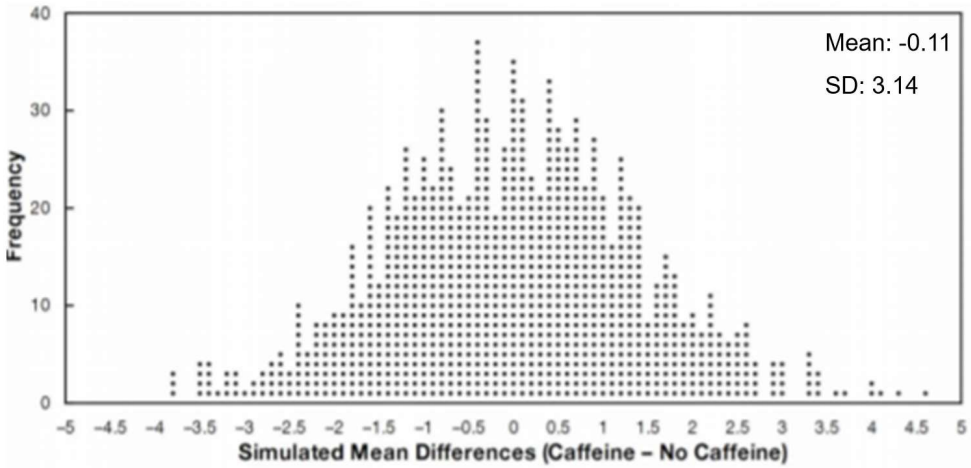
3.

The effects of caffeine on the body have been extensively studied. In one experiment, researchers trained a sample of male college students to tap their fingers at a rapid rate. The sample was then divided at random into two groups of 10 students each. Each student drank the equivalent of about two cups of coffee, which included about 200 mg of caffeine for the students in one group but was decaffeinated coffee for the second group. After a 2-hour period, each student was tested to measure finger tapping rate (taps per minute). The students did not know whether or not their drinks included caffeine and the person measuring the tap rates was also unaware of the groups. The finger-tapping rates measured in this experiment are summarized in the table below.

											Mean
Caffeine	246	248	250	252	248	250	246	248	245	250	248.3
No Caffeine	242	245	244	248	247	248	242	244	246	242	244.8

Calculate the mean difference (Caffeine – No Caffeine) and interpret your answer in the context of the problem.

The researchers then took the twenty finger-tapping rates and rerandomized them 1,000 times using simulation software. The output of the simulation results is shown in the dotplot below.

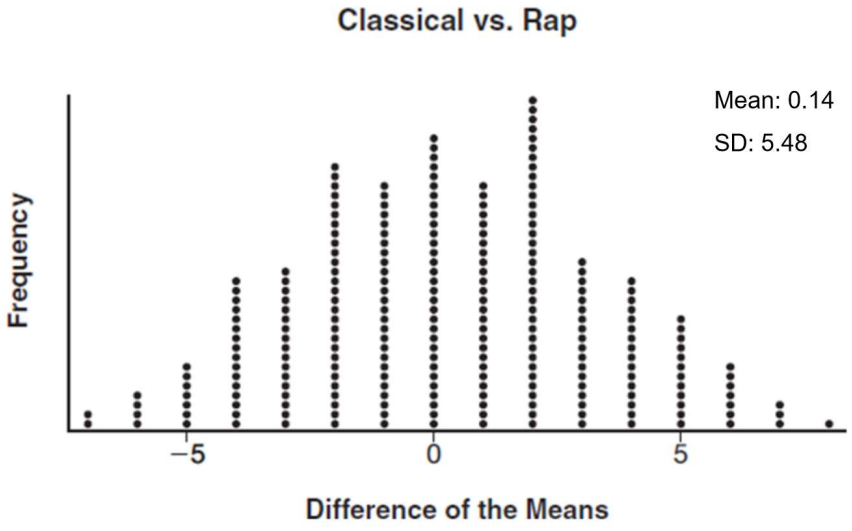


Does the simulation data support the conclusion that caffeine causes an increase in average finger-tapping rate? Justify your answer.

4. To determine if the type of music played while taking a quiz has a relationship to results, 16 students were randomly assigned to either a room softly playing classical music or a room softly playing rap music. The results on the quiz were as follows:

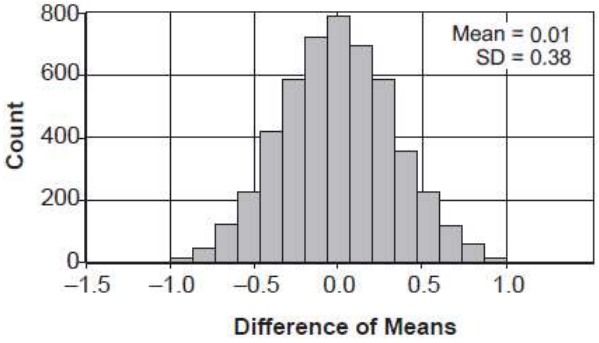
Classical: 74, 83, 77, 77, 84, 82, 90, 89
Rap: 77, 80, 78, 74, 69, 72, 78, 69

John correctly rounded the difference of the means of his experimental groups as 7. How did John obtain this value and what does it represent in the given context? Justify your answer. To determine if there is any significance in this value, John rerandomized the 16 scores into two groups of 8, calculated the difference of the means, and simulated this process 250 times as shown below.

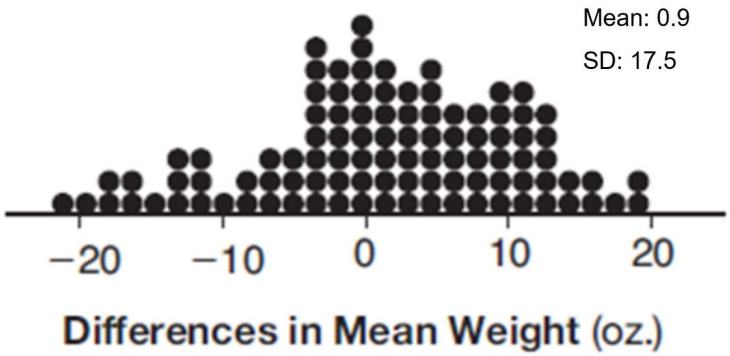


Does the simulation support the theory that there may be a significant difference in quiz scores? Explain.

5. Two classes of students were entered into an experiment to see whether using an interactive whiteboard leads to better grades. It was observed that the mean grade of students in the class with the interactive whiteboard was 0.6 points higher than the class without it. To determine if the observed difference is statistically significant, the classes were rerandomized 5000 times to study these random differences in the mean grades. The output of the simulation is summarized in the histogram below. Determine an interval containing the middle 95% of the simulation results. Round your answer to the *nearest hundredth*. Does the interval indicate that the difference between the classes' grades is significant? Explain.



6. Gabriel performed an experiment to see if planting 13 tomato plants in black plastic mulch leads to larger tomatoes than if 13 plants are planted without mulch. He observed that the average weight of the tomatoes from tomato plants grown in black plastic mulch was 5 ounces greater than those from the plants planted without mulch. To determine if the observed difference is statistically significant, he rerandomized the tomato groups 100 times to study these random differences in the mean weights. The output of his simulation is summarized in the dotplot below. Do you believe that planting in black plastic mulch causes larger tomato size? Explain your answer.



Name _____ Date _____
Mr. Schlansky Algebra II

Statistics Review Sheet

1. A doctor wants to test the effectiveness of a new drug on her patients. She separates her sample of patients into two groups and administers the drug to only one of these groups. She then compares the results. Which type of study *best* describes this situation?

1) census
2) survey
3) observation
4) controlled experiment

2. A market research firm needs to collect data on viewer preferences for local news programming in Buffalo. Which method of data collection is most appropriate?

1) census
2) survey
3) observation
4) controlled experiment

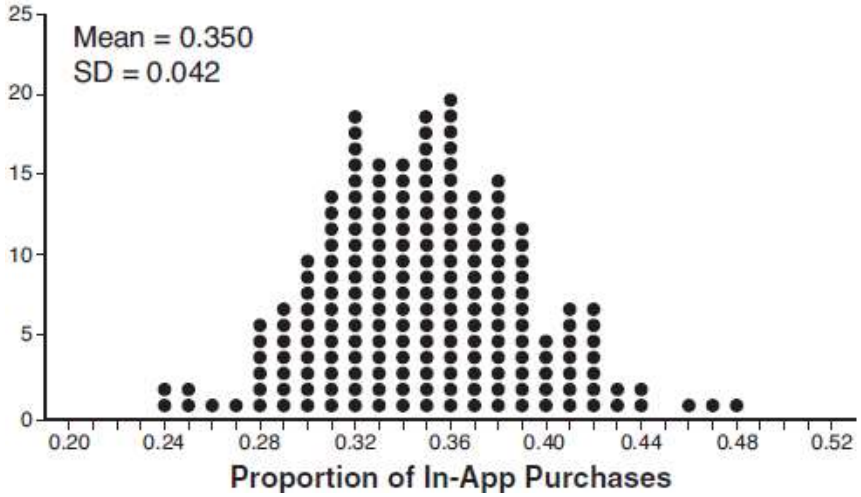
3. Which survey is *least* likely to contain bias?

1) surveying a sample of people leaving a movie theater to determine which flavor of ice cream is the most popular
2) surveying the members of a football team to determine the most watched TV sport
3) surveying a sample of people leaving a library to determine the average number of books a person reads in a year
4) surveying a sample of people leaving a gym to determine the average number of hours a person exercises per week

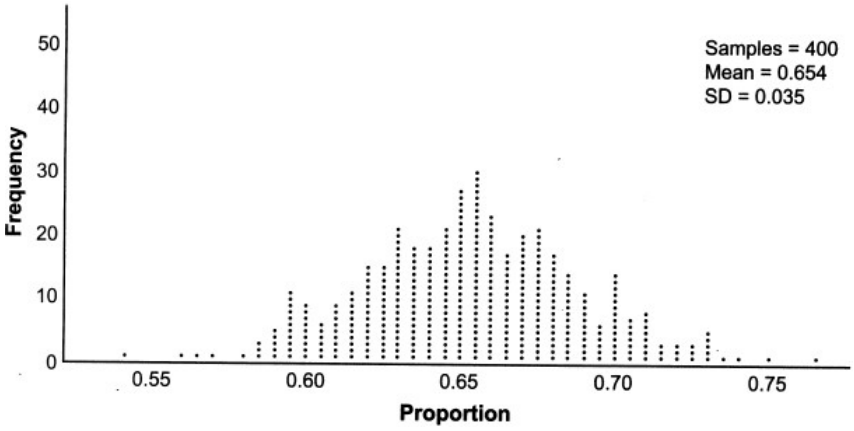
4. A survey is to be conducted in a small upstate village to determine whether or not local residents should fund construction of a skateboard park by raising taxes. Which segment of the population would provide the most unbiased responses?

1) a club of local skateboard enthusiasts
2) senior citizens living on fixed incomes
3) a group opposed to any increase in taxes
4) every tenth person 18 years of age or older walking down Main St.

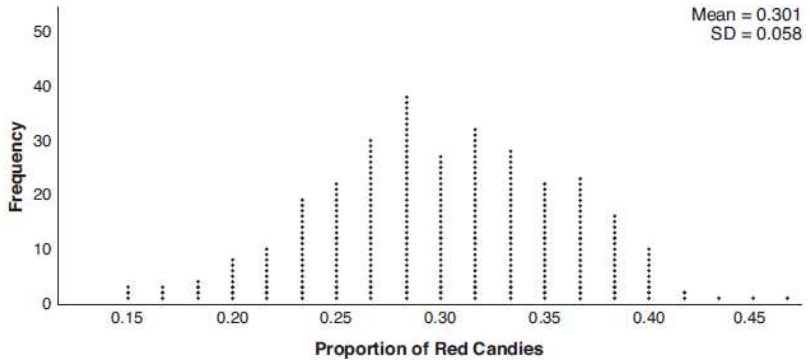
5. Some smart-phone applications contain "in-app" purchases, which allow users to purchase special content within the application. A random sample of 140 users found that 35 percent made in-app purchases. A simulation was conducted with 200 samples of 140 users assuming 35 percent of the samples make in-app purchases. The approximately normal results are shown below. Considering the middle 95% of the data, determine the margin of error, to the *nearest hundredth*, for the simulated results.



6. Betty conducted a survey of her class to see if they like pizza. She gathered 200 responses and 65% of the voters said they did like pizza. Betty then ran a simulation of 400 more surveys, each with 200 responses, assuming that 65% of the voters would like pizza. The output of the simulation is shown below. Considering the middle 95% of the data, what is the margin of error for the simulation?

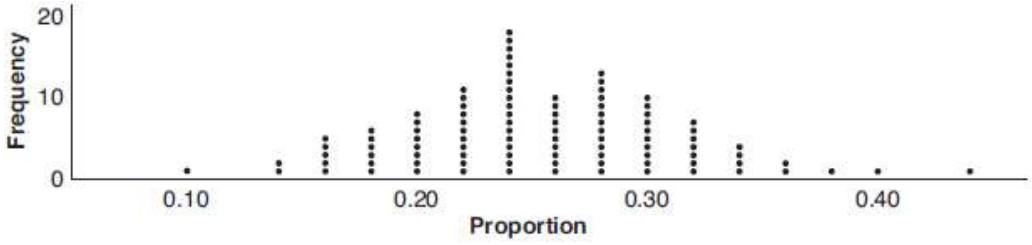


7. Mary bought a pack of candy. The manufacturer claims that 30% of the candies manufactured are red. In her pack, 14 of the 60 candies are red. She ran a simulation of 300 samples, assuming the manufacturer is correct. The results are shown below.



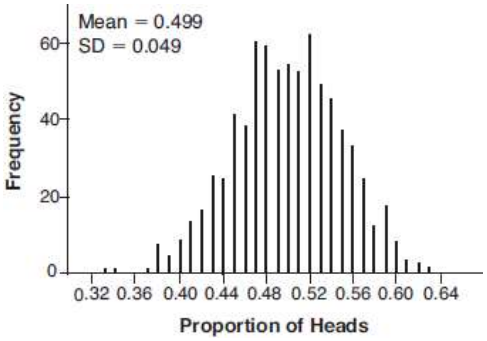
Based on the simulation, determine the middle 95% of plausible values that the proportion of red candies in a pack is within. Based on the simulation, is it unusual that Mary's pack had 14 red candies out of a total of 60? Explain.

8. A group of students was trying to determine the proportion of candies in a bag that are blue. The company claims that 24% of candies in bags are blue. A simulation was run 100 times with a sample size of 50, based on the premise that 24% of the candies are blue. The approximately normal results of the simulation are shown in the dot plot below.



The simulation results in a mean of 0.254 and a standard deviation of 0.060. Based on this simulation, what is a plausible interval containing the middle 95% of the data? A student found that 18 out of 50 of the candies were blue. Use statistical evidence to explain why this is an expected value.

9. Robin flips a coin 100 times. It lands heads up 43 times, and she wonders if the coin is unfair. She runs a computer simulation of 750 samples of 100 fair coin flips. The output of the proportion of heads is shown below. Do the results of the simulation provide strong evidence that Robin's coin is unfair? Explain your answer.

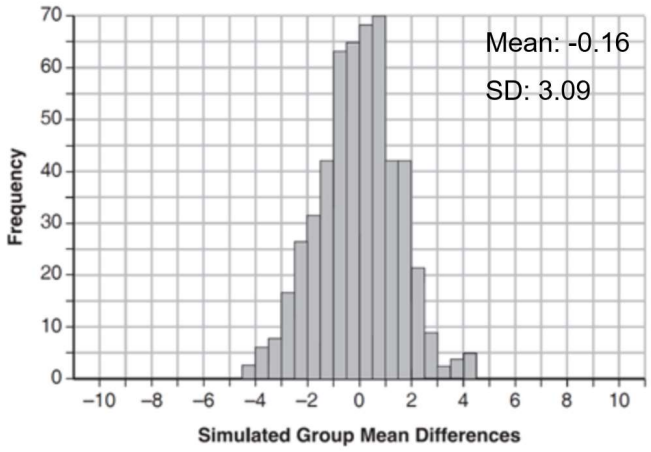


10. Juanita rolls a 6 sided die and recorded that it landed on 6 five times out of 50. She questioned whether the die was fair so she ran a computer simulation of 1000 samples of 50 rolls of a fair die. The mean of the simulation was .159 with a standard deviation of .102. Is her die fair? Explain your answer.

11. Seventy-two students are randomly divided into two equally-sized study groups. Each member of the first group (group 1) is to meet with a tutor after school twice each week for one hour. The second group (group 2), is given an online subscription to a tutorial account that they can access for a maximum of two hours each week. Students in both groups are given the same tests during the year. A summary of the two groups' final grades is shown below:

	Group 1	Group 2
$\bar{x}$	80.16	83.8
S_x	6.9	5.2

Calculate the mean difference in the final grades (group 1 – group 2) and explain its meaning in the context of the problem. A simulation was conducted in which the students' final grades were rerandomized 500 times. The results are shown below.



Use the simulation to determine if there is a significant difference in the final grades. Explain your answer.

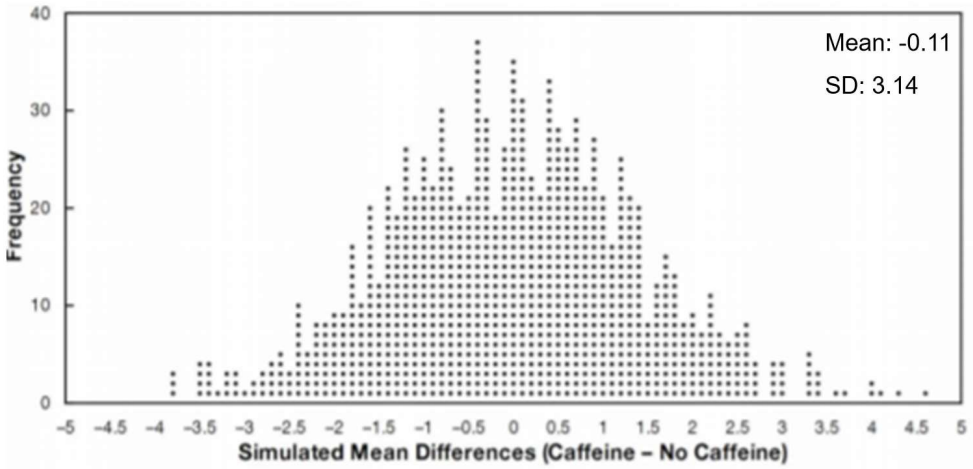
12.

The effects of caffeine on the body have been extensively studied. In one experiment, researchers trained a sample of male college students to tap their fingers at a rapid rate. The sample was then divided at random into two groups of 10 students each. Each student drank the equivalent of about two cups of coffee, which included about 200 mg of caffeine for the students in one group but was decaffeinated coffee for the second group. After a 2-hour period, each student was tested to measure finger tapping rate (taps per minute). The students did not know whether or not their drinks included caffeine and the person measuring the tap rates was also unaware of the groups. The finger-tapping rates measured in this experiment are summarized in the table below.

											Mean
Caffeine	246	248	250	252	248	250	246	248	245	250	248.3
No Caffeine	242	245	244	248	247	248	242	244	246	242	244.8

Calculate the mean difference (Caffeine – No Caffeine) and interpret your answer in the context of the problem.

The researchers then took the twenty finger-tapping rates and rerandomized them 1,000 times using simulation software. The output of the simulation results is shown in the dotplot below.



Does the simulation data support the conclusion that caffeine causes an increase in average finger-tapping rate? Justify your answer.