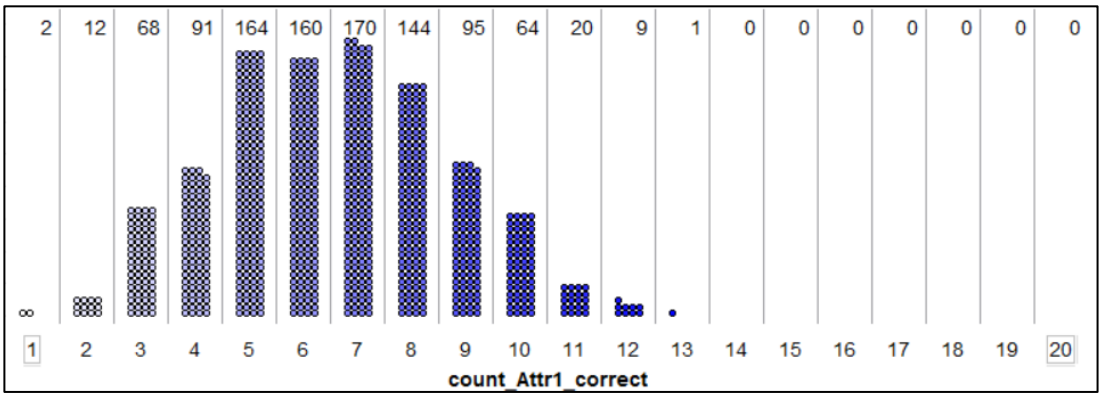
STAT 110 **Name:** _____
Exam 1 (100 PTS)
Fall 2017

True/False Questions: Answer the following true/false questions. (2 pts each)

Questions 1 - 7 concern the following scenario. Suppose a subject claims to suffer from severe short-term memory loss. To investigate this, a researcher randomly selects and shows the subject one of three objects: a blue marker, a green highlighter, or a yellow pencil. She then takes the object away and waits 30 seconds. At that point, she asks the subject to recall which of the objects they were shown. This process is repeated a total of 20 times, and the subject is correct on 4 of the 20 trials.

The researcher believes the subject is exaggerating their memory loss and is intentionally giving wrong answers, and she wants to find statistical evidence that this is the case. The plot below shows 1,000 simulated outcomes which were obtained to investigate her research hypothesis. In each of the 1,000 runs of the simulation, she kept track of and plotted the number of correct answers given (out of 20).



1.	The above outcomes were simulated assuming the subject truly suffered from severe memory loss and was therefore guessing on each trial.	(TRUE) FALSE
2.	Let π = the true long-run probability that the subject answers correctly. The null and alternative hypotheses should be set up as follows: H_o: $\pi = 4/20$, or $\pi = 20\%$ H_a: $\pi < 4/20$, or $\pi < 20\%$	TRUE (FALSE)
3.	Let π = the true long-run probability that the subject answers correctly. The null and alternative hypotheses should be set up as follows: H_o: $\pi = 50\%$ H_a: $\pi < 50\%$	TRUE (FALSE)
4.	If the subject had got 7 out of 20 correct, we would have had statistical evidence to conclude that the subject was exaggerating their memory loss because 7 occurred most often in the simulation study.	TRUE (FALSE)
5.	Recall that the subject was correct on 4 of the 20 trials. The p-value for this investigation would be the proportion of dots <u>at or below 4</u> on this graph.	(TRUE) FALSE

6. Using the 5% rule and the observed result from our study (the subject got 4 out of 20 correct), there is enough statistical evidence to say the subject is intentionally giving wrong answers.	TRUE (FALSE)
7. The binomial distribution with n = 20 and $\pi = 1/3 = 0.3333$ would have approximately the same shape as the above dotplot.	(TRUE) FALSE

Questions 8 - 13 concern the following scenario. On August 17 - 22 of 2017, Quinnipiac University surveyed a random sample of 1,514 American adults. They were asked the following question, "*Which is closer to your point of view? There is too much political correctness in the United States today; or, there is too much prejudice in the United States today.*" Of those surveyed, 833 (or 833/1514 = 55%) answered "*too much prejudice*." The poll reported a margin of error of ± 3% (this is the margin of error associated with a 95% confidence interval).

8. Holding all else constant, if the poll had surveyed only 1,000 American adults instead of 1,514 American adults, the margin of error would have been larger.	(TRUE) FALSE
9. The data that this survey question yields are numerical data because we can calculate the proportion of adults in the study that answered, "*too much prejudice*" (55%).	TRUE (FALSE)
10. In this study, the population of interest is the 1,514 American adults that were surveyed.	TRUE (FALSE)
11. Based on these results, we can say that we are 95% certain that the proportion of all American adults who describe their point of view as being closer to "*too much prejudice*" is somewhere between 52% and 58%.	(TRUE) FALSE
12. It is fair to generalize these results to the opinions of adults *worldwide*.	TRUE (FALSE)
13. This poll <u>does not</u> provide enough statistical evidence that the majority of American adults would describe their point of view as being closer to "*too much prejudice*" because we can't draw such a conclusion unless a p-value is given.	TRUE (FALSE)

Questions 14 – 17 concern the following scenario. Suppose that 500 Winona State students were surveyed to find a 95% confidence interval for the true proportion of all Winona State University students who have cheated on a test. Answer the following questions by circling your response.

14. The sample proportion (i.e., the sample statistic) will <u>definitely</u> be contained in this confidence interval.	(TRUE) FALSE
15. The population proportion (i.e., the true proportion of all Winona State students who have cheated on a test) will <u>definitely</u> be contained in the confidence interval.	TRUE (FALSE)
16. If a 90% confidence interval were constructed instead of a 95% confidence interval, the margin of error would be smaller and the interval would be narrower.	(TRUE) FALSE
17. This confidence interval can be used only to describe the 500 students who were surveyed; it cannot be used to make a statement about all Winona State students.	TRUE (FALSE)

Multiple Choice Questions: (3 pts each)

Questions 18 - 19 concern the following scenario. Suppose a governor is concerned about their "negatives" (i.e., the percentage of state residents who express disapproval with their job performance). Their campaign pays for a series of television ads, hoping that they can keep the negatives below 30%. They use follow-up polling to assess the ads' effectiveness. Their hypotheses are as follows:

H_o: the ads are not effective
H_a: the ads are effective (i.e., the negatives are below 30%)

The negatives come in at 28%, and the p-value obtained is 0.18.

18. Which of the following is the most correctly written conclusion?

(a.) There is not enough evidence to conclude that the ads are effective.
b. We have evidence that 18% of the state residents disapprove of the governor.
c. We have evidence that the ads are effective.
d. We have evidence that the ads are not effective.

19. Recall that the p-value was 0.18. Which of the following is the most correct interpretation of this p-value?

a. There is an 18% chance that a state resident disapproves of the governor.
(b.) There is an 18% chance that the poll could have resulted in 28% or fewer negatives even if the ads were not effective.
c. There is an 18% chance that the ads were effective.
d. There is an 18% chance that the ads were not effective.

20. A random sample of 800 Winona residents were surveyed to estimate the proportion of Winonans who are fans of the Minnesota Vikings (Winona has a population of about 27,000 people). In another poll, a random sample of 800 Minnesotans were surveyed and asked if they were fans of the Vikings (Minnesota has about 5.5 million residents). Which of the following statements is most correct?

a. The margin of error for a 95% confidence interval will be much larger in the poll involving only residents of Winona.
b. The margin of error for a 95% confidence interval will be much larger in the statewide poll.
(c.) The margin of error for a 95% confidence interval will be about the same in both polls.

21. A large supermarket chain was accused of discriminating against women in their promotion practices. Of those eligible for promotion in the company, 60% were female. The company has promoted 50,000 employees to management positions; of these, 29,800 were women (29,800/50,000 = 59.6%). The p-value for testing whether the company was discriminating against women is found to be 0.034. These results are

a. Both statistically and practically significant.
b. Neither statistically nor practically significant.
c. Statistically non-significant but practically significant.
(d.) Statistically significant but not practically significant.

Questions 22 – 24 concern the following scenario. Suppose that a commonly used pain medication is known to cause nausea in about 25% of the patients that use it. A pharmaceutical company has created a new drug with the same pain-killing properties but claims that patients who take it are less likely to experience nausea than with the commonly used medication. In a study involving 100 patients taking the new drug, only 20 experienced nausea.

22. Let π represent the proportion of patients taking the new drug who experience nausea. Which of the following is the null hypothesis for testing the company's claim?

a. $\pi < 0.50$
b. $\pi < 0.25$
c. $\pi < 0.20$
d. $\pi = 0.50$
(e.) $\pi = 0.25$
f. $\pi = 0.20$

23. Which of the following is the alternative hypothesis?

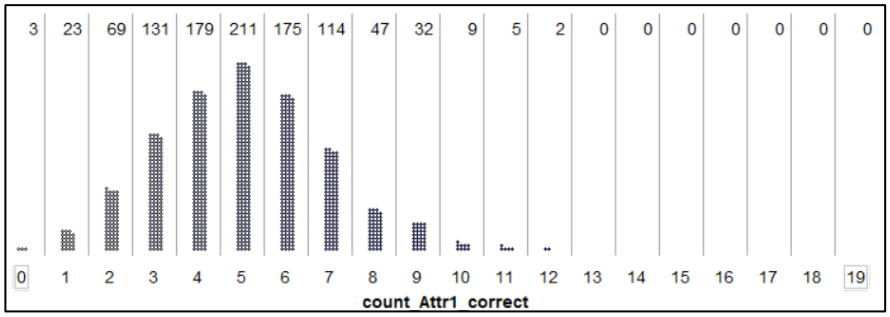
a. $\pi < 0.50$
(b.) $\pi < 0.25$
c. $\pi < 0.20$
d. $\pi = 0.50$
e. $\pi = 0.25$
f. $\pi = 0.20$

24. The results of the study yielded a p-value of 0.15. Which of the following conclusions can the pharmaceutical company draw from this study?

a. They must not be interpreting the results correctly; the p-value should have been below 0.05.
b. There is evidence that the new drug is less likely to cause nausea than the old drug.
c. There is not enough evidence that the new drug is less likely to cause nausea than the old drug; the two drugs must be equally likely to cause nausea.
(d.) There is not enough evidence that the new drug is less likely to cause nausea than the old drug. The two drugs might be equally likely to cause nausea, but it is also possible that the new drug is better and the sample size was just too small to detect this.

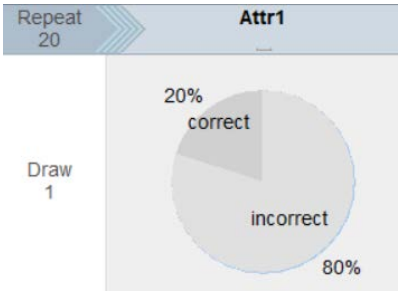
Questions 25 – 26 concern the following scenario. A man claimed to have hysterical blindness. To test his claim, researchers placed him in front of a screen on which four triangles were projected, and there was a switch located under each triangle. In each trial, three of the four triangles were inverted and the fourth (chosen at random) was upright. The subject was told to select the switch below the upright triangle. **This process was repeated 20 times, and the man correctly chose the upright triangle on 4 of the 20 trials.**

A simulation study was conducted using Tinkerplots to get an idea of what to expect if the man were truly blind. The number of trials (out of 20) in which he was correct was recorded for each of the 1,000 simulated runs. The results are shown below.

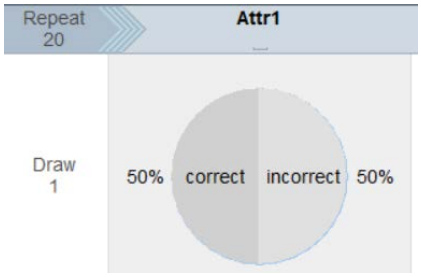


25. Which of the following spinners was used for this simulation study?

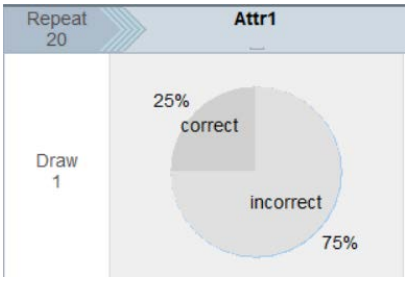
a.



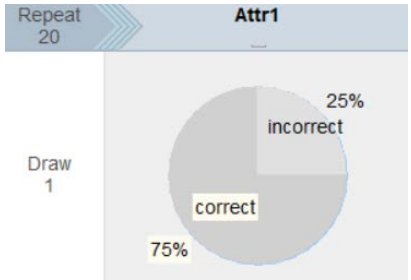
b.



(c.)



d.



26. Recall that the man was correct on 4 of the 20 trials. Which of the following is the estimate of the p-value based on the simulation study?

a. The p-value would be about 0.179 because 179/1000 dots are right at 4.
b. The p-value would be about 0.774 because 774/1000 dots are at 4 or higher.
(c.) The p-value would be about 0.405 because 405/1000 dots are at 4 or below.
d. The p-value would be about 0.20 because he was correct on 4/20 = 0.20 of the trials.

Questions 27 - 30 refer to the following scenario. A study was conducted to investigate whether people preferred bottled water to tap water. They presented 27 subjects with three cups of water. **Two cups contained bottled water, and one cup was filled with tap water.** After tasting all three, each subject reported which of the three cups of water they most preferred (they were forced to choose one). Of the 27 subjects, 6 chose the tap water. The research hypothesis is as follows: "People, in general, prefer bottled water to tap water."

27. Which of the following is the correct alternative hypothesis (H_a)?

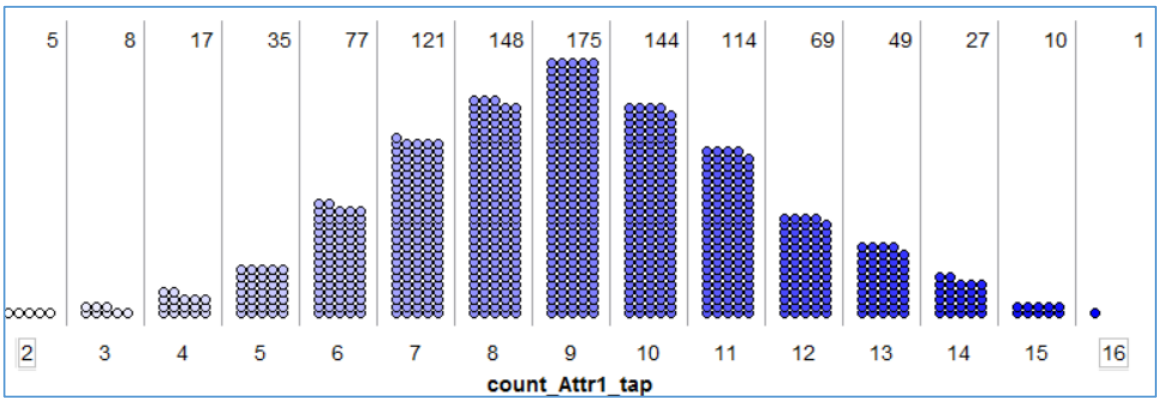
a. In general, people prefer tap water to bottled water.
b. In general, people prefer bottled water to tap water.
c. In general, there is no preference for one type of water over the other.
d. People are just guessing which cup holds the tap water.

Short Answer Questions:

28. You have been asked to conduct a simulation study to investigate this research hypothesis. Below, fill in the blanks to identify the appropriate values for a spinner in Tinkerplots that would allow you to do this. (1 pt each)

a. Box A contains the Repeat value. This should be set to **27**.
b. Box B sets the value for the proportion that chooses the tap water. This value should be set to **1/3**.
c. Box C sets the value for the proportion that chooses the bottled water. This value should be set to **2/3**.

29. Suppose the results of 1,000 simulated trials are as shown in the following dotplot. Note that we kept track of the number that chose the tap water in each simulated trial.



Recall that in the actual study, 6 subjects chose the tap water. Use the dotplot from the previous question to estimate the p-value for this study.

The estimated p-value is **0.142** (2 pts)

30. Which of the following is the <u>most correct</u> statement regarding our research question? *Note that this is a multiple choice question.* (3 pts)

a. In the study, 6/27 = 22% of the subjects chose tap water; thus 78% chose bottled water. We have statistical evidence that people prefer bottled water over tap water simply because a strong majority chose bottled water.
b. In the study, 6/27 = 22% of the subjects chose tap water; since this value is greater than 0.05 (or 5%), we do not have enough statistical evidence that people prefer bottled water over tap water.
c. An outcome as extreme as only 6 subjects choosing the tap water happened more than 5% of the time in our simulation study; thus, there is not enough statistical evidence that people prefer bottled water over tap water.
d. An outcome as extreme as only 6 subjects choosing the tap water happened less than 5% of the time in our simulation study; thus, there is statistical evidence that people prefer bottled water over tap water.

31. A random sample of 300 students is surveyed on a university campus. They are asked if they use a laptop in class to take notes. Suppose that based on the survey, 70 of the 300 students responded "yes." Your goal is to estimate the true proportion of all students on this campus who use a laptop to take notes.

a. Find the sample statistic of interest (also called the point estimate). (1 pt)

70/300 = 0.233

b. Find the margin of error associated with a 95% confidence interval. (3 pts)

$$\mathbf{1.96\sqrt{\frac{(0.233)(1-0.233)}{300}} = 0.048}$$

c. Construct a 95% confidence interval for the true proportion of all students on this campus who use a laptop to take notes. (2 pts)

$0.233 \pm 0.048 \rightarrow 0.185 \leq \pi \leq 0.281$

32. A convenience store offers a "scratch off" promotion. When you enter the store, you are given a card. When you pay, you may scratch off the coating. The store advertises that every card is a winner; moreover, they advertise that half the cards are winners offering immediate cash-back savings of $5 (the others offer a free cup of coffee). You are skeptical of this claim and decide to test it. You go into the store on six different occasions, and you get free coffee all six times. **Your research hypothesis is that more than half of the cards will offer a free cup of coffee instead of the $5 savings.**

a. Let π = the probability that the card offers you a free cup of coffee. Write your null and alternative hypothesis for this investigation. (2 pts)

H_o: **$\pi = 0.50$ (the store's advertising is accurate – half of the cards are winners offering cash-back savings and the other half offer coffee)**

H_a: **$\pi > 0.50$ (the store's advertising is not accurate – more than half of the cards are winners offering free cups of coffee)**

b. When using the binomial distribution to find the p-value, what values should you use for n and π? Keep in mind that π = the probability that the card offers you a free cup of coffee. (3 pts)

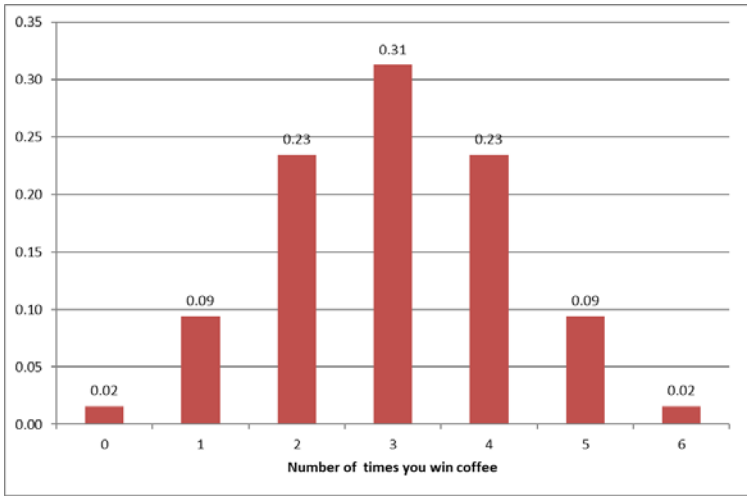
n = **6**

π = **0.50**

c. Recall that you got a free cup of coffee on all six of your visits. Use the binomial probabilities provided below to find the p-value. (3 pts)

p-value: **0.02**

Number of Times You Win Coffee	Binomial Probabilities
0	0.02
1	0.09
2	0.23
3	0.31
4	0.23
5	0.09
6	0.02



d. Write a conclusion to address the research question in the context of the problem. (2 pts)

Since the p-value is less than 0.05, the data do provide evidence that more than half of the cards will offer a free cup of coffee instead of the $5 savings.

33. A company is criticized because only 6 of 31 people in executive-level positions are women (6/31 = 19.4%). The company claims that although this is lower than it might wish, it's not surprising given that only 30% of their employees are women. You decide to conduct a hypothesis test to see whether these data provide sufficient evidence to conclude the company is discriminating against women. **Your research question is as follows: Are there fewer women in executive-level positions in this company than we would expect by chance?**

a. Let π = the probability that someone in an executive-level position in this company is a woman. Write your null and alternative hypothesis for this investigation. (2 pts)

H_o: **$\pi = 0.35$ (the company is not discriminating against women in terms of its executive-level positions)**

H_a: **$\pi < 0.35$ (the company is discriminating against women; i.e., there are fewer women in executive-level positions than we would expect)**

b. When using the binomial distribution to find the p-value, what values should you use for n and π? Keep in mind that the researchers are keeping track of the number of people in executive-level positions who are women, so π = the probability that someone in an executive-level position in this company is a woman. (3 pts)

n = **31**

π = **0.30**

c. Recall that 6 out of 31 people in executive-level positions in this company are women. Use the binomial probabilities provided on the next page to find the p-value. (3 pts)

p-value: **0.14**

d. Write a conclusion to address the research question in the context of the problem. (2 pts)

Since the p-value is above 0.05, there is not enough evidence to conclude that the company is discriminating against women.

e. Complete the following statement using the binomial probabilities from the next page: "Using the 5% rule, we will have evidence that this company has fewer women in executive-level positions than we would expect by chance if we see **6** or fewer women in the 31 executive positions. (2 pts)

Number of Women	Binomial Probabilities
0	0.00
1	0.00
2	0.00
3	0.01
4	0.02
5	0.04
6	0.07
7	0.11
8	0.14
9	0.16
10	0.15
11	0.12
12	0.09
13	0.05
14	0.03
15	0.01
16	0.01
17	0.00
18	0.00
19	0.00
20	0.00
21	0.00
22	0.00
23	0.00
24	0.00
25	0.00
26	0.00
27	0.00
28	0.00
29	0.00
30	0.00
31	0.00