

Chapter 1

Introduction to Statistical Investigations Test Bank

Note: TE = Text entry TE-N = Text entry - Numeric
 Ma = Matching MS = Multiple select
 MC = Multiple choice TF = True-False

 E = Easy, M = Medium, H = Hard

CHAPTER 1 LEARNING OBJECTIVES

- CLO1-1: Use the chance model to determine whether an observed statistic is unlikely to occur.
- CLO1-2: Calculate and interpret a p-value, and state the strength of evidence it provides against the null hypothesis.
- CLO1-3: Calculate a standardized statistic for a single proportion and evaluate the strength of evidence it provides against a null hypothesis.
- CLO1-4: Describe how the distance of the observed statistic from the parameter value specified by the null hypothesis, sample size, and one- vs. two-sided tests affect the strength of evidence against the null hypothesis.
- CLO1-5: Describe how to carry out a theory-based, one-proportion z-test.

Section 1.1: Introduction to Chance Models

- LO1.1-1: Recognize the difference between parameters and statistics.
- LO1.1-2: Describe how to use coin tossing to simulate outcomes from a chance model of the random choice between two events.
- LO1.1-3: Use the One Proportion applet to carry out the coin tossing simulation.
- LO1.1-4: Identify whether or not study results are statistically significant and whether or not the chance model is a plausible explanation for the data.
- LO1.1-5: Implement the 3S strategy: find a statistic, simulate results from a chance model, and comment on strength of evidence against observed study results happening by chance alone.
- LO1.1-6: Differentiate between saying the chance model is plausible and the chance model is the correct explanation for the observed data.

Questions 1 through 4:

Do red uniform wearers tend to win more often than those wearing blue uniforms in Taekwondo matches where competitors are randomly assigned to wear either a red or blue uniform? In a sample of 80 Taekwondo matches, there were 45 matches where the red uniform wearer won.

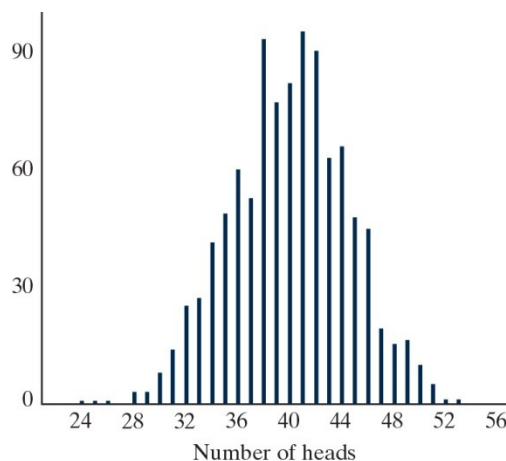
1. What is the parameter of interest for this study?
 - A. The long-run proportion of Taekwondo matches in which the red uniform wearer wins
 - B. The proportion of matches in which the red uniform wearer wins in a sample of 80 Taekwondo matches
 - C. Whether the red uniform wearer wins a match
 - D. 0.50

Ans: A; LO: 1.1-1; Difficulty: Easy; Type: MC

2. What is the statistic for this study?
 - A. The long-run proportion of Taekwondo matches in which the red uniform wearer wins
 - B. The proportion of matches in which the red uniform wearer wins in a sample of 80 Taekwondo matches
 - C. Whether the red uniform wearer wins a match
 - D. 0.50

Ans: B; LO: 1.1-1; Difficulty: Easy; Type: MC

3. Given below is the simulated distribution of the number of “red wins” that could happen by chance alone in a sample of 80 matches. Based on this simulation, is our observed result statistically significant?



- A. Yes, since 45 is larger than 40.
 - B. Yes, since the height of the dotplot above 45 is smaller than the height of the dotplot above 40.
 - C. No, since 45 is a fairly typical outcome if the color of the winner's uniform was determined by chance alone.

FOR INSTRUCTOR USE ONLY

D. No, since we could have observed a value greater than 45 just by random chance.

Ans: C; LO: 1.1-4; Difficulty: Medium; Type: MC

4. What can we conclude from the results of this study? Select all that apply.
- A. The results of this study are something that could easily have happened if the color of the winner's uniform was determined by chance alone.
 - B. We do not have convincing evidence against the "by-chance-alone" model.
 - C. The results of this study prove that the color of the winner's uniform was determined by chance alone.
 - D. We do not have convincing evidence that red uniform wearers tend to win more often than those wearing blue uniforms.

Ans: A, B, D; LO: 1.1-6; Difficulty: Hard; Type: MS

Questions 5 through 8:

Suppose you are testing to see if your dog, Hope, understands pointing towards an object. You place two objects about 2.5 meters away, then you point towards one of the objects. In 20 trials, Hope goes to the correct object 13 times (or 65%).

5. Fill in the blanks with the correct One Proportion applet inputs to carry out an appropriate simulation of this process, if Hope does not understand pointing towards an object and is just guessing.

Probability of success: _____

Sample size: _____

Number of samples: _____

Ans: 0.5 (Tol: 0), 20 (Tol: 0), Any integer as large or larger than 1000; LO: 1.1-3;

Difficulty: Easy; Type: TE-N

6. Match the parts of the real study corresponding to the physical (coin-flipping) simulation:

Coin flip = _____

Heads = _____

Tails = _____

Chance of heads = _____

One repetition = _____

A. 0.5, probability of Hope going to the correct object

B. Hope going to the correct object

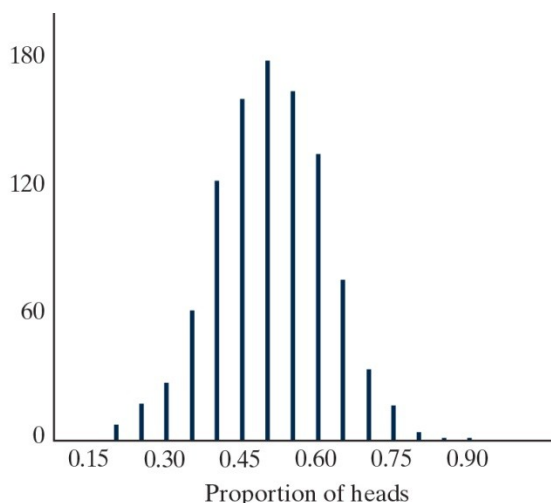
C. Hope going to the incorrect object

D. One set of 20 attempts by Hope

E. Hope going to an object

Ans: E, B, C, A, D; LO: 1.1-2; Difficulty: Medium; Type: Ma

7. Given below is the resulting simulation distribution from the One Proportion applet. Based on this simulation, do we have convincing evidence against the observed study results happening by chance alone?



- A. No, since 0.65 is a fairly typical outcome if Hope was just guessing.
- B. No, since we could have observed a value greater than 0.65 just by random chance.
- C. Yes, since 0.65 is larger than 0.5.
- D. Yes, since the height of the dotplot above 0.65 is smaller than the height of the dotplot above 0.5.

Ans: A; LO: 1.1-5; Difficulty: Medium; Type: MC

8. Which of the following is NOT a step in the 3S Strategy for measuring strength of evidence?
- A. Strength of evidence
 - B. Statistic
 - C. Sample
 - D. Simulate

Ans: C; LO: 1.1-5; Difficulty: Easy; Type: MC

Questions 9 through 12:

True or False?

9. A parameter is a number summarizing the results in the sample.
Ans: False; LO: 1.1-1; Difficulty: Easy; Type: TF
10. A result that is unlikely to occur just by chance alone is called statistically significant.
Ans: True; LO: 1.1-4; Difficulty: Easy; Type: TF
11. If an observed statistic is statistically significant, then we have strong evidence that the chance-alone explanation is correct.
Ans: False; LO: 1.1-5; Difficulty: Medium; Type: TF
12. If an observed statistic is NOT statistically significant, then the chance model is plausible.

FOR INSTRUCTOR USE ONLY

Ans: True; LO: 1.1-6; Difficulty: Medium; Type: TF

Section 1.2: Measuring Strength of Evidence

LO1.2-1: Use appropriate symbols for parameter and statistic.

LO1.2-2: State the null and the alternative hypotheses in words and in terms of the symbol π , the long-run proportion.

LO1.2-3: Explain how to conduct a simulation using a null hypothesis probability that is not 50-50.

LO1.2-4: Use the One Proportion applet to obtain the p-value after carrying out an appropriate simulation.

LO1.2-5: Anticipate the location of the center of the null distribution and how it changes based on whether you are using proportion or count as the statistic.

LO1.2-6: Interpret the p-value.

LO1.2-7: Explain why a smaller p-value provides stronger evidence against the null hypothesis.

LO1.2-8: State a conclusion about the alternative hypothesis and null hypothesis based on the p-value.

Questions 13 through 18:

A survey on 1,500 high school seniors who took the SAT and who completed an optional web survey shows that 55% of high school seniors are fairly certain that they will participate in a study abroad program in college. Does this survey provide convincing evidence that the majority (more than 50%) of all high school seniors who take the SAT are fairly certain they will participate in a study abroad program in college?

13. What is the value of the statistic and its proper notation in this study?

- A. $\pi = 0.55$
- B. $\hat{p} = 0.55$
- C. $\pi = 825$
- D. $\hat{p} = 825$

Ans: B; LO: 1.2-1; Difficulty: Easy; Type: MC

14. Under the null hypothesis, what is the value of the parameter of interest and its proper notation in this study?

- A. $\pi = 0.50$
- B. $\hat{p} = 0.50$
- C. $\pi = 825$
- D. $\hat{p} = 825$

Ans: A; LO: 1.2-1; Difficulty: Easy; Type: MC

15. State the null and alternative hypotheses in proper notation.

A. $H_0 : \pi = 0.50$ vs $H_a : \pi = 0.55$

B.

$H_0 : \pi = 0.55$ vs $H_a : \pi > 0.55$

C.

$H_0 : \pi = 0.50$ vs $H_a : \pi > 0.50$

D.

$H_0 : \hat{p} = 0.50$ vs $H_a : \hat{p} > 0.50$

Ans: C; LO: 1.2-2; Difficulty: Easy; Type: MC

16. Fill in the blanks with the correct values to simulate another survey of 1,500 high school seniors who took the SAT in which the long-run proportion of all high school seniors who take the SAT that are fairly certain they will participate in a study abroad program in college is 0.50.

Create a spinner with _____ percent shaded red, and _____ percent shaded black. Spin the spinner _____ times. Record the _____ of times the spinner lands on red.

Ans: 50 (Tol: 0), 50 (Tol: 0), 1500 (Tol: 0), proportion (or count); LO: 1.2-3; Difficulty: Medium; Type: TE, TE-N

17. Using a *count* as the statistic, where would you expect the null distribution to be centered?

A. 0.55

B. 0.50

C. 825

D. 750

Ans: D; LO: 1.2-5; Difficulty: Medium; Type: MC

18. The p-value for this study is less than 0.001. Interpret this value in the context of the problem.

A. There is less than a 0.1% chance that 50% of all high school seniors who take the SAT are fairly certain they will participate in a study abroad program in college.

B. In less than 0.1% of all samples of 1,500 high school seniors who take the SAT, we would see 55% or more respond that they are fairly certain they will participate in a study abroad program in college, if the true probability is 0.50.

C. In less than 0.1% of all samples of 1,500 high school seniors who take the SAT, we would see 50% or more respond that they are fairly certain they will participate in a study abroad program in college, if the true probability is 0.55.

D. There is less than a 0.1% chance that more than 50% of all high school seniors who take the SAT are fairly certain they will participate in a study abroad program in college.

FOR INSTRUCTOR USE ONLY

Ans: B; LO: 1.2-6; Difficulty: Hard; Type: MC

Questions 19 through 23:

Research done in the mid 1980s indicated that 80% of grizzly bears in the greater ecosystem of Yellowstone National Park entered their den for hibernation by the last day of November. This has come into question as of late. Researchers have hypothesized that climate change has postponed bears entering hibernation until later in the season; that is, researchers believe that a smaller percentage of bears are entering their den for hibernation by the last day of November. Sixty-two Yellowstone grizzly bears were tracked via radio monitors, which allow scientists to pinpoint when the bears enter their dens with a high degree of accuracy. Of the sixty-two Yellowstone grizzly bears being tracked, forty-two entered their den for hibernation by the last day of November.

19. Identify which of the following is the null hypothesis, and which is the alternative hypothesis:

_____ The sample proportion of 62 Yellowstone grizzly bears that enter their den for hibernation by the last day of November is 0.80.

_____ The long-run proportion of Yellowstone grizzly bears that enter their den for hibernation by the last day of November is 0.80.

_____ The long-run proportion of Yellowstone grizzly bears that enter their den for hibernation by the last day of November is less than 0.80.

_____ The long-run proportion of Yellowstone grizzly bears that enter their den for hibernation by the last day of November is greater than 0.80.

Ans: 2nd blank = Null; 3rd blank = Alternative; LO: 1.2-2; Difficulty: Medium; Type: MC

20. What values would you enter for each of the following in the One Proportion applet to conduct a simulation-based hypothesis test with these data?

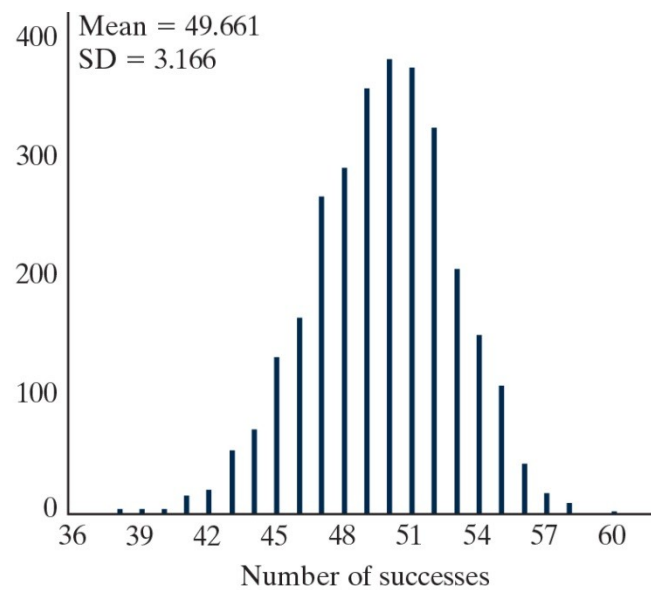
Probability of success: _____

Sample size: _____

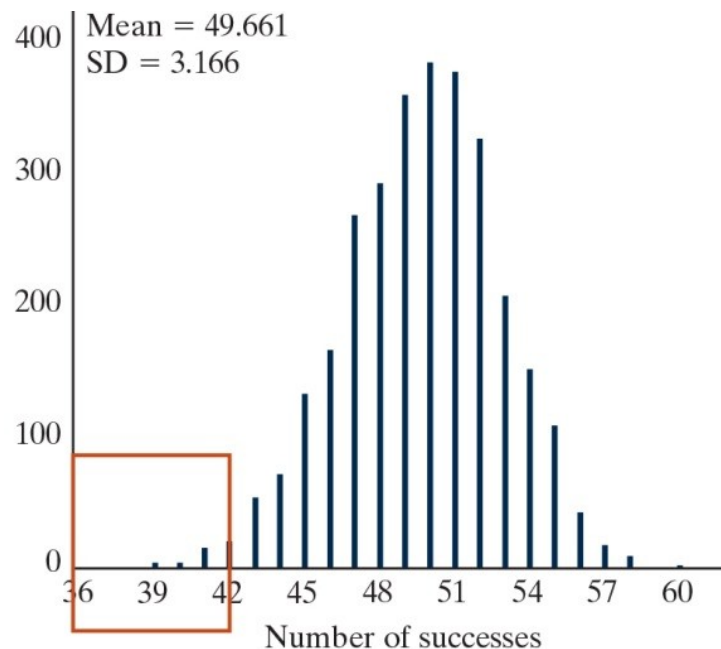
Number of samples: _____

Ans: 0.8 (Tol: 0), 62 (Tol: 0), Any integer as large or larger than 1000; LO: 1.2-4; Difficulty: Easy; Type: TE-N

21. A distribution of simulated "Number of successes," assuming the null hypothesis is true, is shown below. Circle the portion of the plot below that you would count to calculate the p-value.



Ans:



LO: 1.2-4; Difficulty: Medium; Type: Other

22. How could we use coins, a spinner, or cards to simulate one of the samples in the above simulation? Select all that apply.

- A. Flip a coin 62 times, where heads represents “hibernate by November 30.” Count the number of heads in 62 flips.
- B. Create a spinner with 80% shaded to represent “hibernate by November 30.” Spin the spinner 62 times and count the number of times the spinner lands in the shaded area.
- C. Have 8 black cards and 2 red cards where black represents “hibernate by Novem-

FOR INSTRUCTOR USE ONLY

ber 30". Draw with replacement from the cards 62 times and count the number of times a black card is drawn.

- D. Create a spinner with 68% shaded to represent "hibernate by November 30." Spin the spinner 62 times and count the number of times the spinner lands in the shaded area.

Ans: B, C; LO: 1.2-3; Difficulty: Hard; Type: MS

23. The simulation-based p-value for this test is 0.0153. Which of the following conclusions are correct? Select all that apply.

- A. We have strong evidence against the null hypothesis.
- B. We have strong evidence in favor of the null hypothesis.
- C. We have strong evidence against the alternative hypothesis.
- D. We have strong evidence in favor of the alternative hypothesis.

Ans: A, D; LO: 1.2-8; Difficulty: Medium; Type: MS

Questions 24 through 26:

True or False?

24. When using the One Proportion applet, the simulated null distribution should be centered around the observed sample proportion.

Ans: False; LO: 1.2-5; Difficulty: Easy; Type: TF

25. The p-value is the probability that the null hypothesis is true.

Ans: False; LO: 1.2-6; Difficulty: Medium; Type: TF

26. A smaller p-value provides stronger evidence against the null hypothesis.

Ans: True; LO: 1.2-7; Difficulty: Easy; Type: TF

Section 1.3: Alternative Measure of Strength of Evidence

LO1.3-1: Find a standardized statistic from the observed proportion of "successes," the hypothesized mean, and SD of the null distribution as produced by the One Proportion applet.

LO1.3-2: Interpret a standardized statistic.

LO1.3-3: State a conclusion about the alternative hypothesis (and null hypothesis) based on the magnitude of the standardized statistic.

LO1.3-4: Recognize that the standardized statistic is an alternative to the p-value.

Questions 27 through 32:

The first iPhone became available for public purchase in 2007. By 2010, approximately 30% of smartphone users in the US owned an iPhone. Several new companies have entered the smartphone market since then and the folks at Apple want to know how they are stacking up against their competition. In 2019, a representative sample of 529 smartphone users in the US

showed that 47.4% use an Apple iOS device. Does it appear that Apple's share of the US smartphone market has increased since 2010?

27. State the null and alternative hypotheses in proper notation.

- A. $H_0 : \pi = 0.3$ vs $H_a : \pi = 0.474$
- B. $H_0 : \pi = 0.3$ vs $H_a : \pi > 0.3$
- C. $H_0 : \pi = 0.474$ vs $H_a : \pi > 0.474$
- D. $H_0 : \hat{p} = 0.3$ vs $H_a : \hat{p} > 0.3$

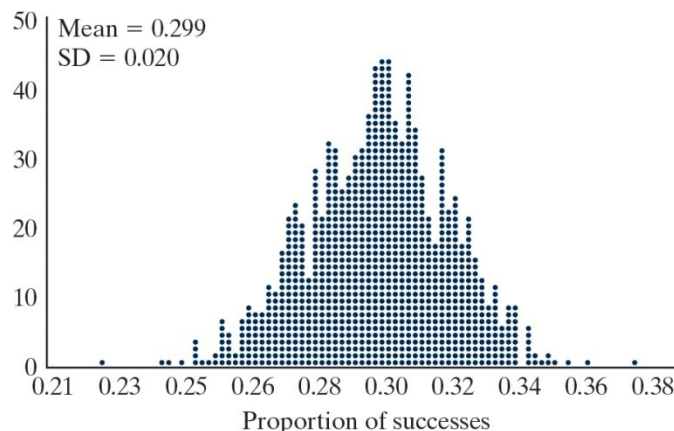
Ans: B; LO: 1.2-2; Difficulty: Easy; Type: MC

28. What is the value of the statistic and its proper notation in this study?

- A. $\pi = 0.529$
- B. $\hat{p} = 0.529$
- C. $\pi = 0.474$
- D. $\hat{p} = 0.474$

Ans: D; LO: 1.2-1; Difficulty: Easy; Type: MC

29. The following simulated null distribution of proportion of "Apple iOS users" that could happen by chance alone in a sample of 529 smartphone users was created with the One Proportion applet. Also, given are the mean and SD for this null distribution.



Calculate and report the standardized statistic.

Ans: 8.7 (Tol: +/- 0.1)

Solution: $(0.474 - 0.3)/0.020 = 8.7$ (Tol = +/-0)

LO: 1.3-1; Difficulty: Medium; Type: TE-N

30. Let X represent the value of your standardized statistic in #29. Select the best contextual interpretation of the value X .

- A. On average, each simulated proportion of the US smartphone market which uses Apple iOS devices is X away from 0.299.
- B. The probability of obtaining a sample proportion of 0.474 or something more extreme, assuming 30% of all US smartphone users use an Apple iOS device, is X .
- C. The proportion of the US smartphone market that uses an Apple iOS device is X .
- D. The 0.474 proportion of the US smartphone users sampled who use Apple iOS devices is X standard deviations above the 2010 Apple US market-share of 0.30.

Ans: D; LO: 1.3-2; Difficulty: Hard; Type: MC

31. Which of the following strength of evidence statements are correct, based on the standardized statistic in #29? Select all that apply. (If you didn't report a standardized statistic in #29, then use $z = 4$ for this question.)

- A. We have very strong evidence against the null hypothesis.
- B. We have very strong evidence against the alternative hypothesis.
- C. We have very strong evidence in favor of the null hypothesis.
- D. We have very strong evidence in favor of the alternative hypothesis.

Ans: A, D; LO: 1.3-3; Difficulty: Medium; Type: MS

32. What can we conclude about this study, based on the standardized statistic in #29? (If you didn't report a standardized statistic in #29, then use $z = 4$ for this question.)

- A. We have very strong evidence that Apple's share of the US smartphone market has remained the same since 2010.
- B. We have very strong evidence that Apple's share of the US smartphone market has increased since 2010.
- C. We do not have very strong evidence that Apple's share of the US smartphone market has increased since 2010.
- D. We have very strong evidence that the proportion of the representative sample of 529 smartphone users in the US using an Apple iOS device is greater than 0.30.

Ans: B; LO: 1.3-3; Difficulty: Medium; Type: MC

33. Match each standardized statistic value with its corresponding p-value:

$z = -0.45$

A. p-value = 0.009

$z = 3.31$

B. p-value = 0.085

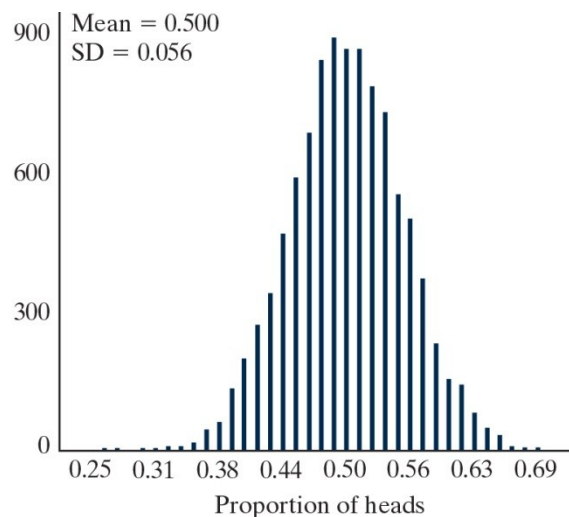
$z = -1.72$

C. p-value = 0.653

Ans: C, A, B; LO: 1.3-4; Difficulty: Medium; Type: Ma

Questions 34 and 35:

Do red uniform wearers tend to win more often than those wearing blue uniforms in Taekwondo matches where competitors are randomly assigned to wear either a red or blue uniform? In a sample of 80 Taekwondo matches, there were 45 matches where the red uniform wearer won. Given below is the simulated null distribution of proportion of “red wins” that could happen by chance alone in a sample of 80 matches. Also, given are the mean and SD for this null distribution.



34. What is the value of the standardized statistic?

- A. $z = 0.5625$
- B. $z = -0.5625$
- C. $z = 1.12$
- D. $z = -1.12$

Ans: C; LO: 1.3-1; Difficulty: Easy; Type: MC

35. Is there evidence that in Taekwondo matches, red uniform wearers tend to win more often than those wearing blue uniforms than would be expected by chance alone?

- A. Yes, since 45 is greater than 30.
- B. No, since we only observed 80 Taekwondo matches.
- C. No, since the standardized statistic is between -1.5 and 1.5 .
- D. Yes, since the standardized statistic is between -1.5 and 1.5 .

Ans: C; LO: 1.3-3; Difficulty: Medium; Type: MC

Questions 36 through 38:

True or False?

36. The standardized statistic represents the number of standard deviations the observed statistic is above or below the hypothesized null value.

Ans: True; LO: 1.3-2; Difficulty: Easy; Type: TF

37. If the standardized statistic is larger than 3, then the p-value will be larger than 0.10.

Ans: False; LO: 1.3-4; Difficulty: Medium; Type: TF

38. A standardized statistic that is closer to zero provides stronger evidence against the null hypothesis.

Ans: False; LO: 1.3-3; Difficulty: Easy; Type: TF

Section 1.4: What Impacts Strength of Evidence?

LO1.4-1: Anticipate and explain why, when everything else remains the same, the p-value is smaller when the observed proportion of successes is farther away from the hypothesized value of the long-run proportion of successes π .

LO1.4-2: Anticipate and explain why, when everything else remains the same, the p-value is larger when the alternative hypothesis is two-sided.

LO1.4-3: Recognize when a two-sided test/alternative hypothesis is suggested by the research question.

LO1.4-4: Anticipate and explain why, when everything else remains the same, the p-value is smaller when the sample size is larger.

Questions 39 through 41:

Research done in the mid 1980s indicated that 80% of grizzly bears in the greater ecosystem of Yellowstone National Park entered their den for hibernation by the last day of November. This has come into question as of late. Researchers have hypothesized that climate change has postponed bears entering hibernation until later in the season; that is, researchers believe that a *smaller* percentage of bears are entering their den for hibernation by the last day of November.

Sixty-two Yellowstone grizzly bears were tracked via radio monitors, which allow scientists to pinpoint when the bears enter their dens with a high degree of accuracy. Of the 62 Yellowstone grizzly bears being tracked, 42 entered their den for hibernation by the last day of November.

For each of the following sample results, circle whether the p-value would increase, decrease, or remain the same if we were to track another sample of Yellowstone grizzly bears.

39. 126 of 186 bears had entered their den for hibernation by the last day of November.

- A. Increase
- B. Decrease
- C. Remain the same

Ans: B; LO: 1.4-4; Difficulty: Medium; Type: MC

40. 45 of 62 bears had entered their den for hibernation by the last day of November.

- A. Increase
- B. Decrease
- C. Remain the same

Ans: A; LO: 1.4-1; Difficulty: Medium; Type: MC

41. 21 of 31 bears had entered their den for hibernation by the last day of November.

- A. Increase
- B. Decrease
- C. Remain the same

Ans: A; LO: 1.4-4; Difficulty: Medium; Type: MC

Questions 42 through 51:

An environmental science teacher at a high school with a large population of students was curious whether the majority (more than half) of students at the school regularly recycle plastic bottles. The teacher selected a random sample of 146 students at the school to interview, then asked each student: "Do you recycle plastic bottles more often than you throw them in the trash?" (Yes/No). Suppose that 89 of the students answered "Yes".

42. What is the parameter of interest for this study?

- A. The proportion of the 146 students in the sample that would answer "Yes"
- B. The proportion of all students at the school that would answer "Yes"
- C. All students at the school
- D. The true mean number of students at the school who would answer "Yes"

Ans: B; LO: 1.1-1; Difficulty: Medium; Type: MC

43. What are the null and alternative hypotheses in appropriate notation?

- A. $H_0 : \pi = 0.50$ vs $H_a : \pi = 0.61$
- B. $H_0 : \pi = 0.61$ vs $H_a : \pi > 0.61$
- C. $H_0 : \pi = 0.50$ vs $H_a : \pi > 0.50$
- D. $H_0 : \hat{p} = 0.50$ vs $H_a : \hat{p} > 0.50$

Ans: C; LO: 1.2-2; Difficulty: Easy; Type: MC

44. Use the One Proportion applet to calculate a p-value for this study.

Ans: 0.005 (Tol = +/- 0.0015); LO: 1.2-4; Difficulty: Hard; Type: TE-N

45. How could we use coins, a spinner, or cards to simulate one of the samples in the simulated null distribution? Select all that apply.

- A. Flip a coin 146 times, where heads represents "answered Yes". Count the number

of heads in 146 flips.

- B. Create a spinner with 61% shaded to represent “answered Yes”. Spin the spinner 146 times and count the number of times the spinner lands in the shaded area.
- C. Have 5 black cards and 5 red cards where black represents “answered Yes”. Draw with replacement from the cards 146 times and count the number of times a black card is drawn.
- D. Create a spinner with 50% shaded to represent “answered Yes”. Spin the spinner 146 times and count the number of times the spinner lands in the shaded area.

Ans: A, C, D; LO: 1.1-2; Difficulty: Easy; Type: MS

46. Would you consider the results of this test to be statistically significant?

- A. Yes, because the simulated distribution is not centered on the observed statistic.
- B. No, because the results are unlikely to occur by random chance alone.
- C. No, because the simulation proves that the null hypothesis is true.
- D. Yes, because the results are unlikely to occur by random chance alone.

Ans: D; LO: 1.1-4; Difficulty: Medium; Type: MC

For each of the following (**questions 47 through 51**), circle whether we would have stronger, weaker, or the same amount of evidence against the null hypothesis, given all other aspects of the study remained the same, or if the answer cannot be determined from the information given

47. The teacher sampled 300 students instead of 146.

- A. Stronger evidence against the null hypothesis
- B. Weaker evidence against the null hypothesis
- C. The same amount of evidence against the null hypothesis
- D. Cannot be determined from the information given

Ans: A; LO: 1.4-4; Difficulty: Medium; Type: MC

48. The sample proportion of students who said “Yes” was 0.75 instead of 0.61.

- A. Stronger evidence against the null hypothesis
- B. Weaker evidence against the null hypothesis
- C. The same amount of evidence against the null hypothesis
- D. Cannot be determined from the information given

Ans: A; LO: 1.4-1; Difficulty: Medium; Type: MC

49. The sample proportion of students who said “Yes” was 0.55 instead of 0.61.

- A. Stronger evidence against the null hypothesis
- B. Weaker evidence against the null hypothesis
- C. The same amount of evidence against the null hypothesis
- D. Cannot be determined from the information given

Ans: B; LO: 1.4-1; Difficulty: Medium; Type: MC

50. We want to know if the proportion of students at the school that would have said “Yes” *differed* from 0.5 rather than was greater than 0.5.

- A. Stronger evidence against the null hypothesis
- B. Weaker evidence against the null hypothesis
- C. The same amount of evidence against the null hypothesis
- D. Cannot be determined from the information given

Ans: B; LO: 1.4-2; Difficulty: Medium; Type: MC

51. We want to know if *less* than half of all students at the school would have said “Yes” rather than greater than half.

- A. Stronger evidence against the null hypothesis
- B. Weaker evidence against the null hypothesis
- C. The same amount of evidence against the null hypothesis
- D. Cannot be determined from the information given

Ans: B; LO: 1.4-2; Difficulty: Hard; Type: MC

Questions 52 through 55:

For each of the following scenarios, determine whether the alternative hypothesis would be one-sided or two-sided.

52. A recent Gallup poll reported that 77% of Americans are at least ‘somewhat’ confident in the government’s ability to handle a coronavirus outbreak. A student researcher is interested if students at their university are more skeptical (less likely to feel confident) than what was reported in the Gallup poll.

- A. one-sided
- B. two-sided

Ans: A; LO: 1.4-3; Difficulty: Medium; Type: MC

53. In 2010, it was believed that 18% of American cell phone owners do their browsing on their phone rather than a computer or other device. A recent survey conducted on a representative sample of 621 American cell phone owners showed 22% now do their browsing on their mobile device. Do these data provide evidence that the proportion of American cell phone owners who do their browsing on their phones rather than their computers or other devices has changed in the last 10 years?

- A. one-sided
- B. two-sided

Ans: B; LO: 1.4-3; Difficulty: Medium; Type: MC

54. According to enrollment data in 2016, 55% of undergraduates at a state university were out-of-state students. Your statistics instructor wonders if the proportion of out-of-state undergraduates at this university has changed. She surveys her students (145 students). The results of the sample showed that 73 were out-of-state students.

- A. one-sided

B. two-sided

Ans: B; LO: 1.4-3; Difficulty: Medium; Type: MC

55. Retired professional tennis players Martina Navratilova, Monica Seles, John McEnroe, and Jimmy Connors are all left-handed. About 10% of the general population is left-handed. Researchers are interested if the proportion of professional tennis players that are left-handed is larger than that of the general population.

A. one-sided

B. two-sided

Ans: A; LO: 1.4-3; Difficulty: Medium; Type: MC

Section 1.5: Inference for a Single Proportion: Theory-based Approach

LO1.5-1: Predict the shape, mean, and standard deviation of the null distribution of sample proportions, as approximately normal (when the sample size is large enough), π , and $\sqrt{\pi(1-\pi)/n}$, respectively.

LO1.5-2: Use the One Proportion applet to find the one-proportion z-test (theory-based; normal approximation-based), p-value, and standardized statistic, z.

LO1.5-3: Explain when simulation and theory will yield different answers.

56. The official four-year graduation rate among US colleges and universities is 57.6%.

Suppose a researcher takes a random sample of 100 college students and tracks their progress through school to determine whether each student graduates in four years or not. Which of the following do you know to be true? Select all that apply.

A. If many samples of 100 students were taken, the distribution of the sample proportions would be bell-shaped and symmetric.

B. If many samples of 100 students were taken, the distribution of the sample proportions would be centered at 0.576.

C. The sample proportion collected by the researcher would be 0.576.

D. If many samples of 100 students were taken, the standard deviation of the sample proportions would be 0.049.

Ans: A, B, D; LO: 1.5-1; Difficulty: Medium; Type: MS

Questions 57 through 63:

A study conducted by the National Center for Health Statistics as part of the Center for Disease Control and Prevention (CDC) found that, as of 2015-2016, 39.6% of Americans over the age of 20 fit the medical definition of being obese (BMI of 30 or higher). A city known for its plentiful outdoor opportunities believes that number is lower among its residents. A representative sample of 80 city residents is collected and the BMI calculated for each (based on reported height and weight). The variable of interest is whether the resident fit the medical definition of obese. The study found 22 people in the sample fit the medical definition of obese.

57. Is a theory-based method appropriate for these data?

- A. Yes, because a representative sample was collected.
- B. Yes, because there were 22 people defined as obese and 58 people not defined as obese in the sample, and both of these values are greater than 10.
- C. No, because the researcher did not collect a random sample.
- D. No, because the national obesity rate is not 0.5.

Ans: B; LO: 1.5-3; Difficulty: Easy; Type: MC

58. Identify which of the following is the null hypothesis.

- A. The sample proportion of 80 city residents that fit the medical definition of obese is equal to 0.396.
- B. The long-run proportion of city residents that fit the medical definition of obese is equal to 0.396.
- C. The long-run proportion of all Americans over the age of 20 that fit the medical definition of obese is equal to 0.396.
- D. The sample proportion of 80 city residents that fit the medical definition of obese is equal to 0.275.

Ans: B; LO: 1.2-2; Difficulty: Medium; Type: MC

59. Identify the correct alternative hypothesis.

- A. $H_a : \pi \neq 0.396$
- B. $H_a : \pi > 0.396$
- C. $H_a : \pi < 0.396$
- D. $H_a : \hat{p} = 0.275$

Ans: C; LO: 1.2-2, 1.4-3; Difficulty: Easy; Type: MC

60. Select the appropriate method of calculating the standard deviation of the sample proportions under the null hypothesis.

- A. $\sqrt{\frac{0.396 \times (1 - 0.396)}{80}}$
- B. $\sqrt{\frac{0.275 \times (1 - 0.275)}{80}}$
- C. $\sqrt{\frac{22 \times (1 - 22)}{80}}$

$$D. \sqrt{\frac{80 \times (1 - 80)}{0.396}}$$

Ans: A; LO: 1.5-1; Difficulty: Easy; Type: MC

61. Let 'A' represent your answer from #60. Which of the following displays the correct method of calculating the standardized statistic?

$$A. \frac{22 - 31.7}{80 \times A}$$

$$B. \frac{0.396 - 0.275}{A}$$

$$C. \frac{0.275 - 0.396}{A}$$

$$D. \frac{A - 0.396}{0.5}$$

Ans: C; LO: 1.5-2; Difficulty: Easy; Type: MC

62. The standardized statistic is equal to -2.213 . Interpret this value in the context of the problem.

- A. On average, each simulated sample proportion of residents that fit the medical definition of obese is 2.213 standard deviations away from 0.396.
- B. The probability of obtaining a sample proportion of 0.275 or something more extreme, assuming 39.6% of all city residents fit the medical definition of obese, is 0.02213.
- C. The observed sample proportion of city residents that fit the medical definition of obese, 0.275, is 2.213 standard deviations below the hypothesized proportion 0.396.
- D. The hypothesized proportion of city residents that fit the medical definition of obese, 0.396, is 2.213 standard deviations below the sample proportion 0.275.

Ans: C; LO: 1.3-2; Difficulty: Hard; Type: MC

63. Once more, let 'A' represent your answer from #60. How should the p-value be calculated for this test? Select all that apply.

- A. Find the area below the standardized statistic on the Standard Normal distribution.
- B. Find the area below 0.275 on a Normal distribution centered at the null value with standard deviation A.
- C. Find the proportion of samples at or below 0.275 on a simulated null distribution.
- D. Find the area below A on a Standard Normal distribution.

Ans: A, B, C; LO: 1.5-2; Difficulty: Hard; Type: MS

64. The World Health Organization reports that the SARS (Severe Acute Respiratory Syndrome) outbreak in 2003 resulted in 8098 probable cases reported, 774 of which were fatalities. A student who is planning to travel to Vietnam is concerned about another virus outbreak and would like to use data on how Vietnam handled SARS cases as a gauge for how well its health care system would handle another virus outbreak. The student plans to test if the fatality rate in Vietnam from SARS is different from the global fatality rate. Of the 89 SARS cases reported in Vietnam in 2003, 5 people died. The student analyzes these data in two ways:

Method 1: The theory-based p-value was calculated to be 0.2061.

Method 2: The simulation-based p-value was calculated to be 0.2797.

Which p-value is most valid?

- A. The theory-based p-value, since the p-value is calculated by finding the exact area under a Normal distribution.
- B. The theory-based p-value, since there are over 10 “failures” (people who survived SARS).
- C. The simulation-based p-value, since there are only 5 “successes” (people who died from SARS).
- D. The simulation-based p-value, since it simulates thousands of samples.

Ans: C; LO: 1.5-3; Difficulty: Medium; Type: MC

Questions 65 through 68:

When President Trump took office in 2016, a Gallup poll reported that 55% of Americans supported stricter laws covering the sale of firearms. In 2020, the same poll was conducted on a representative sample of 943 Americans, and 64% of those surveyed supported stricter laws covering the sale of firearms. Has the percent of Americans supporting stricter laws covering the sale of firearms increased from 2016?

65. Which hypothesis testing method is appropriate for this study?

- A. Theory-based test
- B. Simulation-based test
- C. Either method is appropriate
- D. Neither method is appropriate

Ans: C; LO: 1.5-3; Difficulty: Medium; Type: MC

66. Fill in the appropriate inputs to the Theory-Based Inference applet to conduct a theory-based test for this scenario.

n: _____

count: _____

sample (\hat{p}): _____

Ans: $n = 943$ (Tol: 0), count = 604 (Tol: 0), sample $\hat{p} = 0.64$ (Tol: +/- 0.01); LO: 1.5-2; Difficulty: Easy; Type: TE-N

67. Fill in the blanks to enter the appropriate hypotheses into the Theory-Based Inference applet.

$H_0 : \pi =$ _____ (B1)

$H_s : \pi$ _____ (DD) _____ (B3)

Drop-down (DD) options:

- \neq
- $<$
- $>$

Ans: B1 = 0.55 (Tol: 0), B3 = 0.55 (Tol: 0), DD = $>$; LO: 1.5-2; Difficulty: Easy; Type: TE-N, Drop-down

68. Use the Theory-Based Inference applet to compute the standardized statistic and p-value for this study.

standardized statistic = _____

p-value = _____

Ans: standardized statistic = 5.59, p-value = 0.0000 (tol = 0); LO: 1.5-2; Difficulty: Easy; Type: TE-N

Questions 69 through 71:

True or False?

69. A simulated distribution of sample proportions will always be approximately normal.

Ans: False; LO: 1.5-1; Difficulty: Easy; Type: TF

70. If the sample size increases, the standard deviation of the null distribution of sample proportions will also increase.

Ans: False; LO: 1.5-1; Difficulty: Medium; Type: TF

71. If the sample size increases, the mean of the null distribution of sample proportions will decrease.

Ans: False; LO: 1.5-1; Difficulty: Easy; Type: TF