Chapter 8 Review

MULTIPLE CHOICE.

1) You are going to create a 95% confidence interval for a population proportion and want the margin of error to be no more than 0.05. Historical data indicates that the population proportion has remained constant at 0.7. What is the minimum size random sample you need to construct this interval?
   A) 385 B) 322 C) 274 D) 275 E) 323

2) Which of the following will increase the power of a test?
   A) Increase n
   B) Increase α
   C) Reduce the amount of variability in the sample
   D) Consider an alternative hypothesis further from the null
   E) All of these will increase the power of the test

3) Under a null hypothesis, a sample value yields a P-value of 0.015. Which of the following statements is (are) true?
   I. This finding is statistically significant at the 0.05 level of significance
   II. This finding is statistically significant at the 0.01 level of significance
   III. The probability of getting a sample value as extreme as this one by chance only if the null hypothesis is true is 0.015

   A) I and III only B) I only C) III only
   D) II and III only E) I, II, and III

4) A 95% confidence interval for the difference between two population proportions is found to be $0.07 < p < 0.19$. Which of the following statements is (are) true?
   I. It is unlikely that the two populations have the same proportion.
   II. We are 95% confident that the true difference between the population proportions is between 0.07 and 0.19.
   III. The probability is 0.95 that the true difference between the population proportions is between 0.07 and 0.19.

   A) I only B) II only C) I and II only
   D) I and III only E) II and III only

5) In a test of the null hypothesis H_0: $p = 0.35$ with $\alpha = 0.01$, against the alternative hypothesis H_A: $p < 0.35$, a large random sample produced a z-score of -2.05. Based on this, which of the following conclusions can be drawn?
   A) It is likely that $p < 0.35$.
   B) $p < 0.35$ only 2% of the time.
   C) If the z score were positive instead of negative, we would be able to reject the null hypothesis.
   D) We do not have sufficient evidence to claim that $p < 0.35$.
   E) 1% of the time we will reject the alternative hypothesis in error.

FREE RESPONSE.

6) You attend a large university with approximately 15,000 students. You want to construct a 90% confidence interval estimate, within 5%, for the proportion of students who favor outlawing country music. How large a sample do you need?

7) Two groups of 40 randomly selected students were selected to be part of a study on drop-out rates. Members of one group were enrolled in a counseling program designed to give them skill needed to succeed in school and the other group received no special counseling. Fifteen of the students who received counseling dropped out. Construct a 90% confidence interval for the true difference between the drop-out rates of the two groups. Interpret your answer in the context of the problem.

8) One researcher wants to construct a 99% confidence interval as part of a study. A colleague says such a high level isn't necessary and that a 95% confidence level will suffice. In what ways will these intervals differ?

9) A university is worried that it might not have sufficient housing for its students for the next academic year. It's very expensive to build additional housing, so it is operating under the assumption (hypothesis) that the housing is sufficient, and it will spend the money to build additional housing only if it is convinced it is necessary (that is, it rejects its hypothesis).
   a) For the university's assumption, what is the risk involved in making a Type I error?
   b) For the university's assumption, what is the risk involved in making a Type II error?

10) A flu vaccine is being tested for effectiveness. Three hundred fifty randomly selected people are given that vaccine and observed to see if they develop the flu during the flu season. At the end of the season, 55 of the 350 did get the flu. Construct and interpret a 95% confidence interval for the true proportion of people who will get the flu despite getting the vaccine.

11) A research study gives a 95% confidence interval for the proportion of subjects helped by a new anti-inflammatory drug as (0.56, 0.65).
   a) Interpret this interval in the context of the problem.
   b) What is the meaning of "95%" confidence interval as stated in the problem?

12) A study was conducted to see if attitudes toward travel have changed over the past year. In the prior year, 25% of American families took at least one vacation away from home. In a random sample of 100 families this year, 29 families took a vacation away from home. What is the p-value of getting a finding this different from expected?

13) You want to estimate the proportion of Californians who want to outlaw cigarette smoking in all public places. Generally speaking, by how much must you increase the sample size to cut the margin of error in half?

14) A hypothesis test is conducted with $\alpha = 0.05$ to determine the true difference between the proportion of male and female students enrolled in Statistics (H_0: $p_1 - p_2 = 0$). The P-value of $\hat{p}_1 - \hat{p}_2$ is determined to be 0.03. Is this finding *statistically significant*? Explain what is meant by a statistically significant finding in the context of the problem.

15) Based on the 2000 census, the population of the US was about 281.4 million people, and the population of Nevada was about 2 million. We are interested in generating a 95% confidence interval, with a margin of error of 3%, to estimate the proportion of people who will vote in the next presidential election. How much larger a sample will we need to generate this interval for the United States than for the state of Nevada?

16) The new reality TV show, “I Want to Marry a Statistician,” has been showing on Monday evenings, and ratings show that it has been watched by 55% of the viewing audience each week. The producers are moving the show to Wednesday night buy are concerned that such a move might reduce the percentage of the viewing public watching the show. After the show has been moved, a random sample of 500 people who are watching television on Wednesday night are surveyed and asked what show they are watching. Two hundred fifty-five respond that they are watching “I Want to Marry a Statistician.” Does this finding provide evidence at the 0.01 significance that the percentage of the viewing public watching “I Want to Marry a Statistician” has declined?

17) The director of a large metropolitan airport claim that security procedures are 98% accurate in detecting banned metal objects that passengers may try to carry onto a plane. The local agency charged with enforcing security thinks the security procedures are not as good as claimed. A study of 250 passengers showed that screeners missed nine banned carry-on items. What is the P-value for this test and what conclusion would you draw based on it?

18) An election is bitterly contested between two rivals. In a poll of 750 potential voters taken 4 weeks before the election, 420 indicated a preference for candidate Grumpy over candidate Dopey. Two weeks later, a new poll of 900 randomly selected potential voters found 465 who plan to vote for Grumpy. Dopey immediately began advertising that support for Grumpy was slipping dramatically and that he was going to win the election. Statistically speaking (say at the 0.05 level), how happy should Dopey be (i.e., how sure is he that support for Grumpy has dropped)?

SOLUTIONS

MULTIPLE CHOICE

1)

The correct answer is (e).

$P = 0.7$, $M = 0.05$, $z^* = 1.96$ (for $C = 0.95$) $\Rightarrow$

$n \geq \left(\frac{z^*}{M}\right)^2 (P^*)(1 - P^*) = \left(\frac{1.96}{0.05}\right)^2 (0.7)(0.3) = 322.7$. You need a sample of at least $n = 323$.

2)

The correct answer is (e).

3)

The correct answer is (a). It is not significant at the .01 level because .015 is greater than .01.

4)

The correct answer is (c). Because 0 is not in the interval (0.07, 0.19), it is unlikely to be the true difference between the proportions. III is just plain wrong! We cannot make a probability statement about an interval we have already constructed. All we can say is that the process used to generate this interval has a 0.95 chance of producing an interval that does contain the true population proportion.

5)

The correct answer is (d). To reject the null at the 0.01 level of significance, we would need to have $z < -2.33$.

FREE RESPONSE

6)

$C = 0.90 \Rightarrow z^* = 1.645, M = 0.05$. $n \geq \left(\frac{1.645}{2(0.05)}\right)^2 = 270.6$.

You would need to survey at least 271 students.

7)

This is a two-proportion situation. We are told that the groups were randomly selected, but we need to check that the samples are sufficiently large:

$$\hat{p}_1 = \frac{15}{40} = 0.375, \hat{p}_2 = \frac{23}{40} = 0.575.$$

$$n_1\hat{p}_1 = 40(0.375) = 15, n_1(1-\hat{p}_1) = 40(1-0375) = 25.$$

$n_2\hat{p}_2 = 40(0.575) = 23$, $n_2(1-\hat{p}_2) = 40(1-0.575) = 17$. Since all values are greater than or equal to 5, we are justified in constructing a two-proportion z interval. For a 90% z confidence interval, $z^* = 1.645$.

Thus, $(0.575-0375) \pm 1.645\sqrt{\frac{(0.375)(1-0.375)}{40} + \frac{(0.575)(1-0.575)}{40}} = (0.02, 0.38)$.

We are 90% confident that the true difference between the dropout rates is between 0.02 and 0.38. Since 0 is not in this interval, we have evidence that the counseling program was effective at reducing the number of dropouts.

8)

The 99% confidence interval will be more likely to contain the population value being estimated, but will be wider than a 95% confidence interval.

9)

(a) A Type-I error is made when we mistakenly reject a true null hypothesis. In this situation, that means that we would mistakenly reject the true hypothesis that the available housing is sufficient. The risk would be that a lot of money would be spent on building additional housing when it wasn't necessary.

(b) A Type-II error is made when we mistakenly fail to reject a false hypothesis. In this situation that means we would fail to reject the false hypothesis that the available housing is sufficient. The risk is that the university would have insufficient housing for its students.

10)

(a) We are 95% confident that the true proportion of subjects helped by a new anti-inflammatory drug is (0.56, 0.65).

(b) The process used to construct this interval will capture the true population proportion, on average, 95 times out of 100.

11)

We have $H_0: p = 0.25, H_A: p \neq 0.25, \hat{p} = \frac{29}{100} = 0.29$. Because the hypothesis is two sided, we are concerned about the probability of being in *either* tail of the curve even though the finding was larger than expected.

$$z = \frac{0.29 - 0.25}{\sqrt{\frac{(0.25)0.75}{100}}} = 0.92 \Rightarrow \text{upper tail area}$$

$$= 1 - 0.8212 = 0.1788 \text{ (using Table A).}$$

Then, the P-value $= 2(0.1788) = 0.3576$. Using the TI-83/84, the P-value = `2normalcdf(0.92,100)`.

12)

For a given margin of error using $P^* = 0.5$:

$$n \geq \left(\frac{z^*}{2M}\right)^2.$$

To reduce the margin of error by a factor of 0.5, we have

$$n^* - \left(\frac{z^*}{2\left(M/2\right)}\right)^2 - \left(\frac{2z^*}{2M}\right)^2 - 4\left(\frac{z^*}{2M}\right)^2 - 4n.$$

We need to quadruple the sample size to reduce the margin of error by a factor of $1/2$.

13)

The finding is statistically significant because the P-value is less than the significance level. In this situation, it is unlikely that we would have obtained a value of $\hat{p}_1 - \hat{p}_2$ as different from 0 as we got by chance alone if, in fact, $\hat{p}_1 - \hat{p}_2 = 0$.

14)

The finding is statistically significant because the P-value is less than the significance level. In this situation, it is unlikely that we would have obtained a value of $\hat{p}_1 - \hat{p}_2$ as different from 0 as we got by chance alone if, in fact, $\hat{p}_1 - \hat{p}_2 = 0$.

15)

Trick question! The sample size needed for a 95% confidence interval (or any C-level interval for that matter) is not a function of population size. The sample size needed is given by

$$n \geq \left(\frac{z^*}{M}\right)^2 P^*(1-P^*).$$

n is a function of z^* (which is determined by the confidence level), M (the desired margin of error), and P^* (the estimated value of $\hat{p}$). The only requirement is that the population size be at least 20 times as large as the sample size.

16)

I. Let p = the true proportion of Wednesday night television viewers who are watching "I Want to Marry a Statistician."

H_0: $p = 0.55$.
H_A: $p < 0.55$.

II. We want to use a one-proportion *z-test* at $\alpha = 0.01$. $500(0.55) = 275 > 5$ and $500(1 - 0.55) = 225 > 5$. Thus, the conditions needed for this test have been met.

III. $\hat{p} = \dfrac{255}{500} = 0.51.$

$$z = \frac{0.51-0.55}{\sqrt{\dfrac{0.55(1-0.55)}{500}}} = -1.80 \Rightarrow P\text{-value} = 0.036.$$

(On the TI-83/84, `normalcdf(-100,-1.80)=0.0359`, or you can use `STAT TESTS 1-PropZTest.`)

IV. Because $P > 0.01$, we do not have sufficient evidence to reject the null hypothesis. The evidence is insufficient to conclude at the 0.01 level that the proportion of viewers has dropped since the program was moved to a different night.

17)

H_0: $p \geq 0.98$, H_A: $p < 0.98$, $\hat{p} = \dfrac{241}{250} = 0.964.$

$$z = \frac{0.964-0.98}{\sqrt{\dfrac{(0.98)(0.02)}{250}}} = -1.81,\ P\text{-value} = 0.035.$$

This P-value is quite low and provides evidence against the null and in favor of the alternative that security procedures actually detect less than the claimed percentage of banned objects.

18)

I. Let p_1 = the true proportion of voters who plan to vote for Grumpy 4 weeks before the election.

Let p_2 = the true proportion of voters who plan to vote for Grumpy 2 weeks before the election.

H_0: $p_1 - p_2 = 0$.
H_A: $p_1 - p_2 > 0$.

II. We will use a two-proportion z-test for the difference between two population proportions. Both samples are random samples of the voting populations at the time.

$$\hat{p}_1 = \frac{420}{750} = 0.56,\ \ \hat{p}_2 = \frac{465}{900} = 0.517.$$

Also,

$$n_1\hat{p}_1 = 750(0.56) \approx 420,$$
$$n_1(1-\hat{p}_1) = 750(0.44) \approx 330,$$
$$n_2\hat{p}_2 = 900(0.517) \approx 465,$$
$$n_2(1-\hat{p}_2) = 900(0.483) \approx 435.$$

All values are larger than 5, so the conditions needed for the two-proportion z-test are present.

III. $$\hat{p} = \frac{420+465}{750+900} = \frac{885}{1650} = 0.54,$$

$$z = \frac{0.56 - 0.517}{\sqrt{0.54(0.46)\left(\frac{1}{750} + \frac{1}{900}\right)}} = 1.75 \Rightarrow P\text{-value} = 0.04.$$

(From the TI-83/84, `STAT TESTS 2-PropZTest` yields P-value = 0.039.)

IV. Because $P < 0.05$, we can reject the null hypothesis. Candidate Dopey may have cause for celebration—there is evidence that support for candidate Grumpy is dropping.