AP Statistics Chapter 7-9 Test Study Guide

PART I – CONCEPTS AND VOCABULARY

$$Confidence\ Interval = statistic\ \pm (critical\ value) * (standard\ deviation\ of\ statistic)$$

1. Write out the general formula for a confidence interval for:
   - A proportion: $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$
   - A mean: $\bar{x} \pm t^* \left(\frac{s_x}{\sqrt{n}}\right)$
2. In your own words, define the Central Limit Theorem: As sample size increases, the sampling distribution becomes approximately normal.
3. If I wanted to manipulate the margin of error, what would I have to change? Sample size
   a. If I wanted to **decrease** the margin of error, HOW would I have to change (the above answer)? Increase sample size
4. What does the p-value measure? Assuming the null hypothesis is true, this is the probability that we'd get a statistic ($\bar{x}$ or $\hat{p}$) as small/big by chance alone.

PART II – WHAT TO USE AND WHEN

Word Bank:

Confidence Interval for Means	Confidence Interval for Proportions
One Proportion *z* Test	T-Test
Matched Pairs *t* Test	

1. I want to construct a 95% confidence interval for the true proportion *p* of all first-year students at a university who would identify being well-off as an important personal goal. Confidence Interval for Proportions
2. Jason reads a report that says 80% of U.S. high school students have a computer at home. He believes the proportion is smaller than 0.80 at his large rural high school. Jason chooses an SRS of 60 students and record whether they have a computer at home. One-proportion z test
3. Researchers suspect that Variety A tomato plants have a higher average yield than Variety B. To find out, they randomly select 10 Variety A and 10 Variety B plants. Then, they compare the yield in pounds for the plants. The 10 differences in yield (Variety A – Variety B) are recorded. Matched pairs t test
4. A college professor suspects that students at his school are getting less than 8 hours of sleep a night, on average. To test his belief, the professor asks a random sample of 28 students, "How much sleep did you get last night?" T-test
5. Debra wants to construct a 90% confidence interval for the average number of texts students send during her class. Confidence Interval for Means

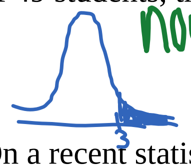
PART III – CHAPTER 7 REVIEW

1. In a certain year, ACT scores had a mean of 18 and a standard deviation of 6. If a simple random sample of 36 students were chosen, what would be the mean and standard deviation of the sampling distribution from which the sample was taken?

   Mean 18 Standard Deviation $6/\sqrt{36} = 6/6 = 1$

2. Last year, Ms. Gunz wanted to figure out what the mean score would be on the AP Statistics test. She gathered data from 49 students in her classes. Suppose the standard deviation of the sampling distribution was 0.25. If the mean score on the AP test was 2.4, what is the probability that in a sample of 49 students, the mean score would be at least a 3?

   don't have to change!

   normalcdf(3, 1E99, 2.4, 0.25) = 0.0082

   normalcdf (49 ≥ 30 ✓)

3. On a recent statistics exam at a large university, 35 of 50 students in one section passed the exam. What is the mean and standard deviation of the proportion of those who passed the exam?

   $\hat{p} = \frac{35}{50} = 0.7$ $\mu_{\hat{p}} = 0.7$ $\sigma_{\hat{p}} = \sqrt{\frac{(.7)(.3)}{50}} = 0.0648$

4. List the conditions for a sample proportion:
   a. Random
   b. 10% → 10n < N
   c. Large Counts → np ≥ 10, np ≥ 10

   n = sample size
   p = proportion

5. List the conditions for a sample mean:
   a. Random
   b. 10%
   c. Large Counts → n ≥ 30 (or distribution must look normal)

PART IV – CHAPTER 8 REVIEW

1. What is the critical value needed to construct a confidence interval about a population proportion at a 94% confidence level based on an SRS of size 125?

   z^*

   area: 1 − .94 = 0.06

   0.06/2 = 0.03

   z^* → invNorm(0.03, 0, 1) = 1.88

2. What is the critical value needed to construct a confidence interval about a population mean at a 99% confidence level based on an SRS of size 58?

   area: 1 − .99 = 0.01 — t^*

   0.01/2 = 0.005

   t^* → invT(0.005, df = 57) = 2.66

3. Interpret the following confidence interval: 90% confidence interval for an unknown population mean has been calculated to be 87 ± 12. → (75, 99)

   We are 90% confident that the interval from 75 to 99 captures the true mean of . . .

z^* because we know population st. dev.

4. Fifth grader Scott is studying the heights of American woman for a science fair project. Scott's teacher told him that the heights of American females are normally distributed with the population standard deviation of 2.4 inches, but he cannot remember the value for the mean. His older brother suggests that he estimate the unknown mean by calculating a confidence interval using data that Scott has collected for the heights of four women, as follows: 63 in, 69 in, 62 in, 66 in. Based on this, what would Scott's 99% confidence interval be?

find $\bar{x}$. $\frac{63+69+62+66}{4}$ $\bar{x}=65$

z^*: invNorm(.005,0,1)

$$65 \pm 2.58\left(\frac{2.4}{\sqrt{4}}\right) \rightarrow (61.495, 68.505)$$

5. High school students who take the SAT Math exam a second time generally score higher than on their first try. Past data suggest that the score increase has a standard deviation of about 50 points. How large a sample of high school students would be needed to estimate the mean change in SAT score to within 2 points with 95% confidence? Show your work.

$$z^*\left(\sigma/\sqrt{n}\right) \leq ME$$

$$\frac{1.96\left(50/\sqrt{n}\right)}{1.96} \leq \frac{2}{1.96}$$

$$\sqrt{n} \cdot \frac{50}{\sqrt{n}} \leq 1.0204\sqrt{n}$$

$$\frac{50}{1.0204} \leq \frac{1.0204\sqrt{n}}{1.0204}$$

$$(49.0003)^2 \leq (\sqrt{n})^2$$

$$2401.04 \leq n$$

ABOUT 2402 Students

6. A flu vaccine is being tested for effectiveness. To test this, 350 randomly selected people are given the vaccine and observed to see if they develop the flu during flu season. At the end of the season, 55 of the 350 did get the flu. $\hat{p} = \frac{55}{350} = 0.157$

a. Construct a 92% confidence interval to estimate the proportion of people who will get the flu despite the vaccine. z^*

invNorm(0.04,0,1) = 1.75

$$0.157 \pm 1.75\left(\sqrt{\frac{0.157(0.843)}{350}}\right) \rightarrow (0.1231, 0.1912)$$

(use 1-prop z int in calc)

b. Suppose the researchers conducting the study want to cut the margin of error in half while maintaining the 92% confidence level. What should they do?

Cut by $\frac{1}{2}$ → everything is under $\sqrt{\ }$ → $\frac{1}{2} = \sqrt{\frac{1}{4}}$ → Therefore, we need the sample size to be QUADRUPLED

PART V – CHAPTER 9 REVIEW

1. Consider a screening test for liver disease that its maker claims will detect the disease in 85% of patients who actually have the disease. One hundred seventy-five patients who have been previously diagnosed with liver disease are given the screening test, and 140 of the patients are identified as having the disease. Does this finding provide evidence that the screening test detects the disease at a lower rate than the 85% rate claimed by its manufacturer?

a. State the null and alternative hypothesis of interest. (Don't forget to define your parameter!)

$H_0: p = 0.85$

$H_a: p < 0.85$

where p = true proportion patients correctly diagnosed by machine.

b. Check that conditions for inference are met.

Random → does not state but must assume randomness to continue

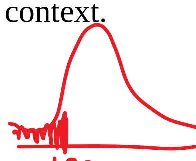
10% → 10(175) = 1750 < all patients who have said disease

LC → 175(.85) = 148.75 ≥ 10 and 175(.15) = 26.25 ≥ 10

c. Calculate and interpret the *p*-value in context.

$\hat{p} = \frac{140}{175} = 0.8$

$z = \frac{0.8 - 0.85}{\sqrt{\frac{.85(.15)}{175}}} = -1.85$



$P(z \leq -1.85) = 0.032$

Assuming Ho is true, the probability of getting $\hat{p} = 0.8$ is 0.032.

d. State the conclusion at the $\alpha = 0.05$ significance level.

Since our p-val of 0.032 is less than $\alpha = 0.05$, we reject Ho. There is convincing evidence that the true proportion of people detected is less than 0.85.

e. Suppose instead the alternative hypothesis was 2-sided. How would this affect the *p*-value of the test and your conclusion at the $\alpha = 0.05$ significance level?

We'd have to multiply the p-value by 2 → 0.038(2) = 0.076

2. A pharmaceutical company claims that no more than 8% of patients who take a new drug for treatment of high blood pressure will experience dizziness or fatigue. After receiving complaints about these side effects from consumers who have used the drug, a consumer advocacy group conducts an experiment. They give the drug to a random sample of 200 patients who have been diagnosed with high blood pressure. In all, 22 of the subjects experience dizziness and fatigue.

a. State the null and alternative hypothesis of interest. (Don't forget to define your parameter!)

$H_0: p = 0.08$

$H_a: p > 0.08$

where p = true proportion experiencing symptoms

b. Describe a Type I error in this setting, and describe a consequence of this error.

Reject Ho when Ho is true → Say more people are experiencing symptoms when there is no increase.

Consequence → Spend money on either improving drug or updating marketing claims when you don't have to.

c. Describe a Type II error in this setting, and describe a consequence of this error.

fail to reject Ho when H_A is true → Say no more than 8% experience symptoms when, in fact, more than 8% did experience symptoms

Consequence → false advertisment

(and people are uncomfortable!)