Chapter 9: Testing A Claim

"A statistical analysis, properly conducted, is a delicate dissection of uncertainties, a surgery of suppositions." M.J. Moroney

Chapter Overview

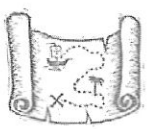
In the last chapter, you learned how to estimate a parameter based on a sample statistic. In this chapter, you will learn how to use data from a random sample to test a claim about a population parameter. Significance tests are the second type of inference you will be introduced to in this course. By the end of this chapter, you should understand the logic behind a significance test, as well as how to use a four-step procedure to carry one out. Specific methods for testing a claim about a proportion and a mean will be explored. In the next few chapters, you'll learn how to test claims about the difference between proportions or means, as well as how to perform significance tests for slopes of regression lines and distributions of categorical variables. In each case, you will use the same logic and the same four-step procedure. Be sure to take the time in this chapter to get a solid understanding of both!

Sections in this Chapter

Section 9.1: Significance Tests: The Basics
Section 9.2: Tests About a Population Proportion
Section 9.3: Tests About a Population Mean

Plan Your Learning

Use the following *suggested* guide to help plan your reading and assignments. Note: your teacher may schedule a different pacing. Be sure to follow his or her instructions!



Read	9.1: pp 527-537	9.1: pp 537-545	9.2: pp 549-555	9.2: pp 556-562
Do	1, 3, 5, 7, 9, 11, 13	15, 19, 21, 23, 25	27-30, 41, 43, 45	47, 49, 51, 53, 55

Read	9.3: pp 565-576	9.3: pp 577-587	Chapter Summary
Do	57-60, 71, 73	75, 77, 89, 94-97, 99-104	Multiple Choice FRAPPY!

Section 9.1: Significance Tests: The Basics

Before You Read: Section Summary

In this section, you will learn the basic ideas and logic behind a significance test. You will be introduced to stating hypotheses, checking conditions, calculating a test statistic, and drawing a conclusion. *P*-values will be introduced as a means of weighing the strength of evidence against a claim. Since we are basing our conclusion on a sample statistic that would vary from sample to sample, we must be aware of the fact that our conclusion could be wrong. This section ends with a discussion of the types of errors that could occur when performing a significance test and how to deal with them. Be sure to get a good understanding of the logic and format for a significance test. You will be using them throughout the rest of the course!

"Where Am I Going?"

Learning Targets:

____ I can state correct hypotheses for a significance test about a population proportion or mean.
____ I can interpret *P*-values in context.
____ I can interpret a Type I error and a Type II error in context, and give the consequences of each.
____ I can describe the relationship between the significance level of a test, *P*(Type II error), and power.

While You Read: Key Vocabulary and Concepts

significance test:

null hypothesis:

alternative hypothesis:

one-sided:

two-sided:

P-value:

reject H_0:

fail to reject H_0:

statistically significant:

Type I error:

Type II error:

power:

After You Read: "Where Am I Now?" Check for Understanding

Concept 1: The Reasoning Behind Significance Tests

A significance test is a procedure that allows us to test a claim about a population parameter by studying a statistic from a random sample. If an observed statistic is "far" away from a hypothesized claim about a parameter, we have some evidence that the claim is wrong. Whether or not the statistic is "far enough" away depends on the sampling distribution of the statistics. That is, statistics will vary from sample to sample. What a significance test does is answers the question, "What are the chances we would observe a sample statistic at least this extreme, assuming the claim about the parameter was true?" If the chances are low, there is evidence that the claim may be wrong. If the chances are pretty good, then there is little evidence to suggest the claim is wrong. Of course, to even begin making that argument, you must start by setting up a null hypothesis (or claim about the parameter you are trying to find evidence against) and an alternative hypothesis (the claim about the parameter you are trying to find evidence for). Note: the null hypothesis will always be a statement of equality. That is, it says "Let's assume the parameter is equal to _____." The alternative hypothesis is set up to test whether or not the actual value of the parameter is greater than, less than, or simply not equal to the value in the null hypothesis.

Check for Understanding: ____ *I can state correct hypotheses for a significance test about a population proportion or mean.*

Suppose you suspect a "chute" of playing cards is not fair. The chute supposedly contains 10 standard decks shuffled together. You are interested in knowing whether there are more hearts than usual. To test this, you deal 12 cards at random and calculate the proportion of hearts in your hand.

a) Describe the parameter of interest in this setting.

b) Write the appropriate null and alternative hypotheses for this situation.

c) Suppose your deal contains 7 hearts. What is the sample proportion of hearts? Is it possible to deal 7 hearts out of 12 cards if the chute really does contain standard decks? Is it likely? Why or why not?

Concept 2: Statistical Significance

To conclude that we have convincing evidence against a null hypothesis, our sample statistic must have a low likelihood of occurring under the assumption that the null hypothesis is true. To determine this, we must consider the sampling distribution of the statistic under the assumption that the null hypothesis is true and calculate the probability of observing a statistic at least as extreme as the one observed. If the conditions for inference are met, we can describe the sampling distribution and calculate a test statistic. This test statistic can then be used to determine a *P*-value. If this *P*-value is low enough, we can say the data are statistically significant and we have evidence to reject the null hypothesis. The general rule of thumb is to use a significance level of 5%, although some situations may specify levels of 1% or even 10%. If the *P*-value is smaller than the chosen significance level, we have enough evidence to reject the null hypothesis and conclude that the alternative is true. The reasoning behind this is that if we observe an outcome that has an extremely low chance of occurring under a given assumption, then there is evidence that the assumption may be false. If we observe an outcome under that has a fairly likely chance of occurring under a given assumption, then we have little reason to doubt the assumption!

Check for Understanding: ____ *I can interpret* P*-values in context.*

Refer to the hypotheses you set in the previous check for understanding. Suppose you conduct a significance test for the proportion of hearts in the chute and find the *P*-value for obtaining 7 hearts in 12 randomly selected cards is 0.004.

a) Explain what it means for the null hypothesis to be true in this setting.

b) Interpret the *P*-value in context.

c) Do these data provide convincing evidence against the null hypothesis at a 5% significance level? Explain.

Concept 3: Type I and Type II Errors

Since our conclusion is based on what we would expect to see in a sampling distribution, there is a chance our sample statistic does not accurately reflect the true value of the parameter. That is, due to sampling variability, it is possible that the random sample that is collected will lead to an incorrect conclusion about the parameter. If we obtain a very small *P*-value and reject the null when, in fact, the null hypothesis is true, we commit a Type I error. The probability of committing a Type I error is equal to the significance level, α. If we obtain a relatively large *P*-value and fail to reject the null when, in fact, the null hypothesis is false, we commit a Type II error. The power of a significance test is the probability that it will correctly reject the null

hypothesis when a given alternative value of the parameter is true. You will not be expected to calculate the probability of Type II errors. However, you should be able to interpret both Type I and Type II errors in the context of the situation and explain the consequences of each.

Understanding: ____ *I can interpret a Type I error and a Type II error in context, and give the consequences of each and* ____ *I can describe the relationship between the significance level of a test, P(Type II error), and power.*

Refer to the previous checks for understanding.

a) Describe a Type I error in this setting.

b) Describe a Type II error in this setting.

Section 9.2: Tests About a Population Proportion

Before You Read: Section Summary

In this section, you will apply the logic of significance tests and the four-step process to test a claim about a population proportion. You will learn how to check the conditions for a significance test as well as how to calculate a test-statistic and *P*-value. Finally, you will learn how to write an appropriate conclusion in context. Your study will also include an introduction to two-sided tests to find evidence that a population proportion is different than a hypothesized value. You will then wrap up the section by studying the connection between confidence intervals and two-sided tests. Again, this section introduces you to concepts that will be used throughout the rest of the course. Make sure to get additional help if you struggle at all with the ideas behind significance testing!

"Where Am I Going?"

Learning Targets:

____ I can check conditions for carrying out a test about a population proportion.
____ I can conduct a significance test about a population proportion.
____ I can use a confidence interval to draw a conclusion for a two-sided test about a population proportion.

While You Read: Key Vocabulary and Concepts

test statistic:

significance tests – four step process:

one-sample *z* test for a proportion:

After You Read: "Where Am I Now?" Check for Understanding

Concept 1: One-Sample z Test for a Proportion

Significance tests and confidence intervals are both based on the same concept—sampling distributions. Therefore, when performing inference with either of these methods, it is important to follow a process that ensures the calculations are justified. To test a claim about a population parameter, the following four-step process should be used:

1) **STATE the parameter of interest and the hypotheses you would like to test.**
When testing a claim about a population proportion, we start by defining hypotheses about the parameter. The null hypothesis assumes that the population proportion is equal to a particular value while the alternative hypothesis is that the population proportion is greater than, less than, or not equal to that value:

$H_0: p = p_0$
$H_a: p > p_0$ OR $p < p_0$ OR $p \neq p_0$

State a significance level.

2) **PLAN: Choose the appropriate inference method and check the conditions.**
We must check the Random, Normal, and Independent conditions before proceeding with inference. To ensure that the sampling distribution of $\hat{p}$ is approximately Normal, check to make sure np_0 and $n(1 - p_0)$ are both at least 10. If sampling is done without replacement, verify that the population is at least 10 times as large as the sample.

3) **DO: If conditions are met, calculate a test statistic and *P*-value.**
Compute the *z* test statistic

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

and find the *P*-value by calculating the probability of observing a *z* statistic at least this extreme in the direction of the alternative hypothesis.

4) **CONCLUDE by interpreting the results of your calculations in the context of the problem.**
If the *P*-value is smaller than the stated significance level, you have significant evidence to reject the null hypothesis. If it is greater than or equal to the significance level, then you fail to reject the null hypothesis.

This four-step procedure can be used with *any* significance test! The only things that will change are the conditions that must be checked and the test statistic calculation. Be sure to get a good grasp of this process here as you will be using it a lot in the coming chapters!

Check for Understanding: ____ *I can check conditions for carrying out a test about a population proportion and* ____ *I can conduct a significance test about a population proportion.*

A study of classic authors uncovered a distinguishable speech pattern that differed from author to author. Plato utilized this pattern in 21.4% of the passages in his works. The owner of a rare bookstore claims to have an original Plato work, but you suspect the speech pattern occurs too frequently to be an original Plato work. A random sample of passages from the work in question was taken and it was found that 136 of the 439 selected passages followed the speech pattern. Do these data provide convincing evidence that the work was not written by Plato? Conduct a one-sample *z* test using the four-step procedure.

Concept 2: Confidence Intervals Give More Information

Significance tests provide evidence that supports or rejects an assumption about a population parameter. If an observed sample statistic is far enough away from a parameter's assumed value, a significance test will suggest there is evidence to question that assumed value. However, the test does not say anything about what the actual parameter value may be. A confidence interval provides more information by suggesting a range of plausible values for the parameter.

Check for Understanding: ____ *I can use a confidence interval to draw a conclusion for a two-sided test about a population proportion.*

A recent study suggested that 77% of teenagers have texted while driving. A random sample of 27 teenage drivers in Atlanta was taken and 15 admitted to texting while driving. Use a 99% confidence interval to determine whether there is convincing evidence that the population proportion of teens who text while driving is different than 77%.

Section 9.3: Tests About a Population Mean

Before You Read: Section Summary

In this section, you'll learn how to perform a significance test about a population mean μ. Just like you did for proportions, you'll learn how to set up hypotheses, check the appropriate conditions, calculate a test statistic and *P*-value, and draw a conclusion in context. You will also learn how tests involving "paired data" can be performed using one-sample *t* procedures.

"Where Am I Going?"

Learning Targets:

____ I can check conditions for carrying out a test about a population mean.
____ I can conduct a one-sample *t* test about a population mean μ.
____ I can use a confidence interval to draw a conclusion for a two-sided test about a population mean.
____ I can recognize paired data and use one-sample *t* procedures to perform significance tests for such data.

While You Read: Key Vocabulary and Concepts

one-sample *t* test:

paired data:

paired *t* procedures:

After You Read: "Where Am I Now?" Check for Understanding

Concept 1: One Sample t Test for μ

The process for testing a claim about a population mean follows the same format as the process used for a population proportion. However, like confidence intervals for means, you usually need to base your calculations on the *t* distributions (unless σ is somehow known).

1) **STATE the parameter of interest and the hypotheses you would like to test.**
   The null hypothesis assumes that the population mean is equal to a particular value while the alternative hypothesis is that the population mean is greater than, less than, or not equal to that value:
   $H_0: \mu = \mu_0$
   $H_a: \mu > \mu_0$ OR $\mu < \mu_0$ OR $\mu \neq \mu_0$
   State a significance level.

2) **PLAN: Choose the appropriate inference method and check the conditions.**
   We must check the Random, Normal, and Independent conditions before proceeding with inference. To ensure that the sampling distribution of $\bar{x}$ is approximately Normal, check that the population distribution is Normal OR $n \geq 30$. If $n < 30$, it is sufficient if the sample data show no signs of strong skewness or outliers. If sampling is done without replacement, verify that the population is at least 10 times as large as the sample.

3) **DO: If conditions are met, calculate a test statistic and *P*-value.**

Compute the *t* test statistic

$$t = \frac{\bar{x} - \mu_0}{\frac{s_x}{\sqrt{n}}}$$

and find the *P*-value by calculating the probability of observing a *t* statistic at least this extreme in the direction of the alternative hypothesis in a *t* distribution with $n - 1$ degrees of freedom.

4) **CONCLUDE by interpreting the results of your calculations in the context of the problem.**

If the *P*-value is smaller than the stated significance level, you have significant evidence to reject the null hypothesis. If it is larger than the significance level, then you fail to reject the null hypothesis.

As with proportions, a confidence interval for a population mean can be used to test a two-sided claim. Further, constructing a confidence interval gives a range of plausible values for μ, while a significance test only allows you to conclude that μ may be different from a particular value.

Check for Understanding: ____ *I can check conditions for carrying out a test about a population mean and* ____ *I can conduct a one-sample* t *test about a population mean* μ.

Humerus bones from the same species of animal tend to have approximately the same length-to-width ratios. When fossils of humerus bones are discovered, archeologists can often determine the species of animal by examining these ratios. It is known that the species Molekius Primatium exhibits a mean ratio of $\mu = 8.9$. Suppose 41 fossils of humerus bones are unearthed at a site on Minnesota's Iron Range, where this species was known to have lived. Researchers are willing to view these as a random sample of all such humerus bones. The length-to-width ratios were calculated and are listed below. Test whether the population mean for the species that left these bones differs from 8.9 at α=.05.

9.73	10.89	9.07	9.2	9.33	9.98	9.84	9.59	9.48	8.71	9.57	9.29
9.94	8.07	8.37	6.85	8.52	8.87	6.23	9.41	6.66	9.35	8.86	9.93
8.91	11.77	10.48	10.39	9.39	9.17	9.89	8.17	8.39	8.8	10.02	8.38
11.67	8.3	9.17	12	9.38							

Concept 2: Paired Data

Studies that involve making two observations on the same individual or making an observation on each of two very similar individuals result in paired data. In these types of studies, we are often interested in analyzing the differences in responses within each pair. If the conditions for inference are met, we can use one-sample *t* procedures to estimate or test a claim about the mean difference μ_d. In this case, we refer to the inference method as a paired *t* procedure.

Check for Understanding: ____ *I can construct and interpret a confidence interval for a population mean.*

A study measured how fast subjects could repeatedly push a button when under the effects of caffeine. Subjects were asked to push a button as many times as possible in two minutes after consuming a typical amount of caffeine. During another test session, they were asked to push the button after taking a placebo. The subjects did not know which treatment they were administered each day and the order of the treatments was randomly assigned. The data, given in presses per two minutes for each treatment follows. Use a paired *t* procedure to determine whether or not caffeine results in a higher rate of beats, on average, per two-minute period.

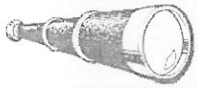
Subject	Beats Caffeine	Beats Placebo	
1	281	201	
2	284	262	
3	300	283	
4	421	290	
5	240	259	
6	294	291	
7	377	354	
8	345	346	
9	303	283	
10	340	391	
11	408	411	

Chapter Summary: Testing a Claim

A significance test tells us whether or not a sample provides convincing evidence against a claim about a population parameter. The test answers the question, "How likely would it have been to observe this particular sample statistic (or one more extreme) if the claim about the parameter was true?" This probability, the *P*-value, gives us an idea just how surprising an observed statistic is under the assumption that the null hypothesis is true. If the *P*-value is small (less than 5% or some other chosen significance level), we have enough evidence to reject the null hypothesis, suggesting the actual parameter value may be greater than, less than, or different from the claim. If the *P*-value is greater than or equal to a specified significance level, we do not have convincing evidence against the claim about the parameter. Keep in mind, however, that since we are basing our decision on a likelihood from a sampling distribution, it is possible that we may be wrong. Type I and Type II errors occur when we mistakenly reject a null hypothesis that is true or fail to reject a null hypothesis that is false, respectively. You should be able to define each of these errors in the context of the situation as well as explain the consequences of each.

This chapter not only introduced you to the logic behind significance tests, but also a four-step procedure for carrying them out. This procedure will be used throughout the rest of the course, so hopefully you are comfortable with it at this point! If not, be sure to practice a few more tests and focus on stating the parameter and hypotheses, checking the appropriate conditions, calculating a test statistic and *P*-value, and concluding in the context of the problem. If you follow those steps, you should have very few problems performing significance tests in the upcoming chapters!

After You Read: "How Can I Close the Gap?"



Complete the vocabulary puzzle, multiple choice questions, and FRAPPY. Check your answers and on your performance on each of the targets.

Target	Got It!	Almost There	Needs Some Work
I can state correct hypotheses for a significance test about a population proportion or mean.			
I can interpret *P*-values in context.			
I can interpret a Type I error and a Type II error in context, and give the consequences of each.			
I can check conditions for carrying out a test about a population proportion.			
I can conduct a significance test about a population proportion.			
I can use a confidence interval to draw a conclusion for a two-sided test about a population proportion.			
I can check conditions for carrying out a test about a population mean.			
I can conduct a one-sample *t* test about a population mean μ.			
I can use a confidence interval to draw a conclusion for a two-sided test about a population mean.			
I can recognize paired data and use one-sample *t* procedures to perform significance tests for such data.			

Did you check "Needs Some Work" for any of the targets? If so, what will you do to address your needs for those targets?

Learning Plan:

Chapter 9 Multiple Choice Practice

Directions. *Identify the choice that best completes the statement or answers the question. Check your answers and note your performance when you are finished.*

1. The average yield of a certain crop is 10.1 bushels per plant. A biologist claims that a new fertilizer will result in a greater yield when applied to the crop. A random sample of 25 of plants given the fertilizer has an average yield of 10.8 bushels and a standard deviation of 2.1 bushels. The appropriate null and alternative hypotheses to test the biologist's claim are

A. H_0: $\mu = 10.8$ against H_a: $\mu > 10.8$.
B. H_0: $\mu = 10.8$ against H_a: $\mu \neq 10.8$.
C. H_0: $\mu = 10.1$ against H_a: $\mu > 10.1$.
D. H_0: $\mu = 10.1$ against H_a: $\mu < 10.1$.
E. H_0: $\mu = 10.1$ against H_a: $\mu \neq 10.1$.

2. An opinion poll asks a random sample of 200 adults how they feel about voting for an amendment in an upcoming election. In all, 150 say they are in favor of the amendment. Does the poll provide evidence that the proportion p of adults who are in favor of the amendment is greater than 60%? The null and alternative hypotheses are

A. H_0: $p = 0.6$ against H_a: $p > 0.6$.
B. H_0: $p = 0.6$ against H_a: $p \neq 0.6$.
C. H_0: $p = 0.6$ against H_a: $p < 0.6$.
D. H_0: $p = 0.6$ against H_a: $p = 0.75$.
E. H_0: $p = 0.75$ against H_a: $p < 0.6$.

3. A test of significance produces a P-value of 0.035. Which of the following conclusions is appropriate?

A. Accept H_a at the $\alpha = 0.05$ level
B. Reject H_a at the $\alpha = 0.01$ level
C. Fail to reject H_0 at the $\alpha = .05$ level
D. Reject H_0 at the $\alpha = 0.05$ level
E. Accept H_0 at the $\alpha = 0.01$ level

4. A Type II error is

A. rejecting the null hypothesis when it is true.
B. failing to reject the null hypothesis when it is false.
C. rejecting the null hypothesis when it is false.
D. failing to reject the null hypothesis when it is true.
E. more serious than a Type I error.

5. A researcher plans to conduct a significance test at the $\alpha = 0.05$ significance level. She designs her study to have a power of 0.85 for a particular alternative value of the parameter. The probability that the researcher will commit a Type II error for the particular alternative value of the parameter at which she computed the power is

A. 0.05.
B. 0.15.
C. 0.80.
D. 0.95.
E. equal to the 1 - (P-value) and cannot be determined until the data have been collected.

6. In hypothesis testing β is the probability of committing a Type II error in a test with significance level α. The probability of committing a Type I error is

A. $1 - \beta$
B. $1 - \alpha$
C. $\beta - \alpha$
D. α
E. Can not be determined

7. A claimed psychic was presented with 200 cards face down and asked to determine if the card was one of five symbols: a star, cross, circle, square, or three wavy lines. The "psychic" was correct in 50 cases. To determine if he has ESP, we test the hypotheses H_0: $p = 0.20$, H_a: $p > 0.20$, where p represents the true proportion of cards for which the psychic would correctly identify the symbol in the long run. Assume the conditions for inference are met. The P-value of this test is

A. between .10 and .05.
B. between .05 and .025.
C. between .025 and .01.
D. between .01 and .001.
E. below .001.

8. The most important condition for drawing sound conclusions from statistical inference is usually

A. that the population standard deviation is known.
B. that at least 30 people are included in the study.
C. that the data come from a random sample or a randomized experiment.
D. that the population distribution is exactly Normal.
E. that no calculation errors are made in the confidence interval or test statistic.

9. The mean weight of a random sample of 35 athletes is found to be 165 pounds with a standard deviation of 20 pounds. It is believed that a mean weight of 160 pounds would be normal for this group. To see if there is evidence that the mean weight of the population of all athletes of this type is significantly higher than 160 pounds, the hypotheses H_0: $\mu = 160$ vs. H_a: $\mu > 160$ are tested. You obtain a P-value of 0.0742. Which of the following is true?

A. At the 5% significance level, you have proved that H_0 is true.
B. You have failed to obtain sufficient evidence against H_0.
C. At the 5% significance level, you have failed to prove that H_0 is true, and a larger sample size is needed to do so.
D. Only 7.42% of the athletes weigh less than 160 pounds.
E. None of the above. A significance test is inappropriate in this setting.

10. A medical researcher wishes to investigate the effectiveness of exercise versus diet in losing weight. Two groups of 25 overweight adult subjects are used, with a subject in each group matched to a similar subject in the other group on the basis of a number of physiological variables. One of the groups is placed on a regular program of vigorous exercise but with no restriction on diet, and the other is placed on a strict diet but with no requirement to exercise. The weight losses after 20 weeks are determined for each subject, and the difference between matched pairs of subjects (weight loss of subject in exercise group - weight loss of matched subject in diet group) is computed. The mean of these differences in weight loss is found to be 2 lb with standard deviation $s_x = 4$ lb. Is this convincing evidence of a difference in mean weight loss for the two methods? To answer this question, you should use

A. one-proportion z test.
B. one-sample z interval for μ_d.
C. one-proportion z interval.
D. one-sample t test for μ_d.
E. none of the above.

Multiple Choice Answers

Problem	Answer	Concept	Right	Wrong	Simple Mistake?	Need to Study More
1	C	Writing Hypotheses				
2	A	Writing Hypotheses				
3	D	Significance Level and *P*-value				
4	B	Type I and Type II Errors				
5	B	Power				
6	D	Type I and Type II Errors				
7	B	Calculating *P*-value for proportions				
8	C	Conditions for Inference				
9	B	Interpreting *P*-value				
10	D	Paired *t*-test				

FRAPPY! Free Response AP Problem, Yay!

The following problem is modeled after actual Advanced Placement Statistics free response questions. Your task is to generate a complete, concise response in 15 minutes. After you generate your response, view two example solutions and determine whether you feel they are "complete," "substantial," "developing" or "minimal." If they are not "complete," what would you suggest to the student who wrote them to increase their score? Finally, you will be provided with a rubric. Score your response and note what, if anything, you would do differently to increase your own score.

During a recent movie promotion, Fruity O's cereal placed mini action figures in some of its boxes. The advertisement on the box states 1 out of every 4 boxes contains an action figure. A group of promotional-toy collectors suspects the proportion of boxes containing the action figure may be lower than 0.25. The group purchased 70 boxes of cereal and found 12 action figures. Assuming the 70 boxes represent a random sample of all of the cereal boxes, is there evidence to support the toy collector's belief that the proportion of boxes containing the figure is less than 0.25? Provide statistical evidence to support your answer.

Student Response 1:

We will perform a one-sample z-test for the proportion of boxes, p, that contain the action figures.

H_0: $p = 0.25$

H_a: $p < 0.25$

We are told the sample is random. Further, both np and $n(1-p)$ are greater than 10.

$z = \dfrac{0.17 - 0.25}{\sqrt{\dfrac{.25(.75)}{70}}} = -1.52$ and our p-value is 0.064.

We do not have significant evidence at the 5% level to reject the null hypothesis. The proportion of boxes that contain the action figure is 0.25.

How would you score this response? Is it substantial? Complete? Developing? Minimal? Is there anything this student could do to earn a better score?

Student Response 2:

Let p = the actual proportion of cereal boxes that contain the action figure

H_0: $p = 0.25$

H_a: $p < 0.25$

We are told the sample is random.
70(0.25) = 17.5 and 70(0.75) = 52.5 Both are greater than 10.
There are more than 700 cereal boxes in the population.

One sample z test for a proportion

$z = z = \dfrac{0.17 - 0.25}{\sqrt{\dfrac{.25(.75)}{70}}} = -1.52$ p-value = 0.064.

Since our p-value is greater than the typical significance level of 5%, we do not have significant evidence to reject the null hypothesis. There is not enough support to suggest that the actual proportion of boxes that contain the figure is less than 0.25.

How would you score this response? Is it substantial? Complete? Developing? Minimal? Is there anything this student could do to earn a better score?

Scoring Rubric

Use the following rubric to score your response. Each part receives a score of "Essentially Correct," "Partially Correct," or "Incorrect." When you have scored your response, reflect on your understanding of the concepts addressed in this problem. If necessary, note what you would do differently on future questions like this to increase your score.

Intent of the Question

The goal of this question is to determine your ability to conduct a significance test for a single proportion.

Solution

The solution should contain 4 parts:

- Hypotheses must be stated appropriately. H_0: $p = 0.25$ and H_a: $p < 0.25$ where p = true proportion of boxes containing the action figure.
- Name of test and conditions: The test must be identified by name or formula as a one-sample z test for a population proportion. **Random** sample is given. **Normal:** 70(0.25) = 17.5 and 70(0.75) = 52.5 Both are at least 10. **Independent:** There are more than 700 cereal boxes in the population.
- Mechanics: $z = -1.52$ and P-value = 0.064
- Conclusion: Because the P-value is greater than the typical significance level of 5%, we fail to reject the null hypothesis. There is not sufficient evidence to suggest that the actual proportion of boxes with the action figure is less than 0.25.

Scoring

Each element scored as essentially correct (E), partially correct (P), or incorrect (I).

Hypotheses is essentially correct if the hypotheses are written correctly. This part is partially correct if the test isn't identified or if the hypotheses are written incorrectly.

Name & Conditions is essentially correct if the response correctly identifies the test by name or formula AND correctly checks the Normal condition and 10% condition. This part is partially correct if the test is correctly identified but only one of the conditions is checked correctly OR if the test is not identified correctly but both conditions are checked correctly.

Mechanics is essentially correct if the response correctly calculates the test statistic and p-value. This part is partially correct if one of the calculations is incorrect.

Conclusion is essentially correct if the response correctly fails to reject the null hypothesis because the P-value is greater than a significance level of 5% OR if the response is to reject the null hypothesis because the P-value is less than a significance level of 10% and provides an interpretation in context. This part is partially correct if the response fails to justify the decision by comparing the P-value to a significance level OR if the conclusion lacks an interpretation in context.

Scoring

This problem has four elements, each receiving an E, P, or I. Assign one point to each E, 0.5 points to each P, and 0 points to each I. Total the points to determine your score. If a score falls between two whole values, consider the strength of the entire response to determine whether to round up or down.