OBJECTIVES

1. Define the null and alternate hypotheses
2. State conclusions to hypothesis tests
3. Distinguish between Type I and Type II errors

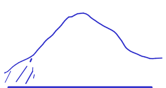
OBJECTIVE 1
DEFINE THE NULL AND ALTERNATE HYPOTHESES

A study published in the *Journal of the Air and Waste Management Association* reported that the mean amount of particulate matter (PM) produced by cars and light trucks in an urban setting is 35 milligrams of PM per mile of travel. Suppose that a new engine design is proposed that is intended to reduce the amount of PM in the air. There are two possible outcomes that could happen with the new engine design: either the new design will reduce the level of PM, or it will not.

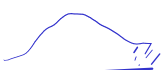
These possibilities are called **hypotheses**. One of the hypotheses is called the **null hypothesis** and the other is called the **alternate hypothesis**.

The null hypothesis about a parameter states that the parameter is equal to a specific value, μ_0. The null hypothesis is denoted H_0.

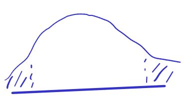
The alternate hypothesis about a parameter states that the parameter differs from the value specified by the null hypothesis, μ_0. The alternate hypothesis is denoted H_1. There are three possible alternate hypotheses:



1. left tailed : parameter < specified value, for example, $H_1 : \mu < \mu_0$



2. Right tailed : Parameter > specified value, for example, $H_1 : \mu > \mu_0$



3. Two tailed : parameter ≠ specified value, for example, $H_1 : \mu \neq \mu_0$

left-tailed and right-tailed hypothesis are called one-tailed hypothesis

EXAMPLE 1: Boxes of a certain kind of cereal are labeled as containing 20 ounces. An inspector thinks that the mean weight may be less than this. State the appropriate null and alternate hypotheses.

SOLUTION: Null hypothesis : $H_0 : \mu = 20$

Alternate hypothesis : $H_1 : \mu < 20$ left-tailed

EXAMPLE 2: Last year, the mean monthly rent for an apartment in a certain city was $800. A real estate agent believes that the mean rent is higher this year. State the appropriate null and alternate hypotheses.

SOLUTION:

$H_0: \mu = 800$

$H_1: \mu > 800$ right-tailed

EXAMPLE 3: Scores on a standardized test have a mean of 70. Some modifications are made to the test, and an educator believes that the mean may have changed. State the appropriate null and alternate hypotheses.

SOLUTION:

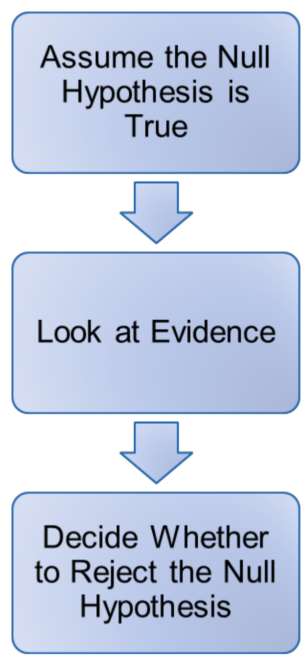
$H_0: \mu = 70$

$H_1: \mu \neq 70$ two-tailed

A HYPOTHESIS TEST IS LIKE A TRIAL

The purpose of a **hypothesis test** is to determine how likely it is that the null hypothesis is true.

The idea behind a hypothesis test is similar to a criminal trial. At the beginning of a trial, the defendant is assumed to be innocent. Then the evidence is presented. If the evidence strongly indicates that the defendant is guilty, we abandon the assumption of innocence and conclude the defendant is guilty. In a hypothesis test, the null hypothesis is like the defendant in a criminal trial.



At the start of a hypothesis test, we assume that the null hypothesis is true.

Then we look at the evidence, which comes from data that have been collected.

If the data strongly indicate that the null hypothesis is false, we abandon our assumption that it is true and believe the alternate hypothesis instead. This is referred to as **rejecting** the null hypothesis.

OBJECTIVE 2
STATE CONCLUSIONS TO HYPOTHESIS TESTS

do not reject

We may either **reject** the null hypothesis or **fail to reject** the null hypothesis. If the null hypothesis is rejected, the conclusion is straightforward: We conclude that the alternate hypothesis, H_1, is true. If the null hypothesis is not rejected, we are saying that there is not enough evidence to conclude that the alternate hypothesis, H_1, is true. We are **not** saying the null hypothesis is true. What we are saying is that the null hypothesis **might** be true.

EXAMPLE 1: Boxes of a certain kind of cereal are labeled as containing 20 ounces. An inspector thinks that the mean weight may be less than this, so he performs a test of H_0: $\mu = 20$ versus H_1: $\mu < 20$. He rejects the null hypothesis. State an appropriate conclusion.

SOLUTION: Because the null hypothesis H_0 is rejected, we conclude that the alternate hypothesis is true. we conclude that there is enough evidence to conclude that the mean weight of cereal boxes is less than 20 ounces

$H_0: \mu = 20$ reject

$H_1: \mu < 20$ accept

EXAMPLE 2: Boxes of a certain kind of cereal are labeled as containing 20 ounces. An inspector thinks that the mean weight may be less than this, so he performs a test of H_0: $\mu = 20$ versus H_1: $\mu < 20$. He does not reject the null hypothesis. State an appropriate conclusion.

SOLUTION: Since H_0 is not rejected, we do not have sufficient evidence to conclude that H_1 is true. In words, we state that there is not enough evidence to conclude that the mean weight of cerel boxes is less than 20 ocunces

$H_0: \mu = 20$ do not reject

$H_1: \mu < 20$ do not accept

OBJECTIVE 3
DISTINGUISH BETWEEN TYPE I AND TYPE II ERRORS

When a hypothesis test is conducted and a decision is made there is a possibility that it is the wrong decision.

There are two ways in which a wrong decision may occur with hypothesis testing:

1. If H_0 is true, we might mistakenly reject it. A type I error occurs when we reject H_0 when it actually true
2. If H_0 is false, we might mistakenly decide not to reject it, A type II error occurs when we do not reject H_0 when it is actually false

EXAMPLE 1: The dean of a business school wants to determine whether the mean starting salary of graduates of her school is greater than \$50,000. She will perform a hypothesis test with the following null and alternate hypotheses:

$$H_0: \mu = \$50{,}000 \qquad H_1: \mu > \$50{,}000$$

Suppose that the true mean is $\mu = \$50{,}000$, and the dean rejects H_0. Is this a Type I error, Type II error, or a correct decision?

SOLUTION: The true mean is $\mu = 50000$, so H_0 is true. Because the dean rejects H_0, this is type I error

EXAMPLE 2: The dean of a business school wants to determine whether the mean starting salary of graduates of her school is greater than \$50,000. She will perform a hypothesis test with the following null and alternate hypotheses:

$$H_0: \mu = \$50{,}000 \qquad H_1: \mu > \$50{,}000$$

Suppose that the true mean is $\mu = \$55{,}000$, and the dean rejects H_0. Is this a Type I error, Type II error, or a correct decision?

SOLUTION: The true mean is $\mu = 550000$, so H_0 is false. Because the dean rejects H_0, this is the correct decision

EXAMPLE 3: The dean of a business school wants to determine whether the mean starting salary of graduates of her school is greater than \$50,000. She will perform a hypothesis test with the following null and alternate hypotheses:

$$H_0: \mu = \$50{,}000 \qquad H_1: \mu > \$50{,}000$$

Suppose that the true mean is $\mu = \$55{,}000$, and the dean does not reject H_0. Is this a Type I error, Type II error, or a correct decision?

SOLUTION: The true mean is $\mu = 550000$, so H_0 is false. Because the dean does not reject H_0, this is a type II error.

YOU SHOULD KNOW …

- How to write the null and alternate hypotheses
- How to determine whether a hypothesis test is left-tailed, right-tailed, or two-tailed
- How to state the conclusion to a hypothesis test
- How to distinguish between Type I and Type II errors and correct decisions

SECTION 9.2: HYPOTHESIS TESTS FOR A POPULATION MEAN, σ KNOWN

OBJECTIVES

1. Perform hypothesis tests with the *P*-value method
2. Describe the relationship between hypothesis tests and confidence intervals
3. Describe the relationship between α and the probability of error
4. Report the *P*-value or the test statistic value
5. Distinguish between statistical significance and practical significance

OBJECTIVE 1
PERFORM HYPOTHESIS TESTS WITH THE P-VALUE VALUE METHOD

PERFORMING A HYPOTHESIS TEST

In the previous section, we discussed how to construct the hypotheses and write the conclusions for a hypothesis test. Now, we turn our attention to actually performing the test.

There are two ways to perform hypothesis tests; both methods produce the same results. The methods are the **Critical Value Method** and the ***P*-Value Method**.

In a hypothesis test, the idea is to select a sample, calculate a statistic such as $\bar{x}$, and compare it to the value in the null hypothesis, H_0. If the difference between the sample mean and the value in H_0 is large, it is less likely to be due to chance, and H_0 is less likely to be true. Otherwise, a small difference may be due to chance and H_0 may well be true.

We must determine how strong the disagreement is between the sample mean and H_0.

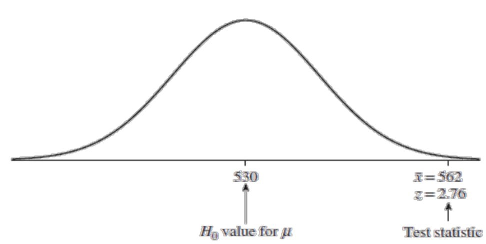
EXAMPLE:

The College Board reported that the mean math SAT score in 2009 was 515, with a standard deviation of 116. Results of an earlier study suggest that coached students should have a mean SAT score of approximately 530. A teacher who runs an online coaching program thinks that students coached by his method have a higher mean score than this. Because the teacher believes that the mean score for his students is greater than 530, the null and alternate hypotheses are:

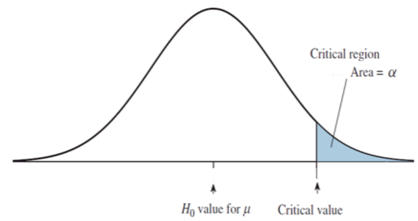
$$H_0: \mu = 530$$
$$H_1: \mu > 530$$

Suppose now that the teacher draws a random sample of 100 students who are planning to take the SAT, and enrolls them in the online coaching program. After completing the program, their sample mean SAT score is $\bar{x}$ = 562. This is higher than the null hypothesis value of 530, but to determine how strong the disagreement is between the sample mean and the null hypothesis μ = 530, we calculate the value of the **test statistic**, which is just the z-score of the sample mean. Assuming that the population standard deviation is σ = 116, the test statistic of the sample mean, $\bar{x}$ is $z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{562-530}{116/\sqrt{100}} = 2.76$.

Visually, we can see from the figure that our observed value $\bar{x}$ = 562 is pretty far out in the tail of the distribution – far from the null hypothesis value μ = 530. Intuitively, it appears that the evidence against H_0 is fairly strong.



The probability that we use to determine whether an event is unusual is called the **significance level**. The significance level is denoted by the letter α. Therefore, the area of critical region is equal to α. The choice of α is determined by how strong we require the evidence against H_0 to be in order to reject it. The smaller the value of α, the stronger we require the evidence to be.



The assumptions for performing a hypothesis test about μ when σ is known are:

Assumption:

1) We have a simple random sample (SRS)

2) The sample size is large ($n > 30$), or the proportion is approximately normal

EXAMPLE: The American Automobile Association reported that the mean price of a gallon of regular grade gasoline in the city of Los Angeles in July 2013 was \$4.04. A recently taken simple random sample of 50 gas stations in Los Angeles had an average price of \$3.99 for a gallon of regular grade gasoline. Assume that the standard deviation of prices is \$0.15. An economist is interested in determining whether the mean price is less than \$4.04. Use the P-value value method to perform a hypothesis test at the $\alpha = 0.05$ level of significance.

SOLUTION: Assumption: SRS ✓ , $n = 50 > 30$ ✓

Thus the assumption are met

Null hypothesis : $H_0 : \mu = 4.04 \leftarrow \mu_0$

Alternate hypothesis : $H_1 : \mu < 4.04$ left-tailed

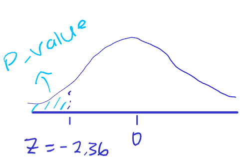
Level of significance : $\alpha = 0.05$

test statistic : $z = \dfrac{\bar{x} - \mu_0}{\frac{\theta}{\sqrt{n}}} = \dfrac{(3.99 - 4.04)}{(0.15/\sqrt{50})} = -2.36$

$\bar{x} = 3.99$

$\theta = 0.05$

$n = 50$



P-value = area to the left of $z = -2.36$

$= \text{normalcdf}(-\infty, -2.36, 0, 1)$

$= 0.00914$

Since the p-value $= 0.00914 < \alpha = 0.05$, we reject H_0

Conclusion: There is enough evidence to conclude that the mean price of a gallon of regular gasoline in Los Angeles is less than \$4.04.

OBJECTIVE 2
PERFORM HYPOTHESIS TESTS WITH THE *P*-VALUE METHOD

THE *P*-VALUE

The ***P*-value** is the probability that a number drawn from the distribution of the sample mean would be as extreme as or more extreme than our observed value of $\bar{x}$.

Unlike the critical value, the *P*-value tells us exactly **how unusual** the test statistic is. For this reason, the *P*-value method is more often used in practice, especially when technology is used to conduct a hypothesis test.

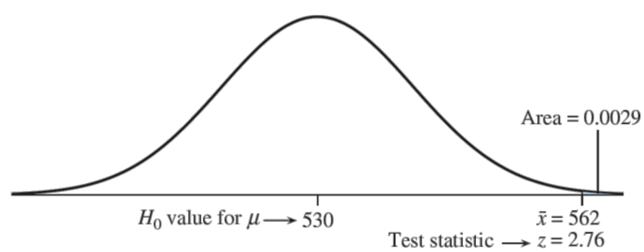
The smaller the *P*-value, the stronger the evidence against H_0.

FINDING *P*-VALUES

Consider again the following example. An online coaching program is supposed to increase the mean SAT math score to a value greater than 530. The null and alternate hypotheses are $H_0: \mu = 530$ & $H_1: \mu > 530$. Now assume that 100 students are randomly chosen to participate in the program, and their sample mean score is $\bar{x} = 562$. Suppose that the population standard deviation is known to be $\sigma = 116$. Does this provide strong evidence against the null hypothesis $H_0: \mu = 530$?

Recall that we begin by assuming that H_0 is true, therefore we assume that the mean of $\bar{x}$ is $\mu = 530$. Since the sample size is large, we know that $\bar{x}$ is approximately normally distributed with standard deviation $\frac{\sigma}{\sqrt{n}} = \frac{116}{\sqrt{100}} = 11.6$. The z-score for $\bar{x}$ is is $z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{562 - 530}{116/\sqrt{100}} = 2.76$.

The *P*-value is the area under the normal curve to the right of $z = 2.76$. This area equals 0.0029 (using technology). Therefore, the *P*-value for this test is 0.0029. Such a small *P*-value strongly suggests that H_0 should be rejected.



CHOOSING A SIGNIFICANCE LEVEL

We have seen that the smaller the *P*-value, the stronger the evidence against H_0. Therefore, to make a decision to reject H_0 when using the *P*-value method, we:

- Choose a significance level α between 0 and 1.
- Compute the *P*-value.
- If $P \le \alpha$, reject H_0. If $P > \alpha$, do not reject H_0.

If $P \le \alpha$, we say that H_0 is rejected at the α level, or that the result is **statistically significant** at the α level.

Example: Suppose we found a P-value that was $P = 0.0122$.

a) Do you reject H_0 at $\alpha = 0.05$ level?

b) Do you reject H_0 at $\alpha = 0.01$ level?

Solution: a) Because $p < 0.05$, we reject H_0 at the $\alpha = 0.05$ level

b) Because $p > 0.01$, we do not reject H_0 at the $\alpha = 0.01$ level

Example: The mean height of adult men in the U.S. is 69.7 inches, with a standard deviation of 3 inches. A sociologist believes that taller men may be more likely to be promoted to positions of leadership, so the mean height μ of male business executives may be greater than the mean height of the entire male population. A simple random sample of 100 male business executives has a mean height of 69.9 in. Assume that the standard deviation of male executive heights is σ = 3 inches. Can we conclude that male business executives are taller on the average than the general male population at the $\alpha = 0.05$ level?

Solution: Assumptions: SRS ✓, $n = 100 > 30$ ✓

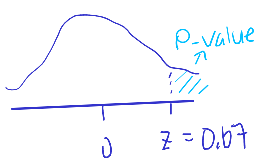
thus, the assumptions are met

$H_0: \mu = 69.7$

$H_1: \mu > 69.7$ (right-tailed)

$\alpha = 0.05$

test statistic: $z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \simeq \dfrac{(69.9 - 69.7)}{(3/\sqrt{100})} = 0.67$



P-value = area to the right of $z = 0.67$

= normalcdf $(0.67, \infty, 0, 1)$

= 0.2514

Since the p-value > 0.05, we do not reject H_0 at the $\alpha = 0.05$ level.

Conclusion: There is not enough evidence to conclude that male executives have greater mean height than adult males in general.

EXAMPLE: At a large company, the attitudes of workers are regularly measured with a standardized test. The scores on the test range from 0 to 100, with higher scores indicating greater satisfaction with their job. The mean score over all of the company's employees was 74, with a standard deviation of $\sigma = 8$. Some time ago, the company adopted a policy of telecommuting. Under this policy, workers could spend one day per week working from home. After the policy had been in place for some time, a random sample of 80 workers was given the test to see whether their mean level of satisfaction had changed since the policy was put into effect. The sample mean was 76. Assume the standard deviation is still $\sigma = 8$. Can we conclude that the mean level of satisfaction is different since the policy change at the $\alpha = 0.05$ level?

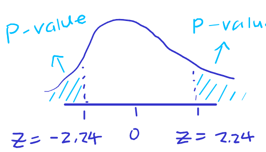
SOLUTION: Assumption : SRS ✓, $n = 80 > 30$ ✓
thus the assumption are met.

$H_0 : \mu = 74$

$H_1 : \mu \neq 74$ (two-tailed)

$\alpha = 0.05$

Test statistic : $z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \dfrac{(76-74)}{(8/\sqrt{80})} = 2.24$



p-value = sum of the area to the right of $z = 2.24$ and to the left of $z = -2.24$

= 2 · normalcdf (2.24, ∞, 0, 1)

= 0.025

Because p-value < $\alpha = 0.05$, we reject H_0 at the $\alpha = 0.05$ level

conclusion : There is enough evidence to conclude that the mean score among employee has changed since the adoption of telecommuting

Calculator check : STAT → Z-Test

μ_0 : 74

σ : 8

$\bar{x}$: 76

n : 80

$\mu \neq \mu_0$

$z = 2.23606$

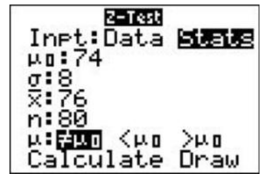
$p = 0.02531$

HYPOTHESIS TESTING ON THE TI-84 PLUS

The **ZTest** command will perform a hypothesis test when the population standard deviation σ is known. This command is accessed by pressing STAT and highlighting the TESTS menu.

If the summary statistics are given the Stats option should be selected for the input option.

If the raw sample data are given, the Data option should be selected.



OBJECTIVE 3
DESCRIBE THE RELATIONSHIP BETWEEN HYPOTHESIS TESTING AND CONFIDENCE INTERVALS

In the previous example we rejected H_0: $\mu = 74$ at α = 0.05. In doing so, we are saying that 74 is not a plausible value for μ.

Another way to express information about μ is through a confidence interval. If we construct a 95% confidence interval for μ about the attitude of workers, we find the interval to be 74.247 < μ < 77.753.

We note that this interval does not contain the null hypothesis value of 74. The confidence interval for μ contains all the plausible values for μ . Because 74 is not in the confidence interval, 74 is not a plausible value for μ.

OBJECTIVE 4
DESCRIBE THE RELATIONSHIP BETWEEN α AND THE PROBABILITY OF ERROR

Recall that a Type I error occurs if we reject H_0 when it is true and a Type II error occurs if we do not reject H_0 when it is false. We would like to make the probabilities of these errors small.

The **probability of a Type I error** is equal to the significance level, which is denoted by α. In other words, if we perform a test at a significance level of α = 0.05, the probability of making a Type I error is 0.05.

Because α is the probability of a Type I error, why don't we always choose a very small value for α? The reason is that the smaller a value we choose for α, the larger the value of β, the probability of making a Type II error.

α = P(type I error)

β = P(type II error)

OBJECTIVE 5
REPORT THE *P*-VALUE OR THE TEST STATISTIC

Sometimes people report only that a test result was statistically significant at a certain level, without giving the *P*-value. It is much better to report the *P*-value along with the decision whether to reject. There are two reasons for this.

There is a big difference between a *P*-value that is just barely small enough to reject and one that is extremely small. For example at the α = 0.05 significance level, there is a huge difference between a *P*-value = 0.049 and *P*-value = 0.0001.

Not everyone may agree with your choice of significance level, α. By reporting the *P*-value, we let people decide for themselves at what level to reject H_0.

When using the critical value method, you should report the value of the test statistic rather simply stating whether the test statistic was in the rejection region.

This will tell the reader whether the value of the test statistic was just barely inside the critical region, or well inside. Also, it provides the reader an opportunity to choose a different critical value and determine whether H_0 can be rejected at a different level.

OBJECTIVE 6
DISTINGUISH BETWEEN STATISTICAL SIGNIFICANCE AND PRACTICAL SIGNIFICANCE

When a result has a small *P*-value, we say that it is "statistically significant." It is therefore tempting to think that statistically significant results must always be important. Sometimes statistically significant results do not have any practical importance or practical significance.

For example, a new study program may raise students' scores by two points on a 100 point scale. This improvement may have statistical significance, but is the improvement important enough to offset the cost of training teachers and the time investment on behalf of the students.

YOU SHOULD KNOW …

- What are the assumptions for performing hypothesis tests about μ when σ is known
- How to find and interpret a *P*-value
- How to perform hypothesis tests about μ when σ is known with the *P*-value method
- How to describe the relationship between hypothesis tests and confidence intervals
- How to describe the relationship between α and the probability error
- How to distinguish between statistical significance and practical significance
- Why it is better to report the *P*-value or the test statistic value along with the decision of a hypothesis test

Section 9.3: Hypothesis Tests for a Population Mean, σ Unknown

Objectives

1. Test a hypothesis about a mean using the *P*-value method

Objective 1
Test a Hypothesis About a Mean Using the *P*-Value Method

We begin this section with an example. Suppose that in a recent medical study, 76 subjects were placed on a low-fat diet. After 12 months, their sample mean weight loss was $\bar{x}$ = 2.2 kilograms, with a sample standard deviation of s = 6.1 kilograms. Can we conclude that the mean weight loss is greater than 0?

If we knew the population standard deviation σ, we would be able to compute the z-score of the sample mean to be $z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$, and use this test statistic to perform a hypothesis test. In this example, as is usually the case, we do not know the population standard deviation. To proceed, we replace σ with the sample standard deviation s, and use the t test statistic instead: $t = \frac{\bar{x}-\mu}{s/\sqrt{n}}$. When the null hypothesis is true, the t statistic has a Student's t distribution with $n-1$ degrees of freedom.

The assumptions for performing a hypothesis test for μ when the population standard deviation σ is unknown are as follows:

1. We have a simple random sample.
2. The sample size is large ($n > 30$), or the population is approximately normal.

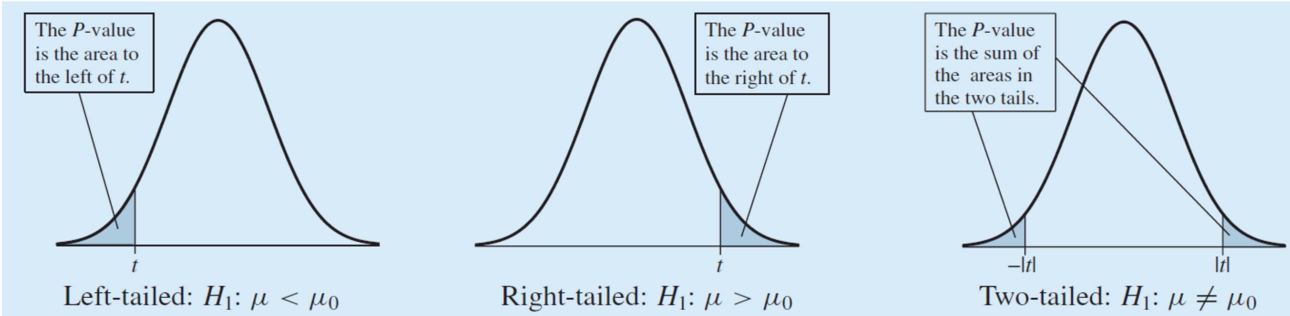
Since we have a simple random sample and the sample size is large, we may proceed with the test. The issue is whether the mean weight loss μ is greater than 0. So the null and alternate hypotheses are $H_0: \mu = 0$ versus $H_1: \mu > 0$.

The test statistic is $t = \frac{\bar{x}-\mu}{s/\sqrt{n}} = \frac{2.2-0}{6.1/\sqrt{76}} = 3.144$. When H_0 is true, the test statistic t has the Student's t distribution with $n - 1 = 76 - 1 = 75$ degrees of freedom. This is a right tail test, so the *P*-value is the area under the Student's t curve to the right of t = 3.144. Using technology, we find the exact *P*-value to be *P* = 0.0012.

Since *P* < 0.05, we reject H_0 at the $\alpha = 0.05$ level. We conclude that the mean weight loss of people who adhered to this diet for 12 months is greater than 0.

Computing *P*-Values

The *P*-value of the test statistic t is the probability, assuming H_0 is true, of observing a value for the test statistic that disagrees as strongly as or more strongly with H_0 than the value actually observed. The *P*-value is an area under the Student's t curve with $n-1$ degrees of freedom. The area is in the left tail, the right tail, or in both tails, depending on the type of alternate hypothesis.



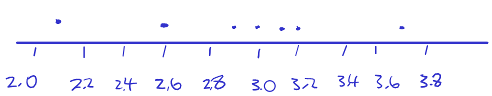
EXAMPLE: Generic drugs are lower-cost substitutes for brand-name drugs. Before a generic drug can be sold in the United States, it must be tested and found to perform equivalently to the brand name product. The U.S. Food and Drug Administration is now supervising the testing of a new generic antifungal ointment. The brand-name ointment is known to deliver a mean of 3.5 micrograms of active ingredient to each square centimeter of skin. As part of the testing, seven subjects apply the ointment. Six hours later, the amount of drug that has been absorbed into the skin is measured. The amounts, in micrograms, are

2.6 3.2 2.1 3.0 3.1 2.9 3.7

How strong is the evidence that the mean amount absorbed differs from 3.5 micrograms? Use the α = 0.01 level of significance.

SOLUTION: Check assumption: SRS ✓ $n = 7 < 30$ ✗

Because the sample is small, the population must be approximately normal, we check this with dotplot.



There is no evidence of strong skewness, and no outliers. We may proceed with the hypothesis testing

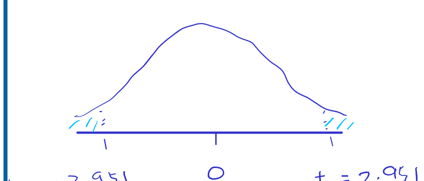
$H_0: \mu = 3.5$ $\bar{x} = 2.9429$

$H_1: \mu \neq 3.5$ (two-tailed) $s = 0.4995$

$\alpha = 0.01$

Test statistic: $t = \dfrac{\bar{x} - \mu_0}{\left(\frac{s}{\sqrt{n}}\right)} = \dfrac{(2.9429 - 3.5)}{\left(\frac{0.4995}{\sqrt{7}}\right)} = -2.951$

P-value $= 2 \cdot \text{tcdf}(\text{lower}, \text{upper}, \underset{n-1}{df}) = 2 \cdot \text{tcdf}(-\infty, -2.951, 6) = 0.02558$



Since the p-value > 0.01, we do not reject H_0.

Therefore, we do not have enough evidence to say that the mean amount of drug absorbed differs from 3.5 micrograms.

HYPOTHESIS TESTING ON THE TI-84 PLUS

The **TTest** command will perform a hypothesis test when the population standard deviation σ is not known. This command is accessed by pressing STAT and highlighting the TESTS menu.

If the summary statistics are given the Stats option should be selected for the input option.

If the raw sample data are given, the Data option should be selected.

EXAMPLE: A computer software vendor claims that a new version of their operating system will crash less than six times per year on average. A system administrator installs the operating system on a random sample of 41 computers. At the end of a year, the sample mean number of crashes is 7.1, with a standard deviation of 3.6. Can you conclude that the vendor's claim is false? Use the α = 0.05 significance level.

SOLUTION: Check assumption: SRS ✓ $n = 41 > 30$ ✓

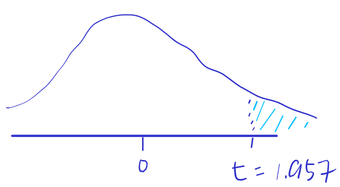
$H_0 : \mu = 6$

$H_1 : \mu > 6$

$\alpha = 0.05$

test statistic : $t = \dfrac{\bar{x} - \mu_0}{\left(\frac{s}{\sqrt{n}}\right)} = \dfrac{7.1 - 6}{\left(\frac{3.6}{\sqrt{41}}\right)} = 1.957$

P-value :



P-value = tcdf (1.957, ∞, 40)

= 0.02368

Since the p-value < α, we reject H_0

There is enough evidence to state that the mean number of crashes is greater than six times per year. Thus, the vendor's claim is false.

YOU SHOULD KNOW …

- The assumptions for hypothesis tests for μ when σ is unknown
- How to perform hypothesis tests for μ when σ is unknown using the *P*-value method

OBJECTIVES

1. Test a hypothesis about a proportion using the *P*-value method

OBJECTIVE 1
TEST A HYPOTHESIS ABOUT A PROPORTION USING THE *P*-VALUE METHOD

sample

In a recent GenX2Z American College Student Survey, 90% of female college students rated the social network site Facebook as "cool." The other 10% rated it as "lame." Assume that the survey was based on a sample of 500 students. A marketing executive at Facebook wants to advertise the site with the slogan "More than 85% of female college students think Facebook is cool." Before launching the ad campaign, he wants to be confident that the slogan is true. Can he conclude that the proportion of female college students who think Facebook is cool is greater than 0.85?

This is an example of a problem that calls for a **hypothesis test about a population proportion**.

We use the following notation:

- p is the population proportion
- x is the number of individuals in the sample
- n is the sample size
- $\hat{p}$ is the sample proportion

$$\hat{p} = \frac{x}{n}$$

point estimate

The method for performing a hypothesis test about a population proportion requires that the sampling distribution be approximately normal. The following assumptions ensure this:

1) SRS
2) the proportion is at least 20 times as large as the sample ($n \leq 0.05N$)
   sample size, population size
3) The items in the population are divided into two categories.
4) The value np_0 and $n(1-p_0)$ are both at least 10

   That is, $np_0 \geq 10$ and $n(1-p_0) \geq 10$

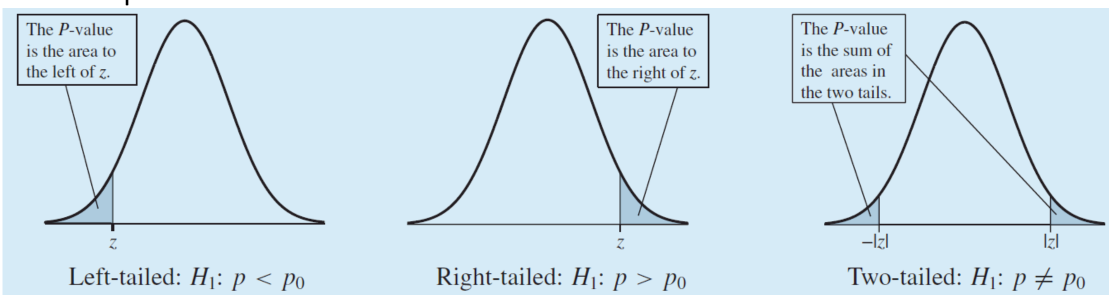
***P*-VALUE METHOD**

Step 1: State the null and alternate hypotheses.

Step 2: If making a decision, choose a significance level α.

Step 3: Compute the test statistic $z = \dfrac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$.

Step 4: Compute the *P*-value.



Step 5: Interpret the *P*-value. If making a decision, reject H_0 if the *P*-value is less than or equal to the significance level α.

Step 6: State a conclusion.

EXAMPLE: In a recent GenX2Z American College Student Survey, 90% of female college students rated the social network site Facebook as "cool." Assume that the survey was based on a random sample of 500 students. A marketing executive at Facebook wants to advertise the site with the slogan "More than 85% of female college students think Facebook is cool." Can you conclude that the proportion of female college students who think Facebook is cool is greater than 0.85? Use the α = 0.05 level of significance.

SOLUTION: Check assumption : SRS ✓ pop > 20 time the sample n = 500

2 categories
- Cool : $n p_0 = 500(0.85) = 425 \geq 10$ ✓
- lame : $n(1-p_0) = 500(1-0.85) = 75 \geq 10$ ✓

$H_0: p = 0.85$

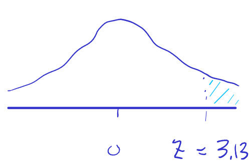
$H_1: p > 0.85$

$\alpha = 0.05$

test statistic : $z = \dfrac{\hat{p} - p}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \dfrac{0.9 - 0.85}{\sqrt{\frac{0.85(1-0.85)}{500}}} = 3.13$

P-value :

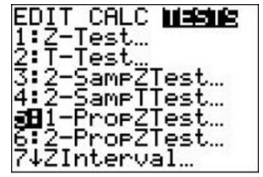
P - value = normalcdf (3.13, ∞, 0, 1) = 0.0008



Since the p-value < α = 0.05, we reject H_0.

There is enough evidence to conclude that more than 85% of female college students think Facebook is cool.

HYPOTHESIS TESTING ON THE TI-84 PLUS



The **1-PropZTest** command will perform a hypothesis test for a population proportion. This command is accessed by pressing **STAT** and highlighting the **TESTS** menu.

The required inputs for the **1-PropZTest** are the values of x and n. If the sample proportion $\hat{p}$ is given in the problem, the value of x can be computed as $x = \hat{p} \cdot n$.

EXAMPLE: A nationwide survey of working adults indicates that only 50% of them are satisfied with their jobs. The president of a large company believes that more than 50% of employees at his company are satisfied with their jobs. To test his belief, he surveys a random sample of 100 employees, and 54 of them report that they are satisfied with their jobs. Can he conclude that more than 50% of employees at the company are satisfied with their jobs? Use the $\alpha = 0.05$ level of significance.

SOLUTION: Check assumptions : SRS ✓ pop > 20 times the sample $n = 100$

2 categories → satisfied $100(0.5) = 50 \geq 10$

→ not satisfied $100(1-0.5) = 50 \geq 10$

$H_0: p = 0.5$

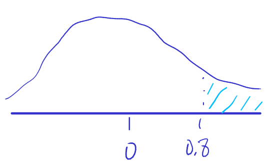
$H_1: p > 0.5$

$\alpha = 0.05$

test statistic : $z = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = \dfrac{0.54 - 0.50}{\sqrt{\dfrac{0.5(1-0.5)}{100}}} = 0.8$

p-value = normalcdf (0.8 , ∞ , 0 , 1)

= 0.21186



Since the p-value > α , we fail to reject H_0 ,

There is not enough evidence to conclude that the company president is correct in his belief that proportion of empolyees who are satisfied with their job is greater than 0.5

YOU SHOULD KNOW …

- The notations used in performing a hypothesis test about a population proportion
- The assumptions for performing a hypothesis test about a population proportion
- How to perform a hypothesis test about a population proportion using the *P*-value method

Determining Which Test of the Population Mean μ to Use

There is more than one procedure for testing a population mean, μ. The tests are

- z-test (Section 9.2)
- t-test (Section 9.3)

The diagram below will help in selecting the appropriate test for population mean, μ.