on the wording of the hypothesis test. However, be aware that many researchers (including one of the co-authors in research work) use = in the null hypothesis, even with > or < as the symbol in the alternative hypothesis. This practice is acceptable because we only make the decision to reject or not reject the null hypothesis.

Example 9.1

H_0: No more than 30% of the registered voters in Santa Clara County voted in the primary election. $p \leq 30$
H_1: More than 30% of the registered voters in Santa Clara County voted in the primary election. $p > 30$

Try It Σ

9.1 A medical trial is conducted to test whether or not a new medicine reduces cholesterol by 25%. State the null and alternative hypotheses.

$H_0: p = 0.25$
$H_1: p \neq 0.25$

Example 9.2

We want to test whether the mean GPA of students in American colleges is different from 2.0 (out of 4.0). The null and alternative hypotheses are:
H_0: $\mu = 2.0$
H_a: $\mu \neq 2.0$

Try It Σ

9.2 We want to test whether the mean height of eighth graders is 66 inches. State the null and alternative hypotheses. Fill in the correct symbol (=, ≠, ≥, <, ≤, >) for the null and alternative hypotheses.

a. H_0: μ = 66
b. H_a: μ ≠ 66

Example 9.3

We want to test if college students take less than five years to graduate from college, on the average. The null and alternative hypotheses are:
H_0: $\mu \geq 5$
H_a: $\mu < 5$

Try It Σ

9.3 We want to test if it takes fewer than 45 minutes to teach a lesson plan. State the null and alternative hypotheses. Fill in the correct symbol (=, ≠, ≥, <, ≤, >) for the null and alternative hypotheses.

a. H_0: μ ≥ 45
b. H_a: μ < 45

Scanned with CamScanner

The Power of the Test is $1 - \beta$. Ideally, we want a high power that is as close to one as possible. Increasing the sample size can increase the Power of the Test.

The following are examples of Type I and Type II errors.

Example 9.5

Suppose the null hypothesis, H_0, is: Frank's rock climbing equipment is safe.

Type I error: Frank thinks that his rock climbing equipment may not be safe when, in fact, it really is safe. **Type II error**: Frank thinks that his rock climbing equipment may be safe when, in fact, it is not safe.

α = **probability** that Frank thinks his rock climbing equipment may not be safe when, in fact, it really is safe. β = **probability** that Frank thinks his rock climbing equipment may be safe when, in fact, it is not safe.

Notice that, in this case, the error with the greater consequence is the Type II error. (If Frank thinks his rock climbing equipment is safe, he will go ahead and use it.)

Try It Σ

9.5 Suppose the null hypothesis, H_0, is: the blood cultures contain no traces of pathogen X. State the Type I and Type II errors. Type 1 error: The researcher thinks the blood cultures do contain traces of pathogen X, when in fact they do not.
Type II error: The researcher thinks the blood cultures do not contain traces of pathogen X, when in fact they do.

Example 9.6

Suppose the null hypothesis, H_0, is: The victim of an automobile accident is alive when he arrives at the emergency room of a hospital.

Type I error: The emergency crew thinks that the victim is dead when, in fact, the victim is alive. **Type II error**: The emergency crew does not know if the victim is alive when, in fact, the victim is dead.

α = **probability** that the emergency crew thinks the victim is dead when, in fact, he is really alive = P(Type I error). β = **probability** that the emergency crew does not know if the victim is alive when, in fact, the victim is dead = P(Type II error).

The error with the greater consequence is the Type I error. (If the emergency crew thinks the victim is dead, they will not treat him.)

Try It Σ

9.6 Suppose the null hypothesis, H_0, is: a patient is not sick. Which type of error has the greater consequence, Type I or Type II? Type 1 error: The patient will not be thought well when, in fact he is not sick.
Type II error: The patient will be thought well when, in fact, he is sick.

Example 9.7

It's a Boy Genetic Labs claim to be able to increase the likelihood that a pregnancy will result in a boy being born. Statisticians want to test the claim. Suppose that the null hypothesis, H_0, is: It's a Boy Genetic Labs has no effect on gender outcome.

Type I error: This results when a true null hypothesis is rejected. In the context of this scenario, we would state that we believe that It's a Boy Genetic Labs influences the gender outcome, when in fact it has no effect. The probability of this error occurring is denoted by the Greek letter alpha, α.

Scanned with CamScanner

Type II error: This results when we fail to reject a false null hypothesis. In context, we would state that It's a Boy Genetic Labs does not influence the gender outcome of a pregnancy when, in fact, it does. The probability of this error occurring is denoted by the Greek letter beta, β.

The error of greater consequence would be the Type I error since couples would use the It's a Boy Genetic Labs product in hopes of increasing the chances of having a boy.

Try It Σ

Ho: μ ≤ 800
Ha : μ > 800

9.7 "Red tide" is a bloom of poison-producing algae–a few different species of a class of plankton called dinoflagellates. When the weather and water conditions cause these blooms, shellfish such as clams living in the area develop dangerous levels of a paralysis-inducing toxin. In Massachusetts, the Division of Marine Fisheries (DMF) monitors levels of the toxin in shellfish by regular sampling of shellfish along the coastline. If the mean level of toxin in clams exceeds 800 μg (micrograms) of toxin per kg of clam meat in any area, clam harvesting is banned there until the bloom is over and levels of toxin in clams subside. Describe both a Type I and a Type II error in this context, and state which error has the greater consequence.

Type II error

Type 1 error: The DMF believes that toxin levels are still too high when, in fact toxin levels are almost 800 ug.

Type II error: The DMF believes that toxin levels are within acceptable levels when in fact toxin levels are still too high

Example 9.8

A certain experimental drug claims a cure rate of at least 75% for males with prostate cancer. Describe both the Type I and Type II errors in context. Which error is the more serious?

Type I: A cancer patient believes the cure rate for the drug is less than 75% when it actually is at least 75%.

Type II: A cancer patient believes the experimental drug has at least a 75% cure rate when it has a cure rate that is less than 75%.

In this scenario, the Type II error contains the more severe consequence. If a patient believes the drug works at least 75% of the time, this most likely will influence the patient's (and doctor's) choice about whether to use the drug as a treatment option.

Try It Σ

9.8 Determine both Type I and Type II errors for the following scenario:

Assume a null hypothesis, H_0, that states the percentage of adults with jobs is at least 88%.

Identify the Type I and Type II errors from these four statements.

a. Not to reject the null hypothesis that the percentage of adults who have jobs is at least 88% when that percentage is actually less than 88%

Type II error
b. Not to reject the null hypothesis that the percentage of adults who have jobs is at least 88% when the percentage is actually at least 88%.

Type I error
c. Reject the null hypothesis that the percentage of adults who have jobs is at least 88% when the percentage is actually at least 88%.

d. Reject the null hypothesis that the percentage of adults who have jobs is at least 88% when that percentage is actually less than 88%.

9.3 | Distribution Needed for Hypothesis Testing

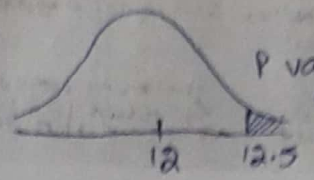
Earlier in the course, we discussed sampling distributions. **Particular distributions are associated with hypothesis**

Scanned with CamScanner

Try It Σ

9.9 A normal distribution has a standard deviation of 1. We want to verify a claim that the mean is greater than 12. A sample of 36 is taken with a sample mean of 12.5.

H_0: $\mu \leq 12$
H_a: $\mu > 12$
The p-value is 0.0013
Draw a graph that shows the p-value.



P value is approximately 0.0013

Decision and Conclusion

A systematic way to make a decision of whether to reject or not reject the **null hypothesis** is to compare the p-value and a **preset or preconceived α (also called a "significance level")**. A preset α is the probability of a **Type I error** (rejecting the null hypothesis when the null hypothesis is true). It may or may not be given to you at the beginning of the problem.

When you make a **decision** to reject or not reject H_0, do as follows:

- If $\alpha > p$-value, reject H_0. The results of the sample data are significant. There is sufficient evidence to conclude that H_0 is an incorrect belief and that the **alternative hypothesis**, H_a, may be correct.
- If $\alpha \leq p$-value, do not reject H_0. The results of the sample data are not significant. There is not sufficient evidence to conclude that the alternative hypothesis, H_a, may be correct.
- When you "do not reject H_0", it does not mean that you should believe that H_0 is true. It simply means that the sample data have failed to provide sufficient evidence to cast serious doubt about the truthfulness of H_0.

Conclusion: After you make your decision, write a thoughtful **conclusion** about the hypotheses in terms of the given problem.

Example 9.10

When using the p-value to evaluate a hypothesis test, it is sometimes useful to use the following memory device

If the p-value is low, the null must go.

If the p-value is high, the null must fly.

This memory aid relates a p-value less than the established alpha (the p is low) as rejecting the null hypothesis and, likewise, relates a p-value higher than the established alpha (the p is high) as not rejecting the null hypothesis.

Fill in the blanks.

Reject the null hypothesis when ______________________________________.

The results of the sample data ______________________________________.

Do not reject the null when hypothesis when ______________________________________.

The results of the sample data ______________________________________.

Solution 9.10

Reject the null hypothesis when **the p-value is less than the established alpha value**. The results of the sample data **support the alternative hypothesis.**

Do not reject the null hypothesis when **the p-value is greater than the established alpha value**. The results of the sample data **do not support the alternative hypothesis.**

Scanned with CamScanner

Try It Σ

9.10 It's a Boy Genetics Labs claim their procedures improve the chances of a boy being born. The results for a test of a single population proportion are as follows:

H_0: $p = 0.50$, H_a: $p > 0.50$

$\alpha = 0.01$

p-value = 0.025

Interpret the results and state a conclusion in simple, non-technical terms.

Interpretation: Since the pvalue is greater than the established value of α, we do not reject the null hyp. There is not enough evidence to support Itsababy Genetics Labs stated claim that their procedures improve the chances of a boy being born.

9.5 | Additional Information and Full Hypothesis Test Examples

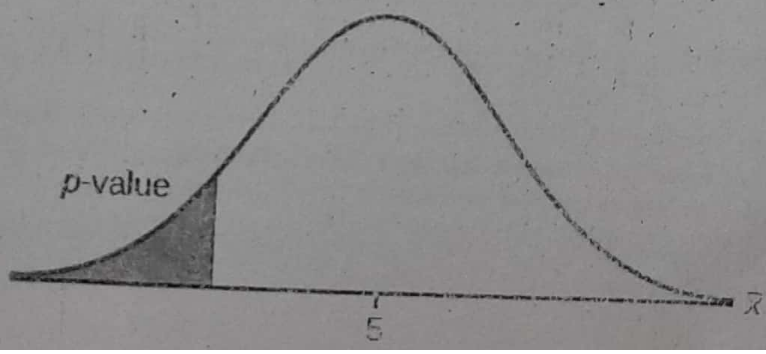
- In a **hypothesis test** problem, you may see words such as "the level of significance is 1%." The "1%" is the preconceived or preset α.
- The statistician setting up the hypothesis test selects the value of α to use **before** collecting the sample data.
- **If no level of significance is given, a common standard to use is $\alpha = 0.05$.**
- When you calculate the p-value and draw the picture, the p-value is the area in the left tail, the right tail, or split evenly between the two tails. For this reason, we call the hypothesis test left, right, or two tailed.
- The **alternative hypothesis**, H_a, tells you if the test is left, right, or two-tailed. It is the **key** to conducting the appropriate test.
- H_a **never** has a symbol that contains an equal sign.
- **Thinking about the meaning of the p-value**: A data analyst (and anyone else) should have more confidence that he made the correct decision to reject the null hypothesis with a smaller p-value (for example, 0.001 as opposed to 0.04) even if using the 0.05 level for alpha. Similarly, for a large p-value such as 0.4, as opposed to a p-value of 0.056 (alpha = 0.05 is less than either number), a data analyst should have more confidence that she made the correct decision in not rejecting the null hypothesis. This makes the data analyst use judgment rather than mindlessly applying rules.

The following examples illustrate a left-, right-, and two-tailed test.

Example 9.11

H_o: $\mu = 5$, H_a: $\mu < 5$

Test of a single population mean. H_a tells you the test is left-tailed. The picture of the p-value is as follows:



Try It Σ

9.11 H_0: $\mu = 10$, H_a: $\mu < 10$

Assume the *p*-value is 0.0935. What type of test is this? Draw the picture of the *p*-value.

Example 9.12

H_0: $p \leq 0.2$ H_a: $p > 0.2$

This is a test of a single population proportion. H_a tells you the test is **right-tailed**. The picture of the *p*-value is as follows:

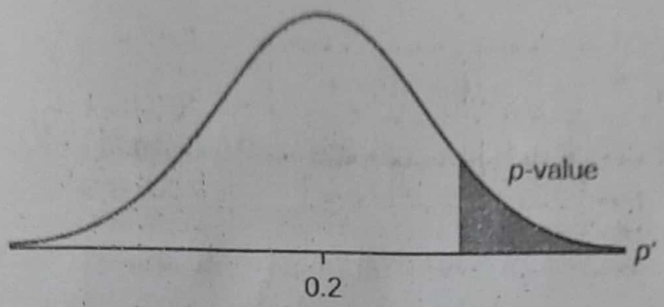
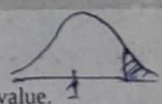


Figure 9.4

Try It Σ

9.12 H_0: $\mu \leq 1$, H_a: $\mu > 1$ Test of a Population Mean 

Assume the *p*-value is 0.1243. What type of test is this? Draw the picture of the *p*-value.

$\alpha = 0{\cdot}05$,
Pvalue $= 0{\cdot}1243$
Since Pvalue $> \alpha$, we do not reject null hyp: $H_0: \mu \leq 1$. We donot have enough evidence to conclude that $\mu > 1$.

Example 9.13

H_0: $p = 50$ H_a: $p \neq 50$

This is a test of a single population mean. H_a tells you the test is **two-tailed**. The picture of the *p*-value is as follows.

Scanned with CamScanner

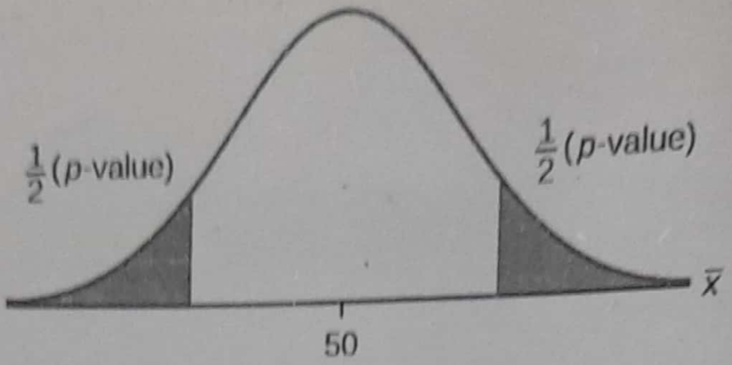
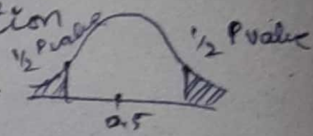


Figure 9.5

Try It Σ

9.13 H_0: $p = 0.5$, H_a: $p \neq 0.5$ Hyp of a single population proportion

Assume the p-value is 0.2564. What type of test is this? Draw the picture of the p-value.



Pvalue = 0.2564, α = 0.05

Since Pvalue > α so we do not reject null hyp (H_0: p = 0.5). We do not have sufficient evidence to conclude H_a: p ≠ 0.5

Full Hypothesis Test Examples

Example 9.14

Jeffrey, as an eight-year old, **established a mean time of 16.43 seconds** for swimming the 25-yard freestyle, with a **standard deviation of 0.8 seconds**. His dad, Frank, thought that Jeffrey could swim the 25-yard freestyle faster using goggles. Frank bought Jeffrey a new pair of expensive goggles and timed Jeffrey for **15 25-yard freestyle swims**. For the 15 swims, **Jeffrey's mean time was 16 seconds. Frank thought that the goggles helped Jeffrey to swim faster than the 16.43 seconds.** Conduct a hypothesis test using a preset $\alpha = 0.05$. Assume that the swim times for the 25-yard freestyle are normal.

Solution 9.14

Set up the Hypothesis Test:

Since the problem is about a mean, this is a **test of a single population mean**.

H_0: $\mu = 16.43$ H_a: $\mu < 16.43$

For Jeffrey to swim faster, his time will be less than 16.43 seconds. The "<" tells you this is left-tailed.

Determine the distribution needed:

Random variable: $\overline{X}$ = the mean time to swim the 25-yard freestyle.

Distribution for the test: $\overline{X}$ is normal (population **standard deviation** is known: $\sigma = 0.8$)

$\overline{X} \sim N\left(\mu, \frac{\sigma_X}{\sqrt{n}}\right)$ Therefore, $\overline{X} \sim N\left(16.43, \frac{0.8}{\sqrt{15}}\right)$

$\mu = 16.43$ comes from H_0 and not the data. $\sigma = 0.8$, and $n = 15$.

Calculate the p-value using the normal distribution for a mean:

when, in fact, he actually swims the 25-yard freestyle, on average, in 16.43 seconds. (Reject the null hypothesis when the null hypothesis is true.)

The Type II error is that there is not evidence to conclude that Jeffrey swims the 25-yard free-style, on average, in less than 16.43 seconds when, in fact, he actually does swim the 25-yard free-style, on average, in less than 16.43 seconds. (Do not reject the null hypothesis when the null hypothesis is false.)

Try It Σ

See below

9.14 The mean throwing distance of a football for Marco, a high school freshman quarterback, is 40 yards, with a standard deviation of two yards. The team coach tells Marco to adjust his grip to get more distance. The coach records the distances for 20 throws. For the 20 throws, Marco's mean distance was 45 yards. The coach thought the different grip helped Marco throw farther than 40 yards. Conduct a hypothesis test using a preset $\alpha = 0.05$. Assume the throw distances for footballs are normal.

First, determine what type of test this is, set up the hypothesis test, find the *p*-value, sketch the graph, and state your conclusion.

Using the TI-83, 83+, 84, 84+ Calculator

Press STAT and arrow over to TESTS. Press 1:Z-Test. Arrow over to Stats and press ENTER. Arrow down and enter 40 for $\mu 0$ (null hypothesis), 2 for σ, 45 for the sample mean, and 20 for *n*. Arrow down to μ: (alternative hypothesis) and set it either as <, ≠, or >. Press ENTER. Arrow down to Calculate and press ENTER. The calculator not only calculates the *p*-value but it also calculates the test statistic (*z*-score) for the sample mean. Select <, ≠, or > for the alternative hypothesis. Do this set of instructions again except arrow to Draw (instead of Calculate). Press ENTER. A shaded graph appears with test statistic and *p*-value. Make sure when you use Draw that no other equations are highlighted in $Y =$ and the plots are turned off.

HISTORICAL NOTE (EXAMPLE 9.11)

The traditional way to compare the two probabilities, α and the *p*-value, is to compare the critical value (*z*-score from α) to the test statistic (*z*-score from data). The calculated test statistic for the *p*-value is –2.08. (From the Central Limit Theorem, the test statistic formula is $z = \frac{\bar{x} - \mu_X}{\left(\frac{\sigma_X}{\sqrt{n}}\right)}$. For this problem, $\bar{x} = 16$, $\mu_X = 16.43$ from the null hypothesis is, $\sigma_X = 0.8$, and $n = 15$.) You can find the critical value for $\alpha = 0.05$ in the normal table (see **15.Tables** in the Table of Contents). The z-score for an area to the left equal to 0.05 is midway between –1.65 and –1.64 (0.05 is midway between 0.0505 and 0.0495). The z-score is –1.645. Since $-1.645 > -2.08$ (which demonstrates that $\alpha >$ *p*-value), reject H_0. Traditionally, the decision to reject or not reject was done in this way. Today, comparing the two probabilities α and the *p*-value is very common. For this problem, the *p*-value, 0.0187 is considerably smaller than α, 0.05. You can be confident about your decision to reject. The graph shows α, the *p*-value, and the test statistic and the critical value.

Try it 9.14 (Solution)

1: Formulation of hyp: $H_0: \mu = 40$
$H_a: \mu > 40$

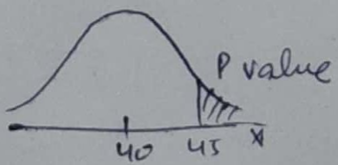
2: Level of Sig
$\alpha = 0.05$

3: Test Statistic
$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

4: Calculation:

$z = \frac{45-40}{2/\sqrt{20}} = 11.1803$

5: P-value:
$= 2.6115 \times 10^{-29}$



6: Conclusion:
Because $P < \alpha$, we reject the null hyp.
There is sufficient evidence to suggest that the change in grip improved Marco's throwing distance.

Scanned with CamScanner