Chapter 8: Hypothesis Testing

A hypothesis is a claim made about a population parameter such as a population mean or a population proportion.

A hypothesis test is a method for using sample data to make a decision about a population parameter.

To perform a hypothesis test, a statistician will:

1. Set up two contradictory hypotheses.
2. Collect sample data.
3. Assuming that one of the hypotheses is true, determine whether the sample data contradict that hypothesis.
4. Make a decision and write a meaningful conclusion.

Null and Alternative Hypotheses

null hypothesis (H_0)- the statement that the value of a population parameter is equal to some claimed value. It is assumed to be true unless it can be shown to be incorrect beyond a reasonable doubt.

alternative hypothesis (H_a)– the statement that you will adopt when the evidence is strong enough to reject H_0.

Types of Tests

Two-tailed:	$H_0: \mu = k$ $H_a: \mu \neq k$	or	$H_0: p = k$ $H_a: p \neq k$
Left-tailed:	$H_0: \mu = k$ $H_a: \mu < k$	or	$H_0: p = k$ $H_a: p < k$
Right-tailed:	$H_0: \mu = k$ $H_a: \mu > k$	or	$H_0: p = k$ $H_a: p > k$

Note that the null hypothesis always contains an equals sign.

Example: A newspaper reporter claims that less than 25% of registered voters plan to vote in the upcoming election.

H_0: $p = 0.25$
H_a: $p < 0.25$ (claim)

Example: A credit card company claims that nationwide percentage of college students using credit cards is 83%.

H_0: $p = 0.83$ (claim)
H_a: $p \neq 0.83$

Example: An administrator believes that takes college students in Massachusetts an average of 4.7 years to complete their degree.

H_0 : $\mu = 4.7$ (claim)
H_a : $\mu \neq 4.7$

Example: A researcher believes that on average, preschoolers are exposed to more than three hours of screen time per day.

H_0: $\mu = 3$
H_a: $\mu > 3$ (claim)

Collect and Analyze Sample Data

When testing a hypothesis about a population mean or population proportion, we will collect sample data and determine whether the sample results contradict our assumption in the null hypothesis.

Example: A newspaper reporter claims that less than 25% of registered voters plan to vote in the upcoming election.

H_0: $p = 0.25$
H_a: $p < 0.25$ (claim)

Suppose that a sample of 1000 registered voters is selected, and 200 of those selected plan to vote in the upcoming election. The sample proportion is therefore 200/1000 = 0.20.

We know that sample proportions vary from sample to sample. Assuming that the statement in the null hypothesis is true (i.e. that the population proportion is 0.25), how likely was it that was obtained a sample proportion of 0.20? Could this result be due to chance? Or might we conclude that the population proportion is actually less than 0.25?

Example: An administrator believes that takes college students in Massachusetts an average of 4.7 years to complete their degree.

H_0 : $\mu = 4.7$ (claim)
H_a : $\mu \neq 4.7$

Suppose that for a sample of 500 Massachusetts college students, the average time to graduate is 5.3 years.

We know that sample means vary from sample to sample. Assuming that the statement in the null hypothesis is true (i.e. that the population mean is 4.7 years), how likely was it that we obtained a sample mean of 5.3 years? Could this result be due to chance? Or might we conclude that the population mean is actually different than 4.7?

To determine how unusual a sample result is, we will calculate a test statistic based on the sampling distribution of the sample mean or sample proportion.

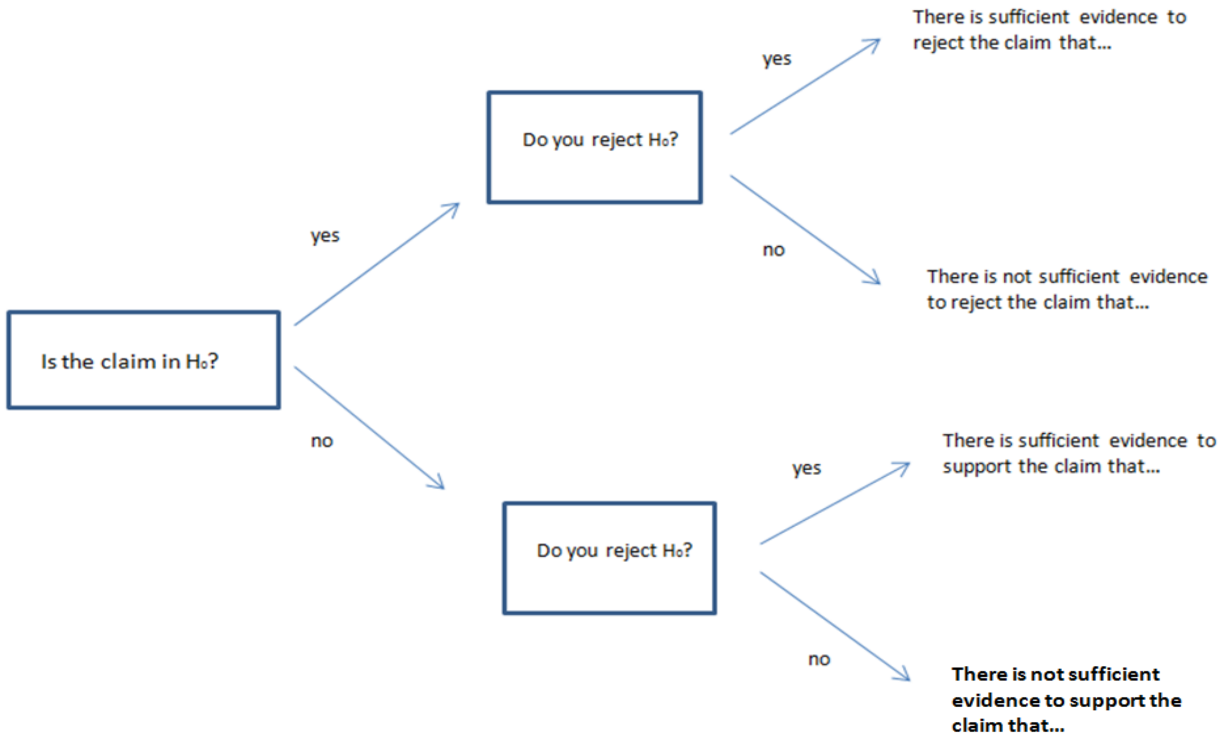
Make a Decision Using the P-value of a Statistical Test

The p-value is the probability that an outcome of the data (for example, the sample mean) will happen purely by chance when the null hypothesis is true. The smaller the p-value, the more unlikely the outcome, and the stronger the evidence is against the null hypothesis.

When you make a decision to reject or fail to reject H_0, do as follows:

- If p-value $\leq \alpha$, reject H_0.
- If p-value $> \alpha$, fail to reject H_0.

Writing a Conclusion



Outcomes and Type I and Type II Errors

When you perform a hypothesis test, there are four possible outcomes.

	H_0 is Actually True	H_0 is Actually False
Reject H_0	Type I error	Correct Outcome
Fail to reject H_0	Correct Outcome	Type II Error

There are two possible errors that could be made:

1. The decision is to reject H_o when, in fact, H_o is true (incorrect decision known as a Type I error).
2. The decision is not to reject H_o when, in fact, H_o is false (incorrect decision known as a Type II error).

Each of the errors occurs with a particular probability. The Greek letters α and β represent the probabilities.

α = probability of a Type I error = P(Type I error) = probability of rejecting the null hypothesis when the null hypothesis is true.

β = probability of a Type II error = P(Type II error) = probability of not rejecting the null hypothesis when the null hypothesis is false.

When conducting a hypothesis test, we will decide in advance how large of a Type I error we're willing to tolerate. This is called the <u>significance level</u> of the test. Often, researchers choose significance levels equal to 0.01, 0.05, or 0.10.

Testing a Claim about a Population Mean (μ) when the Population Standard Deviation (σ) is Known

Recall: Suppose X is a random variable with mean μ and standard deviation σ. For samples of size n, with sample mean $\bar{X}$, the random variable z = $\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$ has a standard normal distribution.

One or both of these conditions must be met:
1. The distribution of X is normal.
2. The sample size is greater than 30.

For all problems of this type, the test statistic will be z = $\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$

Example: A special cable supposedly has a breaking strength of 800 pounds. A researcher selects a sample of 31 cables and finds that the average breaking strength is 793 pounds. At $\propto$ = 0.01, test the claim that the population mean breaking strength is actually less than 800 pounds. Assume that the standard deviation of the population is known to be 12 pounds.

Solution:

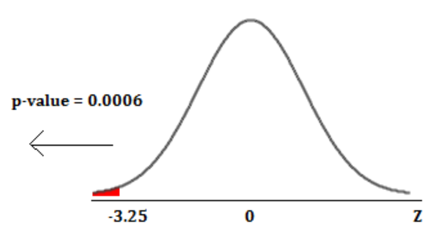
1. State the null and alternative hypothesis.

H_0: μ = 800
H_1: μ < 800 (claim)

2. Compute the test statistic.

$z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = \frac{793-800}{12/\sqrt{31}}$ = -3.25

3. Compute the p-value by plotting the test statistic and shading the left tail because this is a left-tailed test.



p-value = normalcdf(-9999, -3.25, 0, 1) = 0.0006
(Note that your calculator may return an answer in scientific notation.)

4. Since the p-value is less than α, reject H_0.

5. There is sufficient evidence to support the claim that the population mean breaking strength for this type of cable is less than 800 lbs.

Example: Suppose a researcher would like to test the hypothesis that the average number of hours that preschoolers are exposed to media per day is 3 hours. A sample of 500 preschoolers resulted in a sample mean time of 2.95 hours per day. Suppose the standard deviation of the population is known to be 1.05.

Test the claim that the population mean number of hours that preschoolers are exposed to media per day is 3 hours. Use $\alpha = 0.10$.

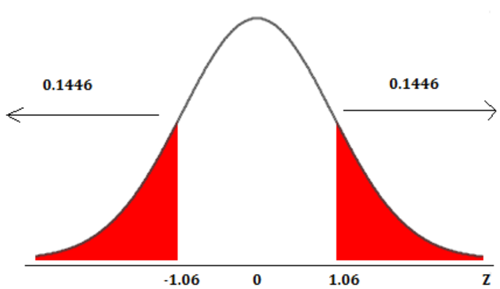
1. State the null and alternative hypothesis.

H_0: $\mu = 3$ (claim)
H_1: $\mu \neq 3$

2. Compute the test statistic.

$$z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = \frac{2.95-3}{1.05/\sqrt{500}} = -1.06$$

3. Compute the p-value by sketching a standard normal curve, plotting the positive and the negative of the test statistic, and shading both tails because this is a two-tailed test.



The area in the left tail is given by normalcdf(-9999, -1.06, 0, 1) = 0.1446. By symmetry, the area in the right tail is 0.1446. The p-value is the sum of these two areas, which is 0.2892.

4. Since the p-value is not less than $\propto$, fail to reject the null hypothesis.

5. There is not enough evidence to reject the claim that the population mean number of hours that preschoolers are exposed to media is equal to 3.

To conduct a hypothesis test for this type of problem by using your calculator,

STAT -> TESTS -> Z-Test -> Stats ->

μ_0: the population mean in H_0
σ: the population standard deviation
$\bar{X}$: the sample mean
n: the sample size
μ: the type of test (two-tailed, left-tailed, or right tailed)

Select Calculate

Testing a Claim about a Population Mean (μ) when the Population Standard Deviation (σ) is Unknown

Suppose X is a random variable with mean μ. For samples of size n, with sample mean $\bar{X}$ and sample standard deviation s, the random variable $t = \frac{\bar{x}-\mu}{s/\sqrt{n}}$ has a Student's t distribution with degrees of freedom d.f. = n – 1.

One or both of these conditions must be met:
1. The distribution of X is normal.
2. The sample size is greater than 30.

The test statistic that we will use for these types of problems is $t = \frac{\bar{x}-\mu}{s/\sqrt{n}}$

Example: Suppose that for a certain type of tree, the average height is 80 centimeters after two years of growth. A researcher would like to test whether the presence of a certain type of soil contaminant will stunt the growth of this type of tree. Twelve trees were grown in contaminated soil, and the sample mean height after two years was of 63 centimeters with a sample standard deviation of 17 centimeters. Test the hypothesis that the population mean height of trees grown in contaminated soil is less than 80 centimeters. Use a significance level of 0.01 and assume the original distribution is normal.

1. State the null and alternative hypothesis.

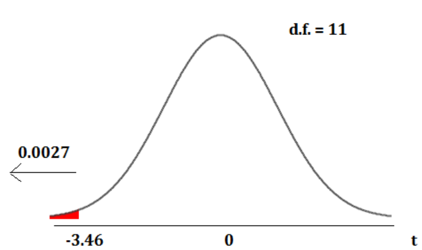
H_0: $\mu = 80$
H_1: $\mu < 80$ (claim)

2. Compute the test statistic.

$t = \frac{\bar{x}-\mu}{s/\sqrt{n}} = \frac{63-80}{17/\sqrt{12}}$ = -3.46

Note that the test statistic comes from a t distribution with d.f. = 12 – 1 = 11.

3. Compute the p-value by plotting the test statistic and shading the left tail because this is a left-tailed test.



p-value = tcdf(left bound, right bound, degrees of freedom) = tcdf(-9999, -3.46, 11) = 0.0027

4. Since the p-value is less than α, reject the null hypothesis.

5. There is sufficient evidence to conclude that the population mean height of trees grown in contaminated soil is less than 80 centimeters.

Example: The average family size in the United States is reported as 3.18. A random sample of 24 families in a particular school district resulted in the following family sizes:

5	4	5	4	4	4	3	6	4	3	3	5
6	3	3	2	7	4	5	2	2	2	3	5

At $\alpha = 0.05$, test the claim that the population mean family size in this school district differs from the national average.

Solution:

1. State the null and alternative hypothesis.

H_0: $\mu = 3.18$
H_1: $\mu \neq 3.18$ (claim)

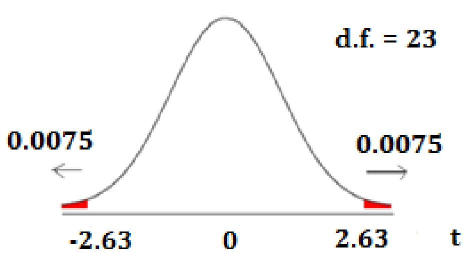
2. Compute the test statistic.

Using 1-Var Stats on the calculator, the sample mean is 3.92 and the sample standard deviation is 1.38.

$$t = \frac{\bar{x}-\mu}{s/\sqrt{n}} = \frac{3.92-3.18}{1.38/\sqrt{24}} = 2.63$$

Note that the test statistic comes from a t distribution with d.f. = 24 – 1 = 23.

3. To find the p-value, plot the positive and negative of the test statistic because this is a two-tailed test. The p-value is the sum of the areas in both tails.



The area to the left of -2.63 is tcdf(-9999, -2.63, 23) = 0.0075. By symmetry, the area to the right of 2.63 is 0.0075. Therefore, the p-value is 0.0075 + 0.0075 = 0.0150.

4. Since the p-value is less than α, reject the null hypothesis.

5. There is enough evidence to support the claim that the population mean family size in this school district differs from the national average.

To conduct a hypothesis test for this type of problem by using your calculator,

Case 1: If you are not given a data set but only summary statistics,

STAT -> TESTS -> T-Test -> Stats

μ_0: the population mean in H_0
$\bar{X}$: the sample mean
S_x: the sample standard deviation
n: the sample size
μ: the type of test (two-tailed, left-tailed, or right tailed)

Select Calculate

Case 2: If you are given a data set,

Enter the data into the statistics editor using STAT -> 1: Edit

STAT -> TESTS -> T-Test -> Data

μ_0: the population mean in H_0
List: L_1
Freq: 1
μ: the type of test (two-tailed, left-tailed, or right tailed)

Select Calculate

Testing a Claim about a Population Proportion (p)

Under certain conditions, the sampling distribution of the sample proportion $\hat{P}$ can be modeled by a normal distribution with a mean of $\mu_{\hat{p}} = p$ and a standard deviation of $\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$

Therefore, the random variable $z = \frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}}$ has a standard normal distribution.

The test statistic that we will use for these types of problems is $z = \frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}}$

Example: For a certain year a study reports that the nationwide percentage of college students using credit cards was 83%. Suppose that you believe that the proportion is different than 83%, and in a sample of 2000 college students, 1701 use credit cards. Test the claim that the proportion of college students using credit cards is different than 83%. Use $\propto = 0.05$.

1. State the null and alternative hypotheses.

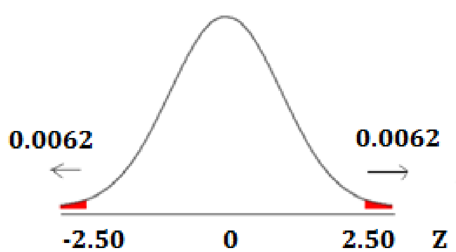
H_0: p = 0.83
H_1: p $\neq$ 0.83 (claim)

2. Compute the test statistic. The sample proportion is $\hat{p}$ = 1701/2000 = 0.851.

$$z = \frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}} = \frac{0.851-0.83}{\sqrt{\frac{(0.83)(0.17)}{2000}}} = 2.50$$

Note that the test statistic comes from a standard normal distribution.

3. Compute the p-value by sketching a standard normal curve, plotting the positive and the negative of the test statistic, and shading both tails because this is a two-tailed test.



The area in the left tail is given by normalcdf(-9999, -2.50, 0, 1) = 0.0062. By symmetry, the area in the right tail is 0.0062. The p-value is the sum of these two areas, which is 0.0124.

4. Reject H_0 because the p-value is less than α.

5. There is enough evidence to support the claim that the population proportion of college students who use credit cards is different than 83%.

Example: In one study of smokers who tried to quit smoking with nicotine patch therapy, 55% of 71 people surveyed were smoking one year after the treatment. Use a 5% significance level to test the claim that among smokers who try to quite with nicotine patch therapy, the majority are smoking a year after treatment.

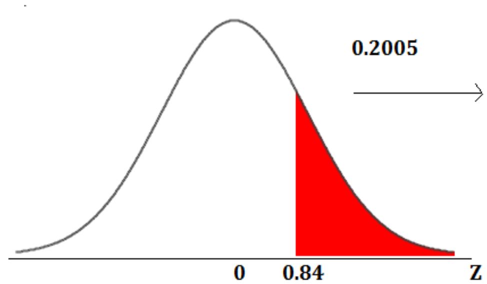
1. State the null and alternative hypothesis. (Note that majority means greater than 50%.)

H_0: p = 0.50
H_1: p > 0.50 (claim)

2. Compute the test statistic. The sample proportion is $\hat{p}$ = 0.55.

$$z = \frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}} = \frac{0.55-0.50}{\sqrt{\frac{(0.50)(0.50)}{71}}} = 0.84$$

3. Compute the p-value by plotting the test statistic and shading the right tail because this is a right-tailed test.



The area in the right tail is given by normalcdf(0.84, 9999, 0, 1) = 0.2005. Therefore, the p-value is 0.2005.

4. Fail to reject the null hypothesis because the p-value is not less than ∝.

5. There is not enough evidence to support the claim that among smokers who try to quit with nicotine patch therapy, the majority are smoking a year after the treatment.

To conduct a hypothesis test for this type of problem by using your calculator,

STAT -> TESTS -> 1-PropZTest

p_0: the population proportion in H_0
X: the number of individuals in the sample having the characteristic of interest**
n: the sample size
$prop$: the type of test (two-tailed, left-tailed, or right tailed)

Select Calculate

**If you are given the sample proportion in the problem, you will have to calculate this. For example, in the previous example, the number of people still smoking one year after treatment (X) is 55% of the 71 people surveyed. Therefore, X = 0.55*71 = 39.05, or 39 people. The calculator will not allow you to enter a decimal for X.