

## Big Picture

Tests of significance are done to test hypotheses. These hypotheses determine some sort of relationship between data. Tests of significance can tell us if these hypotheses are statistically relevant or not.

## Key Terms

**Hypothesis Testing:** Testing the difference between a hypothesized value of a parameter and the test statistic.

**Null Hypothesis ( $H_0$ ):** The default hypothesis, a hypothesis about a parameter that is tested.

**Alternate Hypothesis ( $H_a$ ):** The alternate hypothesis to be accepted if the default hypothesis is rejected.

**Critical Region:** The values of the test statistic that allow us to reject the null hypothesis.

**Level of Significance ( $\alpha$ ):** The strength of the sample evidence needed to reject the null hypothesis.

## Null and Alternate Hypotheses

In **hypothesis testing**, we make a **null hypothesis** and an **alternate hypothesis**.

- The null hypothesis  $H_0$  gives the population parameter that we're looking at a specific value.
  - Null hypothesis always includes the possibility of equality (i.e. includes =,  $\leq$ , or  $\geq$ ).
- The alternate hypothesis  $H_a$  is the opposite of the  $H_0$  hypothesis.
  - One important thing to realize is that the  $H_0$  and  $H_a$  hypothesis can NEVER be true at the same time.

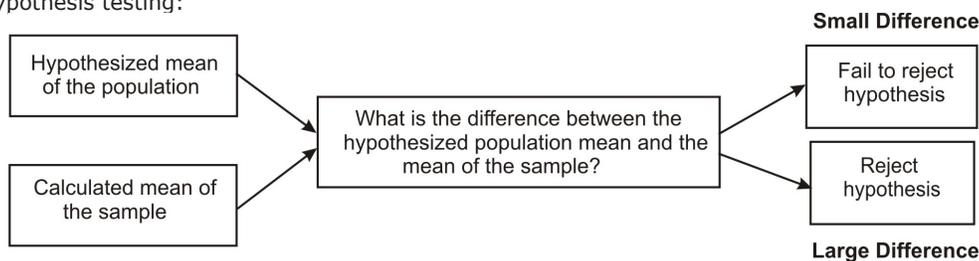
The hypotheses that we make depend on what the problem is asking for. We usually want to see if we have evidence to reject the null hypothesis.

Examples:

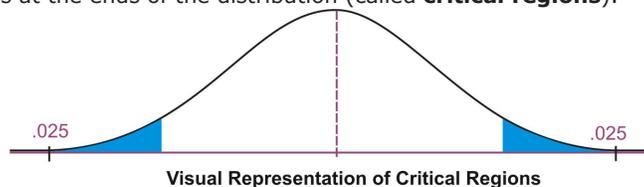
- A new medicine is supposed to have 3 milligrams (mg) of ingredient A.
  - A company wants to know if the new batch of the medicine has the right amount of ingredient A.  $H_0 = 3$  mg.
  - The company doesn't want the medicine to have more or less than specified, so  $H_a \neq 3$  mg.
- An automaker claims that its new car gets 45 miles per gallon (mpg).
  - We want to see if we have evidence to reject this claim, so we want to see if we can reject  $H_0 = 45$  mpg.
  - We assume that the automaker wants to advertise the highest miles per gallon possible, so if we reject the null hypothesis, we believe the mpg will be lower than advertised.  $H_a < 45$  mpg

## Hypothesis Testing

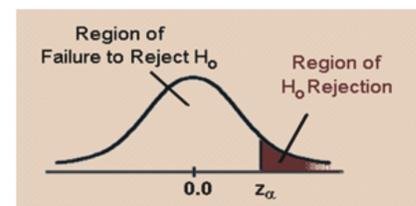
Process of hypothesis testing:



- Example 1 is an example of a two-tailed hypothesis test ( $H_a \neq$ ). We reject the hypothesis if our statistic falls in either of the blue regions at the ends of the distribution (called **critical regions**).



- Example 2 is an example of a one-tailed test ( $H_a <$  or  $H_a >$ ). The region of rejection can be on either side of the distribution (in the diagram on the right, the region of rejection is on the right side).



# TESTS OF SIGNIFICANCE CONT.

## Hypothesis Testing (cont.)

### Types of Errors

There are two types of errors that can be made:

	$H_0$ Is True	$H_0$ Is False
Fail to reject	Correct decision	Error (type II)
Reject	Error (type I)	Correct decision



*We don't "accept" the null hypothesis, we "fail to reject" it. We don't prove that the null hypothesis is true - we just show that we don't have enough evidence to reject it.*

Type I error:

- $H_0$  is true but we reject  $H_0$ .
- We let  $\alpha$  represent the probability of making a type I error.
- $\alpha$  is called the **level of significance**.



*Since being first is winning, being second is like failing to win. Thus, remember that Type II errors are when you FAIL to reject  $H_0$ .*

Type II error:

- $H_0$  is false but we fail to reject  $H_0$ .
- $\beta$  is the probability of making a type II error.
- $1 - \beta$  is called the power of a test.
- Factors that affect the power of a test:
  - Whether  $H_a$  is one- or two-sided
  - The significance level,  $\alpha$
  - Sample size

We usually pick the values of  $\alpha$  and  $\beta$  before looking at any sample data. We want these values to be as small as possible.

- Typical  $\alpha$  values are 0.05 and 0.01.
- Picking  $\alpha = 0.05$ , we are saying that in 100 similar situations, we are willing to accept that on average  $H_0$  will be rejected 5 times when it shouldn't be.

### P Value

The  $P$  value is another way to evaluate a hypothesis. The  $P$  value gives the smallest value of  $\alpha$  needed to reject the null hypothesis for the value of the test statistic that we use. It is the probability that a random variable will be greater than or equal to the observed value by chance:  $P(z > Z_{observed})$ .

- When  $P$  value  $\leq \alpha$ , reject  $H_0$
- When  $P$  value  $> \alpha$ , fail to reject  $H_0$

The  $P$  value is equal to the area to the right of the observed statistic if the test is a right-tailed test and equal to the area to the left in a left-tailed test. For a two-tailed test,  $P$  value is equal to the areas in the two tails.

### Testing a Mean Hypothesis

- Used to test whether there is a difference between a hypothesized value of  $\mu$  and the estimate of that parameter  $\bar{x}$
- For large samples ( $n \geq 30$ ), we say that  $\bar{x}$  distribution is approximately normal (central limit theorem)
- We can use  $z$  values to find critical regions  $z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$ 
  - $z$  = standardized score
  - $\bar{x}$  = sample mean
  - $\mu$  = population mean under null hypothesis
  - $\sigma$  = population standard deviation

1. Form a hypothesis.

- The goal is to find evidence against a claim.
- Null hypothesis: statement that is being tested.  $H_0: \mu =$  or  $\leq$  or  $\geq$  \_\_\_ (value).
- Alternative hypothesis: the statement that we believe is true.  $H_a: \mu >$  or  $<$  or  $\neq$  \_\_\_ (value)

2. Choose  $\alpha$ . Typically  $\alpha = 0.05$  or 0.01.

3. Check conditions and calculate  $\mu$ . Make sure that we can assume the distribution is normal.

4. Calculate critical region(s).

- Find the  $z$  values so that the area in the critical region(s) equals  $\alpha$ .
- If it is a two-tailed test, each critical region needs to equal  $\alpha/2$ .

5. Make a conclusion: reject or fail to reject.

- If the  $z$  value for  $\bar{x}$  falls in the critical region, we have statistically significant evidence to reject the null hypothesis and accept the alternate hypothesis. If the  $z$  value for  $\bar{x}$  is outside the critical region, we do not have statistically significant evidence to reject the null hypothesis.
- Statistically significant means the result is NOT likely to happen by chance alone.

If the sample is small, use the Student's  $t$ -distribution and the  $t$ -test statistic instead of the normal distribution and  $z$ .

