

# 7. HYPOTHESIS TESTING

Introduction, hypothesis test on mean and variance

# Introduction

# Inferential statistics

- Inferential statistics can be divided into two parts.
  - ▣ Parameter estimation
  - ▣ Hypothesis testing
- Hypothesis testing is a huge topic, and we will not cover everything. But hopefully we can introduce the basic concept of it.
- The hypothesis tests that we will cover relates to the population parameter (population mean, variance)

# Hypothesis tests

- The objective of hypothesis tests:
  - ▣ To test whether a sample data support or reject a claim regarding a population.
  
- The process:
  - ▣ Someone makes a claim regarding a population
  - ▣ Collect sample data
  - ▣ Analyse sample data
  - ▣ Make decision whether to support or reject the claim

# Hypothesis tests

- Important concepts:
  - ▣ Null and alternative hypotheses
  - ▣ Critical value, rejection regions, and tails of a test
  - ▣ Type I and Type II errors
  - ▣ Test statistic

# Two hypotheses

- For every hypothesis test, there will be two hypotheses or statements related to the population.
- **Null hypothesis,  $H_0$ :**
  - ▣ is a claim or statement about a population parameter that is assumed to be true until it is declared false
- **Alternative hypothesis,  $H_1$ :**
  - ▣ is a claim or statement about a population parameter that will be declared true if the null hypothesis is declared to be false

# Two hypotheses

- The two hypotheses cannot both be true.
  - ▣ Either  $H_0$  is true, and  $H_1$  is false;
  - ▣ Or  $H_0$  is false, and  $H_1$  is true.
  
- The hypothesis test assumes that the null hypothesis to be true...
  
- And try to find evidence to reject it.

# Example

- Consider as a nonstatistical example, a person has been indicted for committing a crime and is being tried in a court.
- The judge or jury can make one of the two possible decisions:
  - ▣ the person is not guilty
  - ▣ the person is guilty
- During the trial, the person is presumed not guilty. It is the prosecutor's job to prove that he has committed the crime and hence, is guilty.
- In this case, the null and alternative hypotheses are:
  - ▣  $H_0$ : The person is not guilty
  - ▣  $H_1$ : The person is guilty



# Type I and Type II errors

# Type I and Type II errors

- Sample data is used to determine if the null hypothesis should be rejected. Because this decision is based on sample data, there is a possibility that an incorrect decision can be made.
- There are four possibilities:

		Reality	
		$H_0$ is true	$H_0$ is false
Decision	Do not reject $H_0$	Correct decision	Type II error
	Reject $H_0$	Type I error	Correct decision

# Type I and Type II errors

- A **Type I error** occurs when a true null hypothesis is rejected.
- A **Type II error** occurs when a false null hypothesis is not rejected.
- We denote the probabilities of making Type I and Type II errors as  $\alpha$  and  $\beta$ , respectively:  
$$\alpha = P(\text{Type I error}) = P(\text{Reject } H_0 | H_0 \text{ is true})$$
$$\beta = P(\text{Type II error}) = P(\text{Do not reject } H_0 | H_0 \text{ is false})$$
- $\alpha$  is also called the **significance level** of the test and is chosen before hypothesis test is performed. Normally,  $\alpha$  is chosen to be 1%, 5%, or 10%.

# Type I and Type II errors

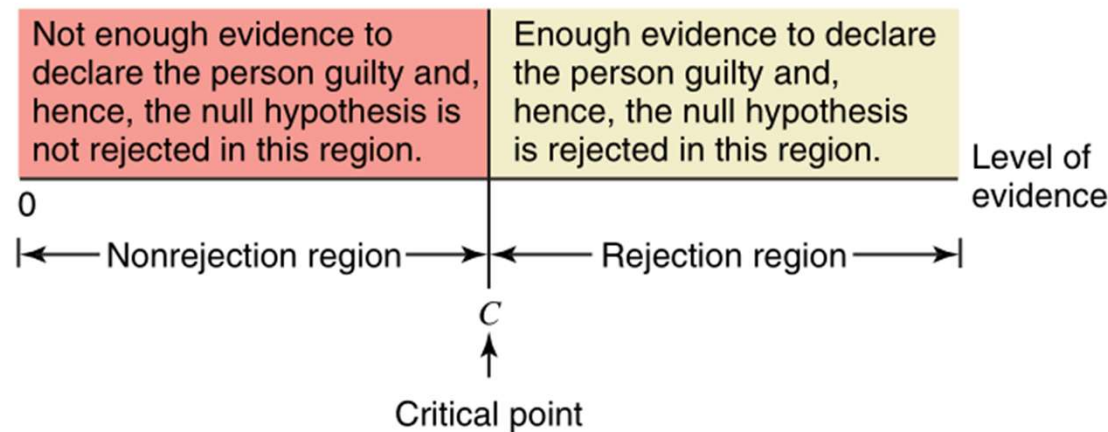
- We want both  $\alpha$  and  $\beta$  to be as low as possible so that making correct decision is more likely.
- However, for a fixed sample size, that is not possible.
  - ▣ If  $\alpha$  decreases,  $\beta$  will increase
  - ▣ If  $\beta$  decreases,  $\alpha$  will increase
- We can reduce both  $\alpha$  and  $\beta$  by increasing sample size.

# Type I and Type II errors

- In the court case example, the hypotheses are:
  - ▣  $H_0$ : The person is not guilty
  - ▣  $H_1$ : The person is guilty
  
- The two possible incorrect judgement or decisions are:
  - ▣ The person is found guilty by the court, although in reality he has not made the crime. (Type I error)
  - ▣ The person is found not guilty by the court, although in reality he has made the crime. (Type II error)

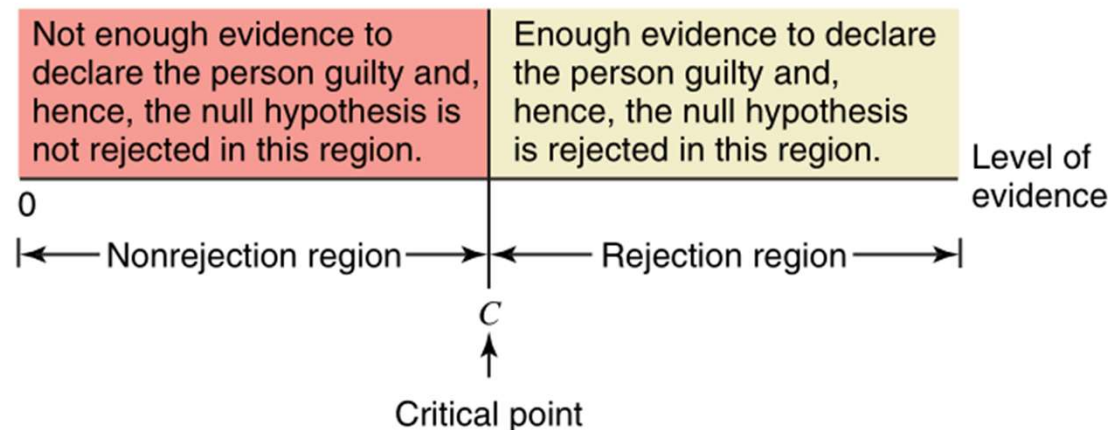
# Critical value, rejection regions, and test statistic

# Rejection and nonrejection region



- The figure represents the court case.
- The 0 on the left marks the point that there is no evidence against the person.
- As we have more and more convincing evidence, we move further to the right.
- We mark a point  $C$  on the horizontal axis to divide how to decide the outcome.

# Rejection and nonrejection region



- The point  $C$  is called **critical point** or **critical value**.
- If “evidence  $< C$ ”, we do not reject  $H_0$ . If “evidence  $> C$ ”, we reject  $H_0$ .
- The region at which the  $H_0$  will be rejected is called **rejection region**.
- The region at which the  $H_0$  will not be rejected is called **nonrejection region**.



# Test statistic

- The **test statistic** is a value calculated from the sample data that is used in the hypothesis test.
- The conclusion of a hypothesis test depends on this value.
- If the test statistic is in the rejection region,  $H_0$  will be rejected.
- Otherwise, if the test statistic is in the nonrejection region,  $H_0$  will not be rejected.
- More on this when we go through specific cases.

# The $p$ -value

- For making conclusion on the test, there are two approaches:
  - ▣ The critical value approach: compare the test statistic and the critical value or rejection region
  - ▣ The  $p$ -value approach: calculate 'p-value' and compare with the significance level
- The  $p$ -value is the probability of obtaining test results at least as extreme as the result observed from data, under the assumption that the null hypothesis is correct.
- If  $p\text{-value} < \alpha$ , then  $H_0$  is rejected. Otherwise, if  $p\text{-value} > \alpha$ , then  $H_0$  is not rejected.

# $p$ -value interpretation

$p\text{-value} > 0.10$	None or very little evidence against $H_0$
$0.05 < p\text{-value} < 0.10$	Weak evidence against $H_0$
$0.01 < p\text{-value} < 0.05$	Strong evidence against $H_0$
$p\text{-value} < 0.01$	Very strong evidence against $H_0$

- The smaller the  $p$ -value the more evidence there is against  $H_0$ .
- For example, if the  $p$ -value is less than 0.0001, then we have extremely strong evidence to reject  $H_0$  and support  $H_1$

# Hypothesis tests on parameter value

# Another example on hypothesis test

- A soft drink company claim on average, its cans contain at least 355 ml of soda. A consumer agency wants to check whether the company's claim is valid and collect a sample of 100 cans. Then it is found that the sample mean is 352 ml.
- Is the company's claim valid?
- Let  $\mu$  be the population mean of soda content in the cans.
- The two hypotheses:
  - ▣  $H_0: \mu \geq 355$  ml (company's claim)
  - ▣  $H_1: \mu < 355$  ml

# Hypothesis tests for parameter value

- In hypothesis tests for parameter value:
  - ▣ Null hypothesis,  $H_0$ 
    - Will always contain the equality:  $=, \leq, \geq$
    - Eg:  $\mu = 0, \mu \leq 5$
  - ▣ Alternative hypothesis,  $H_1$ 
    - Will always contain inequality:  $\neq, <, >$
    - Eg:  $\mu \neq 0, \mu > 5$
- Remember that we are testing on population parameters, not sample statistics.
  - ▣ Eg: we test on  $\mu$  that is unknown, not  $\bar{x}$  that is known

# Exercise

- State the null and alternative hypotheses:
  - a) A manufacturer claims that the average lifetime of its lightbulbs is 8000 hours. A consumer association want to test whether the average lifetime of the lightbulb is less than the manufacturer's claim.
  - b) The diameter of tennis balls manufactured by a factory must have a mean of 3 inches. The factory's manager is worried if the tennis balls manufactured does not meet the specification and wants to test it.
  - c) A dairy processing company claims that the variance of the amount of fat in the whole milk processed by the company is no more than 0.25. You suspect this is wrong and collect a random sample to test the claim.

a)  $\mu$ : average lifetime of lightbulbs

$$H_0: \mu \geq 8000 \quad (\text{manufacturer's claim}) \quad (\mu = 8000)$$

$$H_1: \mu < 8000$$

b)  $\mu$ : mean diameter of tennis balls.

$$H_0: \mu = 3$$

$$H_1: \mu \neq 3$$

c)  $\sigma^2$ : variance of the amount of fat in milk

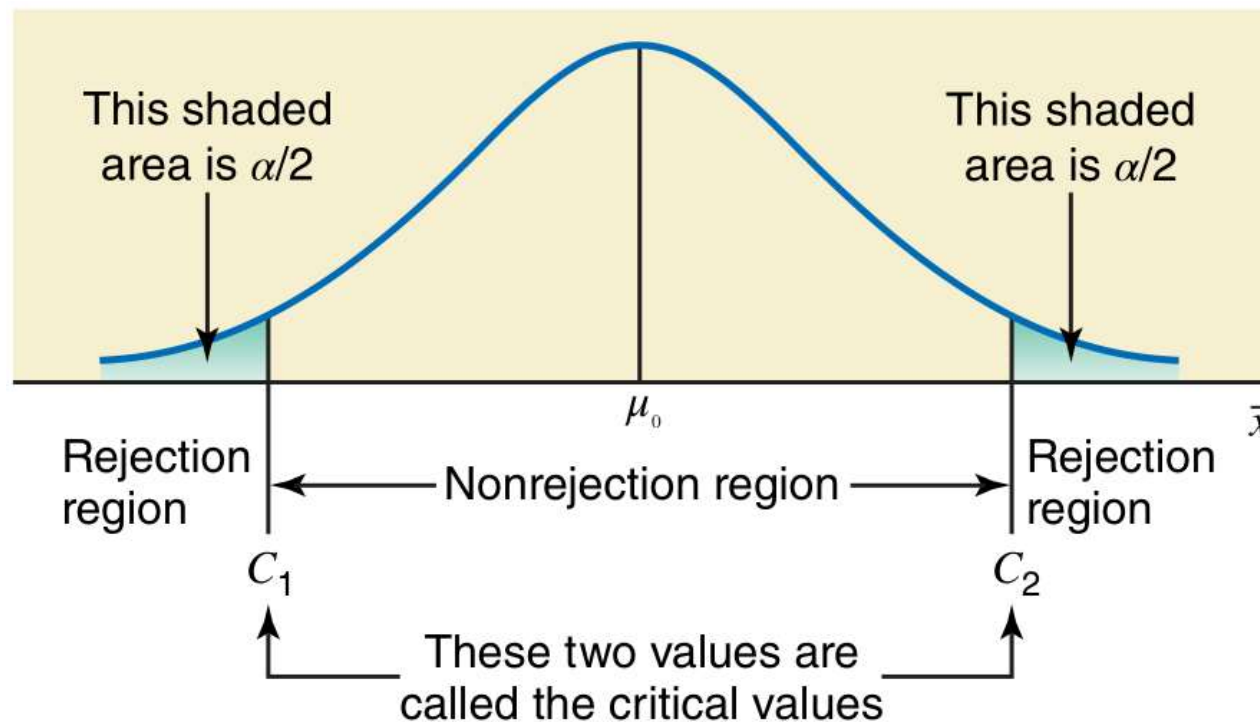
$$H_0: \sigma^2 \leq 0.25 \quad (\text{company's claim})$$

$$H_1: \sigma^2 > 0.25$$



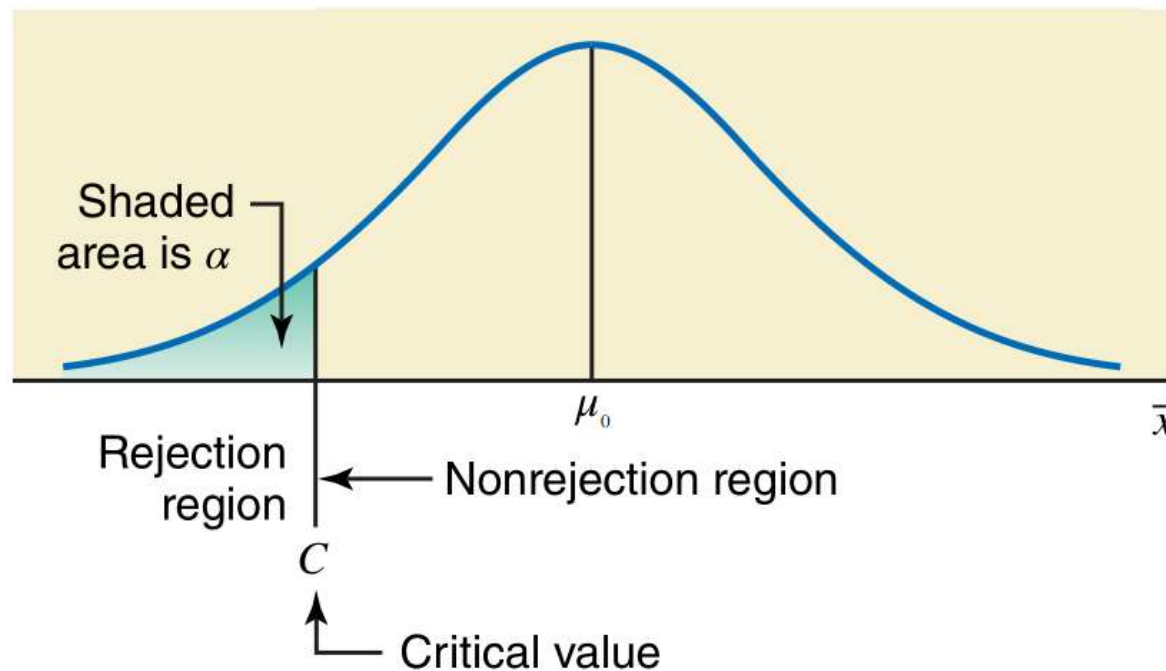
# Rejection regions

- Two-tailed test ( $\mu_0$  is just some value)
  - ▣  $H_0: \mu = \mu_0; H_1: \mu \neq \mu_0$
  - ▣ Reject  $H_0$  if  $|\text{test statistic}| > \text{critical value}$
  - ▣  $p$ -value is twice the area to the right of the positive test statistic, or twice the area to the left of the negative test statistic..



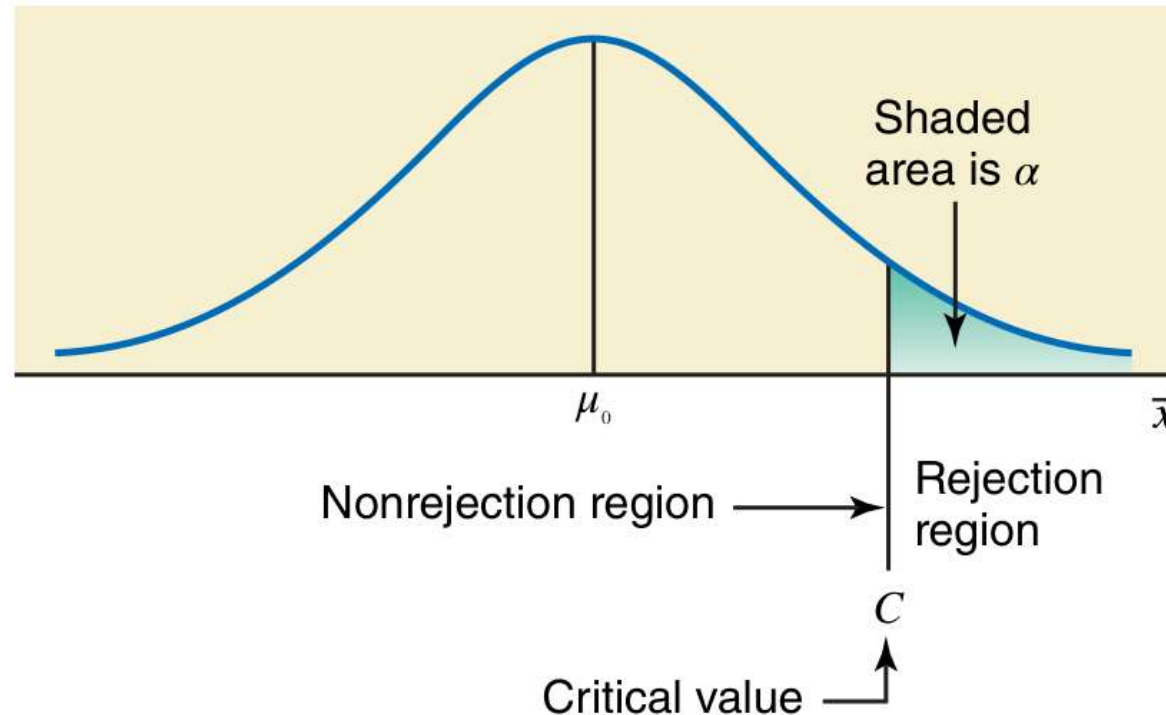
# Rejection regions

- Left-tailed test  
 $(\mu > \mu_0)$ 
  - ▣  $H_0: \mu = \mu_0, H_1: \mu < \mu_0$
  - ▣ Reject  $H_0$  if test statistic  $<$  critical value
  - ▣  $p$ -value is the area to the left of the test statistic



# Rejection regions

- Right-tailed test:
  - ▣  $H_0: \mu = \mu_0; H_1: \mu > \mu_0$
  - ▣ Reject  $H_0$  if test statistic  $>$  critical value
  - ▣  $p$ -value is the area to the right of the test statistic



# Hypothesis test steps

- In general, these are the steps for performing hypothesis test:
  1. Choose the significance level  $\alpha$
  2. State the null and alternative hypotheses
  3. Calculate the test statistic
  4. Make conclusion based on the test statistic
  
- As mentioned, there are two approaches for making conclusion:
  - ▣ The critical value approach
  - ▣ The  $p$ -value approach

# Making conclusion

- In hypothesis test, decision is made whether to reject  $H_0$  or not to reject  $H_0$ .
- If  $H_0$  is rejected, then the test supports  $H_1$ .
- The conclusion of the test can be made on  $H_0$  or  $H_1$ :

Decision	Claim on $H_0$	Claim on $H_1$
Reject $H_0$	There is enough evidence to <b>reject</b> $H_0$	There is enough evidence to <b>support</b> $H_1$
Do not reject $H_0$	There is not enough evidence to reject $H_0$	There is not enough evidence to support $H_1$

# Exercise

- The director of a medical hospital feels that on average, her surgeons perform **fewer** operations per year than the national average of **211**. She selected a random sample of surgeons to test this belief.
  - ▣ State the null and alternative hypotheses.
  - ▣ If the null hypothesis is rejected, state the conclusion that can be made.
  - ▣ If the null hypothesis is not rejected, state the conclusion that can be made.

a)  $\mu$  = average operations per year performed  
by surgeons in the hospital.

$$H_0: \mu \geq 211$$

$$H_1: \mu < 211 \quad (\text{director's belief})$$

b) There is enough evidence to support the director's belief that the average operations per year performed by surgeons in the hospital is less than the national average  
(evidence to support  $H_1$ )

c) There is not enough evidence to support the director's belief.

(not enough evidence to support  $H_1$ ).

# Hypothesis test for population mean (one population)



# Hypothesis test for population mean (one population)

- In this test, we want to test for a normally distributed population, whether:
  - ▣  $H_0: \mu = \mu_0$  vs  $H_1: \mu \neq \mu_0$  (two-tailed test)
  - ▣  $H_0: \mu = \mu_0$  vs  $H_1: \mu < \mu_0$  (left-tailed test)
  - ▣  $H_0: \mu = \mu_0$  vs  $H_1: \mu > \mu_0$  (right-tailed test)for some value  $\mu_0$
  
- Note that we use equality sign “=” for all the null hypotheses above for simplicity.

# Hypothesis test for population mean (one population)

- There are many cases:
  - ▣  $\sigma$  is known
  - ▣  $\sigma$  is unknown but the sample size is large
  - ▣  $\sigma$  is unknown but the sample size is small
- Each case has different test statistic and/or distribution to be used.
- We will focus on the first two cases only that uses normal distribution

# Case 1: $\sigma$ is known

□ If  $H_0$  is assumed to be true, then  $\bar{X} \sim N\left(\mu_0, \frac{\sigma^2}{n}\right)$ .

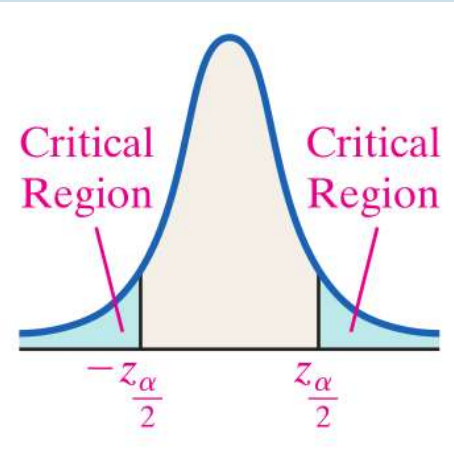
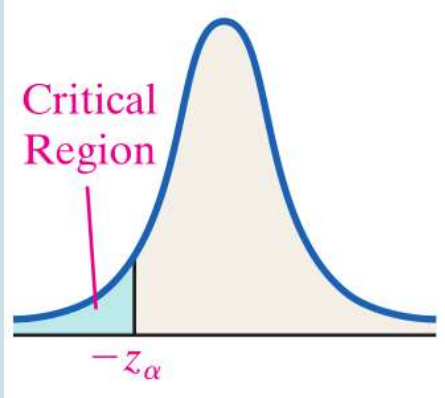
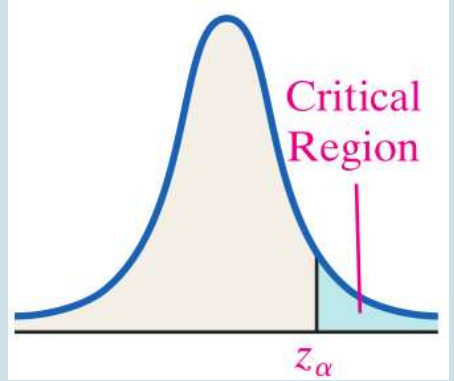
□ If  $\sigma$  is known,

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

can be used as the test statistic. We know that if  $H_0$  is true, then this value follows the standard normal distribution.

□ This test is also called the z-test.

# Case 1: $\sigma$ is known

$H_0$ and $H_1$	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$
Test statistic	$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$		
Decision rule	Reject $H_0$ if $ z  > z_{\alpha/2}$	Reject $H_0$ if $z < -z_\alpha$	Reject $H_0$ if $z > z_\alpha$
$p$ -value	$2P(Z >  z )$	$P(Z < z)$	$P(Z > z)$
Rejection region and critical value			

$$P(z > z_\alpha) = \alpha$$

# Critical values

- For normal distribution, here are the critical values:

Significance level $\alpha$	Left-tailed ( $-z_{\alpha}$ )	Right-tailed ( $z_{\alpha}$ )	Two-tailed ( $z_{\alpha/2}$ )
0.10	-1.28	1.28	$\pm 1.645$
0.05	-1.645	1.645	$\pm 1.96$
0.01	-2.33	2.33	$\pm 2.576$

# Example

- A researcher believes that the mean age of medical doctors in a large hospital system is older than the average age of doctors in the United States, which is 46. Assume the population standard deviation is 4.2 years. A random sample of 30 doctors from the system is selected, and the mean age of the sample is 48.6. Test the claim at  $\alpha = 0.05$ . Given  $z_{0.05} = 1.645$ .
- 

- Step 1: state the hypotheses:

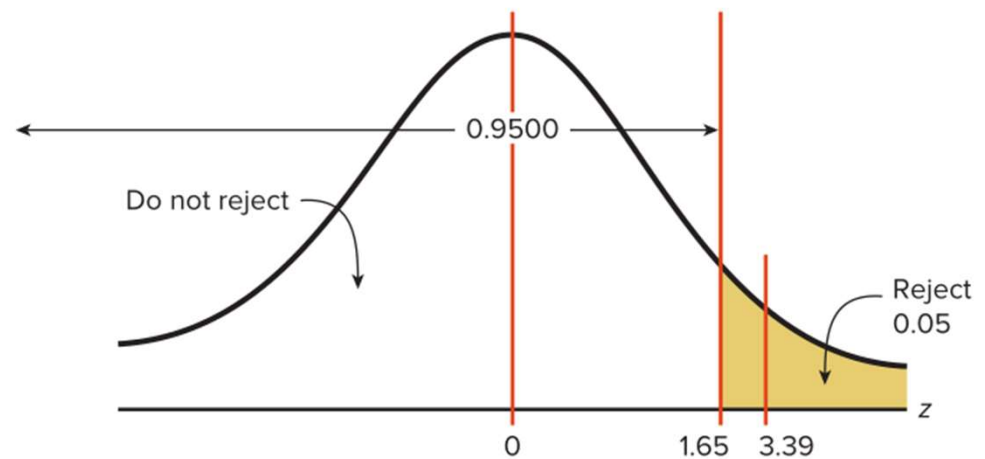
$$H_0: \mu = 46, \quad H_1: \mu > 46$$

- Step 2: calculate the test statistic

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{48.6 - 46}{4.2 / \sqrt{30}} = 3.391$$

# Example

- Step 3: Make conclusion
  - ▣ Reject  $H_0$  if the test statistic  $z > z_\alpha = 1.645$ .
  - ▣ In this case, it is found that  $z > 1.645$ .
  - ▣ Therefore, we have enough evidence to reject  $H_0$ .
  - ▣ There is enough evidence to support the claim that the mean age is more than national average.



# Exercise

- A researcher wishes to test the claim that the average cost of tuition and fees at a four-year public college is greater than \$5700. She selects a random sample of 9 four-year public colleges and finds the mean to be \$5950. The population standard deviation is assumed to be \$659. Is there evidence to support the claim at  $\alpha = 0.05$ ? Use  $z_{0.05} = 1.645$

① Hypotheses :  $H_0: \mu = 5700$

$H_1: \mu > 5700$

② Test statistic :  $z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{5950 - 5700}{659 / \sqrt{9}} = 1.138$



③ Conclusion :

Reject  $H_0$  if  $z > z_\alpha = 1.645$ .

We found that  $z < z_\alpha$ .

we do not have enough evidence to reject  $H_0$ .

There is not enough evidence to support that the average cost of tuition is greater than \$5700.

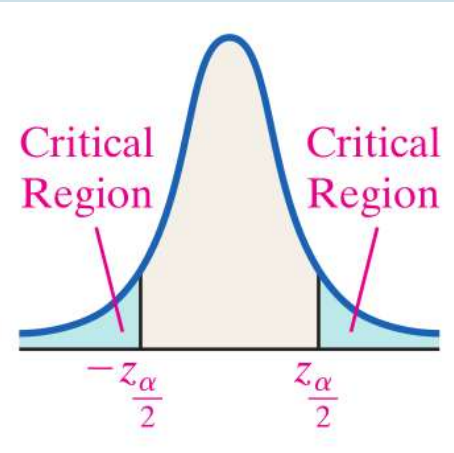
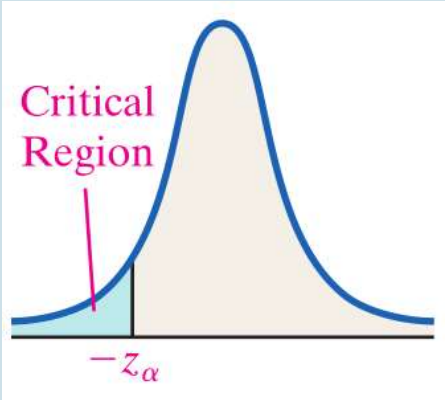
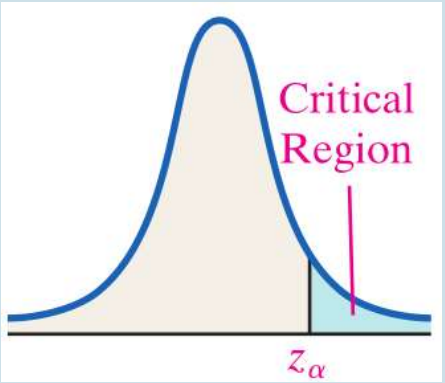
## Case 2: $\sigma$ is unknown but sample size is large

- If  $\sigma$  is not known and  $n \geq 30$ , we can use the sample standard deviation  $s$  as an estimate for  $\sigma$  and use

$$z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

as the test statistic. Under the null hypothesis, this value follows approximately the standard normal distribution.

# Case 2: $\sigma$ is unknown but sample size is large

$H_0$ and $H_1$	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$
Test statistic	$z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$		
Decision rule	Reject $H_0$ if $ z  > z_{\alpha/2}$	Reject $H_0$ if $z < -z_{\alpha}$	Reject $H_0$ if $z > z_{\alpha}$
$p$ -value	$2P(Z >  z )$	$P(Z < z)$	$P(Z > z)$
Rejection region and critical value			

# Example

- A researcher claims that the average wind speed in a certain city is 8 miles per hour. A sample of 32 days has an average wind speed of 8.2 miles per hour and a standard deviation 0.6 mile per hour. At  $\alpha = 0.05$ , is there enough evidence to reject the claim? Use the  $p$ -value approach.
- 

- Step 1: Hypotheses

$$H_0: \mu = 8, \quad H_1: \mu \neq 8$$

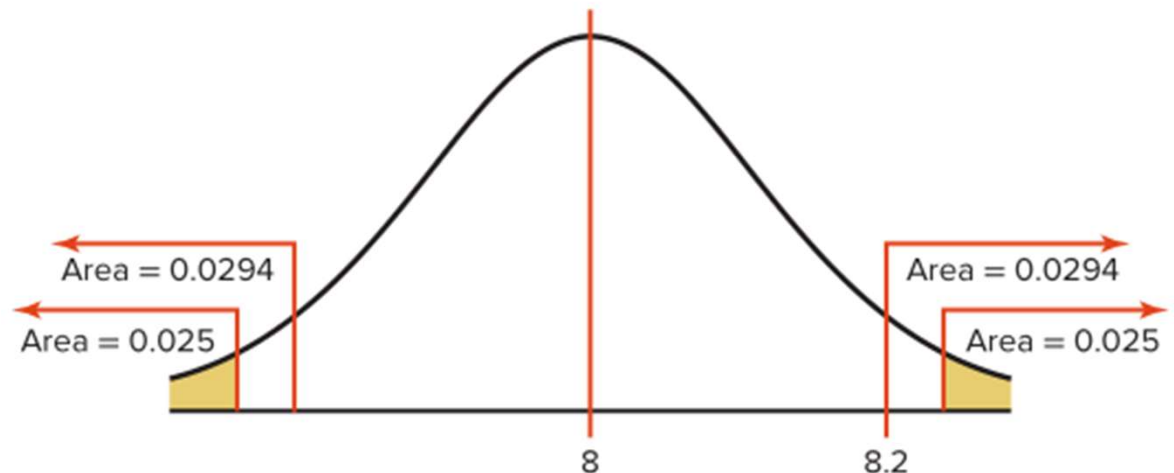
- Step 2: Test statistic

$$z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{8.2 - 8}{0.6/\sqrt{32}} = 1.886$$

# Example

## □ Step 3: Conclusion

- ▣  $p\text{-value} = 2 \times P(Z > |1.886|) = 2(0.0294) = 0.0588.$
- ▣  $p\text{-value} > \alpha = 0.05.$
- ▣ Therefore, we do not have enough evidence to reject  $H_0.$
- ▣ There is not enough evidence to reject the claim that the average wind speed is 8 miles per hour.



# Exercise

- The management of Priority Health Club claims that its members lose an average of 10 pounds or more within the first month after joining the club. A consumer agency that wanted to check this claim took a random sample of 36 members of this health club and found that they lost an average of 9.2 pounds within the first month of membership with the standard deviation of 2.4 pounds. Test the claim using 1% significance level using the  $p$ -value approach.

$$H_0: \mu = 10 \quad (\mu \geq 10) \quad (\text{claim})$$

$$H_1: \mu < 10$$

$$z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{9.2 - 10}{2.4/\sqrt{36}} = -2$$

$$\begin{aligned} p\text{-value} &= P(Z < z) = P(Z < -2) = P(Z > 2) \\ &= 0.02275 \end{aligned}$$

Reject  $H_0$  if  $p\text{-value} < \alpha = 0.01$

In this case  $p\text{-value} > \alpha$ .

We do not have enough evidence to reject  $H_0$ .

There is not enough evidence to reject the claim made by the club.

# Case 3: $\sigma$ is unknown and sample size is small

- If  $\sigma$  is not known and  $n < 30$ , we can use

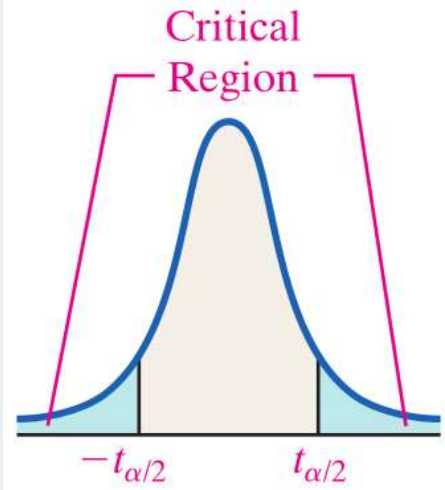
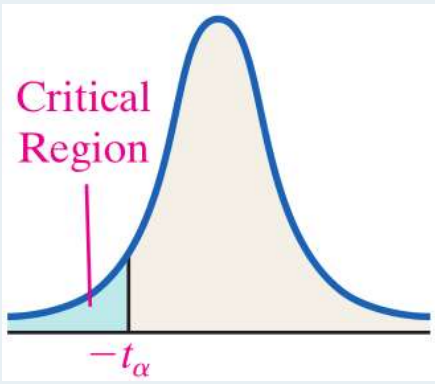
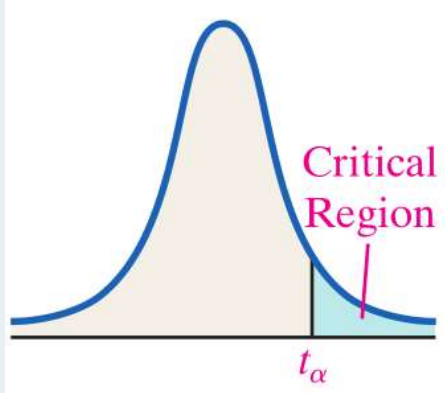
$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$

as the test statistic. This value follows the t-distribution with  $n - 1$  degrees of freedom

- This test is called the t-test.



# Case 3: $\sigma$ is unknown and sample size is small

$H_0$ and $H_1$	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$
Test statistic	$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$		
Decision rule	Reject $H_0$ if $ t  > t_{\alpha/2, n-1}$	Reject $H_0$ if $t < -t_{\alpha, n-1}$	Reject $H_0$ if $t > t_{\alpha, n-1}$
Rejection region and critical value			

# Hypothesis test for population means (two populations)

# Difference between two population means

- Suppose we have two sample means,  $\bar{X}_1$  and  $\bar{X}_2$ 
  - ▣  $\bar{X}_1$  is a sample mean from a  $N(\mu_1, \sigma_1^2)$  population with sample size  $n_1$
  - ▣  $\bar{X}_2$  is a sample mean from a  $N(\mu_2, \sigma_2^2)$  population with sample size  $n_2$
- In this case, we want to compare the two population means,  $\mu_1$  and  $\mu_2$  by taking the difference  $\mu_1 - \mu_2$ .
- Two cases:
  - ▣ Paired/dependent samples
  - ▣ Independent samples

# Dependent samples

- In this case, each observation in the first sample has its pair in the second sample, and the sample size must be the same.
- Example:
  - ▣ Marks for the first and second tests for the same students
  - ▣ Weights before and after a weight loss program for the same people
- To perform hypothesis test, we take the difference for each pair,  $d = x_1 - x_2$ , and perform hypothesis test on the difference like we did in the one population case.

# Independent samples

- Few cases:
  - ▣  $\sigma_1$  and  $\sigma_2$  are known
  - ▣  $\sigma_1$  and  $\sigma_2$  are not known but the sample size is large
  - ▣  $\sigma_1$  and  $\sigma_2$  are not known, sample size is small and  $\sigma_1 = \sigma_2$
  - ▣  $\sigma_1$  and  $\sigma_2$  are not known, sample size is small and  $\sigma_1 \neq \sigma_2$
- We focus on the first two cases only because they use normal distribution

# The hypotheses

- For test involving two population means, we are interested in the difference between the two means.
- Here are the three possible tests, for a given value of  $\delta_0$ :
  - ▣  $H_0: \mu_1 - \mu_2 = \delta_0$  vs  $H_1: \mu_1 - \mu_2 \neq \delta_0$  (two-tailed test)
  - ▣  $H_0: \mu_1 - \mu_2 = \delta_0$  vs  $H_1: \mu_1 - \mu_2 < \delta_0$  (left-tailed test)
  - ▣  $H_0: \mu_1 - \mu_2 = \delta_0$  vs  $H_1: \mu_1 - \mu_2 > \delta_0$  (right-tailed test)
- In most cases, we use  $\delta_0 = 0$ .

# Case 1: variances are known

- If  $X_1 \sim N(\mu_1, \sigma_1^2)$  and  $X_2 \sim N(\mu_2, \sigma_2^2)$ , then  $\bar{X}_1 \sim N\left(\mu_1, \frac{\sigma_1^2}{n_1}\right)$ ,  $\bar{X}_2 \sim N\left(\mu_2, \frac{\sigma_2^2}{n_2}\right)$ , and

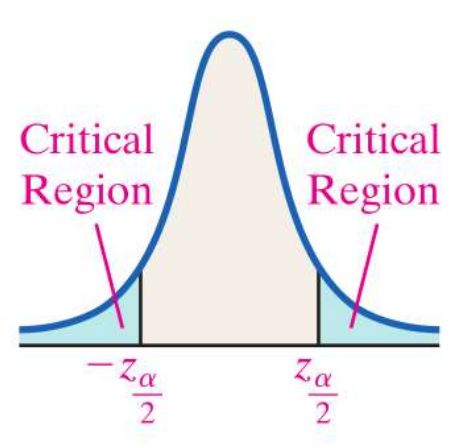
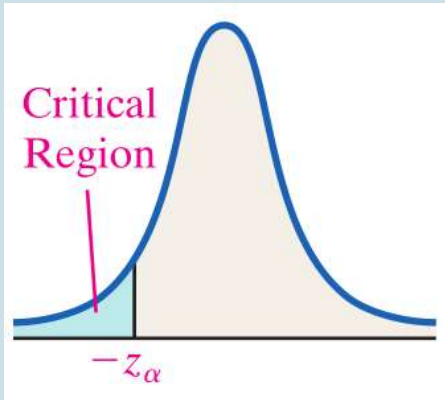
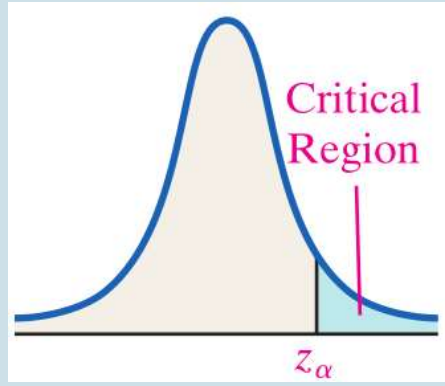
$$\bar{X}_1 - \bar{X}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

- If  $H_0$  is assumed true, then  $\mu_1 - \mu_2 = \delta_0$
- Test statistic:

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - \delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

follows the standard normal distribution

# Case 1: variances are known

$H_0$ and $H_1$	$H_0: \mu_1 - \mu_2 = \delta_0$ $H_1: \mu_1 - \mu_2 \neq \delta_0$	$H_0: \mu_1 - \mu_2 = \delta_0$ $H_1: \mu_1 - \mu_2 < \delta_0$	$H_0: \mu_1 - \mu_2 = \delta_0$ $H_1: \mu_1 - \mu_2 > \delta_0$
Test statistic	$z = \frac{(\bar{x}_1 - \bar{x}_2) - \delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$		
Decision rule	Reject $H_0$ if $ z  > z_{\alpha/2}$	Reject $H_0$ if $z < -z_{\alpha}$	Reject $H_0$ if $z > z_{\alpha}$
$p$ -value	$2P(Z >  z )$	$P(Z < z)$	$P(Z > z)$
Rejection region and critical value			



# Example

- A study using two random samples of 35 people each found that the average amount of time those in the age group of 26–35 years spent per week on leisure activities was 39.6 hours, and those in the age group of 46–55 years spent 35.4 hours. Assume that the population standard deviation for those in the first age group found by previous studies is 6.3 hours, and the population standard deviation of those in the second group found by previous studies was 5.8 hours. At  $\alpha = 0.05$ , can it be concluded that there is a significant difference in the average times each group spends on leisure activities? Use  $z_{0.025} = 1.96$ .

$$n_1 = 35$$

$$n_2 = 35$$

$$\bar{x}_1 = 39.6$$

$$\bar{x}_2 = 35.4$$

$$\sigma_1 = 6.3$$

$$\sigma_2 = 5.8$$

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 - \mu_2 \neq 0$$

$$\begin{aligned}
 z &= \frac{(\bar{x}_1 - \bar{x}_2) - \delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \\
 &= \frac{(39.6 - 35.4) - 0}{\sqrt{\frac{6.3^2}{38} + \frac{5.8^2}{35}}} \\
 &= 2.902
 \end{aligned}$$

Reject  $H_0$  if  $|z| > z_{\alpha/2}$

It is found that

$$|z| > z_{\alpha/2} = 1.96$$

We have enough evidence to reject  $H_0$ .

There is enough evidence to conclude that there is a significance difference in the amount of leisure activities between the two groups.

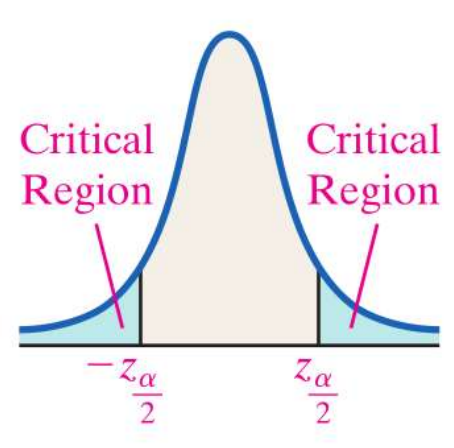
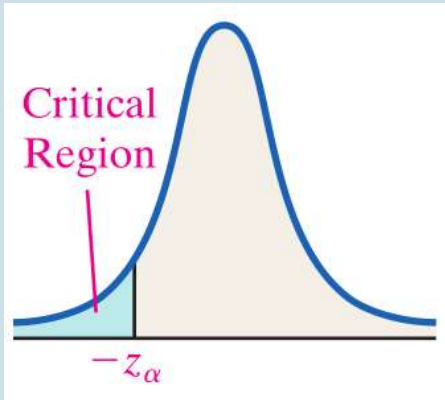
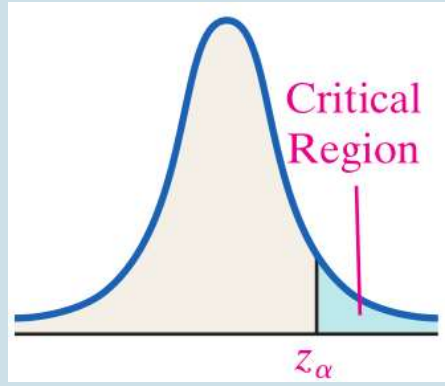
## Case 2: variances unknown but size sample is large

- If the population standard deviations are not known but the sample size is large, the sample standard deviations  $s_1$  and  $s_2$  can be used as estimates for  $\sigma_1$  and  $\sigma_2$ , respectively
- Test statistic:

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - \delta_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

follows the standard normal distribution

# Case 2: variances unknown but size sample is large

$H_0$ and $H_1$	$H_0: \mu_1 - \mu_2 = \delta_0$ $H_1: \mu_1 - \mu_2 \neq \delta_0$	$H_0: \mu_1 - \mu_2 = \delta_0$ $H_1: \mu_1 - \mu_2 < \delta_0$	$H_0: \mu_1 - \mu_2 = \delta_0$ $H_1: \mu_1 - \mu_2 > \delta_0$
Test statistic	$z = \frac{(\bar{x}_1 - \bar{x}_2) - \delta_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$		
Decision rule	Reject $H_0$ if $ z  > z_{\alpha/2}$	Reject $H_0$ if $z < -z_{\alpha}$	Reject $H_0$ if $z > z_{\alpha}$
$p$ -value	$2P(Z >  z )$	$P(Z < z)$	$P(Z > z)$
Rejection region and critical value			

# Example

- A researcher hypothesizes that the average number of sports that colleges offer for males is greater than the average number of sports that colleges offer for females. A random sample of the number of sports offered by 50 colleges for males and 50 colleges for females gives the average of 8.6 sports and 7.9 sports for males and females, respectively, and a standard deviation of 3.3 for both groups. At  $\alpha = 0.10$ , is there enough evidence to support the claim?

$$n_1 = 50$$

$$\bar{x}_1 = 8.6$$

$$s_1 = 3.3$$

(male)

$$n_2 = 50$$

$$\bar{x}_2 = 7.9$$

$$s_2 = 3.3$$

(female)

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 - \mu_2 > 0$$

$$\begin{aligned}
 z &= \frac{(\bar{x}_1 - \bar{x}_2) - \delta_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \\
 &= \frac{(8.6 - 7.9) - 0}{\sqrt{\frac{3.3^2}{50} + \frac{3.3^2}{50}}} \\
 &= 1.06
 \end{aligned}$$

Reject  $H_0$  if  $z > z_\alpha$   
 $= 1.28$

We found that  $z < z_\alpha$ .

We do not have enough evidence to reject  $H_0$ .

There is not enough evidence to support the claim that the average number of sports for males is more than for females.

## Case 3: variances unknown, sample size is small and $\sigma_1 = \sigma_2$

- If the variances are unknown, the size sample is small, but the two variances are (assumed to be) equal, then we can use the following pooled variance as an estimate:

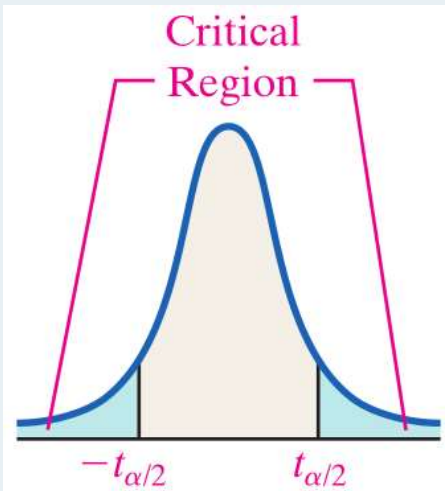
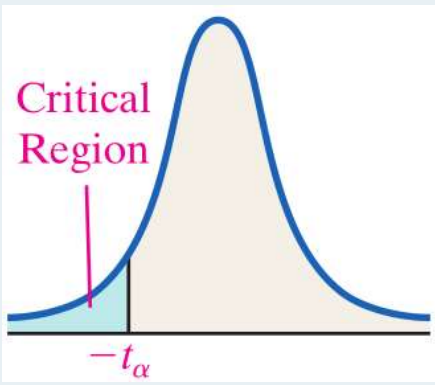
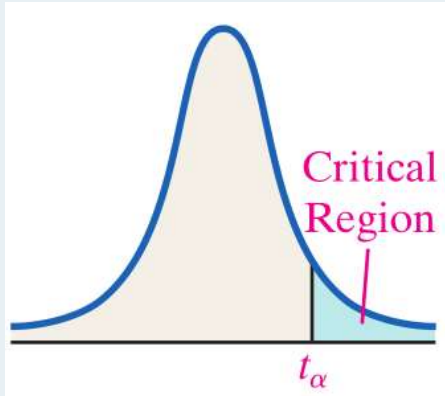
$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

- Test statistic:

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - \delta_0}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

follows the t-distribution with  $n_1 + n_2 - 2$  degrees of freedom.

# Case 3: variances unknown, sample size is small and $\sigma_1 = \sigma_2$

$H_0$ and $H_1$	$H_0: \mu_1 - \mu_2 = \delta_0$ $H_1: \mu_1 - \mu_2 \neq \delta_0$	$H_0: \mu_1 - \mu_2 = \delta_0$ $H_1: \mu_1 - \mu_2 < \delta_0$	$H_0: \mu_1 - \mu_2 = \delta_0$ $H_1: \mu_1 - \mu_2 > \delta_0$
Test statistic	$t = \frac{(\bar{x}_1 - \bar{x}_2) - \delta_0}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \quad \text{where} \quad s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$		
Decision rule	Reject $H_0$ if $ t  > t_{\alpha/2, n_1+n_2-2}$	Reject $H_0$ if $t < -t_{\alpha, n_1+n_2-2}$	Reject $H_0$ if $t > t_{\alpha, n_1+n_2-2}$
Rejection region and critical value			



# Case 4: variance unknown and sample size is small

- Test statistic:

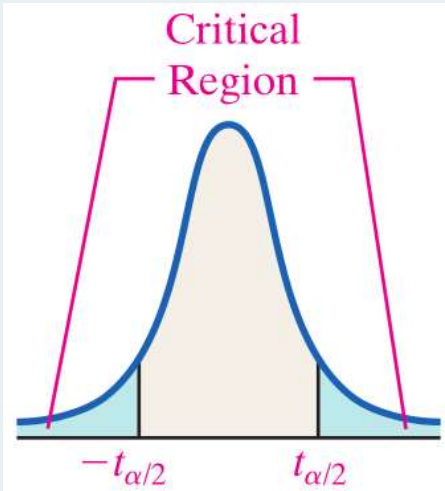
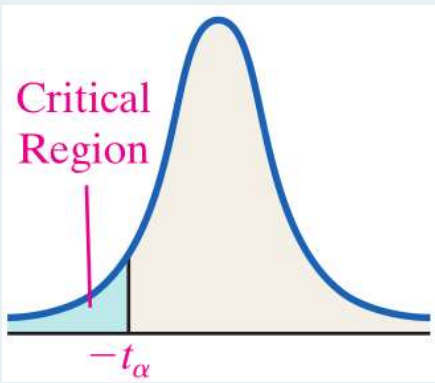
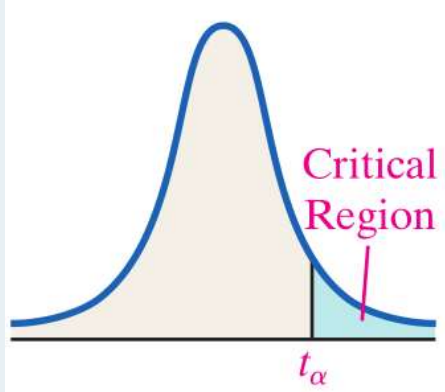
$$t = \frac{(\bar{x}_1 - \bar{x}_2) - \delta_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

follows the t-distribution with the following degrees of freedom:

$$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1} \left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1} \left(\frac{s_2^2}{n_2}\right)^2}$$

- This is also called Welch's t-test.

# Case 4: variance unknown and sample size is small

$H_0$ and $H_1$	$H_0: \mu_1 - \mu_2 = \delta_0$ $H_1: \mu_1 - \mu_2 \neq \delta_0$	$H_0: \mu_1 - \mu_2 = \delta_0$ $H_1: \mu_1 - \mu_2 < \delta_0$	$H_0: \mu_1 - \mu_2 = \delta_0$ $H_1: \mu_1 - \mu_2 > \delta_0$
Test statistic	$t = \frac{(\bar{x}_1 - \bar{x}_2) - \delta_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$		
Decision rule	Reject $H_0$ if $ t  > t_{\alpha/2, \nu}$	Reject $H_0$ if $t < -t_{\alpha, \nu}$	Reject $H_0$ if $t > t_{\alpha, \nu}$
Rejection region and critical value			

# Other hypothesis tests

# Hypothesis test for proportion

- Sometimes we want to test for proportion. Eg:
  - ▣ Does the proportion of customers that prefer the new product higher than the old product?
  - ▣ Is the voter turnout higher than last year?
- Examples of hypotheses:
  - ▣  $H_0: p = 0.5$  vs  $H_1: p > 0.5$
  - ▣  $H_0: p_1 = p_2$  vs  $H_1: p_1 \neq p_2$
- One population test statistic:

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

which follows approximately the standard normal distribution for large  $n$ .

# Hypothesis test for proportion

- Two population proportions:

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}$$

which follows approximately the standard normal distribution for large  $n_1$  and  $n_2$ .

# Hypothesis test for variance

- We can also perform hypothesis test on variance, for normally distributed populations.
- One population variance:
  - ▣  $H_0: \sigma^2 = \sigma_0^2$  vs  $H_1: \sigma^2 \neq \sigma_0^2$  (two-tailed test)
  - ▣  $H_0: \sigma^2 = \sigma_0^2$  vs  $H_1: \sigma^2 < \sigma_0^2$  (left-tailed test)
  - ▣  $H_0: \sigma^2 = \sigma_0^2$  vs  $H_1: \sigma^2 > \sigma_0^2$  (right-tailed test)for a given value  $\sigma_0^2$
- Test statistic (one population):

$$\chi^2 = \frac{(n - 1)S^2}{\sigma_0^2}$$

follows the chi-square distribution with  $n - 1$  degrees of freedom

# Hypothesis test for variance

- Two population variances:
  - ▣  $H_0: \sigma_1^2 = \sigma_2^2$  vs  $H_1: \sigma_1^2 \neq \sigma_2^2$
  - ▣  $H_0: \sigma_1^2 = \sigma_2^2$  vs  $H_1: \sigma_1^2 < \sigma_2^2$
  - ▣  $H_0: \sigma_1^2 = \sigma_2^2$  vs  $H_1: \sigma_1^2 > \sigma_2^2$
  
- Test statistic (two populations):

$$F = \frac{S_1^2}{S_2^2}$$

follows the  $F$ -distribution with  $n_1 - 1$  and  $n_2 - 1$  degrees of freedom

# Test of independence

Group	Prefer new procedure	Prefer old procedure	No preference
Nurses	100	80	20
Doctors	50	120	30

- When given a contingency table, we may want to test whether the variables are independent or not.
- In the table above, suppose a new procedure is administered to patients in a large hospital.
- The researcher then ask the question, Do the doctors feel differently about this procedure from the nurses, or do they feel basically the same way?
  - ▣ Are the opinion about the procedure independent of the profession?



# Test of independence

- Hypotheses:
  - ▣  $H_0$ : The two variables are independent
  - ▣  $H_1$ : The two variables are dependent

- Test statistic:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

follows the chi-square distribution with degrees of freedom equal to (number of rows minus 1) × (number of columns minus 1) where

- ▣  $O$  is the observed frequency
- ▣  $E$  is the expected frequency.  $E = \frac{(\text{row total})(\text{column total})}{\text{grand total}}$

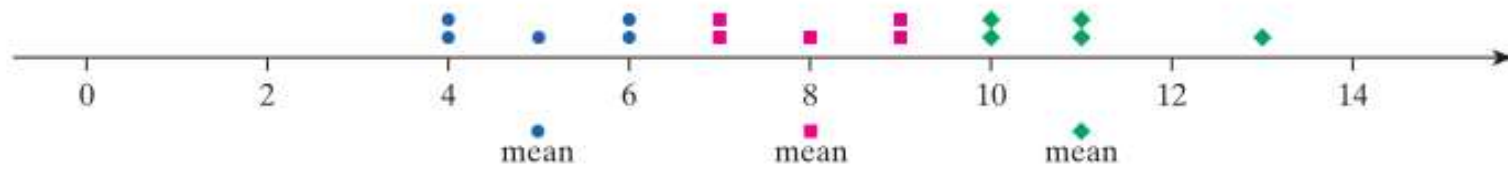
# Analysis of variance (ANOVA)

- The Analysis of variance (ANOVA) is a procedure used to test the equality of means of three or more populations.
- Hypotheses:
  - ▣  $H_0$ : All population means are equal ( $\mu_1 = \mu_2 = \dots = \mu_k$ )
  - ▣  $H_1$ : At least two population means are different
- It uses the test statistic

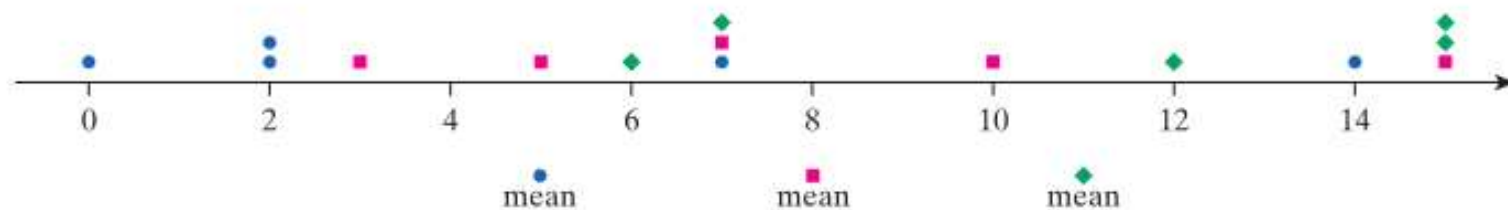
$$F = \frac{\text{Between group variance}}{\text{Within group variance}}$$

which follows the  $F$ -distribution with  $k - 1$  and  $n - 1$  degrees of freedom

# Analysis of variance (ANOVA)



(a)



(b)

# Summary

- Introduction to hypothesis tests
  - ▣ Null and alternative hypotheses
  - ▣ Type I and Type II errors
  - ▣ Critical value, rejection region, test statistic,  $p$ -value
  - ▣ Making conclusion
  
- Hypothesis test for population mean (one population)
  - ▣ Variance known
  - ▣ Variance unknown and sample size is large
  - ▣ Variance unknown and sample size is small

# Summary

- Hypothesis test for population means (two populations)
  - ▣ Paired/dependent samples
  - ▣ Independent samples
    - Variances known
    - Variances unknown and sample size is large
    - Variances unknown, sample size is small and  $\sigma_1 = \sigma_2$
    - Variances unknown, sample size is small and  $\sigma_1 \neq \sigma_2$
- Other hypothesis tests
  - ▣ Tests for proportion
  - ▣ Tests for variance
  - ▣ Test of independence
  - ▣ ANOVA