7. HYPOTHESIS TESTING

Introduction, hypothesis test on mean, independence, ANOVA

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.

- After giving some introduction to hypothesis testing, we will give some examples of it.

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

UNIVERSITI KEBANGSAAN MALAYSIA
The National University of Malaysia

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:

T1 = FN

T2 = FP

		Reality	
		H_0 is true	H_0 is false
Decision	Do not reject H_0	TP Correct decision	Type II error FP
	Reject H_0	Type I error FN	Correct decision TN

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 α and β, respectively:

$$\alpha = P(\text{Type I error}) = P(\text{Reject } H_0 | H_0 \text{ is true})$$

$$\beta = P(\text{Type II error}) = P(\text{Do not reject } H_0 | H_0 \text{ is false})$$

→ aras keertian

- α is also called the significance level of the test and is chosen before hypothesis test is performed. Normally, α is chosen to be 1%, 5%, or 10%.

$\alpha \approx \frac{1}{\beta}$, $\alpha\uparrow$, $\beta\downarrow$

Type I and Type II errors

- We want both α and β 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 α decreases, β will increase
  - If β decreases, α will increase

- We can reduce both α and β 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)

FN

Reject H_0 | H_0 true

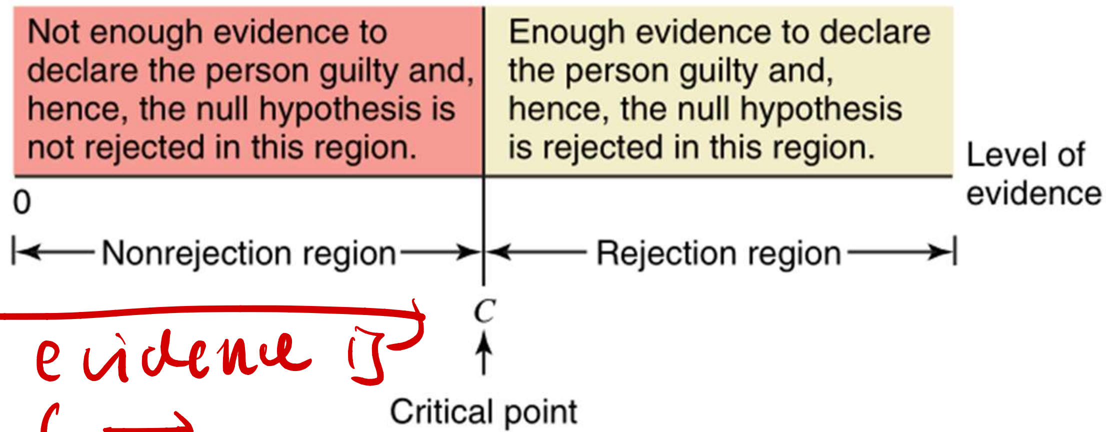
FP

Do not reject H_0 | H_0 is false

Critical value, rejection regions, and test statistic

Rejection and nonrejection region

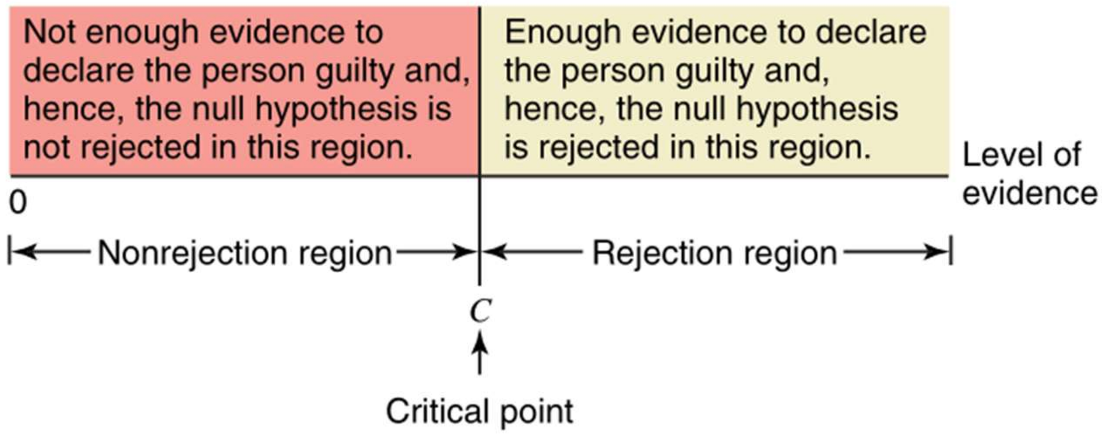
Ho = person is not guilty.



as more evidence is found, C →

- 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 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-value $< \alpha$, then H_0 is rejected. Otherwise, if p-value $> \alpha$, then H_0 is not rejected.

p-value interpretation

p-value > 0.10	None or very little evidence against H_0
$0.05 < p$-value < 0.10	Weak evidence against H_0
$0.01 < p$-value < 0.05	Strong evidence against H_0
p-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 μ be the population mean of soda content in the cans.
- The two hypotheses:
  - H_0: $\mu \geq 355$ ml (company claim).
  - H_1: $\mu < 355$ ml

$\bar{X} = 352$

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 μ 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) Let μ = average life time of light bulbs.

$H_0 : \mu = 8000 \ (\mu \geq 8000)$.

$H_1 : \mu < 8000$

b) Let μ = mean diameter of tennis balls

$H_0 : \mu = 3$ inches

$H_1 : \mu \neq 3$ inches

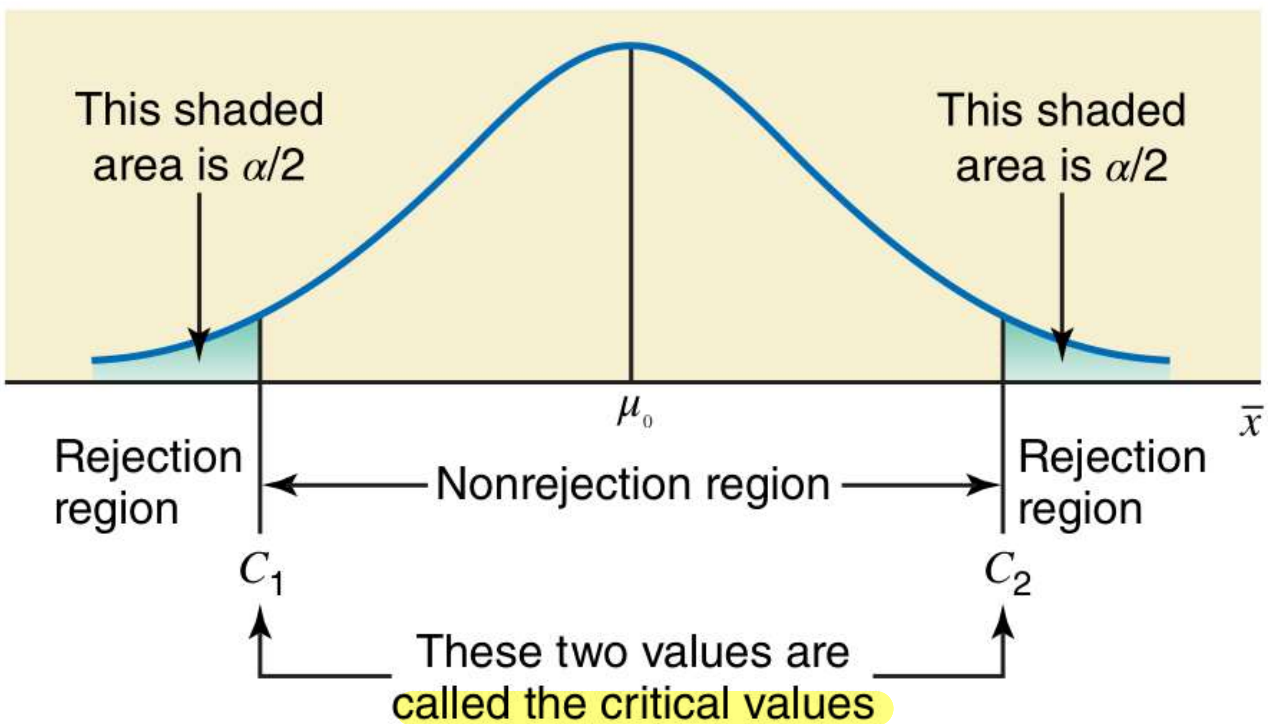
c) Let σ^2 = variance amount of fat in whole milk processed.

$H_0 : \sigma^2 = 0.25$ or $\sigma^2 \leq 0.25$.

$H_1 : \sigma^2 > 0.25$

Rejection regions

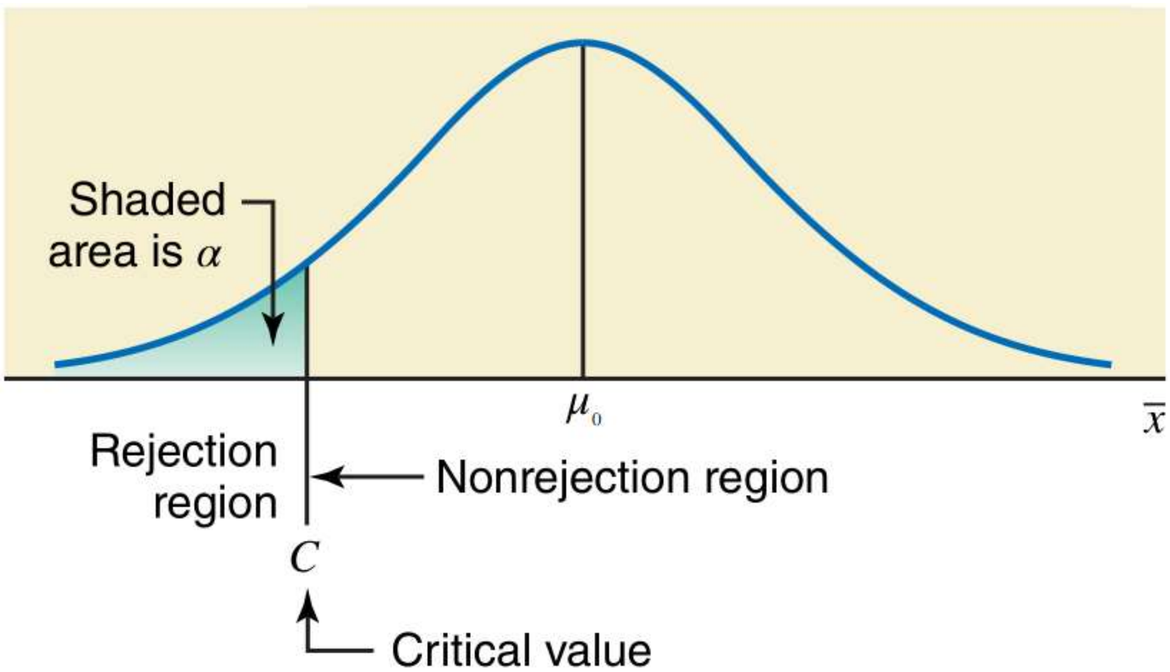
- Two-tailed test
  - $H_0: \mu = \mu_0$; $H_1: \mu \neq \mu_0$
  - Reject H_0 if $|\text{test statistic}| > \text{critical value}$
  - p-value is twice the area to the right of the positive test statistic, or twice the area to the left of the negative test statistic..



p-value = area × 2

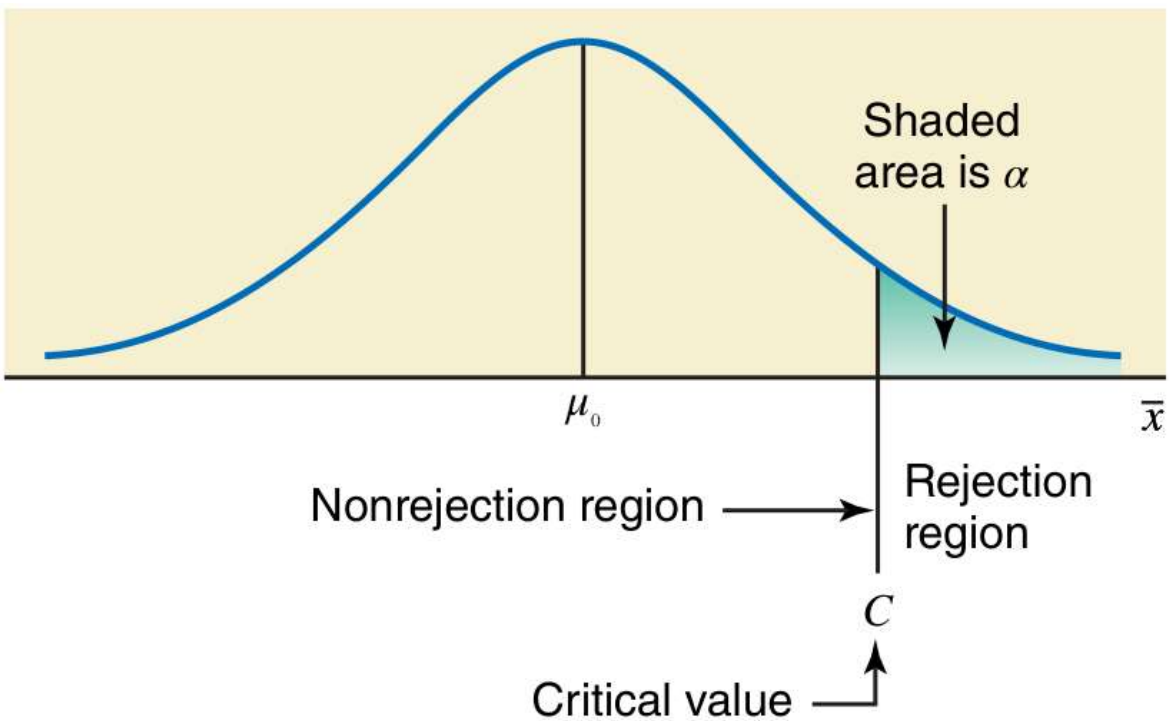
Rejection regions

- Left-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 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 α
  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 $\boldsymbol{H_0}$	Claim on $\boldsymbol{H_1}$
Reject H_0	There is enough evidence to reject H_0	There is enough evidence to support 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) Let μ be the average number of operations per year

$H_0 : \mu \geq 211 \quad (\mu = 211)$.

$H_1 : \mu < 211 \quad$ (director belief).

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

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

Hypothesis test for population mean (one population)

Elements of a hypothesis test

- These are the things that should exist in a hypothesis test:
  - The null and alternative hypotheses
  - The test statistics
  - The distribution of the test statistics
  - The *p*-value or the rejection region (based on the distribution)
  - Conclusion

- In the next sections, you should notice these elements and how they relate to hypothesis test

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 μ_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:
  - σ is known
  - σ is unknown but the sample size is large
  - σ is unknown but the sample size is small

- Each case has different test statistic and/or distribution to be used.

Case 1: σ is known

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

  $\mu = \mu_0$

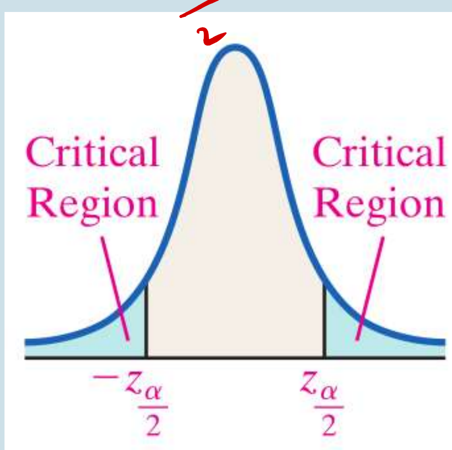
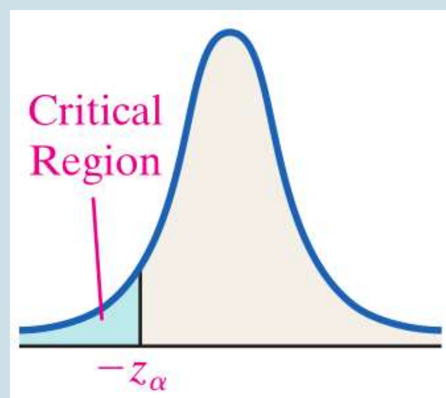
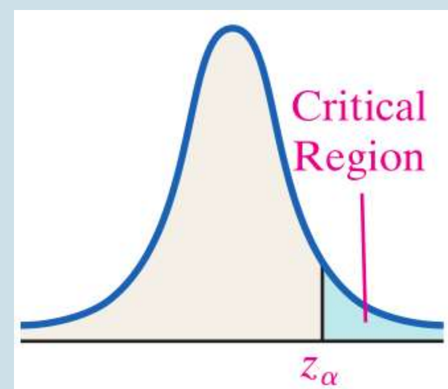
- If σ is known,

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

  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: σ 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	$z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$		
Decision rule	Reject H_0 if $\lvert z \rvert > z_{\alpha/2}$	Reject H_0 if $z < -z_\alpha$	Reject H_0 if $z > z_\alpha$
p-value	$2P(Z > \lvert z \rvert)$ $\frac{1}{2}$	$P(Z < z)$ α	$P(Z > z)$
Rejection region and critical value			

$P(Z > Z_\alpha) = \alpha$

Critical values

- For normal distribution, here are the critical values:

Significance level α	Left-tailed $(-z_\alpha)$	Right-tailed (z_α)	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

UNIVERSITI KEBANGSAAN MALAYSIA
The National University of Malaysia

- 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:

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

- Step 2: calculate the test statistic

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

$\therefore \quad z > 1.645$

$\bar{x} = 48.6$

$\mu = 46$

$n = 30$

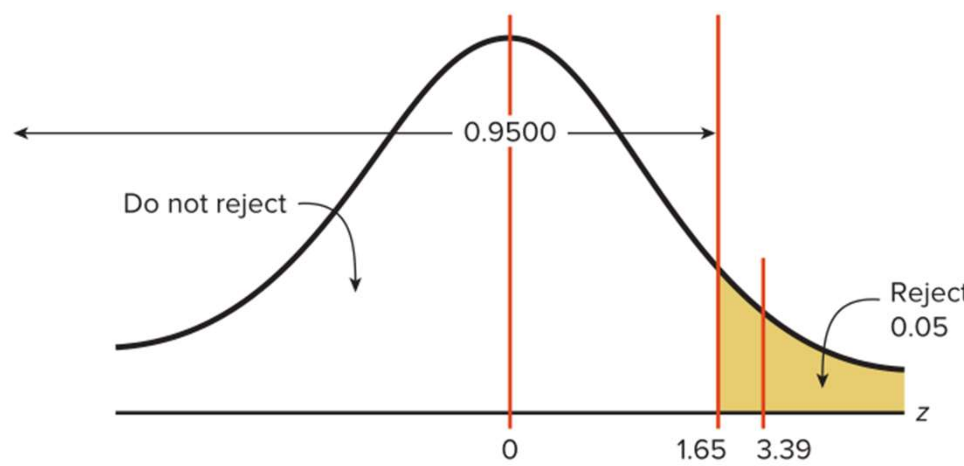
$\sigma = 4.2$

$\alpha = 0.05$

$z = 1.645$

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 $z_{0.05} = 1.645$

$H_0 : \mu = 5700$

$H_1 : \mu > 5700$

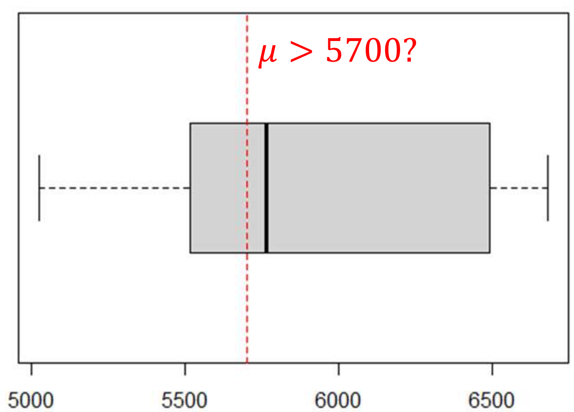
$n = 9$

$\bar{x} = 5950$

$\sigma = 659$

$\alpha = 0.05$

$z = 1.645$



$$z = \frac{\bar{x} - \mu}{\sigma - \sqrt{n}} = \frac{5950 - 5700}{659 / \sqrt{9}} = 1.138$$

$\therefore z < 1.645$

not enough evidence to support the claim that the average cost of tuition at 9 four-year public college is greater than 5700.

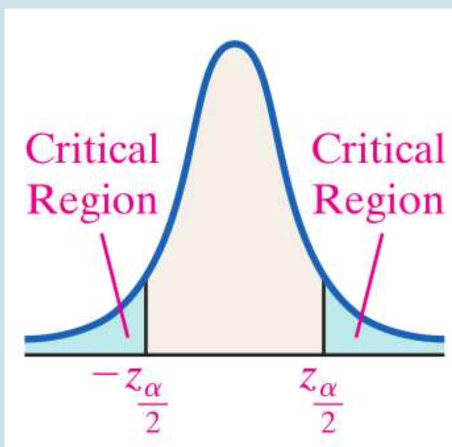
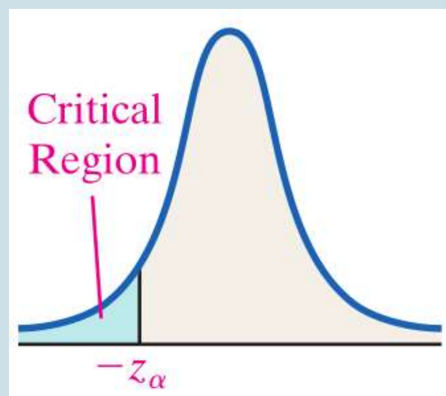
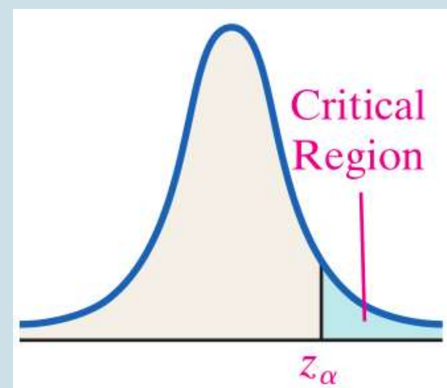
Case 2: σ is unknown but sample size is large

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

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

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

Case 2: σ 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	$z = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim N(0,1)$		
Decision rule	Reject H_0 if $\lvert z \rvert > z_{\alpha/2}$	Reject H_0 if $z < -z_\alpha$	Reject H_0 if $z > z_\alpha$
p-value	$2P(Z > \lvert z \rvert)$	$P(Z < z)$	$P(Z > z)$
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.

$\mu = 8$

$n = 32$

$\bar{x} = 8.2$

$s = 0.6$

- Step 1: Hypotheses

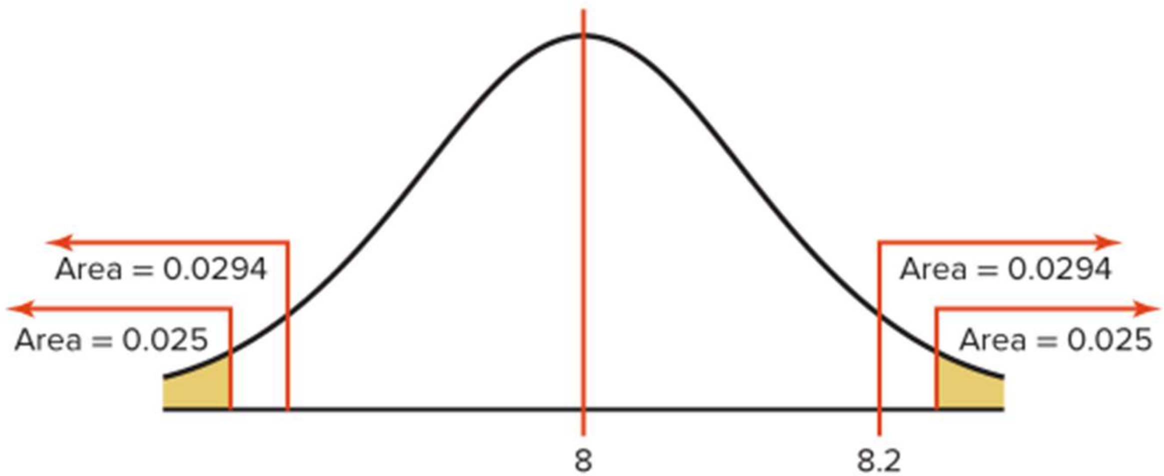
$$H_0: \mu = 8, \qquad H_1: \mu \neq 8$$

- Step 2: Test statistic

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

Example

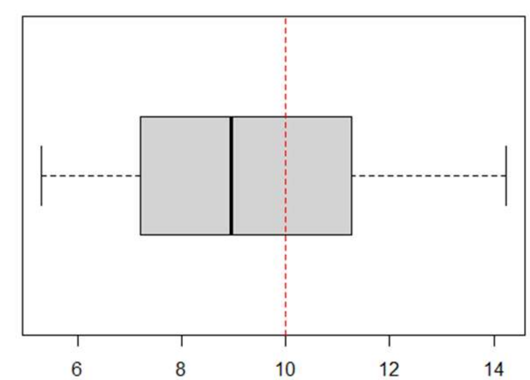
- Step 3: Conclusion
  - p-value $= 2 \times P(Z > |1.886|) = 2(0.0294) = 0.0588.$
  - p-value $> \alpha = 0.05.$
  - Therefore, we do not have enough evidence to reject H_0.
  - There is not enough evidence to reject the claim that the average wind speed is 8 miles per hour.



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.

$\mu < 10?$



$\mu = 10$

$\bar{x} = 9.2$

$n = 36$

$s = 2.4$

$\alpha = 0.01$

$H_0 : \mu \geq 10$

$H_1 : \mu < 10$

$$z = \frac{\bar{x} - \mu}{s/\sqrt{n}} = -2$$

p value = $P(z < 2) = P(z < -2) = P(z > 2)$

p value = 0.02275

p value > α, not enough evidence to reject H_0

Case 3: σ is unknown and sample size is small

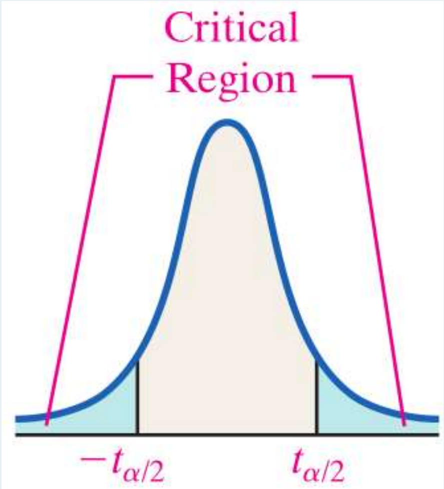
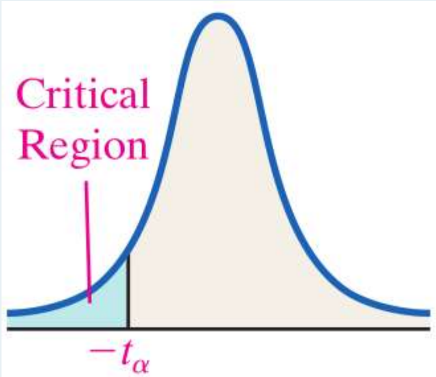
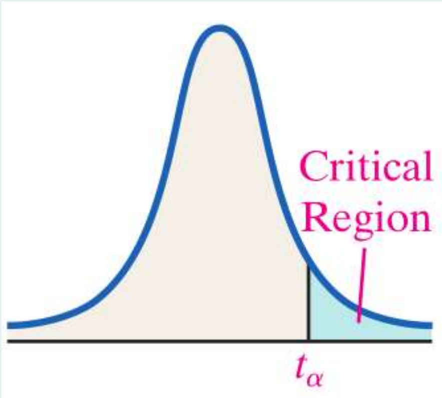
- If σ is not known and $n < 30$, we can use

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

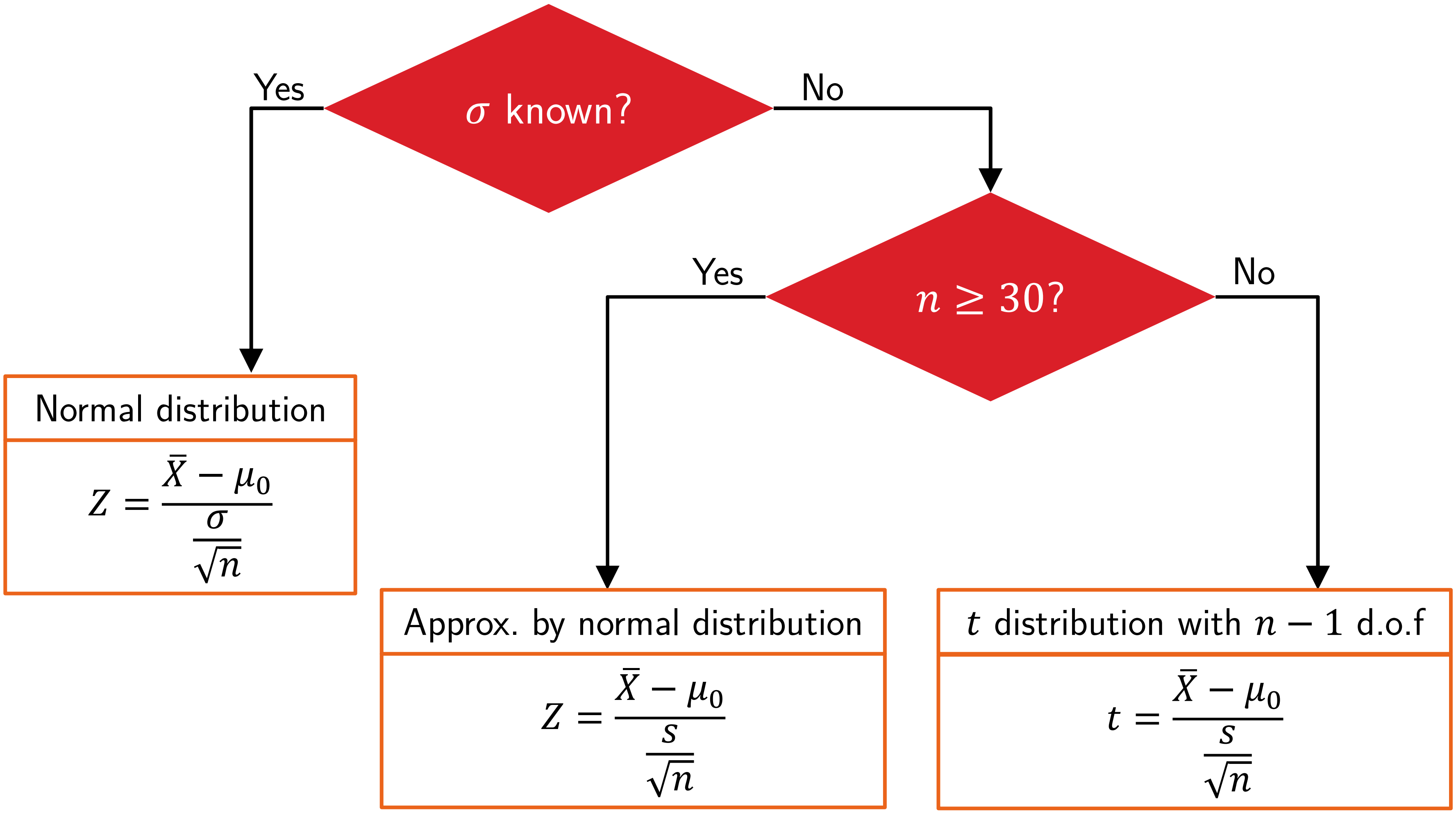
as the test statistic. This value follows the t-distribution with $n - 1$ degrees of freedom

- This test is called the t-test.

Case 3: σ 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	$t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim t_{n-1}$		
Decision rule	Reject H_0 if $\lvert t \rvert > t_{\alpha/2, n-1}$	Reject H_0 if $t < -t_{\alpha, n-1}$	Reject H_0 if $t > t_{\alpha, n-1}$
Rejection region and critical value			

Test statistic (test on mean)



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, μ_1 and μ_2 by taking the difference $\mu_1 - \mu_2$.

- Two cases:
  - Paired/dependent samples
  - Independent samples

Dependent/Paired 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:
  - σ_1 and σ_2 are known
  - σ_1 and σ_2 are not known but the sample size is large
  - σ_1 and σ_2 are not known, sample size is small and $\sigma_1 = \sigma_2$
  - σ_1 and σ_2 are not known, sample size is small and $\sigma_1 \neq \sigma_2$

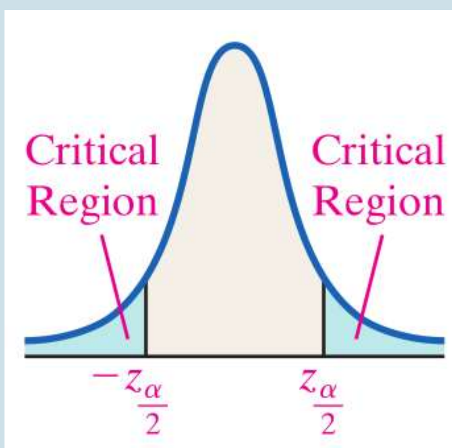
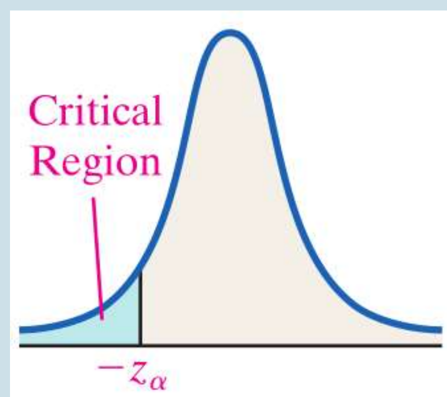
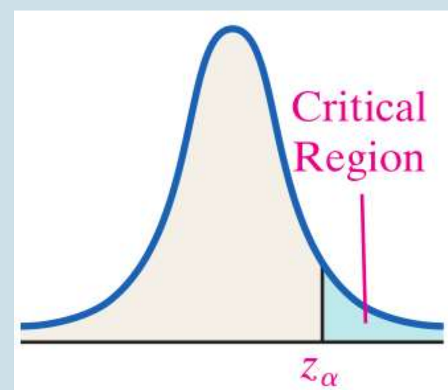
The hypotheses

- 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 δ_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

$\boldsymbol{H_0}$ and $\boldsymbol{H_1}$	$H_0: \mu_1 - \mu_2 = \delta_0$ $H_1: \mu_1 - \mu_2 \neq \delta_0$	$H_0: \mu_1 - \mu_2 = \delta_0$ $H_1: \mu_1 - \mu_2 < \delta_0$	$H_0: \mu_1 - \mu_2 = \delta_0$ $H_1: \mu_1 - \mu_2 > \delta_0$
Test statistic	$z = \dfrac{(\bar{x}_1 - \bar{x}_2) - \delta_0}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim N(0,1)$		
Decision rule	Reject H_0 if $\lvert z \rvert > z_{\alpha/2}$	Reject H_0 if $z < -z_\alpha$	Reject H_0 if $z > z_\alpha$
p-value	$2P(Z > \lvert z \rvert)$	$P(Z < z)$	$P(Z > z)$
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 $z_{0.025} = 1.96$.

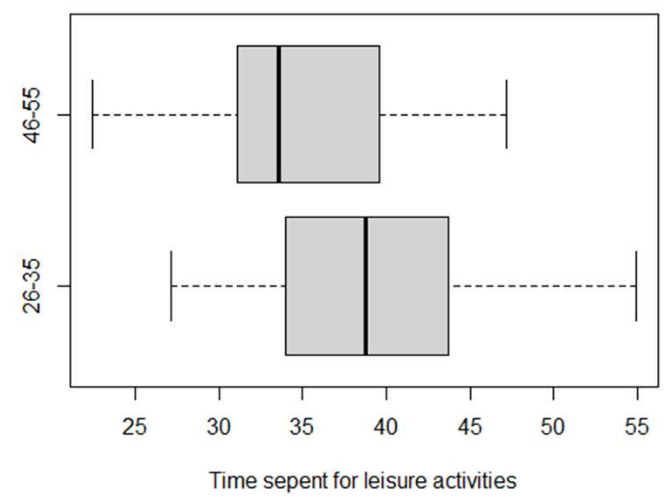
$n_1, n_2 = 35$

$\bar{x}_1 = 39.6, \sigma_1 = 6.3$

$\bar{x}_2 = 35.4, \sigma_2 = 5.8$

$\alpha = 0.05$

$\mu_1 \neq \mu_2$?



$H_0 : \mu_1 - \mu_2 = 0$

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

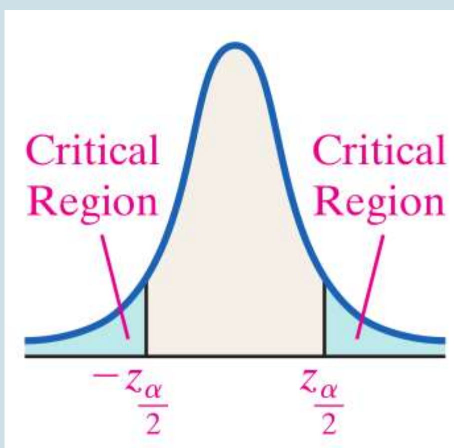
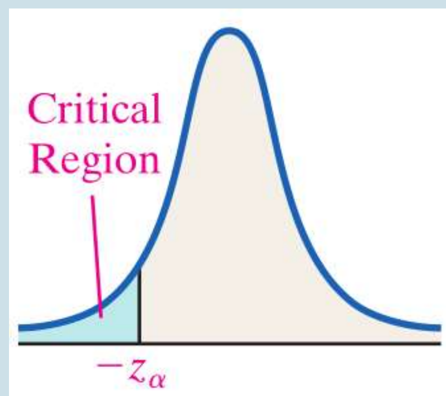
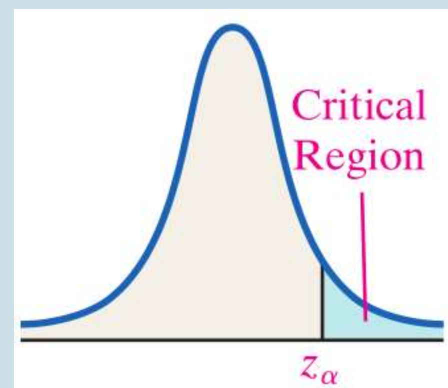
$$Z = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}, \quad \delta = 0$$

$$= \frac{39.6 - 35.4}{\sqrt{\frac{6.3^2}{35} + \frac{5.8^2}{35}}} = 2.902$$

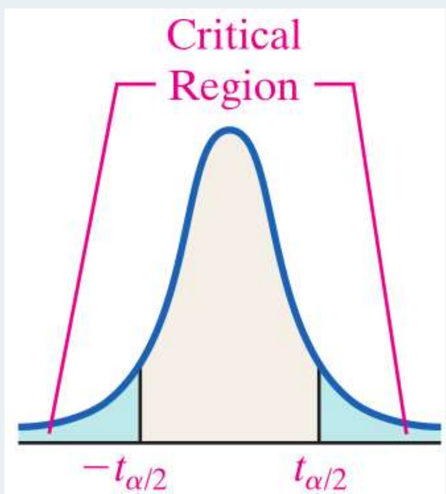
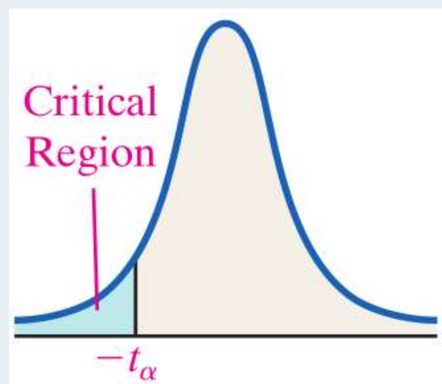
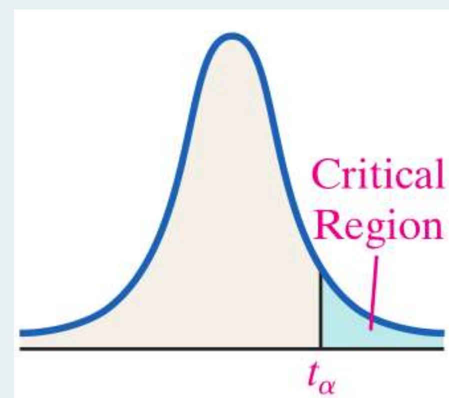
reject H_0 if $|z| > Z_{\alpha/2}$

there is enough evidence to reject the claim

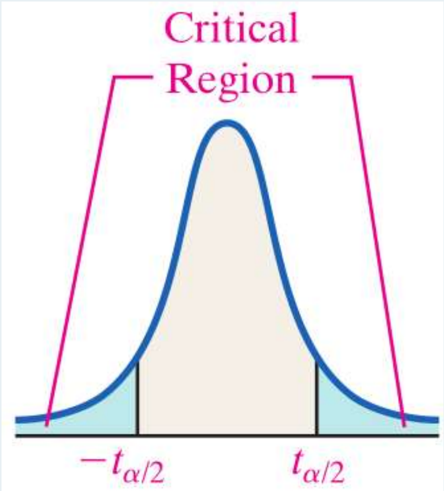
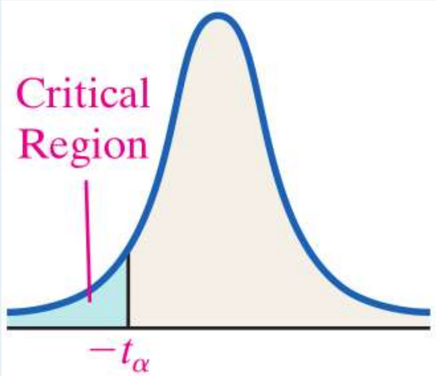
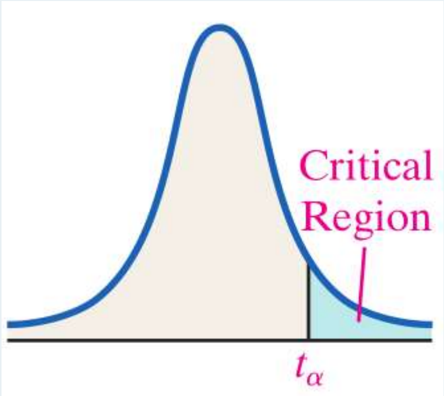
Case 2: variances unknown but size sample is large

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	$z = \dfrac{(\bar{x}_1 - \bar{x}_2) - \delta_0}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} \sim N(0,1)$		
Decision rule	Reject H_0 if $\lvert z \rvert > z_{\alpha/2}$	Reject H_0 if $z < -z_\alpha$	Reject H_0 if $z > z_\alpha$
p-value	$2P(Z > \lvert z \rvert)$	$P(Z < z)$	$P(Z > z)$
Rejection region and critical value			

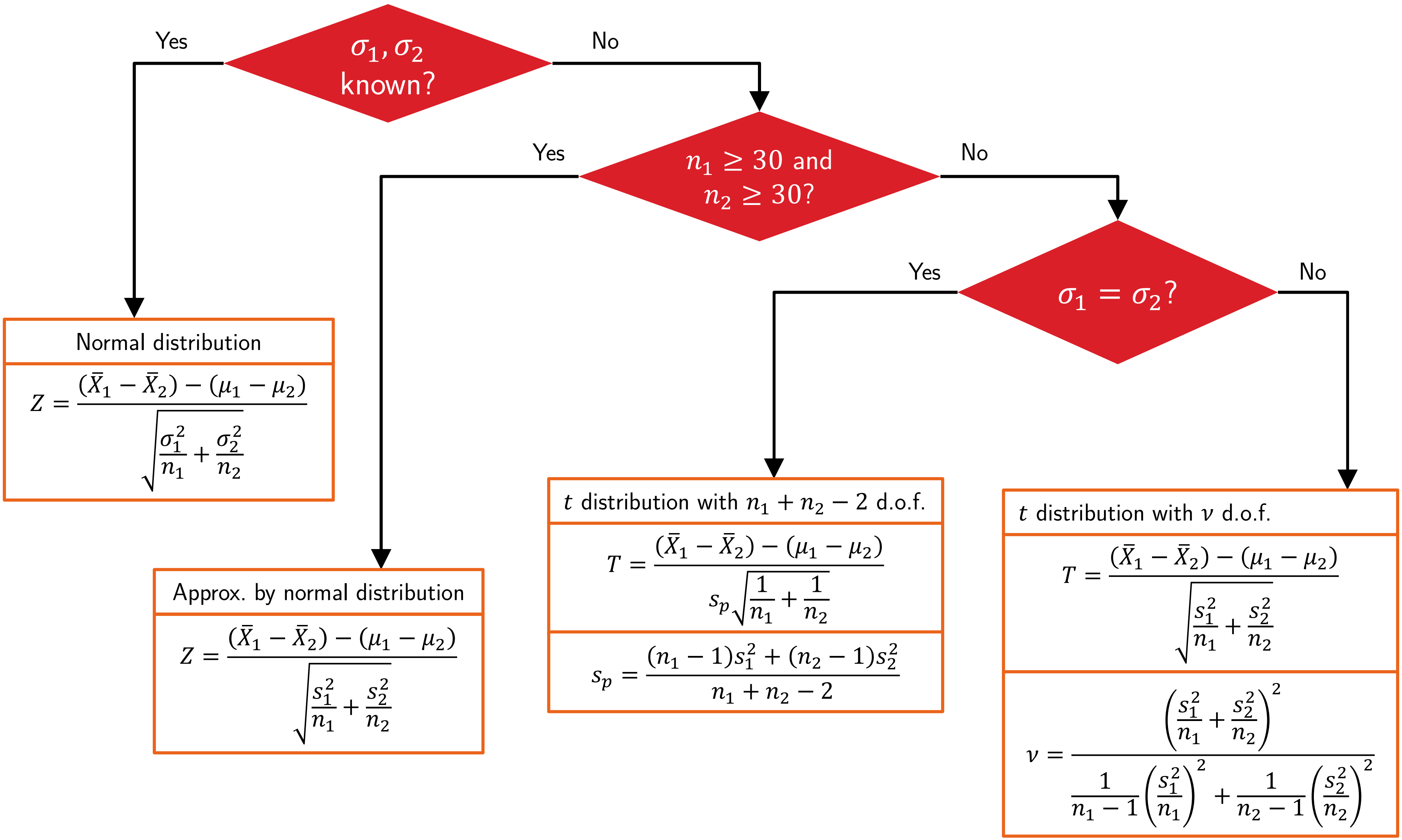
Case 3: variances unknown, sample size is small and $\sigma_1 = \sigma_2$

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	$t = \dfrac{(\bar{x}_1 - \bar{x}_2) - \delta_0}{\sqrt{s_p^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} \sim t_{n_1+n_2-2}$, where $s_p^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$ estimate variance both pop		
Decision rule	Reject H_0 if $\lvert t \rvert > t_{\alpha/2, n_1+n_2-2}$	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

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	$t = \dfrac{(\bar{x}_1 - \bar{x}_2) - \delta_0}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} \sim t_\nu$, where $\nu = \dfrac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\dfrac{1}{n_1 - 1}\left(\dfrac{s_1^2}{n_1}\right)^2 + \dfrac{1}{n_2 - 1}\left(\dfrac{s_2^2}{n_2}\right)^2}$		
Decision rule	Reject H_0 if $\lvert t \rvert > t_{\alpha/2,\nu}$	Reject H_0 if $t < -t_{\alpha,\nu}$	Reject H_0 if $t > t_{\alpha,\nu}$
Rejection region and critical value			

Independent samples



Other hypothesis tests

(Chi-Square)

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 $(R-1) \times (C-1)$ where
  - O is the observed frequency
  - E is the expected frequency. $E = \frac{\text{(row total)(column total)}}{\text{grand total}}$
  - R and C are the number of rows and columns, respectively

- Reject the null hypothesis if $\chi^2 > \chi^2_\alpha$ or the p-value $< \alpha$

Example

- Suppose a new procedure is administered to patients in a large hospital and a survey was conducted to study hospital staffs' opinion on the new procedure. Based on the data, do the doctors feel differently about this procedure from the nurses, or do they feel basically the same way? Conduct a hypothesis test using 5% significance level. It is given that the critical value is $\chi^2_{\alpha,2} = 5.991$.

- Observed frequency:

Group	Prefer new procedure	Prefer old procedure	No preference	Total
Nurses	100	80	20	200
Doctors	50	120	30	200
Total	150	200	50	400

Example

- Hypotheses:

H_0 : the opinion is not related to profession

H_1 : opinion is dependent to profession.

- Expected frequency: $-\ \dfrac{\text{row } T \times \text{col } T}{G\ T}$ $\left(\dfrac{200 \times 150}{400}\right)$

Group	Prefer new procedure	Prefer old procedure	No preference	Total
Nurses	75	100	25	200
Doctors	75	100	25	200
Total	150	200	50	400

Example

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

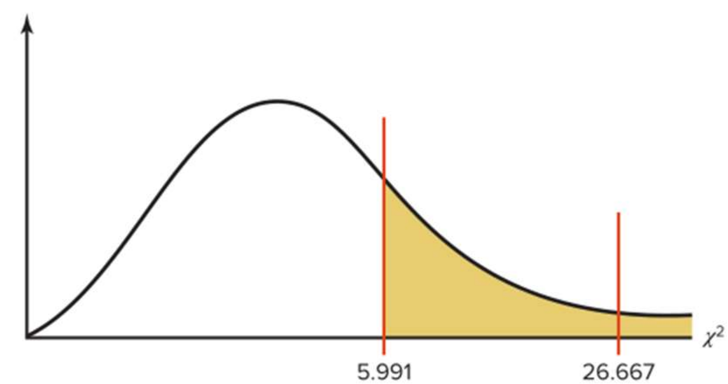
$$\frac{(100-75)^2}{75} + \frac{(80-100)^2}{100} + \ldots = 26.667$$

- Decision: $df = (R-1) * (C-1)$

$$\chi^2_{\alpha,2} = 5.991$$

$$\chi^2 = 26.667$$

$\chi^2 > 5.991$, H_0 is rejected



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 $(\mu_1 = \mu_2 = \cdots = \mu_k)$
  - H_1: At least two population means are different

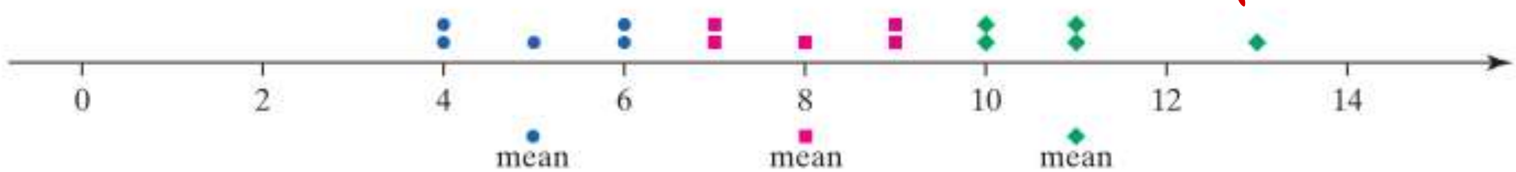
- It uses the test statistic

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

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

Analysis of variance (ANOVA)

within group variance



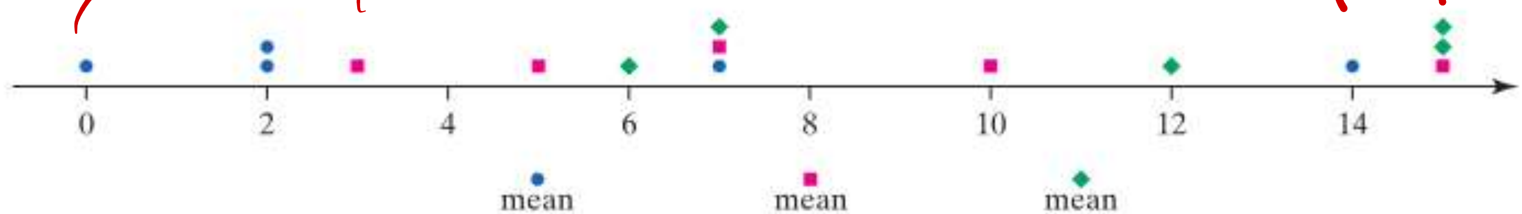
between group variance

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

Analysis of variance (ANOVA)

UNIVERSITI KEBANGSAAN MALAYSIA
The National University of Malaysia

within group



(variance of sample mean) b/w group

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

we expect F to be small and H_0 to not be rejected

Analysis of variance (ANOVA)

UNIVERSITI KEBANGSAAN MALAYSIA
The National University of Malaysia

- ANOVA table:

Table 12.3 ANOVA Table

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	Value of the Test Statistic
Between	$k-1$	SSB	MSB	$F = \frac{\text{MSB}}{\text{MSW}}$
Within	$n-k$	SSW	MSW	
Total	$n-1$	SST		

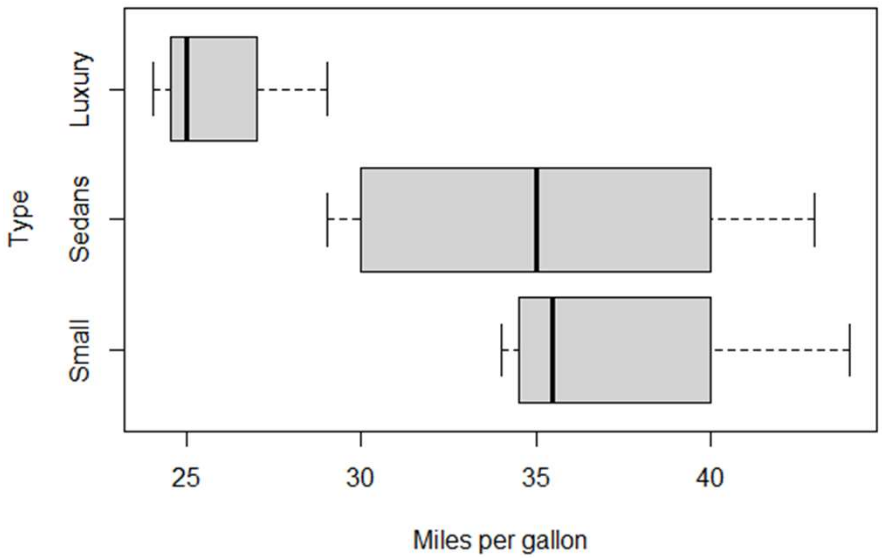
- Reject null hypothesis if $F > F_{\alpha,k-1,n-k}$ or p-value $< \alpha$

Example

UNIVERSITI KEBANGSAAN MALAYSIA
The National University of Malaysia

- A researcher wishes to see if there is a difference in the fuel economy for city driving for three different types of automobiles: small automobiles, sedans, and luxury automobiles. He randomly samples four small automobiles, five sedans, and three luxury automobiles. The miles per gallon for each is shown. At $\alpha = 0.05$, test the claim that there is no difference among the means.

Small	Sedans	Luxury
36	43	29
44	35	25
34	30	24
35	29	
	40	



Example

- Hypotheses:

H_0 : The mean miles per gallon is the same for all cars.

H_1 : atleast 2 population means are different [atleast 1 mean is different compared to others]

- Test statistic:

Source	Sum of squares	d.f.	Mean square	F
Between	242.717	2	121.359	4.83
Within (error)	225.954	9	25.106	
Total	468.671	11		

Example

- Decision:

F > 4.26,

null hypothesis is rejected

can conclude at least

one mean is different from the others.

Summary

- Introduction to hypothesis tests

- Elements of hypothesis test:
  - Null and alternative hypotheses
  - Test statistics
  - The distribution of the test statistics
  - The p-value or the rejection region (based on the distribution)
  - Conclusion

Summary

- Examples of hypothesis tests:
  - Test for population mean (one population)
  - Test for population means (two populations)
  - Test of independence
  - ANOVA