UTM
UNIVERSITI TEKNOLOGI MALAYSIA

CHAPTER 5
Part 3

HYPOTHESIS TESTING
(Two Samples Test)

Example of Questions on Two Samples Test:

- Does taking a small dose of aspirin every day reduce the risk of heart attack?
    - *Compare a group that takes aspirin with one that doesn't.*
- Do men and women in the same occupation have different salaries.
    - *Compare a sample of men with a sample of women.*
- Does fuel A lead to better gas mileage than fuel B?
    - *Compare the gas mileage for a fleet of cars when they use fuel A, to when those same cars use fuel B.*

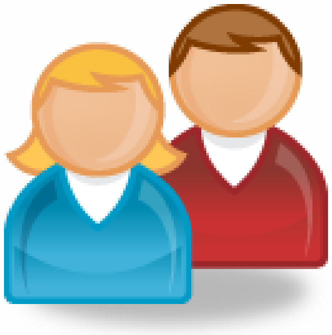
Example:

Suppose we have a population of adult men with a mean height of 71 inches and standard deviation of 2.5 inches. We also have a population of adult women with a mean height of 65 inches and standard deviation of 2.3 inches. In this case the researcher is interested in comparing the mean height between men and women. His research question is: **Does the mean height of men differ from the mean height of women?**. Here, the hypotheses are

$$H_0\colon \mu_1 = \mu_2$$
$$H_1\colon \mu_1 \neq \mu_2$$

where,

μ_1= mean height of men
μ_2= mean height of women



Inferences from Two Samples

- Two Proportions
- Two Means: Independent Samples
- Two Variances or Standard Deviations
- Two Dependent Samples (Matched Pairs)

Two Proportions

Notation Symbols:

For population 1:

p_1 = population proportion

n_1 = size of the sample

x_1 = number of successes in the sample

$$\hat{p}_1 = \frac{x_1}{n_1}$$ (the sample proportion)

$$\hat{q}_1 = 1 - \hat{p}_1$$

The corresponding notations apply to

$p_2, n_2, x_2, \hat{p}_2 \; and \; \hat{q}_2$, which come from population 2.

The pooled sample proportion is given by:

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2};\ \bar{q} = 1 - \bar{p}$$

Requirements

1. We have proportions from two independent simple random samples.
2. For each of the two samples, the number of successes is at least 5 and the number of failures is at least 5.

Test Statistic for Two Proportions

For $H_0 : p_1 = p_2$

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

where (assumed in the null hypothesis) $p_1 - p_2 = 0$

Decision Criteria for Two Proportions

- *P*-value: *P*-values are automatically provided by statistical software (e.g., SPSS). If statistical software is not available, use table.

- Critical values: Use standard normal distribution table.

Example #1

Do people having different spending habits depending on the type of money they have?

89 undergraduates were randomly assigned to two groups and were given a choice of keeping the money or buying gum or mints.

The claim is that “money in large denominations is less likely to be spent relative to an equivalent amount in many smaller denominations”.

Let’s test the claim at the 0.05 significance level.

Example #1 - Solution

Below are the sample data and summary statistics:

	Group 1	Group 2
	Subjects Given \$1	Subjects Given 4 quarters (1 quarter = 25 cent)
Spent the money	x_1=12	x_2=27
Subjects in group	n_1=46	n_2=43

$$\hat{p}_1 = \frac{12}{46} = 0.261; \hat{p}_2 = \frac{27}{43} 0.628 \Rightarrow \bar{p} = \frac{12 + 27}{46 + 43} = 0.438$$

Example #1 – Solution (cont'd)

Requirement Check:

1. The 89 subjects were randomly assigned to two groups, so we consider these independent random samples.

2. The subjects given the $1 bill include 12 who spent it and 34 who did not. The subjects given the quarters include 27 who spent it and 16 who did not. All counts are above 5, so the requirements are all met.

Example #1 – Solution (cont'd)

Step 1:

The claim that "money in large denominations is less likely to be spent" can be expressed as $p_1 < p_2$.

Step 2: If $p_1 < p_2$ is false, then $p_1 \geq p_2$.

Step 3: The hypotheses can be written as:

$$H_0: p_1 = p_2$$
$$H_1: p_1 < p_2$$

Example #1 – Solution (cont'd)

Step 4: The significance level is $\alpha = 0.05$.

Step 5: We will use the normal distribution to run the test with:

$$\hat{p}_1 = \frac{12}{46}; \hat{p}_2 = \frac{27}{43}$$

$$\bar{p} = \frac{12+27}{46+43} = 0.438\ ;\ \bar{q} = 1 - \bar{p} = 0.562$$

Example #1 – Solution (cont'd)

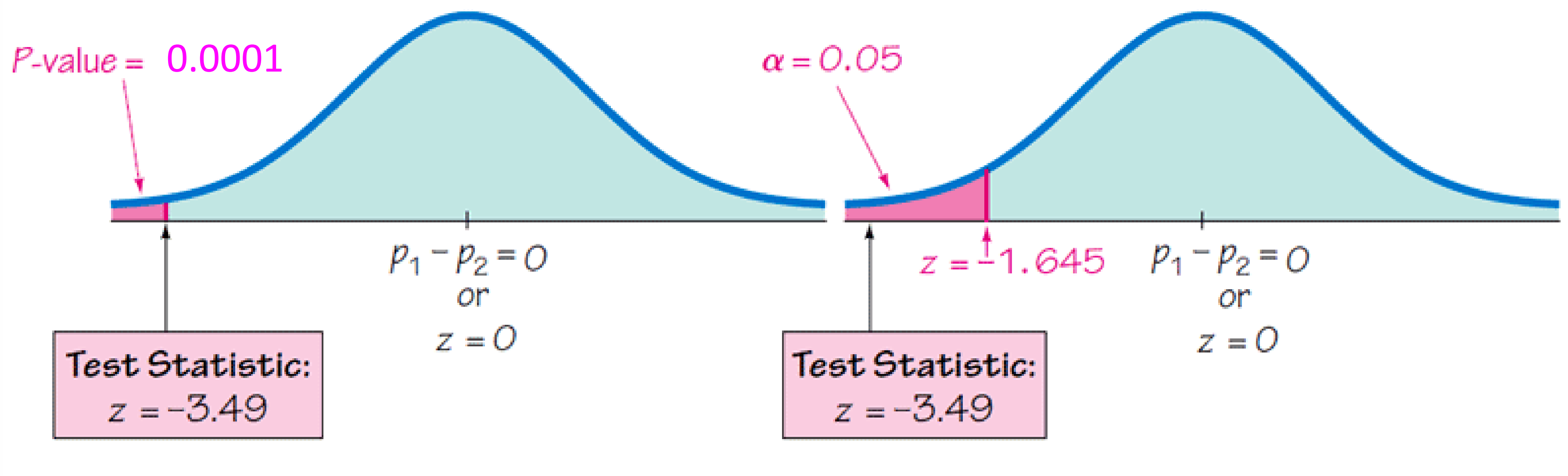
Step 6: Calculate the value of the test statistic:

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\bar{p}\bar{q}(\frac{1}{n_1} + \frac{1}{n_2})}} = \frac{(0.261 - 0.628) - 0}{\sqrt{0.438(0.562)(\frac{1}{46} + \frac{1}{43})}}$$

$$= \frac{-0.367}{0.1052} = -3.49$$

Example #1 – Solution (cont'd)

Step 7: This is a left-tailed test, so the *P*-value is the area to the left of the test statistic $z = -3.49$, or 0.0001. The critical value is also shown below.



(a) *P*-Value Method

(b) Critical Value Method

Example #1 – Solution (cont'd)

Step 8: Because the *P*-value of 0.0001 is less than the significance level of $\alpha = 0.05$, thus we reject the null hypothesis.

There is sufficient evidence to support the claim that people with money in large denominations are less likely to spend relative to people with money in smaller denominations.

It should be noted that the subjects were all undergraduates and care should be taken before generalizing the results to the general population.

Example #2

Time magazine reported the result of a telephone poll of 800 adult Americans. The question posed of the Americans who were surveyed was: "Should the federal tax on cigarettes be raised to pay for health care reform?" The results of the survey were:

Non-smoker	Smoker
$n_1 = 605$	$n_2 = 195$
$y_1 = 351$ said yes	$y_2 = 41$ said yes

Is there sufficient evidence at the α = 0.05 level, to conclude that the two populations (smokers and non-smokers) is differ significantly with respect to their opinions?

Example # 2 - Solutions

- The hypothesis statement:

$$H_0: p_1 = p_2$$
$$H_1: p_1 \neq p_2$$

- The significance level: $\alpha = 0.05$
- Calculate the test statistics:

$$\hat{p}_1 = \frac{351}{605} = 0.58;\ \hat{p}_2 = \frac{41}{195} = 0.21$$

$$\bar{p} = \frac{351 + 41}{605 + 195} = 0.49\ ;\ \bar{q} = 1 - \bar{p} = 0.51$$

Example #2 – Solution (cont'd)

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\bar{p}\bar{q}(\frac{1}{n_1} + \frac{1}{n_2})}} = \frac{(0.58 - 0.21) - 0}{\sqrt{(0.49)(0.51)(\frac{1}{605} + \frac{1}{195})}} = \frac{0.37}{0.0412} = 8.98$$

- This is a two-tailed test, so we reject H_0 if $z \leq -z_{0.025} = -1.96$ or $z \geq z_{0.025} = 1.96$

- Since $z = 8.98 \geq z_{0.025} = 1.96$, we reject H_0. There is sufficient evidence at the 0.05 level to conclude that the two populations differ with respect to their opinions concerning imposing a federal tax to help pay for health care reform.

Inferences About Two Means: Independent Samples

Definition

Two samples are independent if the sample values selected from one population are not related to or somehow paired or matched with the sample values selected from the other population.

Two samples are dependent (or consist of matched pairs) if the members of one sample can be used to determine the members of the other sample.

Example

Independent Samples:

Researchers investigated the reliability of test assessment. On group consisted of 30 students who took proctored tests. A second group consisted of 32 students who took online tests without a proctor.

The two samples are independent, because the subjects were not matched or paired in any way.

Example

Dependent Samples:

Students of the author collected data consisting of the heights (cm) of husbands and the heights (cm) of their wives. Five of those pairs are listed below.

Height of Husband	175	180	173	176	178
Height of Wife	160	165	163	162	166

The data are dependent, because each height of each husband is matched with the height of his wife.

UTM
UNIVERSITI TEKNOLOGI MALAYSIA

Statistical Inferences For Two Sample:

Test On Difference Between Two MEAN

Test on Difference Between Two Mean

Assumptions:-

1. The two population standard deviations are both known.

2. The two samples are independent.
   - $X_{11}, X_{12}, \ , X_{1n1}$ is a random sample from population 1.
   - $X_{21}, X_{22}, \ , X_{2n2}$ is a random sample from population 2

3. Both samples are simple random samples.

4. Either or both of these conditions are satisfied: The two sample sizes are both large (with $n_1 > 30$ and $n_2 > 30$) or both samples come from populations having normal distributions.

The test statistic will be a ***z*** and use the standard normal model.

Test on Mean with σ_1 and σ_2 (variances) are known

- Consider hypothesis testing on the difference in the means $\mu_1 - \mu_2$ of two normal populations.
- The procedure of this testing can be summarized as follow:

Null hypothesis: $H_0\colon \mu_1 - \mu_2 = \Delta_0$

Test statistic: $$Z_0 = \frac{\overline{X}_1 - \overline{X}_2 - \Delta_0}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$$

Alternative Hypotheses	Rejection Criterion
$H_1\colon \mu_1 - \mu_2 \neq \Delta_0$	$z_0 > z_{\alpha/2}$ or $z_0 < -z_{\alpha/2}$
$H_1\colon \mu_1 - \mu_2 > \Delta_0$	$z_0 > z_\alpha$
$H_1\colon \mu_1 - \mu_2 < \Delta_0$	$z_0 < -z_\alpha$

- In the formula, $\bar{X}_1 - \bar{X}_2$ is the observed difference, and expected difference, $\mu_1 - \mu_2$ is 0 when the null hypothesis is $\mu_1 = \mu_2$

- In the comparison of two sample means, the difference maybe due to chance, in which case the null hypothesis will not be rejected, and the researcher can assume that the means of the population are basically the same

Example

A product developer is interested in reducing the drying time of a primer paint. Two formulations of the paint are tested; formulation 1 is the standard chemistry, and formulation 2 has a new drying ingredient that should reduce the drying time. From experience, it is known that the standard deviation of drying time is 8 minutes, and this inherent variability should be unaffected by the addition of the new ingredient. Ten specimens are painted with formulation 1, and another ten specimens are painted with formulation 2; the 20 specimens are painted in random order and normally distributed. The two sample average drying times are $\bar{x}_1$ = 121 minutes and $\bar{x}_2$ = 112 minutes, respectively. What conclusions can the product developer draw about the effectiveness of the new ingredient, using $\alpha = 0.05$?

Example - Solution

1. The quantity of interest is the difference in mean drying times, $\mu_1 - \mu_2$ and $\Delta_0 = 0$.
2. H_0: $\mu_1 - \mu_2 = 0$.
   H_1: $\mu_1 > \mu_2$. We want to reject H_0 if the new ingredient reduces mean drying time.
3. Given, $\alpha = 0.05$.
4. The test statistic is $z_0 = \dfrac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{\dfrac{\sigma_1^2}{n1} + \dfrac{\sigma_2^2}{n2}}}$

where $\sigma_1^2 = \sigma_2^2 = 8^2 = 64$ and $n_1 = n_2 = 10$.

Example – Solution (cont'd)

6. Reject H_0 if $z_0 > z_{0.05} = 1.645$.
7. Computations: Since $\overline{x_1}$ = 121 minutes and $\overline{x_2}$ = 112 minutes, we have

$$z_0 = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{\frac{\sigma_1^2}{n1} + \frac{\sigma_2^2}{n2}}} = \frac{121 - 112}{\sqrt{\frac{64}{10} + \frac{64}{10}}} = 2.52.$$

8. Conclusion: Since z_0 = 2.52 > 1.645, we reject H_0 at the 0.05 level and conclude that adding the new ingredient to the paint significantly reduces the drying time.

Test on Mean, Variance Unknown

- We now consider tests of hypotheses on the difference in means $\mu_1 - \mu_2$ of two normal distributions where the variances σ_1^2 and σ_2^2 are unknown.

- The variances of the two samples may be assumed to be equal or unequal.
  - $\sigma_1^2 = \sigma_2^2 = \sigma^2$ **(CASE 1 –assumed equal variance)**
  - $\sigma_1^2 \neq \sigma_2^2$ **(CASE 2 – assumed unequal variance)**

- A *t*-statistic will be used to test these hypotheses.

Case 1: $\sigma_1^2 = \sigma_2^2 = \sigma^2$

- We wish to test

$$H_0: \mu_1 - \mu_2 = \Delta_0$$
$$H_1: \mu_1 - \mu_2 \neq \Delta_0$$

- Let $X_{11}, X_{12}, \dots, X_{1n_1}$ be a random sample of n_1 observations from the first population.
- Let $X_{11}, X_{12}, \dots, X_{1n_2}$ be a random sample of n_2 observations from the second population.
- Let S_1^2 and S_2^2 be sample variances, respectively.

- The procedure of testing:

Null hypothesis: $H_0\colon \mu_1 - \mu_2 = \Delta_0$

Test statistic: $$T_0 = \frac{\overline{X}_1 - \overline{X}_2 - \Delta_0}{S_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$$

Alternative Hypothesis	Rejection Criterion
$H_1\colon \mu_1 - \mu_2 \neq \Delta_0$	$t_0 > t_{\alpha/2, n_1+n_2-2}$ or $t_0 < -t_{\alpha/2, n_1+n_2-2}$
$H_1\colon \mu_1 - \mu_2 > \Delta_0$	$t_0 > t_{\alpha, n_1+n_2-2}$
$H_1\colon \mu_1 - \mu_2 < \Delta_0$	$t_0 < -t_{\alpha, n_1+n_2-2}$

- A *pooled estimates of the variance* is a weighted average of the variance using the two sample variances and the degree of freedom of each variance as the weights.
- The *pooled estimator* of σ^2, denoted by S_P^2 is defined by:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

- The degree of freedom (d.f) = $(n_1+n_2) - 2$

Example

Two catalysts are being analyzed to determine how they affect the mean yield of a chemical process. Specifically, catalyst 1 is currently in use, but catalyst 2 is acceptable. Since catalyst 2 is cheaper, it should be adopted, providing it does not change the process yield. A test is run in the pilot plant and results in the data shown in Table 1 (next page). Is there any difference between the mean yields? Conduct the t-test with $\alpha = 0.05$, and assume equal variances.

Table 1

Observation Number	Catalyst 1	Catalyst 2
1	91.50	89.19
2	94.18	90.95
3	92.18	90.46
4	95.39	93.21
5	91.79	97.19
6	89.07	97.04
7	94.72	91.07
8	89.21	92.75
	$\bar{x}_1 = 92.255$	$\bar{x}_2 = 92.733$
	$s_1 = 2.39$	$s_2 = 2.98$

Example - Solution

1. The parameters of interest are μ_1 and μ_2, the mean process yield using catalysts 1 and 2, respectively, and we want to know if

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

$$H_1: \mu_1 \neq \mu_2$$

2. Given, α =0.05. The test statistic is

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2 - 0}{s_p\sqrt{\frac{1}{n1} + \frac{1}{n2}}}$$

3. Reject H_0 if $t_0 > t_{0.025,\,14} = 2.145$, or if $t_0 < -\,t_{0.025,\,14} = -2.145$.

4. Calculations: $\bar{x}_1 = 92.255$, $s_1 = 2.39$, $n_1 = 8$,

$\bar{x}_2 = 92.733$, $s_2 = 2.98$, $n_2 = 8$

$$s_p^2 = \frac{(n1-1)s_1^2 + (n2-1)s_2^2}{n1+n2-2} = \frac{(7)(2.39)^2 + (7)(2.398)^2}{8+8-2} = 7.30,$$

$$s_p = \sqrt{7.30} = 2.70,$$

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2 - 0}{2.70\sqrt{\frac{1}{n1} + \frac{1}{n2}}} = \frac{92.255 - 92.733}{2.70\sqrt{\frac{1}{8} + \frac{1}{8}}} = -0.35.$$

5. Conclusion: Since $-2.145 < t_0 = -0.35 < 2.145$, fail to reject H_0. That is, at the 0.05 level of significance, we do not have strong evidence to conclude that catalyst 2 results in a mean yield that differs from the mean yield when catalyst 1 is used.

Case 2: $\sigma_1^2 \neq \sigma_2^2$

- In some situations, we cannot reasonably assume that the unknown σ_1^2 and σ_2^2 are equal.
- There is not an exact t-statistic available for testing

  $H_0: \mu_1 - \mu_2 = \Delta_0$ in this case.
- However, $H_0: \mu_1 - \mu_2 = \Delta_0$ true, the statistic

$$T_0^* = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{\sqrt{\frac{S_1^2}{n1} + \frac{S_2^2}{n2}}}$$

  is distributed approximately as t.

- The degrees of freedom given by

$$v = \frac{\left(\frac{S_1^2}{n1} + \frac{S_2^2}{n2}\right)^2}{\frac{\left(S_1^2/n1\right)^2}{n1-1} + \frac{\left(S_2^2/n2\right)^2}{n2-1}}.$$

Example

Arsenic concentration in public drinking water supplies is a potential health risk. An article in the *Arizona Republic* (Sunday, May 27, 2001) reported drinking water arsenic concentrations in parts per billion (ppb) for 10 metropolitan Phoenix communities and 10 communities in rural Arizona. The data is given as follows:

Metro Phoenix ($\bar{x}_1 = 12.5, s_1 = 7.63$)	Rural Arizona ($\bar{x}_2 = 27.5, s_2 = 15.3$)
Phoenix, 3	Rimrock, 48
Chandler, 7	Goodyear, 44
Gilbert, 25	New River, 40
Glendale, 10	Apachie Junction, 38
Mesa, 15	Buckeye, 33
Paradise Valley, 6	Nogales, 21
Peoria, 12	Black Canyon City, 20
Scottsdale, 25	Sedona, 12
Tempe, 15	Payson, 1
Sun City, 7	Casa Grande, 18

We wish to determine it there is any difference in mean arsenic concentrations between metropolitan Phoenix communities and communities in rural Arizona.

Example – Solution

1. Hypothesis statement:

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2$$

2. Given, α =0.05. The test statistic is

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{12.5 - 27.5 - 0}{\sqrt{\frac{(7.63)^2}{10} + \frac{(15.3)^2}{10}}}$$

$$= \frac{-15}{5.4065} = -2.774$$

3. Calculate the degrees of freedom:

$$v = \frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1-1}+\frac{\left(s_2^2/n_2\right)^2}{n_2-1}} = \frac{\left(\frac{(7.63)^2}{10}+\frac{(15.3)^2}{10}\right)^2}{\frac{\left((7.63)^2/10\right)^2}{9}+\frac{\left((15.3)^2/10\right)^2}{9}} = \frac{854.43}{64.653} = 13.2 \approx 13$$

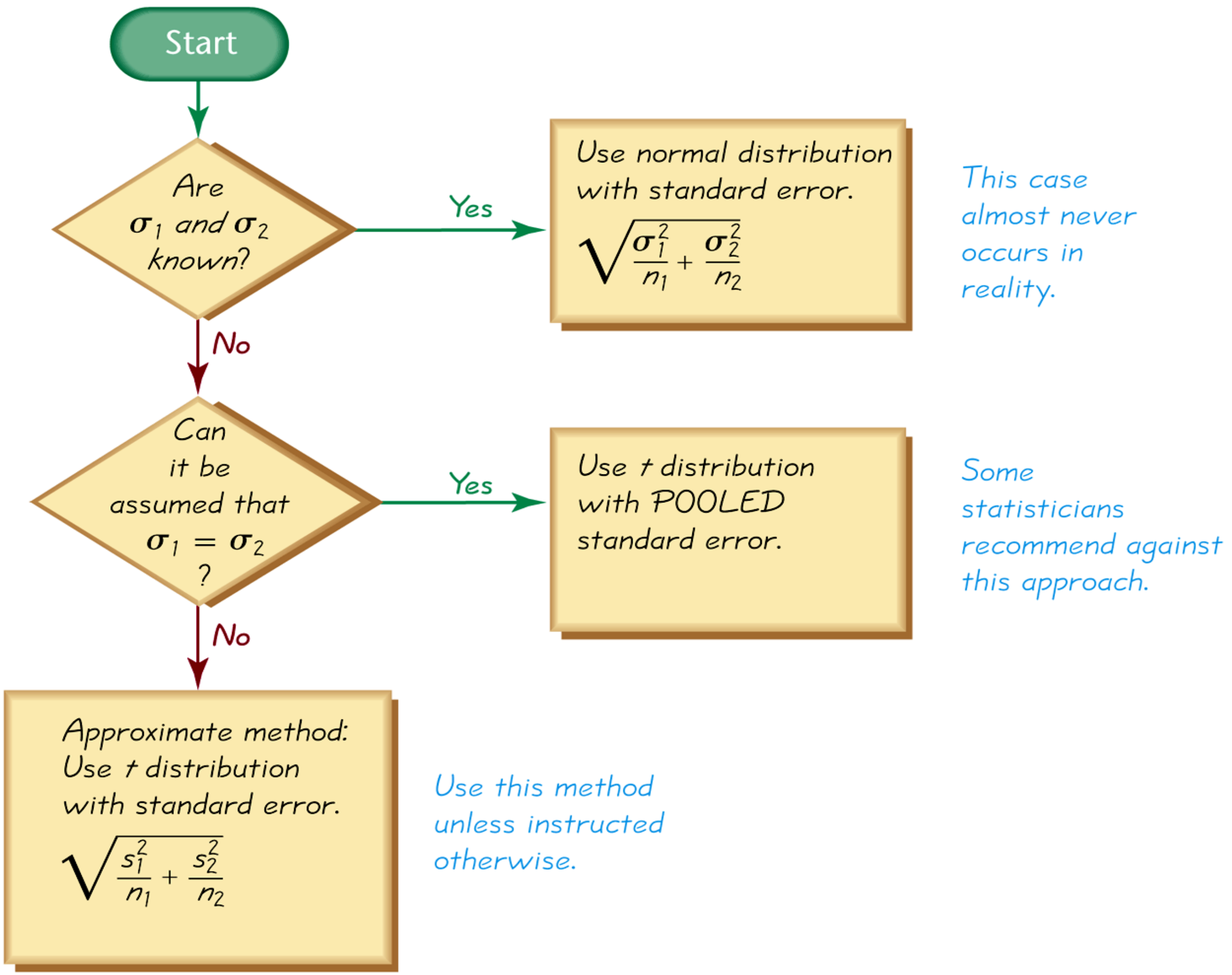
Therefore, using α =0.05, we reject H_0 if

$t_0^* > t_{0.025,13} = 2.16$ or $t_0^* < -t_{0.025,13} = -2.16$

4. Conclusion:

Since, $t_0^* = -2.77 < -t_{0.025,13} = -2.16$, we reject the null hypothesis. There is evidence to conclude that mean arsenic concentration in the drinking water in rural Arizona is different from the mean arsenic concentration in metropolitan Phoenix drinking water.

Summary



Start
Are σ_1 and σ_2 known?
Yes
Use normal distribution with standard error.
$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
This case almost never occurs in reality.
No
Can it be assumed that $\sigma_1 = \sigma_2$?
Yes
Use t distribution with POOLED standard error.
Some statisticians recommend against this approach.
No
Approximate method: Use t distribution with standard error.
$\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
Use this method unless instructed otherwise.

Statistical Inferences For Two Sample:

Test On Difference Between

VARIANCE/STANDARD DEVIATION

Test on Difference Between Two Variances

- In addition to comparing two means, statisticians are interested in comparing two variances or standard deviation. For example, is the variation in the temperature for a certain month for two cities different?
- For the comparison of two variances or standard deviations, an *F* test is used.

When can this test be used?

- There are two samples from two populations
    - These samples can be different size.
- The two samples are independent
- Both population are normally distributed
- Both population variances σ_1^2 and σ_2^2 are unknown
- We wish to test the hypotheses

$$H_0 : \sigma_1^2 = \sigma_2^2$$
$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

Let S_1^2 and S_2^2 be the sample variances.

- The statistic used is the ratio of the two sample variances:

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

- Has an F distribution with degree of freedom

  $v_1 = n_1 - 1$(numerator) and $v_2 = n_2 - 1$ (denominator)

  or written as $F_{(n_1-1, n_2-1)}$

Null hypothesis: $H_0: \sigma_1^2 = \sigma_2^2$

Test statistic: $F_0 = \dfrac{S_1^2}{S_2^2}$

Alternative Hypotheses	Rejection Criterion
$H_1: \sigma_1^2 \neq \sigma_2^2$	$f_0 > f_{\alpha/2, n_1-1, n_2-1}$ or $f_0 < f_{1-\alpha/2, n_1-1, n_2-1}$
$H_1: \sigma_1^2 > \sigma_2^2$	$f_0 > f_{\alpha, n_1-1, n_2-1}$
$H_1: \sigma_1^2 < \sigma_2^2$	$f_0 < f_{1-\alpha, n_1-1, n_2-1}$

Example

A researcher wanted to see if women varied more than men in weight. Nine women and fifteen men were weighed. The variance for the women was 525 and the variance for the men was 142. What can be concluded at the 0.05 level of significance?

Example -Solution

- Since we are testing to see if the variance for the women is larger than the variance for the men, we let S_1^2 be the women's sample variance and S_2^2 the men's sample variance.

- The null and alternative hypotheses are

$$H_0: \sigma_1^2 = \sigma_2^2$$
$$H_1: \sigma_1^2 > \sigma_2^2$$

- Since, $S_1^2 = 525; S_2^2 = 142$, so the test statistic:

$$F = \frac{525}{142} = 3.7$$

- Degree of freedom
  - Numerator, n_1-1=9-1=8
  - Denominator, n_2-1=15-1=14

- Refer to *F*-table, α=0.05

		Degrees of freedom in numerator (df1)								
	p	1	2	3	4	5	6	7	8	12
10	0.100	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.28
	0.050	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	2.91
	0.025	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.62
	0.010	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.71
	0.001	21.04	14.90	12.55	11.28	10.48	9.93	9.52	9.20	8.45
12	0.100	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.15
	0.050	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.69
	0.025	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.28
	0.010	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.16
	0.001	18.64	12.97	10.80	9.63	8.89	8.38	8.00	7.71	7.00
14	0.100	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.05
	0.050	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.53
	0.025	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.05
	0.010	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	3.80
	0.001	17.14	11.78	9.73	8.62	7.92	7.44	7.08	6.80	6.13

$$F_{0.05,8,14} = 2.70$$

- Conclusion:

Since $F = 3.7 > F_{(0.05,8,14)} = 2.7$, we reject the null hypothesis.

Hence, we have significant evidence to conclude that the variance for all female weights is larger than the variance for all male weights.

Statistical Inferences For Two Dependent Samples

(Matched Pair)

Dependent Sample

- Dependent Samples are samples that are paired or matched in some way.
- Example:
  - Samples in which the same subjects are used in a pre-post situation are dependent.
  - Another type of dependent sample are samples matched on the basis of extraneous to the study.

Example

Determine whether the sample in this case study are paired sample or not.

1) An engineering association wants to see if there is a difference in the mean annual salary for electrical engineers and chemical engineers. A random sample of electrical engineers is surveyed about their annual income. Another random sample of chemical engineers is surveyed about their annual income.

No, there is no pairing of individuals, you have two **independent** samples

2) A pharmaceutical company wants to test its new weight-loss drug. Before giving the drug to volunteers, company researchers weigh each person. After a month of using the drug, each person's weight is measured again.

Yes, you have two observations on each individual, resulting in **paired** data.

Requirements

1. The sample data are dependent.

2. The samples are simple random samples.

3. Either or both of these conditions is satisfied: The number of pairs of sample data is large (n > 30) or the pairs of values have differences that are from a population having a distribution that is approximately normal.

Notation for Dependent Samples

D	=	individual difference between the two values of a single matched pair.
μ_D	=	mean value of the differences d for the population of all matched pairs of data.
$\overline{D}$	=	mean value of the differences d for the paired sample data.
s_D	=	standard deviation of the differences d for the paired sample data.
n	=	number of pairs of sample data.

Hypotheses

- Hypotheses:

Two –tailed	Left-tailed	Right-tailed
$H_{0:} : \mu_D = 0$	$H_0 : \mu_D = 0$	$H_0 : \mu_D = 0$
$H_1 : \mu_D \neq 0$	$H_1 : \mu_D < 0$	$H_1 : \mu_D > 0$

- μ_D the expected mean of the differences of the matched pair

General Procedure – Finding the Value of Test Statistic

Step 1: Find the differences of the values of the pairs of data, D.

Step 2: Find the mean of the differences, $\bar{D}$.

$$\bar{D} = \frac{\sum D}{n}$$

Step 3: Find the standard deviation of the differences, S_D

$$S_D = \sqrt{\frac{\sum D^2 - \frac{(\sum D)^2}{n}}{n-1}}$$

Step 4: Find the test value, t

$$t = \frac{\bar{D} - \mu_D}{\frac{S_D}{\sqrt{n}}}$$, with d.f = $n - 1$

Step 5: Calculate the *p*-value

Step 6: Compare α and *p*-value, then make a decision.

Example

A physical education director claims by taking a special vitamin a weight lifter can increase his strength. Eight athletes are selected and given a test of strength, using the standard bench press. After two weeks of regular training, supplemented with the vitamin, they are tested again. Test the effectiveness of the vitamin regime at α=0.05. Each value in these data represents the maximum numbers of pounds the athlete can bench-press. Assume that the variables is approximately normally distributed.

Athlete	1	2	3	4	5	6	7	8
Before (x_1)	210	230	182	205	262	253	219	216
After (x_2)	219	236	179	204	270	250	222	216

Example - Solution

- Step 1 : State the hypotheses. In order for the vitamin to be effective, the before weight must be significantly less than the after weights; hence the mean of the differences must be less than zero

$$H_0 : \mu_D = 0 \quad \text{and} \quad H_1 : \mu_D < 0 \quad \text{(claimed)}$$

- Step 2 : Find the critical value. The degree of freedom, df = 7. The critical value for a left-tailed test with a = 0.05 is -1.895

- Step 3 :Compute the test value

Before(x_1)	After(x_2)	$D=x_1-x_2$	$D^2=(x_1-x_2)^2$
210	219	-9	81
230	236	-6	36
182	179	3	9
205	204	1	1
262	270	-8	64
253	250	-3	9
219	222	3	9
216	216	0	0
		$\sum D = -19$	$\sum D^2 = 209$

$$\overline{D} = \frac{\sum D}{n} = \frac{-19}{8} = -2.375$$

- Calculate the standard deviation of the differences:

$$s_D = \sqrt{\frac{\sum D^2 - \frac{\left(\sum D\right)^2}{n}}{n-1}} = \sqrt{\frac{209 - \frac{(-19)^2}{8}}{8-1}} = 4.84$$

- Calculate the *t*-test value:

$$t = \frac{\overline{D} - \mu_D}{\frac{s_D}{\sqrt{n}}} = \frac{-2.375 - 0}{\frac{4.84}{\sqrt{8}}} = -1.388$$

- Step 4: Make the decision. The decision is not to reject the null hypothesis at $\alpha = 0.05$ since $-1.388 > -1.895$

- Step 5: Summarise the results. There is not enough evidence to support the claim that vitamin increases the strength of weight lifter.

Example

Can playing chess improve your memory? In a study, students who had not previously played chess participated in a program in which they took chess lessons and played chess daily for 9 months. Each student took a memory test before starting the chess program and again at the end of the 9-month period.

Student	1	2	3	4	5	6	7	8	9	10	11	12
Pre-test	510	610	640	675	600	550	610	625	450	720	575	675
Post-test	850	790	850	775	700	775	700	850	690	775	540	680
Difference	-340	-180	-210	-100	-100	-225	-90	-225	-240	-55	35	-5

H_0: $\mu_d = 0$

H_1: $\mu_d < 0$

Where μ_d is the mean memory score difference between students with no chess training and students who have completed chess training

Conduct the hypothesis testing with 0.05 level of significance.

Example – Solution

- Hypothesis: H_0: $\mu_d = 0$; H_1: $\mu_d < 0$

- Test Statistic: $$t = \frac{-144.6 - 0}{109.74 / \sqrt{12}} = -4.56$$

- *t* - value from table: -1.796

- Since -4.56 < -1.796, we reject H_0. There is convincing evidence to suggest that the mean memory score after chess training is higher than the mean memory score before training.

Exercise #1

In a study of memory recall, ten students from a large statistics and data analysis class were selected at random and given 15 minutes to memorize a list of 20 nonsense words. Each was asked to list as many of the words as he or she could remember both 1 hour and 24 hours later. The data are as shown in Table 1. Is there evidence to suggest the mean number of words recalled after 1 hour exceeds the mean recall after 24 hours by more than 3? Use a significance level of 0.05.

Table 1

Student	1	2	3	4	5	6	7	8	9	10
1 hour later	14	12	18	7	11	9	16	15	13	17
24 hour later	10	4	14	5	8	5	11	12	9	10