Unit 8
Inference for the Means of Two Populations

Paired Data

We will often encounter a situation where data are collected in **pairs**. In such situations, we will usually be interested not in the individual observations, but rather in the **differences** in values of some variable for each pair.

Paired Data

The term **paired data** means that the data have been observed in natural pairs. There are several ways in which paired data can occur:

- Two different variables are measured for each individual. We are interested in the difference between the values of the two variables.
- Each individual is measured twice. The two measurements of the same characteristic are made under different conditions (or at different times).
- Similar individuals are placed in pairs and each member of the pair then receives a different treatment. The same response variable is measured and compared for the two individuals in each pair.

Matched Pairs *t* Procedures

The **matched pairs *t* procedures** will be used to help us detect and estimate any differences between responses to the two treatments.

Rather than making just one comparison for the variables of interest, we will make one comparison for **each** of the *n* pairs.

Matched Pairs *t* Procedures

The parameter of interest is μ_d, the true mean of the **differences** of all pairs in the population.

As such, in using the *t* procedures, our assumption is that the **differences follow a normal distribution** with mean μ_d and standard deviation σ_d. We also assume that our pairs represent an SRS of all possible pairs from the population.

Matched Pairs *t* Procedures

The methods for constructing confidence intervals and conducting hypothesis tests in the matched pairs setting are the **same** as the one-sample case, only now we must keep in mind that we are examining **differences** rather than individual observations.

Example

Many drivers buy premium gasoline instead of regular in the belief that they will get better gas mileage. To test this belief, we obtain a sample of eight cars. Each car is run for one tank on regular gas and one tank on premium gas, with the order randomly determined. The mileage (in miles per gallon) is measured for each car for each type of gasoline.

Example

The data are as follows:

Car	Regular	Premium	Difference (P – R)
1	20	24	4
2	26	26	0
3	24	26	2
4	22	23	1
5	23	25	2
6	19	18	– 1
7	22	27	5
8	26	29	3

Example

We will construct a 95% confidence interval for the true mean difference in mileage for premium and regular gasoline. To do this, we must assume that differences d = P – R follow a normal distribution. (Note: It is **not** enough to assume that mileages for premium gas and mileages for regular gas follow normal distributions separately. We must assume **differences** follow a normal distribution.)

We estimate the population mean difference μ_d by the sample mean difference $\bar{x}_d$ and the population standard deviation of differences σ_d by the sample standard deviation of differences s_d.

Dependent vs. Independent

Note that for any given car, mileage for premium gas and mileage for regular gas are **dependent**. In statistics, when we say two variables are dependent, this does not mean that knowing the value of one tells us the value of the other. It simply means they are **related**. (For any two cars, premium mileages are **independent** (i.e., **unrelated**), as are regular mileages for the two cars.)

Example

From the data, we calculate

$$n = 8 \qquad \bar{x}_d = 2.0 \qquad s_d = 2.0$$

Note that the sample size n corresponds to the number of **pairs** in the sample. The 95% confidence interval for μ_d is

$$\bar{x}_d \pm t^* \frac{s_d}{\sqrt{n}} = 2.0 \pm 2.365 \left(\frac{2.0}{\sqrt{8}} \right) = 2.0 \pm 1.67 = (0.33, 3.67)$$

where $t^* = 2.365$ is the upper 0.025 critical value of the t distribution with $n - 1 = 7$ d.f.

R Code

```
> Regular <- c(20, 26, 23, 22, 23, 19, 22, 26)
> Premium <- c(24, 26, 25, 23, 25, 18, 27, 29)
> t.test(Premium, Regular, paired = TRUE,
         conf.level = 0.95)
```

```
data:  Premium and Regular
t = 2.8284, df = 7, p-value = 0.02546
alternative hypothesis: true difference in means is not equal
 to 0
95 percent confidence interval:
 0.3279582 3.6720418
sample estimates:
mean of the differences
                      2
```

Example

We are 95% confident that the true mean difference between gas mileage for premium and regular gasoline is between 0.33 and 3.67.

The interpretation of the interval is analogous to the one-sample case: If we took repeated samples of eight cars and calculated the interval in a similar manner, then 95% of all such intervals would contain the true mean **difference** in gas mileage for premium and regular gasoline.

Example

Note that we could have defined the differences the other way, i.e., Regular – Premium. The signs of the differences (positive or negative) would simply switch in this case, and the interval would have been (–3.67, –0.33), giving us the same information.

Now suppose we would like to test whether these data provide convincing evidence that premium gasoline provides better gas mileage on average than regular gasoline. We will use the P-value approach.

Example

Step 1

Let $\alpha = 0.05$.

Step 2

H_0: Average gas mileage is the same for regular and premium gasoline.

H_a: Average gas mileage is higher for premium gasoline than for regular gasoline.

Equivalently, H_0: $\mu_d = 0$ vs. H_a: $\mu_d > 0$

Example

Note that we are testing whether the mean difference is **greater** than zero. This is because if premium gasoline really is better, we expect the average difference to be positive (since we defined the differences as d = P – R). If we had defined the differences as d = R – P, then this would be a **lower-tailed** test.

Example

Step 3

Reject H_0 if the P-value $\leq \alpha = 0.05$.

Step 4

The test statistic is

$$t = \frac{\bar{x}_d - \mu_{d0}}{s_d / \sqrt{n}} = \frac{\bar{x}_d}{s_d / \sqrt{n}} = \frac{2.0}{2.0 / \sqrt{8}} = 2.83$$

Example

<u>Step 5</u>

The P-value is $P(T(7) \geq 2.83)$. We see from Table 2 that

$$P(T(7) \geq 2.517) = 0.02 \quad \text{and} \quad P(T(7) \geq 2.998) = 0.01$$

Since $2.517 < t = 2.83 < 2.998$, the P-value is between 0.01 and 0.02. For any P-value between 0.01 and 0.02, we reject the null hypothesis, so our conclusion is as follows…

Example

Step 6

Since the P-value < $\alpha = 0.05$, we reject the null hypothesis. At the 5% level of significance, we have sufficient evidence that premium gasoline provides better gas mileage on average than regular gasoline.

If we had defined the differences as $d = \mathrm{R} - \mathrm{P}$, the value of the test statistic would have been $t = -2.83$ and the P-value would be $\mathrm{P}(T(7) \leq -2.83)$, which is the same as the P-value we calculated. The conclusion would be the same.

Example

The interpretation of the P-value is analogous to the one-sample case:

If there was no difference in average gas mileage for regular and premium gasoline, the probability of observing a sample mean difference at least as high as 2.0 mpg would be between 0.01 and 0.02.

R Code

```
> Regular <- c(20, 26, 23, 22, 23, 19, 22, 26)
> Premium <- c(24, 26, 25, 23, 25, 18, 27, 29)
> t.test(Premium, Regular, paired = TRUE,
         alternative = "greater")
```

```
data:  Premium and Regular
t = 2.8284, df = 7, p-value = 0.01273
alternative hypothesis: true difference in means is greater
 than 0
95 percent confidence interval:
 0.6603306       Inf
sample estimates:
mean of the differences
                      2
```

Example

If we had used the critical value method, the decision rule would be to reject the null hypothesis if $t \geq t^* = 1.895$, the upper 0.05 critical value of the t distribution with $n - 1 = 7$ d.f. The conclusion would be to reject H_0, since $t = 2.83 > t^* = 1.895$.

Example

A pharmaceutical company is testing a new blood pressure medication. The systolic blood pressures of a sample of 20 patients with hypertension are measured before and after taking the medication and we calculate the differences d = After – Before. We calculate the sample mean difference to be $\bar{x}_d = -10.9$ and the sample standard deviation of differences to be $s_d = 3.7$. Researchers would like to conduct a hypothesis test to determine whether the medication is successful in reducing systolic blood pressure. We will assume that the differences follow a normal distribution. (Note that for any given patient, blood pressure before and after taking the medication are **dependent**.) We will use the critical value method.

Example

Step 1

Let $\alpha = 0.10$.

Step 2

H_0: The medication has no effect on the mean blood pressure of patients with hypertension.

H_a: The medication reduces the mean blood pressure of patients with hypertension.

Equivalently, $H_0: \mu_d = 0$ vs. $H_a: \mu_d < 0$

Example

Step 3

Reject H_0 if $t \leq -t^* = -1.328$.

Step 4

The test statistic is

$$t = \frac{\bar{x}_d - \mu_{d0}}{s_d / \sqrt{n}} = \frac{\bar{x}_d}{s_d / \sqrt{n}} = \frac{-10.9}{3.7 / \sqrt{20}} = -13.17$$

Example

Step 5

Since $t = -13.17 < -t^* = -1.328$, we reject the null hypothesis. At the 10% level of significance, we have sufficient evidence that the medication reduces the mean blood pressure of patients with hypertension.

Example

Suppose we had instead used the P-value method to conduct the test. The P-value is $P(T(19) \leq -13.17) = P(T(19) \geq 13.17)$ by the symmetry of the t distributions. We see from Table 2 that

$$P(T(19) \geq 3.883) = 0.0005$$

Since $13.17 > 3.883$, the P-value is less than 0.0005, which is less than $\alpha = 0.10$, so we would reject H_0.

Interpretation: If there was no difference in average blood pressure before and after taking the medication, the probability of observing a sample mean difference at least as low as –10.9 would be less than 0.0005.

R Code

We find the exact P-value from R:

```
> pt(-13.17, 19)
[1] 2.650431e-11
```

P-value = 0.00000000002650431

Example

Many psychology studies have examined the differences between the first and second children in a family. Suppose we would like to compare their academic performance. We examine the high school GPAs of a sample of 30 pairs of siblings and calculate the difference in GPAs, d = first child – second child. We calculate an average difference of $\bar{x}_d = 0.12$ and a standard deviation of $s_d = 0.31$. We will assume that differences follow a normal distribution. (Note that for any pair of siblings, high school GPAs are **dependent**.)

Example

A 99% confidence interval for the true mean difference in GPAs is

$$\bar{x}_d \pm t^* \frac{s_d}{\sqrt{n}} = 0.12 \pm 2.756 \left(\frac{0.31}{\sqrt{30}}\right) = 0.12 \pm 0.156 = (-0.036, 0.276)$$

where $t^* = 2.756$ is the upper 0.005 critical value of the t distribution with $n - 1 = 29$ d.f.

Interpretation: If we took repeated samples of 30 pairs of siblings and calculated the interval in a similar manner, 99% of such intervals would contain the true mean difference in high school GPAs for the first and second children in a family.

Example

We will conduct a hypothesis test at the 1% level of significance to determine whether there is evidence of a difference in the average high school academic performance between the first and second children in a family. We will use the P-value method.

Example

Step 1

Let $\alpha = 0.01$.

Step 2

H_0: There is no difference in average high school academic performance for first and second children.

H_a: There is a difference in average high school academic performance for first and second children.

Equivalently, H_0: $\mu_d = 0$ vs. H_a: $\mu_d \neq 0$

Example

Step 3

Reject H_0 if the P-value $\le \alpha = 0.01$.

Step 4

The test statistic is

$$t = \frac{\bar{x}_d - \mu_{d0}}{s_d / \sqrt{n}} = \frac{\bar{x}_d}{s_d / \sqrt{n}} = \frac{0.12}{0.31 / \sqrt{30}} = 2.12$$

Example

Step 5

The P-value is $2P(T(29) \geq 2.12)$. We see from Table 2 that

$P(T(29) \geq 2.045) = 0.025$ and $\mathrm{P}(T(9) \geq 2.150) = 0.02$

Since $2.045 < t = 2.12 < 2.150$, $P(\mathrm{T}(29) \geq 2.12)$ is between 0.02 and 0.025, so the P-value is between 2(0.02) and 2(0.025), i.e., between 0.04 and 0.05. (Can you provide an interpretation of this P-value?)

Example

Step 6

Since the P-value > $\alpha = 0.01$, we fail to reject the null hypothesis. At the 1% level of significance, we have insufficient evidence that there is a difference in average high school academic performance between the first and second children in a family.

If we had defined the difference as $d = 2^{nd} - 1^{st}$, the value of the test statistic would have been $t = -2.12$ and the P-value would be $2P(T(29) \leq -2.12)$, which is the same as the P-value we calculated. The conclusion would be the same.

R Code

We find the exact P-value from R:

```
> 2*pt(2.12, 29, lower.tail = FALSE)
[1] 0.04268472
```

Example

If we had conducted the test using the critical value method, the decision rule would be to reject H_0 if $|t| \geq t^* = 2.756$, the upper 0.005 critical value of the t distribution with 29 d.f. The conclusion would be to fail to reject H_0, since $|t| = 2.12 < t^* = 2.756$.

Since this was a two-sided test with a 1% level of significance, we could have used the 99% confidence interval to conduct the test.

Since $\mu_{d0} = 0$ is contained in the 99% confidence interval for μ_d, we fail to reject H_0 at the 1% level of significance.

Practice Question

Measurements of the left-hand and right-hand gripping strengths of 12 left-handed people are measured. We want to conduct a hypothesis test at the 5% level of significance to determine whether the left-hand is greater on average than the right-handed gripping strength for left-handed people.

Some information that may be helpful is shown below:

Left	Right	Difference (Left – Right)
mean = 109	mean = 104	mean = 5
std. dev. = 15	std. dev. = 13	std. dev. = 6

Practice Question

Which of the following statements is/are **true**?

(I) For each person, left-hand and right-hand gripping strength are dependent.

(II) For each person, left-hand and right-hand gripping strength are independent.

(III) In order to conduct a matched pairs *t* test, we must assume that left-hand and right-hand gripping strength are both normally distributed.

(IV) In order to conduct a matched pairs *t* test, we must assume that the differences in left-hand and right-hand gripping strength are normally distributed.

(A) II (B) I and III (C) I and IV (D) II and III (E) II and IV

Practice Question

(a) Construct a 95% confidence interval for the true mean difference between left-hand and right-hand gripping strengths.

(b) Provide an interpretation of the confidence interval in (a).

(c) Conduct a hypothesis test at the 5% level of significance to determine if left-hand gripping strength is greater on average than right-hand gripping strength. Use the P-value method.

Practice Question

(d) Provide an interpretation of the P-value of the test in (c).

(e) Suppose you had instead used the critical value method to conduct the test in (c). What would be the decision rule and the conclusion?

Comparing Means of Independent Samples

In a matched pairs setting, the observations in each pair are **dependent**.

We now turn our attention to the case of comparing two population means using **independent** samples.

Comparing Means of Independent Samples

We will consider the difference between the means for Population 1 and Population 2

$$\mu_1 - \mu_2$$

We will estimate this quantity by

$$\bar{X}_1 - \bar{X}_2$$

To calculate probabilities for $\bar{X}_1 - \bar{X}_2$, we need to know how this random variable behaves; that is, we need to know its **distribution**.

Comparing Means of Independent Samples

Suppose that $X_1 \sim N(\mu_1, \sigma_1)$ and $X_2 \sim N(\mu_2, \sigma_2)$

It can be shown that the mean of $\bar{X}_1 - \bar{X}_2$ is $\mu_1 - \mu_2$ and the variance of $\bar{X}_1 - \bar{X}_2$ is

$$\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

Note that the mean of $\bar{X}_1 - \bar{X}_2$ is just the difference in means of the two variables, whereas the variance of $\bar{X}_1 - \bar{X}_2$ is the sum of the variances.

Comparing Means of Independent Samples

The standard deviation of $\bar{X}_1 - \bar{X}_2$ is thus

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

It can also be shown that $\bar{X}_1 - \bar{X}_2$ follows a normal distribution. Therefore,

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

Comparing Means of Independent Samples

We know that if a variable follows a normal distribution, then the standardized variable has a standard normal distribution:

$$Z = \frac{variable - mean}{standard\ deviation} \sim N(0,1)$$

Therefore,

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

Sampling Distribution of $\bar{X}_1 - \bar{X}_2$

An illustration of what we mean by "the sampling distribution of $\bar{X}_1 - \bar{X}_2$":

Let X_1 be the weight of a Fuji apple and let X_2 be the weight of a Florida orange. Suppose it is known that

$$X_1 \sim \mathrm{N}(200, 25) \quad \text{and} \quad X_2 \sim \mathrm{N}(150, 20)$$

We take a random sample of four apples and three oranges and calculate the average weights of both types of fruit. What is the distribution of $\bar{X}_1 - \bar{X}_2$?

Sampling Distribution of $\bar{X}_1 - \bar{X}_2$

Imagine taking millions of random samples of 4 apples and 3 oranges, and calculating $\bar{x}_1 - \bar{x}_2$ for each sample:

Sample	# of Apples n_1	Avg. Weight of Apples $\bar{x}_1$	# of Oranges n_2	Avg. Weight of Oranges $\bar{x}_2$	$\bar{x}_1 - \bar{x}_2$
1	4	192	3	148	44
2	4	215	3	137	78
3	4	188	3	165	23
4	4	207	3	170	37
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
1,000,000	4	212	3	157	55
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.

Sampling Distribution of $\bar{X}_1 - \bar{X}_2$

Now make a histogram of all of these millions of values:

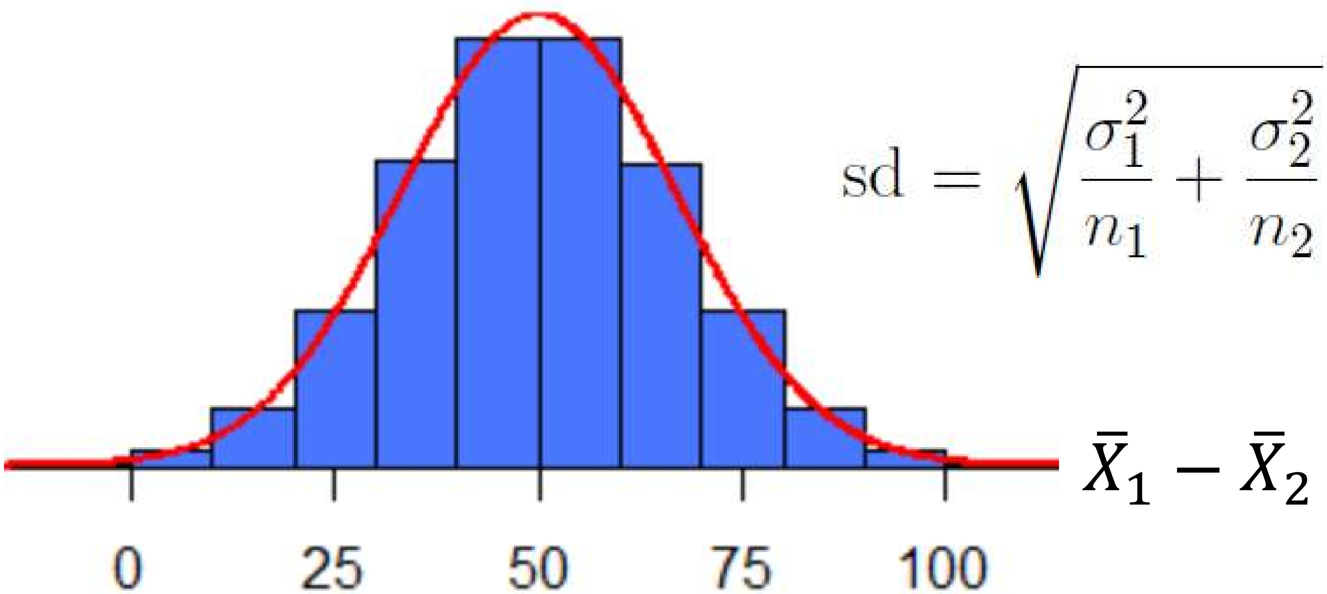
Sample	# of Apples n_1	Avg. Weight of Apples $\bar{x}_1$	# of Oranges n_2	Avg. Weight of Oranges $\bar{x}_2$	$\bar{x}_1 - \bar{x}_2$
1	4	192	3	148	44
2	4	215	3	137	78
3	4	188	3	165	23
4	4	207	3	170	37
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
1,000,000	4	212	3	157	55
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.

Sampling Distribution of $\bar{X}_1 - \bar{X}_2$

And we get the sampling distribution of $\bar{X}_1 - \bar{X}_2$:

$$\text{mean} = \mu_1 - \mu_2 = 200 - 150 = 50$$

$$\text{sd} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{25^2}{4} + \frac{20^2}{3}} = 17$$



Comparing Means of Independent Samples

We will now conduct statistical inference procedures for estimating and testing for the difference between two population means based on independent samples.

We take a simple random sample of size n_1 from Population 1 and measure some variable X_1. We take an **independent** simple random sample of size n_2 from Population 2 and measure the value of some variable X_2.

Two-Sample *t* Procedures

The values of the population standard deviations are usually unknown (since we don't have information about the entire population). In this case, we estimate the population standard deviations σ_1 and σ_2 by the sample standard deviations s_1 and s_2, respectively.

Two-Sample *t* Procedures

In order to compare means for two independent populations, we use the **two-sample *t* statistic**:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

The formulas we will use will differ slightly, depending on whether the population standard deviations σ_1 and σ_2 are equal or not.

Two-Sample *t* Procedures

Question: How do we know whether the population standard deviations σ_1 and σ_2 are equal or not?

Answer: We can never be sure, because we don't have information about the whole population.

Two-Sample *t* Procedures

There are formal hypothesis tests to assess the equality of variances for two populations. In this course, however, we will use the following simple rule:

We divide the higher of the two sample standard deviations by the lower:

$$\frac{\max s_i}{\min s_i} \quad (i = 1,\ 2)$$

If this quantity is 2 or less, we can assume equal population standard deviations, $\sigma_1 = \sigma_2$. If it is greater than 2, we must assume that the population standard deviations differ, $\sigma_1 \neq \sigma_2$.

Two-Sample *t* Procedures

In other words, if the higher of the two sample standard deviations is no more than double the value of the lower sample standard deviation, we will assume the population standard deviations are **equal**.

If the higher of the two sample standard deviations **is** more than double the value of the lower sample standard deviation, we will assume the population standard deviations are **unequal**.

The Case of Unequal Population Standard Deviations

When the population standard deviations σ_1 and σ_2 are **not** equal, the two-sample *t* statistic

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

does **not** have an exact *t* distribution.

The Case of Unequal Population Standard Deviations

We can, however, approximate the distribution of this two-sample t statistic by using a t distribution with degrees of freedom is estimated by

$$df = min\{n_1 - 1, n_2 - 1\}$$

We now examine how we conduct statistical inference to compare two population means in the case where the population standard deviations are unequal.

The Case of Unequal Population Standard Deviations

We take a simple random sample of size n_1 from Population 1 and an independent simple random sample of size n_2 from Population 2, where the population standard deviations σ_1 and σ_2 are **unknown** and **unequal**. A 100(1 – α)% confidence interval for $\mu_1 - \mu_2$ is

$$\bar{x}_1 - \bar{x}_2 \pm t^* \sqrt{\frac{s_1^2}{n1} + \frac{s_2^2}{n_2}}$$

where t^* is the upper α/2 critical value from the t distribution with degrees of freedom equal to $\min\{n_1 - 1, n_2 - 1\}$.

Example

We would like to estimate the difference in the mean gasoline price in Winnipeg and Calgary.

The gasoline prices (in cents/litre) for a random sample of 8 Winnipeg gas stations and 5 Calgary gas stations are recorded one day and are shown below:

Winnipeg:	199.9	202.7	201.7	200.9	
	201.4	202.5	199.5	201.8	
Calgary:	195.4	199.6	198.5	193.4	196.5

Example

Let X_1 be the price of gas at a Winnipeg gas station and let X_2 be the price of gas at a Calgary gas station. From the data, we calculate

$$n_1 = 8, \quad \bar{x}_1 = 201.30, \quad s_1 = 1.145$$
$$n_2 = 5, \quad \bar{x}_2 = 196.68, \quad s_2 = 2.463$$

We see that

$$\frac{max\ s_i}{min\ s_i} = \frac{2.463}{1.145} = 2.15 > 2$$

so we will conduct our inference using the method for unequal standard deviations. We must assume that gas prices in both cities follow normal distributions.

Example

We estimate the degrees of freedom: $df = \min\{n_1 - 1, n_2 - 1\} = \min\{7, 4\} = 4$. A 95% confidence interval for the difference in mean gas prices for the two cities $\mu_1 - \mu_2$ is

$$\bar{x}_1 - \bar{x}_2 \pm t^* \sqrt{\frac{s_1^2}{n1} + \frac{s_2^2}{n_2}} = 201.30 - 196.68 \pm 2.776 \sqrt{\frac{(1.145)^2}{8} + \frac{(2.463)^2}{5}}$$

$$= 4.62 \pm 3.26 = (1.36, 7.88)$$

where $t^* = 2.776$ is the upper 0.025 critical value of the t distribution with 4 d.f.

Example

We interpret the interval as follows:

If we were to take repeated samples of 8 Winnipeg gas stations and 5 Calgary gas stations and calculate the interval in a similar manner, then 95% of such intervals would contain the difference in the true mean gas prices for the two cities.

The Case of Unequal Population Standard Deviations

Now suppose we would like to conduct a hypothesis test comparing the means for Population 1 and 2. We will be testing the null hypothesis H_0: $\mu_1 = \mu_2$ (or equivalently, H_0: $\mu_1 - \mu_2 = 0$). Our test statistic is

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - 0}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} \quad \Rightarrow \quad t = \frac{(\overline{X}_1 - \overline{X}_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

Standard Error of $\bar{X}_1 - \bar{X}_2$

The quantity in the denominator of the test statistic is the **standard error of $\bar{X}_1 - \bar{X}_2$**.

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

Recall that the standard error of an estimator is the estimated standard deviation.

Example

We would like to test at the 5% level of significance whether there is a difference in mean gas prices for the two cities. We will use the P-value method.

Let $\alpha = 0.05$.

We are testing the hypotheses

$$\mathrm{H}_0 : \mu_1 = \mu_2 \quad \text{vs.} \quad \mathrm{H}_a : \mu_1 \neq \mu_2$$

Example

We will reject H_0 if the P-value $\leq \alpha = 0.05$.

The test statistic is

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{201.30 - 196.68}{\sqrt{\frac{(1.145)^2}{8} + \frac{(2.463)^2}{5}}} = 3.94$$

Example

The P-value is $2P(T(4) \geq 3.94)$.

We see from Table 2 that

$P(T(4) \geq 3.747) = 0.01$ and $P(T(4) \geq 4.604) = 0.005$.

Since $t = 3.94$ is between these two values, it follows that $P(T(4) \geq 3.94)$ is between 0.005 and 0.01. But our P-value is double this probability, so the P-value of the test is between 0.01 and 0.02.

Example

We interpret the P-value as follows:

If the true mean gas prices were the same for the two cities, the probability of observing a difference in sample means at least as extreme as 4.62 cents per litre would be between 0.01 and 0.02.

Since the P-value $< \alpha = 0.05$, we reject H_0. At the 5% level of significance, we have sufficient evidence to conclude that there is a difference in mean gas prices in Winnipeg and Calgary.

R Code

```
> Wpg <- c(199.9, 202.7, 201.7, 200.9, 201.4,
           202.5, 199.5, 201.8)
> Cgy <- c(195.4, 199.6, 198.5, 193.4, 196.5)
> t.test(Wpg, Cgy, alternative = "two.sided",
         var.equal = FALSE, conf.level = 0.95)
```

R Code

```
data:  Wpg and Cgy
t = 3.9366, df = 5.1006, p-value = 0.01057
alternative hypothesis: true difference in means is not equal
 to 0
95 percent confidence interval:
 1.620983 7.619017
sample estimates:
mean of x mean of y
   201.30    196.68
```

Note that R uses a different (more precise) estimate of the degrees of freedom than we do, so the confidence limits and P-value may differ slightly from our calculations.

Example

If we had instead used the critical value method, the decision rule would be to reject the null hypothesis if $| t | \geq t^* = 2.776$. Since $| t | = 3.94 > t^* = 2.776$, we would reject H_0.

Example

Since this is a two-sided test, and since the confidence level of our confidence interval and the significance level of our test add up to 1, we could have used the 95% confidence interval to conduct this test.

We are testing the null hypothesis H_0: $\mu_1 = \mu_2$ or equivalently, H_0: $\mu_1 - \mu_2 = 0$. The 95% confidence interval for $\mu_1 - \mu_2$ does not contain the value 0, so we reject H_0 at the 5% level of significance.

Example

By estimating our degrees of freedom as the smaller of $n_1 - 1$ and $n_2 - 1$, we are taking a **conservative** approach in the sense that:

- Confidence intervals we calculate will be slightly **wider** than the confidence intervals we would obtain using the exact distribution (and so our confidence level will be slightly **higher** than the specified value).
- Critical values and P-values will be slightly **higher** than those we would obtain using the exact distribution. If we reject H_0 using the estimated degrees of freedom, we would definitely reject H_0 using the exact distribution.

Matched Pairs vs. Two Independent Samples

Question: How was this example different from the matched pairs case?

Answer: In the matched pairs case, samples were **dependent**. It made sense to examine the difference for each pair of individuals **separately** before making an overall comparison. In the above example, observations are **not** in pairs. The first gas station in the sample for Winnipeg is not related to the first gas station in the sample for Calgary any more than it is to any other Calgary gas station.

As such, it makes no sense to look at individual differences. Rather than examining the **mean difference** (as we did in the matched pairs t procedures), we examine the **difference in means.**

The Case of Equal Population Standard Deviations

When the population standard deviations σ_1 and σ_2 are **equal**, the two-sample *t* statistic will have an **exact** *t* distribution. In this case, we estimate the common variance σ^2 ($= \sigma_1{}^2 = \sigma_2{}^2$) by the quantity

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

where s_p^2 is called the **pooled sample variance** of X_1 and X_2. Note that the pooled variance is just a weighted average of the variances of X_1 and X_2, where the weights are the respective degrees of freedom.

The Case of Equal Population Standard Deviations

We take a simple random sample of size n_1 from Population 1 and an independent simple random sample of size n_2 from Population 2, where the population standard deviations σ_1 and σ_2 are **unknown** and **equal**. A $100(1 - \alpha)\%$ confidence interval for $\mu_1 - \mu_2$ is

$$\bar{x}_1 - \bar{x}_2 \pm t^* \sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

where t^* is the upper $\alpha/2$ critical value from the t distribution with $\boldsymbol{n_1 + n_2 - 2}$ **degrees of freedom**.

The Case of Equal Population Standard Deviations

Now suppose we would like to conduct a hypothesis test comparing the means for Population 1 and 2. We will be testing the null hypothesis H_0: $\mu_1 = \mu_2$ (or equivalently, H_0: $\mu_1 - \mu_2 = 0$).

Our test statistic is

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - 0}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad \Rightarrow \quad t = \frac{(\overline{X}_1 - \overline{X}_2)}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Under the null hypothesis, the test statistic has a t distribution with $n_1 + n_2 - 2$ degrees of freedom.

Standard Error of $\bar{X}_1 - \bar{X}_2$

The quantity in the denominator of the test statistic is the **standard error of $\bar{X}_1 - \bar{X}_2$**.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Recall that the standard error of an estimator is the estimated standard deviation.

Example

We would like to determine if the true mean size of all farms in western Saskatchewan is greater than the true mean size of all farms in eastern Saskatchewan. The sizes (in acres) of random samples of farms in both regions are shown below:

West:	1012	894	743	952	1135	880		
East:	891	615	695	818	1070	940	860	743

Sizes of farms in both regions are known to follow normal distributions.

Example

Let X_1 be the size of a farm in western Saskatchewan and let X_2 be the size of a farm in eastern Saskatchewan. From the data, we calculate

$$\begin{array}{lll} n_1 = 6, & \bar{x}_1 = 936, & s_1 = 132.5 \\ n_2 = 8, & \bar{x}_2 = 829, & s_2 = 144.7 \end{array}$$

We see that

$$\frac{max\ s_i}{min\ s_i} = \frac{144.7}{132.5} = 1.09 < 2$$

so we will conduct our inference using the method for equal standard deviations.

Example

We will construct a 95% confidence interval for the difference in the true mean farm sizes for the two regions. Let X_1 be the size of a farm in western Saskatchewan and let X_2 be the size of a farm in eastern Saskatchewan.

First, we calculate the pooled sample variance:

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

$$= \frac{5(132.5)^2 + 7(144.7)^2}{8 + 6 - 2} = \frac{234347.88}{12} = 19529$$

Example

The 95% confidence interval for $\mu_1 - \mu_2$ is:

$$\bar{x}_1 - \bar{x}_2 \pm t^* \sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

$$= 936 - 829 \pm 2.179 \sqrt{19529 \left(\frac{1}{6} + \frac{1}{8} \right)}$$

$$= 107 \pm 164.5 = (-57.5, 271.5)$$

where $t^* = 2.179$ is the upper 0.025 critical value of the t distribution with $6 + 8 - 2 = 12$ d.f.

R Code

```
> West <- c(1012, 894, 743, 952, 1135, 880)
> East <- c(891, 615, 695, 818, 1070, 940, 860, 743)
> t.test(West, East, var.equal = TRUE, conf.level = 0.95)
```

```
data:  West and East
t = 1.418, df = 12, p-value = 0.1816
alternative hypothesis: true difference in means is not equal
 to 0
95 percent confidence interval:
 -57.40685 271.40685
sample estimates:
mean of x mean of y 
      936       829 
```

Example

We interpret the interval as follows:

If we took repeated samples of the same sizes from the same populations and constructed the interval in a similar manner, then 95% of such intervals would contain the difference in the true mean sizes of all farms in the two regions of Saskatchewan.

We now conduct the hypothesis test to determine whether there is sufficient evidence that the average size of farms in western Saskatchewan is greater than the average size of farms in eastern Saskatchewan. We will use the P-value method.

Example

Let $\alpha = 0.05$.

We are testing the hypotheses

H_0: The true mean sizes of farms are equal in western and eastern Saskatchewan.
H_a: The true mean size of farms in western Sask. is greater than that for eastern Sask.

Equivalently, $H_0 : \mu_1 = \mu_2$ vs. $H_a : \mu_1 > \mu_2$

Example

We will reject H_0 if the P-value $\leq \alpha = 0.05$.

The test statistic is

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{936 - 829}{\sqrt{19529\left(\frac{1}{6} + \frac{1}{8}\right)}} = 1.42$$

Example

The P-value is $P(T(12) \geq 1.42)$.

We see from Table 2 that

$P(T(12) \geq 1.356) = 0.10$ and $P(T(12) \geq 1.782) = 0.05$.

Since $t = 1.42$ is between these two values, it follows that the P-value of the test is between 0.05 and 0.10.

Example

We interpret the P-value as follows:

If the true mean sizes of farms in the two regions were equal, the probability of observing a difference in sample means at least as high as 107 acres would be between 0.05 and 0.10.

Since the P-value > α = 0.05, we fail to reject the null hypothesis. At the 5% level of significance, we have insufficient evidence to conclude that the true mean sizes of farms in western Saskatchewan is greater than that for farms in eastern Saskatchewan.

R Code

```
> West <- c(1012, 894, 743, 952, 1135, 880)
> East <- c(891, 615, 695, 818, 1070, 940, 860, 743)
> t.test(West, East, var.equal = TRUE, alternative =
         "greater")
```

```
data:  West and East
t = 1.418, df = 12, p-value = 0.09081
alternative hypothesis: true difference in means is greater
 than 0
95 percent confidence interval:
 -27.48621       Inf
sample estimates:
mean of x mean of y 
      936       829 
```

Example

Suppose we had instead used the critical value method to conduct the test. The decision rule would be to reject H_0 if $t \geq t^* = 1.782$, the upper 0.05 critical value of the t distribution with 12 d.f. Since $t = 1.42 < t^* = 1.728$, we would fail to reject H_0.

Example

We would like to compare the effectiveness of two pain relief medications. 29 people who experience frequent headaches volunteer to participate in an experiment. 15 volunteers are randomly assigned to take Tylenol the next time they have a headache, and the other 14 will take Advil. The time until each individual experiences pain relief will be recorded. We will assume that relief times follow a normal distribution for both Tylenol and Advil.

Example

After the experiment is concluded, it is calculated that the mean time for the 15 people taking Tylenol to experience relief was 133 minutes and the standard deviation was 35 minutes. The mean time for the 14 people taking Advil to experience relief was 117 minutes, and the standard deviation was 28 minutes.

Example

Let us construct a 98% confidence interval for the difference in the true mean times to experience relief for the two medications. Let X_1 be the relief time for a person taking Tylenol and let X_2 be the relief time for a person taking Advil. Since

$$\frac{max\ s_i}{min\ s_i} = \frac{35}{28} = 1.25 < 2$$

we will use the pooled method, assuming equal population standard deviations.

Example

First, we calculate the pooled sample variance:

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

$$= \frac{14(35)^2 + 13(28)^2}{15 + 14 - 2} = \frac{27342}{27} = 1012.67$$

Example

The 98% confidence interval for $\mu_1 - \mu_2$ is:

$$\bar{x}_1 - \bar{x}_2 \pm t^* \sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

$$= 133 - 117 \pm 2.473 \sqrt{1012.67 \left(\frac{1}{15} + \frac{1}{14} \right)}$$

$$= 16 \pm 29.24 = (-13.24, 45.24)$$

where $t^* = 2.473$ is the upper 0.01 critical value of the t distribution with $15 + 14 - 2 = 27$ d.f.

Example

We interpret the interval as follows:

If we took repeated samples of the same sizes from the same populations and constructed the interval in a similar manner, then 98% of such intervals would contain the difference in the true mean relief times for the two medications.

We now conduct the hypothesis test to determine whether there is sufficient evidence that the average relief time for Tylenol differs from that of Advil. We will use the critical value method.

Example

Let $\alpha = 0.02$.

We are testing the hypotheses

H_0: The true mean relief times for Tylenol and Advil are equal.
H_a: The true mean relief times for Tylenol and Advil are different.

Equivalently, $H_0 : \mu_1 = \mu_2$ vs. $H_a : \mu_1 \neq \mu_2$

Example

We will reject H_0 if $|t| \geq t^* = 2.473$.

The test statistic is

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{133 - 117}{\sqrt{1012.67\left(\frac{1}{15} + \frac{1}{14}\right)}} = 1.35$$

Since $-2.473 < t = 1.35 < 2.473$, we fail to reject H_0. At the 2% level of significance, we have insufficient evidence to conclude that the true mean relief times for the two medications differ.

Example

Suppose we had instead used the P-value method to conduct the test. The P-value is $2P(T(27) \geq 1.35)$.

We see from Table 2 that

$P(T(27) \geq 1.314) = 0.10$ and $P(T(27) \geq 1.703) = 0.05$.

Since $t = 1.35$ is between these two values, it follows that $P(T(27) \geq 1.35)$ is between 0.05 and 0.10. But our P-value is double this probability, and so the P-value of the test is between 0.10 and 0.20.

Example

Since the P-value > $\alpha = 0.02$, we fail to reject the null hypothesis.

We interpret the P-value as follows:

If the true mean relief times for Tylenol and Advil were equal, the probability of observing a difference in sample means at least as extreme as 16 minutes would be between 0.10 and 0.20.

R Code

We find the exact P-value from R:

```
> 2*pt(1.35, 27, lower.tail = FALSE)
[1] 0.1882257
```

Example

Since this is a two-sided test, and since the confidence level of our interval and the significance level of our test add up to 1, we could have used the 98% confidence interval to conduct this test.

We are testing the null hypothesis H_0: $\mu_1 = \mu_2$ or equivalently, H_0: $\mu_1 - \mu_2 = 0$. The 98% confidence interval for $\mu_1 - \mu_2$ contains the value 0, so we fail to reject H_0 at the 2% level of significance.

Example

Do drivers slow down when they know there is a radar camera at an intersection? We measure the speeds (in km/h) for a random sample of 31 cars driving through an intersection with no camera, and a random sample of 50 cars driving through a nearby intersection with a camera. The speed limit at both intersections is 50 km/h.

The mean and standard deviation for the cars at the intersection with no camera are 56.7 and 7.7 km/h, respectively. The mean and standard deviation for the cars at the intersection with a camera are 51.6 and 3.5, respectively.

Example

Let X_1 be the speed of a car at the intersection with no camera and let X_2 be the speed of a car at the intersection with a camera. We assume that both X_1 and X_2 follow normal distributions

We see that

$$\frac{max\ s_i}{min\ s_i} = \frac{7.7}{3.5} = 2.2 > 2$$

so we will conduct our inference using the method for unequal standard deviations.

Example

We estimate the degrees of freedom: $df = \min\{n_1 - 1, n_2 - 1\} = \min\{30, 49\} = 30$. A 99% confidence interval for the difference in mean speeds at the two intersections $\mu_1 - \mu_2$ is

$$\bar{x}_1 - \bar{x}_2 \pm t^* \sqrt{\frac{s_1^2}{n1} + \frac{s_2^2}{n_2}} = 56.7 - 51.6 \pm 2.750 \sqrt{\frac{(7.7)^2}{31} + \frac{(3.5)^2}{50}}$$

$$= 5.1 \pm 4.04 = (1.06, 9.14)$$

where $t^* = 2.750$ is the upper 0.005 critical value of the t distribution with 30 d.f. (Can you interpret this interval?)

Example

We would like to test at the 1% level of significance if the true mean speed of cars at the intersection with no camera is greater than that of cars at the intersection with a camera. We will use the critical value method.

Let $\alpha = 0.01$.

We are testing the hypotheses

$$\mathrm{H}_0 : \mu_1 = \mu_2 \quad \text{vs.} \quad \mathrm{H}_a : \mu_1 > \mu_2$$

Example

We will reject H_0 if $t \geq t^* = 2.457$.

The test statistic is

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{56.7 - 51.6}{\sqrt{\frac{(7.7)^2}{31} + \frac{(3.5)^2}{50}}} = 3.47$$

Example

Since $t = 3.47 > t^* = 2.457$, we reject H_0. At the 1% level of significance, we have sufficient evidence to conclude that the true mean speed at the intersection without a camera is greater than the true mean speed at the intersection with a camera.

Example

Suppose we had instead used the P-value method to conduct the test. The P-value is $P(T(30) \geq 3.47)$.

We see from Table 2 that

$P(T(30) \geq 3.385) = 0.001$ and $P(T(30) \geq 3.646) = 0.0005$.

Since $t = 3.74$ is between these two values, it follows that the P-value is between 0.0005 and 0.001.

Since the P-value $< \alpha = 0.01$, we reject H_0. (Can you interpret this P-value?)

Example

Note that we assumed speeds at both intersections follow a normal distribution. This is typically not true; speeds usually follow a right skewed distribution.

Regardless, our sample sizes for both groups are high, so the sampling distributions of the sample means are very close to normal. The use of the t distribution is therefore justified.

Practice Question

The prices of a sample of five items are recorded at two grocery stores. The data are shown in the table below, as well as the calculation of some summary statistics:

Item	2L Pepsi	Dempster's Bread	Kleenex Tissues	Dairyland 4L Milk	Quaker Granola Bars	$\bar{x}$	s
Store 1	\$3.29	\$4.99	\$3.79	\$5.35	\$3.19	\$4.12	\$0.99
Store 2	\$2.89	\$4.25	\$3.49	\$5.89	\$2.50	\$3.80	\$1.34

We would like to conduct a hypothesis test to determine if prices differ on average at the two stores.

Practice Question

Item	2L Pepsi	Dempster's Bread	Kleenex Tissues	Dairyland 4L Milk	Quaker Granola Bars	$\bar{x}$	s
Store 1	\$3.29	\$4.99	\$3.79	\$5.35	\$3.19	\$4.12	\$0.99
Store 2	\$2.89	\$4.25	\$3.49	\$5.89	\$2.50	\$3.80	\$1.34

Assuming the necessary normality conditions are satisfied, the appropriate test of significance is:

(A) a matched pairs t test with 4 d.f.

(B) a matched pairs t test with 8 d.f.

(C) a pooled two-sample t test with 8 d.f.

(D) a pooled two-sample t test with 9 d.f.

(E) an unpooled two-sample t test with 4 d.f.

Practice Question

Random samples of offensive and defensive players in the NFL are selected, and their weights (in pounds) are recorded. The data are shown in the table below with some summary statistics.

							mean	variance
Offense	275	189	326	210	304	266	261.7	2815.5
Defense	294	226	317	248			270.3	1057.6

We would like to conduct a hypothesis test to determine if the average weight of all offensive NFL players differs from that of all defensive NFL players.

Practice Question

							mean	variance
Offense	275	189	326	210	304	266	261.7	2815.5
Defense	294	226	317	248			270.3	1057.6

Assuming the necessary normality conditions are satisfied, the appropriate test of significance is:

(A) a matched pairs t test with 5 d.f.

(B) a pooled two-sample t test with 8 d.f.

(C) a pooled two-sample t test with 9 d.f.

(D) an unpooled two-sample t test with 3 d.f.

(E) an unpooled two-sample t test with 8 d.f.

Practice Question

Trace metals in drinking water can affect the flavour of the water and unusually high concentrations can pose a health risk. Water in wells may vary in the concentration of the trace metals depending on where it is drawn from. In the paper "Trace Metals of South Indian River Region" (Environmental Studies, 1982), trace metal concentrations (in mg/L) of zinc were measured in water drawn from the bottom and the top of each of six wells. The data are as follows:

Location	1	2	3	4	5	6
Bottom	0.430	0.266	0.567	0.531	0.707	0.717
Top	0.415	0.238	0.390	0.410	0.605	0.609

Practice Question

Some information that may be helpful is shown below:

Bottom (B)	Top (T)	Difference (d = Bottom – Top)
mean = 0.536	mean = 0.445	mean = 0.092
std. dev. = 0.172	std. dev. = 0.142	std. dev. = 0.061

Differences in trace metal concentration (Bottom – Top) are known to follow a normal distribution.

(a) Construct a 95% confidence interval for the true mean difference in trace metal concentrations between the bottom and top of a well.

Practice Question

(b) Conduct a hypothesis test at the 5% level of significance to determine if there is a difference in the mean trace metal concentration between the bottom and top of a well. Use the P-value method.

(c) Could the confidence interval in (a) have been used to conduct the test in (b)? Explain.

Practice Question

Which of the following conditions is **not** required for a pooled two-sample t test to be valid?

(A) We have two independent random samples.
(B) The sample standard deviations must be equal.
(C) The population standard deviations must be equal.
(D) The populations must be normally distributed.
(E) All of the above

Practice Question

The sugar contents (in grams) for random samples of two different varieties of peaches are measured. Some summary statistics are shown below:

Variety	n	$\bar{x}$	s
A	16	14.4	1.8
B	12	13.7	1.5

Assume that sugar contents follow a normal distribution for both varieties of peaches.

Practice Question

(a) Construct a 98% for the difference in the true mean sugar contents for the two varieties of peaches.

(b) Provide an interpretation of the confidence interval in (a).

(c) Conduct a hypothesis test at the 2% level of significance to determine if there is a difference in the true mean sugar contents for the two varieties of peaches. Use the P-value method.

Practice Question

(d) Provide an interpretation of the P-value of the test in (c).

(e) Suppose you had instead used the critical value method to conduct the test in (c). What would be the decision rule and the conclusion?

(f) Could the confidence interval in (a) have been used to conduct the test in (c)? Why or why not? If the interval could have been used, what would be the conclusion, and why?

Practice Question

We would like to compare the weight losses of people following two different diet programs. Twenty-one overweight adults volunteer to participate in a weight loss study. Each person is randomly assigned to follow either Diet 1 or Diet 2 for one year. The weight loss (in pounds) is measured for each volunteer at the end of the year. Weight losses for both diet programs are known to follow normal distributions. Some summary statistics are shown below:

Diet	n	$\bar{x}$	s
1	11	32.6	10.5
2	10	21.3	4.7

Practice Question

(a) Construct a 90% for the difference in the true mean weight loss for the two diet programs.

(b) Provide an interpretation of the confidence interval in (a).

(c) Conduct a hypothesis test at the 10% level of significance to determine if People following Diet Program 1 lose more weight on average than those following Diet Program 2. Use the P-value method.

Practice Question

(d) Provide an interpretation of the P-value of the test in (d).

(e) Suppose you had instead used the critical value method to conduct the test in (c). What would be the decision rule and the conclusion of the test?

(f) Could the confidence interval in (a) have been used to conduct the test in (c)? Why or why not? If the interval could have been used, what would be the conclusion, and why?