# Stat 250 Gunderson Lecture Notes
# 9: Learning about the Difference in Population Means

### Part 1: Distribution for a Difference in Sample Means

## The Independent Samples Scenario

Recall that two samples are said to be **independent samples** when the measurements in one sample are not related to the measurements in the other sample. Independent samples are generated in a variety of ways. Some common ways:

- **Random samples are taken separately from two populations** and the same response variable is recorded for each individual.
- **One random sample** is taken and a variable is recorded for each individual, but then **units are categorized as belonging to one population or another**, e.g. male/female.
- **Participants are randomly assigned to one of two treatment conditions**, such as diet or exercise, and the same response variable, such as weight loss, is recorded for each individual unit.

If the response variable is quantitative, a researcher might compare two independent groups by looking at the **difference between the two means**.

## Sampling Distribution for the Difference in Two Sample Means

### Family Dinners and Teen Stress

A study was conducted to look at the relationship between the number of times a teen has dinner with their family and level of stress in the teen's life. Teens were asked to rate the level of stress in their lives on a point scale of 0 to 100.

The researcher would like to estimate the difference in the population mean stress level for teens who have frequent family dinners (group 1) versus the population mean stress level for teens who have infrequent family dinners (group 2).

**A Typical Summary of Responses for a Two Independent Samples Problem**

| Population | Sample Size | Sample Mean | Sample Standard Deviation |
|---|---|---|---|
| 1 Frequent | **10** | **53.5** | **15.7** |
| 2 Infrequent | **10** | **65.5** | **14.6** |

Let $\mu_1$ be the population mean stress level for all teens who have frequent family dinners.

Let $\mu_2$ be the population mean stress level for all teens who have infrequent family dinners.

We want to learn about $\mu_1$ and $\mu_2$ and how they compare to each other. We could estimate the difference in population means $\mu_1 - \mu_2$ with the difference in the sample means $\bar{x}_1 - \bar{x}_2$. Will it be a good estimate?

Can anyone say how close this observed difference in sample mean stress levels $\bar{x}_1 - \bar{x}_2$ of -12 points is to the true difference in population means $\mu_1 - \mu_2$? _________

If we were to repeat this survey (with samples of the same sizes), would we get the same value for the difference in sample means? ___________________

Is a difference in the sample means of 12 points large enough to convince us that there is a real difference in the means for the populations of teens?

So what are the possible values for the difference in sample means $\bar{x}_1 - \bar{x}_2$ if we took many sets of independent random samples of the same sizes from these two populations? What would the distribution of the possible $\bar{x}_1 - \bar{x}_2$ values look like?

What can we say about the **distribution of the difference in two sample means**?

Using results from how to handle differences of independent random variables and the results for the sampling distribution for a single sample mean, the sampling distribution of the difference in two sample means $\bar{x}_1 - \bar{x}_2$ can be determined.

First recall that when working with the difference in two independent random variables:

- the mean of the difference is just the difference in the two means
- the variance of the difference is the sum of the variances

Next, remember that the standard deviation of a sample mean is $\frac{\sigma}{\sqrt{n}}$.

So what would the *variance* of a single sample mean be?

So let's apply these ideas to our newest parameter of interest, the difference in two sample means $\bar{x}_1 - \bar{x}_2$.

**Sampling Distribution of the Difference in Two (Indep) Sample Means**

If the two populations are normally distributed (or sample sizes are both large enough),

Then $\bar{x}_1 - \bar{x}_2$ is (approximately)

Since the population standard deviations of $\sigma_1$ and $\sigma_2$ are generally not known, we will use the data to compute the standard error of the difference in sample means.

**Standard Error of the Difference in Sample Means**

$$\text{s.e.}(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

where $s_1$ and $s_2$ are the two sample standard deviations

**The standard error of $\bar{x}_1 - \bar{x}_2$ estimates, roughly, the average distance of the possible $\bar{x}_1 - \bar{x}_2$ values from $\mu_1 - \mu_2$.** The possible $\bar{x}_1 - \bar{x}_2$ values result from considering all possible independent random samples of the same sizes from the same two populations.

Moreover, we can use this standard error to produce a range of values that we are very confident will contain the difference in the population means $\mu_1 - \mu_2$, namely, $\bar{x}_1 - \bar{x}_2 \pm$ (a few)s.e.($\bar{x}_1 - \bar{x}_2$). This is the basis for confidence interval for the difference in population means discussed in Part 2.

**Looking ahead:**
Do you think the 'few' in the above expression will be a $z^*$ value or a $t^*$ value?
What do you think will be the degrees of freedom?

We will use the standard error of the difference in the sample means to compute a standardized test statistic for testing hypotheses about the difference in the population means $\mu_1 - \mu_2$, namely,

$$\frac{\textbf{Sample statistic – Null value}}{\textbf{(Null) standard error}}$$

This is the basis for testing covered in Part 3.

**Looking ahead:**
Do you think the standardized test statistic will be a *z* statistic or a *t* statistic?
What do you think will be the most common null value used?

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

# Stat 250 Gunderson Lecture Notes
# Learning about the Difference in Population Means

### Part 2: Confidence Interval for a Difference in Population Means

## Confidence Interval for the Difference in Two Population Means

### General (Unpooled) Approach

- We have two populations or groups from which independent samples are available, (or one population for which two groups can be formed using a categorical variable).
- The response variable is quantitative and we are interested in comparing the means for the two populations.

**A Typical Summary of the Responses for a Two Independent Samples Problem:**

| Population | Sample Size | Sample Mean | Sample Standard Deviation |
|---|---|---|---|
| 1 | $n_1$ | $\bar{x}_1$ | $s_1$ |
| 2 | $n_2$ | $\bar{x}_2$ | $s_2$ |

Let $\mu_1$ be the mean response for the first population and $\mu_2$ be the mean response for the second population.

**Parameter of interest:** the difference in the population means $\mu_1 - \mu_2$.

**Sample estimate:** the difference in the sample means $\bar{x}_1 - \bar{x}_2$.

**Standard error:** $\text{s.e.}(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ where $s_1$ and $s_2$ are the sample standard deviations.

So we have our estimate of the difference in the two population means, namely $\bar{x}_1 - \bar{x}_2$, and we have its standard error. To make our confidence interval, we need to know the multiplier. The **multiplier *t**** is a *t*-value such that the area between $-t^*$ and $t^*$ equals the desired confidence level. The degrees of freedom for the *t*-distribution will depend on whether we use an *ugly* formula (used by software packages) or we use a conservative "by-hand" approach.

***<u>General</u>* Two Independent-Samples *t* Confidence Interval for $\mu_1$ - $\mu_2$**

$$(\bar{x}_1 - \bar{x}_2) \pm t^*(\text{s.e.}(\bar{x}_1 - \bar{x}_2))$$

where $\text{s.e.}(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ and $t^*$ is the appropriate value for a *t*-distribution, and the df can be found using an approximation or conservatively as df = smaller of ($n_1 - 1$ or $n_2 - 1$)

**This interval requires we have independent random samples from normal populations.**
If the sample sizes are large (both > 30), the assumption of normality is not so crucial and the result is approximate.

## The Pooled Approach

If we can further ***assume the population variances are (unknown but) equal***, then there is a procedure for which the $t^*$ multiplier is easier to find using an exact (not approximate) $t$-distribution. It involves pooling the sample variances for an overall estimate and updating the standard error accordingly.

It sometimes may be reasonable to assume that the measurements in the two populations have the same variances …

so that ____________________ where _____ denotes the **common population variance**.

Since both sample variances would be estimating the common population variance, it would make sense to combine or ***pool*** **the two sample variances together** to form an overall estimate.

**Pooled standard deviation:** $s_p = \sqrt{\dfrac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$

**Notes:**

(1) Each sample variance is weighted by the corresponding degrees of freedom.
So a larger sample size will result in a larger weight for that sample variance.

(2) The denominator gives the total degrees of freedom:
df =

Replacing the individual standard deviations $s_1$ and $s_2$ with the pooled version $s_p$ in the formula for the standard error leads to the pooled standard error of $\bar{x}_1 - \bar{x}_2$ is given by:

Pooled s.e.( $\bar{x}_1 - \bar{x}_2$ ) =

***<u>Pooled</u>*** **Two Independent-Samples *t* Confidence Interval for $\mu_1$ - $\mu_2$**

$$(\bar{x}_1-\bar{x}_2) \pm t^*(\text{pooled s.e.}(\bar{x}_1-\bar{x}_2))$$

where $\text{pooled s.e.}(\bar{x}_1-\bar{x}_2) = s_p\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}$

and $s_p = \sqrt{\dfrac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$ and $t^*$ is the appropriate value for a $t(n_1 + n_2 - 2)$ distribution.

**This interval requires we have independent random samples from normal populations with equal population variances.** If the sample sizes are large(both>30), the assumption of normality is not so crucial and the result is approximate.

## Notes:

(1) Some computer software packages will provide results for both the unpooled and the pooled two independent samples *t* procedures automatically. Others, such as R, will require you to explore the data in appropriate ways to help decide which method you wish to use up front as you request the analysis.

(2) First **compare the sample standard deviations**. If the sample standard deviations are similar, the assumption of common population variance is reasonable and the pooled procedure can be used. If the sample sizes happen to be the same, the pooled and unpooled standard errors are equal anyway. The advantage for the pooled version is that finding the df is simpler.

(3) A graphical tool to help assess if equal population variances is reasonable is **side-by-side boxplots**. If the lengths of the boxes (the IQRs) and overall ranges between the two groups are very different, the pooled method may not be reasonable.

(4) Some computer software also provide or allow you to produce first the results of a **Levene's test for assessing if the population variances can be assumed equal**.

The null hypothesis for this test is that the population variances are equal. So a small *p*-value for Levene's test would lead to rejecting that null hypothesis and concluding that the pooled procedure should not be used.

Often a significance level of 10% is used for this condition checking. Your lab workbook provides more details about Levene's test. We will see Levene's test results in some of our examples ahead.

**Bottom-line:** Pool if reasonable; but if the sample standard deviations are not similar, we have the unpooled procedure that can be used.

## Try It! Comparing Stress Levels Scores

A study was conducted to look at the relationship between the number of times a teen has dinner with their family and level of stress in the teen's life. Teens were asked to rate the level of stress in their lives on a point scale of 0 to 100.

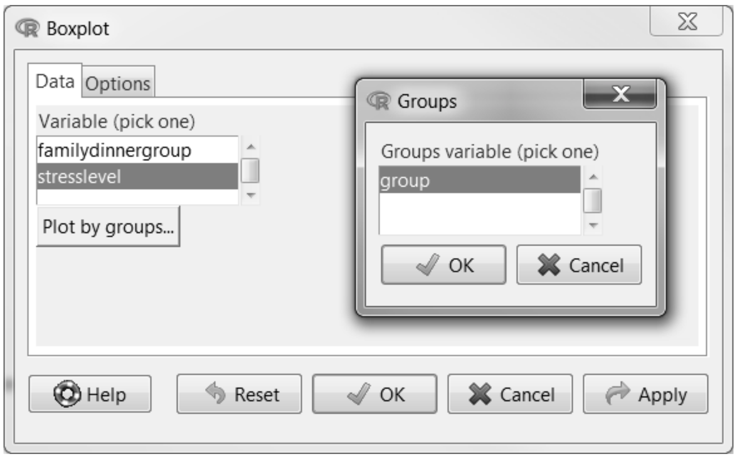
The researcher would like to estimate the difference in the population mean stress level for teens who have frequent family dinners (group 1) versus the population mean stress level for teens who have infrequent family dinners (group 2). Here is a partial listing of the data in R. Note there are two Dinner Group variables, one is numerical (as was in the original data

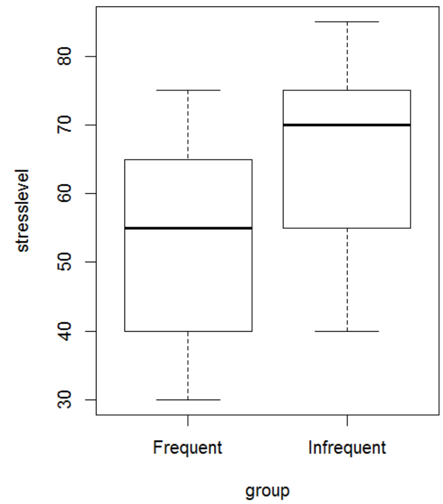
Dataset

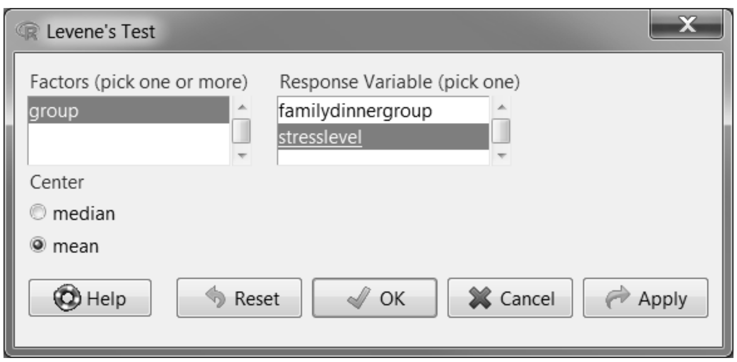
| | familydinnergroup | stresslevel | group |
|---|---|---|---|
| 1 | 1 | 75 | Frequent |
| 2 | 1 | 70 | Frequent |
| 3 | 1 | 65 | Frequent |

set) and the other is categorical (needed for R).
Some descriptive summaries, side-by-side boxplots, and Levene's Test results are provided first.

| Population | Sample Size | Sample Mean | Sample Standard Deviation |
|---|---|---|---|
| 1 Frequent | **10** | **53.5** | **15.7** |
| 2 Infrequent | **10** | **65.5** | **14.6** |



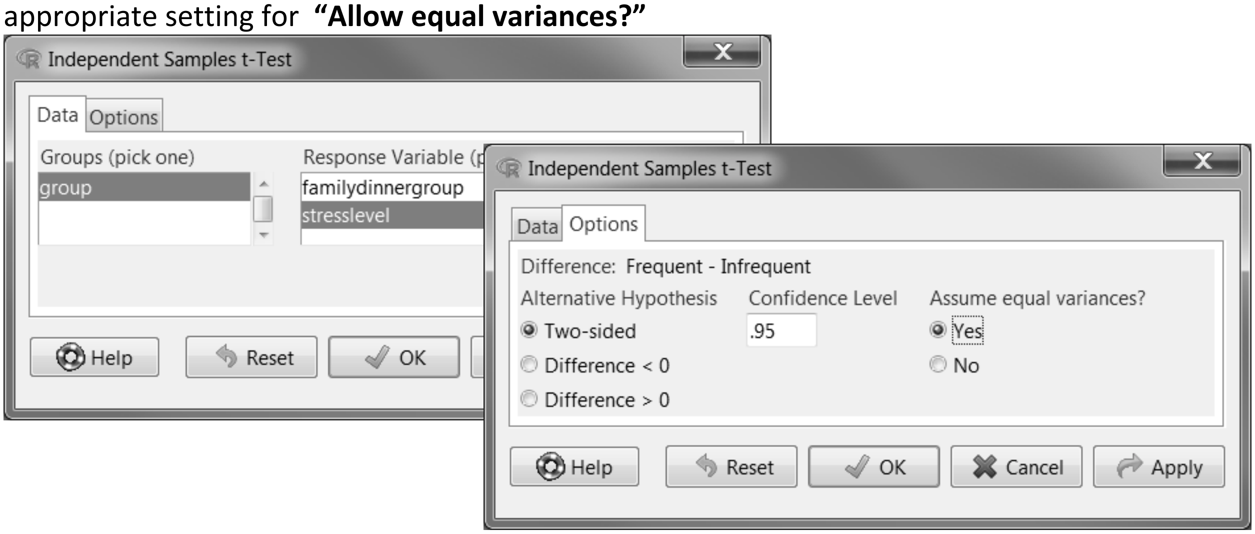

```
> with(Dataset, tapply(stresslevel, group, var,
na.rm=TRUE))
  Frequent Infrequent
  244.7222   213.6111

> leveneTest(stresslevel ~ group, data=Dataset,
center="mean")
Levene's Test for Homogeneity of Variance (center =
"mean")
      Df F value Pr(>F)
group  1  0.3958 0.5372
      18
```

a. One of the assumptions for the pooled two independent samples confidence interval to be valid is that the two populations (from which we took our samples) have the same standard deviation. Look at the two sample standard deviations, the boxplots, and the Levene's test result. Does the assumption seem to hold (at the 10% level)? Explain.

b. Give a 95% confidence interval for the difference in the population mean stress levels, that is, for $\mu_1 - \mu_2$. Show all work.

c. Based on the interval, does there appear to be a difference in the mean stress levels for the two populations? Explain.

We could use R Commander to generate the *t*-test output using **Statistics > Means > Independent-Samples T Test**. Under the **Options** tab, since we want a (two-sided) confidence interval, we select two-sided for the alternate direction. Set the confidence level and the appropriate setting for **"Allow equal variances?"**




```
> t.test(stresslevel~group, alternative='two.sided', conf.level=.95,
+   var.equal=TRUE, data=Dataset)

   Two Sample t-test

data:  stresslevel by group
t = -1.7725, df = 18, p-value = 0.09323
alternative hypothesis: true difference in means is not equal to
   0
95 percent confidence interval:
 -26.223309   2.223309
```

```
sample estimates:
  mean in group Frequent mean in group Infrequent
                    53.5                       65.5
```

**Try It! Stroop's Word Color Test**

In Stroop's Word Color Test, words that are color names are shown in colors different from the word. For example, the word red might be displayed in blue. The task is to correctly identify the display color of each word; in the example just given the correct response would be blue.

Gustafson and Kallmen (1990) recorded the time needed to complete the Color Test for *n* = 16 individuals after they had consumed alcohol and for *n* = 16 other individuals after they had consumed a placebo drink flavored to taste as if it contained alcohol. Each group was balanced with 8 men and 8 women.

In the alcohol group, the mean completion time was 113.75 seconds and standard deviation was 22.64 seconds. In the placebo group, the mean completion time was 99.87 seconds and standard deviation was 12.04 seconds.

| ***Group*** | ***Sample size*** | ***Sample mean*** | ***Sample standard deviation*** |
|---|---|---|---|
| 1 = alcohol | 16 | 113.75 | 22.64 |
| 2 = placebo | 16 | 99.87 | 12.04 |

We can assume that the two samples are independent random samples, that the model for completion time is normal for each population.

a. What graph(s) would you make to check the normality condition? Be specific.

b. How did the two groups compare descriptively?

c. Which procedure? Pooled or unpooled? Why?

d. Calculate a **95% confidence interval for the difference in population means**.

e. Based on the confidence interval, can we conclude that the population means for the two groups are different? Why or why not?

**What if?**

Suppose the researchers Gustafson and Kallmen were convinced (based on past results) that the underlying population variances were equal, so they prefer that a pooled confidence interval be constructed.

The estimate of the common population standard deviation would be:

$$s_p = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}} = \sqrt{\frac{(16-1)(22.64)^2+(16-1)(12.04)^2}{16+16-2}} = \sqrt{328.77} = 18.13$$

The pooled standard error for the difference in the two sample means would be:

$$\text{pooled s.e.}(\bar{x}_1-\bar{x}_2) = s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} = 18.13\sqrt{\frac{1}{16}+\frac{1}{16}} = 6.41$$

which is the same as the unpooled standard error since the sample sizes were equal.

The $t^*$ multiplier would be based on df = 16 + 16 – 2 = 30, so $t^*$ = 2.04 (from Table A.2).

The 95% Pooled Confidence Interval would be:

$$(\bar{x}_1-\bar{x}_2) \pm t^*(\text{pooled s.e.}(\bar{x}_1-\bar{x}_2))$$

→ (13.88) ± (2.04)(6.41) → 13.88 ± 13.08 → (0.80, 26.96)

This interval still does not include 0, so the same decision would be made; however, the interval is a bit narrower. In this example, the unpooled interval may be a bit conservative (wider) but the evidence is still strong to state the two population means appear to differ.

## Stat 250 Formula Card

| Two Population Means | |
|---|---|
| General | Pooled |
| **Parameter** $\mu_1-\mu_2$ | **Parameter** $\mu_1-\mu_2$ |
| **Statistic** $\bar{x}_1-\bar{x}_2$ | **Statistic** $\bar{x}_1-\bar{x}_2$ |
| **Standard Error** $\text{s.e.}(\bar{x}_1-\bar{x}_2)=\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$ | **Standard Error** $\text{pooled s.e.}(\bar{x}_1-\bar{x}_2)=s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$ where $s_p=\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$ |
| **Confidence Interval** $(\bar{x}_1-\bar{x}_2)\pm t^*(\text{s.e.}(\bar{x}_1-\bar{x}_2))$ df = $\min(n_1-1, n_2-1)$ | **Confidence Interval** $(\bar{x}_1-\bar{x}_2)\pm t^*(\text{pooled s.e.}(\bar{x}_1-\bar{x}_2))$ df = $n_1+n_2-2$ |
| **Two-Sample *t*-Test** $t=\frac{\bar{x}_1-\bar{x}_2-0}{\text{s.e.}(\bar{x}_1-\bar{x}_2)}=\frac{\bar{x}_1-\bar{x}_2}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$ df = $\min(n_1-1, n_2-1)$ | **Pooled Two-Sample *t*-Test** $t=\frac{\bar{x}_1-\bar{x}_2-0}{\text{pooled s.e.}(\bar{x}_1-\bar{x}_2)}=\frac{\bar{x}_1-\bar{x}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ df = $n_1+n_2-2$ |

# Stat 250 Gunderson Lecture Notes
# 9: Learning about the Difference in Population Means

### Part 3: Testing about a Difference in Population Means

## Testing Hypotheses about the Difference in Two Population Means

- We have two populations or groups from which independent samples are available, (or one population for which two groups can be formed using a categorical variable).
- The response variable is quantitative and we are interested in testing hypotheses about the means for the two populations.

**A Typical Summary of the Responses for a Two Independent Samples Problem:**

| Population | Sample Size | Sample Mean | Sample Standard Deviation |
|---|---|---|---|
| 1 | $n_1$ | $\bar{x}_1$ | $s_1$ |
| 2 | $n_2$ | $\bar{x}_2$ | $s_2$ |

Let $\mu_1$ be the mean response for the first population and $\mu_2$ be the mean response for the second population.

**Parameter of interest:** the difference in the population means $\mu_1 - \mu_2$.

**Sample estimate:** the difference in the sample means $\bar{x}_1 - \bar{x}_2$.

**Standard error:** $$\text{s.e.}(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

where $s_1$ and $s_2$ are the two sample standard deviations

**Pooled standard error:** $$\text{pooled s.e.}(\bar{x}_1 - \bar{x}_2) = s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

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

Recall there are two methods for conducting inference for the difference between two population means for independent samples – the **General (Unpooled) Case** and the **Pooled Case**. Both require we have independent random samples from normal populations (but if the sample sizes are large, the assumption of normality is not so crucial). Both will result in a t-test statistic, but the standard error used in the denominator differ as well as the degrees of freedom used for computing the *p*-value using a t-distribution.

Here is the summary for these two significance tests:
**Possible null and alternative hypotheses.**

1. **$H_0$:** **versus $H_a$:**

2. **$H_0$:** **versus $H_a$:**

3. **$H_0$:** **versus $H_a$:**

**Test statistic = Sample statistic – Null value / Standard error**

| Two Population Means | |
|---|---|
| General | Pooled |
| **Parameter** $\mu_1 - \mu_2$ | **Parameter** $\mu_1 - \mu_2$ |
| **Statistic** $\bar{x}_1 - \bar{x}_2$ | **Statistic** $\bar{x}_1 - \bar{x}_2$ |
| **Standard Error** $\text{s.e.}(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ | **Standard Error** $\text{pooled s.e.}(\bar{x}_1 - \bar{x}_2) = s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ where $s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$ |
| **Confidence Interval** $(\bar{x}_1 - \bar{x}_2) \pm t^*(\text{s.e.}(\bar{x}_1 - \bar{x}_2))$ $\quad$ $\text{df} = \min(n_1 - 1, n_2 - 1)$ | **Confidence Interval** $(\bar{x}_1 - \bar{x}_2) \pm t^*(\text{pooled s.e.}(\bar{x}_1 - \bar{x}_2))$ $\quad$ $\text{df} = n_1 + n_2 - 2$ |
| **Two-Sample *t*-Test** $t = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\text{s.e.}(\bar{x}_1 - \bar{x}_2)} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ $\quad$ $\text{df} = \min(n_1 - 1, n_2 - 1)$ | **Pooled Two-Sample *t*-Test** $t = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\text{pooled s.e.}(\bar{x}_1 - \bar{x}_2)} = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ $\quad$ $\text{df} = n_1 + n_2 - 2$ |

Recall the **guidelines** to assess which version to use:

(1) First **compare the sample standard deviations**. If the sample standard deviations are similar, the assumption of common population variance is reasonable and the pooled procedure can be used.

(2) A graphical tool to help assess if equal population variances is reasonable is **side-by-side boxplots**. If the lengths of the boxes (the IQRs) and overall ranges between the two groups are very different, the pooled method may not be reasonable.

(3) Examine the results of a **Levene's test for assessing if the population variances can be assumed equal**. The null hypothesis for this test is that the population variances are equal. So a small *p*-value for Levene's test would lead to rejecting that null hypothesis and concluding that the pooled procedure should not be used.

**Bottom-line:** Pool if reasonable; but if the sample standard deviations are not similar, we have the unpooled procedure that can be used.

## Try It! Effect of Beta-blockers on pulse rate

Do beta-blockers reduce the pulse rate? In a study of heart surgery, 60 subjects were randomly divided into two groups of 30. One group received a beta-blocker while the other group was given a placebo. The pulse rate at a particular time during the surgery was measured. The results are given below.

| *Group* | *Sample size* | *Sample mean* | *Sample standard deviation* |
|---|---|---|---|
| **1=beta-blockers** | 30 | 65.2 | 7.8 |
| **2=placebo** | 30 | 70.3 | 8.4 |

a. State the hypotheses to assess if beta-blockers reduce pulse rate on average.

   $H_0$: ______________________ versus $H_a$: _________________________

b. Which test will you perform – the pooled or unpooled test? Explain.

c. Perform the t-test. Show all steps. Are the results significant at a 5% level?

## Try It! Does the Drug Speed Learning?

In an animal-learning experiment, a researcher wanted to assess if a particular drug ***speeds*** learning. One group of 5 rats (Group 1 = control) is required to learn to run a maze without use of the drug, whereas a second independent group of 8 rats (Group 2 = experimental) is administered the drug. The running times (time to complete the maze) for the rats in each group were entered into R.

| ***Summary Statistics*** | | | | | |
|---|---|---|---|---|---|
| **Group** | **Mean** | **Std. Dev** | **Sample Size** | **General Std. Error** | **Pooled Std. Error** |
| **Control** | 46.80 | 3.42 | 5 | 2.28 | 2.47 |
| **Experimental** | 38.38 | 4.78 | 8 | | |

```
> leveneTest(stresslevel ~ group, data=Dataset,
center="mean")
Levene's Test for Homogeneity of Variance (center =
"mean")
      Df F value Pr(>F)
group  1  1.09   0.32
      11
```

| ***Two Sample T Results*** | | | |
|---|---|---|---|
| | ***t*** | ***df*** | ***p-value*** |
| ***Unpooled*** | 3.70 | 10.653 | 0.002 |
| ***Pooled*** | 3.41 | 11 | 0.003 |

Conduct the test to address the theory of the researcher (state the null and alternate hypotheses, report the test statistic, *p*-value, and state your decision and conclusion at the 5% level of significance).

$H_0$: ___________________________ $H_a$: ___________________________

Test statistic: ______________________ *p*-value: ____________________________

Decision: (circle one) **Fail to reject $H_0$** **Reject $H_0$**

Thus …

**Try It! Eat that Dark Chocolate**

An Ann Arbor News article entitled: Dark Chocolate may help blood flow, reported the results of a study in which researchers fed a small 1.6-ounce bar of dark chocolate to each of 22 volunteers daily for two weeks. Half of the subjects were randomly selected and assigned to receive bars containing dark chocolate's typically high levels of flavonoids, and the other half received placebo bars with just trace amounts of flavonoids. The ability of the brachial artery to dilate *significantly* improved for those in the high-flavonoid group compared to those in the placebo group.

Let $\mu_1$ represent the population average improvement in blood flow for those on the high-flavonoid diet and $\mu_2$ represent the population average improvement in blood flow for those on the placebo diet. The researchers tested that the high-flavonoid group would have a higher average improvement in blood flow.

a. State the null and alternate hypotheses

   $H_0$: ______________________ versus $H_a$: __________________________

b. The researchers conducted a pooled two sample t-test. The two assumptions about the data are that the two samples are independent random samples.

   i. Clearly state one of the remaining two assumptions regarding the populations that are required for this test to be valid.

   ii. Explain how you would use the data to assess if the above assumption in part (i) is reasonable. (Be specific.)

c. A significance level of 0.05 was used. Based on the statements reported above, what can you say about the *p*-value? Clearly circle your answer:

   ***p*-value > 0.05** ***p*-value ≤ 0.05** **can't tell**

d. The researchers also found that concentrations of the cocoa flavonoid epicatechin soared in blood samples taken from the group that received the high-flavonoid chocolate, rising from a baseline of25.6 nmol/L to 204.4 nmol/L. In the group that received the low-flavonoid chocolate, concentrations of epicatechin decreased slightly, from a baseline of 17.9 nmol/L to 17.5 nmol/L. The average improvement for the high-flavonoid group of 204.4 – 25.6 = 178.8 nmol/L is a … (circle all correct answers):

   **parameter** **statistic** **sample mean** **population mean** **sampling distribution**

## Name That Scenario

Now that we have covered a number of inference techniques, let's think about some questions to ask to help dentify the appropriate procedure based on the research question(s) of interest.

1. **Is the response variable measured quantitative or categorical?**

Categorical → Proportions, percentages

$p$: One population proportion

$p_1$- $p_2$: Difference between two population proportions

Quantitative → Means

$\mu$: One population mean

$\mu_d$: Paired difference population mean

$\mu_1$ - $\mu_2$: Difference between two population means

2. **How many samples?**

If two sets of measurements – are they paired or independent?

3. **What is the main purpose?**

To estimate a numerical value of a parameter? → confidence interval

To make a 'maybe not' or 'maybe yes' type of conclusion about a specific hypothesized value? → hypothesis test

---

**Additional Notes**

A place to … jot down questions you may have and ask during office hours, take a few extra notes, write out an extra problem or summary completed in lecture, create your own summary about these concepts.

# Stats 250 Formula Card Summary

| Population Proportion | Two Population Proportions | Population Mean |
|---|---|---|
| **Parameter** $p$ | **Parameter** $p_1 - p_2$ | **Parameter** $\mu$ |
| **Statistic** $\hat{p}$ | **Statistic** $\hat{p}_1 - \hat{p}_2$ | **Statistic** $\bar{x}$ |
| **Standard Error**<br>$\text{s.e.}(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ | **Standard Error**<br>$\text{s.e.}(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$ | **Standard Error**<br>$\text{s.e.}(\bar{x}) = \frac{s}{\sqrt{n}}$ |
| **Confidence Interval**<br>$\hat{p} \pm z^* \text{s.e.}(\hat{p})$<br>**Conservative Confidence Interval**<br>$\hat{p} \pm \frac{z^*}{2\sqrt{n}}$ | **Confidence Interval**<br>$(\hat{p}_1 - \hat{p}_2) \pm z^* \text{s.e.}(\hat{p}_1 - \hat{p}_2)$ | **Confidence Interval**<br>$\bar{x} \pm t^* \text{s.e.}(\bar{x})$ df $= n - 1$<br>**Paired Confidence Interval**<br>$\bar{d} \pm t^* \text{s.e.}(\bar{d})$ df $= n - 1$ |
| **Large-Sample $z$-Test**<br>$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$ | **Large-Sample $z$-Test**<br>$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$<br>where $\hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}$ | **One-Sample $t$-Test**<br>$t = \frac{\bar{x} - \mu_0}{\text{s.e.}(\bar{x})} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$ df $= n - 1$<br>**Paired $t$-Test**<br>$t = \frac{\bar{d} - 0}{\text{s.e.}(\bar{d})} = \frac{\bar{d}}{s_d/\sqrt{n}}$ df $= n - 1$ |
| **Sample Size**<br>$n = \left(\frac{z^*}{2m}\right)^2$ | | |

| Two Population Means | |
|---|---|
| General | Pooled |
| **Parameter** $\mu_1 - \mu_2$ | **Parameter** $\mu_1 - \mu_2$ |
| **Statistic** $\bar{x}_1 - \bar{x}_2$ | **Statistic** $\bar{x}_1 - \bar{x}_2$ |
| **Standard Error**<br>$\text{s.e.}(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ | **Standard Error**<br>pooled $\text{s.e.}(\bar{x}_1 - \bar{x}_2) = s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$<br>where $s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}}$ |
| **Confidence Interval**<br>$(\bar{x}_1 - \bar{x}_2) \pm t^*(\text{s.e.}(\bar{x}_1 - \bar{x}_2))$ df $= \min(n_1 - 1, n_2 - 1)$ | **Confidence Interval**<br>$(\bar{x}_1 - \bar{x}_2) \pm t^*(\text{pooled s.e.}(\bar{x}_1 - \bar{x}_2))$ df $= n_1 + n_2 - 2$ |
| **Two-Sample $t$-Test**<br>$t = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\text{s.e.}(\bar{x}_1 - \bar{x}_2)} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ df $= \min(n_1 - 1, n_2 - 1)$ | **Pooled Two-Sample $t$-Test**<br>$t = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\text{pooled s.e.}(\bar{x}_1 - \bar{x}_2)} = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ df $= n_1 + n_2 - 2$ |