Chapter 10: BUSN 210

- Inference about Difference Between 2 Pop Means
- Experimental Design ➔
- Analysis of Variance ➔

---

**Chapter 8:**
- Create Confidence Interval for 1 Pop. Mean
  - $95\%$ Interval
  - $M \pm z_{\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}$

**Chapter 9:**
- Hypothesis Testing for 1 Pop. Mean
  - Hypothesis Testing
  - Critical Values
  - $5\%$ risk that we Reject $H_0$ & accept $H_a$, even though $H_0$ was TRUE.
  - Z or T test statistic $p$-value

**Chapter 10:**
- Hypothesis Testing & confidence Intervals for Difference Between 2 Pop. Means
- We are interested in whether 2 Pop means are different

**Example 1:**
- Mean Income in Bradford = $38,010$
- Mean Income in Kane = $35,006$
- Difference = $3,004$

Is difference due to chance, or is it a real difference (statistically significant)?
Example 2:

Sample Mean Imported Compact Car MPG = $\bar{X}_1 = 36.25$
Sample Mean Domestic compact Car MPG = $\bar{X}_2 = 33.82$

$\text{Difference} = \bar{X}_1 - \bar{X}_2 = 2.69 \text{ MPG}$

Is the difference due to chance or is a statistically significant difference?

1. $\bar{X}_1 - \bar{X}_2$ is point estimate of:

   Population Difference $= \mu_1 - \mu_2$

2. What we would like to do is take

   $\bar{X}_1 - \bar{X}_2$ and compare against

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

   and then use either:

   $Z$ - Distribution or $t$ - Distribution

   we must remember $\rightarrow$ chapter 7
Chapter 7
Sampling Distribution of $\bar{X}$
Probability Distribution of all possible values of $\bar{X}$

- $\bar{X}$ is the Random Variable
- Mean = $M_{\bar{X}}$
- SD = $\sigma_{\bar{X}}$

Mean of sampling Distribution of $\bar{X}$

$M = M_{\bar{X}} = E(\bar{X}) = \frac{\text{sum of all possible } \bar{X}}{\text{count of all possible } \bar{X}}$

If we are able to select all possible samples of size $n$ from a given population, then the mean of the sampling Distribution of $\bar{X}$ is equal to the pop. mean $M$, that is: $M = M_{\bar{X}}$

Standard Deviation of Sampling Distribution of $\bar{X}$, standard Error:

$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$

Chapter 10
Sampling Distribution of $\bar{X}_1 - \bar{X}_2$
Probability Distribution of all possible values of $\bar{X}_1 - \bar{X}_2$

- $\bar{X}_1 - \bar{X}_2$ is the Random Variable
- Mean = $M_{\bar{X}_1} - M_{\bar{X}_2}$
- SD = $\sigma_{\bar{X}_1 - \bar{X}_2}$

Mean of sampling Distribution of $\bar{X}_1 - \bar{X}_2$

$M_1 - M_2 = M_{\bar{X}_1} - M_{\bar{X}_2} = E(\bar{X}_1 - \bar{X}_2) = \frac{\text{sum of all possible } \bar{X}_1 - \bar{X}_2}{\text{count of all possible } \bar{X}_1 - \bar{X}_2}$

If we are able to select all possible sample differences of size $n$ from a given population, then the mean of the sampling Distribution of $\bar{X}_1 - \bar{X}_2$ is equal to the pop. mean $M_1 - M_2$, that is: $M_1 - M_2 = M_{\bar{X}_1} - M_{\bar{X}_2}$

Standard Deviation of Sampling Distribution of $\bar{X}_1 - \bar{X}_2$, standard Error:

$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \cdot \sqrt{\frac{N-n_1}{N-1} \cdot \frac{N-n_2}{N-1}}$

Finite Pop. correction Factor not need when $n/n_1 \leq 0.05$
Sampling Distribution of $\bar{X}_1 - \bar{X}_2$

(\text{SD of } \bar{X}_1 - \bar{X}_2)

We can take the difference between 2 samples and compare to SD of $\bar{X}_1 - \bar{X}_2$

Here? $M_{\bar{X}_1 - \bar{X}_2}$

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

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

Let's go over to Excel
1. \[ \{ \text{Difference Between Pop. Means} \} = M_1 - M_2 = M_{\bar{x}_1 - \bar{x}_2} = E(\bar{x}_1 - \bar{x}_2) = \{ \text{Mean of All Sample Differences} \} \]

2. Sampling Distribution of \( \bar{x}_1 - \bar{x}_2 \) tends to be symmetrical.

3. Less variation in Sampling Distribution of \( \bar{x}_1 - \bar{x}_2 \)

4. \[ \{ \text{Standard Deviation of } \bar{x}_1 - \bar{x}_2 \} = \{ \text{Standard Error of } \bar{x}_1 - \bar{x}_2 \} = \sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \]

* Finite Population Correction Factor not used in any of text book problems.

4. Central Limit Theorem

If both populations have a normal distribution, or if the sample size, \( n \), is big enough (\( n \geq 30 \), skewed \( n \geq 125 \))

then we can use:

1. \( \text{SD of } \bar{x}_1 - \bar{x}_2 \} \rightarrow \text{Build Confidence Intervals} \)
2. \( \text{SE} = \sigma_{\bar{x}_1 - \bar{x}_2} \} \rightarrow \text{Do Hypothesis Testing} \)
confidence interval for $M_1 - M_2$

when $\sigma_1$ and $\sigma_2$ are known

\[
\begin{align*}
\{ \text{point estimate of } M_1 - M_2 \} &= \bar{X}_1 - \bar{X}_2 \\
\{ \text{standard deviation/standard error of } \} &= \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \\
\{ \text{margin of error } \} &= \frac{Z_{\text{upper}} \sqrt{\sigma_1^2 + \sigma_2^2}}{\sqrt{n_1 + n_2}} \\
\{ \text{confidence interval } \} &= \bar{X}_1 - \bar{X}_2 \pm \frac{Z_{\text{upper}} \sqrt{\sigma_1^2 + \sigma_2^2}}{\sqrt{n_1 + n_2}}
\end{align*}
\]

Variables:

$\bar{X}_1$ = sample mean from population 1
$n_1$ = sample size from pop. 1
$\sigma_1$ = standard deviation from pop. 1
$\sigma_2^2$ = variation from pop. 1

$\bar{X}_2$ = sample mean from pop. 2
$n_2$ = sample size from pop. 2
$\sigma_2$ = standard deviation from pop. 2
$\sigma_2^2$ = variation from pop. 2

$M_1 - M_2$ = pop. difference between 2 pop. means

2 independent random samples (taken separately & independently so no bias.)
Two cities, Bradford and Kane are separated only by the Conewango River

The local paper recently reported that the mean household income in Bradford is $38,010 from a sample of 40 households. The population standard deviation (past data) is $6,000.

The same article reported the mean income in Kane is $35,006 from a sample of 35 households. The population standard deviation (past data) is $7,000.

From the sample data, create a 95% Confidence Interval to estimate population difference.

![Confidence Interval Example](image)

Bradford Data = 1

- \( \bar{x}_1 = \$38,010 \)
- \( n_1 = 40 \)
- \( \sigma_1 = \$6,000 \)
- \( \sigma^2_1 = 6000^2 = 36,000,000 \)

Kane Data = 2

- \( \bar{x}_2 = \$35,006 \)
- \( n_2 = 35 \)
- \( \sigma_2 = \$7,000 \)
- \( \sigma^2_2 = 7000^2 = 49,000,000 \)

Point Estimate of \( \mu_1 - \mu_2 \) = \( \bar{x}_1 - \bar{x}_2 = 38,010 - 35,006 = \$3,004 \)

\( SE of \ \text{SD of } \bar{x}_1 - \bar{x}_2 \) = \[ \sqrt{\frac{36000000 + 49000000}{40} + \frac{49000000}{35}} = 1516.575089 = 1516.57 \] \[ = \frac{0.125 \times \sigma}{\sqrt{n_1 n_2}} \]

Zupper = NORM.S.INV(1 - 0.025) = 1.959963985

Margin of Error = 1.959963985 \times 1516.575089 = 2972.43

Confidence Interval = \( \bar{x}_1 - \bar{x}_2 \pm Zupper \times SE = 3004 \pm 2972.43 \)

Point Estimate Margin of Error Lower = \$31.57 Upper = \$5976.43

We are 95% sure that population difference between the Bradford mean Income & Kane mean Income is between about \$32 and \$5976.
Hypothesis Testing about $M_1 - M_2$ with $\sigma_1^2$ and $\sigma_2^2$ known

- Same 5 steps as Chapter 9

- $M_1 - M_2 = D_0 = \text{Hypothesized Difference}$
- Possible $H_0$ & $H_a$:

<table>
<thead>
<tr>
<th>Z-tail</th>
<th>1-tail upper</th>
<th>1-tail lower</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0 : M_1 - M_2 = D_0$</td>
<td>$H_0 : M_1 - M_2 \leq D_0$</td>
<td>$H_0 : M_1 - M_2 \geq D_0$</td>
</tr>
<tr>
<td>$H_a : M_1 - M_2 \not\leq D_0$</td>
<td>$H_a : M_1 - M_2 &gt; D_0$</td>
<td>$H_a : M_1 - M_2 &lt; D_0$</td>
</tr>
</tbody>
</table>

- Test statistic $\sigma$ known for Hypothesis Testing about $M_1 - M_2$

\[ Z = \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \]

Remember $\sigma_1$ & $\sigma_2$ are SD, $\sigma_1^2$ & $\sigma_2^2$ are variance.
Hypothesis Testing continued:

- $M_1 - M_2$ = difference between 2 pop means.
- $\sigma_1 = \text{Pop SD}_1$, $\sigma_2 = \text{Pop SD}_2$
- $\sigma_1^2 = \text{Pop variance}_1$, $\sigma_2^2 = \text{Pop variance}_2$
- $n = \text{sample size}$
- If populations are both normal or $n$ is big enough to invoke Central Limit Theorem, we can use standard Normal curves.

To get statistics for Hypothesis Testing we can use:

Excel Data Ribbon Tab, Data Analysis, Descriptive statistics:

- Z - Test: Two Samples for Means
  - 2 samples must be independent random samples (so no bias)
Two cities, Bradford and Kane are separated only by the Conewango River.

The local paper recently reported that the mean household income in Bradford is $38,010 from a sample of 40 households. The population standard deviation (past data) is $6,000.

The same article reported the mean income in Kane is $35,006 from a sample of 35 households. The population standard deviation (past data) is $7,000.

At the .05 significance level can we conclude the mean income in Bradford is more?

Step 1: List Null and Alternative Hypotheses

- Null Hypothesis ($H_0$): $\mu_1 - \mu_2 \leq 0$
- Alternative Hypothesis ($H_a$): $\mu_1 - \mu_2 > 0$

Step 2: Level of Significance = Alpha = Risk of rejecting $H_0$ when it is TRUE

$\alpha = 0.05$

Step 3: Sample, Calculate, Draw Picture, Calculate Test Statistic

1. Use $Z$ because $\sigma$ known

$Z = \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$

$\bar{X}_1 = 38,010$  
$\bar{X}_2 = 35,006$  
$\sigma_1 = 6,000$  
$\sigma_2 = 7,000$  
$n_1 = 40$  
$n_2 = 35$

$Z = \frac{38,010 - 35,006}{\sqrt{\frac{6^2}{40} + \frac{7^2}{35}}} = 3.004$

Step 4: Create Rules for Critical Value and p-value

- Critical value upper = 1.6449 = $\text{NORM.S.INV}(1-.05)$
- P-value = $1 - \text{NORM.S.DIST}(1.9808) = 0.0238$

Step 5: Conclude

Because our test statistic > critical value & p-value < alpha, we reject $H_0$ & accept $H_a$. The evidence suggests that the mean income in Bradford is more than in Kane.
\[ z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \]

Critical value: 1.645

Point estimate

\[ \bar{x} = 380.10, \sigma = 35.06 \]

\[ H_0: \mu = \bar{M} - M_2 \]

\[ H_a: \mu > \bar{M} - M_2 \]

\[ 0 = \bar{M} - M_2 \]

\[ 0 < \bar{M} - M_2 \]

\[ p \text{-value} = 0.0238 \]
Confidence Interval for $M_1 - M_2$, $\sigma_1$ & $\sigma_2$ unknown

- $M_1 - M_2 =$ difference between 2 pop. means
  $\sigma$ (sigma, pop standard deviation) is not known, so we use sample standard deviation, $s$, in place of sigma, $\sigma$, & use $t$ distribution rather than $Z$ distribution.
- If populations are both normal or sample size, $n$, is big enough to invoke Central Limit Theorem, we can use $t$ Distributions.

$\overline{x} = x_bar =$ sample mean
$s = $ sample standard deviation
$s^2 = $ sample variance
$n = $ sample size
$\alpha = $ alpha = risk that pop mean is Not in Interval.
$1 - \alpha =$ confidence Limit/Coefficient

Point Estimate for $M_1 - M_2$

\[
\{\text{Upper \& Lower Limit for confidence Interval}\} \quad \overline{x}_1 - \overline{x}_2 \pm t_{upper} \sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}}
\]

$t_{upper}$ Estimates
$\# \, \text{of} \, \text{standard deviations}$

$SE = $ Estimates
$\text{Standard Deviation of sampling distribution}$
$\text{of } x_bar - x_bar \text{ for } t \text{ distribution}$
Degrees of Freedom, \( df \), for \( t \) Distribution with 2 Independent Random Samples:

\[
(df) = \frac{\left(\frac{S_1^2}{n_1}\right)^2 + \left(\frac{S_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1} + \frac{1}{n_2 - 1}}
\]

\( S_1^2 \) = Variance from Sample 1 Data
\( S_2^2 \) = Variance from Sample 2 Data
\( n_1 = \) Sample size taken from Pop. 1
\( n_2 = \) Sample size taken from Pop. 2.

**Note 1:** Best to round down \( df \) to get more conservative Interval, *Excel T functions will automatically truncate \( df \) (Rounddown).

**Note 2:** Excel Hint: Because \( S^2 \) is used so often in formula, use helper cell to calculate \( \bar{S}^2 \) & then refer to it in formula with cell reference.

**Note 3:** Does not require that \( \sigma_1 = \sigma_2 \) (like some other \( df \) formulas do). Formula works whether \( \sigma_1 = \sigma_2 \) or not!!

**Formula not used in this class:**

If you can assume \( \sigma_1 = \sigma_2 \) then:

\[
\left\{ \frac{\text{pooled variance}}{\text{weighted mean of 2 variances}} \right\} = \frac{\bar{S}^2}{\bar{S}_p} = \frac{(n_1 - 1) * S_1^2 + (n_2 - 1) * S_2^2}{n_1 + n_2 - 2}
\]

\[
t \text{ test statistic} = \frac{(\bar{X}_1 - \bar{X}_2) - \Delta}{\sqrt{S_{SP} \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}
\]
A recent EPA study compared the highway fuel economy of domestic and imported compact cars. A sample of 20 domestic compact cars revealed a mean of 33.83 MPG with a standard deviation of 2.33 MPG. A sample of 20 imported compact cars revealed a mean of 36.52 MPG with a standard deviation of 3.43 MPG. Samples are independent, distributions for samples are normal.

At the 95% confidence level, construct a confidence interval for the population difference between imported auto MPG and domestic auto MPG.

\[
\text{Imported compact Auto} \\
\text{MPG = Data set 1} \\
\begin{align*}
\bar{X}_1 &= 36.5185 \text{ MPG} \\
S_1 &= 3.4329 \text{ MPG} \\
S_1^2 &= \text{Var}_1 = 11.7852 \\
S_1^2/n_1 &= 11.7852/20 = 0.5893
\end{align*}
\]

\[
\text{Domestic compact Auto} \\
\text{MPG = Data set 2} \\
\begin{align*}
\bar{X}_2 &= 33.825 \text{ MPG} \\
S_2 &= 2.337 \text{ MPG} \\
S_2^2 &= \text{Var}_2 = 5.4462 \\
S_2^2/n_2 &= 5.4462/20 = 0.2723
\end{align*}
\]

\[
\text{Point Estimate of } \mu_1 - \mu_2 = \bar{X}_1 - \bar{X}_2 = 36.5185 - 33.825 = 2.6935 \text{ MPG}
\]

\[
SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{0.5893 + 0.2723} = 0.9282 \text{ MPG}
\]

\[
df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}} = \frac{(0.5893 + 0.2723)^2}{19 + 19} = 33 \text{ df}
\]

\[
\text{upper } t = T. \text{ INV}(1 - 0.95, 33) = 2.0345 = \# \text{ standard deviations}
\]

\[
\text{Margin of Error} = t \times SE = 2.0345 \times 0.9282 = 1.8884 \text{ MPG}
\]

\[
\text{Confidence Interval} = \bar{X}_1 - \bar{X}_2 \pm \text{upper } t \times SE = 2.6935 \pm 1.8884 \text{ MPG}
\]

We are 95% sure that pop. diff. between MPG for imported & domestic compact autos is between 0.81 & 4.58 MPG.
Hypothesis Testing About $M_1 - M_2$, $\sigma_1$ and $\sigma_2$ unknown

- Same 5 tests for Hypothesis Testing as ch 9
- $M_1 - M_2 = D_0 =$ Hypothesized Difference
- 2 Independent Random samples.

Test statistic for Hypothesis Testing about $M_1 - M_2$:

$$t = \frac{(\bar{x}_{1} - \bar{x}_{2}) - D_0}{\sqrt{\frac{S^2_1}{n_1} + \frac{S^2_2}{n_2}}}$$

$df = \frac{\left(\frac{S^2_1}{n_1}\right)^2}{\frac{(S^2_1)^2}{n_1-1}} + \frac{\left(\frac{S^2_2}{n_2}\right)^2}{\frac{(S^2_2)^2}{n_2-1}}$

$\text{Excel T functions}$

*Big Excel Hint 1*: Calculate $\frac{S^2}{n}$ is separate cell & refer to it in formulas with cell references.
General Notes about Hypothesis Testing:

* Build Alternative Hypothesis, $H_a$, so that you can make the desired conclusion (because we control for Type I Error).

For example:

If you want to test whether $M_1 > M_2$

then, when we accept $H_a$:

$$H_a: M_1 - M_2 > 0$$

becomes

$$M_1 > M_2$$

Notes for sample size for $t$ test:

* Whenever possible, equal sample sizes of $n_1 = n_2$ are recommended, page 448 in text.

* Even small sample sizes of $n_1 + n_2 = 20$, tend to give good results. Page 448 in text
A recent EPA study compared the highway fuel economy of domestic and imported compact cars.

- A sample of 20 domestic compact cars revealed a mean of 33.83 MPG with a standard deviation of 2.33 MPG.
- A sample of 20 imported compact cars revealed a mean of 36.52 MPG with a standard deviation of 3.43 MPG.

Assume: 1) Samples are random and independent, 2) Distributions for samples are normal.

At the .05 significance level can the EPA conclude that the MPG is higher on the imported cars?

**Step 1: List Null and Alternative Hypotheses**

- $H_0: \mu_1 - \mu_2 \geq 0$
- $H_a: \mu_1 - \mu_2 < 0$

**Step 2: Pick Level of Significance = Alpha (Risk that $H_0$ is TRUE, but we Reject it)**

$\alpha = 0.05$

**Step 3: Sample, Calculate, Draw Picture, Calculate Test Statistic**

$t = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$

$t = \frac{33.83 - 36.52}{\sqrt{\frac{2.33^2}{20} + \frac{3.43^2}{20}}} = \frac{-2.9}{0.2723 + 0.5893} = -2.9$

$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\left(\frac{s_1^2}{n_1}\right)^2 + \left(\frac{s_2^2}{n_2}\right)^2} = \frac{(0.2723 + 0.5893)^2}{19} = 3.3$

$t = -2.9$, $df = 32$

Reject $H_0$ if $t < -1.69$

Critical value: $t < -1.69$

$p-value < 0.05$

Fail to Reject $H_0$

$p-value > 0.05$
Step 4: Create Rules for Critical Value and p-value

{Critical value} = T.INV(0.05, 33) = -1.69

{P-value} = T.DIST(-2.9, 33, 1) = 0.0032

**Critical value Rule:**
If our calculated test statistic is greater than our critical value, we Reject Ho & accept Ha, otherwise we fail to Reject Ho.

**p-value Rule:**
If the p-value is less than our alpha, we Reject Ho & accept Ha, otherwise we fail to Reject Ho.

**Steps: Conclusion**
Because our calculated test statistic of -2.9 is less than our critical value of -1.69, we reject Ho & accept Ha. The sample evidence strongly suggests that the MPG is higher for Imported compact cars than it is for Domestic compact cars.
Sampling Distribution of $X_{bar1} - X_{bar2}$

$H_0: M_1 - M_2 > 0$
$H_a: M_1 - M_2 < 0$

$SE = \sqrt{\frac{s^2_1}{n_1} + \frac{s^2_2}{n_2}} = 3.5$

$p$-value: $0.0032$

$\alpha = 0.05$

Reject $H_0$ if $t < -2.365$.
Inference (Hypothesis Testing or Confidence Interval) about $M_1 - M_2$ Matched Samples

**Matched Samples:**

1. Each sampled element provides a pair of data values.

**Examples:**

1. One worker tries two different production methods and provides two different times to complete production.

2. One person gives a rating for a product (0 to 10) before watching an ad and after watching the ad.

3. Two different (but similar) workers provide times for production method 1 & production method 2.

The pair of matched sample values are then subtracted to get the difference.
Compare & contrast Matched & Independent Samples:

**Independent Sample Design:**
Samples selected from two populations in such a way that the elements making up one sample are chosen independently of the elements making up the other sample.

**Matched Sample Design:**
One simple random sample of elements is selected & 2 data values are obtained for each element (or similar elements).

Matched sample design often leads to smaller sampling error than the independent design because variation between sampled items (like difference in production workers) is significantly reduced as a source of sample error.
Key to analysis of matched sample design: we consider only column of differences.

Example of Matched Samples:

1. Established method for making Quad Boomerangs \( \times \) = Sample 1 = Method 1
2. New method for making Quad boomerangs \( \times \) = Sample 2 = Method 2
3. Simple random sample to pick 10 employees to try both methods
4. We time both methods for each worker:

<table>
<thead>
<tr>
<th>Worker</th>
<th>Method 1 time (minutes)</th>
<th>Method 2 time (minutes)</th>
<th>Difference (diff)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12</td>
<td>11</td>
<td>12-11 = -1.0</td>
</tr>
<tr>
<td>2</td>
<td>11</td>
<td>11.3</td>
<td>11-11.3 = -0.3</td>
</tr>
<tr>
<td>3</td>
<td>13</td>
<td>10.5</td>
<td>13-10.5 = 2.5</td>
</tr>
<tr>
<td>4</td>
<td>12.2</td>
<td>12</td>
<td>12.2-12 = 0.2</td>
</tr>
<tr>
<td>5</td>
<td>12</td>
<td>12</td>
<td>12-12 = 0.0</td>
</tr>
<tr>
<td>6</td>
<td>13</td>
<td>13.1</td>
<td>13-13.1 = -0.1</td>
</tr>
<tr>
<td>7</td>
<td>11.2</td>
<td>10.4</td>
<td>11.2-10.4 = 0.8</td>
</tr>
<tr>
<td>8</td>
<td>10.9</td>
<td>10.8</td>
<td>10.9-10.8 = 0.1</td>
</tr>
<tr>
<td>9</td>
<td>12.5</td>
<td>12</td>
<td>12.5-12 = 0.5</td>
</tr>
<tr>
<td>10</td>
<td>12.4</td>
<td>11.8</td>
<td>12.4-11.8 = 0.6</td>
</tr>
</tbody>
</table>

Total =
Formulas

1. \{ \text{mean of matched sample differences} \} = \bar{d} = \frac{\sum d_i}{n} \\
   \text{Excel: AVERAGE}

2. \{ \text{standard deviation of matched sample differences} \} = S_d = \sqrt{\frac{\sum (d_i - \bar{d})^2}{n-1}} \\
   \text{Excel: STDEV.S}

3. \{ \text{test statistic} \} = t = \frac{\bar{d} - \mu_d}{\frac{S_d}{\sqrt{n}}}

\bar{d} = \text{mean of matched sample diff.}

\mu_d = \text{pop mean of differences}

n = \text{sample size}

S_d = \text{standard deviation of matched sample differences}

Excel Data Ribbon Tab, Data Analysis, "t-test: paired 2 sample for means"

Excel

5. Same rules for confidence interval & hypothesis testing as earlier in class.
Hypothesis Test Matched Sample Example:

Question: Is there a difference between 2 methods?

Step 1: 
- **Ho**: $M_d = 0$
- **Ha**: $M_d \neq 0$

Step 2:
- **Alpha** = 0.05
- **Alpha/2** = 0.025

Step 3:
- $\overline{d} = \frac{\sum d_i}{n} = 0.53$
- $S_d = \sqrt{\frac{\sum (d_i - \overline{d})^2}{n-1}} = 0.8056$
- $t = \frac{\overline{d} - M_d}{\frac{S_d}{\sqrt{n}}} = 2.08$
- $df = 10 - 1 = 9$
- $p$-value = T.DIST.2T(2.08, 9) = 0.0672
- Upper t = T.INV(1 - 0.025, 9) = 2.26

Rules:
- If our calculated test statistic is greater than 2.26 or less than -2.26, we reject $H_0$ & accept $H_a$, otherwise we Fail to Reject $H_0$.
- If our $p$-value is smaller than our alpha, we Reject $H_0$ and Accept $H_a$, otherwise we Fail to Reject $H_0$.

Step 5:
Because our calculated test statistic is between -2.26 & 2.26 AND because our $p$-value is bigger than our alpha, we Fail to Reject $H_0$. → more
steps continued:

The sample evidence does not support the conclusion that the new method is faster (different from) than the old method. We do run a 5% risk of failing to reject $H_0$, even though $H_a$ was TRUE (new method was different from old method).

$$\text{Confidence Interval for Matched Samples:}$$

$$\text{Confidence Interval} = \bar{x} \pm t_{a/2} \cdot \frac{S_d}{\sqrt{n}}$$

- $\bar{x}$: Mean of differences
- $t_{a/2}$: Critical value of $t$ for $a/2$ level of significance
- $S_d$: Standard deviation of differences
- $n$: Sample size

$$= 0.53 \pm 2.26 \cdot \frac{0.8056}{\sqrt{10}}$$

$$= 0.53 \pm 0.5763 \Rightarrow -0.0463 \, \text{to} \, 1.1063$$

We are 95% sure that the population difference lies between $-0.0463 \, \text{to} \, 1.1063$ minutes difference.
Final Note about Matched Samples:

when comparing 2 population means, the matched sample procedure generally provides better precision than the independent sample approach, and therefore it is the recommended design when possible.

But what do we do if we are comparing more than 2 population means?
Completely Randomized Design (CRD) & Analysis of Variance (ANOVA)

Before we define these, let's look at an example.

Example of CRD & ANOVA

1. Boomerang manufacturer is considering 4 different methods for manufacturing Quads.
2. A random sample of 20 workers is selected.
3. The workers are randomly assigned 1 of 4 manufacturing methods.
4. Each worker is timed making the Quad with the assigned manufacturing method.
5. Time is recorded in minutes rounded to 2 decimals.
6. The mean time for each method is recorded.
7. We want to determine if there is a significant difference between the 4 mean times.
8. Collected data (times & means):

<table>
<thead>
<tr>
<th>Method 1 (Mins)</th>
<th>Method 2 (Mins)</th>
<th>Method 3 (Mins)</th>
<th>Method 4 (Mins)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.56</td>
<td>11.87</td>
<td>15.57</td>
<td>12.94</td>
</tr>
<tr>
<td>15.92</td>
<td>12.94</td>
<td>10.64</td>
<td>12.01</td>
</tr>
<tr>
<td>10.42</td>
<td>11.23</td>
<td>14.05</td>
<td>10.21</td>
</tr>
<tr>
<td>14.41</td>
<td>15.97</td>
<td>13.08</td>
<td>16.18</td>
</tr>
<tr>
<td>10.32</td>
<td>17.14</td>
<td>16.71</td>
<td>11.21</td>
</tr>
</tbody>
</table>

Xbar: 12.73, 13.83, 14.01, 11.31

5²: 6.04, 6.73, 5.49, 1.41

(sample mean)

(sample variation)
From this sample data we want to ask the question: \[ M_1 = M_2 = M_3 = M_4 \]

we have:
1. Method 1 mean time = 12.73 mins
2. Method 2 mean time = 13.83 mins
3. Method 3 mean time = 14.01 mins
4. Method 4 mean time = 11.31 mins

If we were to compare each mean, one at a time, we would have to run six tests:
(1 vs. 2), (1 vs. 3), (1 vs. 4), (2 vs. 3), (2 vs. 4), (3 vs. 4)

If we ran six tests we would build up a lot of Type 1 Error (Risk of Rejecting Ho, when it is TRUE).

Example: @ \( \alpha = 0.05 \)

\[ P(\text{All correct}) = 0.95 \times 0.95 \times 0.95 \times 0.95 \times 0.95 \times 0.95 = 0.735 \]

\[ P(\text{at least 1 NOT correct}) = 1 - 0.735 = 0.265 \]

we would also have to do many calculations if we tested all six sets of means.

so, we have to learn a different statistical method to test:
\[ M_1 = M_2 = M_3 = M_4 \]
Analysis of Variance ANOVA

- A statistical method to check whether K population means are equal $M_1 = M_2 = \ldots = M_K$

- Define Terms used in ANOVA:
  
  **Factor**: independent variable of interest
  
  **Treatment**: different populations = Different levels of Factor
  
  **Response Variable**: Dependent variable

  **Single Factor Experiment**: 1 Factor

  **Experimental unit**: Object of interest in an experiment → our example = People making Quad Booms.

<table>
<thead>
<tr>
<th>Treatment 1</th>
<th>Treatment 2</th>
<th>Treatment 3</th>
<th>Treatment 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Method 1 (mins)</strong></td>
<td><strong>Method 2 (mins)</strong></td>
<td><strong>Method 3 (mins)</strong></td>
<td><strong>Method 4 (mins)</strong></td>
</tr>
<tr>
<td>12.36</td>
<td>11.87</td>
<td>15.57</td>
<td>12.34</td>
</tr>
<tr>
<td>15.92</td>
<td>12.94</td>
<td>10.64</td>
<td>12.01</td>
</tr>
<tr>
<td>10.42</td>
<td>11.23</td>
<td>14.05</td>
<td>10.21</td>
</tr>
<tr>
<td>14.41</td>
<td>15.97</td>
<td>13.08</td>
<td>10.18</td>
</tr>
<tr>
<td>10.32</td>
<td>17.14</td>
<td>16.71</td>
<td>11.21</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sample mean Xbar</th>
<th>Sample Variance S²</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.73</td>
<td>6.04</td>
</tr>
<tr>
<td>13.83</td>
<td>6.73</td>
</tr>
<tr>
<td>14.01</td>
<td>5.49</td>
</tr>
<tr>
<td>11.31</td>
<td>1.41</td>
</tr>
</tbody>
</table>

4 pops. = 4 Treatments

Time to make Quad in mins
But we must be careful to make sure we have a random experiment so we will use:

**Completely Randomized Design (CRD)**

- An experimental design in which the treatments are randomly assigned to experimental units
- For us CRD means randomly assign manufacturing methods to Workers.

<table>
<thead>
<tr>
<th>Employees Making Boomerangs:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annette Caldwell</td>
</tr>
<tr>
<td>Sophia Wise</td>
</tr>
<tr>
<td>Florence Ferguson</td>
</tr>
<tr>
<td>Jorge Reid</td>
</tr>
<tr>
<td>Belinda Ruiz</td>
</tr>
<tr>
<td>Everett Marish</td>
</tr>
<tr>
<td>Guy Tettler</td>
</tr>
<tr>
<td>Angel Clayton</td>
</tr>
<tr>
<td>Johnnie Zimmerman</td>
</tr>
<tr>
<td>Teresa Rios</td>
</tr>
<tr>
<td>Alvin McCoy</td>
</tr>
<tr>
<td>Carole Sullivan</td>
</tr>
<tr>
<td>Winifred Ortega</td>
</tr>
<tr>
<td>Lucas Ortiz</td>
</tr>
<tr>
<td>Sharon Richardson</td>
</tr>
<tr>
<td>Martha Gutierrez</td>
</tr>
<tr>
<td>Jodi Houston</td>
</tr>
<tr>
<td>Ora Kim</td>
</tr>
<tr>
<td>Loren Freeman</td>
</tr>
<tr>
<td>Danny Mullins</td>
</tr>
<tr>
<td>Jared Young</td>
</tr>
<tr>
<td>Sylvester Hodges</td>
</tr>
<tr>
<td>Daniel Watson</td>
</tr>
<tr>
<td>Nathan Beck</td>
</tr>
<tr>
<td>Lamar Hammond</td>
</tr>
<tr>
<td>Tommie Turner</td>
</tr>
<tr>
<td>Garrett King</td>
</tr>
<tr>
<td>Bradford Boone</td>
</tr>
<tr>
<td>Julian Quinn</td>
</tr>
<tr>
<td>Phyllis Waters</td>
</tr>
<tr>
<td>Carlton Matthews</td>
</tr>
<tr>
<td>Kimberly Palmer</td>
</tr>
<tr>
<td>Arlene Baker</td>
</tr>
<tr>
<td>Jose Goodman</td>
</tr>
<tr>
<td>Floyd Erickson</td>
</tr>
<tr>
<td>Faith Salazar</td>
</tr>
<tr>
<td>Kathy Dean</td>
</tr>
<tr>
<td>Ernestine Weaver</td>
</tr>
<tr>
<td>Judy McKenzie</td>
</tr>
</tbody>
</table>

From List Of 100, randomly select 20 Workers >>

<table>
<thead>
<tr>
<th>Workers</th>
<th>Randomly Assign Manufacturing Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sophia Wise</td>
<td>Method 4</td>
</tr>
<tr>
<td>Alexandra Moren</td>
<td>Method 4</td>
</tr>
<tr>
<td>Herbert Garrett</td>
<td>Method 3</td>
</tr>
<tr>
<td>Sharon Richardson</td>
<td>Method 4</td>
</tr>
<tr>
<td>Robert Valdez</td>
<td>Method 2</td>
</tr>
<tr>
<td>Winifred Ortega</td>
<td>Method 3</td>
</tr>
<tr>
<td>Carole Sullivan</td>
<td>Method 3</td>
</tr>
<tr>
<td>Willard Payne</td>
<td>Method 1</td>
</tr>
<tr>
<td>Irvin Owens</td>
<td>Method 2</td>
</tr>
<tr>
<td>Constance Henderson</td>
<td>Method 3</td>
</tr>
<tr>
<td>Jodi Houston</td>
<td>Method 1</td>
</tr>
<tr>
<td>Danny Mullins</td>
<td>Method 4</td>
</tr>
<tr>
<td>Jaime Gibson</td>
<td>Method 2</td>
</tr>
<tr>
<td>Arlene Baker</td>
<td>Method 3</td>
</tr>
<tr>
<td>Gay Frazier</td>
<td>Method 1</td>
</tr>
<tr>
<td>Van Simmons</td>
<td>Method 4</td>
</tr>
<tr>
<td>Floyd Erickson</td>
<td>Method 1</td>
</tr>
<tr>
<td>Denise Farmer</td>
<td>Method 4</td>
</tr>
<tr>
<td>Reginald Montgomery</td>
<td>Method 3</td>
</tr>
<tr>
<td>Edwin Watkins</td>
<td>Method 2</td>
</tr>
</tbody>
</table>

Time and record data >>>

<table>
<thead>
<tr>
<th>Workers</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sophia Wise</td>
<td>Method 4</td>
</tr>
<tr>
<td>Alexandra Moren</td>
<td>Method 4</td>
</tr>
<tr>
<td>Herbert Garrett</td>
<td>Method 3</td>
</tr>
<tr>
<td>Sharon Richardson</td>
<td>Method 4</td>
</tr>
<tr>
<td>Robert Valdez</td>
<td>Method 2</td>
</tr>
<tr>
<td>Winifred Ortega</td>
<td>Method 3</td>
</tr>
<tr>
<td>Carole Sullivan</td>
<td>Method 3</td>
</tr>
<tr>
<td>Willard Payne</td>
<td>Method 1</td>
</tr>
<tr>
<td>Irvin Owens</td>
<td>Method 2</td>
</tr>
<tr>
<td>Constance Henderson</td>
<td>Method 3</td>
</tr>
<tr>
<td>Jodi Houston</td>
<td>Method 1</td>
</tr>
<tr>
<td>Danny Mullins</td>
<td>Method 4</td>
</tr>
<tr>
<td>Jaime Gibson</td>
<td>Method 2</td>
</tr>
<tr>
<td>Arlene Baker</td>
<td>Method 3</td>
</tr>
<tr>
<td>Gay Frazier</td>
<td>Method 1</td>
</tr>
<tr>
<td>Van Simmons</td>
<td>Method 1</td>
</tr>
<tr>
<td>Floyd Erickson</td>
<td>Method 1</td>
</tr>
<tr>
<td>Denise Farmer</td>
<td>Method 4</td>
</tr>
<tr>
<td>Reginald Montgomery</td>
<td>Method 3</td>
</tr>
<tr>
<td>Edwin Watkins</td>
<td>Method 2</td>
</tr>
</tbody>
</table>

### Experimental units = Workers who make Quads

### Treatments = populations = methods to make Quad 1, 2, 3, 4.

60 more in list below

---

P.30
Before we use ANOVA to test, we must check what assumptions are necessary:

1. Assumptions for ANOVA (Analysis of Variance):

1. The population for the response variables (time to make quad) follow normal distribution.
2. The variance of response variables ($\sigma^2$) are the same for all populations.
3. Observations must be independent (not related or linked or influenced by other observations).

Conceptual overview of ANOVA:

Sample means are "close to each other" because there is only one sampling distribution when $H_0$ is TRUE.

Sampling means come from different sampling distributions and are not close together when $H_0$ is FALSE.
Conceptual overview of ANOVA continued:

1. with ANOVA we will develop 2 independent estimates of the common population variance & then compare them with division

   △ MSTR = estimate of variability between the samples
   
   MSE = estimate of variability within each sample

2. 

   △ \frac{MSTR}{MSE} too big we don't think means are equal.
   
   \frac{MSTR}{MSE} close to 1 then reasonable to assume means equal.

   (F critical value will tell us how close to 1 we need to be.)

6. Hypothesis Testing with ANOVA

   \[ H_0 : M_1 = M_2 = M_3 = M_4 \]
   
   \[ H_a : "Not all population means are equal" \]

   → If H₀ rejected, we cannot conclude that all means are different, only that "at least 2 pop. means are different"
   
   → ANOVA is procedure to determine if we reject H₀.
F Distribution used in ANOVA table

1. Family of F Distributions (many F Distributions)
2. \( df \) in numerator = 1st estimate of variance
   \( df \) in denominator = 2nd estimate of variance
3. F Distribution is continuous from 0 to \( \infty \)
4. F Distribution cannot be negative (0 1st number)
5. Positive skew (long tail to right, as both \( df \) increase, approach normal)
6. Used when you want to test whether two samples are from populations having equal variances.
7. Used when calculating ANOVA table
8. For F Critical value in Excel use: \( =F.INV.RT(\alpha, \text{df}_1, \text{df}_2) \)
9. For F Hypothesis Test in Excel use: \( =F.DIST.RT(MSTR, \text{df}_1, MSE, \text{df}_2) \)
8) How to do ANOVA in Excel
1. $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$
2. $H_a = \text{Not all population means are the same (at least two are different)}$
3. Decision Rule: $p\text{-value} < \alpha$, Reject $H_0$ & Accept $H_a$ or $\text{Test Statistic } F > \text{Critical } F$,
4. Reject $H_0$ & Accept $H_a$, otherwise we fail to Reject $H_0$

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population 1</td>
<td>Population 2</td>
<td>Population 3</td>
<td>Population 4</td>
<td></td>
</tr>
<tr>
<td>Treatment 1 Method 1</td>
<td>Treatment 2 Method 1</td>
<td>Treatment 3 Method 1</td>
<td>Treatment 4 Method 1</td>
<td></td>
</tr>
<tr>
<td>12.56</td>
<td>11.87</td>
<td>15.57</td>
<td>12.94</td>
<td></td>
</tr>
<tr>
<td>15.92</td>
<td>12.94</td>
<td>10.64</td>
<td>12.01</td>
<td></td>
</tr>
<tr>
<td>10.42</td>
<td>11.23</td>
<td>14.05</td>
<td>10.21</td>
<td></td>
</tr>
<tr>
<td>14.41</td>
<td>15.97</td>
<td>13.08</td>
<td>10.18</td>
<td></td>
</tr>
<tr>
<td>10.32</td>
<td>17.14</td>
<td>16.71</td>
<td>11.21</td>
<td></td>
</tr>
<tr>
<td>Sample Size $= n =$</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Sample Size $= \bar{X}_n =$</td>
<td>12.726</td>
<td>13.83</td>
<td>14.01</td>
<td>11.31</td>
</tr>
<tr>
<td>Sample Variance $= S_2 =$</td>
<td>6.04288</td>
<td>6.73235</td>
<td>5.48675</td>
<td>1.41095</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

| Grand Over All Count $= n_T =$ | 20 |
| Grand Over All Sample Mean $= \bar{X}_{all} =$ | 12.969 |
| # Treatments $= k =$ | 4 |
| Calculations for Total Variance for All Observations $= |
| Sum of Squares Total $= SST =$ | 101.87338 |
| $df_T =$ | 19 |
| Calculations for Variance Between Treatments $= |
| Sum of Squares of Treatments $= SSTR =$ | 23.18166 |
| $df_{TR} =$ | 3 |
| Estimate of Variance Between Treatments $= MSTR =$ | \( \frac{SSTR}{df_{TR}} \) |
| 7.72722 |
| Calculations for Variance Within Treatments $= |
| Sum of Squares Error $= SSE =$ | 78.69172 |
| $df_{E} =$ | 16 |
| Estimate of Variance Within Treatments $= MSE =$ | \( \frac{SSE}{df_{E}} \) |
| 4.9182325 |
| F Test Statistic $= MSTR/MSE =$ | 1.571137599 |
| Alpha $= 0.05$ |
| p-value $= 0.235295845$ |
| F Critical $= 3.238871517$ |
| 35 |
| 36 Because the F Calculated Test Statistic is not past our F Critical, We Fail To Reject $H_0$. |
| 37 It seems reasonable to assume that the pop means are the same. |

---

1. Quad Example Final Results
<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>Population 1</td>
<td>Population 2</td>
<td>Population 3</td>
<td>Population 4</td>
</tr>
<tr>
<td>7</td>
<td>Treatment 1</td>
<td>Treatment 2</td>
<td>Treatment 3</td>
<td>Treatment 4</td>
</tr>
<tr>
<td>8</td>
<td>Method 1</td>
<td>Method 2</td>
<td>Method 3</td>
<td>Method 4</td>
</tr>
<tr>
<td>9</td>
<td>12.56</td>
<td>11.87</td>
<td>15.57</td>
<td>12.94</td>
</tr>
<tr>
<td>10</td>
<td>15.92</td>
<td>12.94</td>
<td>10.64</td>
<td>12.01</td>
</tr>
<tr>
<td>11</td>
<td>10.42</td>
<td>11.23</td>
<td>14.05</td>
<td>10.21</td>
</tr>
<tr>
<td>12</td>
<td>14.41</td>
<td>15.97</td>
<td>13.08</td>
<td>10.18</td>
</tr>
<tr>
<td>13</td>
<td>Sample Size = n =</td>
<td>=COUNT(B8:B12)</td>
<td>=COUNT(B8:C12)</td>
<td>=COUNT(B8:B12)</td>
</tr>
<tr>
<td>14</td>
<td>Sample Size = Xbar =</td>
<td>=AVERAGE(B8:B12)</td>
<td>=AVERAGE(B8:C12)</td>
<td>=AVERAGE(B8:B12)</td>
</tr>
<tr>
<td>15</td>
<td>Sample Variance = s² =</td>
<td>=VAR.S(B8:B12)</td>
<td>=VAR.S(B8:C12)</td>
<td>=VAR.S(B8:B12)</td>
</tr>
<tr>
<td>16</td>
<td>Grand Over All Count = n =</td>
<td>=COUNT(B8:E12)</td>
<td>=COUNT(B8:D12)</td>
<td>=COUNT(E8:E12)</td>
</tr>
<tr>
<td>17</td>
<td>Grand Over All Sample Mean = XbarG =</td>
<td>=AVERAGE(B8:E12)</td>
<td>=AVERAGE(B8:D12)</td>
<td>=AVERAGE(E8:E12)</td>
</tr>
<tr>
<td>18</td>
<td># Treatments = k =</td>
<td>=COUNTA(B7:E7)</td>
<td>=COUNTA(B7:D12)</td>
<td>=COUNTA(E7:E12)</td>
</tr>
<tr>
<td>19</td>
<td>Calculations for Total Variance for All Observations =</td>
<td>=SUMPRODUCT(((B8:E12-B14:E14)^2))</td>
<td>=SUMPRODUCT(((B14:E14-B14:E14)^2))</td>
<td>=SUMPRODUCT(((B8:E12-B8:E12)^2))</td>
</tr>
<tr>
<td>20</td>
<td>df₁ =</td>
<td>=B17-1</td>
<td>=B19-1</td>
<td>=B24/B25</td>
</tr>
<tr>
<td>21</td>
<td>Calculations for Variance Between Treatments =</td>
<td>=SUMPRODUCT((B14:E14-B14:E14)^2)</td>
<td>=SUMPRODUCT((B8:E12-B8:E12)^2)</td>
<td>=SUMPRODUCT((B8:E12-B8:E12)^2)</td>
</tr>
<tr>
<td>22</td>
<td>Estimates of Variance Between Treatments = MSTR = SSTR/df₁ =</td>
<td>=B24/B25</td>
<td>=B18/B19</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Calculations for Variance Within Treatments =</td>
<td>=SUMPRODUCT((B15:E15,B13:E13-1))</td>
<td>=SUMPRODUCT((B8:E12-B14:E14)^2)</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Sum of Squares Error = SSE =</td>
<td>=B26/B30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>df₁ =</td>
<td>=B17-1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Estimates of Variance Within Treatments = MSE = SSE/df₁ =</td>
<td>=B28/B29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>F Test Statistic = MSTR/MSE =</td>
<td>=B26/B30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Alpha =</td>
<td>0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>p-value =</td>
<td>=F.DIST.RT(B21,B25,B29)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>F Critical =</td>
<td>=F.INV.RT(B32,B25,B29)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Check: =SUMPRODUCT(B13:E13,B14:E14)/SUM(B13:E13)
H₀: μ₁ = μ₂ = μ₃ = μ₄
Ha = Not all population means are the same (at least two are different)

Decision Rule: p-value < alpha, Reject H₀ & Accept Ha or Test Statistic F > Critical F, Reject H₀ & Accept Ha, otherwise we fail to Reject H₀

<table>
<thead>
<tr>
<th>Population 1</th>
<th>Population 1</th>
<th>Population 3</th>
<th>Population 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment 1</td>
<td>Treatment 2</td>
<td>Treatment 3</td>
<td>Treatment 4</td>
</tr>
<tr>
<td>Method 1</td>
<td>Method 2</td>
<td>Method 3</td>
<td>Method 4</td>
</tr>
<tr>
<td>12.56</td>
<td>11.87</td>
<td>15.57</td>
<td>12.94</td>
</tr>
<tr>
<td>15.92</td>
<td>12.94</td>
<td>10.64</td>
<td>12.01</td>
</tr>
<tr>
<td>10.42</td>
<td>11.23</td>
<td>14.05</td>
<td>10.21</td>
</tr>
<tr>
<td>14.41</td>
<td>15.97</td>
<td>13.08</td>
<td>10.18</td>
</tr>
<tr>
<td>10.32</td>
<td>17.14</td>
<td>16.71</td>
<td>11.21</td>
</tr>
</tbody>
</table>

Because our p-value is larger than, We Fail To Reject H₀. It seems reasonable to assume that the pop means are the same.
Analysis of Variance to test equality of K population means for completely randomized design. Book formulas:

1. \( H_0: \mu_1 = \mu_2 = \ldots = \mu_K \)
   \( H_a: \) Not all population means are equal

2. Variables:
   - \( j = \) particular treatment "jth treatment"
   - \( \mu_j = \) pop. mean of jth population
   - \( K = \) # of treatments/populations
   - \( n_j = \) Random sample size for jth population
   - \( X_{ij} = \) Value of observation i for treatment j
   - \( \bar{X}_j = \) Sample mean for treatment j
   - \( S_j^2 = \) Sample Variance for treatment j
   - \( S_j = \) Sample Standard Deviation for treatment j
   - \( \bar{X} = \) Overall sample mean

3. Formulas:
   \[ \bar{X}_j = \frac{\sum_{i=1}^{n_j} X_{ij}}{n_j} \]
   \[ S_j^2 = \frac{\sum_{i=1}^{n_j} (X_{ij} - \bar{X}_j)^2}{n_j - 1} \]

\[ \bar{X} = \frac{\sum_{j=1}^{k} \sum_{i=1}^{n_j} X_{ij}}{n_T} \]

or if all \( n_j \) are same:

\[ \bar{X} = \frac{\sum_{j=1}^{k} \bar{X}_j}{k} \]

\(* \bar{X} \rightarrow i.e \) sample means provided sample size not same...

\[ \bar{X} = \frac{n_1 \bar{X}_1 + n_2 \bar{X}_2 + \ldots + n_k \bar{X}_k}{n_T} \]
3. Formulas continued:

Between - Treatments Estimate of population Variance \( (\sigma^2) \) (Sample sizes equal)

\[
\text{Variability Between (measured Among Samples)} = \text{Mean Square due to Treatments} = MSTR = \frac{\sum_{j=1}^{K} n_j (\bar{x}_j - \bar{x})^2}{K - 1}
\]

4. \( SS_{T} = \{ \text{Sum of squares due to Treatments} \} \)

5. \( \{ \text{Degrees of Freedom Associated with SS}_{T} \} = K - 1 \)

6. \( \{ \text{Mean Square due to Treatments} \} = MSTR = \frac{SS_{T}}{K - 1} \)

**If Ho is TRUE \( (M_1 = M_2 = \ldots = M_K) \)**

**MSTR is unbiased estimate of \( \sigma^2 \)**

**If Ha is TRUE (Not all pop. means are equal)***

**1. MSTR is NOT unbiased estimate of \( \sigma^2 \)**

**2. MSTR will overestimate \( \sigma^2 \)**

\*if each sample size is same:

\[
MSTR = \frac{\sum_{j=1}^{K} n_j (\bar{x}_j - \bar{x})^2}{K - 1} = \eta \left[ \frac{\sum_{j=1}^{K} (\bar{x}_j - \bar{x})^2}{K - 1} \right] = n \bar{S}_x^2
\]
Within Treatments Estimate of Pop. Variance ($\sigma^2$) (sample sizes equal)

\[
\text{Variability within each sample} = \text{mean square due to Error} = M SE = \frac{1}{n_T - K} \sum_{j=1}^{K} (n_j - 1) S_j^2
\]

7. \[
\text{Sum of squares due to Error} = SSE = \sum_{j=1}^{K} \sum_{i=1}^{n_j} (x_{ij} - \bar{X}_j)^2
\]

8. \[
\text{Degrees of Freedom Associated w/ SSE} = n_T - K
\]

9. \[
\text{Mean square due to Error} = M SE = \frac{SSE}{n_T - K}
\]

Note 1: M SE is based on the variation within each of the treatments.

Note 2: It is not influenced by whether H_0 is TRUE.

Note 3: Therefore, M SE always provides an unbiased estimate of common pop. Variance $\sigma^2$.

*If sample size is same*:

Thus:

\[
M SE = \frac{1}{K(n-1)} \sum_{j=1}^{K} (n_j - 1) S_j^2 = \frac{1}{K(n-1)} \sum_{j=1}^{K} S_j^2 = \frac{1}{K} \sum_{j=1}^{K} S_j^2 = \{\text{Average of K sample variances}\}
\]

then: $n_T = Kn$ & $n_T - k = K(n-1)$
Comparing common Pop. Variance $\sigma^2$

Estimates $\text{MSTR}$ & $\text{MSE}$ with $F$ Test:

1. If $H_0$ TRUE ($\mu_1=\mu_2=\ldots=\mu_k$) $\text{MSTR}$ & $\text{MSE}$ provide $2$ independent, unbiased estimates of $\sigma^2$

2. If three assumptions for ANOVA (each response variable has normal distribution, $\sigma^2$ for response variable is same for all pops., observations, $x_{ij}$ are independent) are valid:
   
   1) Sampling distribution of $\frac{\text{MSTR}}{\text{MSE}}$ is an $F$ distribution
   
   2) Numerator $(\text{MSTR})$ $dF = k-1$
   
   3) Denominator $(\text{MSE})$ $dF = n_k - k$
   
   4) If $H_0$ TRUE, $\frac{\text{MSTR}}{\text{MSE}}$ should appear to have been selected from $F$ distribution
   
   4) If $H_0$ FALSE, $\text{MSTR}$ overestimates $\sigma^2$ & $\frac{\text{MSTR}}{\text{MSE}}$ will be large

Test statistic for equality of $K$ population means

$$F = \frac{\text{MSTR}}{\text{MSE}}$$

Total sum of squares $TSS = \text{SSST} + \text{SSSE}$
3 Formulas continued:

\[ \{ \text{Total sum of squares} \} = \text{SST} = \text{SSTR} + \text{SSE} = \]
\[ = \sum_{j=1}^{K} \sum_{i=1}^{n_j} (x_{ij} - \bar{x})^2 = \]

If we took \( \frac{\text{SST}}{n_T-1} \), we would get overall sample variance for all observations.

**ANOVA Table for completely randomized design:**

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Square</th>
<th>F</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatments</td>
<td>SSTR</td>
<td>K - 1</td>
<td>MSTR = ( \frac{\text{SSTR}}{K-1} )</td>
<td>( \frac{\text{MSTR}}{\text{MSE}} )</td>
<td>( F_{\text{critical}} )</td>
</tr>
<tr>
<td>Error</td>
<td>SSE</td>
<td>n_T - K</td>
<td>MSE = ( \frac{\text{SSE}}{n_T-K} )</td>
<td>( \frac{\text{MSE}}{n_T-K} )</td>
<td>( F_{\text{critical}} )</td>
</tr>
<tr>
<td>Total</td>
<td>SST</td>
<td>n_T - 1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ P\text{-value} = P \left( F \geq \frac{\text{MSTR}}{\text{MSE}} \right) \] if \( p\text{-value} \leq \alpha \)

\[ \text{Critical value} = F_{\alpha} \] if \( F \geq F_{\alpha} \)
Testing for Equality of K Pop. Means for an **observational Study**

Same procedures for experimental studies, but we don't have as much control.
H_0: \mu_1 - \mu_2 = 0
H_a: \mu_1 - \mu_2 \neq 0

Sample distribution of \(X_{bar1} - X_{bar2}\)

Standard deviation of \(X_{bar1} - X_{bar2}\) = \(\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\)

Test statistic = \(Z = \frac{X_{bar1} - X_{bar2}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\)

P-value = \(P(Z > 1.81) = 0.0352\)

Decision: Reject \(H_0\)

\(\alpha = 0.05\)

\(0.0352 \leq 0.05\)

Fail to reject \(H_0\)
Sampling Distribution of $X_{\bar{r}} - X_{\bar{w}}$

$H_0 : \mu_{\bar{r}} - \mu_{\bar{w}} = 0$

$H_a : \mu_{\bar{r}} - \mu_{\bar{w}} \neq 0$

Critical values:
- $z_{1/2} = 0.025$
- $z_{1/2} = 0.025$

$P$-value:

\[ P \left( X_{\bar{r}} - X_{\bar{w}} \geq 2.46 \right) = 0.0138 \]

Test statistic:

\[ \bar{x}_{\bar{r}} - \bar{x}_{\bar{w}} = 2.46 \]
Sample: Distribution of Differences

$\text{t-distribution}$

$\overset{\text{Reject } H_0}{\rightarrow}$

$\overset{\text{Accept } H_0}{\leftarrow}$

$\text{Test Statistic} = -2.4$

$p$-value: $p(t \leq -2.4) = 0.0691$

Critical Value: $t = 1.86$

Fail to Reject $H_0$

$\sqrt{n} \cdot (\bar{x} - \mu) = 2$

$\text{New - New}_2$

$\text{New - New}_1$

$\overline{x}_1 - \overline{x}_2$

$\mu_1 - \mu_2 \\ \theta_1 - \theta_2$

$\text{HW} 14$
Sampling Distribution of Differences $X_{bar1} - X_{bar2}$

$t$ Distribution

fails to reject $H_0$

Reject $H_0$
Accept $H_A$

$\alpha = 0.05$

$p$-value = $P(t \geq 1.81) = 0.04167$

$test statistic = 1.81$

Critical Value $t = 1.71$
H₀ : \( M₁ - M₂ = 0 \)
Hₐ : \( M₁ - M₂ < 0 \)

Sampling Distribution of \( \bar{X}_{bar} \), \(- \bar{X}_{bar₂}\)

Test statistic \( = 0.32 \)

P-value = \( P(\text{t} \leq -0.72 \text{ or } \text{t} \geq 0.72) = 0.7535 \)

Critical values \( \pm 2.02 \)
Fail to Reject $H_0$

$\alpha = 0.05$

Reject $H_0$

Accept $H_a$

Critical value

1.895

t = 1.357

$p$-value = $P(Z \geq 1.357) = 0.1084$
Test statistic: \( t = -1.42 \)

\[ \text{p-value: } p(-1.42 < t < 1.73) = 0.173 \]