Class 26

Daniel B. Rowe, Ph.D.

Department of Mathematics, Statistics, and Computer Science
Agenda:

Recap Chapter 9.3 and 9.4

Lecture Chapter 10.1-10.3

Go over Exam 6
Recap Chapter 9.3 and 9.4
9: Inferences Involving One Population

9.2 Inference about the Binomial Probability of Success

We talked about a Binomial experiment with two outcomes.

\[ P(x) = \frac{n!}{x!(n-x)!} p^x (1 - p)^{n-x} \]

\( n = 1, 2, 3, \ldots \)

\( x = 0, 1, \ldots, n \)

\( 0 \leq p \leq 1 \)

\( n = \# \) of trials, \( x = \# \) of successes, \( p = \) prob. of success

Sample Binomial Probability

\[ p' = \frac{x}{n} \]

i.e. number of H out of n flips

(9.3)

where \( x \) is the number of successes in \( n \) trials.
9: Inferences Involving One Population
9.2 Inference about the Binomial Probability of Success

In Statistics, $\text{mean}(cx) = c \mu$ and $\text{variance}(cx) = c^2 \sigma^2$.

With $p' = \frac{x}{n}$, the constant is $c = \frac{1}{n}$, and

\[
\text{mean} \left( \frac{x}{n} \right) = \left( \frac{1}{n} \right) \text{mean}(x) = \left( \frac{1}{n} \right) np = p = \mu_{p'}
\]

and the variance of $p' = \frac{x}{n}$ is variance

\[
\text{variance} \left( \frac{x}{n} \right) = \frac{\sigma^2}{n^2} = \frac{p(1-p)}{n}
\]

standard error of $p' = \frac{x}{n}$ is

\[
\sigma_{p'} = \sqrt{\frac{\sigma^2}{n^2}} = \sqrt{\frac{p(1-p)}{n}}.
\]
9: Inferences Involving One Population
9.2 Inference about the Binomial Probability of Success

That is where 1. and 2. in the green box below come from.

If a random sample of size $n$ is selected from a large population with $p = P(\text{success})$, then the sampling distribution of $p'$ has:

1. A mean $\mu_p$, equal to $p$

2. A standard error $\sigma_p$, equal to $\sqrt{\frac{p(1-p)}{n}}$

3. An approximately normal distribution if $n$ is sufficiently “large.”
For a confidence interval, we would use

**Confidence Interval for a Proportion**

\[
p' - z(\alpha / 2)\sqrt{\frac{p'q'}{n}} \quad \text{to} \quad p' + z(\alpha / 2)\sqrt{\frac{p'q'}{n}}
\]

(9.6)

where \( p' = \frac{x}{n} \) and \( q' = (1 - p') \).

Since we didn’t know the true value for \( p \), we estimate it by \( p' \).

This is of the form \( \text{point estimate} \pm \text{some amount} \).
9: Inferences Involving One Population
9.2 Inference about the Binomial Probability of Success
Determining the Sample Size

Using the error part of the CI, we determine the sample size \( n \).

Maximum Error of Estimate for a Proportion

\[
E = z(\alpha / 2) \sqrt{\frac{p'(1 - p')}{n}}
\]  

(9.7)

Sample Size for 1- \( \alpha \) Confidence Interval of \( p \)

\[
n = \frac{[z(\alpha / 2)]^2 p^* (1 - p^*)}{E^2}
\]

(9.8)

From prior data, experience, gut feelings, séance. Or use 1/2.

where \( p^* \) and \( q^* \) are provisional values used for planning.
9: Inferences Involving One Population
9.2 Inference about the Binomial Probability of Success

Hypothesis Testing Procedure

We can perform hypothesis tests on the proportion

\[ H_0: p \geq p_0 \text{ vs. } H_a: p < p_0 \]
\[ H_0: p \leq p_0 \text{ vs. } H_a: p > p_0 \]
\[ H_0: p = p_0 \text{ vs. } H_a: p \neq p_0 \]

**Test Statistic for a Proportion \( p \)**

\[
z^* = \frac{p' - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}} \quad \text{with} \quad p' = \frac{x}{n}
\]

(9.9)
9: Inferences Involving One Population
9.3 Inference about the Variance and Standard Deviation

We can perform hypothesis tests on the variance.

\[ H_0: \sigma^2 \geq \sigma_0^2 \text{ vs. } H_a: \sigma^2 < \sigma_0^2 \]
\[ H_0: \sigma^2 \leq \sigma_0^2 \text{ vs. } H_a: \sigma^2 > \sigma_0^2 \]
\[ H_0: \sigma^2 = \sigma_0^2 \text{ vs. } H_a: \sigma^2 \neq \sigma_0^2 \]

For this hypothesis test, use the \( \chi^2 \) distribution

ignore \( \mu = df \)
\[ \sigma^2 = 2df \]

1. \( \chi^2 \) is nonnegative
2. \( \chi^2 \) is not symmetric, skewed to right
3. \( \chi^2 \) is distributed to form a family each determined by \( df = n-1 \).

Figure from Johnson & Kuby, 2012.
9: Inferences Involving One Population

9.4 Inference about the Variance and Standard Deviation

Test Statistic for Variance (and Standard Deviation)

\[ \chi^2* = \frac{(n-1)s^2}{\sigma^2_0}, \]  
with \( df=n-1. \)  

(9.10)

Will also need critical values.

\[ P\left( \chi^2 > \chi^2(df, \alpha) \right) = \alpha \]

Table 8
Appendix B
Page 721

Figure from Johnson & Kuby, 2012.
9: Inferences Involving One Pop.

**Example:** Find $\chi^2(20, 0.05)$.

Table 8, Appendix B, Page 721.

<table>
<thead>
<tr>
<th>df</th>
<th>0.005</th>
<th>0.01</th>
<th>0.025</th>
<th>0.05</th>
<th>0.10</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.90</th>
<th>0.95</th>
<th>0.975</th>
<th>0.99</th>
<th>0.995</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00000393</td>
<td>0.000157</td>
<td>0.000982</td>
<td>0.00393</td>
<td>0.0158</td>
<td>0.102</td>
<td>0.455</td>
<td>1.32</td>
<td>2.71</td>
<td>3.84</td>
<td>5.02</td>
<td>6.63</td>
<td>7.88</td>
</tr>
<tr>
<td>2</td>
<td>0.0100</td>
<td>0.0201</td>
<td>0.0506</td>
<td>0.103</td>
<td>0.211</td>
<td>0.575</td>
<td>1.39</td>
<td>2.77</td>
<td>4.61</td>
<td>5.99</td>
<td>7.38</td>
<td>9.21</td>
<td>10.6</td>
</tr>
<tr>
<td>3</td>
<td>0.0717</td>
<td>0.115</td>
<td>0.216</td>
<td>0.352</td>
<td>0.584</td>
<td>1.21</td>
<td>2.37</td>
<td>4.11</td>
<td>6.25</td>
<td>7.81</td>
<td>9.35</td>
<td>11.3</td>
<td>12.8</td>
</tr>
<tr>
<td>4</td>
<td>0.2077</td>
<td>0.297</td>
<td>0.484</td>
<td>0.711</td>
<td>1.06</td>
<td>1.92</td>
<td>3.35</td>
<td>5.39</td>
<td>7.78</td>
<td>9.49</td>
<td>11.1</td>
<td>13.3</td>
<td>14.9</td>
</tr>
<tr>
<td>5</td>
<td>0.412</td>
<td>0.554</td>
<td>0.831</td>
<td>1.15</td>
<td>1.61</td>
<td>2.67</td>
<td>4.35</td>
<td>6.63</td>
<td>9.24</td>
<td>11.1</td>
<td>12.8</td>
<td>15.1</td>
<td>16.7</td>
</tr>
<tr>
<td>6</td>
<td>0.676</td>
<td>0.872</td>
<td>1.24</td>
<td>1.64</td>
<td>2.20</td>
<td>3.45</td>
<td>5.35</td>
<td>7.84</td>
<td>10.6</td>
<td>12.6</td>
<td>14.4</td>
<td>16.8</td>
<td>18.5</td>
</tr>
<tr>
<td>7</td>
<td>0.989</td>
<td>1.24</td>
<td>1.69</td>
<td>2.17</td>
<td>2.83</td>
<td>4.25</td>
<td>6.35</td>
<td>9.04</td>
<td>12.0</td>
<td>14.1</td>
<td>16.0</td>
<td>18.5</td>
<td>20.3</td>
</tr>
<tr>
<td>8</td>
<td>1.34</td>
<td>1.65</td>
<td>2.18</td>
<td>2.73</td>
<td>3.49</td>
<td>5.07</td>
<td>7.34</td>
<td>10.2</td>
<td>13.4</td>
<td>15.5</td>
<td>17.5</td>
<td>20.1</td>
<td>22.0</td>
</tr>
<tr>
<td>9</td>
<td>1.73</td>
<td>2.09</td>
<td>2.70</td>
<td>3.33</td>
<td>4.17</td>
<td>5.90</td>
<td>8.34</td>
<td>11.4</td>
<td>14.7</td>
<td>16.9</td>
<td>19.0</td>
<td>21.7</td>
<td>23.6</td>
</tr>
<tr>
<td>10</td>
<td>2.16</td>
<td>2.56</td>
<td>3.25</td>
<td>3.94</td>
<td>4.87</td>
<td>6.74</td>
<td>9.34</td>
<td>12.5</td>
<td>16.0</td>
<td>18.3</td>
<td>20.5</td>
<td>23.2</td>
<td>25.2</td>
</tr>
<tr>
<td>11</td>
<td>2.60</td>
<td>3.05</td>
<td>3.82</td>
<td>4.57</td>
<td>5.58</td>
<td>7.58</td>
<td>10.34</td>
<td>13.7</td>
<td>17.3</td>
<td>19.7</td>
<td>21.9</td>
<td>24.7</td>
<td>26.8</td>
</tr>
<tr>
<td>12</td>
<td>3.07</td>
<td>3.57</td>
<td>4.40</td>
<td>5.23</td>
<td>6.30</td>
<td>8.44</td>
<td>11.34</td>
<td>14.8</td>
<td>18.5</td>
<td>21.0</td>
<td>23.3</td>
<td>26.2</td>
<td>28.3</td>
</tr>
<tr>
<td>13</td>
<td>3.57</td>
<td>4.11</td>
<td>5.01</td>
<td>5.89</td>
<td>7.04</td>
<td>9.30</td>
<td>12.34</td>
<td>16.0</td>
<td>19.8</td>
<td>22.4</td>
<td>24.7</td>
<td>27.7</td>
<td>29.8</td>
</tr>
<tr>
<td>14</td>
<td>4.07</td>
<td>4.66</td>
<td>5.63</td>
<td>6.57</td>
<td>7.79</td>
<td>10.2</td>
<td>13.34</td>
<td>17.1</td>
<td>21.1</td>
<td>23.7</td>
<td>26.1</td>
<td>29.1</td>
<td>31.3</td>
</tr>
<tr>
<td>15</td>
<td>4.60</td>
<td>5.23</td>
<td>6.26</td>
<td>7.26</td>
<td>8.55</td>
<td>11.0</td>
<td>14.34</td>
<td>18.2</td>
<td>22.3</td>
<td>25.0</td>
<td>27.5</td>
<td>30.6</td>
<td>32.8</td>
</tr>
<tr>
<td>16</td>
<td>5.14</td>
<td>5.81</td>
<td>6.91</td>
<td>7.96</td>
<td>9.31</td>
<td>11.9</td>
<td>15.34</td>
<td>19.4</td>
<td>23.5</td>
<td>26.3</td>
<td>28.8</td>
<td>32.0</td>
<td>34.3</td>
</tr>
<tr>
<td>17</td>
<td>5.70</td>
<td>6.41</td>
<td>7.56</td>
<td>8.67</td>
<td>10.1</td>
<td>12.8</td>
<td>16.34</td>
<td>20.3</td>
<td>24.8</td>
<td>27.6</td>
<td>30.2</td>
<td>33.4</td>
<td>35.7</td>
</tr>
<tr>
<td>18</td>
<td>6.26</td>
<td>7.01</td>
<td>8.23</td>
<td>9.39</td>
<td>10.9</td>
<td>13.7</td>
<td>17.34</td>
<td>21.6</td>
<td>26.0</td>
<td>28.9</td>
<td>31.5</td>
<td>34.8</td>
<td>37.2</td>
</tr>
<tr>
<td>19</td>
<td>6.84</td>
<td>7.63</td>
<td>8.91</td>
<td>10.1</td>
<td>11.7</td>
<td>14.6</td>
<td>18.34</td>
<td>22.7</td>
<td>27.2</td>
<td>30.1</td>
<td>32.9</td>
<td>36.2</td>
<td>38.6</td>
</tr>
<tr>
<td>20</td>
<td>7.43</td>
<td>8.26</td>
<td>9.59</td>
<td>10.9</td>
<td>12.4</td>
<td>15.5</td>
<td>19.34</td>
<td>23.8</td>
<td>28.4</td>
<td>31.4</td>
<td>34.2</td>
<td>37.6</td>
<td>40.0</td>
</tr>
</tbody>
</table>

Figures from Johnson & Kuby, 2012.

Rowe, D.B.
Chapter 9: Inferences Involving One Population

Questions?

Homework: Chapter 9 # 5, 7, 9, 21, 23, 27, 28, 35, 37, 43, 49, 55, 71, 85, 89, 91, 93, 95, 97, 99, 105, 109, 119, 121, 129, 131, 135, 139, 145
Lecture Chapter 10
Chapter 10: Inference Involving Two Populations

Daniel B. Rowe, Ph.D.

Department of Mathematics, Statistics, and Computer Science
10: Inferences Involving Two Populations
10.1 Dependent and Independent Samples

In this chapter we will have samples from two populations.

The two populations can either be dependent or independent.

**Dependent Samples:** If samples have related pairs.
Random sample of married couples.
Male Height vs. Female Height

**Independent Samples:** If samples are unrelated.
Random sample of males, Random Sample of females.
Male Height vs. Female Height
When we have dependent samples, there is a commonality between the two items in the pair. Quite often before and after.

Population 1: \( \mu_1 = \mu_c + \mu_{\text{before}} \)

Population 2: \( \mu_2 = \mu_c + \mu_{\text{after}} \)

But we’re interested in the difference in means:

\[
\mu_1 - \mu_2 = (\mu_c + \mu_{\text{before}}) - (\mu_c + \mu_{\text{after}})
= \mu_{\text{before}} - \mu_{\text{after}}
\]
10: Inferences Involving Two Populations
10.2 Inference for Mean Difference Two Dependent Samples

We form a paired difference from the data

\[
d = x_1 - x_2
\]  

(10.1)

This means that we are subtracting the sample value from population 2 from the sample value from population 1.
Imagine that we have paired data \((x_{1,1}, x_{2,1}), \ldots, (x_{1,n}, x_{2,n})\)
\(x_{j,i}\), population \(j\), observation \(i\).

We form a paired difference from the data \(d_i = x_{1,i} - x_{2,i}\)
\(d_1 = x_{1,1} - x_{2,1}\), \(d_2 = x_{1,2} - x_{2,2}\), \ldots, \(d_n = x_{1,n} - x_{2,n}\).

When paired observations are randomly selected from normal populations, the paired difference, \(d_i = x_{1,i} - x_{2,i}\)
will be approximately normally distributed about a mean \(\mu_d\)
with a standard deviation \(\sigma_d\).
So if the $d_i$’s are approximately normally distributed

with a mean of $\mu_d$ and a standard deviation of $\sigma_d$, then

$$ \bar{d} = -\frac{1}{n} \sum_{i=1}^{n} d_i $$

is normally distributed (recall CLT)

with a mean $\mu_{\bar{d}} = \mu_d$, and standard deviation $\sigma_{\bar{d}} = \frac{\sigma_d}{\sqrt{n}}$. 
10: Inferences Involving Two Populations
10.2 Inference for Mean Difference Two Dependent Samples

This would allow us to form a \( z \) statistic for the mean of differences \( \bar{d} \), 
\[
    z = \frac{\bar{d} - \mu_d}{\sigma_d / \sqrt{n}}
\]
with a standard normal distribution.

We can then look up probabilities in the table, find critical values \( z(\alpha) \), construct confidence intervals
\[
    \bar{d} \pm z(\alpha / 2) \frac{\sigma_d}{\sqrt{n}}
\]
and test hypotheses using
\[
    z^* = \frac{\bar{d} - \mu_{0d}}{\sigma_d / \sqrt{n}}
\]

Figure from Johnson & Kuby, 2012.
10: Inferences Involving Two Populations

10.2 Inference for Mean Difference Two Dependent Samples

However, as in Inferences for One Population, we never know the true value of $\sigma_d$. So we estimate it with sample standard deviation $s_d$. This changes the formula to

$$ z = \frac{\bar{d} - \mu_d}{\sigma_d / \sqrt{n}} $$

to

$$ t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} $$

and the distribution from standard normal to Student $t$ with $df=n-1$ where

$$ s_d^2 = \frac{1}{n-1} \sum_{i=1}^{n} (d_i - \bar{d})^2. $$
10: Inferences Involving Two Populations

10.2 Inference for Mean Difference Two Dependent Samples

Confidence Interval Procedure

With $\sigma_d$ unknown, a $1-\alpha$ confidence interval for $\mu_d$ is:

Confidence Interval for Mean Difference (Dependent Samples)

$$\bar{d} - t(df, \alpha / 2) \frac{s_d}{\sqrt{n}} \quad \text{to} \quad \bar{d} + t(df, \alpha / 2) \frac{s_d}{\sqrt{n}} \quad \text{where } df = n - 1$$

(10.2)

$$\bar{d} = \frac{1}{n} \sum_{i=1}^{n} d_i \quad (10.3)$$

$$s_d^2 = \frac{1}{n-1} \sum_{i=1}^{n} (d_i - \bar{d})^2 \quad (10.4)$$
Example:
Construct a 95% CI for mean difference in Brand B – A tire wear.

d_i’s: 8, 1, 9, –1, 12, 9

\[ \bar{d} = 6.3 \quad \alpha = 0.05 \]

\[ s_d = 5.1 \]

\[ \bar{d} \pm t(df, \alpha / 2) \frac{s_d}{\sqrt{n}} \rightarrow (0.090, 11.7) \]

Figure from Johnson & Kuby, 2012.
We can test for differences in the population means:

\[ H_0: \mu_1 \geq \mu_2 \text{ vs. } H_a: \mu_1 < \mu_2 \quad \rightarrow \quad H_0: \mu_1 - \mu_2 \geq 0 \text{ vs. } H_a: \mu_1 - \mu_2 < 0 \]

\[ H_0: \mu_1 \leq \mu_2 \text{ vs. } H_a: \mu_1 > \mu_2 \quad \rightarrow \quad H_0: \mu_1 - \mu_2 \leq 0 \text{ vs. } H_a: \mu_1 - \mu_2 > 0 \]

\[ H_0: \mu_1 = \mu_2 \text{ vs. } H_a: \mu_1 \neq \mu_2 \quad \rightarrow \quad H_0: \mu_1 - \mu_2 = 0 \text{ vs. } H_a: \mu_1 - \mu_2 \neq 0 \]

\[ \mu_d = \mu_1 - \mu_2 \quad \rightarrow \quad H_0: \mu_d \geq 0 \text{ vs. } H_a: \mu_d < 0 \]

\[ (\mu_d = \mu_{\text{before}} - \mu_{\text{after}}) \quad H_0: \mu_d \leq 0 \text{ vs. } H_a: \mu_d > 0 \]

\[ H_0: \mu_d = 0 \text{ vs. } H_a: \mu_d \neq 0 \]
10: Inferences Involving Two Populations
10.2 Inference for Mean Difference Two Dependent Samples

Hypothesis Testing Procedure

With \( \sigma_d \) unknown, the test statistic for \( \mu_d \) is:

\[
 t^* = \frac{\bar{d} - \mu_{d0}}{s_d / \sqrt{n}} \quad \text{where} \; df = n - 1
\]

Go through the same five hypothesis testing steps.
10: Inferences Involving Two Populations

10.2 Inference for Mean Difference Two Dependent Samples

\[ n = 6 \quad 8, 1, 9, -1, 12, 9 \]

<table>
<thead>
<tr>
<th>Car</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brand A</td>
<td>125</td>
<td>64</td>
<td>94</td>
<td>38</td>
<td>90</td>
<td>106</td>
</tr>
<tr>
<td>Brand B</td>
<td>133</td>
<td>65</td>
<td>103</td>
<td>37</td>
<td>102</td>
<td>115</td>
</tr>
</tbody>
</table>

Example:

Test mean difference of Brand B minus Brand A is zero.

Step 1 \( H_0: \mu_d = 0 \) vs. \( H_a: \mu_d \neq 0 \)

Step 2

\[ d.f. = 5 \quad t^* = \frac{\bar{d} - \mu_{d0}}{s_d / \sqrt{n}} \]

\[ \alpha = .05 \]

Step 3 \( \bar{d} = 6.3 \]
\( s_d = 5.1 \)
\[ t^* = \frac{6.3 - 0}{5.1 / \sqrt{6}} = 3.03 \]

Step 4 \( t(df, \alpha / 2) = 2.57 \)

Step 5 Since \( t^* > t(df, \alpha / 2) \), reject \( H_0 \)

Conclusion: Significant difference in tread wear at .05 level.

Figures from Johnson & Kuby, 2012.
10: Inferences Involving Two Populations
10.3 Inference for Mean Difference Two Independent Samples

For Normal said that mean($\bar{x}$) = $\mu$ and variance($\bar{x}$) = $\frac{\sigma^2}{n}$.

We are often interested in comparisons between means $\bar{x}_1 - \bar{x}_2$.

There’s a rule that says that if $\bar{x}_1$ and $\bar{x}_2$ have means $\mu_1$ and $\mu_2$, and variances $\sigma_1^2$ and $\sigma_2^2$,

then mean($\bar{x}_1 - \bar{x}_2$) = $\mu_1 - \mu_2$

and variance($\bar{x}_1 - \bar{x}_2$) = $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$

if $x_1$ & $x_2$ independent

Rowe, D.B.
10: Inferences Involving Two Populations

10.3 Inference for Mean Difference Two Independent Samples

Means Using Two Independent Samples

If two populations are independent we can construct confidence intervals and test hypotheses for the difference in their means.

If independent samples of sizes $n_1$ and $n_2$ are drawn ... with means $\mu_1$ and $\mu_2$ and variances $\sigma^2_1$ and $\sigma^2_2$, then the sampling distribution of $\bar{x}_1 - \bar{x}_2$ ... has

1. mean $\mu_{\bar{x}_1 - \bar{x}_2} = \mu_1 - \mu_2$ and

2. standard error $\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{ \left( \frac{\sigma^2_1}{n_1} \right) + \left( \frac{\sigma^2_2}{n_2} \right) }$

If both pops. are normal, then $\bar{x}_1 - \bar{x}_2$ is normal . \hspace{1cm} (10.6)

Actually the CLT works here for $\bar{x}$‘s.
10: Inferences Involving Two Populations
10.3 Inference for Mean Difference Two Independent Samples

However, the true population variances are never truly known so we estimate \( \sigma_1^2 \) and \( \sigma_2^2 \) by \( s_1^2 \) and \( s_2^2 \) and the standard error

\[
\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)} \quad (10.6)
\]

by

\[
s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)} \quad . \quad (10.7)
\]
10: Inferences Involving Two Populations

10.3 Inference for Mean Difference Two Independent Samples

Confidence Interval Procedure

With $\sigma_1$ and $\sigma_2$ unknown, a 1-$\alpha$ confidence interval for $\mu_1 - \mu_2$ is:

$$
(\bar{x}_1 - \bar{x}_2) - t(df, \alpha / 2) \sqrt{ \left( \frac{s_1^2}{n_1} \right) + \left( \frac{s_2^2}{n_2} \right) } \quad \text{to} \quad (\bar{x}_1 - \bar{x}_2) + t(df, \alpha / 2) \sqrt{ \left( \frac{s_1^2}{n_1} \right) + \left( \frac{s_2^2}{n_2} \right) }
$$

where $df$ is either calculated or smaller of $df_1$, or $df_2$  

$$
df = \left( \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right) \left( \frac{s_1^2 / n_1}{n_1 - 1} + \frac{s_2^2 / n_2}{n_2 - 1} \right)^{1/2}
$$

Actually, this is for $\sigma_1 \neq \sigma_2$.

Next larger number than

If using a computer program.

If not using a computer program.
Example:
Interested in difference in mean heights between men and women. The heights of 20 females and 30 males is measured. Construct a 95% confidence interval for $\mu_m - \mu_f$, $\sigma_m$ & $\sigma_f$ unknown.

$$(\bar{x}_m - \bar{x}_f) \pm t(df, \alpha / 2) \sqrt{\left(\frac{s_m^2}{n_m}\right) + \left(\frac{s_f^2}{n_f}\right)}$$

$$= (69.8 - 63.8) \pm 2.09 \sqrt{\left(\frac{(1.92)^2}{30}\right) + \left(\frac{(2.18)^2}{20}\right)}$$

Therefore $4.75$ to $7.25$.
10: Inferences Involving Two Populations
10.3 Inference for Mean Difference Two Independent Samples

Hypothesis Testing Procedure

We can test for differences in the population means:

$H_0: \mu_1 \geq \mu_2$ vs. $H_a: \mu_1 < \mu_2 \quad \rightarrow \quad H_0: \mu_1 - \mu_2 \geq 0$ vs. $H_a: \mu_1 - \mu_2 < 0$

$H_0: \mu_1 \leq \mu_2$ vs. $H_a: \mu_1 > \mu_2 \quad \rightarrow \quad H_0: \mu_1 - \mu_2 \leq 0$ vs. $H_a: \mu_1 - \mu_2 > 0$

$H_0: \mu_1 = \mu_2$ vs. $H_a: \mu_1 \neq \mu_2 \quad \rightarrow \quad H_0: \mu_1 - \mu_2 = 0$ vs. $H_a: \mu_1 - \mu_2 \neq 0$
10: Inferences Involving Two Populations
10.3 Inference for Mean Difference Two Independent Samples

Hypothesis Testing Procedure

With $\sigma_1$ and $\sigma_2$ unknown, the test statistic for $\mu_1 - \mu_2$ is:

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

where $df$ is either calculated or smaller of $df_1$, or $df_2$ (10.9)

Go through the same five hypothesis testing steps.
10: Inferences Involving Two Populations

10.3 Inference for Mean Difference Two Independent Samples

Hypothesis Testing Procedure

Height vs. Weight

Rowe, D.B.
10: Inferences Involving Two Populations
10.3 Inference for Mean Difference Two Independent Samples

Hypothesis Testing Procedure

Rowe, D.B.
10: Inferences Involving Two Populations

10.3 Inference for Mean Difference Two Independent Samples

Hypothesis Testing Procedure

Step 1

\( H_0 : \mu_f = \mu_m \) vs. \( H_a : \mu_f \neq \mu_m \)

Step 2

\[ t^* = \frac{(\bar{x}_m - \bar{x}_f) - (\mu_m - \mu_f)}{\sqrt{\left(\frac{s^2_m}{n_m}\right) + \left(\frac{s^2_f}{n_f}\right)}} \]

\( df = 26 \)
\( \alpha = 0.05 \)

Step 3

\[ t^* = \frac{(71.7 - 64.4) - (0)}{\sqrt{\left(\frac{9.8}{27}\right) + \left(\frac{7.2}{53}\right)}} = 10.3 \]

Step 4

\[ t(df, \alpha / 2) = 2.06 \]

Step 5

Reject \( H_0 \) if \( 10.3 > 2.06 \), height \( m \)ales \( \neq \)height females

Rowe, D.B.
Chapter 10: Inferences Involving Two Populations

Questions?