Class 29

Daniel B. Rowe, Ph.D.

Department of Mathematics, Statistics, and Computer Science
Agenda:

Recap Chapter 11.1 and 11.2

Lecture Chapter 11.3

Review for Final Exam
3: Descriptive Analysis and Bivariate Data

3.1 Bivariate Data

**Bivariate data**: The values of two different variables that are obtained from the same population element.

Qualitative-Qualitative
Qualitative-Quantitative
Quantitative-Quantitative

When Qualitative-Qualitative
Cross-tabulation tables or contingency tables
Sometimes called $r$ by $c$ ($r \times c$)
3: Descriptive Analysis and Bivariate Data

3.1 Bivariate Data: two qualitative

Example:
Construct a $2 \times 3$ table.

<table>
<thead>
<tr>
<th>Name</th>
<th>Gender</th>
<th>Major</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adams</td>
<td>M</td>
<td>LA</td>
</tr>
<tr>
<td>Argento</td>
<td>F</td>
<td>BA</td>
</tr>
<tr>
<td>Baker</td>
<td>M</td>
<td>LA</td>
</tr>
<tr>
<td>Bennett</td>
<td>F</td>
<td>LA</td>
</tr>
<tr>
<td>Brand</td>
<td>M</td>
<td>T</td>
</tr>
<tr>
<td>Brock</td>
<td>M</td>
<td>BA</td>
</tr>
<tr>
<td>Chun</td>
<td>F</td>
<td>LA</td>
</tr>
<tr>
<td>Crain</td>
<td>M</td>
<td>T</td>
</tr>
<tr>
<td>Cross</td>
<td>F</td>
<td>BA</td>
</tr>
<tr>
<td>Ellis</td>
<td>F</td>
<td>BA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name</th>
<th>Gender</th>
<th>Major</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feeney</td>
<td>M</td>
<td>T</td>
</tr>
<tr>
<td>Flanigan</td>
<td>M</td>
<td>LA</td>
</tr>
<tr>
<td>Hodge</td>
<td>F</td>
<td>LA</td>
</tr>
<tr>
<td>Holmes</td>
<td>M</td>
<td>T</td>
</tr>
<tr>
<td>Jopson</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>Kee</td>
<td>M</td>
<td>BA</td>
</tr>
<tr>
<td>Kleeberg</td>
<td>M</td>
<td>LA</td>
</tr>
<tr>
<td>Light</td>
<td>M</td>
<td>BA</td>
</tr>
<tr>
<td>Linton</td>
<td>F</td>
<td>LA</td>
</tr>
<tr>
<td>Lopez</td>
<td>M</td>
<td>T</td>
</tr>
<tr>
<td>McGowan</td>
<td>M</td>
<td>BA</td>
</tr>
<tr>
<td>Mowers</td>
<td>F</td>
<td>BA</td>
</tr>
<tr>
<td>Ornt</td>
<td>M</td>
<td>T</td>
</tr>
<tr>
<td>Palmer</td>
<td>F</td>
<td>LA</td>
</tr>
<tr>
<td>Pullen</td>
<td>M</td>
<td>T</td>
</tr>
<tr>
<td>Rattan</td>
<td>M</td>
<td>BA</td>
</tr>
<tr>
<td>Sherman</td>
<td>F</td>
<td>LA</td>
</tr>
<tr>
<td>Small</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>Tate</td>
<td>M</td>
<td>BA</td>
</tr>
<tr>
<td>Yamamoto</td>
<td>M</td>
<td>LA</td>
</tr>
</tbody>
</table>

M = male
F = female
LA = liberal arts
BA = business admin
T = technology

Figures from Johnson & Kuby, 2012.

Rowe, D.B.
11: Applications of Chi-Square
11.3 Inferences Concerning Contingency Tables

Example: Construct a $2 \times 3$ table.

Each in group of 300 students identified as male or female and asked if preferred classes in math-science, social science, or humanities.

Sample Results for Gender and Subject Preference

<table>
<thead>
<tr>
<th>Gender</th>
<th>Math-Science (MS)</th>
<th>Social Science (SS)</th>
<th>Humanities (H)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male (M)</td>
<td>37</td>
<td>41</td>
<td>44</td>
<td>122</td>
</tr>
<tr>
<td>Female (F)</td>
<td>35</td>
<td>72</td>
<td>71</td>
<td>178</td>
</tr>
<tr>
<td>Total</td>
<td>72</td>
<td>113</td>
<td>115</td>
<td>300</td>
</tr>
</tbody>
</table>

Figure from Johnson & Kuby, 2012.

Rowe, D.B.
11: Applications of Chi-Square
11.3 Inferences Concerning Contingency Tables

*Test of Independence*

Is “Preference for math-science, social science, or humanities” … “independent of the gender of a college student?”

Sample Results for Gender and Subject Preference

<table>
<thead>
<tr>
<th>Gender</th>
<th>Math-Science (MS)</th>
<th>Social Science (SS)</th>
<th>Humanities (H)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male (M)</td>
<td>37</td>
<td>41</td>
<td>44</td>
<td>122</td>
</tr>
<tr>
<td>Female (F)</td>
<td>35</td>
<td>72</td>
<td>71</td>
<td>178</td>
</tr>
<tr>
<td>Total</td>
<td>72</td>
<td>113</td>
<td>115</td>
<td>300</td>
</tr>
</tbody>
</table>

Figure from Johnson & Kuby, 2012.
11: Applications of Chi-Square

11.3 Inferences Concerning Contingency Tables

Test of Independence

Is “Preference for math-science, social science, or humanities” … “independent of the gender of a college student?”

There is a Hypothesis test (of independence) to determine this. Is Favorite Subject independent of Gender.

Similar to Example with the Die, now have rows $i$ and columns $j$.

Observed values, $O_{ij}$’s.

\[
\chi^2 = \sum_{\text{all cells}} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}
\]

What are $E_{ij}$’s?

<table>
<thead>
<tr>
<th>Gender</th>
<th>Math–Science (MS)</th>
<th>Social Science (SS)</th>
<th>Humanities (H)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male (M)</td>
<td>37</td>
<td>41</td>
<td>44</td>
<td>122</td>
</tr>
<tr>
<td>Female (F)</td>
<td>35</td>
<td>72</td>
<td>71</td>
<td>178</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>72</strong></td>
<td><strong>113</strong></td>
<td><strong>115</strong></td>
<td><strong>300</strong></td>
</tr>
</tbody>
</table>

Figure from Johnson & Kuby, 2012.
11: Applications of Chi-Square
11.3 Inferences Concerning Contingency Tables

Test of Independence

\[ \chi^2* = \sum_{all\ cells} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \]

D of F for Contingency Tables:

\[ df = (r - 1)(c - 1) \quad (11.4) \]

Where does this formula for \( E_{ij} \)'s come from?

\[ E_{ij} = \frac{\text{row total} \times \text{column total}}{\text{grand total}} = \frac{R_iC_j}{n} \quad (11.5) \]

Expected Frequencies for Contingency Tables

\[ r > 1, c > 1 \]
4: Probability

4.5 Independent Events

**Independent events**: Two events are independent if the occurrence or nonoccurrence of one gives us no information about the likeliness of occurrence for the other.

In algebra:

\[ P(A) = P(A \mid B) = P(A \mid \text{not } B) \]

In words:

1. Prob of \( A \) unaffected by knowledge that \( B \) has occurred, not occurred, or no knowledge.
2. …
3. …
4: Probability
4.5 Independent Events

Two events $A$ and $B$ are independent if the probability of one is not “influenced” by the occurrence or nonoccurrence of the other.

Two Events $A$ and $B$ are independent if:

1. $P(A) = P(A|B)$
2. $P(B) = P(B|A)$
3. $P(A \text{ and } B) = P(A) \cdot P(B)$

Examples:?
11: Applications of Chi-Square

11.3 Inferences Concerning Contingency Tables

*Test of Independence*

Where does this formula for \( E_{ij} \)'s come from?

\[
E_{ij} = \frac{R_i C_j}{n}
\]

<table>
<thead>
<tr>
<th>Gender</th>
<th>Math–Science (MS)</th>
<th>Social Science (SS)</th>
<th>Humanities (H)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male (M)</td>
<td>37</td>
<td>41</td>
<td>44</td>
<td>122</td>
</tr>
<tr>
<td>Female (F)</td>
<td>35</td>
<td>72</td>
<td>71</td>
<td>178</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>72</strong></td>
<td><strong>113</strong></td>
<td><strong>115</strong></td>
<td><strong>300</strong></td>
</tr>
</tbody>
</table>

If Favorite Subject (column variable) is independent of Gender (row variable), then

\[
P(\text{MS} \mid M) = P(\text{MS} \mid F) = P(\text{MS})
\]

\[
P(A) = P(A \mid B)
\]

\[
P(A \text{ and } B) = P(A) \cdot P(B)
\]

Figure from Johnson & Kuby, 2012.
11: Applications of Chi-Square
11.3 Inferences Concerning Contingency Tables

Test of Independence

Where does this formula for $E_{ij}$’s come from?

<table>
<thead>
<tr>
<th></th>
<th>MS</th>
<th>SS</th>
<th>H</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>29.28</td>
<td>45.95</td>
<td>46.77</td>
<td>122.00</td>
</tr>
<tr>
<td>Female</td>
<td>42.72</td>
<td>67.05</td>
<td>68.23</td>
<td>178.00</td>
</tr>
<tr>
<td>Total</td>
<td>72.00</td>
<td>113.00</td>
<td>115.00</td>
<td>300.00</td>
</tr>
</tbody>
</table>

If Favorite Subject is independent of Gender, then

\[
P(M \text{ and } MS) = P(M)P(MS) = (122/300)(72/300)
\]

\[
E(M \text{ and } MS) = nP(M)P(MS) = 300(122/300)(72/300)
\]

\[
E(M \text{ and } MS) = 122 \times 72 / 300
\]

Figure from Johnson & Kuby, 2012.
11: Applications of Chi-Square
11.3 Inferences Concerning Contingency Tables

Test of Independence

Where does this formula for $E_{ij}$’s come from?

$$E_{ij} = \frac{R_i C_j}{n}$$

<table>
<thead>
<tr>
<th>Gender</th>
<th>MS</th>
<th>SS</th>
<th>H</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>37 (29.28)</td>
<td>41 (45.95)</td>
<td>44 (46.77)</td>
<td>122</td>
</tr>
<tr>
<td>Female</td>
<td>35 (42.72)</td>
<td>72 (67.05)</td>
<td>71 (68.23)</td>
<td>178</td>
</tr>
<tr>
<td>Total</td>
<td>72</td>
<td>113</td>
<td>115</td>
<td>300</td>
</tr>
</tbody>
</table>

If Favorite Subject is independent of Gender, then

$$\chi^2* = \sum_{all\ cells} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} < \chi^2(2, 0.05)$$

$$\alpha = 0.05$$

$$\chi^2* = 4.604 < \chi^2(2, 0.05) = 5.99$$

$$df = (r - 1)(c - 1) = (2 - 1)(3 - 1)$$

Figure from Johnson & Kuby, 2012.
11: Applications of Chi-Square
11.3 Inferences Concerning Contingency Tables

**Test of Independence**

Expected Frequencies for an $r \times c$ Contingency Table

<table>
<thead>
<tr>
<th>Row</th>
<th>1</th>
<th>2</th>
<th>$j$th column</th>
<th>...</th>
<th>$c$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\frac{R_1 \times C_1}{n}$</td>
<td>$\frac{R_1 \times C_2}{n}$</td>
<td>...</td>
<td>$\frac{R_1 \times C_i}{n}$</td>
<td>...</td>
<td>$\frac{R_1 \times C_c}{n}$</td>
</tr>
<tr>
<td>2</td>
<td>$\frac{R_2 \times C_1}{n}$</td>
<td>$\frac{R_2 \times C_2}{n}$</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$i$th row</td>
<td>$\frac{R_i \times C_1}{n}$</td>
<td>$\frac{R_i \times C_2}{n}$</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$r$</td>
<td>$\frac{R_r \times C_1}{n}$</td>
<td>$\frac{R_r \times C_2}{n}$</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Total: $C_1$ $C_2$ ... $C_i$ ... ... $n$

$$E_{ij} = \frac{R_i C_j}{n}$$

$$\chi^2 = \sum_{all \ cells} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

$$< \chi^2((r - 1)(c - 1), \alpha)$$

Figure from Johnson & Kuby, 2012.
11: Applications of Chi-Square

11.3 Inferences Concerning Contingency Tables

*Test of Homogeneity*

Is the distribution within all rows the same for all rows?

<table>
<thead>
<tr>
<th>Residence</th>
<th>Governor’s Proposal</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Favor</td>
<td>Oppose</td>
<td>Total</td>
</tr>
<tr>
<td>Urban</td>
<td>143</td>
<td>57</td>
<td>200</td>
</tr>
<tr>
<td>Suburban</td>
<td>98</td>
<td>102</td>
<td>200</td>
</tr>
<tr>
<td>Rural</td>
<td>13</td>
<td>87</td>
<td>100</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>254</strong></td>
<td><strong>246</strong></td>
<td><strong>500</strong></td>
</tr>
</tbody>
</table>

If so, then

\[ P(F \text{ and } Urban) = P(F)P(U) \]

\[ E(F \text{ and } Urban) = nP(F)P(U) \]

\[ E(F \text{ and } Urban) = 500 \left( \frac{254}{500} \right) \left( \frac{200}{500} \right) \]
11: Applications of Chi-Square

11.3 Inferences Concerning Contingency Tables

*Test of Homogeneity*

Is the distribution within all rows the same for all rows?

\[ E_{ij} = \frac{R_i C_j}{n} \]

<table>
<thead>
<tr>
<th>Residence</th>
<th>Governor’s Proposal</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Favor</td>
<td>Oppose</td>
</tr>
<tr>
<td>Urban</td>
<td>143</td>
<td>57</td>
</tr>
<tr>
<td>Suburban</td>
<td>98</td>
<td>102</td>
</tr>
<tr>
<td>Rural</td>
<td>13</td>
<td>87</td>
</tr>
<tr>
<td>Total</td>
<td>254</td>
<td>246</td>
</tr>
</tbody>
</table>

\[ \chi^2* = \sum_{all \ cells} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} < \chi^2((r - 1)(c - 1), \alpha) \]

\[ \alpha = 0.05 \]

\[ df = (r - 1)(c - 1) = (3 - 1)(2 - 1) \]
Chapter 11: Applications of Chi-Square

Questions?

Homework: Chapter 11 # 3, 5, 11, 15, 21, 49, 53
Review Chapters 9 and 10
(Final Exam Chapters)

Just the highlights!
Recap Chapter 9
9: Inferences Involving One Population
9.1 Inference about the Mean $\mu$ ($\sigma$ Unknown)

In Chapter 8, we performed hypothesis tests on the mean by

1) assuming that $\bar{x}$ was normally distributed ($n$ “large”),

2) assuming the hypothesized mean $\mu_0$ were true,

3) assuming that $\sigma$ was known, so that we could form

$$z^* = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

which with 1) – 3) has standard normal dist.
9: Inferences Involving One Population

9.1 Inference about the Mean $\mu$ ($\sigma$ Unknown)

However, in real life, we never know $\sigma$ for so we would like to estimate $\sigma$ by $s$, then use

$$z^* = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

so we would like to estimate $\sigma$ by $s$, then use

$$t^* = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}.$$

But $t^*$ does not have a standard normal distribution.

It has what is called a Student $t$-distribution.
9: Inferences Involving One Population
9.1 Inference about the Mean $\mu$ ($\sigma$ Unknown)
Using the $t$-Distribution Table

Finding critical value from a Student $t$-distribution, $df=n-1$

t($df, \alpha$), $t$ value with $\alpha$ area larger than it

with $df$ degrees of freedom

Table 6
Appendix B
Page 719.

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

9.1 Inference about the Mean $\mu$ ($\sigma$ Unknown)

**Example:** Find the value of $t(10,0.05)$, $df=10$, $\alpha=0.05$.

---

**Table 6**

Appendix B

Page 719.

Go to 0.05 One Tail column and down to 10 $df$ row.

Figures from Johnson & Kuby, 2012.
9: Inferences Involving One Population
9.1 Inference about the Mean $\mu$ ($\sigma$ Unknown)

Recap 9.1:
Essentially have new critical value, $t(df, \alpha)$ to look up in a table when $\sigma$ is unknown. Used same as before.

$\sigma$ assumed known

$$
\bar{x} \pm z(\alpha / 2) \frac{\sigma}{\sqrt{n}}
$$

$z^* = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$

$\sigma$ assumed unknown

$$
\bar{x} \pm t(df, \alpha / 2) \frac{s}{\sqrt{n}}
$$

$$
t^* = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}
$$
We talked about a Binomial experiment with two outcomes.

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

where:
- \( n \) is the number of trials,
- \( x \) is the number of successes,
- \( p \) is the probability of success.

For a sample binomial probability:

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

where \( x \) is the number of successes in \( n \) trials.
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

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

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.”

---

Rowe, D.B.
9: Inferences Involving One Population
9.2 Inference about the Binomial Probability of Success

For a confidence interval, we would use

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

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**

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

(9.7)

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

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

(9.8)

where \( p^* \) and \( q^* \) are provisional values used for planning.

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

\( H_0: \sigma^2 \geq \sigma_0^2 \) vs. \( H_a: \sigma^2 < \sigma_0^2 \)

\( H_0: \sigma^2 \leq \sigma_0^2 \) vs. \( H_a: \sigma^2 > \sigma_0^2 \)

\( H_0: \sigma^2 = \sigma_0^2 \) vs. \( H_a: \sigma^2 \neq \sigma_0^2 \)

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

\[ \mu = df \]

\[ \sigma^2 = 2df \]

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)

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

(9.10)

Will also need critical values.

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

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>0.0000393</td>
<td>0.000157</td>
<td>0.000982</td>
<td>0.00393</td>
<td>0.0158</td>
<td>0.0102</td>
<td>0.0455</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>0.207</td>
<td>0.297</td>
<td>0.484</td>
<td>0.711</td>
<td>1.06</td>
<td>1.92</td>
<td>3.36</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>15.3</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>26.7</td>
<td>29.7</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>20.5</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 # 7, 21, 23, 35, 37, 39, 47, 55, 67, 73, 75, 93, 95, 97, 103, 117, 119, 121, 129, 131, 135
Recap Chapter 10
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 $\mu_d=(\mu_1-\mu_2)$ is:

Confidence Interval for Mean Difference (Dependent Samples)

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

where $df=n-1$  \(10.2\)
10: Inferences Involving Two Populations

10.2 Inference for Mean Difference Two Dependent Samples

Example:

Construct a 95% CI for mean difference in Brand B – A tire wear.

\[ \bar{d} = \frac{1}{n} \sum_{i=1}^{n} d_i \]

\[ df = \frac{\sum (d_i - \bar{d})^2}{\sqrt{n}} \]

\[ s_d = 5.1 \quad \alpha = 0.05 \]

\[ n = 6 \quad t(df, \alpha / 2) = 2.57 \]

\[ \bar{d} = 6.3 \]

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

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

10.2 Inference for Mean Difference Two Dependent Samples

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

\[
\begin{align*}
df &= 5 \\
t^* &= \frac{\bar{d} - \mu_{d0}}{s_d / \sqrt{n}}
\end{align*}
\]

$\alpha = .05$

Step 3

\[
\begin{align*}
\bar{d} &= 6.3 \\
s_d &= 5.1
\end{align*}
\]

$t^* = \frac{6.3 - 0}{5.1 / \sqrt{6}} = 3.03$

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
Confidence Interval Procedure

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

Confidence Interval for Mean Difference (Independent Samples)

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

where $df$ is either calculated or smaller of $df_1$, or $df_2$  \hspace{1cm} (10.8)

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

Next larger number than

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

If using a computer program.
If not using a computer program.
# 10: Inferences Involving Two Populations

## 10.3 Inference Mean Difference

### Confidence Interval

**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 $7.25$

---

<table>
<thead>
<tr>
<th>Sample</th>
<th>Number</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female (f)</td>
<td>$n_1 = 20$</td>
<td>$\bar{x}_f = 63.8$</td>
<td>$s_f = 2.18$</td>
</tr>
<tr>
<td>Male (m)</td>
<td>$n_2 = 30$</td>
<td>$\bar{x}_m = 69.8$</td>
<td>$s_m = 1.92$</td>
</tr>
</tbody>
</table>

Figure from Johnson & Kuby, 2012.
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
\[ t^* = \frac{(\bar{x}_m - \bar{x}_f) - (\mu_m - \mu_f)}{\sqrt{\left(\frac{s_m^2}{n_m}\right) + \left(\frac{s_f^2}{n_f}\right)}} \]

\( df = 7 \)
\( \alpha = .05 \)

Step 3
\[ t^* = \frac{(71.4 - 65.2) - (0)}{\sqrt{\left(\frac{7.4}{8}\right) + \left(\frac{8.2}{24}\right)}} = 5.5 \]

Step 4
\[ t(df, \alpha / 2) = 2.36 \]

Step 5
Reject \( H_0 \) if \( 5.5 > 2.36 \), height males \( \neq \) height females.
10: Inferences Involving Two Populations

10.4 Inference for Difference between Two Proportions

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

If independent samples of size $n_1$ and $n_2$ are drawn … with $p_1 = P_1(\text{success})$ and $p_2 = P_2(\text{success})$, then the sampling distribution of $p'_1 - p'_2$ has these properties:

1. mean $\mu_{p'_1 - p'_2} = p_1 - p_2$

2. standard error $\sigma_{p'_1 - p'_2} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$ (10.10)

3. approximately normal dist if $n_1$ and $n_2$ are sufficiently large.

ie I $n_1, n_2 > 20$  II $n_1 p_1, n_1 q_1, n_2 p_2, n_2 q_2 > 5$  III sample <10% of pop
10: Inferences Involving Two Populations
10.4 Inference for Difference between Two Proportions
Confidence Interval Procedure

Assumptions for ... difference between two proportions

\( p_1 - p_2 \): The \( n_1 \) ... and \( n_2 \) random observations ... are selected independently from two populations that are not changing

Confidence Interval for the Difference between Two Proportions \( p_1 - p_2 \)

\[
(p_1' - p_2') - z(\alpha / 2) \sqrt{\frac{p_1'q_1'}{n_1} + \frac{p_2'q_2'}{n_2}} \quad \text{to} \quad (p_1' - p_2') + z(\alpha / 2) \sqrt{\frac{p_1'q_1'}{n_1} + \frac{p_2'q_2'}{n_2}}
\]

where \( p_1' = \frac{x_1}{n_1} \) and \( p_2' = \frac{x_2}{n_2} \).  

(10.11)
10: Inferences Involving Two Populations

10.4 Inference for Difference between Two Proportions

Confidence Interval Procedure

Example:
Construct a 99% CI for proportion of female A’s minus male A’s difference $p_f - p_m$.

40 values

$\begin{align*}
  n_m &= 9 \\
  n_f &= 31 \\
  x_m &= 2 \\
  x_f &= 11
\end{align*}$

$\begin{align*}
  p'_f &= \frac{x_f}{n_f} = \frac{11}{31} = .35 \\
  p'_m &= \frac{x_m}{n_m} = \frac{2}{9} = .22
\end{align*}$

$\begin{align*}
  z(\alpha / 2) &= 2.58 \\
  (p'_f - p'_m) &\pm z(\alpha / 2) \sqrt{\frac{p'_f q'_f}{n_f} + \frac{p'_m q'_m}{n_m}} \\
  (.35 - .22) &\pm 2.05 \sqrt{\frac{(.35)(.65)}{31} + \frac{(.22)(.78)}{9}} \\
  &\approx -.287 \text{ to } .553
\end{align*}$

Rowe, D.B.
10: Inferences Involving Two Populations
10.4 Inference for Difference between Two Proportions

Hypothesis Testing Procedure

We can perform hypothesis tests on the proportion

\[ H_0: p_1 \geq p_2 \text{ vs. } H_a: p_1 < p_2 \]

\[ H_0: p_1 \leq p_2 \text{ vs. } H_a: p_1 > p_2 \]

\[ H_0: p_1 = p_2 \text{ vs. } H_a: p_1 \neq p_2 \]

Test Statistic for the Difference between two Proportions - Population Proportions Known

\[
z^* = \frac{(p_1' - p_2') - (p_{10} - p_{20})}{\sqrt{pq \left[ \frac{1}{n_1} + \frac{1}{n_2} \right]}}
\]

\[
p_1' = \frac{x_1}{n_1}, \quad p_2' = \frac{x_2}{n_2}
\]

when \( p_1 = p_2 = p \).
10: Inferences Involving Two Populations
10.4 Inference for Difference between Two Proportions

Hypothesis Testing Procedure

where we assume \( p_1 = p_2 \) and use pooled estimate of proportion \( \hat{p} \).

Test Statistic for the Difference between two Proportions \( \text{UnKnown} \)

\[
Z^* = \frac{(p_1' - p_2') - (p_{10} - p_{20})}{\sqrt{\frac{p_1'q_1'}{n_1} + \frac{p_2'q_2'}{n_2}}}
\]

\[\hat{p}_p, \text{ estimated}\]

\[
p_1' = \frac{x_1}{n_1} \quad p_2' = \frac{x_2}{n_2} \quad \frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2} = pq \left[ \frac{1}{n_1} + \frac{1}{n_2} \right] \quad p_p' = \frac{x_1 + x_2}{n_1 + n_2}
\]
10: Inferences Involving Two Populations

10.4 Inference for Difference between Two Proportions

Hypothesis Testing Procedure

Step 1

\[ H_0: p_s - p_c \leq 0 \text{ vs. } H_a: p_s - p_c > 0 \]

Step 2

\[ z^* = \frac{(p'_s - p'_c) - (p_{0s} - p_{0c})}{\sqrt{p'_p q'_p \left( \frac{1}{n_s} + \frac{1}{n_c} \right)}} \]

\( \alpha = .05 \)

Step 3

\[ z^* = \frac{(0.10 - 0.04) - (0)}{\sqrt{(0.07)(0.93) \left[ \frac{1}{150} + \frac{1}{150} \right]}} = 2.04 \]

Step 4

\[ z(\alpha) = 1.65 \]

\[ .02 < p-value < .023 \text{ or } 2.04 > 1.65 \]

Step 5

Reject \( H_0 \)

Rowe, D.B.
10: Inferences Involving Two Populations
10.5 Inference for Ratio of Two Variances Two Ind. Samples

Hypothesis Testing Procedure

We can perform hypothesis tests on two variances

\[ H_0: \sigma_1^2 \geq \sigma_2^2 \quad \text{vs.} \quad H_a: \sigma_1^2 < \sigma_2^2 \]

\[ H_0: \sigma_1^2 \leq \sigma_2^2 \quad \text{vs.} \quad H_a: \sigma_1^2 > \sigma_2^2 \]

\[ H_0: \sigma_1^2 = \sigma_2^2 \quad \text{vs.} \quad H_a: \sigma_1^2 \neq \sigma_2^2 \]

Assumptions: Independent samples from normal distribution

Test Statistic for Equality of Variances

\[ F^* = \frac{S_n^2}{S_d^2} \quad \text{with} \quad df_n = n_n - 1 \quad \text{and} \quad df_d = n_d - 1. \]  

(10.16)

Use new table to find areas for new statistic.
10: Inferences Involving Two Pops.
10.5 Inference Ratio of Two Variances

Example: Find $F(5, 8, 0.05)$.

$$df_n = n_n - 1 \quad df_d = n_d - 1$$

Table 9, Appendix B, Page 722.

<table>
<thead>
<tr>
<th>Degrees of Freedom for Denominator $df_d$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>18.5</td>
<td>19.0</td>
<td>19.2</td>
<td>19.2</td>
<td>19.3</td>
<td>19.3</td>
<td>19.4</td>
<td>19.4</td>
<td>19.4</td>
<td>19.4</td>
</tr>
<tr>
<td>3</td>
<td>10.1</td>
<td>9.55</td>
<td>9.28</td>
<td>9.12</td>
<td>9.01</td>
<td>8.94</td>
<td>8.89</td>
<td>8.85</td>
<td>8.81</td>
<td>8.79</td>
</tr>
<tr>
<td>4</td>
<td>7.71</td>
<td>6.94</td>
<td>6.59</td>
<td>6.39</td>
<td>6.26</td>
<td>6.16</td>
<td>6.09</td>
<td>6.04</td>
<td>6.00</td>
<td>5.96</td>
</tr>
<tr>
<td>5</td>
<td>6.61</td>
<td>5.79</td>
<td>5.41</td>
<td>5.19</td>
<td>5.05</td>
<td>4.95</td>
<td>4.88</td>
<td>4.82</td>
<td>4.77</td>
<td>4.74</td>
</tr>
<tr>
<td>6</td>
<td>5.99</td>
<td>5.14</td>
<td>4.76</td>
<td>4.53</td>
<td>4.39</td>
<td>4.28</td>
<td>4.21</td>
<td>4.15</td>
<td>4.10</td>
<td>4.06</td>
</tr>
<tr>
<td>7</td>
<td>5.59</td>
<td>4.74</td>
<td>4.35</td>
<td>4.12</td>
<td>3.97</td>
<td>3.87</td>
<td>3.79</td>
<td>3.73</td>
<td>3.68</td>
<td>3.64</td>
</tr>
<tr>
<td>8</td>
<td>5.32</td>
<td>4.46</td>
<td>4.07</td>
<td>3.84</td>
<td>3.69</td>
<td>3.58</td>
<td>3.50</td>
<td>3.44</td>
<td>3.39</td>
<td>3.35</td>
</tr>
<tr>
<td>9</td>
<td>5.12</td>
<td>4.26</td>
<td>3.86</td>
<td>3.63</td>
<td>3.48</td>
<td>3.37</td>
<td>3.29</td>
<td>3.23</td>
<td>3.18</td>
<td>3.14</td>
</tr>
<tr>
<td>10</td>
<td>4.96</td>
<td>4.10</td>
<td>3.71</td>
<td>3.48</td>
<td>3.33</td>
<td>3.22</td>
<td>3.14</td>
<td>3.07</td>
<td>3.02</td>
<td>2.98</td>
</tr>
</tbody>
</table>

$\alpha = 0.05$

$F(5, 8, 0.05) = 3.69$

Figures from Johnson & Kuby, 2012.
10: Inferences Involving Two Populations
10.5 Inference for Ratio of Two Variances Two Ind. Samples

Hypothesis Testing Procedure

One tailed tests: Arrange $H_0$ & $H_a$ so $H_a$ is always “greater than”

$H_0: \sigma_1^2 \geq \sigma_2^2$ vs. $H_a: \sigma_1^2 < \sigma_2^2$ \quad \rightarrow \quad H_0: \frac{\sigma_2^2}{\sigma_1^2} \leq 1$ vs. $H_a: \frac{\sigma_2^2}{\sigma_1^2} > 1$ \quad \quad F* = \frac{s_2^2}{s_1^2}$

$H_0: \sigma_1^2 \leq \sigma_2^2$ vs. $H_a: \sigma_1^2 > \sigma_2^2$ \quad \quad H_0: \frac{\sigma_1^2}{\sigma_2^2} \leq 1$ vs. $H_a: \frac{\sigma_1^2}{\sigma_2^2} > 1$ \quad F* = \frac{s_1^2}{s_2^2}$

Reject $H_0$ if $F* = \frac{s_n^2}{s_d^2} > F(df_n, df_d, \alpha)$.

Two tailed tests: put larger sample variance $s^2$ in numerator

$H_0: \sigma_1^2 = \sigma_2^2$ vs. $H_a: \sigma_1^2 \neq \sigma_2^2$ \quad \rightarrow \quad H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1$ vs. $H_a: \frac{\sigma_1^2}{\sigma_2^2} \neq 1$

$\frac{\sigma_n^2}{\sigma_d^2} = \frac{\sigma_1^2}{\sigma_2^2}$ if $s_1^2 > s_2^2$ \quad \quad \frac{\sigma_n^2}{\sigma_d^2} = \frac{\sigma_2^2}{\sigma_1^2}$ if $s_2^2 > s_1^2$

Reject $H_0$ if $F* = \frac{s_n^2}{s_d^2} > F(df_n, df_d, \alpha/2)$.
10: Inferences Involving Two Populations

10.5 Inference for Ratio of Two Variances Two Ind. Samples

Hypothesis Testing Procedure

Step 1

\[ H_0 : \sigma_m^2 \geq \sigma_f^2 \quad \text{vs.} \quad H_a : \sigma_m^2 < \sigma_f^2 \]

\[ H_0 : \sigma_f^2 \leq \sigma_m^2 \quad \text{vs.} \quad H_a : \sigma_f^2 > \sigma_m^2 \]

\[ H_0 : \sigma_f^2 / \sigma_m^2 \leq 1 \quad \text{vs.} \quad H_a : \sigma_f^2 / \sigma_m^2 > 1 \]

Step 2

\[ F^* = \frac{s_f^2}{s_m^2} \quad \text{df}_f = 23 \]

\[ F^* = \frac{s_f^2}{s_m^2} \quad \text{df}_f = 7 \]

Step 3

\[ \alpha = .01 \]

\[ F^* = \frac{8.2}{7.4} = 1.12 \]

Step 4

\[ F(23, 7, .01) = 6.09 \]

Step 5

Fail to Reject \( H_0 \) \[ 1.12 < 6.09 \]
Chapter 10: Inferences Involving Two Populations

Questions?