Chapter 13

Comparing Three or More Means

13.1 Comparing Three or More Means (One-Way Analysis of Variance)

A one-way ANOVA procedure tests multiple averages using test statistic

\[ F_0 = \frac{\text{MST}}{\text{MSE}} \]

where

\[ \text{MST} = \frac{\text{SST}}{k-1}, \quad \text{MSE} = \frac{\text{SSE}}{n-k} \]

where \( k \) is number of groups, \( n \) is total number of observations, and where

\[ \text{SST} = n_1(\bar{x}_1 - \bar{x})^2 + n_2(\bar{x}_2 - \bar{x})^2 + \cdots + n_k(\bar{x}_k - \bar{x})^2 \]

\[ \text{SSE} = (n_1-1)s_1^2 + (n_2-1)s_2^2 + \cdots + (n_k-1)s_k^2. \]

which has a \( F \) distribution, with \((n-k, k)\) df, where \( n \) data points sampled randomly from \( k \) treatments (populations), and where \( k \) samples are independent of one another, each normally distributed and all with same variance, \( \sigma^2 \), and where MST is mean square due to treatment and MSE is mean square due to error.

Normality is checked using normal probability plots and equal variance is (roughly) checked by assuring largest SD is no more than twice the smallest SD or if sample size in each treatment group is the same.

Exercise 13.1 (Comparing Three or More Means (One-Way ANOVA))

1. Test comparing multiple means, ANOVA: average drug responses A.
   Fifteen different patients, chosen at random, subjected to three drugs. Test if at least one of the three mean patient responses to drug is different at \( \alpha = 0.05 \).
<table>
<thead>
<tr>
<th>drug 1</th>
<th>drug 2</th>
<th>drug 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.90</td>
<td>5.51</td>
<td>5.01</td>
</tr>
<tr>
<td>5.92</td>
<td>5.50</td>
<td>5.00</td>
</tr>
<tr>
<td>5.91</td>
<td>5.50</td>
<td>4.99</td>
</tr>
<tr>
<td>5.89</td>
<td>5.49</td>
<td>4.98</td>
</tr>
<tr>
<td>5.88</td>
<td>5.50</td>
<td>5.02</td>
</tr>
</tbody>
</table>

\[ \bar{x}_1 \approx 5.90 \quad \bar{x}_2 \approx 5.50 \quad \bar{x}_3 \approx 5.00 \]

(Blank data table. Relabel var1 drug 1, var2 drug 2, var3 drug 3. Type data into these three columns. Data, Save data, 13.1.1 drug same data.)

(a) **Statement.** Choose one.

i. \( H_0 : \mu_1 = \mu_2 = \mu_3 \quad \text{vs} \quad H_1 : \mu_1 \neq \mu_2, \mu_1 = \mu_3. \)

ii. \( H_0 : \mu_1 = \mu_2 = \mu_3 \quad \text{vs} \quad H_1 : \mu_1 \neq \mu_3, \mu_1 \neq \mu_2. \)

iii. \( H_0 : \mu_1 = \mu_2 = \mu_3 \quad \text{vs} \quad H_1 : \mu_i \neq \mu_j, i \neq j; i, j = 1, 2, 3. \)

iv. \( H_0 : \text{means same} \quad \text{vs} \quad H_1 : \text{at least one of the means different} \)

No matter how the question is worded, null is *always “means same” and alternative is always “at least one of the means different”.*

(b) **Test.**

p–value = (circle one) \( 0.00 / 0.035 / 0.043. \)

(Stat, ANOVA, One Way, choose Compare selected columns drug 1, drug 2, drug 3, Calculate.)

Level of significance \( \alpha = \) (choose one) \( 0.01 / 0.05 / 0.10. \)

(c) **Conclusion.**

Since p–value = 0.00 < \( \alpha = 0.05, \)

(circle one) **do not reject** / reject null \( H_0 : \text{means same}. \)

Data indicates (circle one)

**average drug responses same**

**at least one of average drug responses different**

(d) **Related question.**

“\( H_1 : \text{at least one of the means different} \)” means: (choose one or more)

i. \( \mu_1 \neq \mu_2, \text{but } \mu_2 = \mu_3 \)

ii. \( \mu_1 \neq \mu_3, \text{but } \mu_2 = \mu_3 \)

iii. \( \mu_2 \neq \mu_3, \text{but } \mu_1 = \mu_3 \)

iv. \( \mu_1 \neq \mu_2, \mu_1 \neq \mu_3 \text{ and } \mu_2 \neq \mu_3 \)

2. **Test comparing multiple means, ANOVA: average drug responses B.**

Fifteen different patients, chosen at random, subjected to three drugs. Test if at least one of the of three mean patient responses to drug is different at \( \alpha = 0.05. \)
Section 1. Comparing Three or More Means (One-Way Analysis of Variance) (Lecture Notes 13)

<table>
<thead>
<tr>
<th>drug 1</th>
<th>drug 2</th>
<th>drug 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.90</td>
<td>6.31</td>
<td>4.52</td>
</tr>
<tr>
<td>4.42</td>
<td>3.54</td>
<td>6.93</td>
</tr>
<tr>
<td>7.51</td>
<td>4.73</td>
<td>4.48</td>
</tr>
<tr>
<td>7.89</td>
<td>7.20</td>
<td>5.55</td>
</tr>
<tr>
<td>3.78</td>
<td>5.72</td>
<td>3.52</td>
</tr>
</tbody>
</table>

\( \bar{x}_1 \approx 5.90 \quad \bar{x}_2 \approx 5.50 \quad \bar{x}_3 \approx 5.00 \)

(Blank data table. Relabel var1 drug 1, var2 drug 2, var3 drug 3. Type data into these three columns. Data, Save data, 13.1.2 drugs different data.)

(a) Statement. Choose none, one or more.
   i. \( H_0: \mu_1 = \mu_2 = \mu_3 \) vs \( H_1: \mu_1 \neq \mu_2, \mu_1 = \mu_3 \).
   ii. \( H_0: \mu_1 = \mu_2 = \mu_3 \) vs \( H_1: \mu_1 \neq \mu_3, \mu_1 \neq \mu_2 \).
   iii. \( H_0: \mu_1 = \mu_2 = \mu_3 \) vs \( H_1: \mu_i \neq \mu_j, i \neq j; i, j = 1, 2, 3 \).
   iv. \( H_0: \) means same vs \( H_1: \) at least one of the means different

(b) Test.
   p–value = (circle one) 0.00 / 0.35 / 0.66.
   (Stat, ANOVA, One Way, choose Compare selected columns drug 1, drug 2, drug 3, Calculate.)
   Level of significance \( \alpha = (\text{choose one}) 0.01 / 0.05 / 0.10 \).

(c) Conclusion.
   Since p–value = 0.66 > \( \alpha = 0.05 \),
   (circle one) do not reject / reject null \( H_0: \) means same.
   Data indicates (circle one)
   average drug responses same
   at least one of average drug responses different

(d) Related question. ANOVA table is

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum Of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Squares</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>2.033</td>
<td>2</td>
<td>1.0167</td>
<td>0.436</td>
</tr>
<tr>
<td>Error</td>
<td>27.985</td>
<td>12</td>
<td>2.3321</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>30.018</td>
<td>14</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The “SS” is “sum of squares”, “MS” is “mean squares” and “df” is “degrees of freedom”.
so test statistic is

\[
F_0 = \frac{\text{MST}}{\text{MSE}} = \frac{1.0167}{2.3321} =
\]

(circle one) 0.345 / 0.436 / 0.767,
so, with \( k-1 = 3-1 = 2 \) and \( n-k = 15-3 = 12 \) df,

\[
p–value = P(F \geq 0.436) =
\]
(circle one) $0.00 / 0.35 / 0.66$.

(Stat, Calculators, $F$, Num. DF: 2, Den. DF: 12, $\text{Prob}(X \geq ) 0.436 = 
\frac{2}{7}$). The $F$ statistic and $p$–value calculated here by hand may not match values calculated using accurate StatCrunch because of round–off error.


(a) Comparing averages, $\bar{x}_1$, $\bar{x}_2$ and $\bar{x}_3$, in two drug response data sets.

Of following two data sets, data set $A$,

<table>
<thead>
<tr>
<th>drug 1</th>
<th>drug 2</th>
<th>drug 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.90</td>
<td>5.51</td>
<td>5.01</td>
</tr>
<tr>
<td>5.92</td>
<td>5.50</td>
<td>5.00</td>
</tr>
<tr>
<td>5.91</td>
<td>5.50</td>
<td>4.99</td>
</tr>
<tr>
<td>5.89</td>
<td>5.49</td>
<td>4.98</td>
</tr>
<tr>
<td>5.88</td>
<td>5.50</td>
<td>5.02</td>
</tr>
</tbody>
</table>

$\bar{x}_1 \approx 5.90$ $\bar{x}_2 \approx 5.50$ $\bar{x}_3 \approx 5.00$

$s_1 \approx 0.12$ $s_2 \approx 0.004$ $s_3 \approx 0.01$

and data set $B$,

<table>
<thead>
<tr>
<th>drug 1</th>
<th>drug 2</th>
<th>drug 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.90</td>
<td>6.31</td>
<td>4.52</td>
</tr>
<tr>
<td>4.42</td>
<td>3.54</td>
<td>6.93</td>
</tr>
<tr>
<td>7.51</td>
<td>4.73</td>
<td>4.48</td>
</tr>
<tr>
<td>7.89</td>
<td>7.20</td>
<td>5.55</td>
</tr>
<tr>
<td>3.78</td>
<td>5.72</td>
<td>3.52</td>
</tr>
</tbody>
</table>

$\bar{x}_1 \approx 5.90$ $\bar{x}_2 \approx 5.50$ $\bar{x}_3 \approx 5.00$

$s_1 \approx 1.82$ $s_2 \approx 1.27$ $s_3 \approx 1.30$

three average patient responses within drugs in data set $A$ are (circle one) smaller than / same as / larger than three average patient responses within drugs in data set $B$.

(b) Comparing SDs, $s_1$, $s_2$ and $s_3$, in two drug response data sets.

The standard deviations in patient responses within drugs in data set $A$ are (circle one) smaller than / the same as / larger than the standard deviations in patient responses within drugs in data B.

(c) Comparing averages, taking into account SDs.

As shown in figure below, since standard deviations within drugs in data set $A$ are smaller than they are for data set $B$, we are “more certain” about where averages are in data set $A$, than we are about where averages are in data set $B$. Consequently, it is “easiest” to tell if average patient responses are different from one another in data set (choose one) $A$ / $B$. This is why $p$–value is smaller for data set $A$, than it is for data set $B$ even though averages are same in both sets.
Section 1. Comparing Three or More Means (One-Way Analysis of Variance) (Lecture Notes 13)

Data set A
small SD (MSE)

Data set B
large SD (MSE)

Figure 13.1 (Comparing average drug responses)

(d) **Comparing $F$ test statistics.**
Since mean square treatment (MST) same for data sets A and B, but mean square error (MSE) smaller for data set A than for data set B,

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

is (choose one) **smaller** / **same as** / **larger** for A than for B, so greater chance of rejecting null (means same) for A than for B.

4. **Test comparing multiple means, ANOVA: yet more drug responses.**
Twelve different patients, chosen at random, are subjected to three drugs. Test if at least one of the three average patient responses to drug is different at $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>drug 1</th>
<th>drug 2</th>
<th>drug 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.90</td>
<td>5.50</td>
<td>5.01</td>
</tr>
<tr>
<td>5.92</td>
<td>5.53</td>
<td>5.00</td>
</tr>
<tr>
<td>5.91</td>
<td></td>
<td>4.99</td>
</tr>
<tr>
<td>5.89</td>
<td></td>
<td>4.98</td>
</tr>
<tr>
<td>5.88</td>
<td></td>
<td>5.02</td>
</tr>
</tbody>
</table>

$x_1 \approx 5.90$  $x_2 \approx 5.52$  $x_3 \approx 5.00$
$s_1 \approx 0.016$  $s_2 = 0.021$  $s_3 \approx 0.016$

(Blank data table. Relabel var1 drug 1, var2 drug 2, var3 drug 3. Type data into these three columns. Data, Save data, 13.1.3 drugs unequal sample size.)

(a) **Check assumptions.**
Figure 13.2 (Normal probability plots for drug responses.)

i. **Data normal?**

Normal probability plots for *all* three samples indicates **normal / not normal** because data within dotted bounds.

(Graphics, QQ Plot, Select Columns: drug 1, drug 2, drug 3, Create Graph! It is also possible to check normality of residuals from ANOVA model.)

ii. **Constant (equal) variance (SD)?**

Since maximum SD, $s_2 \approx 0.021$, is less than twice minimum SD, $s_1 = s_3 \approx 0.016$, $\frac{s_2}{s_1} \approx \frac{0.021}{0.016} \approx 1.3 < 2$, data indicates **equal / unequal** SD for different drug responses.

(Stat, Summary Stats, Columns, Select Columns: drug 1, drug 2, drug 3, Calculate.)

(b) **Statement.** Choose none, one or more.

i. $H_0 : \mu_1 = \mu_2 = \mu_3$ vs $H_1 : \mu_1 \neq \mu_2, \mu_1 = \mu_3$.
ii. $H_0 : \mu_1 = \mu_2 = \mu_3$ vs $H_1 : \mu_1 \neq \mu_3, \mu_1 \neq \mu_2$.
iii. $H_0 : \mu_1 = \mu_2 = \mu_3$ vs $H_1 : \mu_i \neq \mu_j, i \neq j$; $i, j = 1, 2, 3$.
iv. $H_0 : \text{means same}$ vs $H_1 : \text{at least one of the means different}$

(c) **Test.**

p–value = (circle one) **0.00 / 0.035 / 0.043**.

(Stat, ANOVA, One Way, choose Compare selected columns drug 1, drug 2, drug 3, Calculate.)

Level of significance $\alpha = (\text{choose one}) 0.01 / 0.05 / 0.10$.

(d) **Conclusion.**

Since p–value $= 0.00 < \alpha = 0.05$, (circle one) **do not reject / reject** null $H_0 : \text{means same}$.

The data indicates (circle one) **drug response means same**

at least one of drug response means different

5. **Test comparing multiple means, ANOVA: rats city.**

Rats are counted at twelve different city locations, chosen at random. Test claim mean rat count per square meter in at least one of the three city areas is different at 1%.
Section 1. Comparing Three or More Means (One-Way Analysis of Variance) (Lecture Notes 13)

(Blank data table. Relabel var1 sewers, var2 parks, var3 city hall. Type data into these three columns. Data, Save data, 13.1.4 rats data.)

(a) **Statement.** Choose none, one or more.
   i. $H_0 : \mu_1 = \mu_2 = \mu_3$ vs $H_1 : \mu_1 \neq \mu_2, \mu_1 = \mu_3$.
   ii. $H_0 : \mu_1 = \mu_2 = \mu_3$ vs $H_1 : \mu_1 \neq \mu_3, \mu_1 \neq \mu_2$.
   iii. $H_0 : \mu_1 = \mu_2 = \mu_3$ vs $H_1 : \mu_i \neq \mu_j, i \neq j; i, j = 1, 2, 3$.
   iv. $H_0 :$ means same vs $H_1 :$ at least one of the means different

(b) **Test.**
   p–value = (circle one) 0.002 / 0.035 / 0.043.
   (Stat, ANOVA, One Way, choose Compare selected columns city hall, parks, sewers, Calculate.)
   Level of significance $\alpha =$ (choose one) 0.01 / 0.05 / 0.10.

(c) **Conclusion.**
   Since p–value = 0.002 < $\alpha =$ 0.01,
   (circle one) do not reject / reject null $H_0 :$ means same.
   The data indicates (circle one)
   rat count means same
   at least one of rat count means different

(d) **How are averages different?**

   ![Boxplots](Figure 13.3 (Side-by-side boxplots for rat count.))

   (Graphics, Boxplot, Select Columns: sewers, parks, city hall, Next, check use fences to identify outliers,
   Draw boxes horizontally, Create Graph!)
   Rat count, is, from smallest to largest (choose one)
Chapter 13. Comparing Three or More Means (Lecture Notes 13)

i. sewers, parks, city hall
ii. parks, sewers, city hall
iii. city hall, sewers, parks

Twelve different patients, chosen at random, are subjected to three drugs. Test if at least one of the three average patient responses to drug is different at $\alpha = 0.05$.

<table>
<thead>
<tr>
<th></th>
<th>Drug 1</th>
<th>Drug 2</th>
<th>Drug 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{x}_1$</td>
<td>5.90</td>
<td>$\bar{x}_2$</td>
<td>5.52</td>
</tr>
<tr>
<td>$s^2_1$</td>
<td>0.000256</td>
<td>$s^2_2$</td>
<td>0.000441</td>
</tr>
<tr>
<td>$n_1$</td>
<td>5</td>
<td>$n_2$</td>
<td>2</td>
</tr>
</tbody>
</table>

Since StatCrunch does not deal with summarized data for ANOVA, this analysis must be done “by hand”.

(a) Statement. Choose none, one or more.

i. $H_0: \mu_1 = \mu_2 = \mu_3$ vs $H_1: \mu_1 \neq \mu_2, \mu_1 = \mu_3$.

ii. $H_0: \mu_1 = \mu_2 = \mu_3$ vs $H_1: \mu_1 \neq \mu_3, \mu_1 \neq \mu_2$.

iii. $H_0: \mu_1 = \mu_2 = \mu_3$ vs $H_1: \mu_i \neq \mu_j, i \neq j; \ i, j = 1, 2, 3$.

iv. $H_0: \text{means same}$ vs $H_1: \text{at least one of the means different}$

(b) Test.
Since grand average is

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2 + n_3\bar{x}_3}{n_1 + n_2 + n_3} \approx \frac{5 \cdot 5.90 + 2 \cdot 5.52 + 5 \cdot 5.00}{5 + 2 + 5} \approx$$

(Circle one) 4.56 / 5.46 / 6.45 and

$$SST = n_1(\bar{x}_1 - \bar{x})^2 + n_2(\bar{x}_2 - \bar{x})^2 + n_3(\bar{x}_3 - \bar{x})^2$$
$$= 5(5.90 - 5.46)^2 + 2(5.52 - 5.46)^2 + 5(5.00 - 5.46)^2 \approx$$

(Circle one) 0.56 / 1.06 / 2.03 and

$$SSE = (n_1 - 1)s^2_1 + (n_2 - 1)s^2_2 + (n_3 - 1)s^2_3$$
$$= (5 - 1)0.000256 + (2 - 1)0.000441 + (5 - 1)0.000256 \approx$$

(Circle one) 0.0006 / 0.0016 / 0.0025

and so $MST = \frac{SST}{k-1} = \frac{2.03}{3-1} = $
Section 2. Post Hoc Tests on One-Way Analysis of Variance (Lecture Notes 13) 235

(circle one) 0.53 / 0.81 / 1.02
and MSE = \( \frac{\text{SSE}}{n-k} = \frac{0.0025}{12-3} \)
(Calculator: 0.0025/9.)
(circle one) 0.00027 / 0.0027 / 0.027
which gives ANOVA table

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum Of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Squares</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>2.03</td>
<td>2</td>
<td>1.02</td>
<td>3778</td>
</tr>
<tr>
<td>Error</td>
<td>0.0025</td>
<td>9</td>
<td>0.00027</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2.0325</td>
<td>11</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

where, notice, test statistic is

\[ F_0 = \frac{\text{MST}}{\text{MSE}} = \frac{1.02}{0.00027} \approx \]
(Calculator: 1.02/0.00027.)
(circle one) 1908 / 2310 / 3778,
so, with \( k - 1 = 3 - 1 = 2 \) and \( n - k = 12 - 3 = 9 \) df,

\[ \text{p-value} = P(F \geq 3778) = \]
(circle one) 0.00 / 0.35 / 0.66.
(Stat, Calculators, F, Num. DF: 2, Den. DF: 9, Prob(X \( \geq \)) 3778 = \[\_\_\_\_\_\_
Level of significance \( \alpha = \) (choose one) 0.01 / 0.05 / 0.10.

(c) Conclusion.
Since p-value = 0.00 < \( \alpha = 0.05 \),
(circle one) do not reject / reject null \( H_0 : \) means same.
Data indicates (circle one)
drug response means same
at least one of drug response means different

13.2 Post Hoc Tests on One-Way Analysis of Variance
Not covered.

13.3 The Randomized Complete Block Design
Not covered.
13.4 Two-Way Analysis of Variance

Not covered.

Summary Section

<table>
<thead>
<tr>
<th>HYPOTHESIS TESTS</th>
<th>mean $\mu$</th>
<th>variance $\sigma^2$</th>
<th>proportion $p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>one sample</td>
<td>chapter 10</td>
<td>chapter 10</td>
<td>chapter 10</td>
</tr>
<tr>
<td>two sample</td>
<td>chapter 11</td>
<td>chapter 11</td>
<td>chapter 11</td>
</tr>
<tr>
<td>multiple</td>
<td>$F_0 = \frac{MS_{A}}{MS_{B}}$</td>
<td>not covered</td>
<td>chapter 12</td>
</tr>
</tbody>
</table>