Cognitive development study

- Cognitive test administered to grades 1-5
- 15 children per grade
- Average scores increase approximately linearly
- Correlation between grade and $Y = 0.97$
- Use ANOVA to evaluate group differences

```r
> load(file=url("http://pnb.mcmaster.ca/bennett/psy710/datasets/contrasts.rda"))
> df3$grade <- factor(df3$grade, ordered=FALSE)
> options(contrasts=c("contr.sum","contr.poly"))  # IMPORTANT!!
> aov.01 <- aov(score~grade,df3)
> anova(aov.01)

Analysis of Variance Table

Response: score
  Df  Sum Sq Mean Sq F value Pr(>F)
grade   4 1361.2 340.31   2.2578 0.07152 .
Residuals 70 10550.9 150.73
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

- The effect of grade was not significant
- Do not reject the null hypothesis of no difference among group means
Cognitive development study

Note that 1-sided 95% CIs for effect size & association strength include zero

```r
> library(effectsize)
> cohens_f(aov.01)
# Effect Size for ANOVA
Parameter | Cohen's f | 95% CI
----------|-----------|---------
grade     | 0.36      | [0.00, Inf]
```

```r
> omega_squared(aov.01)
# Effect Size for ANOVA
Parameter | Omega2 | 95% CI
----------|--------|---------
grade     | 0.06   | [0.00, 1.00]
```

Note that 1-sided 95% CIs for effect size & association strength include zero

Cognitive development study

estimate power assuming medium effect size (f = 0.25)

```r
> library(pwr)
> pwr.anova.test(k=5,n=15,f=0.25,sig.level=.05,power=NULL)
```

Balanced one-way analysis of variance power calculation

- **k** = 5
- **n** = 15
- **f** = 0.25
- **sig.level** = 0.05
- **power** = 0.35

NOTE: n is number in each group

Cognitive development study

estimate power assuming f = 0.36

```r
> library(pwr)
> pwr.anova.test(k=5,n=15,f=0.36,sig.level=.05,power=NULL)
```

Balanced one-way analysis of variance power calculation

- **k** = 5
- **n** = 15
- **f** = 0.36
- **sig.level** = 0.05
- **power** = 0.67

NOTE: n is number in each group

Cognitive development study

check constant variance assumption

```r
> bartlett.test(score~grade,df3)
Bartlett test of homogeneity of variances
```

Bartlett test of homogeneity of variances

- **data**: score by grade
- **data**: Bartlett's K-squared = 6.6227, df = 4, p-value = 0.1572

Do not reject null hypothesis that variances are equal
Cognitive development study

check normality assumption

```r
> shapiro.test(residuals(aov.01))
Shapiro-Wilk normality test
data:  residuals(aov.01)
W = 0.9871, p-value = 0.6482
```

Do not reject null hypothesis that residuals are Normal

Cognitive development study

alternatives to ANOVA

```r
> oneway.test(score~grade,data=df3)
One-way analysis of means (not assuming equal variances)
data:  score and grade
F = 3, num df = 4, denom df = 35, p-value = 0.04

> kruskal.test(score~grade,data=df3)
Kruskal-Wallis rank sum test
data:  score by grade
Kruskal-Wallis chi-squared = 9, df = 4, p-value = 0.07
```

K-W Null Hypothesis:
- groups were sampled from the same distribution
- If we assume distributions have same shape & scale:
  - then H0 is that group medians are equal

Omnibus vs. Focused F tests

\[ H_0: \alpha_1 = \alpha_2 = \cdots = \alpha_a = 0 \]
\[ H_1: \alpha_j \neq 0 \]

- A significant omnibus F test supports a very general hypothesis
  - not all means are equal
  - not all group effects are zero
- Significant F doesn’t tell us how group means differ
- Generality of omnibus F often comes at cost of reduced power

linear contrasts/comparisons
Omnibus vs. Focused F tests

Omnibus F test is not significant:

```r
> lm.01 <- lm(score~grade,data=df3)
> anova(lm.01)
```

Analysis of Variance Table

Response: score

<table>
<thead>
<tr>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>grade</td>
<td>4</td>
<td>1361.2</td>
<td>340.31</td>
<td>2.2578</td>
</tr>
<tr>
<td>Residuals</td>
<td>70</td>
<td>10550.9</td>
<td>150.73</td>
<td></td>
</tr>
</tbody>
</table>

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

```r
> fractions( contrasts(df3$grade) )
```

```r
c1   c2   c3  c4  
g1 -1/2   -1    0    0  
g2 -1/2    1    0    0  
g3  1/3    0   -1    0  
g4  1/3    0  1/2   -1  
g5  1/3    0  1/2    1
```

```r
> aov.02 <- aov(score~grade,data=df3)
> summary(aov.02,
+     split=list(grade=list(c1=1,c2=2,c3=3,c4=4)))
```

Df | SS | MS | F | Pr(>F)
---|----|----|---|--------|
grade | 4 | 1361 | 340.3 | 2.258 | 0.0715 . |
grade: c1 | 1 | 753 | 752.9 | 4.995 | 0.029 * |
grade: c2-c4 | 3 | 608 | 202.7 | 1.345 | 0.269 |
Residuals | 70 | 10551 | 150.7 |

Linear Contrast Example

H0: \( (\text{means of grades 3,4,5}) = (\text{means of grades 1,2}) \)

H1: \( (\text{means of grades 3,4,5}) \neq (\text{means of grades 1,2}) \)

```r
> summary(aov.trends,
+     split=list(grade=list(Lin=1,NonLin=2:4)))
```

Df | SS | MS | F | Pr(>F)
---|----|----|---|--------|
grade | 4 | 1361 | 340.3 | 2.258 | 0.07152 |
grade: Lin | 1 | 1270 | 1269.9 | 8.425 | 0.00495 ** |
grade: NonLin | 3 | 68 | 22.67 | 0.035 | 0.89460 |
Residuals | 70 | 10551 | 150.7 |

Trend Analysis Example

Linear contrasts can be more powerful & more appropriate tests of null hypothesis

H0: linear trend of score across grade = 0

H1: linear trend of score across grade \neq 0

<table>
<thead>
<tr>
<th>Lin</th>
<th>Quad</th>
<th>Cubic</th>
<th>Quartic</th>
</tr>
</thead>
<tbody>
<tr>
<td>(g1)</td>
<td>-0.632</td>
<td>0.535</td>
<td>-0.316</td>
</tr>
<tr>
<td>(g2)</td>
<td>-0.316</td>
<td>-0.267</td>
<td>0.632</td>
</tr>
<tr>
<td>(g3)</td>
<td>0.000</td>
<td>-0.535</td>
<td>-4.10e-16</td>
</tr>
<tr>
<td>(g4)</td>
<td>0.316</td>
<td>-0.267</td>
<td>-6.32e-01</td>
</tr>
<tr>
<td>(g5)</td>
<td>0.632</td>
<td>0.535</td>
<td>3.16e-01</td>
</tr>
</tbody>
</table>

Linear Contrasts (Comparisons)

- Contrasts allow us to evaluate focussed hypotheses
  - evaluate specific pattern of differences among group means
- Each contrast is defined by a set of contrast weights
  - weights \( c_1, c_2, \ldots, c_a \) specify a pattern of group means
- value of contrast, \( \psi \), is a weighted combination of group means
  \[
  \psi = c_1 \bar{Y}_1 + c_2 \bar{Y}_2 + c_3 \bar{Y}_3 + c_4 \bar{Y}_4 + \ldots + c_a \bar{Y}_a
  \]
Hypotheses tested with Linear Contrasts

- Linear contrasts are defined by weights
  - must sum to zero
  - \( \sum(1/2, 1/2,-1/3,-1/3,-1/3) = 0 \)
  - Multiplying weights by constant produces an equivalent linear contrast
  - \( w_1 = (1/2, 1/2,-1/3,-1/3,-1/3) \)
  - \( w_2 = 6 \times w_1 = (3,3,-2,-2,-2) \)
  - \( w_1 \) is equivalent to \( w_2 \)

Contrasts are defined by weights

each contrast sums to zero!

```r
my.weights0 <- c(3,3,-2,-2,-2)
```

Linear contrasts are defined by

\[
\begin{align*}
H_0 & : \frac{1}{2} \mu_1 + \frac{1}{2} \mu_2 - \frac{1}{3} \mu_3 - \frac{1}{3} \mu_4 - \frac{1}{3} \mu_5 = 0 \\
H_1 & : \frac{1}{2} \mu_1 + \frac{1}{2} \mu_2 - \frac{1}{3} \mu_3 - \frac{1}{3} \mu_4 - \frac{1}{3} \mu_5 \neq 0
\end{align*}
\]

```r
contrast weights
H0: \( \frac{1}{2} \mu_1 + \frac{1}{2} \mu_2 - \frac{1}{3} \mu_3 - \frac{1}{3} \mu_4 - \frac{1}{3} \mu_5 = 0 \)
H1: \( \frac{1}{2} \mu_1 + \frac{1}{2} \mu_2 - \frac{1}{3} \mu_3 - \frac{1}{3} \mu_4 - \frac{1}{3} \mu_5 \neq 0 \)
```

### Hypotheses Evaluated by a Contrast

```r
> my.weights0 <- c(3,3,-2,-2,-2)
N.B. my.weights0 is equivalent to c(1/2, 1/2, -1/3, -1/3, -1/3)
```

\[
\begin{align*}
H_0 & : 3 \mu_1 + 3 \mu_2 - 2 \mu_3 - 2 \mu_4 - 2 \mu_5 = 0 \\
& \quad = 3(\mu_1 + \mu_2) - 2(\mu_3 + \mu_4 + \mu_5) = 0 \\
& \quad = 3(\mu_1 + \mu_2) - 2(\mu_3 + \mu_4 + \mu_5) \\
& \quad = \frac{3(\mu_1 + \mu_2)}{3} - \frac{2(\mu_3 + \mu_4 + \mu_5)}{3} \\
& \quad = \frac{(\mu_1 + \mu_2)}{3} - \frac{(\mu_3 + \mu_4 + \mu_5)}{3}
\end{align*}
\]

```r
H0: \( 3 \mu_1 + 3 \mu_2 - 2 \mu_3 - 2 \mu_4 - 2 \mu_5 = 0 \)
```

### Hypotheses Evaluated by a Contrast

```r
> my.weights <- c(-1, -1, -1, -1, -1, 6)
```

\[
\begin{align*}
H_0 & : -1(\mu_1 + \mu_2 + \mu_3 + \mu_4 + \mu_5) + 6\mu_6 = 0 \\
& \quad = -1(\mu_1 + \mu_2 + \mu_3 + \mu_4 + \mu_5) + 6\mu_6 = 0 \\
& \quad = \frac{1}{6}(\mu_1 + \mu_2 + \mu_3 + \mu_4 + \mu_5)
\end{align*}
\]

```r
H0: \(-1(\mu_1 + \mu_2 + \mu_3 + \mu_4 + \mu_5) + 6\mu_6 = 0\)
```

\[
\begin{align*}
H_1 & : \mu_6 \neq \frac{1}{6}(\mu_1 + \mu_2 + \mu_3 + \mu_4 + \mu_5)
\end{align*}
\]

```r
H1: \( \mu_6 \neq \frac{1}{6}(\mu_1 + \mu_2 + \mu_3 + \mu_4 + \mu_5) \)
```
Hypotheses Evaluated by a Contrast

\[ H_0: \frac{1}{2}(\mu_1 + \mu_2) + 0 \times \mu_3 + \frac{1}{4}(\mu_4 + \mu_5 + \mu_6) = 0 \]

\[ \frac{(\mu_1 + \mu_2 + \mu_3)}{4} - \frac{(\mu_1 + \mu_2)}{2} = 0 \]

\[ \frac{(\mu_4 + \mu_5 + \mu_6)}{4} - \frac{(\mu_1 + \mu_2)}{2} = 0 \]

\[ H_1: \frac{(\mu_1 + \mu_2 + \mu_3)}{4} - \frac{(\mu_1 + \mu_2)}{2} \neq 0 \]

\[ \frac{(\mu_4 + \mu_5 + \mu_6)}{4} - \frac{(\mu_1 + \mu_2)}{2} \neq 0 \]

General Form of Linear Contrast

\[ H_0: c_1\mu_1 + c_2\mu_2 + \cdots + c_n\mu_n = \Psi = 0 \]

weighted sum of population means equals zero

\[ \sum_{j=1}^{n} c_j = 0 \]

sum of weights must equal zero

\[ SS_{\text{contrast}} = MS_{\text{contrast}} \]

\[ \Psi = \sum_{j=1}^{n} (c_j Y_j) \]

value of contrast equals weighted sum of group means

Evaluate comparison with F:

With equal n per group:

\[ F = \frac{(\Psi^2)/\sum_{j=1}^{n}(c_j^2/n_j)}{MS_W} \]

\[ F = \frac{(n\Psi^2)/\sum_{j=1}^{n}(c_j^2)}{MS_W} \]

\[ \text{df} = (1, N-n) \]

Hypotheses Evaluated by a Contrast

\[ w_1 \leftarrow c(-2,-2,0,1,1,1,1) \]

\[ (w_2 \leftarrow w_1/4) \]

\[ > \# sum(w_1) = \text{sum}(w_2) = 0 \]

\[ H_0: 3\mu_1 - 0\mu_2 - 1\mu_3 - 1\mu_4 = 0 \]

\[ 3\mu_1 - 1\mu_3 - 0\mu_4 = 0 \]

\[ 3\mu_1 = 1\mu_3 + \mu_4 \]

\[ \mu_3 = \frac{1}{3}\mu_1 + \mu_4 + \mu_1 \]

\[ H_1: \mu_3 \neq \frac{1}{3}(\mu_1 + \mu_4 + \mu_1) \]

Hypotheses tested with Linear Contrasts

2-tailed tests

\[ H_0: \frac{1}{2}\mu_1 + \frac{1}{2}\mu_2 - \frac{1}{3}\mu_3 - \frac{1}{3}\mu_4 - \frac{1}{3}\mu_5 = 0 \]

\[ H_1: \frac{1}{2}\mu_1 + \frac{1}{2}\mu_2 - \frac{1}{3}\mu_3 - \frac{1}{3}\mu_4 - \frac{1}{3}\mu_5 \neq 0 \]

1-tailed tests

\[ H_0: \frac{1}{2}\mu_1 + \frac{1}{2}\mu_2 - \frac{1}{3}\mu_3 - \frac{1}{3}\mu_4 - \frac{1}{3}\mu_5 \geq 0 \]

\[ H_1: \frac{1}{2}\mu_1 + \frac{1}{2}\mu_2 - \frac{1}{3}\mu_3 - \frac{1}{3}\mu_4 - \frac{1}{3}\mu_5 < 0 \]

\[ H_0: \frac{1}{2}\mu_1 + \frac{1}{2}\mu_2 - \frac{1}{3}\mu_3 - \frac{1}{3}\mu_4 - \frac{1}{3}\mu_5 \geq 0 \]

\[ H_1: \frac{1}{2}\mu_1 + \frac{1}{2}\mu_2 - \frac{1}{3}\mu_3 - \frac{1}{3}\mu_4 - \frac{1}{3}\mu_5 < 0 \]
General Form of Linear Contrast
(directional tests evaluated with $t$ statistic)

$$\Psi = \sum_{j=1}^{a} (c_j Y_j)$$

$$F = \frac{\left(\Psi^2\right) / \sum_{j=1}^{a} (c_j^2/n_j)}{MS_W}$$

$$t = \frac{\Psi / \sqrt{\sum_{j=1}^{a} (c_j^2/n_j)}}{\sqrt{MS_W}} \quad df = N-a$$

$$t^2 = F$$

$t$ statistic more useful for 1-tailed tests

$H_0: \frac{\mu_1 + \mu_2}{2} \geq \frac{\mu_3 + \mu_4}{3}$

$H_1: \frac{\mu_1 + \mu_2}{2} < \frac{\mu_3 + \mu_4}{3}$

$t$ statistic: the sign of weights matters!

$w = \{1/2, -1/2, 1/3, 1/3, -1/3\}$

$H_0: \frac{1}{2} \mu_1 + \frac{1}{2} \mu_2 - \frac{1}{3} \mu_3 - \frac{1}{3} \mu_4 \geq 0$

$H_1: \frac{1}{2} \mu_1 + \frac{1}{2} \mu_2 - \frac{1}{3} \mu_3 - \frac{1}{3} \mu_4 < 0$

$t$ statistic: the sign of weights determines the direction of the test

$w = \{1/2, 1/2, -1, -1, -1\}$

$H_0: \frac{1}{2} \mu_1 + \frac{1}{2} \mu_2 - \frac{1}{3} \mu_3 - \frac{1}{3} \mu_4 \geq 0$

$H_1: \frac{1}{2} \mu_1 + \frac{1}{2} \mu_2 - \frac{1}{3} \mu_3 - \frac{1}{3} \mu_4 < 0$

$t$ statistic: the sign of weights matters!

$w = \{-1/2, -1/2, 1/3, 1/3, 1/3\}$

$H_0: \frac{1}{2} \mu_1 + \frac{1}{2} \mu_2 + \frac{1}{3} \mu_3 + \frac{1}{3} \mu_4 \leq \frac{1}{2} \mu_4$

$H_1: \frac{1}{2} \mu_1 + \frac{1}{2} \mu_2 + \frac{1}{3} \mu_3 + \frac{1}{3} \mu_4 > \frac{1}{2} \mu_4$

equivalent!

**Calculating contrasts with R aov()**

```r
> contrasts(df3$grade) <- cMat
> fractions( contrasts(df3$grade) )
> contrasts(df3$grade) <- cMat

converting contrasts with aov & emmeans

conducting ANOVA with aov

Write ANOVA table with summary(), but split results for grouping variable into separate lines for different contrasts

split = list(factor.name=list(contrast.name.1=1, contrast.name.2=2, ...))

linear contrasts assessed with F tests

Store contrast weights as columns in a matrix and then assign contrast weights to grouping variable

Perform ANOVA with aov

Store contrast weights as columns in a matrix and then assign contrast weights to grouping variable

Perform ANOVA with aov

Write ANOVA table with summary(), but split results for grouping variable into separate lines for different contrasts

split = list(factor.name=list(contrast.name.1=1, contrast.name.2=2, ...))

Linear contrasts assessed with F tests

Store contrast weights as columns in a matrix and then assign contrast weights to grouping variable

Perform ANOVA with aov

Write ANOVA table with summary(), but split results for grouping variable into separate lines for different contrasts

split = list(factor.name=list(contrast.name.1=1, contrast.name.2=2, ...))

Linear contrasts assessed with F tests

Store contrast weights as columns in a matrix and then assign contrast weights to grouping variable

Perform ANOVA with aov

Write ANOVA table with summary(), but split results for grouping variable into separate lines for different contrasts

split = list(factor.name=list(contrast.name.1=1, contrast.name.2=2, ...))

Linear contrasts assessed with F tests

Store contrast weights as columns in a matrix and then assign contrast weights to grouping variable

Perform ANOVA with aov

Write ANOVA table with summary(), but split results for grouping variable into separate lines for different contrasts

split = list(factor.name=list(contrast.name.1=1, contrast.name.2=2, ...))

Linear contrasts assessed with F tests

Store contrast weights as columns in a matrix and then assign contrast weights to grouping variable

Perform ANOVA with aov

Write ANOVA table with summary(), but split results for grouping variable into separate lines for different contrasts

split = list(factor.name=list(contrast.name.1=1, contrast.name.2=2, ...))

Linear contrasts assessed with F tests

Store contrast weights as columns in a matrix and then assign contrast weights to grouping variable

Perform ANOVA with aov

Write ANOVA table with summary(), but split results for grouping variable into separate lines for different contrasts

split = list(factor.name=list(contrast.name.1=1, contrast.name.2=2, ...))

Linear contrasts assessed with F tests

Store contrast weights as columns in a matrix and then assign contrast weights to grouping variable

Perform ANOVA with aov

Write ANOVA table with summary(), but split results for grouping variable into separate lines for different contrasts

split = list(factor.name=list(contrast.name.1=1, contrast.name.2=2, ...))

Linear contrasts assessed with F tests

Store contrast weights as columns in a matrix and then assign contrast weights to grouping variable

Perform ANOVA with aov

Write ANOVA table with summary(), but split results for grouping variable into separate lines for different contrasts

split = list(factor.name=list(contrast.name.1=1, contrast.name.2=2, ...))

Linear contrasts assessed with F tests

Store contrast weights as columns in a matrix and then assign contrast weights to grouping variable

Perform ANOVA with aov

Write ANOVA table with summary(), but split results for grouping variable into separate lines for different contrasts

split = list(factor.name=list(contrast.name.1=1, contrast.name.2=2, ...))

Linear contrasts assessed with F tests
Conducting Contrasts with emmeans

emmeans = estimated marginal means

Very statisticians: Getting started with emmeans

Conducting Contrasts with linear.comparison

linear.comparison() & emmeans() yield same results

Trend Analysis

trends are linear contrasts

- the analysis of trends uses the same methods as linear contrasts
- weights are designed to evaluate specific differences across groups:
  - linear, quadratic, cubic, etc.
- weights must sum to zero
- weights can be calculated using R's contr.poly function
  - useful when differences between groups are not constant
**Trend Analysis Example**

- Trends are linear contrasts.
- `contr.poly` is used to set polynomial contrasts as default for ordered factors.
- `order()` is used to sort unique values of `df3$grade`.
- `contrasts(df3$grade) <- contr.poly(n=5,scores=1:5)` sets polynomial contrasts as default for ordered factors.
- `summary(aov.trends)` provides an ANOVA table showing the overall effect of group, which does not depend on the code used to represent the group variable.
- The results are exactly the same as the previous one. Why? Because the ANOVA table shows the overall effect of group, which does not depend on the code used to represent the group variable.

**Figure 2:** Linear, quadratic, cubic, and quartic trends.

- The results are exactly the same regardless of whether group is represented as an ordered factor.
- Now we simply do an ANOVA with our new ordered variable using `aov(score ~ grade, data=df3)`.
- The ANOVA table shows the overall effect of group, which does not depend on the code used to represent the group variable.

**Trend Analysis Example**

- Trends are linear contrasts.
- `contrasts(df3$grade) <- contr.poly(n=5,scores=1:5)` sets polynomial contrasts as default for ordered factors.
- `summary(aov.trends)` provides an ANOVA table showing the overall effect of group, which does not depend on the code used to represent the group variable.
- The results are exactly the same regardless of whether group is represented as an ordered factor.
- Now we simply do an ANOVA with our new ordered variable using `aov(score ~ grade, data=df3)`.

**Figure 2:** Linear, quadratic, cubic, and quartic trends.
## Trend Analysis Example

Trends are linear contrasts can evaluate all higher-order, nonlinear trends with a single F test

```r
> summary(aov.trends)
Call: aov(formula = trends ~ grade, data = df3)
Residuals:
   Min    1Q Median    3Q   Max
-3.750 -0.750  0.000  0.750  3.750

Coefficients: Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.000    0.785   2.567  0.0343 *
c1          1.500    0.785   1.923  0.0935 .
c2          0.500    0.785   0.641  0.5305
```

### Effect Size

Cohen’s d calculation with emmeans & linear.comparison

```r
> library(emmeans)
> aov.01 <- aov(score ~ grade, data=df3)
> sigma <- sigma(aov.01) # sqrt(MS.resid)
> edf <- df.residual(aov.01) # residual df
> aov.em <- emmeans(aov.01, specs="grade")
> myContrasts <- list(c1=myC1,c2=myC2,c3=myC3,c4=myC4)
> # Calculate Cohen's d for each contrast:
> contrast effect.size SE df lower.CL upper.CL
> c1 0.53 0.24 70 0.05 1.01
> c2 0.47 0.37 70 -0.26 1.20
> c3 0.37 0.32 70 -0.26 1.01
> c4 0.36 0.37 70 -0.37 1.09
```

---

**Trend Analysis Example**

- **trends are linear contrasts**
- Can evaluate all higher-order, nonlinear trends with a single F test

```r
> summary(aov.trends)
Call: aov(formula = trends ~ grade, data = df3)
Residuals:
   Min    1Q Median    3Q   Max
-3.750 -0.750  0.000  0.750  3.750

Coefficients: Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.000    0.785   2.567  0.0343 *
c1          1.500    0.785   1.923  0.0935 .
c2          0.500    0.785   0.641  0.5305
```

### Effect Size

Cohen’s d calculation with emmeans & linear.comparison

```r
> library(emmeans)
> aov.01 <- aov(score ~ grade, data=df3)
> sigma <- sigma(aov.01) # sqrt(MS.resid)
> edf <- df.residual(aov.01) # residual df
> aov.em <- emmeans(aov.01, specs="grade")
> myContrasts <- list(c1=myC1,c2=myC2,c3=myC3,c4=myC4)
> # Calculate Cohen’s d for each contrast:
> contrast effect.size SE df lower.CL upper.CL
> c1 0.53 0.24 70 0.05 1.01
> c2 0.47 0.37 70 -0.26 1.20
> c3 0.37 0.32 70 -0.26 1.01
> c4 0.36 0.37 70 -0.37 1.09
```
Effect Size

Cohen's d calculation with emmeans & linear.comparison

Association Strength for a Linear Comparison

\[ R^2_{alerting} = \frac{SS_{contrast}}{SS_B} \]
- Proportion of Between-Groups variation accounted for by contrast

\[ R^2_{effectsize} = \frac{SS_{contrast}}{SS_{Total}} \]
- Proportion of total variation accounted for by contrast

\[ R^2_{contrast} = \frac{SS_{contrast}}{(SS_{contrast} + SS_W)} \]
- Variation accounted for by contrast relative to the sum of contrast-variation and within-group (error) variation
- Not affected by groups that are weighted zero
- More resistant to changes in experimental design (e.g., adding or removing groups).

unequal variances
Unequal Group Variances

- So far our tests assume equal variance in different groups
- F/t tests for contrasts are not robust to violation of equal variance assumption
- When groups have unequal variances, use a different method to calculate F/t denominator, which is an estimate of population error variance
- Correcting for unequal var reduces denominator df (and, hence, power)

\[ F = \frac{(\Psi^2) / \sum_{j=1}^{a} (c_j/n_j)}{\sum_{j=1}^{a} (c_j^2/n_j) / \sum_{j=1}^{a} (c_j^2/n_j)} \]

\[ df = \frac{\left[ \sum_{j=1}^{a} (c_j^2/n_j)^2/(n_j - 1) \right]^2}{\sum_{j=1}^{a} (c_j^2/n_j)^2} \]

Contrasts with unequal variances

`linear.comparison()` can correct for unequal variances

```r
> source("http://psych.mcmaster.ca/bennett/psy710/Scripts/linear_contrast_v2.R")
> "loading function linear.comparison"
> y <- df33grade
> myContrast1 <- linear.comparison(y, g, c.weights = myContrasts, var.equal=T)
> myContrast2 <- linear.comparison(y, g, c.weights = myContrasts, var.equal=F)
> [1] "C1: F=4.995, t=2.235, p=0.029, psi=6.467, CI=(-0.367,12.568), adj.CI= (-0.952,13.887)"
> [1] "C2: F=1.662, t=1.289, p=0.202, psi=5.780, CI=(-6.176,11.175), adj.CI= (-5.714,17.274)"
> [1] "C3: F=1.396, t=1.82, p=0.04, psi=4.688, CI=(-2.544,11.719), adj.CI= (-5.966,14.542)"
> [1] "C4: F=0.977, t=0.989, p=0.326, psi=4.432, CI=(-2.955,11.819), adj.CI= (-7.062,15.926)"
```

Orthogonal Contrasts

For orthogonal contrasts, the sums of products of coefficients across groups must be zero:

- For equal n:
  \[ \sum_{j=1}^{a} c_{1j}c_{2j} = 0 \]
- For unequal n:
  \[ \sum_{j=1}^{a} c_{1j}c_{2j}/n_j = 0 \]

A set of contrasts is mutually orthogonal if all pairs of contrasts are orthogonal.

Orthogonal contrasts evaluate independent questions about group means.
**Complete Set of Mutually Orthogonal Contrasts**

If there are \(a\) groups, then the largest set of mutually orthogonal contrasts will have \((a-1)\) contrasts, and:

\[
\sum_{j=1}^{a-1} SS_{contrast,j} = SS_B
\]

- A complete set of orthogonal contrasts divides \(SS_B\) into independent pieces of variation,
- the sum of the \((a-1)\) \(SS_{contrast}\) will equal \(SS_B\),
- and the average of the contrast \(F\) values will equal the omnibus \(F\).

**Multiple Comparisons of Group Means**

**Complete set of orthogonal contrasts**

breaks \(SS_{group}\) into separate pieces

\[
\text{> cMat <- contrasts(df3$grade)}
\]

\[
\text{> fractions(cMat)}
\]

\[
\text{myC1 myC2 myC3 myC4}
\]

\[
g1 -1/2 -1 0 0
g2 -1/2 1 0 0
g3 0 1/3 0 -1
\]

\[
g4 0 1/3 0 1/2 -1
\]

\[
g5 0 1/3 0 1/2 1
\]

> if these contrasts/columns are mutually orthogonal:

> round(t(cMat) %*% cMat,digits=2)

<table>
<thead>
<tr>
<th>g1</th>
<th>g2</th>
<th>g3</th>
<th>g4</th>
<th>g5</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1/2</td>
<td>-1/2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1/2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1/3</td>
<td>0</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1/3</td>
<td>0</td>
<td>1/2</td>
<td>-1</td>
</tr>
<tr>
<td>0</td>
<td>1/3</td>
<td>0</td>
<td>1/2</td>
<td>1</td>
</tr>
</tbody>
</table>

N.B. Each element in this matrix is the sum of cross-products.

**Multiple comparisons**

\[
F(\text{at least one Type I error}) = \alpha_{FW} = 1 - (1 - \alpha_{PC})^{1/C}
\]

if \(\alpha_{PC} = 0.05\) and \(C = 100\), then \(\alpha_{FW} = 0.994\)

- Multiple comparisons inflate Type I error rate
- Generally want to control family-wise Type I error rate by adjusting the per-comparison Type I error rate
- for \(C = 100\) comparisons
  - if \(\alpha_{PC} = 0.0051\), then \(\alpha_{FW} = 0.05\)
- there are several methods for adjusting \(\alpha_{PC}\)
Controlling False Discovery Rate

• Instead of controlling $\alpha_{FW}$, control **False Discovery Rate** (FDR):
  - $Q = (\text{# of false H0 rejections}) / (\text{total # H0 rejections})$
  - FDR = Expected Value[$Q$]

• When all H0 are true, controlling $\alpha_{FW}$ and FDR are equivalent
• When some H0 are false, FDR-based methods are more powerful

Corrections for Multiple Comparisons

• Controlling $\alpha_{FW}$ by adjusting $\alpha_{PC}$:
  - Bonferroni Adjustment (aka Dunn’s Procedure)
  - Holm’s Sequential Bonferroni Test
• Controlling False Discovery Rate (FDR):
  - Benjamini & Hochberg’s (1995) Linear Step-Up Procedure (FDR)
• Relative Power: FDR > Holm’s > Bonferroni

Multiple Comparisons in R

adjust p values with `p.adjust()`

```r
> my.p.values <- c(.127,.08,.03,.032,.02,.001,.01,.005,.025)
  > sort(my.p.values)
[1] 0.001 0.005 0.010 0.020 0.025 0.030 0.032 0.080 0.127
> p.adjust(sort(my.p.values),method='bonferroni')
[1] 0.009 0.045 0.090 0.180 0.225 0.270 0.288 0.720 1.000
> p.adjust(sort(my.p.values),method='holm')
[1] 0.009 0.040 0.070 0.125 0.125 0.160 0.160
> p.adjust(sort(my.p.values),method='fdr')
[1] 0.009 0.0225 0.0300 0.04114 0.04114 0.04114 0.04114 0.090 0.127
```

Significant tests (alpha/FDR = .05) are highlighted in orange font.
N.B. Sorting p-values is not required.

Controlling Type I error rate

`p.adjust()`

```r
> aov.vp <- aov(visPref~complexity,data=df4)
> summary(aov.vp)
Df Sum Sq Mean Sq F value    Pr(>F)
complexity   4  1.271  0.3177   6.214 0.000691 ***
Residuals   35  1.790  0.0511
> summary(aov.vp,split=list(complexity=list(L=1,Q=2,C=3,q4=4)))
Df    SS         MS          F          Pr(>F)
complexity          4    1.2709   0.3177    6.214    0.000691 ***
complexity: L    1   0.7441    0.7441  14.552    0.000532 ***
complexity: Q   1   0.4357    0.4357    8.521    0.006100 **
complexity: C   1   0.0477    0.0477    0.933    0.340714
complexity: q4  1   0.0434    0.0434    0.848    0.363286
Residuals         35   1.7897    0.0511
```

Significant tests (alpha/FDR = .05) are highlighted in orange font.
Controlling Type I error rate

```r
> aov.vp <- aov(visPref~complexity,data=df4)
> summary(aov.vp)
Df Sum Sq Mean Sq F value Pr(>F)
complexity  4  1.271  0.3177  6.214 0.000691 ***
Residuals  35  1.790  0.0511
> vp.em <- emmeans(aov.vp,specs="complexity")
> # contrast(vp.em,method="poly",adjust="bonferroni")
> # contrast(vp.em,method="poly",adjust="holm")
> contrast(vp.em,method="poly",adjust="fdr")

contrast estimate   SE   df      t          p
linear          0.964  0.253 35  3.815  0.0021
quadratic   -0.873  0.299 35 -2.919  0.0122
  cubic         -0.244  0.253 35 -0.966  0.3633
quartic        0.616  0.669 35  0.921  0.3633

P value adjustment: fdr method for 4 tests
```

Setting family-wise alpha and FDR

- Generally, $\alpha_{FW}$ and FDR are set to 0.01 or 0.05
- larger $\alpha_{FW}$ may be justified for small number of orthogonal comparisons
  - Bonferroni & Holm tests may reduce power too much
  - perhaps set $\alpha_{PC}$ to 0.05 or 0.01
  - family-wise Type I error will increase but Type II error will decrease
- Note: we do this with factorial ANOVA already...

All pairwise tests (Tukey HSD)

- Tukey HSD evaluates all pairwise differences between groups
- Is more powerful than Bonferroni method (for between-subj designs)
- Tukey HSD:
  - NOT necessary to evaluate omnibus F prior to Tukey test
  - assumes equal n per group & equal variances
  - Tukey-Kramer is valid with sample sizes are unequal
  - Dunnett’s T3 test is better with unequal n & unequal variances
    [see Kirk (1995, pp. 146-50) for more details]

Tukey HSD (all pairwise differences)

optimal method for evaluating all pairwise differences

```r
> TukeyHSD(aov.vp,which="complexity")

Tukey multiple comparisons of means 95% family-wise confidence level
Fit: aov(formula = visPref ~ complexity, data = df4)

$complexity
 contrast estimate   SE  df t.ratio p.adjust
p2-p1  0.1663 0.159  3.815 0.0021
p3-p1  0.4620 0.137  3.399 0.0057
p4-p1  0.4569 0.132  3.480 0.0067
p5-p1  0.3369 0.076  3.379 0.0103
p3-p2  0.2957 0.076  3.226 0.0093
p4-p2  0.2906 0.076  3.226 0.0093
p5-p2  0.1706 0.076  3.226 0.0093
p4-p3 -0.0051 0.076 -0.066  1.0000
p5-p3 -0.1250 0.076 -1.654  0.8475
p5-p4 -0.1199 0.076 -1.606  0.8558
  

P value adjustment method: single-step
```

---

```r
> library(PMCMRplus)
> dunnettT3Test(x=df4$visPref,g=df4$complexity)

Pairwise comparisons using Dunnett’s T3 test for multiple comparisons with unequal variances

data: df4$visPref and df4$complexity

<table>
<thead>
<tr>
<th></th>
<th>p1</th>
<th>p2</th>
<th>p3</th>
<th>p4</th>
</tr>
</thead>
<tbody>
<tr>
<td>p1</td>
<td></td>
<td>0.8475</td>
<td>0.9563</td>
<td></td>
</tr>
<tr>
<td>p2</td>
<td>0.1934</td>
<td></td>
<td>0.9321</td>
<td></td>
</tr>
<tr>
<td>p3</td>
<td>0.8475</td>
<td>0.9563</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p4</td>
<td>0.8475</td>
<td>0.9563</td>
<td>0.9321</td>
<td></td>
</tr>
</tbody>
</table>

P value adjustment method: single-step alternative hypothesis: two.sided
```
Tukey HSD (all pairwise differences)
emmeans (assumes equal variances)

> summary(new G.lm.01)
# print coefficients & t-tests

> new G.lm.01 <- lm(y ~ new G)
> new G <- g;

[1] 0.002387

> c.4 vs 7 [[1]] p.2 tailed
[1] 3.149

> c.4 vs 7 [[1]] t
[1] 9.915

> c.4 vs 7 [[1]] F

[c.4 vs 7 - linear comparison (y, g, weights = my.contrast)]

> my.contrast <- list(c(0,0,0,1,0,0,-1,0));
> c.l.close (get Connection(lc.source))

> lc.source <- url("http://psyserver.mcmaster.ca/bennett/psy710/Scripts/linear_contrast_v2.R")

In the denominator and therefore be more powerful. As before, the estimated error variance likely to be more accurate, but the test will have many more degrees of freedom across groups. One advantage of this method is it uses all groups to derive an estimate of error variance, whereas between groups 4 and 7.

The results are significant (t=4.165, df=18, p-value=0.00058), so I reject the null hypothesis of no difference.

Performing a Single Comparison
After plotting data, I decide to compare means of groups 4 & 7 using a t-test:

Two Sample t-test
data: y.4 and y.7

   t = 4.165, df = 18, p-value = 0.0005813
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
13.84 42.01
sample estimates:
mean of x mean of y
111.33  83.41

Performing a Single Comparison
Next I use a linear contrast which uses all groups to derive estimate of error variance:

> my.contrast <- list(c(0,0,0,1,0,0,-1,0));
> c.l.close (get Connection(lc.source))

Loading function linear.comparison

> lc.source <- url("http://psyserver.mcmaster.ca/bennett/psy710/Scripts/linear_contrast_v2.R")
What was wrong with the preceding analyses?

**Answer:** I performed the analyses after inspecting the data and choosing to compare groups 4 & 7 because they looked different which, obviously, inflates Type I error.

---

**Planned vs. Post-hoc Comparisons**

- Previous comparisons were **planned**
- Last 2 comparisons, made after looking at data, were **post-hoc**
- Scheffe method is preferred for post-hoc linear contrasts
  - compute contrast with normal procedures
  - evaluate observed $F$ with new critical value:
    \[ F_{Scheffe} = (a-1) \times F_{(df1= a-1; \, df2 = N-a)} \]
    - $a = $ number of groups
    - $F_{(df1)}$ is the $F$ value required for desired alpha
    - $F_{Scheffe}$ is "normal" omnibus $F \times (a-1)$
  - alternatively, keep standard $F$ & adjust $p$ values using Scheffe adjustment
- Scheffe method and omnibus $F$ test are **mutually consistent**

---

**Scheffe test**

for post-hoc comparisons

These methods compute normal $F$ and adjust the $p$ value to be consistent with Scheffe method.

```r
> (con.poly <- contrast(vp.em, method="poly", adjust="none"))
contrast  estimate    SE df t.ratio p.value
linear       0.964 0.253 35   3.815  0.0005
quadratic   -0.873 0.299 35  -2.919  0.0061
cubic       -0.244 0.253 35  -0.966  0.3407
quartic      0.616 0.669 35   0.921  0.3633
> summary(con.poly, adjust="scheffe", scheffe.rank=4)
contrast  estimate    SE df t.ratio p.value
linear       0.964 0.253 35   3.815  0.0140
quadratic   -0.873 0.299 35  -2.919  0.0977
cubic       -0.244 0.253 35  -0.966  0.9178
quartic      0.616 0.669 35   0.921  0.9300
P value adjustment: scheffe method with rank 4
```

---

```r
> c1 <- c(-3, -3, 2, 2, 2)
c2 <- c(-1,1,0,0,0)
> (con1 <- contrast(vp.em, method=list(c1, c2), adjust="none"))
contrast           estimate    SE df t.ratio p.value
C(-3, -3, 2, 2, 2)    2.013 0.438 35   4.596  0.0001
C(-1, 1, 0, 0, 0)     0.166 0.113 35   1.471  0.1503
> summary(con1, adjust="scheffe", scheffe.rank=4)
contrast           estimate    SE df t.ratio p.value
C(-3, -3, 2, 2, 2)    2.013 0.438 35   4.596  0.0019
C(-1, 1, 0, 0, 0)     0.166 0.113 35   1.471  0.7068
P value adjustment: scheffe method with rank 4
```

Scheffe rank should be set to degrees of freedom for grouping factor (i.e., $a-1$)