Lab Tutorial 4: Linear Contrasts, Trend Analysis and Focused Tests

Chris M. Fiacconi

Omnibus Tests Vs. Focused Tests

In the previous tutorial we noted that a standard one-way ANOVA is an example of an *omnibus* test because it looks for *any* difference(s) among the different groups/conditions. In other words, in a 3-group design, a significant ANOVA does not tell us whether group 1 differs from group 2, group 2 differs from group 3, or group 1 differs from group 3. It simply suggests that at least one of the group means differs from one of the other group means. Although we can narrow in on where the difference lies by conducting *post-hoc* tests, such tests can be considerably less powerful than conducting a *focused* test from the start. Focused tests are sometimes referred to as *linear contrasts* because they test for specific pattern of mean differences. *Planned contrasts* are linear contrasts that are specified *a priori* based on a particular hypothesis about the pattern of differences among condition means. If possible, it is almost always better to specifiy a hypothesized pattern of group differences and test this hypothesis using a focused test. Such an approach can substantially increase your chances of detecting a true effect.

Focused Tests: Planned Contrasts

Example #1 - Equal Variances

To specify a particular pattern of group differences, we need to create a set of *contrast weights* that correspond to the pattern of means we wish to test for. Consider a drug study with 4 groups - one control group, and three treatment groups each given one of three doses of a drug. The data are depicted in Figure 1.

First, let's conduct a standard omnibus one-way ANOVA and store MS_W as a separate variable. We'll need it later.

```
options(contrasts=c("contr.sum","contr.poly")) # Set sum-to-zero contrasts

ex.data<-read.table("Ch4E11.txt",header=T) # Load in data file
str(ex.data)
```

```
## 'data.frame':    24 obs. of  2 variables:
##  $ dv  : int  9 8 7 7 5 9 7 8 7 7 ...
##  $ cond: Factor w/ 4 levels "Control","Treat1",..: 2 2 2 2 2 2 3 3 3 3 ...
```

```
ex.aov<-aov(dv~cond,data=ex.data) # Do ANOVA
print(summary(ex.aov)) # Print results
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## cond         3  10.46   3.486    2.39  0.099 .
## Residuals   20  29.17   1.458
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```
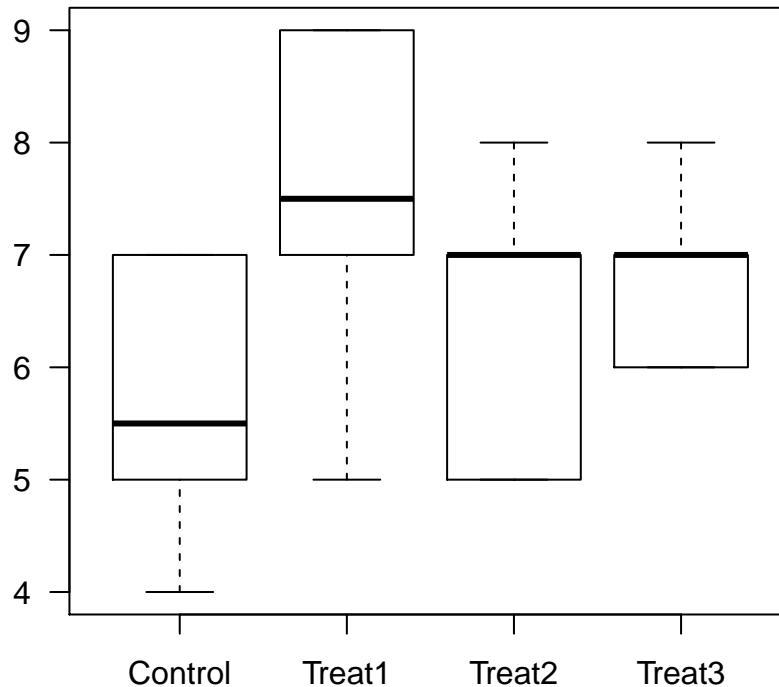
Figure 1: Boxplot of data derived from MDK Chapter 4 exercise 11

```
MS.w<-1.46
```

You'll notice that $F(3,20) = 2.39, p \approx .10$, so we fail to reject the null hypothesis that the means for all four conditions are equal. However, if we were interested in whether patients given *any* form of treatment were better off than those in the control condition, we could aggregate the scores from the three treatment conditions and test whether this aggregate score differs from the control condition. In other words, we could test the null hypothesis that:

$$H_0 : \mu_1 = \frac{1}{3}(\mu_2 + \mu_3 + \mu_4)$$

$$H_0 : \frac{1}{3}(\mu_2 + \mu_3 + \mu_4) - \mu_1 = 0$$

In contrast, the alternative hypothesis is:

$$H_a : \frac{1}{3}(\mu_2 + \mu_3 + \mu_4) - \mu_1 \neq 0$$

Note that we're testing whether the mean of the control group is equal to the *average* of the remaining treatment groups. To do this, we need to specify a set of contrast weights that capture this hypothesized pattern of mean differences. In this case the weights could be, $c_1 = -3$, $c_2 = 1$, $c_3 = 1$, $c_4 = 1$. Notice that the weight for the control group ($c_1 = -3$) is negative, as this is the group that we would expect to have a smaller mean than the other three treatment groups.

Importantly, the chosen contrast weights for each condition must sum to zero:

$$\sum_{j=1}^{k} c_j = 0$$

Once we have chosen our contrast weights, we need to test the null hypothesis using an F-test of the form:

$$F = \frac{\psi^2 / \sum_{j=1}^{k} (c_j^2 / n_j)}{MS_W}$$

The formula to calculate ψ is:

$$\psi = \sum_{j=1}^{k} c_j \bar{Y}_j$$

Importantly, the numerator of the equation to calculate F represents the sums-of-squares associated with our contrast $(SS_{Contrast})$. This means that we can re-express the equation for the F statistic as:

$$F = \frac{SS_{Contrast}}{MS_W}$$

Because the numerator df for a contrast always $= 1$, $SS_{Contrast} = MS_{Contrast}$.

Given these equations, we can now re-express the null hypothesis as:

$$H_0 : \psi = \sum_{j=1}^{k} c_j \mu_j = 0$$

Using R to perform these calculations is relatively straightforward:

```r
library(tidyverse)

# Get means for each condition
cond.means<-ex.data %>% group_by(cond) %>% summarize(Mean=mean(dv))
print(cond.means)
```

```
## # A tibble: 4 x 2
##    cond     Mean
##    <fct>    <dbl>
## 1 Control  5.67
## 2 Treat1   7.5
## 3 Treat2   6.5
## 4 Treat3   6.83
```

```r
n.pergroup<-6
my.weights<-c(-3,1,1,1) # Define contrast weights
psi<-sum(my.weights*cond.means$Mean) # Calculate psi
F<-(psi^2)/(sum((my.weights^2)/n.pergroup))/MS.w # Use MSw from omnibus ANOVA

p.value<-1-pf(F,df1=1,df2=20) # df1 = 1; df2 = N - k
print(round(p.value,digits=2))
```

```
## [1] 0.04
```

Using this focused test approach, we obtained a result, $F(1, 20) = 5.03, p = .04$, so we reject the null hypothesis that:

$$H_0 : \frac{1}{3}(\mu_2 + \mu_3 + \mu_4) - \mu_1 = 0$$

or equivalently that:

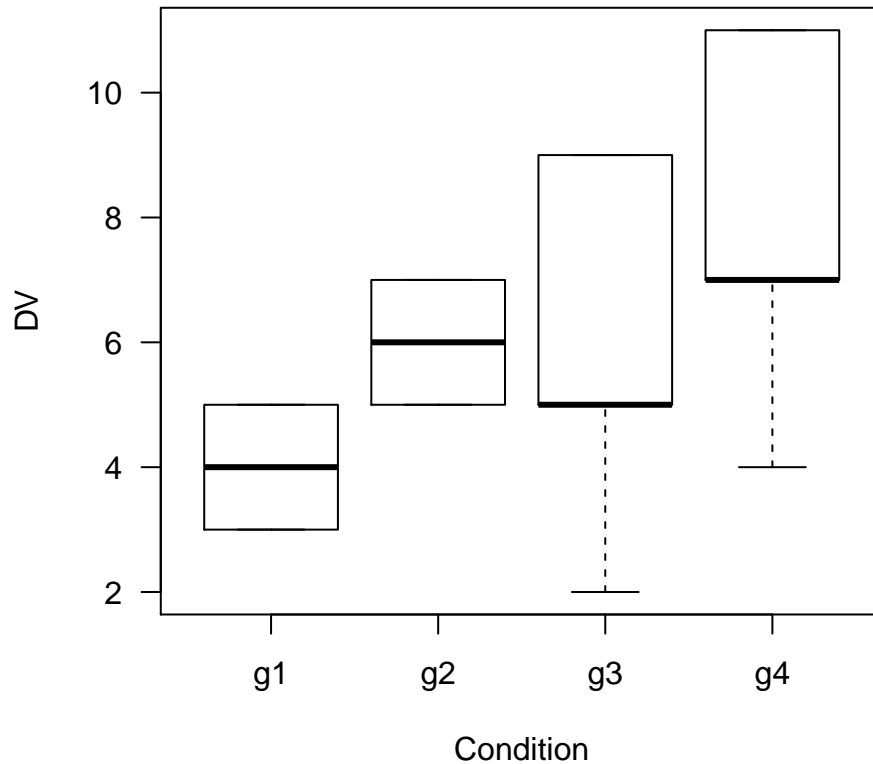$$H_0 : \psi = \sum_{j=1}^{k} c_j \mu_j = 0$$

Figure 2: Boxplot depicting data from MDK Chapter 4 Exercise 13

This example illustrates the power of focused tests: even though we failed to reject the null hypothesis using the omnibus test, once we honed in on testing a particular pattern of mean differences, we found a significant effect. Importantly, the null hypothesis that is tested with the omnibus test is not the same null hypothesis that is being tested with the focused test. Therefore, a significant focused test does not mean the same thing as a significant omnibus test. Conclusions based on the significance of a focused test pertain only to the particular contrast, defined by your contrast weights, that you tested.

Important: The formulas shown here are only correct when the homogeneity of variance assumption is met. If the groups do not have equal variance, the calculations become slightly more complex. Unlike ANOVA which is generally robust to violations of the homogeneity of variance assumption, linear contrasts are not. See your textbook for more details.

Example #2 - Unequal Variances

When the group variances are unequal, the math becomes a bit unwieldy, and doing the calculations using the basic R functions can become cumbersome. Instead we can use the **linear.comparison** function developed by Patrick J. Bennett (McMaster University) to do the heavy lifting. Consider the following dataset taken directly from Chapter 4 of your textbook. Let's load in the data file and get some descriptive statistics:

```
ex2.data<-read.table(file="Ch4E13.txt",header=T)
str(ex2.data)
```

```
## 'data.frame':    20 obs. of  2 variables:
##  $ dv  : int  3 4 5 5 3 7 5 6 5 7 ...
##  $ cond: Factor w/ 4 levels "g1","g2","g3",..: 1 1 1 1 1 2 2 2 2 2 ...
```

```
ex2.data$cond<-as.factor(ex2.data$cond) # Ensure cond variable is treated as a factor


ex2.descriptives<-ex2.data %>% group_by(cond) %>% summarize(Mean=mean(dv),S.Dev=sd(dv))
print(ex2.descriptives)
```

```
## # A tibble: 4 x 3
##   cond   Mean S.Dev
##   <fct> <dbl> <dbl>
## 1 g1        4     1
## 2 g2        6     1
## 3 g3        6     3
## 4 g4        8     3
```

From looking at the boxplot and the SDs for each group, it would appear that the group variances may not
be equal. Let's formally evaluate the null hypothesis that the group variances are in fact equal using the
bartlett.test function:

```
with(ex2.data, bartlett.test(dv~cond))
```

```
##
##  Bartlett test of homogeneity of variances
##
## data:  dv by cond
## Bartlett's K-squared = 7.4022, df = 3, p-value = 0.06013
```

Although our p-value $= .06$, which is $>.05$, we're cutting it a bit too close for comfort. If we wished to test
the null hypothesis that the mean of groups 1 and 2 is smaller than the mean of groups 3 and 4

$$H_0 : \frac{1}{2}(\mu_1 + \mu_2) = \frac{1}{2}(\mu_3 + \mu_4)$$

$$H_0 : \frac{1}{2}(\mu_3 + \mu_4) - \frac{1}{2}(\mu_1 + \mu_2) = 0$$

we need to choose the following contrast weights, $c_1 = -1$, $c_2 = -1$, $c_3 = 1$, $c_4 = 1$. Note that once again

$$\sum_{j=1}^{k} c_j = 0$$

Now let's use the **linear.comparison** function to test this contrast:

```
source("lin_comp.R") # Make sure lin_comp.R is in your working directory
ex2.weights<-c(-1,-1,1,1)

linear.comparison(y=ex2.data$dv,group=ex2.data$cond,c.weights=ex2.weights,
                  var.equal=FALSE) # Note var.equal = FALSE!!!
```

```
## [1] "computing linear comparisons assuming unequal variances among groups"
## [1] "C 1: F=4.000, t=2.000, p=0.074, psi=4.000, CI=(-0.471,8.471), adj.CI= (-0.471,8.471)"
```

```
## [[1]]
## [[1]]$contrast
## [1] -1 -1  1  1
##
## [[1]]$F
## [1] 4
##
## [[1]]$t
## [1] 2
##
## [[1]]$df1
## [1] 1
##
## [[1]]$df2
## [1] 9.756098
##
## [[1]]$p.2tailed
## [1] 0.07409955
##
## [[1]]$psi
## [1] 4
##
## [[1]]$confinterval
## [1] -0.471425  8.471425
##
## [[1]]$adj.confint
## [1] -0.471425  8.471425
##
## [[1]]$alpha
## [1] 0.05
##
## [[1]]$SS.contrast
## [1] 20
##
## [[1]]$d.effect.size
## [1] 0.8944272
##
## [[1]]$R2.alerting
## [1] 0.5
##
## [[1]]$R2.effect.size
## [1] 0.1666667
##
## [[1]]$R2.contrast
## [1] 0.8333333
```

We obtained the following results: $F(1, 9.76) = 4.00, p = .074$, so we fail to reject the null hypothesis here. Note that the denominator degrees of freedom ($df2$) has changed when assuming unequal group variances. You'll also notice that there is quite a bit of additional output from this function, including measures of effect size which we will turn to next.

Measures of Effect Size

The following sections illustrate how to compute effect size measures using the data from Example #1.

Confidence Interval for Psi

Given that the null hypothesis states that $\psi = 0$, it is sometimes useful to know the precision of our sample estimate of ψ. To capture this precision, we can calculate a 95% CI around ψ. Recall the formula for ψ:

$$\psi = \sum_{j=1}^{k} c_j \bar{Y}_j$$

The 95% CI for ψ is given as:

$$\psi \pm \sqrt{F_{.95;1,N-k}} \sqrt{MS_W \sum_{j=1}^{k} (c_j^2/n_j)}$$

Let's calculate the 95% CI on ψ using R:

```r
print(psi)
```

```
## [1] 3.833333
```

```r
critical.F<-qf(p=.95,df1=1,df2=20)
ci.interval<-sqrt(critical.F)*sqrt(MS.w*sum((my.weights^2)/n.pergroup))
print(ci.interval)
```

```
## [1] 3.564496
```

```r
psi.interval<-c(psi-ci.interval, psi+ci.interval)
print(psi.interval)
```

```
## [1] 0.2688374 7.3978292
```

Therefore, we can be reasonably confident that the true population contrast value lies between .27 and 7.40. In other words, the true population difference between the control group and the *average* of the treatment groups will likely fall within this interval.

R-squared Alerting

Another way of capturing the effect size of our linear contrast is to compute $R^2_{Alerting}$. To understand why this measure is useful, it should be noted that $SS_{Contrast} \leq SS_{Between}$. Therefore, the variability accounted for by our particular contrast of group means, $SS_{Contrast}$, can be thought of as a subset of the total variation among group means, $SS_{Between}$. The $R^2_{Alerting}$ measure indicates how much of the total between group variability can be attributed to our linear contrast:

$$R^2_{Alerting} = \frac{SS_{Contrast}}{SS_{Between}}$$

As you might expect, $R^2_{Alerting}$ can range anywhere between 0 and 1.

```r
SS.contrast<-(psi^2)/(sum((my.weights^2)/n.pergroup)) # Numerator of F-statistic
SS.between<-10.46 # Taken from omnibus ANOVA table

r2.alerting<-SS.contrast/SS.between
print(r2.alerting)
```

```
## [1] 0.7024113
```

Therefore, our linear contrast accounts for $\approx 70\%$ of the variability among condition means.

R-squared Effect Size

Although $R^2_{Alerting}$ is a very useful measure of effect size, it does not factor in within-group variability. In some situations, it is important to know the proportion of variability attributable to our contrast relative to the total variability among observations in our sample. This measure is known as $R^2_{EffectSize}$, and is calculated as shown below:

$$R^2_{EffectSize} = \frac{SS_{Contrast}}{SS_{Total}}$$

where

$$SS_{Total} = SS_{Between} + SS_{Within}$$

```r
SS.within<-29.17 # Taken from omnibus ANOVA table
SS.total<-SS.between+SS.within

r2.effectsize<-SS.contrast/SS.total
print(r2.effectsize)
```

```
## [1] 0.1853955
```

So although our contrast accounts for $\approx 70\%$ of the variability among condition means, it only accounts for $\approx 19\%$ of the total variability of observations in our sample.

Multiple Linear Contrasts and Orthogonality

Although we will cover the issues surrounding multiple comparisons in greater depth in the Lab 5 Tutorial, we will introduce the concept here briefly to illustrate the concept of *orthogonal contrasts*. Suppose you'd like to conduct a second linear contrast, which corresponds to a different hypothesis about the pattern of difference between population means. The original contrast and the new contrast are said to be orthogonal if, and only if, the sum of the product of their contrast weights equals zero:

$$\sum c_{1j}c_{2j} = 0$$

Why is this important? If a set of contrasts are not orthogonal, it means that each contrast is partially redundant with the other. In other words, each contrast does not capture a *unique* portion of the variability among condition means. Remember, it is important to limit the number of additional tests conducted because each additional test is associated with an increase in family-wise Type I error. Testing only contrasts that are orthogonal allows you to get the biggest "bang for your buck" because each test can tell you something novel about the underlying pattern of mean differences. Keep in mind though, that you shouldn't test a

contrast simply because it is orthogonal to those that you've already tested. Test only those contrasts that are theoretically interesting or informative.

When there are k groups, it is possible to define a set of $k - 1$ orthogonal contrasts. For instance, consider the following set of contrasts:

$$\psi_1 = \mu_1 - \mu_2$$

$$\psi_2 = \frac{1}{2}(\mu_1 + \mu_2) - \mu_3$$

$$\psi_3 = \frac{1}{3}(\mu_1 + \mu_2 + \mu_3) - \mu_4$$

Let's evaluate whether this set of three contrasts is orthogonal:

```
c1<-c(1,-1,0,0)
c2<-c(1,1,-2,0) # Note the scale change to make whole numbers
c3<-c(1,1,1,-3) # Note the scale change to make whole numbers

print(sum(c1*c2))
```

```
## [1] 0
```

```
print(sum(c2*c3))
```

```
## [1] 0
```

```
print(sum(c1*c3))
```

```
## [1] 0
```

Because the sum of the products of the weights in each contrast all equal 0 for each pairwise comparison, we can conclude that this set of contrasts is indeed orthogonal. Orthogonal contrast weights will be important when we discuss factorial ANOVA.

Trend Analysis: Polynomial Contrasts

So far, we've covered how to test focused hypotheses when the independent variable (IV) is categorical. This type of IV is considered **nominal**, as it is not quantitative in nature. Rather, each level represents a distinct category that does not share a quantitative relationship with the other levels. However, it is easy to imagine many different IVs where the levels *do* have an ordinal relationship to each other. For instance, consider a developmental psychologist who is interested in understanding how cognitive functioning changes with Age. This researcher might want to examine cognitive performance as a function of four different levels of Age (e.g., 4-, 6-, 8-, and 10-year old children). Note that there is an ordinal relationship between levels of the Age variable: $4 < 6 < 8 < 10$. In addition, the researcher may want to test a very particular hypothesis that cognitive functioning should increase or improve with Age. Note that an omnibus test would not provide a specific test of this hypothesis. Instead, the researcher may want to conduct a focused test that captures the hypothesized linear relationship between Age and cognitive functioning. Let's learn how to implement such a focused test using R.

In some sense, not much is new here. You might surmise that we could proceed by defining a set of contrast weights that capture the hypothesized linear relationship. You would be correct! For an experiment with $k = 4$ groups (e.g, 4-, 6-, 8-, and 10-year old children), the appropriate contrast weights could be

$c_1 = -1.5, c_2 = -0.5, c_3 = 0.5, c_4 = 1.5$. Note that the weights for a contrast that tests for a linear trend are spaced apart by equal intervals, and that:

$$\sum_{j=1}^{k} c_j = 0$$

You can also use the command **contr.poly** in R to get the appropriate contrast weights:

```
k<-4 # Number of groups in the study
p.contr.weights<-contr.poly(n=k) # Calculate coefficients for polynomial contrasts with k groups
print(p.contr.weights)
```

```
##                  .L    .Q          .C
## [1,] -0.6708204   0.5 -0.2236068
## [2,] -0.2236068  -0.5  0.6708204
## [3,]  0.2236068  -0.5 -0.6708204
## [4,]  0.6708204   0.5  0.2236068
```

Note that there are three columns in the output here. The first column contains weights for a trend analysis examining a *linear* trend in the data, whereas columns 2 and 3 represent weights for examining *quadratic* and *cubic* trends, respectively. You'll also note that the specific weights from column 1 differ from the four weights that I chose. It turns out it doesn't matter too much, so long as the weights are equally spaced. Although the specific values of the weights won't alter your F or p-values, the value of ψ does depend on the specific weights chosen. Here is a graph depicting the pattern of means that correspond to linear, quadratic, and cubic trends:

```
par(mfrow=c(2,2))
plot(x=1:4,y=p.contr.weights[,1],type="b",cex=1.5,xlab="Group",ylab="Contrast Weight",las=1,
    main="Linear Trend")
plot(x=1:4,y=p.contr.weights[,2],type="b",cex=1.5,xlab="Group",ylab="Contrast Weight",las=1,
    main="Quadratic Trend")
plot(x=1:4,y=p.contr.weights[,3],type="b",cex=1.5,xlab="Group",ylab="Contrast Weight",las=1,
    main="Cubic Trend")
```
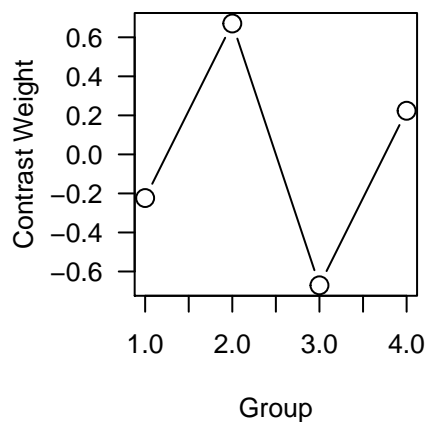
Linear Trend
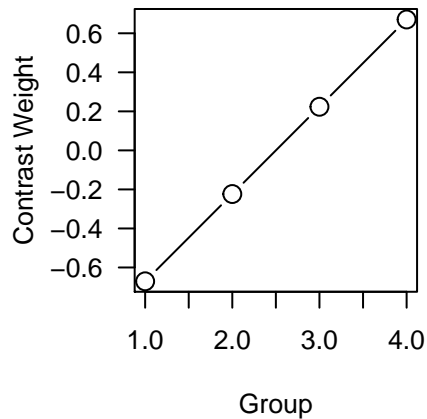
Quadratic Trend
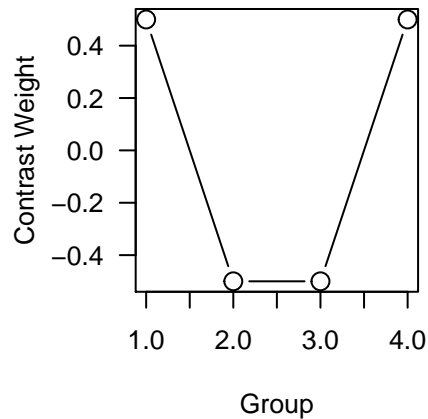
Cubic Trend

At this point, we could in principle use the **linear.comparison** function that we've been using so far in this tutorial, but it is also possible to conduct a linear trend analysis using the **aov** command. We just need to set up a few parameters first.

Let's load in the data from Chapter 6 Table 1 from your textbook and perform a linear trend analysis. In this experiment, the researcher examined how memory for words was affected by amount of study time (1-, 2-, 3-, or 4-minutes). A different group of subjects was assigned to each study time condition. The dependent variable in this study is the number of words recalled following a brief delay.

```
lt.data<-read.table(file="Ch6T1.txt",header=TRUE) # Load in data file
str(lt.data) # Look at structure of the data file
```

```
## 'data.frame':    24 obs. of  2 variables:
##  $ Recall : int  2 3 1 2 0 4 6 8 5 3 ...
##  $ Minutes: int  1 1 1 1 1 1 2 2 2 2 ...
```

```
boxplot(Recall~Minutes,data=lt.data,las=1,xlab="Study Time (min)",ylab="Words Recalled")
```
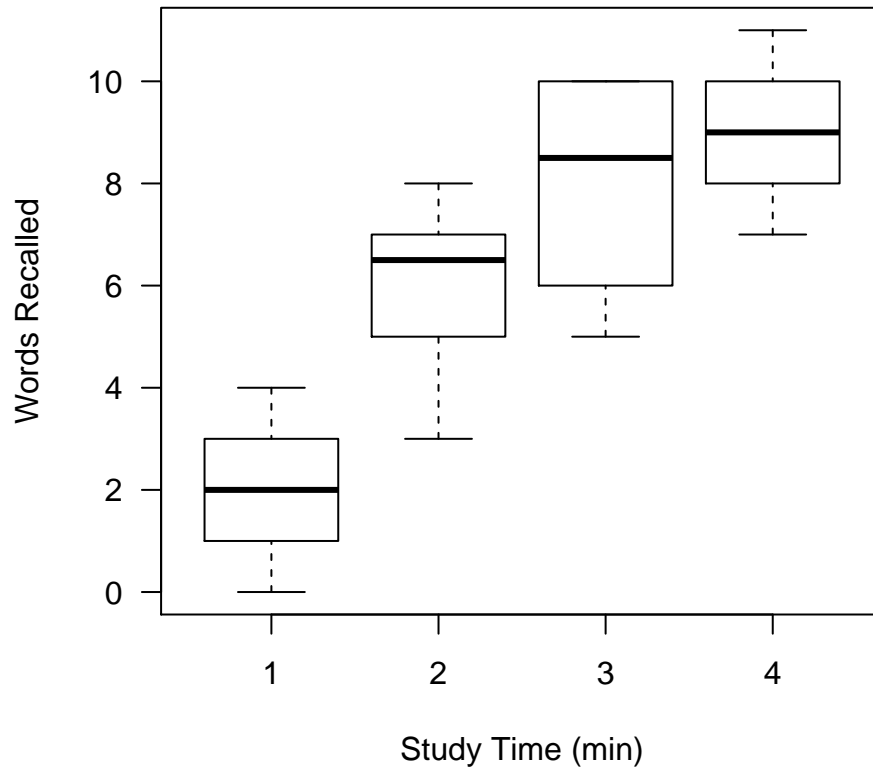
Figure 3: Boxplot of data adapted from MDK Chapter 6 Table 1

```r
# Convert Minutes variable into an ordered factor
lt.data$Minutes<-factor(lt.data$Minutes,labels="m",ordered=TRUE)
class(lt.data$Minutes) # Show variable type for Minutes
```

```
## [1] "ordered" "factor"
```

```r
unique(lt.data$Minutes) # Look at levels of the Minutes variable
```

```
## [1] m1 m2 m3 m4
## Levels: m1 < m2 < m3 < m4
```

Note how *lt.data$Minutes* is now an ordered factor. This is also shown by the sequential level ordering (m1 < m2 < m3 < m4). Now *R* knows that Minutes is an IV with an ordinal structure and will automatically generate the appropriate weights for the linear, quadratic, and cubic trends. We're now ready to use the **aov** command to examine whether there is a significant linear trend among the group means:

```r
options(contrasts=c("contr.sum","contr.poly")) # Important! Set sum-to-zero contrasts
lt.aov<-aov(Recall~Minutes,data=lt.data)

# Show ANOVA results split into linear, quadratic and cubic components
summary(lt.aov, split=list(Minutes=list(linear=1,quadratic=2,cubic=3)))
```

```
##                 Df Sum Sq Mean Sq F value   Pr(>F)
## Minutes          3  172.5    57.5  19.828 3.31e-06 ***
```

12

```
##   Minutes: linear      1  158.7   158.7  54.724 3.82e-07 ***
##   Minutes: quadratic   1   13.5    13.5   4.655   0.0433 *
##   Minutes: cubic       1    0.3     0.3   0.103   0.7511
## Residuals             20   58.0     2.9
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Not only did we get a test of the linear trend in the *lt.data* set, but *R* also shows us the results of the quadratic and cubic trend contrasts as well. As you can see, the linear trend was highly significant, but there was also a statistically significant quadratic trend among the pattern of group means.

We can also get the value of ψ for each contrast as well as its 95% CI with the commands, **coef**, and **confint**:

```
coef(lt.aov)
```

```
## (Intercept)   Minutes.L   Minutes.Q   Minutes.C
##   6.2500000   5.1429563  -1.5000000   0.2236068
```

```
confint(lt.aov)
```

```
##                  2.5 %       97.5 %
## (Intercept)   5.524896   6.97510362
## Minutes.L     3.692749   6.59316358
## Minutes.Q    -2.950207  -0.04979276
## Minutes.C    -1.226600   1.67381403
```

Also note that the linear, quadratic, and cubic contrasts are orthogonal, and therefore:

$$F_{omnibus} = \frac{1}{3}(F_{linear} + F_{quadratic} + F_{cubic})$$

and that:

$$SS_{Ominibus} = SS_{Linear} + SS_{Quadratic} + SS_{Cubic}$$

For any experiment with k groups, there will always be $k-1$ orthogonal polynomial contrasts, each with their own F and p-values.