Interaction Contrasts and their Calculation in R

Chris Williams

9/11/2020

Interaction contrasts

One approach to understanding interactions in factorial ANOVA models is to use interaction contrasts. If an orthogonal set of interaction contrasts can be selected, they can partition an overall interaction test into an independent set of contrasts, giving insight about the interaction. The approach taken here is meant to explain the ideas and present code that is similar to the code also used in SAS. There are probably easier ways to obtain these same results.

```
oex9_2 <- read.table("c:/temp/OEx92.txt",header=TRUE)

oex9_2$E <- as.factor(oex9_2$E)
oex9_2$F <- as.factor(oex9_2$F)

head(oex9_2)
```

```
##        E       F    y
## 1 lowlow  lowlow 26.1
## 2 lowlow  lowlow 27.5
## 3 lowlow lowhigh 23.5
## 4 lowlow lowhigh 21.1
## 5 lowlow highlow 22.8
## 6 lowlow highlow 23.8
```

Factorial ANOVA

You can also embed plots, for example:

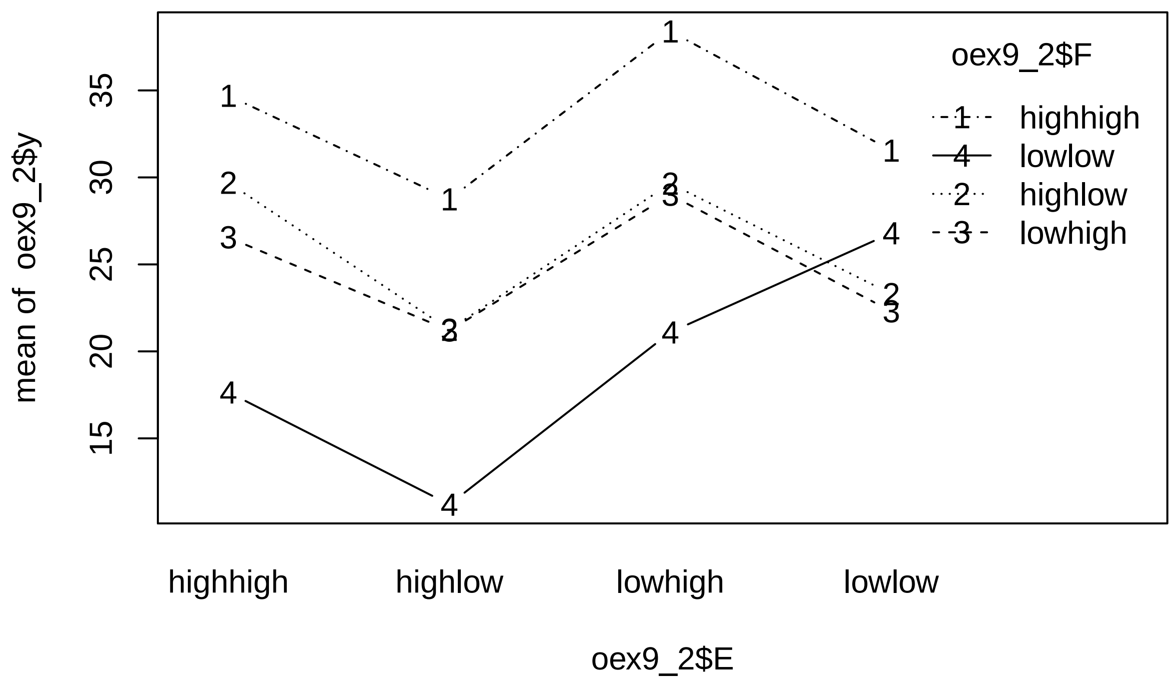
```
oex9_2.lm1  <- lm(y ~ E +F +E:F, data=oex9_2)

anova(oex9_2.lm1)
```

```
## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq F value    Pr(>F)
## E          3 342.92 114.307 111.045 6.361e-11 ***
## F          3 814.67 271.555 263.806 7.868e-14 ***
## E:F        9 209.61  23.290  22.625 1.832e-07 ***
## Residuals 16  16.47   1.029
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Examine the interaction

```
interaction.plot(oex9_2$E,oex9_2$F,
                 oex9_2$y,type="b")
```



The text calls this situation a 'one-cell interaction' meaning that a single combination mean is so different than others so that the interaction is essentially due to it. Viewing the interaction plot, we see that the 'lowlow' level of factor F is different for the level 'lowlow' of factor E compared to the other levels of E. One way to characterize this is to consider the simple main effect contrast comparing level 'lowlow' of F to the other three levels of F, at the 'lowlow' level of E:

$$\frac{\mu_{hi,hi} + \mu_{hi,lo} + \mu_{lo,hi}}{3} - \mu_{lo,lo} \text{ at } E = lo, lo$$

or

$$e(lo, lo) = \frac{\mu_{lo,lo,hi,hi} + \mu_{lo,lo,hi,lo} + \mu_{lo,lo,lo,hi}}{3} - \mu_{lo,lo,lo,lo}$$

Where e(lo,lo) is a notation reminding us that this contrast in F is at the 'lo,lo' level of E. This particular contrast at E = lo,lo looks different than at the other levels of E. We can create an interaction contrast by comparing this contrast among levels of F at E = lo,lo to an average of the value of the contrast at the other three levels of E:

$$\frac{e(hi, hi) + e(hi, lo) + e(lo, hi)}{3} - e(lo, lo)$$

If we insert the values of the means μ into the formula above and multiply by 9 to remove fractions, then the set of 16 coefficients for this contrast are:

$$(-1, -1, -1, 3, -1, -1, -1, 3, -1, -1, -1, 3, 3, 3, 3, -9)$$

We can use the emmeans package to define this contrast and perform a t-test for it.

The contrast for the empty cell

```
library(emmeans)
```

```
## Welcome to emmeans.
## NOTE -- Important change from versions <= 1.41:
##     Indicator predictors are now treated as 2-level factors by default.
##     To revert to old behavior, use emm_options(cov.keep = character(0))
```

```
lsmFE <- lsmeans(oex9_2.lm1,c('F','E'))

emptycellcontrast <- list(eccontr = c(-1,-1,-1,3,-1,-1,-1,3,-1,-1,-1,3,3,3,3,-9))

summary(contrast(lsmFE,emptycellcontrast))
```

```
##  contrast estimate   SE df t.ratio p.value
##  eccontr      -119 8.61 16 -13.834 <.0001
```

This t statistic gives the same test as the SAS F test, which can be seen by noting that the square of the t value equals the F value found from SAS.

An orthogonal set of three contrasts

It turns out that we can define two more interaction contrasts that are orthogonal and examine whether other contrasts in factor F vary between the E = lo,lo level and the other levels of factor E.

```
contrast2 <- list(c2 = c(1, 1, -2, 0, 1, 1, -2, 0, 1, 1, -2, 0, -3, -3, 6, 0))

contrast3 <- list(c3 = c(1, -1, 0, 0, 1, -1, 0, 0, 1, -1, 0, 0, -3, 3, 0, 0) )

summary(contrast(lsmFE,contrast2))
```

```
##  contrast estimate   SE df t.ratio p.value
##  c2           -1.8 6.09 16 -0.296  0.7713
```

```
summary(contrast(lsmFE,contrast3))
```

```
##  contrast estimate   SE df t.ratio p.value
##  c3           -3.5 3.51 16 -0.996  0.3341
```

The result of those two tests is that these other contrasts in factor F do not vary across the levels of E. In other words, those contrasts in F do not interact with the E factor.

We can verify that the set of 3 contrasts is orthogonal by checking the condition (also called a vector dot product) between each pair of contrasts:

```
# are the empty cell contrast and the second contrast orthogonal?
t(as.matrix(emptycellcontrast$eccontr)) %*% as.matrix(contrast2$c2)
```

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

```
# are the empty cell contrast and the third contrast orthogonal?
t(as.matrix(emptycellcontrast$eccontr)) %*% as.matrix(contrast3$c3)
```

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

```
# are the second contrast and the third contrast orthogonal?
t(as.matrix(contrast2$c2)) %*% as.matrix(contrast3$c3)
```

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

```
# the zero values for these calculations indicate that the contrasts are orthogonal
```

Interpreting the sums of squares from these three interaction contrasts

Of the original 9 degrees of freedom for the E:F interaction, these three contrasts account for the variation (and d.f.) removed when we remove the 'low,low' level of E from the data set and reanalyze the data. In other words, if we take the SS for interaction in the reduced data analysis (11.49 on 6 d.f.) and add the SS from each of these three contrasts (197.01, 0.09, and 1.02, respectively each with 1 d.f., results shown in SAS) they total to the SS for interaction in the original analysis shown above (209.61).