

Contrasts

Tim Hanson

Department of Statistics
University of South Carolina

January, 2017

Modified from originals by Gary W. Oehlert

A contrast is a linear combination of treatment means or treatment effects where the coefficients add to 0.

$$\sum_{i=1}^g w_i \mu_i = \sum_{i=1}^g w_i \alpha_i \quad \text{where} \quad \sum_{i=1}^g w_i = 0$$

Contrasts do not depend on how you parameterize the effects α_i ! When we said that important things do not depend on parameterization, we meant contrasts.

Contrasts let us look at specific ways that treatment means differ.

In this chapter we will always be considering the separate means model, i.e.

$$Y_{ij} = \mu^* + \alpha_i + \epsilon_{ij}.$$

Longevity of male fruit flies

Problem 3.2 had five treatments: no companions, one virgin companion, eight virgin companions, one pregnant companion, and eight pregnant companions. We could set up contrasts:

Question	coefficients
Companions vs no companions	-1, .25, .25, .25, .25
One companion vs eight companions	0, .5, -.5, .5, -.5
Virgin companions vs pregnant companions	0, -.5, -.5, .5, .5

Choose contrast coefficients to answer a specific question.

We want to estimate

$$\sum_{i=1}^g w_i \mu_i = \sum_{i=1}^g w_i \alpha_i$$

We just use

$$\sum_{i=1}^g w_i \hat{\mu}_i = \sum_{i=1}^g w_i \hat{\alpha}_i$$

with

$$\hat{\mu}_i = \bar{y}_{i\bullet}$$

or any of our parameterizations of the treatment effects.

Confidence intervals

The basic form of a confidence interval for a mean or mean-like parameter is

$$\text{estimate} \pm t\text{-multiplier} \times \text{standard error of estimate}$$

The standard deviation of the estimate is

$$\sqrt{\sigma^2 \sum_{i=1}^g \frac{w_i^2}{n_i}}$$

Use MS_E as an estimate of σ^2 to get the SE

$$\sqrt{MS_E \sum_{i=1}^g \frac{w_i^2}{n_i}}$$

Confidence intervals

The t-multiplier depends on a degrees of freedom, and we use the df of MS_E , our estimate of the error variance. In this case, df is $N - g$.

Putting it all together, our confidence interval is

$$\sum_{i=1}^g w_i \hat{\mu}_i \pm t_{\mathcal{E}/2, N-g} \sqrt{MS_E \sum_{i=1}^g \frac{w_i^2}{n_i}}$$

where \mathcal{E} is the error rate of the confidence interval.

```
library(cfcdae) # contains linear.contrast function, otherwise...
source("http://people.stat.sc.edu/hansont/stat506/cfcdae.R")
d=read.table("http://users.stat.umn.edu/~gary/book/fcdae.data/pr3.2",header=T)
attach(d)
ftrt=factor(trt)
f=lm(days~ftrt) # parameterization doesn't matter, here alpha1=0
linear.contrast(f,ftrt,c(-1,.25,.25,.25,.25)) # companion vs. no companion
linear.contrast(f,ftrt,c(0,.5,-.5,.5,-.5)) # one companion vs. eight companions
linear.contrast(f,ftrt,c(0,-.5,-.5,.5,.5)) # virgin companions vs. pregnant companions

cm=matrix(c(-1,.25,.25,.25,.25,0,.5,-.5,.5,-.5,0,-.5,-.5,.5,.5),5,3)
cm # matrix of contrasts (each column is contrast)
linear.contrast(f,ftrt,cm,jointF=T) # tests all three contrasts=0 simultaneously w/ F-test
linear.contrast(f,ftrt,allpairs=T) # can control FER w/ Scheffe or Bonferroni (Chap 5)
pairwise(f,ftrt) # default is Tukey HSD (Chap 5)
```

If you want to test $H_0 : \sum_{i=1}^g w_i \mu_i = \delta$, then you can form a t-test via:

$$\frac{\sum_{i=1}^g w_i \hat{\mu}_i - \delta}{\sqrt{MS_E \sum_{i=1}^g \frac{w_i^2}{n_i}}}$$

Compare this to a t-distribution with df from the MS_E (here $N - g$).

You can have one or two-sided alternatives depending on your question. Default with `linear.contrast` function is two-sided; cut p-value in half if one-sided (being careful that estimate is in the right direction!) Look at fruit fly output again...

If you want to test $H_0 : \sum_{i=1}^g w_i \mu_i = 0$ with a two-sided alternative, you can instead compute the SS for the contrast:

$$SS_w = \frac{(\sum_{i=1}^g w_i \hat{\mu}_i)^2}{\sum_{i=1}^g \frac{w_i^2}{n_i}}$$

and then compute the F-test (note that this F is the square of the analogous t-test).

$$F = \frac{SS_w}{MS_E}$$

This has 1 and $N - g$ df.

Contrasts always have one degree of freedom.

Orthogonal contrasts

Some pairs of contrasts have a special property called orthogonality. Two contrasts $\{w_i\}$ and $\{w_i^*\}$ are orthogonal if

$$\sum_{i=1}^g \frac{w_i w_i^*}{n_i} = 0$$

Assuming equal samples sizes, all of the Problem 3.2 contrast examples shown above are orthogonal to each other. The SS for each contrast add up to the total model SS...exactly.

Orthogonal contrasts produce statistically independent results.

If you have g groups, you can find $g - 1$ contrasts that are all orthogonal to each other, but you cannot find another contrast orthogonal to all of the others.

There are many sets of $g - 1$ orthogonal contrasts.

This is exactly analogous to being able to find many sets of two perpendicular lines in a plane, but not being able to find a third line perpendicular to the other two.

If you have a complete set of $g - 1$ orthogonal contrasts, then their SS partition the SS between treatments SS_M with $g - 1$ df (Pythagorean theorem again).

Use contrasts that address the questions that you have.

If these contrasts happen to be orthogonal, lucky you, you get a couple of additional minor advantages.

It is much more important to answer your questions than to achieve orthogonality for orthogonality's sake.

In days of yore, fitting polynomial models was an immense pain.

There is always a contrast that has a sum of squares equal to the “improvement” SS of a monomial in a polynomial model.

If the quantitative treatment values z_i are equally spaced and the sample sizes n_i are all the same, then the coefficients of these polynomial contrasts are simple. For example, they are tabulated in Table D.6 of the text.

Thus, when Prof. Oehlert was a younger man, these contrasts would be used to test terms in polynomial models.

Othogonal polynomial example

g	Order	Coefficients						
		1	2	3	4	5	6	7
3	1	-1	0	1				
	2	1	-2	1				
4	1	-3	-1	1	3			
	2	1	-1	-1	1			
	3	-1	3	-3	1			
5	1	-2	-1	0	1	2		
	2	2	-1	-2	-1	2		
	3	-1	2	0	-2	1		
	4	1	-4	6	-4	1		