## More on Contrasts and Orthogonality

Under some conditions, contrast sum of squares (SS) may add up to less than the treatment SS or they may add up to MORE than the treatment SS. The most satisfying condition is when they sum to equal the treatment SS. This is not necessarily a problem, as long as the contrasts are testing the hypotheses that you are interested in testing.

If we do only a few contrasts, fewer than the d.f. for the treatments, the contrast SS will probably add up to less than the treatment SS. No problem.

If we do MANY contrasts, more than the number of d.f. for treatments, the contrast SS will probably add up to more than the treatment SS. You are data-dredging? Consider a Scheffé adjustment.

If you do a number of contrasts equal to the number of treatment d.f., then the contrast SS can add up to more or less than the treatment SS. However, if the contrasts are orthogonal they will sum to exactly the treatment SS.

Contrasts are orthogonal if all their pairwise cross products sum to zero.

The cross products of a set of paired numbers is simply the product of the pairs. For example, take the following contrasts. Where the treatment levels are A1, A2, A3 and A4, write contrasts for A1 versus A2, A1 and A2 versus A3 and A4, and A3 versus A4.

These contrasts are given below.

| Contrast | a1 | a2 | a3 | a4 | Sum |
|---|---|---|---|---|---|
| a1 v a2 | –1 | 1 | 0 | 0 | 0 |
| a1&a2 v a3&a4 | –1 | –1 | 1 | 1 | 0 |
| a3 v a4 | 0 | 0 | –1 | 1 | 0 |
| **Cross products** | | | | | |
| c1 & c2 | 1 | –1 | 0 | 0 | 0 |
| c1 & c3 | 0 | 0 | 0 | 0 | 0 |
| c2 & c3 | 0 | 0 | –1 | 1 | 0 |

These contrasts are orthogonal. How about the set below?

Where the treatment levels are A1, A2, A3 and A4, write contrasts for A1 versus A2 and A3, A1 and A2 versus A3 and A4, and A3 versus A4.

If any one set of cross products do not sum to zero, the contrasts are not orthogonal. Orthogonality is a nice property, but not necessary. Write the contrasts that you want to test, orthogonal if possible.

Remember the ANOVA source table with its d.f. and Expected mean squares?

Well, a more "modern" approach involves estimating the variance components directly (PROC MIXED).

| Source | d.f. | EMS Random | EMS Fixed |
|---|---|---|---|
| Treatment | t–1 | $\sigma_\varepsilon^2 + n\sigma_\tau^2$ | $\sigma_\varepsilon^2 + n\dfrac{\Sigma \tau_i^2}{t-1}$ |
| Error | t(n–1) | $\sigma_\varepsilon^2$ | $\sigma_\varepsilon^2$ |
| Total | tn–1 | | |

Since the components are estimated directly there is no "sum of squares" for each line in the table. The model is fitted interatively (maximum likelihood).

## Traditional ANOVA table

```
                        Sum of       Mean
     Source         DF  Squares     Square  F Value  Pr > F
     Model           4  838.5976  209.6494    15.38  0.0001
     Error          20  272.6680   13.6334
     Corrected Total 24 1111.2656
```

Te results of the tests and contrasts are usually the same. However, the mixed model analysis is capable of addressing issues tht PROC GLM cannot, so when differences exist in the analysis PROC MIXED is likely to give the better result.

```
     PROC MIXED   ANOVA table
            Type 3 Tests of Fixed Effects
                     Num      Den
     Effect           DF       DF    F Value     Pr > F
     treatmnt          4     7.08      25.64     0.0003

                         Contrasts
                     Num      Den
     Label            DF       DF    F Value    Pr > F
     3 low vrs 2 high   1     12.7      21.59    0.0005
     odd vrs even       1     5.55      11.66    0.0161
     1st vrs 2nd        1     7.21      19.87    0.0027
```

## Summary

Understand the post-hoc tests. The range tests and contrasts. Be able to interpret these from SAS output.

Understand the differences between the post-hoc tests (error rates). Only one is correct for a particular objective.

Understand that contrasts are best done as *a priori* tests, and there is less concern with inflated Type I error rates if these are *a priori* tests. What is the error rate for contrasts by the way? The ANOVA was summarized. Note those aspects that I consider most important.

Understand Orthogonality.

Understand Expected mean squares. These will become extremely important in discussing larger designs. Fortunately SAS will give us the EMS (later), we need only understand them.

## The Factorial Treatment Arrangement

Also known as "two-way" ANOVA, this analysis has two (or more) treatments. For example, treatment A with two levels (a1 and a2) and treatment B with two levels (b1 and b2). The treatments are cross-classified such that each level of one treatment occurs in combination with each level of the other treatment (e.g. $a_1b_1$, $a_1b_2$, $a_2b_1$, $a_2b_2$).

Each treatment may be fixed or random (independently).

The combinations of treatments are still assigned at random to experimental units, so the design is still a CRD. For example, the 4 combinations in the example given ($a_1b_1$, $a_1b_2$, $a_2b_1$, $a_2b_2$)

would be assigned at random to the available experimental units, preferably in equal numbers to achieve a balanced design.

This treatment arrangement is called a "factorial", and the dimensions are usually given as 2 by 2 (above), 2 by 3, 3 by 3, etc. A schematic of a 3 by 3 factorial is given below.

| Treatments | A1 | A2 | A3 |
|---|---|---|---|
| B1 | a1b1 | a2b1 | a3b1 |
| B2 | a1b2 | a2b2 | a3b2 |
| B3 | a1b3 | a2b3 | a3b3 |

## Interactions

The principle treatments (A and B in the previous examples) are called main effects. The main effect for treatment A will be calculated from the marginal means or sums of the A treatment, averaged across the B treatment. Likewise, the main effect of treatment B will be calculated from the marginal means for treatment B average across the levels of A.

Marginal sums or means are used to evaluate the main effects.

| Treatments | A1 | A2 | A3 | B Means |
|---|---|---|---|---|
| B1 | a1b1 | a2b1 | a3b1 | b1 mean |
| B2 | a1b2 | a2b2 | a3b2 | b2 mean |
| B3 | a1b3 | a2b3 | a3b3 | b3 mean |
| A Means | a1 mean | a2 mean | a3 mean | |

Calculations for the main effects (Uncorrected treatment SS) are the same as for the CRD. There is however one new issue. It is possible for the same main effects to arise from various different cell patterns.
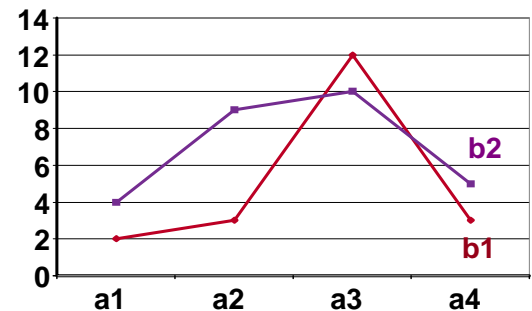
Plotting the means for the first case.

| Treatment | a1 | a2 | a3 | a4 | Means |
|---|---|---|---|---|---|
| b1 | 2 | 5 | 10 | 3 | 5 |
| b2 | 4 | 7 | 12 | 5 | 7 |
| Means | 3 | 6 | 11 | 4 | |
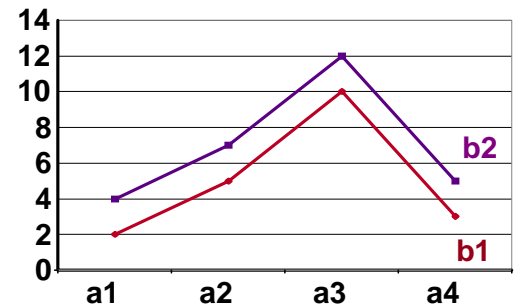


Plotting the means for the second case.

| Treatment | a1 | a2 | a3 | a4 | Means |
|---|---|---|---|---|---|
| b1 | 2 | 3 | 12 | 3 | 5 |
| b2 | 4 | 9 | 10 | 5 | 7 |
| Means | 3 | 6 | 11 | 4 | |



This lack of consistency in the cells is caused when the marginal means are not strictly additive. When additivity exists if some treatment marginal mean (#1) is larger by 2 units than some other mean (#2), each cell will in treatment #1 be 2 units higher than the corresponding mean of the treatment #2. This would represent additivity, or no interaction between the treatments.

If, however, the increases and decreases are not consistent, with the marginal means, then there is an interaction, or a lack of additivity. The marginal means (or sums) are used to calculate the main effects of the treatments. The cell to cell variation is used to measure the interaction

(after adjusting for the main effects). If we plot the treatment means, as done previously, and the lines do not appear parallel, then there is some interaction.

However, the lines are never perfectly parallel. Is the departure from additivity significant or not? To determine this, we test the interaction. This is normally done in ANOVA for all factorial designs.

### Interpreting interactions

Sometimes the main effects are very important relative to the interactions and may tell you most of what you need to know to interpret the results. Sometimes interactions can be important. Significant interactions indicate that the main effects are somehow inconsistent. You should determine how this inconsistency affects your eventual conclusions. Significant interactions should not be ignored.

## Factorial contrasts

Factorial experiments, also called two-way ANOVAs, are usually done in SAS by entering two class variables and their interaction in the model.

```
PROC GLM; CLASSES A B;
MODEL Y = A B A*B; RUN;
```

However, it is also possible to do factorials as contrasts, setting up the treatments as a one-way ANOVA. For a simple 2 by 2 factorial, with treatments A and B, we have a total of 4 cells and 3 degrees of freedom. The 4 combinations are of the treatments are $a_1b_1$, $a_2b_1$, $a_1b_2$ and $a_2b_2$. We can test the A main effect with a contrast, likewise the B main effect. To test the interaction, calculate the cross-product of the A and B contrasts.

| Treatment levels | a1b1 | a1b2 | a2b1 | a2b2 | Sum |
|---|---|---|---|---|---|
| A | –1 | –1 | 1 | 1 | 0 |
| B | –1 | 1 | –1 | 1 | 0 |
| A*B Interaction | 1 | –1 | –1 | 1 | 0 |

For larger designs the pattern is similar, for example, examine the $2 \times 2 \times 2$ factorial below. Treatment A has two levels (a and A), B has two levels (b and B) and C has levels (c and C). All contrasts consist of plus ones or minus ones, so only the + or – is shown.

| Tmt | abc | Abc | aBc | ABc | abC | AbC | aBC | ABC | Sum |
|---|---|---|---|---|---|---|---|---|---|
| A main | – | + | – | + | – | + | – | + | 0 |
| B main | – | – | + | + | – | – | + | + | 0 |
| A * B | + | – | – | + | + | – | – | + | 0 |
| C main | – | – | – | – | + | + | + | + | 0 |
| A * C | + | – | + | – | – | + | – | + | 0 |
| B * C | + | + | – | – | – | – | + | + | 0 |
| A * B * C | – | + | + | – | + | – | – | + | 0 |

A larger factorial. with more than 2 levels in some treatment, would have more than 2 d.f. in some treatment. This would require a 2 or 3 or more d.f. contrast. These can be done in SAS but we will not discuss these this semester.

## Summary

Factorials, or two-way ANOVA, was covered.

A factorial is a way of entering two or more treatments into an analysis.

The description of a factorial usually includes a measure of size, a 2 by 2, 3 by 4, 6 by 3 by 4, 2 by 2 by 2, etc.

Interactions were discussed.

Interactions test additivity of the main effects

Interactions are a measure of inconsistency in the behavior or the cells relative to the main effects.

Interactions are tested along with the main effects

Interactions should not be ignored if significant.

Factorial analyses can be done as two-way ANOVAs in SAS, or they can be done as contrasts.

# The Randomized Block Design

This analysis is similar in many ways to a "two-way" ANOVA

The CRD is defined by the linear model, $Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$. The simplest version of the CRD has one treatment and one error term. The factorial treatment arrangement discussed previously occurred within a CRD, and it had several different treatments, $Y_{ijk} = \mu + \tau_{1i} + \tau_{2j} + \tau_{1i}\tau_{2j} + \varepsilon_{ijk}$. This model has two treatments and one error. It could have many more treatments, and it would still be a factorial design. Designs having a single treatment or multiple treatments can all occur within a CRD and are referred to as different treatment arrangements.

There are other modifications of a CRD that could be done. Instead of multiple treatments we may find it necessary to subdivide the error term.

Why would we do this? Perhaps there is some variation that is not of interest. If we ignore it, that variation will go to the error term. For example, suppose we had a large agricultural experiment, and had to do our experiment in 8 different fields, or due to space limitations in a greenhouse experiment we had to separate our experiment into 3 different greenhouses or 5 different incubators. Now there is a source of variation that is due to different fields, or different greenhouses or incubators!

If we do it as a CRD, we put our treatments in the model, but if there is some variation due to field, greenhouse or incubator it will go to the error term. This would inflate our error term an make it more difficult to detect a difference (we would lose power).

How do we prevent this? First, make sure each treatment occurs in each field, greenhouse or incubator (preferably balanced). Then we would factor the new variation out of the error term by putting it in the model.

$$Y_{ijk} = \mu + \beta_i + \tau_j + \beta_i\tau_j + \varepsilon_{ijk}$$

This is not a new treatment. We will call it a BLOCK. This looks like a factorial, but it is not because the blocks are not a source of variation that we are interested in discussing.

Also, in a factorial the interaction term is likely to be something of interest. In a block design the interaction is an error term, representing random variation of experimental units across treatments.

Another difference, treatments can be either fixed or random. If both treatments are fixed, the interaction is fixed. However, blocks are usually random, and the block interaction is always random.

So why are we blocking?

It is usually used to add replication to an experiment. Additional replicates are added in another field, another greenhouse. On the one hand, the larger experiment should add power. On the other hand, if we do not take measures to keep the new variation out of the error term, we may lose power due to the larger error.

So, how does this affect our analysis?

We still have treatments with the test of treatments in the ANOVA (an F test).

We can still do post-hoc tests on the treatments.

There is only one new issue, the error term. To examine this we will need to look at the expected mean squares (EMS) for the Randomized Block Design.

## RBD EMS

We will examine two possible types of models.

In the first model we have treatments and blocks and nothing else. Each treatment occurs in each block ONCE. The experiment is similar to a factorial in some regards, but not many.

The model is

$$Y_{ij} = \mu + \beta_i + \tau_j + \varepsilon_{ij}$$

In this model the error term ($\varepsilon_{ij}$) actually comes from the block by treatment interactions ($\beta\tau_{ij}$). This is the only error available, but that is OK. It is usually a good error term because it represents random variation among the experimental units.

| Blocks \ Treatments | A1 | A2 | A3 |
|---|---|---|---|
| Block 1 | $a_1b_1$ | $a_2b_1$ | $a_3b_1$ |
| Block 2 | $a_1b_2$ | $a_2b_2$ | $a_3b_2$ |
| Block 3 | $a_1b_3$ | $a_2b_3$ | $a_3b_3$ |

This looks like a factorial.

The analysis is the same as the factorial, we get marginal sums or means and proceed to calculate the SS for blocks and treatments and "interaction" as before.

However, there is one big difference. If this was a factorial we would have Treatment A, Treatment B and the A*B interaction.

What would you use as an error term? We would not have one. A factorial ANOVA must have an error term for testing treatments and interaction.

However, since the "interaction" in a block design is assumed to be random variation among experimental units, it serves as an error term.

So the model works for Block designs.

$$Y_{ij} = \mu + \beta_i + \tau_j + \varepsilon_{ij}$$

The "interaction" term is a useful and respectable error term.