BIO4259G/BIO9915B

11. Orthogonality: Did I really say that?

Orthogonality

"Two factors are said to be orthogonal if the estimated effects for each factor are the same regardless of whether the other term is included [in the model] or not." (pg. 56)

- Blocks and treatments are orthogonal if the estimated treatment effects are the same whether or not we include or ignore the block.
  - So, why do we block?
- Practically speaking, estimates for the treatment effects are simple means when blocks and treatments are orthogonal.
  - Thank goodness we now have computers!

Orthogonality

Consider an experiment with two factors:

- Level r of Factor 1 occurs $n_{r.}$ times in total.
- Level s of Factor 2 occurs $n_{s.}$ times in total.
- Treatment combination rs is repeated n_{rs} times.

r= first factir
s=second factor.
this design will be orthogonal,

The experiment will be orthogonal if:

I50*50)/ 100
orthogonal if teratment being

$$n_{rs} = \frac{n_{r.} \times n_{.s}}{N}$$

= experimental units (total).

Where N is the total sample size.

experiment that he gave : experimental units, how many are there? 100, because 50 each.

Example:

In the example we have $n_r = n_s = 50$ for $r = 1,2$ and s=1,2.

- Implies that $N = 100$.

The design would be orthogonal if:

$$n_{rs} = \frac{50 \times 50}{100} = 25$$

for all r and s. However:

$$n_{11} = n_{22} = 49 \text{ and } n_{12} = n_{21} = 1$$

NO LONGER BALANCED OR ORTHOLOGONAL AND THAT IS WHAT IS CAUSING THE PORBLEM HERE.

Orthogonality: ANOVA

When all factors in an ANOVA are orthogonal the order in which the sums of squares are computed doesn't matter and the partitioning of the variance is unique.

When factors are not orthogonal the sums of squares depend on the order in which and the partitioning of the variance depends on the order in which we consider the variables.

WHY DOES IT ALL BREAK DWON
HOW ARE THE SUMS OF SQUARES COMPUTED?
NEXT SLIDE

Orthogonality: Sums of Squares

How are sums of squares computed?

Model Sums of Squares:

- Recall the identity:
$$TotSS = TrtSS + ResSS$$
- As we add more terms into the model $TrtSS$ increases and $ResSS$ decreases.
- The sum of squares for a particular term in the ANOVA table is equal to the increase in the $TrtSS$ (equivalently the decrease in the $ResSS$) when that term is added to the model.

Example:

In the example there are five possible models we can fit to the data:

1. Null Model: $Response = Intercept + Error$
2. Main effect of A: $Response = A + Error$
3. Main effect of B: $Response = B + Error$
4. Both main effects: $Response = A + B + Error$
5. Interaction: $Response = A * B + Error$

Sums of squares for the main effects and interaction are computed by comparing the treatment sums of squares for these models.

Example:

Model	TrtSS*	
Intercept	34.374	
A	41.953	
B	42.686	
A + B	42.689	0.166
A * B	42.855	

A+b and a gives a different value compared to the one where A is reduced frm the intercept.

*Note that the book refer to these as model sums of squares (ModSS).

Example:

If we add A to the model first then we get the following sequence of sums of squares:

Model	TrtSS*	Term SS
Intercept	34.374	
A	41.953	7.579
A + B	42.689	0.735
A * B	42.855	0.166

Example:

However, if we add B to the model first then we get the following sequence of sums of squares:

Model	TrtSS*	Term SS
Intercept	34.374	
B	42.686	8.312
A + B	42.689	0.003
A * B	42.855	0.166

Type I Sums of Squares

Sums of squares computed in this way are called the Type I Sums of Squares.

Advantage:

- Type I Sums of Squares partition the total sums of squares exactly (they sum to the overall $TrtSS$).

Disadvantages:

- The partition is not unique if the factors are not orthogonal and depends on the order in which the terms enter the model.

Type III Sums of Squares

Type II Sums of Squares are computed by comparing the models with and without a given term.

- SS(A:B)=ModSS(A*B) – ModSS(A + B)
- SS(A)=ModSS(A * B) - ModSS(B + A:B)
- SS(B)=ModSS(A * B) – ModSS(A + A:B)

Advantage:

- Type III SS do not depend on ordering.

Disadvantages:

- Type III SS do not partition the variance exactly
- Type III SS do not obey the principle of marginality.

what is it?
we have effect in the model, all marginals are in the model as well.
A and B are marginal to the interaction,

Type II Sums of Squares

Type II Sums of Squares are computed by comparing the models with and without a given term – accounting for the principle of marginality.*

- SS(A:B)=ModSS(A*B) – ModSS(A + B)
- SS(A)=ModSS(A + B) - ModSS(B)
- SS(B)=ModSS(A + B) – ModSS(A)

Advantage:

- Type II SS do not depend on ordering.
- Type II SS obey the principle of marginality.

Disadvantages:

- Type III SS do not partition the variance exactly. TYPE 2 NOT 3

* Note that this definition is not used consistently. This definition is used by the Anova() function in R and differs from the definition in the text.

Hypothesis Tests

Tests of significance can be computed from any of these sums of squares, but they answer different questions:

- Type I SS: Does the addition of a term significantly improve estimation of the treatment means given the terms already in the model?

DOES IT IMPROBE?
YES

- Type II SS: Does the addition of a term significantly improve estimate of the treatment means given all other terms, adjusting for marginality?

THREE.

NAH

- Type II SS: Does the addition of a term significantly improve estimate of the treatment means given all other terms?

BUT WE ARENT ADJUSTING FOR THE PRICNIP;E OF MARGINALITY

Comparing Type of SS

Note:

- When factors are orthogonal the different sums of squares are equivalent and there is a unique partition of the variance.
- The different sums of squares for the highest order interaction are always equal.
- Type II and III sums of squares are equivalent for any term in the model which is marginal to no other term.

Comparing Types of SS

My suggestion:

- Use Type I SS if the design is orthogonal.

- Use Type II SS if the design is non-orthogonal, but simplify the model by removing non-significant terms (according to the principle of marginality).

- Always think about what question your tests are answering.

Example

My analysis of the example data would conclude that:

1. There is no interaction between Factors A and B ($p=.434$).
2. Factor A is significant by itself ($p<.001$) but not once Factor B is in the model ($p=.920$).
3. Factor B is significant by itself ($p<.001$) but not once Factor A is in the model ($p=.101$).

The final two conclusions would lead me to note that the factors are not orthogonal (if I didn’t know this already) and their effect cannot be disentangled.

Summary

- Sums of squares are unique only if the factors are orthogonal.
- If the factors are not orthogonal then the Type I SS depend on the order in which terms enter the model.
- This is addressed by Type II and III SS, but the resulting partition of the variance is not exact.
- Always think about what you are doing!