Statistics 230

Section 4

Factorial Treatment Structure

The slides for this class are adapted from multiple sources:

- The main outline comes from *ANOVA and Mixed Models: A Short Introduction Using R*, by Lukas Meier (https://stat.ethz.ch/~meier/teaching/anova/index.html)
- *Introduction to Design and Analysis of Experiments*, by George W. Cobb
- Notes from prior semesters created by William Christensen and Dennis Tolley

Learning Outcomes

This section and the associated R examples and assignments achieve the following course expected learning outcomes:

- **Data Import:** Create datasets in R from space-, comma-, tab-delimited files
- **Summary Statistics:** Compute summary statistics from R datasets
- **Create Graphics:** Create graphics in R for exploratory data analysis and communicating results
- **Variability:** Understand the concept of variability in data and the attempt to identify sources of that variability
- **Writing Statistical Models:** Practice writing statistical models
- **Analyze Data:** Analyze data from 'Treatment-Control' or 'A/B' experiments using professional statistical software
- **Randomized Design:** Analyze data from completely randomized designs using professional statistical software
- **Two-Factor Factorial Design:** Analyze data from two-factor factorial designs using professional statistical software
- **Writing Statistical Models:** Practice writing statistical models

Remember,

- Our test statistic is generally a **signal-to-noise ratio**.
  - The treatment structure specifies the **signal**
  - The design structure specifies the **noise**

Example: ANOVA F-statistic

$$F = \frac{MSTrt}{MSE} = \frac{Treatment\ Variability}{Natural\ Variability} = \frac{Treatment\ Structure}{Design\ Structure} = \frac{Signal}{Noise}$$

Example: One-sample t-test

$$t = \frac{\bar{y} - \mu_o}{s/\sqrt{n}} = \frac{Systematic\ Variability}{Natural\ Variability} = \frac{Treatment\ Structure}{Design\ Structure} = \frac{Signal}{Noise}$$

So far, we have focused on the **completely randomized design** (CRD) for comparing means from different groups. We will stick with that ***design structure***. But we will discuss different **treatment structures**.

- Treatments are often combinations of levels of multiple factors
- ***Treatment structure*** describes the combinations of the treatments

- Remember: Our test statistic is generally a **signal-to-noise ratio**.
  - The treatment structure specifies the **signal**
  - The design structure specifies the **noise**

Consider a study where the researcher was interested in what affects the flavor of a bag of microwave popcorn.

Factor: The variable that can cause changes in the response. The factor has multiple levels (ex. “type of oil” is the factor with two levels, “canola” and “buttery”)

This study has 3 factors:

- Salt (low vs. high)
- Oil (canola vs. butter)
- Brand (orville vs. generic)

Example with 3 factors

Salt	Oil	Brand	Flavor
low	canola	orville	36
low	butter	orville	28
high	canola	orville	63
high	butter	orville	81
low	canola	generic	47
low	butter	generic	42
high	canola	generic	67
high	butter	generic	85

When multiple factors are included, there is a question about the nature of the combinations of these factors. These combinations can be described as **crossed** or **nested**.

Factor A is said to be ***crossed*** with Factor B if for every level of Factor A there are observations at every level of Factor B.

We will look at ***nested*** in a later section.

Example with 3 factors

Salt	Oil	Brand	Flavor
low	canola	orville	36
low	butter	orville	28
high	canola	orville	63
high	butter	orville	81
low	canola	generic	47
low	butter	generic	42
high	canola	generic	67
high	butter	generic	85

Example:

A lab experiment was performed to compare mountain bike tires of two different brands, 1 and 2, on three different terrains, soft, rocky and extreme. Each combination of brand and terrain was performed three times (using a new tire each time). The response was the number of kilometers traveled until the tread depth was reduced to a specified amount. The experiments were performed in random order.

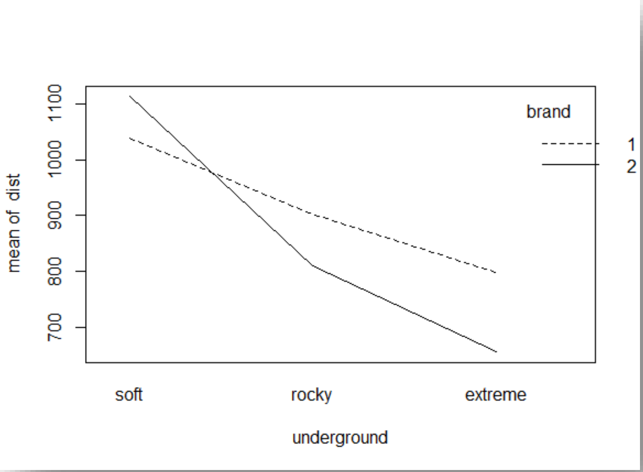
Brand	Soft	Rocky	Extreme
1	1014, 1062, 1040	884, 929, 893	824, 778, 792
2	1095, 1116, 1127	818, 794, 820	642, 633, 692

We can see from the data that Brand and Terrain are crossed factors because every level of one factor has data in every level of the other factor.

< Section 4 Two-Factors ANOVA.qmd >

Interaction plot of the mountain bike tires data:

How would you interpret the interaction shown in this graph?



Two-Way ANOVA Model

- Balanced Design implies the same number of replicates, n, for every combination of factor levels

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \epsilon_{ijk}$$

- Y_{ijk} Individual responses
- μ Overall mean response (fixed)
- α_i Fixed main effect of factor A
- β_j Fixed main effect of factor B
- $\alpha\beta_{ij}$ Fixed interaction effect between A and B
- ϵ_{ijk} Random error, $\epsilon_{ijk} \sim i.i.d.\ N(0, \sigma^2)$

Note: *i* has a levels, *j* has b levels, *k* has *n* levels

If we remove the interaction term, we have an "additive" model:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk}$$

Two-Way ANOVA Model

The expected values for the different cells in the bike tire example:

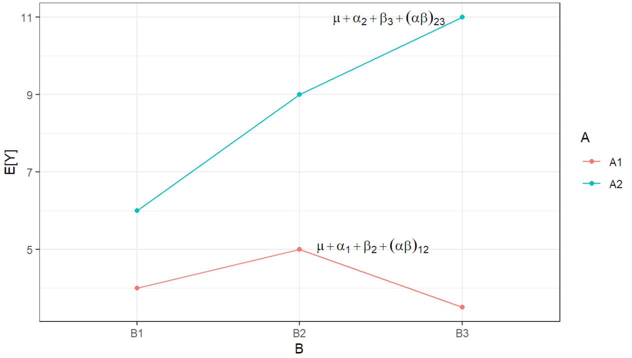
Which of the variables in the following table represents Factor A (Brand or Terrain)?

Brand	Soft	Rocky	Extreme
1	$\mu + \alpha_1 + \beta_1 + \alpha\beta_{11}$	$\mu + \alpha_1 + \beta_2 + \alpha\beta_{12}$	$\mu + \alpha_1 + \beta_3 + \alpha\beta_{13}$
2	$\mu + \alpha_2 + \beta_1 + \alpha\beta_{21}$	$\mu + \alpha_2 + \beta_2 + \alpha\beta_{22}$	$\mu + \alpha_2 + \beta_3 + \alpha\beta_{23}$

Each cell has its own mean (means model)

Interaction Plots

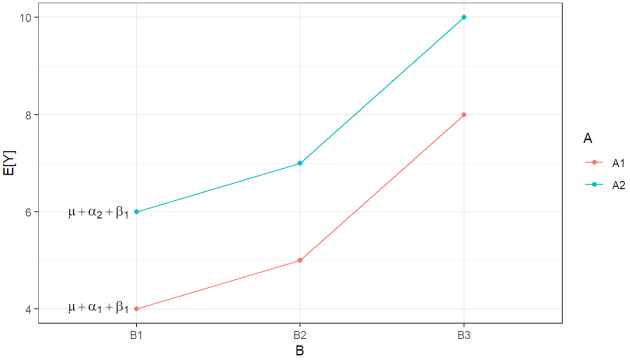
Example of an interaction plot. Anytime the lines are not PERFECTLY parallel, there is potential for interaction. However, the interaction in this graph is referred to as "ordinal interaction" because the A2 means are always higher than A1 means. Ordinal interactions are often not of any business value. Keep in mind that these plots are based on sample data. The dots (means) are typically sample means, not population means.



Interaction Plots

Example of no interaction (additive model)

- The effect of factor A does not depend on the levels of factor B
- The effect of factor B does not depend on the levels of factor A
- This picture is somewhat artificial because the lines are perfectly parallel. This rarely occurs in practice.



Parameter Estimation

Using the principle of least squares and the sum-to-zero constraints, here are the parameter estimates:

Parameter	Estimate
μ	$\widehat{\mu} = \overline{y}_{...}$
α_i	$\widehat{\alpha}_i = \overline{y}_{i..} - \overline{y}_{...}$
β_j	$\widehat{\beta}_j = \overline{y}_{.j.} - \overline{y}_{...}$
$(\alpha\beta)_{ij}$	$\widehat{(\alpha\beta)}_{ij} = \overline{y}_{ij.} - \widehat{\mu} - \widehat{\alpha}_i - \widehat{\beta}_j$

TABLE 4.4: Parameter estimates for the two-way ANOVA model.

Note that the estimate for the main effect of A ignores all aspects of the other factor (B).

Parameter Estimation

Using the principle of least squares and the sum-to-zero constraints, here are the parameter estimates:

Parameter	Estimate
μ	$\widehat{\mu} = \bar{y}_{...}$
α_i	$\widehat{\alpha}_i = \bar{y}_{i..} - \bar{y}_{...}$
β_j	$\widehat{\beta}_j = \bar{y}_{.j.} - \bar{y}_{...}$
$(\alpha\beta)_{ij}$	$\widehat{(\alpha\beta)}_{ij} = \bar{y}_{ij.} - \widehat{\mu} - \widehat{\alpha}_i - \widehat{\beta}_j$

TABLE 4.4: Parameter estimates for the two-way ANOVA model.

Also note that the expected/estimated value of the response for A level 1 and B level 2 would be

$$\hat{\mu} + \hat{\alpha}_1 + \hat{\beta}_2 + \widehat{\alpha\beta}_{12} = \bar{y}_{12.}$$

which is a cell mean (expected response for that cell)

ANOVA Model in R

- Use lm function to fit the ANOVA model

```
fit.bike <- lm(dist ~ brand * underground, data = bike)
```

- The asterisk (*) requests main effects and interactions

```
fit.bike <- lm(dist ~ brand + underground + brand:underground, data = bike)
```

- Colon (:) indicates an interaction effect. The above scripts produce equivalent results.

Table of parameter estimates

Use this table to calculate parameter estimates using the sum to zero constraint.

Brand	Soft	Rocky	Extreme	$\bar{y}_{i\cdot\cdot}$
1	1014, 1062, 1040	884, 929, 893	824, 778, 792	$\bar{y}_{1\cdot\cdot} = 912.89$
2	1095, 1116, 1127	818, 794, 820	642, 633, 692	$\bar{y}_{2\cdot\cdot} = 859.67$
$\bar{y}_{\cdot j\cdot}$	$\bar{y}_{\cdot 1\cdot} = 1075.67$	$\bar{y}_{\cdot 2\cdot} = 856.33$	$\bar{y}_{\cdot 3\cdot} = 726.83$	$\bar{y}_{\cdots} = 886.28$

TABLE 4.5: Bike data set including row-, column-wise and overall mean.

Parameter	Estimate
μ	$\widehat{\mu} = \bar{y}_{\cdots}$
α_i	$\widehat{\alpha}_i = \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdots}$
β_j	$\widehat{\beta}_j = \bar{y}_{\cdot j\cdot} - \bar{y}_{\cdots}$
$(\alpha\beta)_{ij}$	$\widehat{(\alpha\beta)}_{ij} = \bar{y}_{ij\cdot} - \widehat{\mu} - \widehat{\alpha}_i - \widehat{\beta}_j$

TABLE 4.4: Parameter estimates for the two-way ANOVA model.

Example: $\hat{\mu} = \bar{y}_{\dots} = 886.28$

Calculate $\hat{\alpha}_1, \hat{\beta}_2, \widehat{\alpha\beta}_{12}$

Table of parameter estimates

Use this table to calculate parameter estimates using the sum to zero constraint.

Brand	Soft	Rocky	Extreme	$\bar{y}_{i\cdot\cdot}$
1	1014, 1062, 1040	884, 929, 893	824, 778, 792	$\bar{y}_{1\cdot\cdot} = 912.89$
2	1095, 1116, 1127	818, 794, 820	642, 633, 692	$\bar{y}_{2\cdot\cdot} = 859.67$
$\bar{y}_{\cdot j\cdot}$	$\bar{y}_{\cdot 1\cdot} = 1075.67$	$\bar{y}_{\cdot 2\cdot} = 856.33$	$\bar{y}_{\cdot 3\cdot} = 726.83$	$\bar{y}_{\cdots} = 886.28$

TABLE 4.5: Bike data set including row-, column-wise and overall mean.

Parameter	Estimate
μ	$\widehat{\mu} = \bar{y}_{\cdots}$
α_i	$\widehat{\alpha}_i = \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdots}$
β_j	$\widehat{\beta}_j = \bar{y}_{\cdot j\cdot} - \bar{y}_{\cdots}$
$(\alpha\beta)_{ij}$	$\widehat{(\alpha\beta)}_{ij} = \bar{y}_{ij\cdot} - \widehat{\mu} - \widehat{\alpha}_i - \widehat{\beta}_j$

TABLE 4.4: Parameter estimates for the two-way ANOVA model.

$$\hat{\mu} = \bar{y}_{\dots} = 886.28$$
$$\hat{\alpha}_1 = \bar{y}_{1\cdot\cdot} - \bar{y}_{\dots} = 26.61$$
$$\hat{\beta}_2 = \bar{y}_{\cdot 2\cdot} - \bar{y}_{\dots} = -29.95$$
$$\widehat{\alpha\beta}_{12} = \bar{y}_{12\cdot} - \hat{\mu} - \hat{\alpha}_1 - \hat{\beta}_2 = 19.06$$

Compare these results to the computer output.

ANOVA Table

The ANOVA table partitions the sums of squares according to the different model components:

Source	*d.f.*	*SS*	*MS*	*F-statistic*	*p-value*
Factor A	df(A)	SS(A)	MS(A)	F(A)	p-value(A)
Factor B	df(B)	SS(B)	MS(B)	F(B)	p-value(B)
AB interaction	df(AB)	SS(AB)	MS(AB)	F(AB)	p-value(AB)
Error (within)	df(E)	SS(E)	MS(E)		
Total	df(total)	SS(total)			

ANOVA Table

The ANOVA table partitions the sums of squares according to the different model components:

Source	*d.f.*	*SS*	*MS*	*F-statistic*	*p-value*
Factor A	a-1	SS(A)	$MS(A) = \frac{SS(A)}{df_A}$	$F(A) = \frac{MS(A)}{MS(E)}$	P(f > F(A))
Factor B	b-1	SS(B)	$MS(B) = \frac{SS(B)}{df_B}$	$F(B) = \frac{MS(B)}{MS(E)}$	P(f > F(B))
AB interaction	(a-1)(b-1)	SS(AB)	$MS(AB) = \frac{SS(AB)}{df_{AB}}$	$F(AB) = \frac{MS(AB)}{MS(E)}$	P(f > F(AB))
Error (within)	ab(n-1)	SS(within)	$MS(E) = \frac{SS(within)}{df_{within}}$		
Total	abn-1	SS(total)			

The Sliding Plate Phenomenon

The treatment structure does not impact the design structure. The treatment structure dictates the nature of the treatments, and the design structure dictates the nature of the experimental units and error terms.

Source	*d.f.*
Treatments	t-1
Error (within)	t(n-1)
Total	tn-1

Source	*d.f.*
Factor A	a-1
Factor B	b-1
AB interaction	(a-1)(b-1)
Error (within)	ab(n-1)
Total	abn-1

Just like the helicopter experiment

Dr. Stats, Ph.D.

As with a sliding name plate, the treatment structure on the right fits within treatment designation on the left. Additionally, the totals of both *df* and *SS* on the right will match those on the left.

Calculations for sums of squares

Notice the patterns:

$$SS(A) = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot\cdot\cdot})^2 = \sum_{i=1}^{a} bn(\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot\cdot\cdot})^2 = \sum_{i=1}^{a} bn(\hat{\alpha}_i)^2$$

$$SS(B) = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(\bar{y}_{\cdot j\cdot} - \bar{y}_{\cdot\cdot\cdot}\right)^2 = \sum_{j=1}^{b} an\left(\bar{y}_{\cdot j\cdot} - \bar{y}_{\cdot\cdot\cdot}\right)^2 = \sum_{j=1}^{b} an\left(\hat{\beta}_j\right)^2$$

$$SS(AB) = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot j\cdot} + \bar{y}_{\cdot\cdot\cdot}\right)^2 = \sum_{i=1}^{a}\sum_{j=1}^{b} n\left(\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot j\cdot} + \bar{y}_{\cdot\cdot\cdot}\right)^2 = \sum_{i=1}^{a}\sum_{j=1}^{b} n\left(\widehat{\alpha\beta}_{ij}\right)^2$$

$$SS(Within) = SSE = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(\bar{y}_{ijk} - \bar{y}_{ij\cdot}\right)^2$$

$$SS(Total) = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(\bar{y}_{ijk} - \bar{y}_{\cdot\cdot\cdot}\right)^2$$

Calculations for sums of squares

Notice the patterns:

Overall mean

$$SS(A) = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(\bar{y}_{i..} - \bar{y}_{...})^2 = \sum_{i=1}^{a} bn(\bar{y}_{i..} - \bar{y}_{...})^2 = \sum_{i=1}^{a} bn(\hat{\alpha}_i)^2$$

Factor A level means

$$SS(B) \quad \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} \quad = \sum_{j=1}^{b} an\left(\bar{y}_{.j.} - \bar{y}_{...}\right)^2 = \sum_{j=1}^{b} an\left(\hat{\beta}_j\right)^2$$

$$SS(AB) = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...}\right)^2 = \sum_{i=1}^{a}\sum_{j=1}^{b} n\left(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...}\right)^2 = \sum_{i=1}^{a}\sum_{j=1}^{b} n\left(\widehat{\alpha\beta}_{ij}\right)^2$$

$$SS(Within) = SSE = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(\bar{y}_{ijk} - \bar{y}_{ij.}\right)^2$$

$$SS(Total) = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(\bar{y}_{ijk} - \bar{y}_{...}\right)^2$$

Hypotheses

$$H_0\text{: } \alpha\beta_{ij} = 0 \; \forall i, j$$
$$H_A\text{: } \alpha\beta_{ij} \neq 0 \; for\ at\ least\ one\ i, j$$

$$H_0\text{: } \alpha_i = 0 \; \forall i$$
$$H_A\text{: } \alpha_i \neq 0 \; for\ at\ least\ one\ i$$

$$H_0\text{: } \beta_j = 0 \; \forall j$$
$$H_A\text{: } \beta_j \neq 0 \; for\ at\ least\ one\ j$$

The test of the interaction is a test to determine if our model is additive vs. multiplicative. If there are only main effects in the model, it is considered additive. If there are interaction effects in the model, the model is considered multiplicative.

Testing Strategy

1. The interaction is tested first. If significant, we present the results in terms of the interaction between the two factors. The main effects then become less interesting.

2. If the interaction was not found to be significant, we then test the main effects. We indicate the interaction was not found to be significant and present the results in terms of the main effects.

3. Use multiple comparisons to see where differences are significant.

4. This all gets more complicated with 3+ factors.

Example: 3-way treatment structure with factors A, B, and C

- First, we examine the highest-order interaction: **ABC**. The result of the hypothesis test is not significant.
- Next, we examine all the next-highest-order interactions: **AB, AC, BC**. The results indicate that AB is significant but AC and BC are not. Now we know that Factors A and B have combined effect on the response. So, we have no need to inspect the main effects of A and B. But Factor C is not a part of a significant interaction.
- We next inspect the main effect of Factor **C** to see if it is significant. It is.
- So, **AB**-interaction and **C**-main effect are significant.
- We interpret the results in terms of **AB** and **C** and their respective levels.
- **We ignore ABC, AC, and BC** because they were not significant.
- **We ignore A and B** because their contribution is covered in the AB interaction.

< Section 4 Two-Factors ANOVA.qmd >

What can we do if there is no replication (only one measurement unit per factor combination)?

Model with interaction:

```
             Df Sum Sq Mean Sq F value Pr(>F)
prop          2 1608.6   804.3
filler        1 5280.7  5280.7
prop:filler   2 1493.6   746.8
Residuals     0    0.0
Warning message:
In anova.lm(wear.lm) :
  ANOVA F-tests on an essentially perfect fit are unreliable
```

Model without interaction:

```
          Df Sum Sq Mean Sq F value Pr(>F)
prop       2 1608.6   804.3  1.0770 0.4815
filler     1 5280.7  5280.7  7.0711 0.1171
Residuals  2 1493.6   746.8
```

How do we know if it is safe to drop the interaction term from the model?
If wrong, our tests and confidence intervals will have reduced power

Added complexities when data are unbalanced

- Sums of squares do not have unique partitions.
- We must decide how to partition the variability.
- Type I: Sequential fitting of a two-way effects model.
  - The sum of squares obtained at each step, is the reduction in the error sum of squares for including that term in the model (typical in regression analysis). The order of the main effects matters.
  - SS(A | 1)
  - SS(B | 1, A)
  - SS(AB | 1, A, B)
  - Type I is the default for the *aov* function.
- Type II: Hierarchical fitting of a two-way effects model.
  - Controlling for all other terms of the same hierarchical level. Interaction terms account for every effect that makes up that interaction.
  - SS(A | 1, B)
  - SS(B | 1, A)
  - SS(AB | 1, A, B)
  - The function *Anova* and package *car* can be used

Added complexities when data are unbalanced

- Type III: Controlling for every other term in the model.
  - SS(A | 1, B, AB)
  - SS(B | 1, A, AB)
  - SS(AB | 1, A, B)
  - We can use the *drop1* command to get Type III SS.
  - Makes comparisons for main effects and interaction effects as if they had equal weights. So, it adjusts unbalanced designs. Use when you want equal weight among levels of main effects and interactions.
- Type IV: Same as Type III when no data is missing. But makes adjustments when cells (factor levels or interaction levels) are empty.
- For designed experiments, use Type III / Type IV.

Added complexities when data are unbalanced

- Type I: Sequential fitting of a two-way effects model.
- Type II: Hierarchical fitting of a two-way effects model.
- Type III: Controlling for every other term in the model.
- Type IV: Same as Type III when no data is missing.

Examples: Which type of sums of squares should we use?

1. A study involves a **designed experiment** for a field trial. It is assumed that no data will be missing.
2. A data science problem involves modeling available data to see which if any of a number of factors and interactions effects the mean response. We wish to **first test all main effects** then all 2-way interactions and finally all 3-way interactions.
3. A factorial design has cells with unequal sample sizes. We would like to **test all effects at the same level equally** (no order of importance).
4. In a behavioral science study, we want to see the added impact of various factors on wellness. There is an inherent order of importance to these factors. Therefore, we would like to **hierarchically** test the model components in the order of their importance.

< Section 4 Sums of Squares.qmd >

More than two factors?

When we have more than two factors, we can have 3-way interactions or even 4-way interactions.

Can we interpret a 4-way interaction?

Often, we just don't include higher-order interaction terms in the model reasoning that they are likely just measuring random error.