Factorial Designs

Design of Experiments - Montgomery
Sections 5-1 - 5-3

Two Factor Analysis of Variance

- Trts often different levels of one factor
- What if interested in combinations of two factors
  - Temperature and Pressure
  - Seed variety and Fertilizer
  - Diet and Exercise Regime
- Could treat each combination as trt and do ANOVA
  - a levels of factor A and b levels of factor B
  - ab total treatments each with n observations
  - Use contrasts to study specific effects
  - Recall Problem 6 of Homework 3

Example

An experiment is conducted to study the effect of hormones injected into test rats. There are two distinct hormones (A,B) each with two distinct levels. For purposes here, we will consider this to be four different treatments labeled {A,a,B,b}. Each treatment is applied to six rats with the response being the amount of glycogen (in mg) in the liver.

Treatment	Responses					
A	106	101	120	86	132	97
a	51	98	85	50	111	72
B	103	84	100	83	110	91
b	50	66	61	72	85	60

Three contrasts are of interest. They are:

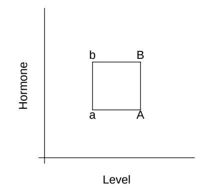
Comparison	A	a	B	b
Hormone A vs Hormone B	1	1	-1	-1
Low level vs High level	1	-1	1	-1
Equivalence of level effect	1	-1	-1	1

These contrasts are orthogonal so the SS_{Trt} is broken up into separate sums of squares. Can we redo the experiment in such a way that these sum of squares are already separated?

Factorial Experiment

- Break up trts into the two factors (two levels each)
- Known as a 2^2 factorial
  - Factor A: Hormone (A or B)
  - Factor B: Level (L or H)
- Factorial - investigates all combinations of factors
- Single replicate of factorial involves ab trials
- Design often illustrated as table or graphically

Hormone	Level: Low	Level: High
A	XXXXXX	XXXXXX
B	XXXXXX	XXXXXX



Statistical Model

- Statistical model is

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk} \quad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, n \end{cases}$$

  μ - grand mean

  τ_i - ith level effect of factor A (ignores B)

  β_j - jth level effect of factor B (ignores A)

  $(\tau\beta)_{ij}$ - interaction effect of combination ij

  Explains variation not described by main effects

  $\epsilon_{ijk} \sim \text{N}(0, \sigma^2)$

- Over-parameterized model.

- Must include $a + b + 1$ model constraints. Typically

$$\sum_i \tau_i = 0 \quad \sum_j \beta_j = 0 \quad \sum_i (\tau\beta)_{ij} = 0 \quad \sum_j (\tau\beta)_{ij} = 0$$

Model Estimates

- Previous constraints result in estimates

  $\widehat{\mu} = \overline{y}_{...}$

  $\widehat{\tau}_i = \overline{y}_{i..} - \overline{y}_{...}$

  $\widehat{\beta}_j = \overline{y}_{.j.} - \overline{y}_{...}$

  $\widehat{(\tau\beta)}_{ij} = \overline{y}_{ij.} - \overline{y}_{i..} - \overline{y}_{.j.} + \overline{y}_{...}$

- The predicted value is ij combination average, so

  $\widehat{y}_{ijk} = \overline{y}_{ij.}$ and $\widehat{\epsilon}_{ijk} = y_{ijk} - \overline{y}_{ij.}$

- NOTE: From model (ignoring parameter constraints)

$$\begin{aligned} \overline{y}_{i..} &= \mu + \tau_i + \overline{\beta}_. + \overline{(\tau\beta)}_{i.} \\ \overline{y}_{i..} - \overline{y}_{i'..} &= \tau_i - \tau_{i'} + \overline{(\tau\beta)}_{i.} - \overline{(\tau\beta)}_{i'.} \end{aligned}$$

  Difference in trt effects depends on $(\tau\beta)$ constraints

  $\tau_i - \tau_{i'}$ is non-estimable (if interaction)

- Using typical constraints,

$$\begin{aligned} \overline{y}_{i..} &= \mu + \tau_i \\ \overline{y}_{i..} - \overline{y}_{i'..} &= \tau_i - \tau_{i'} \end{aligned}$$

- Caution if interaction present

- Should always test interaction first

Partitioning the Sum of Squares

- Rewrite observation as:

$$y_{ijk} = \overline{y}_{...} + (\overline{y}_{i..} - \overline{y}_{...}) + (\overline{y}_{.j.} - \overline{y}_{...}) + (\overline{y}_{ij.} - \overline{y}_{i..} - \overline{y}_{.j.} + \overline{y}_{...}) + (y_{ijk} - \overline{y}_{ij.})$$

- Can look at $\sum (y_{ijk} - \overline{y}_{...})^2$

- Right hand side simplifies to

$$\begin{aligned} &bn \sum_i (\overline{y}_{i..} - \overline{y}_{...})^2 + \\ &an \sum_j (\overline{y}_{.j.} - \overline{y}_{...})^2 + \\ &n \sum_i \sum_j (\overline{y}_{ij.} - \overline{y}_{i.} - \overline{y}_{.j} + \overline{y}_{..})^2 + \ldots \end{aligned}$$

- $\text{SS}_\text{A} + \text{SS}_\text{B} + \text{SS}_\text{AB} + \text{SS}_\text{E}$

- Under normality, all SS/σ^2 independent

Hypothesis Testing

- Can show: Fixed Case

  $\text{E}(\text{MS}_\text{E}) = \sigma^2$

  $\text{E}(\text{MS}_\text{A}) = \sigma^2 + bn \sum \tau_i^2 / (a-1)$

  $\text{E}(\text{MS}_\text{B}) = \sigma^2 + an \sum \beta_j^2 / (b-1)$

  $\text{E}(\text{MS}_\text{AB}) = \sigma^2 + n \sum (\tau\beta)_{ij}^2 / (a-1)(b-1)$

- Use F-test to test equality of A, B, and AB effects

$$F_0 = \frac{\text{SS}_\text{A}/(a-1)}{\text{SS}_\text{E}/(ab(n-1))}$$

$$F_0 = \frac{\text{SS}_\text{B}/(b-1)}{\text{SS}_\text{E}/(ab(n-1))}$$

$$F_0 = \frac{\text{SS}_\text{AB}/(a-1)(b-1)}{\text{SS}_\text{E}/(ab(n-1))}$$

Analysis of Variance Table

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Factor A	SS_A	$a-1$	MS_A	F_0
Factor B	SS_B	$b-1$	MS_B	F_0
Interaction	SS_{AB}	$(a-1)(b-1)$	MS_{AB}	F_0
Error	SS_E	$ab(n-1)$	MS_E	
Total	SS_T	$abn-1$		

$$SS_T = \sum\sum y_{ijk}^2 - y_{...}^2/abn$$

$$SS_A = \frac{1}{bn}\sum y_{i..}^2 - y_{...}^2/abn$$

$$SS_B = \frac{1}{an}\sum y_{.j.}^2 - y_{...}^2/abn$$

$$SS_{Sub} = \frac{1}{n}\sum\sum y_{ij.}^2 - y_{...}^2/abn$$

$$SS_{AB} = SS_{Sub} - SS_A - SS_B$$

SS_E=Subtraction

$df_E > 0$ only if $n > 1$. When $n = 1$, cannot separate interaction from error (confounded). Recall typical RCBD uses $n = 1$. Assuming no interaction allows us to estimate error and test for treatment differences.

Rat Hormone Example

$\sum\sum\sum y_{ijk} = 2074$ and $\sum\sum\sum y_{ij}^2 = 191022$

$y_{12.} = 642$, $y_{22.} = 571$, $y_{11.} = 467$ and $y_{21.} = 394$

$SS_T = 191022 - 2074^2/24 = 11793.83$

$SS_A = (1109^2 + 965^2)/12 - 2074^2/24 = 864.00$

$SS_B = (861^2 + 1213^2)/12 - 2074^2/24 = 5162.67$

$SS_{Sub} = (642^2 + 467^2 + 571^2 + 394^2)/6 - 2074^2/24 = 6026.83$

$SS_{AB} = 6026.83 - 5162.67 - 864.00 = .16$

$SS_E = 11793.83 - 6026.83 = 5767.00$

$F_0^{AB} = (.16/1)/(5767/20) \approx 0.0$ (Not Significant)

$F_0^{A} = (864/1)/(5767/20) \approx 3.0$ (Not Significant)

$F_0^{B} = (5162.67/1)/(5767/20) \approx 17.9$ (Significant)

Example : Comparing Factors

- The mean at each combination is

Level	Hormone A	Hormone B
L	77.83	65.67
H	107.00	95.17

- Can look at one factor by averaging out other
- Only interpretable if no interaction
- Comparing Hormone A to Hormone B
  - Average out the level

$$\frac{(77.83 + 107.00)}{2} - \frac{(65.67 + 95.17)}{2} \approx 12.0$$

- Comparing Low Level to High Level
  - Average out the hormone

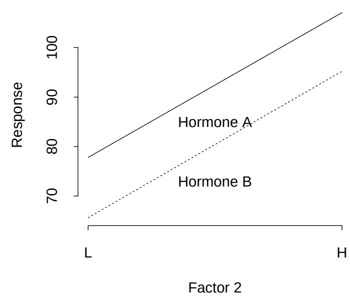
$$\frac{(77.83 + 65.67)}{2} - \frac{(107.00 + 95.17)}{2} \approx -29.335$$

Two Factor Experiment

- Interaction
  - Diff in response of one factor not constant over other factor levels
  - Diff in level of Hormone A vs Diff in level of Hormone B

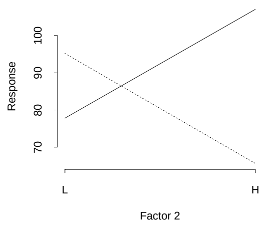
$$(107.00 - 77.83) - (95.17 - 65.67) \approx -0.33$$

- Interaction Plot
  - Can view changes in response difference
  - Are lines parallel or not?

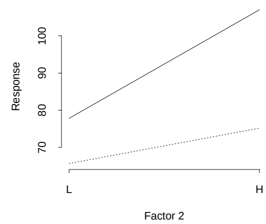


Two Factor Experiment

- Completely opposite behavior (no Factor 2 effect?)



- Increase but not same amount (still Factor 2 effect?)



Regression/Response Surface Approach

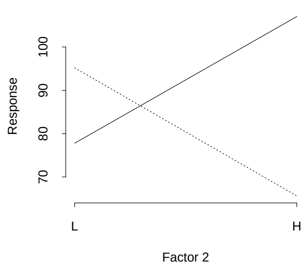
$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \epsilon$$

- $x_1 = 1$ if Hormone A and $x_1 = -1$ if Hormone B
- $x_2 = 1$ if High level and $x_2 = -1$ if Low Level
- Trt A: E(y)=$\beta_0 + \beta_1 + \beta_2 + \beta_{12}$, Trt a: E(y)=$\beta_0 + \beta_1 - \beta_2 - \beta_{12}$
- Trt B: E(y)=$\beta_0 - \beta_1 + \beta_2 - \beta_{12}$, Trt b: E(y)=$\beta_0 - \beta_1 - \beta_2 + \beta_{12}$

- β_1 estimates Hormone effect
- β_2 estimates Level effect
- β_{12} estimates Interaction
- Response surface described by $-1 \leq x_i \leq 1$
- If β_{12}=0 then surface is a plane (additive model)
- If $\beta_{12} \neq 0$ then surface "curved" (pgs 172-173)
- Recall non-additivity residual plot diagnostic

Interaction in Two Factor Experiment

- If interaction, how to interpret main effects
  - In regression, if $\beta_{12} \neq 0$, leave x_1 and x_2 in model
  - Main effect result depends on level of other factor
  - Cannot average over other factor
  - Compare $\overline{y}_{ij.}$'s instead
  - For example, if opposite behavior, effect will cancel out



  - May still be able to discuss main effect (but not estimate)
  - Sometimes interaction due to only a few combos

Hidden Replication Feature: Main Effects

- One factor approach
  - Can estimate main effects using 3 combinations (A, a, b)
  - Main effect Level: $y_A - y_a$, Main effect Hormone: $y_a - y_b$
  - Must replicate to have variance estimate
  - Cannot estimate interaction without 4th combination
  - If $n = 2$ so $N = 6$, Var(Effect) = $2\sigma^2/2 = \sigma^2$
- Factorial approach
  - Estimate effects using $N = 4$ observations (A, a, B, b)
  - Have variance estimate if no interaction
  - Main effect Hormone: $.5(y_A + y_a) - .5(y_B + y_b)$
  - Var(Effect) = σ^2
  - Replication provides ability to estimate interaction

Factorial gives same accuracy with less ($N = 6$ vs $N = 4$)