STAT 5000

Statistical Methods I

Week 9

Fall 2024

Dr. Danica Ommen

Unit 2

Two-Way ANOVA: Effects

Larvae Example: Examine the effects of different concentrations of copper and zinc in water on the ability of minnow larvae to produce protein

- Factor A: Concentration of copper (0 or 150 ppm)
- Factor B: Concentration of zinc (0, 750 or 1500 ppm)
- Treatments: All 6 combination of 2 levels of copper and 3 levels of zinc (complete/full factorial treatment design)
- Experimental units: 12 water tanks containing minnow larvae
- Experimental design: CRD, experimental units are randomly assigned to the 6 treatments with 2 units per treatment
- Response Variable: protein content (μg/tank)

Effects Model: Larvae Example

$$Y_{ijk} = \mu + \alpha_i + \tau_j + (\alpha\tau)_{ij} + \epsilon_{ijk}$$

$$\begin{bmatrix} Y_{111} \\ Y_{112} \\ Y_{121} \\ Y_{122} \\ Y_{131} \\ Y_{132} \\ Y_{211} \\ Y_{212} \\ Y_{221} \\ Y_{222} \\ Y_{231} \\ Y_{232} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \tau_1 \\ \tau_2 \\ \tau_3 \\ (\alpha\tau)_{11} \\ (\alpha\tau)_{12} \\ (\alpha\tau)_{13} \\ (\alpha\tau)_{21} \\ (\alpha\tau)_{22} \\ (\alpha\tau)_{23} \end{bmatrix} + \begin{bmatrix} \epsilon_{111} \\ \epsilon_{112} \\ \epsilon_{121} \\ \epsilon_{122} \\ \epsilon_{131} \\ \epsilon_{132} \\ \epsilon_{211} \\ \epsilon_{212} \\ \epsilon_{221} \\ \epsilon_{222} \\ \epsilon_{231} \\ \epsilon_{232} \end{bmatrix}$$

Effects Model

- Estimate $a \times b$ Treatment means with
  - μ
  - a effects for Factor A
  - b effects for Factor B
  - $a \times b$ interaction effects for Factors A and B
- Impose constraints on main effects and interaction effects to reduce number of parameters to $a \times b$

Baseline Constraints Effects Model

$$Y_{ijk} = \mu + \alpha_i + \tau_j + (\alpha\tau)_{ij} + \epsilon_{ijk}$$

- Set $\alpha_a = 0$ (level a of Factor A)
- Set $\tau_b = 0$ (level b of Factor B)
- Set $(\alpha\tau)_{aj} = 0$ for all $j = 1, \ldots, b$
  (All interaction effects with level a of Factor A)
- Set $(\alpha\tau)_{ib} = 0$ for all $i = 1, \ldots, a$
  (All interaction effects with level b of Factor B)

Baseline Constraints Effects Model

- $\mu = \mu_{ab}$
  mean response at level *a* of Factor A and level *b* of Factor B
- $\alpha_i = \mu_{ib} - \mu_{ab}$
  simple effect of the i^{th} level of Factor A when Factor B is at level *b*
- $\tau_j = \mu_{aj} - \mu_{ab}$
  simple effect of the j^{th} level of Factor B when Factor A is at level *a*
- $(\alpha\tau)_{ij} = \mu_{ij} - \mu_{ib} - \mu_{aj} + \mu_{ab}$
  interaction effect

Baseline Constraints Effects Model: Larvae Example

$$\begin{bmatrix} Y_{111} \\ Y_{112} \\ Y_{121} \\ Y_{122} \\ Y_{131} \\ Y_{132} \\ Y_{211} \\ Y_{212} \\ Y_{221} \\ Y_{222} \\ Y_{231} \\ Y_{232} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \tau_1 \\ \tau_2 \\ (\alpha\tau)_{11} \\ (\alpha\tau)_{12} \end{bmatrix} + \begin{bmatrix} \epsilon_{111} \\ \epsilon_{112} \\ \epsilon_{121} \\ \epsilon_{122} \\ \epsilon_{131} \\ \epsilon_{132} \\ \epsilon_{211} \\ \epsilon_{212} \\ \epsilon_{221} \\ \epsilon_{222} \\ \epsilon_{231} \\ \epsilon_{232} \end{bmatrix}$$

Baseline Constraints Effects Model: Larvae Example

Least squares estimate of β is given by

$$
\begin{aligned}
\hat{\beta} &= (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y} \\
&= \begin{pmatrix}
\bar{Y}_{23\cdot} \\
\bar{Y}_{13\cdot} - \bar{Y}_{23\cdot} \\
\bar{Y}_{21\cdot} - \bar{Y}_{23\cdot} \\
\bar{Y}_{22\cdot} - \bar{Y}_{23\cdot} \\
\bar{Y}_{11\cdot} - \bar{Y}_{13\cdot} - \bar{Y}_{21\cdot} + \bar{Y}_{23\cdot} \\
\bar{Y}_{12\cdot} - \bar{Y}_{13\cdot} - \bar{Y}_{22\cdot} + \bar{Y}_{23\cdot}
\end{pmatrix}
\end{aligned}
$$

Baseline Constraints Effects Model: Larvae Example

Copper Conc.	Zinc Concentration 0 ppm	750 ppm	1500 ppm
0 ppm	$\bar{Y}_{11\cdot} = 193.5$	$\bar{Y}_{12\cdot} = 167.5$	$\bar{Y}_{13\cdot} = 119.5$
150 ppm	$\bar{Y}_{21\cdot} = 172.5$	$\bar{Y}_{22\cdot} = 170.5$	$\bar{Y}_{23\cdot} = 111$

- Plug these values into β from previous slide

Sum-to-Zero Constraints Effects Model

$$Y_{ijk} = \mu + \alpha_i + \tau_j + (\alpha\tau)_{ij} + \epsilon_{ijk}$$

- Set $\sum_{i=1}^{a} \alpha_i = 0$
- Set $\sum_{j=1}^{b} \tau_j = 0$
- Set $\sum_{i=1}^{a} (\alpha\tau)_{ij} = 0$ for all j
- Set $\sum_{j=1}^{b} (\alpha\tau)_{ij} = 0$ for all i

Sum-to-Zero Constraints Effects Model

- $\mu = \bar{\mu}_{..}$
  overall mean response
- $\alpha_i = \bar{\mu}_{i.} - \bar{\mu}_{..}$
  related to main effect of the i^{th} level of Factor A
- $\tau_j = \bar{\mu}_{.j} - \bar{\mu}_{..}$
  related to main effect of the j^{th} level of Factor B
- $(\alpha\tau)_{ij} = \mu_{ij} - \bar{\mu}_{i.} - \bar{\mu}_{.j} + \bar{\mu}_{..}$
  interaction effect

Sum-to-Zero Constraints Effects Model: Larvae Example

- Set $\alpha_2 = -\alpha_1$
- Set $\tau_3 = -\tau_1 - \tau_2$
- Set $(\alpha\tau)_{13} = -(\alpha\tau)_{11} - (\alpha\tau)_{12}$
- Set $(\alpha\tau)_{21} = -(\alpha\tau)_{11}$
- Set $(\alpha\tau)_{22} = -(\alpha\tau)_{12}$
- Set $(\alpha\tau)_{23} = (\alpha\tau)_{11} + (\alpha\tau)_{12}$

Sum-to-Zero Constraints Effects Model: Larvae Example

$$\begin{bmatrix} Y_{111} \\ Y_{112} \\ Y_{121} \\ Y_{122} \\ Y_{131} \\ Y_{132} \\ Y_{211} \\ Y_{212} \\ Y_{221} \\ Y_{222} \\ Y_{231} \\ Y_{232} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 0 & -1 & 0 \\ 1 & -1 & 1 & 0 & -1 & 0 \\ 1 & -1 & 0 & 1 & 0 & -1 \\ 1 & -1 & 0 & 1 & 0 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \tau_1 \\ \tau_2 \\ (\alpha\tau)_{11} \\ (\alpha\tau)_{12} \end{bmatrix} + \begin{bmatrix} \epsilon_{111} \\ \epsilon_{112} \\ \epsilon_{121} \\ \epsilon_{122} \\ \epsilon_{131} \\ \epsilon_{132} \\ \epsilon_{211} \\ \epsilon_{212} \\ \epsilon_{221} \\ \epsilon_{222} \\ \epsilon_{231} \\ \epsilon_{232} \end{bmatrix}$$

Sum-to-Zero Constraints Effects Model: Larvae Example

$$
\begin{aligned}
\hat{\beta} &= (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y} \\
&= \begin{pmatrix}
\bar{Y}_{\cdots} \\
\bar{Y}_{1\cdot\cdot} - \bar{Y}_{\cdots} \\
\bar{Y}_{\cdot 1\cdot} - \bar{Y}_{\cdots} \\
\bar{Y}_{\cdot 2\cdot} - \bar{Y}_{\cdots} \\
\bar{Y}_{11\cdot} - \bar{Y}_{1\cdot\cdot} - \bar{Y}_{\cdot 1\cdot} + \bar{Y}_{\cdots} \\
\bar{Y}_{12\cdot} - \bar{Y}_{1\cdot\cdot} - \bar{Y}_{\cdot 2\cdot} + \bar{Y}_{\cdots}
\end{pmatrix}
\end{aligned}
$$

Sum-to-Zero Constraints Effects Model: Larvae Example

Copper Conc.	Zinc Concentration 0 ppm	750 ppm	1500 ppm	Mean
0 ppm	$\bar{Y}_{11\cdot} = 193.5$	$\bar{Y}_{12\cdot} = 167.5$	$\bar{Y}_{13\cdot} = 119.5$	$\bar{Y}_{1\cdot\cdot} = 160.1667$
150 ppm	$\bar{Y}_{21\cdot} = 172.5$	$\bar{Y}_{22\cdot} = 170.5$	$\bar{Y}_{23\cdot} = 111$	$\bar{Y}_{2\cdot\cdot} = 151.3333$
Mean	$\bar{Y}_{\cdot1\cdot} = 182$	$\bar{Y}_{\cdot2\cdot} = 169$	$\bar{Y}_{\cdot3\cdot} = 115.25$	$\bar{Y}_{\cdot\cdot\cdot} = 155.75$

- Plug these values into β from previous slide

Effects Tests: Larvae Example

Question 2: Are the response means the same for the two levels of copper, averaging across zinc levels?

- $H_0 : \bar{\mu}_{1\cdot} = \bar{\mu}_{2\cdot}$ where

$$\begin{aligned} \bar{\mu}_{1\cdot} &= \mu + \alpha_1 + \frac{\sum_{j=1}^{b} \tau_j}{b} + \frac{\sum_{j=1}^{b} (\alpha\tau)_{1j}}{b} \\ \bar{\mu}_{2\cdot} &= \mu + \alpha_2 + \frac{\sum_{j=1}^{b} \tau_j}{b} + \frac{\sum_{j=1}^{b} (\alpha\tau)_{2j}}{b} \end{aligned}$$

- Equivalent to testing

$$\alpha_1 + \frac{\sum_{j=1}^{b} (\alpha\tau)_{1j}}{b} = \alpha_2 + \frac{\sum_{j=1}^{b} (\alpha\tau)_{2j}}{b}$$

Effects Tests: Larvae Example

Question 2: Are the response means the same for the two levels of copper, averaging across zinc levels?

- Baseline Constraints:

$$H_0 : \alpha_1 + \frac{\sum_{j=1}^{b} (\alpha\tau)_{1j}}{b} = 0$$

- Sum-to-Zero Constraints:

$$H_0 : \alpha_1 = 0$$

Effects Tests: Larvae Example

Question 3: Are the response means the same for the three levels of zinc, averaging across copper levels?

- $H_0 : \bar{\mu}_{\cdot 1} = \bar{\mu}_{\cdot 2} = \bar{\mu}_{\cdot 3}$ where

$$\bar{\mu}_{\cdot 1} = \mu + \frac{\sum_{i=1}^{a} \alpha_i}{a} + \tau_1 + \frac{\sum_{i=1}^{a} (\alpha\tau)_{i1}}{a}$$

$$\bar{\mu}_{\cdot 2} = \mu + \frac{\sum_{i=1}^{a} \alpha_i}{a} + \tau_2 + \frac{\sum_{i=1}^{a} (\alpha\tau)_{i2}}{a}$$

$$\bar{\mu}_{\cdot 3} = \mu + \frac{\sum_{i=1}^{a} \alpha_i}{a} + \tau_3 + \frac{\sum_{i=1}^{a} (\alpha\tau)_{i3}}{a}$$

- Equivalent to testing

$$H_0 : \tau_1 + \frac{\sum_{i=1}^{a} (\alpha\tau)_{i1}}{a} = \tau_2 + \frac{\sum_{i=1}^{a} (\alpha\tau)_{i2}}{a} = \tau_3 + \frac{\sum_{i=1}^{a} (\alpha\tau)_{i3}}{a}$$

Effects Tests: Larvae Example

Question 3: Are the response means the same for the three levels of zinc, averaging across copper levels?

- Baseline Constraints:

$$H_0 : \tau_1 + \frac{\sum_{i=1}^{a} (\alpha\tau)_{i1}}{a} = \tau_2 + \frac{\sum_{i=1}^{a} (\alpha\tau)_{i2}}{a} = 0$$

- Sum-to-Zero Constraints:

$$H_0 : \tau_1 = \tau_2 = 0$$

Effects Tests: Larvae Example

Question 4: Are differences in mean responses between copper levels consistent across zinc levels?

- $H_0 : \mu_{11} - \mu_{21} = \mu_{12} - \mu_{22} = \mu_{13} - \mu_{23}$ where

$$\mu_{11} - \mu_{21} = [\mu + \alpha_1 + \tau_1 + (\alpha\tau)_{11}] - [\mu + \alpha_2 + \tau_1 + (\alpha\tau)_{21}]$$
$$\mu_{12} - \mu_{22} = [\mu + \alpha_1 + \tau_2 + (\alpha\tau)_{12}] - [\mu + \alpha_2 + \tau_2 + (\alpha\tau)_{22}]$$
$$\mu_{13} - \mu_{23} = [\mu + \alpha_1 + \tau_3 + (\alpha\tau)_{13}] - [\mu + \alpha_2 + \tau_3 + (\alpha\tau)_{23}]$$

- Equivalent to testing

$$H_0 : (\alpha\tau)_{11} - (\alpha\tau)_{21} = (\alpha\tau)_{12} - (\alpha\tau)_{22} = (\alpha\tau)_{13} - (\alpha\tau)_{23}$$

Effects Tests: Larvae Example

Question 4: Are differences in mean responses between copper levels consistent across zinc levels?

- Baseline Constraints:

$$H_0 : (\alpha\tau)_{11} = (\alpha\tau)_{12} = 0$$

- Sum-to-Zero Constraints:

$$H_0 : (\alpha\tau)_{11} = (\alpha\tau)_{12} = 0$$

Treatment Means

- Treatment mean: $\mu_{ij} = \mu + \alpha_i + \tau_j + (\alpha_i\tau_j)_{ij}$
- Estimate: $\bar{Y}_{ij\cdot}$
- Std. Error: $\sqrt{MS_{\text{error}}/n}$

  Needed for inferences of "Simple Effects" ...

Inference for Simple Effects

- Simple Effects: Difference in treatment means between two levels of one factor at a specific level of the other factor.
  - For Factor A: $\mu_{ij} - \mu_{rj}$
  - For Factor B: $\mu_{ij} - \mu_{is}$
- Estimate:
  - For Factor A: $\bar{Y}_{ij\cdot} - \bar{Y}_{rj\cdot}$
  - For Factor B: $\bar{Y}_{ij\cdot} - \bar{Y}_{is\cdot}$
- Std. Error: $\sqrt{MS_{\text{error}}\left(\frac{2}{n}\right)}$

Marginal Means

- Marginal Means: Mean of level of a Factor

Factor	A	B
Marginal Mean	$\bar{\mu}_{i\cdot}$	$\bar{\mu}_{\cdot j}$
Estimate	$\bar{Y}_{i\cdot\cdot}$	$\bar{Y}_{\cdot j\cdot}$
Std. Error	$\sqrt{MS_{\text{error}}\left(\frac{1}{nb}\right)}$	$\sqrt{MS_{\text{error}}\left(\frac{1}{na}\right)}$

Needed for inferences of "Main Effects" ...

Inference for Main Effects

- Main Effects: Difference in means of two levels of a Factor

Factor	A	B
Main Effect	$\bar{\mu}_{i\cdot} - \bar{\mu}_{r\cdot}$	$\bar{\mu}_{\cdot j} - \bar{\mu}_{\cdot s}$
Estimate	$\bar{Y}_{i\cdot\cdot} - \bar{Y}_{r\cdot\cdot}$	$\bar{Y}_{\cdot j\cdot} - \bar{Y}_{\cdot s\cdot}$
Std. Error	$\sqrt{MS_{\text{error}}\left(\frac{2}{nb}\right)}$	$\sqrt{MS_{\text{error}}\left(\frac{2}{na}\right)}$

Inference for Interaction Effects

- Interaction effects: Differences of simple effects
  - e.g. $(\mu_{ij} - \mu_{rj}) - (\mu_{il} - \mu_{rl})$
- Estimate: $(\bar{Y}_{ij\cdot} - \bar{Y}_{rj\cdot}) - (\bar{Y}_{il\cdot} - \bar{Y}_{rl\cdot})$
- Std. Error: $\sqrt{MS_{\text{error}}\left(\frac{4}{n}\right)}$
- Interaction effects are the least precisely estimated effects

Multiple Comparisons

- Adjust for multiple comparisons depending on desired analysis
  - Tukey HSD: all pairs of treatment or marginal means
  - Scheffe: all linear contrasts
  - Bonferroni: predetermined comparisons
  - Dunnett: compare many levels to one baseline

Unit 2

Two-Way ANOVA: Diagnostics

Model Assumptions

$$\epsilon_{ijk} \text{ are i.i.d. } N(0, \sigma^2)$$

- Independence
- Homogeneous (Equal) Variance
- Normality

3 x 3 Sugarcane Experiment (from Snedecor and Cochrane)

- Experimental units: 36 plots
- Factor 1: Three varieties of sugar cane (V1, V2, V3)
- Factor 2: Three levels of nitrogen:
  N1: 150 lbs/acre
  N2: 210 lbs/acre
  N3: 270 lbs/acre
- Response: Yield of sugar cane (tons/acre)
- Randomization: Completely randomized experiment, each of the 9 combinations of factor levels (treatments) is randomly assigned to 4 plots.
- Sources of variability?

3 x 3 Sugarcane Experiment: Data

	150 lb/a	210 lb/a	270 lb/a
Variety 1	$Y_{111} = 70.5$	$Y_{121} = 67.3$	$Y_{131} = 79.7$
	$Y_{112} = 67.5$	$Y_{122} = 75.9$	$Y_{132} = 72.8$
	$Y_{113} = 63.9$	$Y_{123} = 72.2$	$Y_{133} = 64.8$
	$Y_{114} = 64.2$	$Y_{124} = 60.5$	$Y_{134} = 86.3$
Variety 2	$Y_{211} = 58.6$	$Y_{221} = 64.3$	$Y_{231} = 64.4$
	$Y_{212} = 65.2$	$Y_{222} = 48.3$	$Y_{232} = 67.3$
	$Y_{213} = 70.2$	$Y_{223} = 74.0$	$Y_{233} = 78.0$
	$Y_{214} = 51.8$	$Y_{224} = 63.6$	$Y_{234} = 72.0$
Variety 3	$Y_{311} = 65.8$	$Y_{321} = 64.1$	$Y_{331} = 56.3$
	$Y_{312} = 66.3$	$Y_{322} = 64.8$	$Y_{332} = 54.7$
	$Y_{313} = 72.7$	$Y_{323} = 70.9$	$Y_{333} = 66.2$
	$Y_{314} = 67.6$	$Y_{324} = 58.3$	$Y_{334} = 54.4$

3 x 3 Sugarcane Experiment: Model

Effects model:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$$

with baseline constraints

- $\alpha_3 = 0$, $\beta_3 = 0$
- $(\alpha\beta)_{i3} = 0$ for $i = 1, 2, 3$
- $(\alpha\beta)_{3j} = 0$ for $j = 1, 2, 3$
- $\epsilon_{ijk} \sim N(0, \sigma^2)$

3 x 3 Sugarcane Experiment: Summary Statistics

	150 lb/a	210 lb/a	270 lb/a	Variety means
V1	$\bar{Y}_{11\cdot} = 66.53$ $S^2_{11} = 9.683$	$\bar{Y}_{12\cdot} = 68.98$ $S^2_{12} = 44.33$	$\bar{Y}_{13\cdot} = 75.95$ $S^2_{13} = 85.66$	$\bar{Y}_{1\cdot\cdot} = 70.48$
V2	$\bar{Y}_{21\cdot} = 61.45$ $S^2_{21} = 63.95$	$\bar{Y}_{22\cdot} = 62.55$ $S^2_{22} = 112.0$	$\bar{Y}_{23\cdot} = 70.43$ $S^2_{23} = 35.31$	$\bar{Y}_{2\cdot\cdot} = 64.81$
V3	$\bar{Y}_{31\cdot} = 68.60$ $S^2_{31} = 8.58$	$\bar{Y}_{32\cdot} = 64.53$ $S^2_{32} = 26.55$	$\bar{Y}_{33\cdot} = 57.90$ $S^2_{33} = 31.31$	$\bar{Y}_{3\cdot\cdot} = 63.68$
Nitrogen means	$\bar{Y}_{\cdot 1\cdot} = 65.53$	$\bar{Y}_{\cdot 2\cdot} = 65.35$	$\bar{Y}_{\cdot 3\cdot} = 68.09$	$\bar{Y}_{\cdot\cdot\cdot} = 66.32$

3 x 3 Sugarcane Experiment: ANOVA Table

source of variation	degrees of freedom	Sums of Squares
Varieties	$3-1$	$12\sum_{i=1}^{3}(\bar{Y}_{i..}-\bar{Y}_{...})^2$ =319.37
Nitrogen levels	$3-1$	$12\sum_{j=1}^{b}(\bar{Y}_{.j.}-\bar{Y}_{...})^2$=56.54
V$\times$N Interaction	$(2)(2)$	$4\sum_i\sum_j(\bar{Y}_{ij.}-\bar{Y}_{i..}-\bar{Y}_{.j.}+\bar{Y}_{...})^2$ = 559.79
Error	$(3)(3)(4-1)$	$\sum_i\sum_j\sum_k(Y_{ijk}-\bar{Y}_{ij.})^2$ = 1254.46
Total	$36-1$	$\sum_i\sum_j\sum_k(Y_{ijk}-\bar{Y}_{...})^2$=2190.16

3 x 3 Sugarcane Experiment: ANOVA

- Compute the F-test for interaction

$$F = \frac{MS_{interaction}}{MS_{error}} = \frac{139.95}{46.46} = 3.01$$

  on (4,27) df with p-value=0.036
- Examine profile plots
- Compare mean yields for nitrogen levels

$$F = \frac{MS_{nitrogen}}{MS_{error}} = \frac{28.27}{46.46} = 0.61$$

  on (2,27) df with p-value$>$0.50
- Averaging across varieties, the mean yields are not significantly different for the three levels of nitrogen

3 x 3 Sugarcane Experiment: Inference for Nitrogen Effect

- Since there is a significant interaction, nitrogen has some effect on the mean yield of sugar cane, but it is not the same for all varieties.
- Variance of a difference between nitrogen means

$$Var(\bar{Y}_{.j.} - \bar{Y}_{.k.}) = Var(\bar{Y}_{.j.}) + Var(\bar{Y}_{.k.}) = \frac{\sigma^2}{12} + \frac{\sigma^2}{12} = \frac{\sigma^2}{6}$$

- Standard error: $S_{\bar{Y}_{.j.} - \bar{Y}_{.k.}} = \sqrt{MS_{error}\frac{1}{6}}$

3 x 3 Sugarcane Experiment: Inference for Variety Effect

- Compare mean yields for varieties

$$F = \frac{MS_{varieties}}{MS_{error}} = \frac{159.7}{46.46} = 3.44$$

  on (2,27) df with p-value = 0.047
- Averaging across nitrogen levels, mean yields are different for the three varieties
- Variance of a difference between two variety means

$$Var(\bar{Y}_{i..} - \bar{Y}_{k..}) = Var(\bar{Y}_{i..}) + Var(\bar{Y}_{k..}) = \frac{\sigma^2}{12} + \frac{\sigma^2}{12} = \frac{\sigma^2}{6}$$

- Standard error: $S_{\bar{Y}_{i..} - \bar{Y}_{k..}} = \sqrt{MS_{error}\frac{1}{6}} = 2.78$

3 x 3 Sugarcane Experiment: Orthogonal Contrasts - Variety

- Compare mean yields for varieties 1 and 2

$$\hat{C}_{V1-V2} = (1)\bar{Y}_{1\cdot\cdot} + (-1)\bar{Y}_{2\cdot\cdot} + (0)\bar{Y}_{3\cdot\cdot} = 5.67$$

- $SS_{\hat{C}_{V1-V2}} = \dfrac{(\hat{C}_{V1-V2})^2}{\frac{(1)^2}{12} + \frac{(-1)^2}{12} + \frac{0}{12}} = 192.9$
- Compare mean yield for varieties 3 against the combined mean yield for varieties 1 and 2

$$\hat{C}_{V3-(V1+V2)/2} = (-.5)\bar{Y}_{1\cdot\cdot} + (-.5)\bar{Y}_{2\cdot\cdot} + (1)\bar{Y}_{3\cdot\cdot} = -3.965$$

- $SS_{\hat{C}_{V3-(V1+V2)/2}} = \dfrac{(\hat{C}_{V3-(V1+V2)/2})^2}{\frac{(-.5)^2}{12} + \frac{(-.5)^2}{12} + \frac{(1)^2}{12}} = 125.8$

3 x 3 Sugarcane Experiment: Orthogonal Contrasts - Variety

Decomposition of $SS_{varieties}$

source of variation	df	SS	MS	F	p-value
Varieties	2	319.37	159.69	3.44	0.046
$\hat{C}_{V1-V2}$	1	192.9	192.9	4.2	0.051
$\hat{C}_{V3-(V1+V2)/2}$	1	125.8	125.8	2.7	0.11

3 x 3 Sugarcane Experiment: Orthogonal Contrasts - Nitrogen

- Linear trend in mean yields across nitrogen levels

$$\hat{C}_{N:lin} = (-1)\bar{Y}_{\cdot 1 \cdot} + (0)\bar{Y}_{\cdot 2 \cdot} + (1)\bar{Y}_{\cdot 3 \cdot} = 2.56$$

- Quadratic trend in mean yields across nitrogen levels

$$\hat{C}_{N:quad} = (1)\bar{Y}_{\cdot 1 \cdot} + (-2)\bar{Y}_{\cdot 2 \cdot} + (1)\bar{Y}_{\cdot 3 \cdot} = 2.92$$

- Decomposition of $SS_{nitrogen}$

source of variation	df	SS	MS	F	p-value
N-levels	2	56.54	28.27	0.6	
$\hat{C}_{N:lin}$	1	39.3	39.3	0.8	
$\hat{C}_{N:quad}$	1	17.1	17.1	0.4	

3 x 3 Sugarcane Experiment: Orthogonal Contrasts - Interaction

- Difference in linear trends in mean yields across nitrogen levels for varieties 1 and 2

$$\begin{aligned}\hat{C}_{int,1} &= [(-1)\bar{Y}_{11\cdot} + (0)\bar{Y}_{12\cdot} + (1)\bar{Y}_{13\cdot}] \\ &\quad -[(-1)\bar{Y}_{21\cdot} + (0)\bar{Y}_{22\cdot} + (1)\bar{Y}_{23\cdot}] = 0.45\end{aligned}$$

- Contrast coefficients

Nitrogen levels

Varieties		-1	0	1
V1	1	-1	0	1
V2	-1	1	0	-1
V3	0	0	0	0

3 x 3 Sugarcane Experiment: Orthogonal Contrasts - Interaction

- Difference in linear trends in mean yields across nitrogen levels for variety 3 vs the average of varieties 1 and 2

$$\begin{aligned}\hat{C}_{int,2} &= [(-1)\bar{Y}_{31\cdot} + (0)\bar{Y}_{32\cdot} + (1)\bar{Y}_{33\cdot}] \\ &\quad -\frac{1}{2}[(-1)\bar{Y}_{11\cdot} + (0)\bar{Y}_{12\cdot} + (1)\bar{Y}_{13\cdot}] \\ &\quad -\frac{1}{2}[(-1)\bar{Y}_{21\cdot} + (0)\bar{Y}_{22\cdot} + (1)\bar{Y}_{23\cdot}] = -19.90\end{aligned}$$

3 x 3 Sugarcane Experiment: Orthogonal Contrasts - Interaction

- Contrast coefficients

		Nitrogen levels		
Varieties		-1	0	1
V1	-0.5	0.5	0	-0.5
V2	-0.5	0.5	0	-0.5
V3	1	-1	0	1

3 x 3 Sugarcane Experiment: Orthogonal Contrasts - Interaction

- Difference in quadratic trends in mean yields across nitrogen levels for varieties 1 and 2

$$\begin{aligned}\hat{C}_{int,3} &= [(1)\bar{Y}_{11\cdot} + (-2)\bar{Y}_{12\cdot} + (1)\bar{Y}_{13\cdot}] \\ &\quad -[(1)\bar{Y}_{21\cdot} + (-2)\bar{Y}_{22\cdot} + (1)\bar{Y}_{23\cdot}] = -2.26\end{aligned}$$

- Contrast coefficients

Nitrogen levels

Varieties		1	-2	1
V1	1	1	-2	1
V2	-1	-1	2	1
V3	0	0	0	0

3 x 3 Sugarcane Experiment: Orthogonal Contrasts - Interaction

- Difference between the quadratic trend in mean yields across nitrogen levels for variety 3 and the average of the quadratic trends for varieties 1 and 2

$$\begin{aligned}\hat{C}_{int,4} &= [(1)\bar{Y}_{31\cdot} + (-2)\bar{Y}_{32\cdot} + (1)\bar{Y}_{33\cdot}] \\ &\quad -\frac{1}{2}[(-0.5)\bar{Y}_{11\cdot} + (1)\bar{Y}_{12\cdot} + (-0.5)\bar{Y}_{13\cdot}] \\ &\quad -\frac{1}{2}[(-0.5)\bar{Y}_{21\cdot} + (1)\bar{Y}_{22\cdot} + (-0.5)\bar{Y}_{23\cdot}] = -8.2\end{aligned}$$

3 x 3 Sugarcane Experiment: **Orthogonal Contrasts - Interaction**

- Contrast coefficients

		Nitrogen levels		
Varieties		1	-2	1
V1	-0.5	-0.5	1	-0.5
V2	-0.5	-0.5	1	-0.5
V3	1	1	-2	1

3 x 3 Sugarcane Experiment: Orthogonal Contrasts - Interaction

- Decomposition of $SS_{interaction}$

source of variation	df	SS	MS	F	p-value
Interaction	4	599.79	139.95	3.01	0.036
$\hat{C}_{int,1}$	1	0.025	0.2025	0.004	0.948
$\hat{C}_{int,2}$	1	528.01	528.01	11.36	0.0023
$\hat{C}_{int,3}$	1	1.69	1.69	4.2	0.850
$\hat{C}_{int,4}$	1	29.88	28.88	0.64	0.430

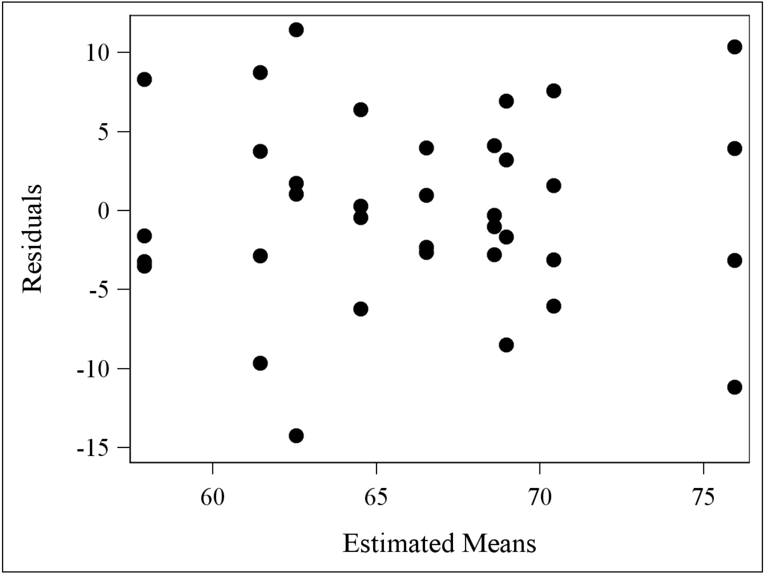
3 x 3 Sugarcane Experiment: Summary

- There is no significant quadratic trend in mean yields across nitrogen levels for any variety.
- Varieties 1 and 2 show increasing linear trends in mean yields across nitrogen levels with about the same slope.
- Variety 1 consistently yields about 5.7 tons/acre more than variety 2
- Variety 3 exhibits a decreasing linear trend in mean yields across nitrogen levels. It may be the best variety for low nitrogen levels.

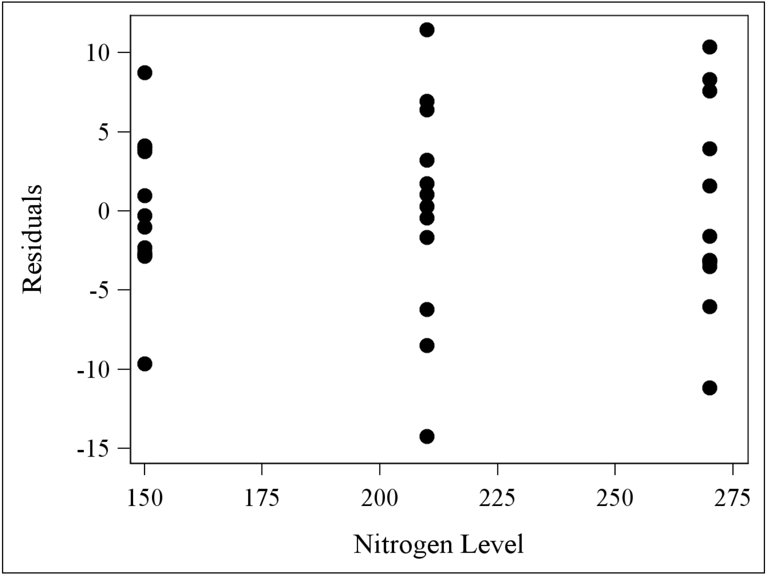
Model Diagnostics

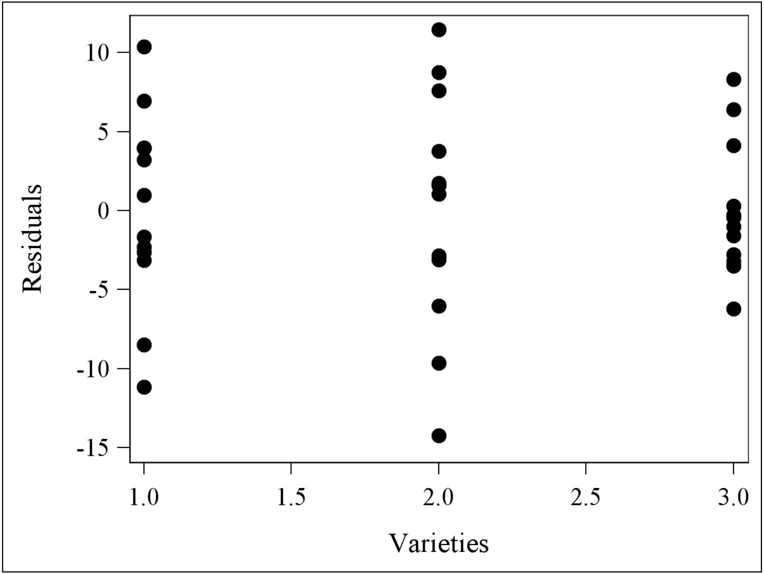
- Independence - check details of data collection
- Homogeneous variance
  - Plot residuals vs. estimated means
  - Plot residuals vs. levels of each factor
- Normality
  - Normal probability plot for residuals
  - Histogram of residuals
  - Tests for Normality

Two-Way ANOVA Diagnostics



Two-Way ANOVA Diagnostics




10
5
0
-5
-10
-15
Residuals
1.0
1.5
2.0
2.5
3.0
Varieties

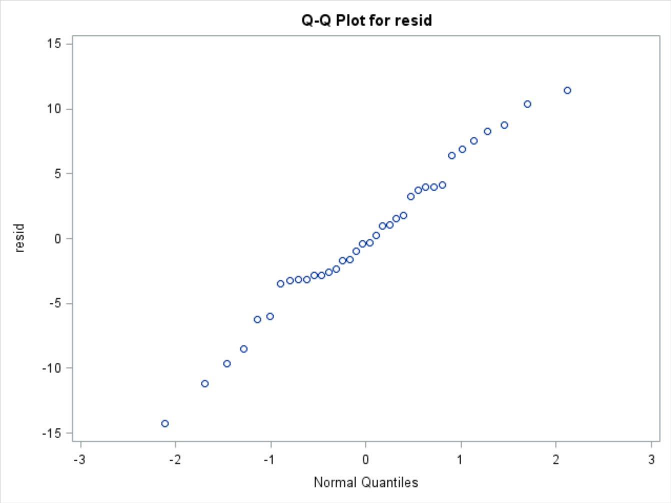
The UNIVARIATE Procedure
Variable: resid

Moments			
N	36	**Sum Weights**	36
Mean	0	**Sum Observations**	0
Std Deviation	5.98679499	**Variance**	35.8417143
Skewness	-0.1835349	**Kurtosis**	-0.0641013
Uncorrected SS	1254.46	**Corrected SS**	1254.46
Coeff Variation	.	**Std Error Mean**	0.99779917

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.982138	**Pr < W**	0.8150
Kolmogorov-Smirnov	**D**	0.112735	**Pr > D**	>0.1500
Cramer-von Mises	**W-Sq**	0.041065	**Pr > W-Sq**	>0.2500
Anderson-Darling	**A-Sq**	0.252323	**Pr > A-Sq**	>0.2500

Two-Way ANOVA Diagnostics

The UNIVARIATE Procedure
Variable: resid



Remedies

- Transformation remedy for non-normality is commonly used
- Remember, transformation changes
  - Error properties
  - The model for the mean responses
  - Can eliminate (reduce) or introduce (enhance) interactions
- Randomization tests
- Rank tests

Unit 2

Additional Factorial Designs

Blocking

- RCBD with full factorial treatment design ($r = ab$ treatments)
- Different random assignment of units to treatments in each of the n blocks
  - One experimental unit for each block each treatment
  - Assume no interaction between block and treatment effects
  - Model: $Y_{ijk} = \mu + \beta_i + \alpha_j + \tau_k + (\alpha\tau)_{jk} + \epsilon_{ijk}$

Blocking

Variation	d.f	Sums of Squares
Block	$n-1$	$ab\sum_{i=1}^{n}(\bar{Y}_{i\cdot\cdot}-\bar{Y}_{\cdot\cdot\cdot})^2$
Factor A	$a-1$	$nb\sum_{j=1}^{a}(\bar{Y}_{\cdot j\cdot}-\bar{Y}_{\cdot\cdot\cdot})^2$
Factor B	$b-1$	$na\sum_{k=1}^{b}(\bar{Y}_{\cdot\cdot k}-\bar{Y}_{\cdot\cdot\cdot})^2$
A×B interaction	$(a-1)(b-1)$	$n\sum_j\sum_k(\bar{Y}_{\cdot jk}-\bar{Y}_{\cdot j\cdot}-\bar{Y}_{\cdot\cdot k}+\bar{Y}_{\cdot\cdot\cdot})^2$
Error	$(ab-1)(n-1)$	SS_{error}
Corrected total	$abn-1$	$\sum_i\sum_j\sum_k(Y_{ijk}-\bar{Y}_{\cdot\cdot\cdot})^2$

$$
\begin{aligned}
SS_{\text{error}} &= b\sum_i\sum_j(\bar{Y}_{ij\cdot}-\bar{Y}_{i\cdot\cdot}-\bar{Y}_{\cdot j\cdot}+\bar{Y}_{\cdot\cdot\cdot})^2 \\
&\quad +a\sum_i\sum_k(\bar{Y}_{i\cdot k}-\bar{Y}_{i\cdot\cdot}-\bar{Y}_{\cdot\cdot k}+\bar{Y}_{\cdot\cdot\cdot})^2 \\
&\quad +\sum_i\sum_j\sum_k(Y_{ijk}-\bar{Y}_{ij\cdot}-\bar{Y}_{i\cdot k}-\bar{Y}_{\cdot jk}+\bar{Y}_{i\cdot\cdot}+\bar{Y}_{\cdot j\cdot}+\bar{Y}_{\cdot\cdot k}-\bar{Y}_{\cdot\cdot\cdot})^2
\end{aligned}
$$

Three Factors: Notation

- Factor A with a levels: $i = 1, \ldots, a$
- Factor B with b levels: $j = 1, \ldots, b$
- Factor C with c levels: $k = 1, \ldots, c$
- n replications for each treatment: $l = 1, \ldots, n$

$$Y_{ijkl} = \mu + \alpha_i + \tau_j + \delta_k + (\alpha\tau)_{ij} + (\alpha\delta)_{ik} + (\tau\delta)_{jk} + (\alpha\tau\delta)_{ijk} + \epsilon_{ijkl}$$

Three Factors: Effects

- α_i = Factor A effect
- τ_j = Factor B effect
- δ_k = Factor C effect
- $(\alpha\tau)$ = interaction of Factors A and B
- $(\alpha\delta)$ = interaction of Factors A and C
- $(\tau\delta)$ = interaction of Factors B and C
- $(\alpha\tau\delta)$ = interaction of Factors A, B, and C

Three Factors: Larvae Example

- Experimental units: 36 water tanks containing minnow larvae
- Factors: Factor A = Nickel (3 levels), Factor B = Copper (2 levels), Factor C = Zinc (3 levels)
- Replication: Experimental units are randomly assigned to the 18 treatments = 2 units per treatment.
- Response Variable: protein content (μg/larva)

Three Factors: ANOVA Table

source of variation	degrees of freedom	sums of squares
factor A	$a-1$	$nbc\sum_i(\bar{Y}_{i\cdots}-\bar{Y}_{\cdots\cdot})^2$
factor B	$b-1$	$nac\sum_j(\bar{Y}_{\cdot j\cdot\cdot}-\bar{Y}_{\cdots\cdot})^2$
factor C	$c-1$	$nab\sum_k(\bar{Y}_{\cdot\cdot k\cdot}-\bar{Y}_{\cdots\cdot})^2$
interaction AB	$(a-1)(b-1)$	$nc\sum_i\sum_j(\bar{Y}_{ij\cdot\cdot}-\bar{Y}_{i\cdots}-\bar{Y}_{\cdot j\cdot\cdot}+\bar{Y}_{\cdots\cdot})^2$
interaction AC	$(a-1)(c-1)$	$nb\sum_i\sum_k(\bar{Y}_{i\cdot k\cdot}-\bar{Y}_{i\cdots}-\bar{Y}_{\cdot\cdot k\cdot}+\bar{Y}_{\cdots\cdot})^2$
interaction BC	$(b-1)(c-1)$	$na\sum_j\sum_k(\bar{Y}_{\cdot jk\cdot}-\bar{Y}_{\cdot j\cdot\cdot}-\bar{Y}_{\cdot\cdot k\cdot}+\bar{Y}_{\cdots\cdot})^2$
interaction ABC	$(a-1)(b-1)(c-1)$	SS_{ABC}
error	$abc(n-1)$	$\sum_i\sum_j\sum_k\sum_l(Y_{ijkl}-\bar{Y}_{ijk\cdot})^2$
total	$abcn-1$	$\sum_i\sum_j\sum_k\sum_l(Y_{ijkl}-\bar{Y}_{\cdots\cdot})^2$

$$SS_{ABC}=n\sum_i\sum_j\sum_k(\bar{Y}_{ijk\cdot}-\bar{Y}_{ij\cdot\cdot}-\bar{Y}_{i\cdot k\cdot}-\bar{Y}_{\cdot jk\cdot}+\bar{Y}_{i\cdots}+\bar{Y}_{\cdot j\cdot\cdot}+\bar{Y}_{\cdot\cdot k\cdot}-\bar{Y}_{\cdots\cdot})^2$$

Three Factors: Interpreting Effects

- Main effects and simple effects are defined as before.
- Main effects are differences (or contrasts) between levels of one factor averaged over all levels of the other factors. Correspond to marginal means that are averages over "left out" factors and replicates.
  - e.g., difference between two copper levels averaged over all Zinc levels and Nickel levels
- Simple effects correspond to differences between levels of one factor at specific levels of the other factors.
  - e.g., difference between two copper levels at Zinc level 1 and Nickel level 1.

Three Factors: Interpreting with Interactions

- Interpretations of factor effects are complicated by the presence of three-factor interaction
- Example: A, B, C
  - AB interaction: interaction between factor A and factor B (averaging across the levels of factor C)
  - ABC interaction: nature of the AB interaction depends on the level of factor C
    - Remember 2-way interaction: Are the simple effects of A the same at every level of B?
    - Three-way interaction generalizes this: Are the 2-way A*B interaction effects the same for each level of C?

Three Factors: Testing Interactions

Example: Is there 3-way interaction?

The cell means:

level of C	level of B	Cell mean for: A=1	A=2
1	1	6	6
1	2	1	3
2	1	2	4
2	2	1	5

- When C = 1, the interaction effect for A*B is: (6-1)-(6-3)=2
- When C = 2, the interaction effect for A*B is: (2-1)-(4-5) = 2

Three Factors: Testing Interactions

- Because the 2-way interaction A*B does not depend on the level of C, there is no 3-way interaction.
- Four-way interaction (if four factors): Are the 3-way A*B*C interaction effects the same for each level of D?
- Such concepts extend to many factors.
- Often (not always) magnitude of main effect $>$ that of 2-way interactions $>$ 3-way $>$ ... $>$ high order interaction

No Replication - 2^K Studies: Set-up

- K factors with $j = 1, 2, ..., r_k$ levels for the kth factor
- Known as a $r_1 \times r_2 \times \cdots \times r_k$ factorial designs
- There are $r_1 \times r_2 \times \cdots \times r_k$ experimental units, and exactly one unit is assigned to each treatment.
- If all possible interactions are included in the model, there are no degrees of freedom left for computing MS_{error}
- We will only consider the special situation of K factors with exactly two levels for each factor - 2^K factorial designs

No Replication - 2^K Studies: Special Features

- All main effect and interaction contrasts have 1 d.f.
- Set up the model with
  - One column in the model matrix X for each main effect using +1/-1 coding
  - Interaction columns are obtained by multiplication of appropriate main effect columns
  - All columns of X are orthogonal

No Replication - 2^K Studies: $MS_{\textbf{error}}$

- With no replication there are no degrees of freedom for error when you fit all possible main effects and interactions
- Pool sums of squares from non-significant interaction terms to obtain a mean square error value

No Replication - 2^K Studies

Determination of non-significant terms:

- Construct a normal Q-Q plot: Order estimated effects (except the overall mean/intercept) from smallest to largest and plot against expected quantiles from a N(0,1) distribution
- All estimated effects must have same variance
- The slope of the line provides an estimate of the common standard deviation
- Estimates far from the line indicate “non-zero” values that do not correspond to random error
- Estimates close to the line reflect random error; pool their corresponding sums of squares to obtain a mean square error value
- Use this estimate of σ^2 to test the significance of the “non-zero” contrasts, construct CI’s, etc...

No Replication - 2^K **Studies: Chemical Process Study Example**

- Examine the effects of four factors on a chemical process
  - Determine which factors affect the conversion percentage (response)
  - Look for interactions
  - Randomization: Order in which various combinations of factor levels were run was randomized

No Replication - 2^K Studies: Chemical Process Study Example

- Factor 1: Amount of a catalyst (α)
  ($i = 1$) 10 lb (coded as 1)
  ($i = 2$) 15 lb (coded as -1)
- Factor 2: Temperature (τ)
  ($j = 1$) 220^oC (coded as 1)
  ($j = 2$) 240^oC (coded as -1)
- Factor 3: Pressure (γ)
  ($k = 1$) 50 psi (coded as 1)
  ($k = 2$) 80 psi (coded as -1)
- Factor 4: Concentration of substance to be converted (δ)
  ($l = 1$) 10% (coded as 1)
  ($l = 2$) 15% (coded as -1)

No Replication - 2^K Studies: Chemical Example Data

Catalyst	Temp.	Pressure	Conc.	Y	Run order
10	220	50	10	71	8
15	220	50	10	61	2
10	240	50	10	90	10
15	240	50	10	82	4
10	220	80	10	68	15
15	220	80	10	61	9
10	240	80	10	87	1
15	240	80	10	80	13
10	220	50	15	61	16
15	220	50	15	50	5
10	240	50	15	89	11
15	240	50	15	83	14
10	220	80	15	59	3
15	220	80	15	51	12
10	240	80	15	85	6
15	240	80	15	78	7

No Replication - 2^K Studies: Chemical Example Model

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon$$

$$\begin{bmatrix} Y_{1111} \\ Y_{2111} \\ Y_{1211} \\ Y_{2211} \\ Y_{1121} \\ Y_{2121} \\ Y_{1221} \\ Y_{2221} \\ Y_{1112} \\ Y_{2112} \\ Y_{1212} \\ Y_{2212} \\ Y_{1122} \\ Y_{2122} \\ Y_{1222} \\ Y_{2222} \end{bmatrix} = \mathbf{X} \begin{bmatrix} \mu \\ \alpha_1 \\ \tau_1 \\ \gamma_1 \\ \delta_1 \\ (\alpha\tau)_{11} \\ (\alpha\gamma)_{11} \\ (\alpha\delta)_{11} \\ (\tau\gamma)_{11} \\ (\tau\delta)_{11} \\ (\gamma\delta)_{11} \\ (\alpha\tau\gamma)_{111} \\ (\alpha\tau\delta)_{111} \\ (\alpha\gamma\delta)_{111} \\ (\tau\gamma\delta)_{111} \\ (\alpha\tau\gamma\delta)_{1111} \end{bmatrix} + \begin{bmatrix} \epsilon_{1111} \\ \epsilon_{2111} \\ \epsilon_{1211} \\ \epsilon_{2211} \\ \epsilon_{1121} \\ \epsilon_{2121} \\ \epsilon_{1221} \\ \epsilon_{2221} \\ \epsilon_{1112} \\ \epsilon_{2112} \\ \epsilon_{1212} \\ \epsilon_{2212} \\ \epsilon_{1122} \\ \epsilon_{2122} \\ \epsilon_{1222} \\ \epsilon_{2222} \end{bmatrix}$$

No Replication - 2^K Studies: Chemical Example Model

$$
\mathbf{X} = \begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & 1 & -1 \\
1 & 1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 \\
1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 & 1 \\
1 & 1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 & -1 \\
1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \\
1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & 1 \\
1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \\
1 & 1 & 1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 & -1 & -1 \\
1 & -1 & 1 & 1 & -1 & -1 & -1 & 1 & 1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 \\
1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 \\
1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & 1 & -1 \\
1 & 1 & 1 & -1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & -1 & -1 & 1 & 1 & 1 \\
1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 & -1 \\
1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 \\
1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & 1
\end{bmatrix}
$$

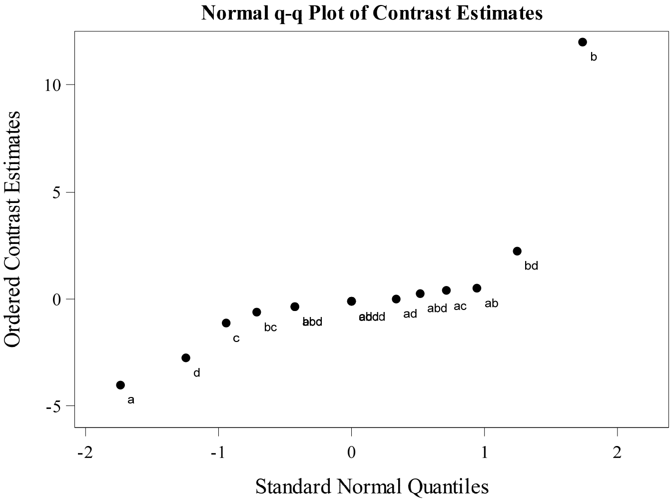
No Replication - 2^K Studies: **Chemical Example**

- Least Squares Estimates: $\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$ with covariance matrix $Var(\hat{\beta}) = \sigma^2(\mathbf{X}^T\mathbf{X})^{-1} = \frac{\sigma^2}{2^K}\,\mathbf{I}$
- Order the $2^K - 1$ estimates (other than the overall mean) and plot the i-th smallest against $q_i = \Phi(\frac{i-.375}{2^K-1+0.25})$ (Blom approximation)
- Points that deviate from a straight line pattern indicate significant effects
- Slope of line fit to the remaining points provides an estimate of $\sqrt{\sigma^2/2^K}$

No Replication - 2^K Studies: Chemical Example Estimates

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\alpha}_1 = -4.000 \\ \hat{\delta}_1 = -2.750 \\ \hat{\gamma}_1 = -1.125 \\ \widehat{(\tau\gamma)}_{11} = -0.625 \\ \widehat{(\alpha\tau\gamma)}_{111} = -0.375 \\ \widehat{(\tau\gamma\delta)}_{111} = -0.375 \\ \widehat{(\gamma\delta)}_{11} = -0.125 \\ \widehat{(\alpha\gamma\delta)}_{111} = -0.125 \\ \widehat{(\alpha\tau\gamma\delta)}_{1111} = -0.125 \\ \widehat{(\alpha\delta)}_{11} = 0.000 \\ \widehat{(\alpha\tau\delta)}_{111} = 0.250 \\ \widehat{(\alpha\gamma)}_{11} = 0.375 \\ \widehat{(\alpha\tau)}_{11} = 0.500 \\ \widehat{(\tau\delta)}_{11} = 2.250 \\ \hat{\tau}_1 = 12.000 \end{bmatrix}$$

No Replication - 2^K Studies: Chemical Example



No Replication - 2^K Studies: Chemical Example

ANOVA Table for Selected Effects

variation	df	SS	MS	F	p
Catalyst	1	256	256	136.5	$< .0001$
Temp.	1	2304	2304	1228.8	$< .0001$
Pressure	1	20.25	20.25	10.8	.0082
Conc.	1	121	121	64.5	$< .0001$
Temp$\times$Conc	1	81	81	43.2	$< .0001$
Residuals	10	18.75	1.875		
total	15	2801.00			

Unit 2

This is the end ...

Midterm Exam #2 Coming Up!

Questions?

Contact me:

Email: dmommen@iastate.edu
Student Office Hours: Thursdays @ 10-11 am