Workshop 7.6a: Factorial ANOVA

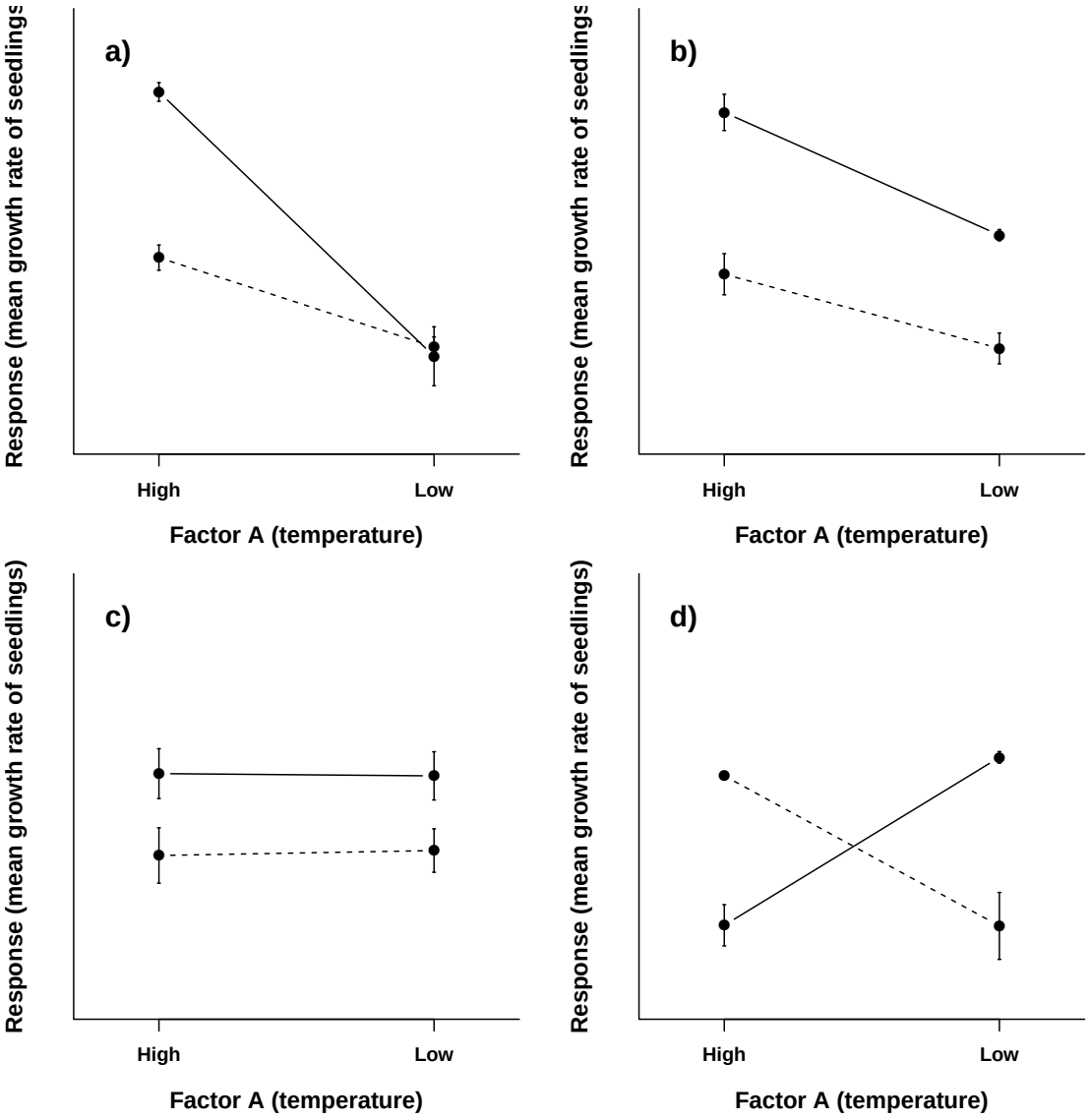
Murray Logan
July 19, 2017

Table of contents

1. Background

1.1. Factorial ANOVA

1.2. Factorial ANOVA



1.3. The linear model

Two-factor

	Low N	Medium N	High N
Low temp.	XXX	XXX	XXX
High temp	XXX	XXX	XXX

$$y_{ijk} = \mu + \alpha_i + \beta_j + \alpha_i\beta_j + \varepsilon_{ijk}$$

- α_i is the effect of the i_{th} temperature
- β_j is the effect of the j_{th} nitrogen level
- $\alpha_i\beta_j$ is the effect of the ij_{th} interaction.

1.4. The linear model

Two-factor

	Low N	Medium N	High N
Low temp.	XXX	XXX	XXX
High temp	XXX	XXX	XXX

Temp	Nitrogen
Low	Low
Low	Low
Low	Low
Low	Medium
Low	Medium
Low	Medium
Low	High
Low	High
Low	High
High	Low
High	Low
High	Low
High	Medium
High	Medium
High	Medium
High	High
High	High
High	High

$$y_i = \beta_{0i} + \beta_{1i} + \beta_{2i} + \beta_{3i} + \beta_{4i} + \beta_{5i} + \beta_{6i} + \varepsilon i$$

1.5. The linear model

Two-factor

T	N	NA	(Intercept)	THigh	NMedium	NHigh	THigh:NMedium	THigh:NHigh
Low	Low	NA	1	0	0	0	0	0
Low	Low	NA	1	0	0	0	0	0
Low	Low	NA	1	0	0	0	0	0
Low	Medium	NA	1	0	1	0	0	0
Low	Medium	NA	1	0	1	0	0	0
Low	Medium	NA	1	0	1	0	0	0
Low	High	NA	1	0	0	1	0	0
Low	High	NA	1	0	0	1	0	0
Low	High	NA	1	0	0	1	0	0
High	Low	NA	1	1	0	0	0	0
High	Low	NA	1	1	0	0	0	0

```
High  Low    NA        1          1          0          0          0          0
High Medium  NA        1          1          1          0          1          0
High Medium  NA        1          1          1          0          1          0
High Medium  NA        1          1          1          0          1          0
High  High   NA        1          1          0          1          0          1
High  High   NA        1          1          0          1          0          1
High  High   NA        1          1          0          1          0          1
```

1.6. The linear model

Two-factor

	Low N	Medium N	High N
Low temp.	XXX	XXX	XXX
High temp	XXX	XXX	XXX

$$y_i = \beta_{0i} + \beta_{1i} + \beta_{2i} + \beta_{3i} + \beta_{4i} + \beta_{5i} + \beta_{6i} + \varepsilon i$$

- β_0 is the mean of the $T_L : N_L$ group
- β_1 is the **difference** between $T_H : N_L$ and $T_L : N_L$
- β_2 is the **difference** between $T_L : N_M$ and $T_L : N_L$

1.7. The linear model

Two-factor

	Low N	Medium N	High N
Low temp.	XXX	XXX	XXX
High temp	XXX	XXX	XXX

$$y_{ijk} = \mu + \alpha_i + \beta_j + \alpha_i\beta_j + \varepsilon_{ijk}$$

- α_i is the effect of the i_{th} temperature **at the base level of** β

. . .

1.8. Factorial ANOVA

Factor	MS	F-ratio (both fixed)	F-ratio (A fixed, B random)	F-ratio (both random)
A	MS_A	MS_A/MS_{Resid}	$MS_A/MS_{A:B}$	$MS_A/MS_{A:B}$
B	MS_B	MS_B/MS_{Resid}	MS_B/MS_{Resid}	$MS_B/MS_{A:B}$
A:B	$MS_{A:B}$	$MS_{A:B}/MS_{Resid}$	$MS_{A:B}/MS_{Resid}$	$MS_{A:B}/MS_{Resid}$

2. Design Balance

2.1. Balance

When **balanced**

$$SS_{TOTAL} = SS_A + SS_B + SS_{A:B} + SS_{Resid}$$

2.2. Factoral ANOVA

2.2.1. Design balance

- When **balanced**

$$SS_{TOTAL} = SS_A + SS_B + SS_{A:B} + SS_{Resid}$$

- When **not balanced**

$$SS_{TOTAL} \neq SS_A + SS_B + SS_{A:B} + SS_{Resid}$$

2.3. Factorial ANOVA

2.4. Factoral ANOVA

2.4.1. Design balance

- When **balanced**

$$SS_{TOTAL} = SS_A + SS_B + SS_{A:B} + SS_{Resid}$$

- When **not balanced**

$$SS_{TOTAL} \neq SS_A + SS_B + SS_{A:B} + SS_{Resid}$$

    - **can't use** sequential SS (Type I SS)
    - should use either
        - hierarchical (Type II SS)
        - marginal (**Type III SS**)

2.5. Factorial ANOVA

2.6. Factorial ANOVA

2.6.1. Assumptions

- Normality
- Homogeneity of variance
- Independence
- Considerations for **Balance**

3. Worked examples

3.1. Worked examples

```
> #Worked examples
> stehman <- read.csv('../data/stehman.csv', strip.white=T)
```

```
Error in file(file, "rt"): cannot open the connection
```

```
> head(stehman)
```

```
Error in head(stehman): object 'stehman' not found
```

3.2. Worked Examples

Question: what effects do pH and health have on the bud emergence rating of spruce seedlings

Linear model:

$$Buds_{ijk} = \mu + \alpha_i + \beta_j + \alpha_i\beta_j + \varepsilon_{ijk} \qquad \varepsilon \sim \mathcal{N}(0, \sigma^2)$$

3.3. Worked Examples

```
Error in file(file, "rt"): cannot open the connection
```

```
Error in head(quinn): object 'quinn' not found
```

3.4. Worked Examples

Question: what effects do season and density have on barnacle recruitment

Linear model:

$$Recruits_{ijk} = \mu + \alpha_i + \beta_j + \alpha_i\beta_j + \varepsilon_{ijk} \qquad \varepsilon \sim \mathcal{N}(0, \sigma^2)$$