5.1 Fixed vs. Random Effects

**Fixed:** want conclusions about factor levels we used  
**Random:** want conclusions about population those levels are sampled from  

The type of effect you want depends on your goals.

Fixed Effect Examples:
- vitamin C vs. no vitamin C
- three varieties of cotton

Random Effect Examples:
- 240 subjects in a clinical trial
- 8 rolls of ethylene vinyl acetate film

Fixed: \( Y_{ij} = \mu + \alpha_i + \epsilon_{ij} \) \( \sum \alpha_i = 0 \) \( i = 1,\ldots,I \) \( j = 1,\ldots,n \)  
Random: \( Y_{ij} = \mu + a_i + \epsilon_{ij} \) \( \alpha_i \sim N(0, \sigma^2_A) \) \( i = 1,\ldots,I \) \( j = 1,\ldots,n \)

5.2 ANOVA

\[ SS_{total} = SS_{treatment} + SS_{error} = SS_{between} + SS_{within} \]

\[ \sum_i \sum_j (Y_{ij} - \overline{Y}_.)^2 = \sum_i \sum_j (Y_{i.} - \overline{Y}_.)^2 + \sum_i \sum_j (Y_{ij} - \overline{Y}_{i.})^2 \]

\[ MS_{treatment} = \frac{SS_{treatment}}{I-1} \quad MS_{error} = \frac{SS_{error}}{I(n-1)} \]

\[ F = \frac{MS_{treatment}}{MS_{error}} \sim F_{I-1,I(n-1)} \]

Under \( H_0 : \sum \alpha^2_i = 0 \) or \( \sigma_A = 0 \)

Under \( H_A : F \sim (1 + \frac{n\sigma_A^2}{\sigma^2_E} X F_{I-1,I(k-1)}) \)
5.3 Blocks

5.3.1 Examples

Example 1:
- 4 varieties of tomato
- $y =$ yield
- Try all varieties on each farm to balance out farm quality
- e.g. Farm 1: ADBC, ..., Farm n: CDBA

Example 2:
- 3 diets for mice
- e.g. Litter 1: BCA, ..., Litter n: ABC

Example 3:
- Semiconductors - 24 wafers

5.3.2 Why block?

- Remove unwanted source of variation
- Reduce cost - could be cheaper to block, e.g. 4 at a time

Risk: blocking on unimportant variables is harmful in that you lose degrees of freedom

5.3.3 Model

$Y_{ti} = \mu + \beta_i + \tau_t + \epsilon_{ti}$ \quad t = 1,....,k treatments \quad i = 1,....,n blocks

Estimates:
- Grand mean = $\mu \approx \bar{Y}_.$
- Block effect = $\beta_i \approx \bar{Y}_{i.} - \bar{Y}_.$
- Treatment = $\tau_t \approx \bar{Y}_{.t} - \bar{Y}_.$
- Error = $\epsilon_{ti} \approx Y_{ti} - \bar{Y}_{i.} - \bar{Y}_{.t} + \bar{Y}_.$

Degrees of freedom:
- $SS_{blocks} : n - 1$
- $SS_{treatment} : k - 1$
- $SS_{error} : (n - 1)(k - 1)$

$F = \frac{MS_{treatment}}{MS_{error}} \sim F_{k-1,(n-1)(k-1)}$

Interactions cause lost power since they increase $MS_{error}$

5-2
5.3.4 Latin Square

Block to remove 2 sources of additive effects. Each row and each column defines one treatment.

Example:

\[
\begin{array}{cccc}
A & D & B & C \\
D & C & A & B \\
C & B & D & A \\
B & A & C & D \\
\end{array}
\]

Rows and columns could be geography, machine x operator, etc.

Additive model: \( Y_{ijt} = \mu + \beta_i + \delta_j + \tau_t + \epsilon_{ijt} \)

• Uses only \( k^2 \) observations, not \( k^3 \)

Method:

• Select a standard square
• Permute rows
• Permute columns
• Assign treatments to letters at random

OEIS Integer Sequence A000315

| \( k \) | \# standard squares | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 4 | 56 | 9408 | 16942080 |

Cyclic Square:

\[
\begin{array}{cccc}
A & B & C & D & \ldots \\
B & C & D & E & \ldots \\
C & D & E & F & \ldots \\
\ldots
\end{array}
\]

Be sure to permute! E.g. suppose row = subject, column = time. Then most people get D right after C unless we permute.
ANOVA:

- $\hat{\mu} = \bar{Y}_{..}$
- $\hat{\beta}_i = \bar{Y}_{i..} - \bar{Y}_{..}$
- $\hat{\delta}_j = \bar{Y}_{..j} - \bar{Y}_{..}$
- $\hat{\tau}_t = \bar{Y}_{..t} - \bar{Y}_{..}$

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
</tr>
</thead>
<tbody>
<tr>
<td>row</td>
<td>$k - 1$</td>
</tr>
<tr>
<td>column</td>
<td>$k - 1$</td>
</tr>
<tr>
<td>treatment</td>
<td>$k - 1$</td>
</tr>
<tr>
<td>error</td>
<td>$(k^2 - 1) - 3(k - 1) = (k - 2)(k - 1)$</td>
</tr>
<tr>
<td>total</td>
<td>$k^2 - 1$</td>
</tr>
</tbody>
</table>

Good: small experiment
Bad: fewer df, vulnerable to non-additivity and missing data NB: error degrees of freedom by subtraction