Settlershawite's Approximate F test

Construct 2 new MS terms

\[ MS' = MS_1 + \ldots + MS_n \]
\[ MS'' = MS_v + \ldots + MS_w \]

so that \( E(\text{MS}') - E(\text{MS}'') = \) constant times the desired variance component

For yesterday's example

Try \( MS' = MS_{A+B} + MS_{A\cdot B} \)
\[ MS'' = MS_{A+B} + MS_{A\cdot B} \]

Settlershawite proved that \( \frac{MS'}{MS''} \) has an approximate F distribution with \( p + q \) df,

where \[ p = \frac{(MS')^2}{\frac{MS_1^2}{df_1} + \ldots + \frac{MS_n^2}{df_n}} \]
\[ q = \frac{(MS'')^2}{\frac{MS_v^2}{df_v} + \ldots + \frac{MS_w^2}{df_w}} \]
Model for a nested design:

\[ y_{ijk} = \mu + \tau_i + \beta_{ij} + \epsilon_{ijk} \]

Parameter estimates, assuming both effects are fixed:

Assume \( \sum_{i=1}^{a} \tau_i = 0 \) and \( \sum_{j=1}^{b} \beta_{ij} = 0 \) for all \( i \)

\[ SSE = \sum_{i} \sum_{j} \sum_{k} (y_{ijk} - (\mu + \tau_i + \beta_{ij}))^2 \]
\[ \frac{\partial \text{SE}}{\partial \mu} = \sum_{i,j,k} \sum_{l} \sum_{m} \left[ y_{ijk} - (\mu + \tau_i + \beta_{jlm}) \right] < \gamma > \overset{\text{set}}{=} 0 \tag{5} \]

\[ y_{...} - N\mu - 0 - 0 = 0 \]

\[ \therefore \hat{\mu} = \frac{y_{...}}{N} = \bar{y}_{...} \]

\[ \frac{\partial \text{SE}}{\partial \tau_i} = \sum_{j,k} \sum_{l} \left[ y_{ijk} - (\mu + \tau_i + \beta_{jlm}) \right] < \gamma > \overset{\text{set}}{=} 0 \]

\[ y_{i...} - bn\mu - bn\tau_i - 0 = 0 \]

\[ \hat{\tau}_i = \frac{y_{i...} - bn\bar{y}_{...}}{bn} = \bar{y}_{i...} - \bar{y}_{...} \]

\[ \frac{\partial \text{SE}}{\partial \beta_{jlm}} = \sum_{k} \left[ y_{ijk} - (\mu + \tau_i + \beta_{jlm}) \right] < \gamma > \overset{\text{set}}{=} 0 \tag{6} \]

\[ y_{ij...} - n\mu - n\tau_i - n\beta_{jlm} = 0 \]

\[ \hat{\beta}_{jlm} = \frac{y_{ij...} - n\bar{y}_{...} - n(\bar{y}_{i...} - \bar{y}_{...})}{n} \]

\[ = \bar{y}_{ij...} - \bar{y}_{...} - \bar{y}_{i...} + \bar{y}_{...} \]

\[ = \bar{y}_{ij...} - \bar{y}_{i...} \]
Compare to the crossed model:
\[ \hat{\mu} = \bar{y}, \quad \hat{\tau}_i = \bar{y}_i - \bar{y} \]
\[ \hat{\beta}_j = \bar{y}_{j.} - \bar{y}, \quad \hat{\tau}_{i,j} = \bar{y}_{i,j} - \bar{y}_i - \bar{y}_{j.} + \bar{y} \]

\[ SS = df, \text{ using rule from Tuesday's class} \]
\[ df_a = a - 1 \quad \tau_i \]
\[ df_{B(A)} = (b - 1)a = ab - a \quad \beta_{j(c)} \]

Compare to crossed:
\[ df_a = a - 1, \quad df_b = b - 1, \quad df_{A,B} = (a - 1)(b - 1) \]

\[ SS_a = \sum \frac{y_{i.}^2}{bn} - \frac{\bar{y}^2}{n} \]

\[ SS_{B(A)} = \sum \sum \frac{y_{i,j}^2}{n} - \sum \frac{y_{i.}^2}{n} \]

Example:

- **A**: machine fixed, a levels
- **B**: operator random, nested within machine, b operators
- n replicates
<table>
<thead>
<tr>
<th>i</th>
<th>j</th>
<th>k</th>
<th>EMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>b \sigma_p^2 + \sigma^2</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>n</td>
<td>n \sigma_p^2 + \sigma^2</td>
</tr>
<tr>
<td>1</td>
<td>b</td>
<td>n</td>
<td>b n \sqrt{\frac{\sum x_i^2}{a-1}} + n \sigma_p^2 + \sigma^2</td>
</tr>
</tbody>
</table>

\[ F \text{ tests for } A: \frac{MS_A}{MS_{B(a)}} \]

\[ B(a): \frac{MS_{B(a)}}{MS_e} \]

Example: y measures hardness of material

2 different alloy formulations
3 different heats for each alloy
2 replicates of each alloy
2 replicates

A: alloy \( a = 2 \) fixed
B: heat \( b = 3 \) fixed nested within A
C: rejections \( c = 2 \) random nested with both A & B
\( n = 2 \)
\[ \begin{array}{cccc|c}
\text{EMS} & 12 \sum x^2 + 2 \sigma_y^2 + \sigma^2 \\
\text{EMS} & 4 \frac{\sum x^2}{\delta} + 2 \sigma_y^2 + \sigma^2 \\
\text{EMS} & 2 \sigma_y^2 + \sigma^2 \\
\text{EMS} & \sigma^2 \\
\text{EMS} & \frac{\text{MSSA}}{\text{MSE}} & \text{df} = 1, 6 \\
\text{EMS} & \frac{\text{MSB(A)}}{\text{MSE}} & \text{df} = 4, 6 \\
\text{EMS} & \frac{\text{MS(C(A))}}{\text{MSE}} & \text{df} = 6, 12 \\
\end{array} \]

\[
\begin{array}{ll}
\text{Source} & \text{df} \\
A & 1 \\
B(A) & 4 \\
C(A) & 6 \\
\text{E(R)} & 12 \\
\text{TOT} & 23 \\
\end{array}
\]

\[ \text{HW} \# 2 \]
\[ \# 13.17 \]
13.17. Consider a four-factor factorial experiment where factor A is at a levels, factor B is at b levels, factor C is at c levels, factor D is at d levels, and there are n replicates. Write down the sums of squares, the degrees of freedom, and the expected mean squares for the following cases. Assume the restricted model for all mixed models. You may use a computer package such as Minitab.

(a) A, B, C, and D are fixed factors.

(b) A, B, C, and D are random factors.

(c) A is fixed and B, C, and D are random.

(d) A and B are fixed and C and D are random.

(e) A, B, and C are fixed and D is random.

Do exact tests exist for all effects? If not, propose test statistics for those effects that cannot be directly tested.