13. The Cochran-Satterthwaite Approximation for Linear Combinations of Mean Squares
Suppose $M_1, \ldots, M_k$ are independent mean squares and that

$$\frac{d_i M_i}{E(M_i)} \sim \chi^2_{d_i} \quad \forall \ i = 1, \ldots, k.$$ 

It follows that

$$E \left[ \frac{d_i M_i}{E(M_i)} \right] = d_i, \quad \text{Var} \left[ \frac{d_i M_i}{E(M_i)} \right] = 2d_i, \quad \text{and} \quad M_i \sim \frac{E(M_i)}{d_i} \chi^2_{d_i}$$

for all $i = 1, \ldots, k$. 
Consider the random variable

\[ M = a_1 M_1 + a_2 M_2 + \cdots + a_k M_k, \]

where \( a_1, a_2, \ldots, a_k \) are known constants in \( \mathbb{R} \).

Note that \( M \) is a linear combination of scaled \( \chi^2 \) random variables.

The Cochran-Satterthwaite approximation works by assuming that \( M \) is approximately distributed as a scaled \( \chi^2 \), just like each of the variables in the linear combination.

\[
\frac{dM}{E(M)} \sim \chi_d^2 \iff M \sim \frac{E(M)}{d} \chi_d^2.
\]

What choice for \( d \) makes the approximation most reasonable?
If

\[ M \sim \frac{E(M)}{d} \chi_d^2, \]

then

\[
\text{Var}(M) \approx \left( \frac{E(M)}{d} \right)^2 \text{Var} \left( \chi_d^2 \right) = \left( \frac{E(M)}{d} \right)^2 (2d) = 2 \left[ \frac{E(M)}{d} \right]^2 \frac{d}{d} \\
\approx \frac{2M^2}{d}.
\]
Now note that

\[
\text{Var}(M) = a_1^2 \text{Var}(M_1) + \cdots + a_k^2 \text{Var}(M_k)
\]

\[
= a_1^2 \left[ \frac{E(M_1)}{d_1} \right]^2 2d_1 + \cdots + a_k^2 \left[ \frac{E(M_k)}{d_k} \right]^2 2d_k
\]

\[
= 2 \sum_{i=1}^k \frac{a_i^2 [E(M_i)]^2}{d_i}
\]

\[
\approx 2 \sum_{i=1}^k \frac{a_i^2 M_i^2}{d_i}.
\]
Equating these two variance approximations yields

\[
\frac{2M^2}{d} = 2 \sum_{i=1}^{k} a_i^2 M_i^2 / d_i.
\]

Solving for \(d\) yields

\[
d = \frac{M^2}{\sum_{i=1}^{k} a_i^2 M_i^2 / d_i} = \left( \frac{\sum_{i=1}^{k} a_i M_i}{\sum_{i=1}^{k} a_i^2 M_i^2 / d_i} \right)^2.
\]

This is the Cochran-Satterthwaite formula for the approximate degrees of freedom associated with the linear combination of mean squares defined by \(M\).
Recall the first example from the last slide set.

\[ y = \begin{bmatrix} y_{111} \\ y_{121} \\ y_{211} \\ y_{212} \end{bmatrix}, \quad X = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad Z = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \]

\[ X_1 = 1, \quad X_2 = X, \quad X_3 = Z \]

\[ y' (P_2 - P_1) y + y' (P_3 - P_2) y + y' (I - P_3) y = y' (I - P_1) y \]
Expected Mean Squares

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>EMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>trt</td>
<td>$1.5\sigma_u^2 + \sigma_e^2 + (\tau_1 - \tau_2)^2$</td>
</tr>
<tr>
<td>xu(trt)</td>
<td>$\sigma_u^2 + \sigma_e^2$</td>
</tr>
<tr>
<td>ou(xu, trt)</td>
<td>$\sigma_e^2$</td>
</tr>
</tbody>
</table>

\[
E(1.5MS_{xu(trt)} - 0.5MS_{ou(xu, trt)}) = 1.5(\sigma_u^2 + \sigma_e^2) - 0.5\sigma_e^2 = 1.5\sigma_u^2 + \sigma_e^2
\]
An Approximate $F$ Test

The statistic

$$F = \frac{MS_{trt}}{1.5MS_{xu(trt)} - 0.5MS_{ou(xu, trt)}}$$

is approximately $F$ distributed with 1 numerator degree of freedom and denominator degrees of freedom approximated by the Cochran-Satterthwaite Method:

$$d = \frac{(1.5MS_{xu(trt)} - 0.5MS_{ou(xu, trt)})^2}{(1.5)^2 [MS_{xu(trt)}]^2 + (-0.5)^2 [MS_{ou(xu, trt)}]^2}.$$
SAS Code for Example

data d;
  input trt xu y;
  cards;
1 1 6.4
1 2 4.2
2 1 1.5
2 1 0.9
;
run;
SAS Code for Example

```sas
proc mixed method=type1;
  class trt xu;
  model y=trt / ddfm=satterthwaite;
  random xu(trt);
run;
```
The Mixed Procedure

Model Information

Data Set WORK.D
Dependent Variable y
Covariance Structure Variance Components
Estimation Method Type 1
Residual Variance Method Factor
Fixed Effects SE Method Model-Based
Degrees of Freedom Method Satterthwaite

Class Level Information

Class Levels Values

trt 2 1 2
xu 2 1 2
## Dimensions

- Covariance Parameters: 2
- Columns in X: 3
- Columns in Z: 3
- Subjects: 1
- Max Obs Per Subject: 4

### Number of Observations

- Number of Observations Read: 4
- Number of Observations Used: 4
- Number of Observations Not Used: 0
Type 1 Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>trt</td>
<td>1</td>
<td>16.810000</td>
<td>16.810000</td>
</tr>
<tr>
<td>xu(trt)</td>
<td>1</td>
<td>2.420000</td>
<td>2.420000</td>
</tr>
<tr>
<td>Residual</td>
<td>1</td>
<td>0.180000</td>
<td>0.180000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>Expected Mean Square</th>
<th>Error Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>trt</td>
<td>Var(Residual)+1.5</td>
<td>1.5 MS(xu(trt))</td>
</tr>
<tr>
<td></td>
<td>Var(xu(trt))+Q(trt)</td>
<td>- 0.5 MS(Residual)</td>
</tr>
<tr>
<td>xu(trt)</td>
<td>Var(Residual)+Var(xu(trt))</td>
<td>MS(Residual)</td>
</tr>
<tr>
<td>Residual</td>
<td>Var(Residual)</td>
<td>.</td>
</tr>
</tbody>
</table>
Degrees of Freedom for Satterthwaite Approximation

\[ d = \frac{(1.5MS_{xu(trt)} - 0.5MS_{ou(xu, trt)})^2}{(1.5)^2 [MS_{xu(trt)}]^2 + (-0.5)^2 [MS_{ou(xu, trt)}]^2} \]

\[ = \frac{(1.5 \times 2.42 - 0.5 \times 0.18)^2}{(1.5)^2 [2.42]^2 + (-0.5)^2 [0.18]^2} \]

\[ = 0.9504437 \]
Type 1 Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>trt</td>
<td>0.9504</td>
<td>4.75</td>
<td>0.2840</td>
</tr>
<tr>
<td>xu(trt)</td>
<td>1</td>
<td>13.44</td>
<td>0.1695</td>
</tr>
<tr>
<td>Residual</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

Covariance Parameter Estimates

<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>xu(trt)</td>
<td>2.2400</td>
</tr>
<tr>
<td>Residual</td>
<td>0.1800</td>
</tr>
</tbody>
</table>
Concluding Remarks

This example was chosen to be small so that we could write out all the data and see how each observation was involved in the analysis.

Because of the very low sample size, it would be surprising if the approximate $F$ test worked well for this example.

It would be difficult to draw any meaningful conclusions with 4 observations of the response on 3 experimental units.

We will see more practically relevant examples where the samples sizes are larger and the approximate $F$-based inferences may be reasonable.
Concluding Remarks

In more complicated examples, there may be more than one linear combination of mean squares with the desired expectation.

In such cases, linear combinations with non-negative coefficients are recommended over those with a mix of positive and negative coefficients.