13. The Cochran-Satterthwaite Approximation for Linear Combinations of Mean Squares

Suppose $M_1, ..., M_k$ are independent mean squares and that

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

It follows that

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

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 χ^2 random variables.

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

$$\frac{dM}{E(M)} \overset{\bullet}{\sim} \chi^2_d \iff M \overset{\bullet}{\sim} \frac{E(M)}{d}\chi^2_d.$$

What choice for d makes the approximation most reasonable?

If

$$M \overset{\bullet}{\sim} \frac{E(M)}{d}\chi^2_d,$$

then

$$\begin{aligned}
\mathrm{Var}(M) &\approx \left(\frac{E(M)}{d}\right)^2 \mathrm{Var}\left(\chi^2_d\right) \\
&= \left(\frac{E(M)}{d}\right)^2 (2d) \\
&= \frac{2\left[E(M)\right]^2}{d} \\
&\approx \frac{2M^2}{d}.
\end{aligned}$$

Now note that

$$\begin{aligned}
\text{Var}(M) &= a_1^2 \text{Var}(M_1) + \cdots + a_k^2 \text{Var}(M_K) \\
&= a_1^2 \left[\frac{E(M_1)}{d_1} \right]^2 2d_1 + \cdots + a_k^2 \left[\frac{E(M_k)}{d_k} \right]^2 2d_k \\
&= 2 \sum_{i=1}^{k} \frac{a_i^2 [E(M_i)]^2}{d_i} \\
&\approx 2 \sum_{i=1}^{k} a_i^2 M_i^2 / d_i .
\end{aligned}$$

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

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

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.

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

$$\boldsymbol{X}_1 = \boldsymbol{1}, \qquad \boldsymbol{X}_2 = \boldsymbol{X}, \qquad \boldsymbol{X}_3 = \boldsymbol{Z}$$

$$\boldsymbol{y}'(\boldsymbol{P}_2 - \boldsymbol{P}_1)\boldsymbol{y} + \boldsymbol{y}'(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{y} + \boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P}_3)\boldsymbol{y} = \boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P}_1)\boldsymbol{y}$$

Expected Mean Squares

SOURCE	EMS
trt	$1.5\sigma_u^2 + \sigma_e^2 + (\tau_1 - \tau_2)^2$
$xu(trt)$	$\sigma_u^2 + \sigma_e^2$
$ou(xu, trt)$	σ_e^2

$$\begin{aligned} E(1.5MS_{xu(trt)} - 0.5MS_{ou(xu,trt)}) &= 1.5(\sigma_u^2 + \sigma_e^2) - 0.5\sigma_e^2 \\ &= 1.5\sigma_u^2 + \sigma_e^2 \end{aligned}$$

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:

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

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

```
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

                            Sum of
Source          DF         Squares      Mean Square

trt              1       16.810000        16.810000
xu(trt)          1        2.420000         2.420000
Residual         1        0.180000         0.180000


Source     Expected Mean Square          Error Term

trt        Var(Residual)+1.5             1.5 MS(xu(trt))
           Var(xu(trt))+Q(trt)           - 0.5 MS(Residual)

xu(trt)    Var(Residual)+Var(xu(trt))    MS(Residual)

Residual   Var(Residual)                 .
```

Degrees of Freedom for Satterthwaite Approximation

$$
\begin{aligned}
d &= \frac{(1.5MS_{xu(trt)} - 0.5MS_{ou(xu,trt)})^2}{(1.5)^2\left[MS_{xu(trt)}\right]^2 + (-0.5)^2\left[MS_{ou(xu,trt)}\right]^2} \\
&= \frac{(1.5 \times 2.42 - 0.5 \times 0.18)^2}{(1.5)^2\,[2.42]^2 + (-0.5)^2\,[0.18]^2} \\
&= 0.9504437
\end{aligned}
$$

Type 1 Analysis of Variance

Source	Error DF	F Value	Pr > F
trt	0.9504	4.75	0.2840
xu(trt)	1	13.44	0.1695
Residual	.	.	.

Covariance Parameter Estimates

Cov Parm	Estimate
xu(trt)	2.2400
Residual	0.1800

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.