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

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

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

It follows that

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

for all $i = 1, \ldots, k$.

Scaled $\chi^2_{d_i}$

Chapt 12 : $1.5\, MS(xu(trt)) - 0.5\, MSE$

Consider the random variable

$$M = a_1 M_1 + a_2 M_2 + \cdots + a_k M_k, \qquad (1)$$

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.

On slide 2 we state that $\frac{d_i M_i}{E(M_i)} \sim \chi^2_{d_i}$

while we now switch to approx

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

Why? df d_i on slide 2 are integer values. They have to be integer values by definition of a χ^2-random variable.

The a_i in $M = a_1 M_1 + \ldots + a_k M_k$ are typically not integers $\Rightarrow$ hence, we only have an approx. χ^2-r.v.

What choice for d makes the approximation most reasonable?

The Cochran-Satterthwaite formula for the approximate degrees of freedom associated with the linear combination of mean squares defined by M is

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

Why?

Consider the definition/approximation of M in (1) and (2):

- $M = a_1 M_1 + a_2 M_2 + \cdots + a_k M_k$ and (1)

- $M \overset{approx.}{\sim} \frac{\mathrm{E}(M)}{d} \chi^2_d.$ (2)

- Calculate $\mathrm{Var}(M)$ using M as defined in (1) and also using its approximation given in (2)

- Both variances are functions of d. We equate both variances and solve for d.

- The details are shown on the next two slides.

from (2) on slide 5

see comment added on slide 3

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

slide 2

* Technically, the last line is "=", however, the overall result is approximate because that is how we started out in line 1.

And

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

Equating these two variance approximations yields

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

and solving for d yields

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

Recall the first example from the last slide set.

account for trt 2

exp. unit 1 exp. unit 3

trt 1

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

trt 2

exp. unit 2

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

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

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

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

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:

from SAS, R,

$$d = \frac{(1.5MS_{xu(trt)} - 0.5MS_{ou(xu,trt)})^2}{(1.5)^2\left[MS_{xu(trt)}\right]^2/1 + (-0.5)^2\left[MS_{ou(xu,trt)}\right]^2/1}.$$

df are 1 here due instead of 2=tn-t
because the data are not balanced

SAS Code for Example

Secondly, we only have 4 data points, allowing for "only" df=3 that we can use for testing in the ANOVA table; see table on slide 16

```
data d;
   input trt xu y;
   cards;
1 1 6.4
1 2 4.2
2 1 1.5
2 1 0.9
;
run;
```

two distinct exp. units

tell SAS by properly setting up the random statement

SAS Code for Example

Proc Glimmix or Proc Mixed

random int / subject = xu

does the same

Methods of Moments

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

tells SAS that xu are nested within treatment

random effects

```
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
```

$\sigma_u^2 \quad \sigma_e^2$

Dimensions

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

SAS recognizes the 3 distinct exp. units

Number of Observations

Number of Observations Read	4
Number of Observations Used	4
Number of Observations Not Used	0

Type 1 Analysis of Variance

$\hat{\sigma}_u^2 = 2.42 - 0.18$
$= 2.24$

Source	DF	Sum of Squares	Mean Square
trt	1	16.810000	16.810000
xu(trt)	1	2.420000	2.420000 $= \hat{\sigma}_e^2 + \hat{\sigma}_u^2$
Residual	1	0.180000	0.180000 $\hat{\sigma}_e^2$

We only have 3 degrees of freedom to use $\Rightarrow$ 1 df for each Source

estimate

Source	Expected Mean Square	Error Term
trt	Var(Residual)+1.5 Var(xu(trt))+Q(trt) $(\tau_1-\tau_2)^2$	1.5 MS(xu(trt)) - 0.5 MS(Residual)
xu(trt)	Var(Residual)+Var(xu(trt)) $\sigma_e^2 + \sigma_u^2$	MS(Residual)
Residual	Var(Residual) σ_e^2	

estimate

Degrees of Freedom for Satterthwaite Approximation

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

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	.	.	.

d using the CS approximat.

Covariance Parameter Estimates

Cov Parm	Estimate
xu(trt)	2.2400
Residual	0.1800

$\hat{\sigma}_u^2$

$\hat{\sigma}_e^2$

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.