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

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

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

Consider the random variable

$$M = a_1M_1 + a_2M_2 + \cdots + a_kM_k,$$

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.

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

What is a good choice for d that will make this approximation reasonable?

If

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

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

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

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 M.