Power in Mixed Effects

Gary W. Oehlert

School of Statistics
University of Minnesota

December 1, 2014

Power is an important aspect of designing an experiment; we now return to power in mixed effects.

We will compute power for "old school" tests. This technique works for balanced designs; it will be exact for some situations with REML and approximate in other situations.

Modeling Assumptions

For pure random terms $\alpha\beta_{ij}$ where all contributing factors are random, we assume all $\alpha\beta_{ij}$s are independent of each other.

For mixed terms $\alpha\beta_{ij}$, say where A is fixed and B is random, we have a choice:

Unrestricted assumptions say that all elements of $\alpha\beta_{ij}$ are independent.

Restricted assumptions say that elements of $\alpha\beta_{ij}$ will add to zero across any fixed subscript (i in this case) but are otherwise independent.

Restricted assumptions induce negative correlation among some random effects:

Under restricted assumptions, two random effects from the same mixed term are negatively correlated if all of their subscripts corresponding to random factors are the same.

Otherwise, they are independent.

The text usually defaults to restricted assumptions, but either could be appropriate depending on the situation, or something else could be better still.

It is usually very tricky to decide which mixed modeling assumptions are appropriate.

In general, unrestricted assumptions are more conservative (typically more difficult to reject the null.)

The `lme` and `lmer` functions in R fit using the unrestricted model assumptions.

Via heroic effort, one can make `lme` fit some models under the restricted assumptions.

Give R predilections, we will concentrate on the unrestricted approach.

Anova and Expected Mean Squares

Old school testing in mixed effects proceeds as follows:

1. Compute an Anova table as if everything in the model is a fixed effect.
2. Compute the expectation of every mean square (the expected mean squares) using the complete Hasse diagram.
3. Use the Hasse diagram (or EMS) to determine the correct ratio of MS (correct F test) for every term of interest.
4. Do the tests.

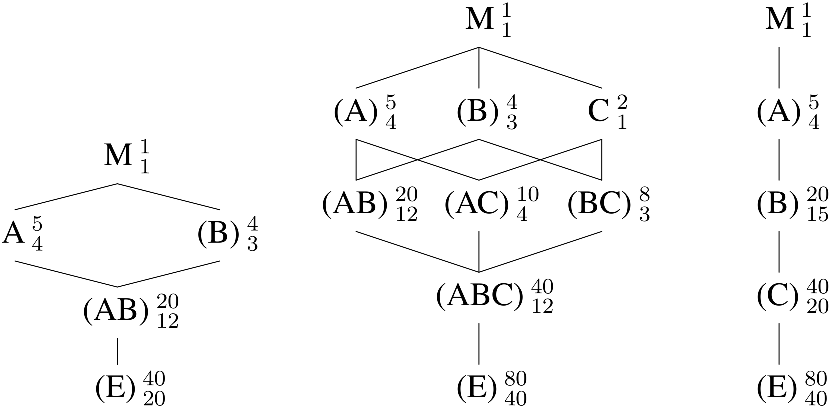
We need the EMS and DF for the F-test to do power.

The complete Hasse diagram includes super- and subscripts on every term.

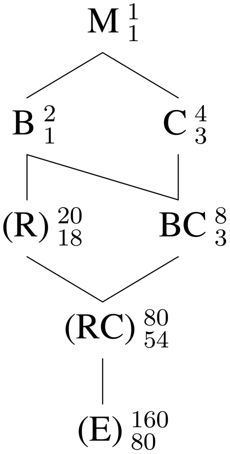
For each node on the diagram, add a superscript that indicates the number of different levels of the effect in that term.

For each node on the diagram, add a subscript that indicates the degrees of freedom. Compute the df for a term U by starting with the superscript for U and subtracting the subscripts (df) for all terms above U.

A has 5 levels, B has 4 levels, C has 2 levels, 2 replications.



Cheese raters



1. The representative element for a random term is its variance.

2. The representative element for a fixed term is the sum of the squared fixed effects divided by degrees of freedom.

3. Contribution from a term is N, divided by the superscript, times the representative element.

4. Using unrestricted model assumptions, the EMS for a term is the contribution from that term and all random terms below it.

A fixed, B random, crossed (first diagram)

$$
\begin{array}{llcl}
MS_E & \frac{40}{40}\sigma^2 & = & \sigma^2 \\
MS_{AB} & \frac{40}{40}\sigma^2 + \frac{40}{20}\sigma^2_{\alpha\beta} & = & \sigma^2 + 2\sigma^2_{\alpha\beta} \\
MS_A & \frac{40}{40}\sigma^2 + \frac{40}{20}\sigma^2_{\alpha\beta} + \frac{40}{5}\frac{\sum_{i=1}^5 \alpha_i{}^2}{4} & = & \sigma^2 + 2\sigma^2_{\alpha\beta} + 8\frac{\sum_{i=1}^5 \alpha_i{}^2}{4} \\
MS_B & \frac{40}{40}\sigma^2 + \frac{40}{20}\sigma^2_{\alpha\beta} + \frac{40}{4}\sigma^2_{\beta} & = & \sigma^2 + 2\sigma^2_{\alpha\beta} + 10\sigma^2_{\beta}
\end{array}
$$

A random, B random, C fixed, crossed (second diagram)

$$
\begin{array}{llcl}
MS_E & \frac{80}{80}\sigma^2 & = & \sigma^2 \\
MS_{ABC} & \frac{80}{80}\sigma^2 + \frac{80}{40}\sigma^2_{\alpha\beta\gamma} & = & \sigma^2 + 2\sigma^2_{\alpha\beta\gamma} \\
MS_{AB} & \frac{80}{80}\sigma^2 + \frac{80}{40}\sigma^2_{\alpha\beta\gamma} + \frac{80}{20}\sigma^2_{\alpha\beta} & = & \sigma^2 + 2\sigma^2_{\alpha\beta\gamma} + 4\sigma^2_{\alpha\beta} \\
MS_{BC} & \frac{80}{80}\sigma^2 + \frac{80}{40}\sigma^2_{\alpha\beta\gamma} + \frac{80}{8}\sigma^2_{\beta\gamma} & = & \sigma^2 + 2\sigma^2_{\alpha\beta\gamma} + 10\sigma^2_{\beta\gamma} \\
MS_{AC} & \frac{80}{80}\sigma^2 + \frac{80}{40}\sigma^2_{\alpha\beta\gamma} + \frac{80}{10}\sigma^2_{\alpha\gamma} & = & \sigma^2 + 2\sigma^2_{\alpha\beta\gamma} + 8\sigma^2_{\alpha\gamma}
\end{array}
$$

A random, B random, C fixed, crossed (second diagram), continued.

$$MS_C \quad \frac{80}{80}\sigma^2 + \frac{80}{40}\sigma^2_{\alpha\beta\gamma} + \frac{80}{10}\sigma^2_{\alpha\gamma} + \frac{80}{8}\sigma^2_{\beta\gamma} + \frac{80}{2}\frac{\sum_{i=1}^{2}{\gamma_i}^2}{1} =$$

$$\sigma^2 + 2\sigma^2_{\alpha\beta\gamma} + 8\sigma^2_{\alpha\gamma} + 10\sigma^2_{\beta\gamma} + 40\frac{\sum_{i=1}^{2}{\gamma_i}^2}{1}$$

$$MS_B \quad \frac{80}{80}\sigma^2 + \frac{80}{40}\sigma^2_{\alpha\beta\gamma} + \frac{80}{20}\sigma^2_{\alpha\beta} + \frac{80}{8}\sigma^2_{\beta\gamma} + \frac{80}{4}\sigma^2_{\beta} =$$

$$\sigma^2 + 2\sigma^2_{\alpha\beta\gamma} + 4\sigma^2_{\alpha\beta} + 10\sigma^2_{\beta\gamma} + 20\sigma^2_{\beta}$$

$$MS_A \quad \frac{80}{80}\sigma^2 + \frac{80}{40}\sigma^2_{\alpha\beta\gamma} + \frac{80}{20}\sigma^2_{\alpha\beta} + \frac{80}{10}\sigma^2_{\alpha\gamma} + \frac{80}{5}\sigma^2_{\alpha} =$$

$$\sigma^2 + 2\sigma^2_{\alpha\beta\gamma} + 4\sigma^2_{\alpha\beta} + 8\sigma^2_{\alpha\gamma} + 16\sigma^2_{\alpha}$$

A random, B random, C random, fully nested (third diagram).

$$
\begin{array}{llcl}
MS_E & \frac{80}{80}\sigma^2 & = & \sigma^2 \\
MS_C & \frac{80}{80}\sigma^2 + \frac{80}{40}\sigma_\gamma^2 & = & \sigma^2 + 2\sigma_\gamma^2 \\
MS_B & \frac{80}{80}\sigma^2 + \frac{80}{40}\sigma_\gamma^2 + \frac{80}{20}\sigma_\beta^2 & = & \sigma^2 + 2\sigma_\gamma^2 + 4\sigma_\beta^2 \\
MS_A & \frac{80}{80}\sigma^2 + \frac{80}{40}\sigma_\gamma^2 + \frac{80}{20}\sigma_\beta^2 + \frac{80}{5}\sigma_\alpha^2 & = & \sigma^2 + 2\sigma_\gamma^2 + 4\sigma_\beta^2 + 16\sigma_\alpha^2
\end{array}
$$

Cheese raters.

$$MS_E \quad \frac{160}{160}\sigma^2 = \sigma^2$$

$$MS_{RC} \quad \frac{160}{160}\sigma^2 + \frac{160}{80}\sigma^2_{\rho\gamma} = \sigma^2 + 2\sigma^2_{\rho\gamma}$$

$$MS_{BC} \quad \frac{160}{160}\sigma^2 + \frac{160}{80}\sigma^2_{\rho\gamma} + \frac{160}{8}\frac{\sum_{i,j=1}^{2,4}\beta\gamma^2_{jk}}{3} = \sigma^2 + 2\sigma^2_{\rho\gamma} + 20\frac{\sum_{i,j=1}^{2,4}\beta\gamma^2_{jk}}{3}$$

$$MS_C \quad \frac{160}{160}\sigma^2 + \frac{160}{80}\sigma^2_{\rho\gamma} + \frac{160}{4}\frac{\sum_{k=1}^{4}\gamma^2_k}{3} = \sigma^2 + 2\sigma^2_{\rho\gamma} + 40\frac{\sum_{j=1}^{4}\gamma^2_k}{3}$$

$$MS_R \quad \frac{160}{160}\sigma^2 + \frac{160}{80}\sigma^2_{\rho\gamma} + \frac{160}{20}\sigma^2_{\rho} = \sigma^2 + 2\sigma^2_{\rho\gamma} + 8\sigma^2_{\rho}$$

$$MS_B \quad \frac{160}{160}\sigma^2 + \frac{160}{80}\sigma^2_{\rho\gamma} + \frac{160}{20}\sigma^2_{\rho} + \frac{160}{2}\frac{\sum_{j=1}^{2}\beta^2_j}{1}$$

$$= \sigma^2 + 2\sigma^2_{\rho\gamma} + 8\sigma^2_{\rho} + 80\frac{\sum_{j=1}^{2}\beta^2_j}{1}$$

A has a levels, B has b levels, C has c levels, n replications.

Notice the pattern[1] of the integer multiplier for regular models like these:

A	nbc
B	nac
C	nab
AB	nc
AC	nb
BC	na
ABC	n

The multiplier is the product of the levels not in the term.

[1] Also notice the pattern that computing EMS is pretty damn tedious.

F Tests

In the old school approach, we test a null hypothesis such as $\sigma_\alpha^2 = 0$ or $0 = \sum \beta_j^2$ by

- Finding two Mss with EMSs that differ by a multiple of the item of interest.
- Computing the F ratio of those two MS and using the df for the two MSs to find a p-value.

I asserted that as if it were always possible; this is <u>not</u> always possible.

A fixed, B random, crossed (first diagram)

Item	Num. MS	Den. MS
$\sum \alpha_i^2$	MS_A	MS_{AB}
σ_β^2	MA_B	MS_{AB}
$\sigma_{\alpha\beta}^2$	MS_{AB}	MS_E

No trouble here.

Fully nested design (third diagram)

Item	Num. MS	Den. MS
σ^2_α	MS_A	MS_B
σ^2_β	MA_B	MS_C
σ^2_γ	MS_C	MS_E

No trouble here.

Cheese raters

Item	Num. MS	Den. MS
$\sum_j \beta_j^2$	MS_B	MS_R
$\sum_k \gamma_k^2$	MS_C	MS_{RC}
$\sum_{j,k} \beta\gamma_{jk}^2$	MS_{BC}	MS_{RC}
σ_ρ^2	MS_R	MS_{RC}
$\sigma_{\rho\gamma}^2$	MS_{RC}	MS_E

No trouble here.

And that brings us to the second diagram.

Item	Num. MS	Den. MS
$\sum_k \gamma_k^2$	—	—
σ_β^2	—	—
σ_α^2	—	—
$\sigma_{\alpha\beta}^2$	MS_{AB}	MS_{ABC}
$\sigma_{\alpha\gamma}^2$	MS_{AC}	MS_{ABC}
$\sigma_{\beta\gamma}^2$	MS_{BC}	MS_{ABC}
$\sigma_{\alpha\beta\gamma}^2$	MS_{ABC}	MS_E

In the second diagram/model, there are no ordinary F-tests for main effects!

Looking at the Hasse diagrams, the denominator for a term (using unrestricted model assumptions) is the first random term below the term of interest.

If there is more than one random term you can get to without going through another random term, then there is no exact test.

We will eventually get to approximate tests for these cases.

Power for Fixed Effects

Recall that the noncentrality parameter controls power for fixed effects (along with degrees of freedom and the error rate).

The expected mean square for a fixed has the form:

$$\text{random stuff} + \frac{N}{\text{superscript}} \times \frac{\text{sum of squared effects}}{df}$$

The noncentrality parameter is then

$$\frac{\frac{N}{\text{superscript}} \times \text{sum of squared effects}}{\text{random stuff}}$$

The random stuff goes to the denominator and the df disappears.

Try this approach on Chapter 7 problems . . . it works.

For testing A in the first diagram, the EMS is

$$\sigma^2 + 2\sigma^2_{\alpha\beta} + 8\frac{\sum_{i=1}^{5} {\alpha_i}^2}{4}$$

and the noncentrality parameter is

$$\frac{8\sum_{i=1}^{5} {\alpha_i}^2}{\sigma^2 + 2\sigma^2_{\alpha\beta}}$$

Note that you need to know, or make assumptions about, two variances to compute the NCP.

In more general form, the NCP for this test is

$$\frac{nb\sum_{i=1}^{5}{\alpha_i}^2}{\sigma^2 + n\sigma^2_{\alpha\beta}}$$

You cannot always make a noncentrality parameter in a mixed effects model arbitrarily large by increasing n.

For this problem, you need to increase b, the number of levels of the random term B, to make the power go to 1.

Power for Random Effects

Power for random effects is actually easier than power for fixed effects.

Suppose we want to test $H_0 : \sigma_\eta^2 = 0$.

We have two MS with $EMS_1 = \tau + k\sigma_\eta^2$ and $EMS_2 = \tau$.

The F test is MS_1/MS_2 with ν_1 and ν_2 df.

Under the null,

$$\frac{MS_1}{MS_2} \sim \frac{\tau + k \times 0}{\tau} F_{\nu_1,\nu_2} = F_{\nu_1,\nu_2}$$

Under the alternative,

$$\frac{MS_1}{MS_2} \sim \frac{\tau + k\sigma_\eta^2}{\tau} F_{\nu_1,\nu_2}$$

We reject H_0 if $MS_1/MS_2 > F_{\mathcal{E},\nu_1,\nu_2}$.

Looking at the distribution under the alternative, we reject when

$$\frac{\tau + k\sigma_\eta^2}{\tau} F_{\nu_1,\nu_2} > F_{\mathcal{E},\nu_1,\nu_2}$$

or, put another way, when

$$F_{\nu_1,\nu_2} > \frac{\tau}{\tau + k\sigma_\eta^2} F_{\mathcal{E},\nu_1,\nu_2}$$

Consider testing $H_0 : \sigma^2_{\alpha\gamma} = 0$ in the second model (fully crossed three-way design).

The test is MS_{AC}/MS_{ABC}. The df are 4 and 12.

Test at $\mathcal{E} = .01$ assuming $\sigma^2 = 1$, $\sigma^2_{\alpha\beta\gamma} = 2$, and $\sigma^2_{\alpha\gamma} = .5$.

The EMS are:
$EMS_{AC} = \sigma^2 + 2\sigma^2_{\alpha\beta\gamma} + 8\sigma^2_{\alpha\gamma} = 9$
$EMS_{ABC} = \sigma^2 + 2\sigma^2_{\alpha\beta\gamma} = 5$

$F_{.01,4,12} = 5.41$ and $5/9 \times 5.41 = 3.01$

Power is the probability that F with 4 and 12 df is larger than 3.01, which is .062 (which is not much power).

Approximate Tests

When there is no F test we can usually construct an approximate test.

We want to find a sum of two MS for the numerator and a sum of two MS for the denominator such that the sum of the EMS on the top is equal to the sum of the EMS on the bottom plus the term of interest.

On the Hasse diagram, if there are two random terms immediately below the term of interest, the bottom will be the sum of those two random terms, and the top will be the sum of the term of interest plus the term where the denominator terms "intersect."

For testing σ^2_α in our second example, the numerator MS are:

$$MS_A,\ EMS_A = \sigma^2 + 2\sigma^2_{\alpha\beta\gamma} + 4\sigma^2_{\alpha\beta} + 8\sigma^2_{\alpha\gamma} + 16\sigma^2_\alpha$$
$$MS_{ABC},\ EMS_{ABC} = \sigma^2 + 2\sigma^2_{\alpha\beta\gamma}$$

The denominator MS are:

$$MS_{AB},\ EMS_{AB} = \sigma^2 + 2\sigma^2_{\alpha\beta\gamma} + 4\sigma^2_{\alpha\beta}$$
$$MS_{AC},\ EMS_{ABC} = \sigma^2 + 2\sigma^2_{\alpha\beta\gamma} + 8\sigma^2_{\alpha\gamma}$$

Numerator and denominator sums are:

$$EMS_A + EMS_{ABC} = 2\sigma^2 + 4\sigma^2_{\alpha\beta\gamma} + 4\sigma^2_{\alpha\beta} + 8\sigma^2_{\alpha\gamma} + 16\sigma^2_\alpha$$
$$EMS_{AB} + EMS_{AC} = 2\sigma^2 + 4\sigma^2_{\alpha\beta\gamma} + 4\sigma^2_{\alpha\beta} + 8\sigma^2_{\alpha\gamma}$$

The approximate test is a decent statistic, the problem is that it doesn't follow an F distribution (or any other standard distribution).

What we do is treat it as if it follows an F distribution, and compute some approximate degrees of freedom separately for the numerator and denominator.

The df approximation is called the Satterthwaite approximation.

Suppose we're trying to get an approximate df for $MS_1 + MS_2$ with ν_1 and ν_2 df.

Let $E_i = MS_i$ and $V_i = 2E_i^2/\nu_i$.

The the approximate df for the sum is:

$$\frac{2(E_1 + E_2)^2}{V_1 + V_2}$$

For computing power, let $E_i = EMS_i$. (In fact, even for data E_i is supposed to be EMS_i, we're just using MS_i to estimate EMS_i.)