

Statistics 514: Design of Experiments

Topic 10 Summary

Random Effects Factors

- Use EMS to decide tests (denominator should have components except component for testing) and estimates (via subtraction and dividing by scalar)
- `proc mixed` and `proc varcomp (type1)` give unrestricted analysis.
- Use rules for finding expected mean squares (and estimators and tests)
 - restricted model
- Can also use Restricted Maximum Likelihood Estimation
- Hasse diagrams useful graphical method for forming tests for both restricted and unrestricted models

Restricted Model

- Interactions act as (correlated) random effects with fixed restrictions.

Other Considerations

- *Usual diagnostics* – may want to check if random effects normally distributed
- *Multiple Comparisons* – use MS (and associated df) to math expected value of SE
- *Sample Size* – like usual factorial except df_{den} change

Restricted vs. Unrestricted Models

- *Unrestricted*
 - Done in SAS
 - Random effects use interaction denominator.
- Interpretation
 - Expect negative correlations among interaction terms?

Methods for mixed models

- Think in terms of linear model – patterns in covariance
- Use (restricted) maximum likelihood, not LS

Approximate Tests

- Use criteria for deleting variables
- Pool SS and df into error
- Test largest interaction first

Approximate Tests (Satterthwaite)

- Get numerator and denominator in appropriate form (using expected mean squares).
 - Try to avoid subtraction.
- Get approximate degrees of freedom with Satterthwaites formula.
- Use approximation with χ^2 to get CI on variance components.

Restricted vs. unrestricted

- Affects inference on random effects (variance component vs. equal means)
- Restricted applied to “limited resource” contexts.

Issues with Mixed Models

- Sometimes it’s easier to deny the possibility of random effects.
- Likelihood are sometimes easy to work with... sometimes not.
 - Depends on random/fixed pattern, constraints

Nesting

- The absence of crossing
 - Order of levels doesn’t matter.
- Looks like lots of interaction (with no pattern)
- Already seen with replication (omitted in `model` statement)
- *Subsampling*
 - Can’t measure entire EU (but still interested in inference on EU).
 - Several stages of sampling – EU’s and subsamples within EU’s.
 - Use summary to get EU observation value.
 - Nesting models allow SAS analysis

- Different variance components (similar testing) and SE's
 - Interested in inference on EU's.
- Design/cost question about number of levels, subsamples (more efficient sampling)
- *Mechanics*: DF and SS – $B(A) = B + AB$
- Use rules and formulas to get EMS and variance component estimates.
- Use Haas diagram to get tests
 - Used to answer where to sample for more power.
- Can get results from factorial with interaction.
- Looks like lots of interaction with little marginal effect.
 - Depends on interpretation and expectation
 - Is direction the same?
 - Are scales comparable?
- Can *stagger* nesting for design with several nesting levels to improve power (error df).
- Can have nesting and crossing in same design.

Split Plot Design

- (At least) three factors: two treatments, blocking
- Treatments are crossed
- Block restricts randomization of subplot factor (baking soda, variety)
 - Can every combination of factors appear “next to” every other possible combination?
- Each blocking level is a *whole plot* (baking period, farm), average of which is EU value for whole plot factor (oven temperature, fertilizer).
- Two situations:
 - Each whole plot contains just one level of whole plot factor. (whole plot replication)

```

model resp = fert farm(fert) var fert*var;
test h = fert e = farm(fert);

```

- Each whole plot contains combination of treatments. (Replication is whole plot/whole plot factor interaction.)

$$y_{i,j,k} = \mu + \alpha_i + \eta_k + (\alpha\eta)_{i,k} + \beta_j + (\alpha\beta)_{i,j} + ((\beta\eta)_{j,k} + \epsilon_{k(i,j)})$$

	<i>a</i>	<i>n</i>	<i>b</i>	1		
Source of Variation	F <i>i</i>	R <i>k</i>	F <i>j</i>	R <i>l</i>	Degrees of Freedom	Expected Mean Square
Block (<i>R</i>)	<i>a</i>	1	<i>b</i>	1	<i>n</i> – 1	$ab\sigma_R^2 + \sigma^2$
WPF (<i>A</i>)	0	<i>n</i>	<i>b</i>	1	<i>a</i> – 1	$nb\phi_A + b\sigma_{RA}^2 + \sigma^2$
Bl × WP	0	1	<i>b</i>	1	$(a - 1)(n - 1)$	$b\sigma_{RA}^2 + \sigma^2$
Sub. Fac. (<i>B</i>)	<i>a</i>	<i>n</i>	0	1	$(b - 1)$	$an\phi_B + \sigma^2$
WP × SP	0	<i>n</i>	0	1	$(a - 1)(b - 1)$	$n\phi_{AB}^2 + \sigma^2$
Error	1	1	1	1	$a(b - 1)(n - 1)$	σ^2

Error (subplot interaction with block *RB*) may not be pooled if block random.

- Factorial, covariate, or incomplete block possible for whole factor.
- Any other factorial or blocking or covariate structure possible for subplot factor.
- One replicate (of whole plot factor) case: may look like tested over subplot error since no DF for wholeplot error.
- Pooling whole plot error into subplot error can ruin validity of testing both factors.

Extensions

- *Split-split plot*: subplots act as blocks for subsubplots
- *Strip plots*: two subplot factors and blocking factor has row and column effects.
 - `effect(row column)` gives replication error adjusted for row and column effect (same effect as *block* × *WP* interaction)
 - same as `row*effect` or `column*effect` or `row*column*effect` after `row column effect` in model statement.

Computing Standard Errors

- *First step*: Write out model for each observation.
 - What are the factors?
 - What are their distributions?
- *Second step*: Write out test statistic in terms of parameters.

- What is being summed over?
- What is denominator under the sum?
- Heuristic: Match number of dots to number of letters in denominator.
- Do terms cancel out in subtraction?
- *Third step:* Compute the variance.
 - Fixed effects do not contribute to variance.
 - Variance of a mean is variance of components divided by denominator of sum.
- *Fourth step:* Find (a combination of) *MS* with expected value that matches variance.
 - Use Satterthwaite to calculated degrees of freedom.

Interpreting Models

- List experimental factors and cross/nesting.
- What is a block/replicate? Is there a Latin Square design?
- Is there a split-plot structure?

Repeated Measures Analysis

- Have already looked at cross-over designs (different treatment level with each time).
- Might use single summary over time.
- What about development as time goes on?
- *Split plot – subject* is whole plot; averages over time.
- Is assumption of *compound symmetry* appropriate?
- Can use split-plot (with adjustments) under *sphericity assumption*.
- Use `repeated` to do several analyses: MANOVA, split plot, and orthogonal polynomial
- Use `proc mixed` to model assess more general correlation structures.