Statistics 514: Design of Experiments

Topic 8 Summary

Multi-factor Analysis

- Will spend most of the rest of the semester on multiple factor designs
- Naive one-factor-at-a-time approach inefficient and hard to interpret.

Randomized Complete Block Design

- nuisance factor known ahead of experiment (allocated, not assigned)
- one treatment level with no replicates per block (which are homogeneous)
- think of as restriction on randomization/generalization of paired t-test
- reduced error variance (increases power)
- additive model (no interaction) with SS analysis and block source
- additional restrictions on block effects
- Get "orthogonal" design, which refers to properties of design matrix
- treatment SS unaffected by block effect
- no inference on block effects

Diagnostics

- Usual checking residuals
- Check additivity assumption

Multiple Comparisons

- Blocks represent replicates.
- Error degrees of freedom change.

Randomization

- Permute within blocks
- Leads to F-test in limit

Missing Values

- Can use Type III SS (nonorthogonal design)
- Can impute missing values (change df)

Randomized Complete Block Design

- *Additivity assumption* – check with plot and Tukey's test
  - Doesn't catch all types of interactions
  - Can transform away interaction ("transformable nonadditivity")
- *Random blocks* – no change in inference; use `proc mixed` if testing individual cell means
- *Replicates within blocks* – simplification of analysis
  - no interpretation of interaction
  - Use if not much variety in EU's among blocks

Relative Efficiency and Blocking

- How well did blocking work (vs. just randomizing)?
- Quantifies tradeoff between summed squared error and degrees of freedom
- Get comparison of sample sizes (and costs) needed to achieve same results

Factorial Design

- Can test factors with *contrasts*
- *Interaction*
  - Explains non-additive effects
  - Affects estimation of main effects (damping, for instance)
  - Makes estimation (based on averages) conditional on other variables
  - Test first
  - Check for parallelness (and patterns/local interaction)
  - Can fill out response curve (not flat)
  - SAS uses regression set-up for calculation

Factorial Designs

- Factorial designs are efficient (same precision with fewer design points) even if no interaction is assumed.
  - Exploits constant variance assumption
- *Diagnostics:* sometimes interested in more specific questions
- For multiple comparisons, use `lsmeans` slice option and adjust by hand, if necessary.

Sample Size

- Want sample size to test all effects
  - Usually interaction is more important
  - Use model error df's
  - Other factor levels act as replicates

General Factorial Models

- Can include many factors in factorial design
- Check higher-level interaction first
- Higher-level interactions and random factors complicate.
- Working with quantitative factors means interpolation is possible.
  - Use regression output (done in Linear Models material) to get response surfaces for different factors.
  - May need to check significance of regression coefficients
- Blocking factors – can include factorial design in each block
  - More additivity assumptions

Unbalanced Designs

- Lose orthogonality
- Use higher type of SS
  - Inference based on more terms
  - Often (not always) less significant

Latin Square Design

- Two blocking factors (# block levels = # treatment levels)
- Generate with `proc plan`
- Look at treatment F_0.

Missing Data

- Can look at Type III SS
- Substitution makes interpretation easier

Random blocks

- Inference shouldn't change.
- Use `proc mixed` to get cell mean SE's correct

Correlated Observations

- Can use `proc mixed` with `repeated` option to test different correlation structures
- Use `Fit Statistics` to decide on fit of model
- Similar to what we'll see in Repeated Measures Designs

Replicated Latin Squares

- Usual Latin Squares efficient but not necessarily precise
- Replicate square (to increase precision)
  - Include replicate effect in model
- Replicate square and shuffle rows
  - Row effect nested within square.
  - Uses more blocking factor levels
- Replicate square and shuffle rows and columns
  - Two nested effects
  - Efficient use of many row/column levels
- Latin Rectangle
  - No replication

- Not every treatment/column combination appears in each row.
- Useful for crossover designs

Crossover Designs

- Block on period and subject
- Want no longterm effects
  - Can use washout period, replicated periods
- Can test for residual period (difference of sums)
- If no residual effect, can estimate treatment effect by difference of differences
  - Period block effects cancels
- Can model and estimate residual effect (for example, additive over one period)
- Balance order of pairs to improve precision

Graeco-Latin Squares/Latin Hypersquares

- Several block effects in p^2 EU's
- Superimposing several alphabets
- Orthogonality means each pair of letters appears once.

Relative Efficiency

- Check against RCB, CRD
- Diminishing returns as block factors added

Balanced Incomplete Block Designs

- What if number of treatments per block is less than total number of treatments?
- Balance pairs of treatment occurences
- a treatments, b blocks, k treatments per blocks
- r replicates of each treatment ($kb = ar$)
- λ replicates of each pair ($\frac{r(k-1)}{a-1}$)

Analysis

- Similar to unbalanced RCBD (no interaction)
- Use Q_i's ({sum of obs with trt i} – {sum of block means with trt i}) to estimate τ_i's (SE's given as multiple of σ^2)
- *Power* – choose λ to get b to get power
  - Since estimated treatment effects (based on adjusted values Q_i) are correlated, effective number of replicates is not r, but $\lambda a/k$.
- *Multiple Comparisons* – Use Q_i's in estimating contrasts (and `lsmeans` to get right SE)

Interblock/Intrablock Analysis

- $\hat{\tau}_i$ based on difference of block totals from treatment totals
- $\tilde{\tau}_i$ based on difference of grand mean from block totals (generated by contrasts with `model` on `block` only)
- Can combine (with degrees of freedom correction) to improve estimate
- Compare page 74 to page 76 (Topic 8) SAS output
  - Treatment effect: $p = 0.0107$ vs. $p = 0.0112$
  - Group 3 vs. 4 contrast: -3.0 vs. -2.97 ($p = 0.0077$ vs. $p = 0.0080$)
  - Group 1 vs. 2 contrast: $p = 0.739$ vs. $p = 0.7829$
  - $\hat{\sigma}^2_\beta = 8.02$, $\hat{\sigma}^2 = 0.65$

Extensions

- *Youden Squares* – BIBD for 2 block effects
- *PBIBD* – pick λ_i for each class i
- *Cyclic Designs* – easily generated
- *Lattice Designs* – For large number of treatments, keeps orthogonality of blocks