STAT 368 - Final review

The final exam is like the sample exam, in terms of what is emphasized (in a general sense) and what is not. **Emphasis on**:

- basic concepts of experimental design
  - blocking, randomization, replication etc. - what are they and why are they employed?
  - What are the standard kinds of designs?
  - HOW are blocking, randomization, replication implemented in these designs?

- basic concepts in analysis of experimental data
  - Principles of analysis - least squares estimation, comparing full and reduced models through their SS_E in order to test hypotheses

  - Using appropriate averages of data values to estimate effects (often a simple hand calculation)

  - Effects models which allow for the estimation of all factors which might help explain the variation (related to the design used)

  - Making inferences about the (significantly different) levels of fixed factors - construction of simultaneous confidence intervals, interpretation of graphical output

  - Random effects models - how is the analysis different?

- Interpreting R output

  - ANOVA, normal plots, residual plots, etc.

  - drawing appropriate conclusions (in terms of the experimental situation, rather than merely something like 'reject H_0').

Emphasis is NOT on:

- Lengthy hand calculations
- Plugging numbers into formulas
- Memorization

Basic designs which have been covered:

- Completely randomized. One factor at a levels, n observations made at each level. All an runs in completely random order (or anything else that can be randomized). Effects model has terms τ_i for treatment effects only (as well as the overall mean μ and random error ε_{ij}, of course).

- Randomized complete block. One factor at a levels, but we bock the runs in order to control for a nuisance factor, which is at b levels. Randomize within blocks only (so each block looks like 1 replicate of a CRD). Effects model has terms τ_i for treatment effects and β_j for block effects.

- Latin squares. One factor at p levels. A Latin square is a novel technique in which we block in two ways, to control two nuisance factors each at p levels, and we do all this in just p^2 runs rather

than p^3. But only a small number of d.f. are available for estimating σ^2, and all of these are used up if we try to estimate interactions. So this design is really only useful if we are sure that there are no interactions of significance.

- Graeco-Latin squares. An extreme case of Latin squares. *Three* nuisance factors at p levels; still only p^2 runs. But even fewer d.f. for error.

- Balanced Incomplete blocks. Incompleteness necessary if blocks are too small to allow us to run all treatment levels in a block. 'Balance' means that, despite the incompleteness, we achieve ... .

Factorial designs

- Two factors at a and b levels; all ab runs in random order. The full effects model with all main effects and interactions uses all available $(ab-1)$ d.f., leaving nothing for error. So estimate less (drop terms - possibly using half-normal plots to decide which) or observe more (replicate - then replicates are treated as blocks).

- Basic principles extend to 3 or more factors.

- 2^k factorials - k factors, each at two levels (high/low, ± 1)

  - Notation $(1), a, bc$, etc., tables of $\pm$ signs (both in *standard order*)

  - Hand calculations of effects from the table of $\pm$ signs and the data

- Blocking factorials

  - Complete blocks: Here we are replicating an entire factorial. We view the b replicates as blocks, using $b-1$ d.f. to estimate the block effects (how?).

  - Incomplete blocks: Here only a fraction of the 2^k runs are made in a block.

    * For *two* blocks, 1 d.f. will be used to estimate the block effects, so something has to be sacrificed. We can decide what - we pick one effect to *confound* with blocks, put all runs where that effect is '$+$' in one block, all where it is '-' in the other. Alternatively we can compute the defining contrast, and do arithmetic mod 2 to get the blocks.

* For *four* blocks we pick two effects to confound, and form blocks corresponding to the 4 combinations $(+,+)$, $(+,-)$, $(-,+)$, $(-,-)$ of signs for these two effects. Alternatively we can compute the two defining contrasts, and form blocks corresponding to the 4 combinations $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$ of values (mod 2) of these contrasts.

* Blocking doesn't address the problem of not having enough observations, in one replicate, to estimate all of the effects (in fact it makes it worse). We can replicate, thus getting an estimate of error. We don't have to confound the same effect(s) in each replicate - we can *partially confound* effects (how?).

- Fractional factorials

  - We can run only one half, or one quarter, ... of the combinations in a 2^k factorial. Then the table of $\pm$ signs has only one half, or one quarter, ... as many rows. Use the available rows in the usual way to get estimates of effects (you should review the notation $[A]$, $[ABC]$, etc.). For a half fraction, any such estimate will be (possibly) biased, with an expectation equal to that of the intended effect, plus others with which it is *aliased*.

  - We can choose the effects to be aliased; from these we get the appropriate design. The *principal fraction* has defining relation $I = AB \cdots K$, the *complementary fraction* has instead $I = -AB \cdots K$. Doing algebra (mod 2) on these relationships yields all the other aliases (how?) in this one-half fraction.

  - These are not the only halves which might be run, but they give the largest resolution. (What is that? Why does one want a large resolution?)

Nested designs

- Designs in which the levels of one factor make sense only within those of another have the factors *nested*. Then interaction between the factors makes no sense (why not?) and the effects model (what is it?) contains only main effects.

- Nesting and factorials are not mutually exclusive - we looked at a 3 factor example with one factor nested in another, but crossed with a third. (Then what effects can reasonably be estimated? Write down the effects model.)

Split plot designs

- Special features - restrictions on the randomization, two levels of replication. Effects model includes replicates and replicate$\times$whole plot treatment interaction (this latter SS is used to test

for the significance of whole plot treatments) and the usual 2-way model with effects of whole plot treatments, subplot treatments, and their interaction.

Repeated measures designs

- The discussion of repeated measures designs was intended to give some appreciation for the breadth of the field of experimental designs. However, it will not be on the final exam.

A good way to review these designs and models: Go through the exercises in the book. Read the descriptions of the experiments and see if you can identify the type of design being used, and an appropriate model of the data (without looking at the chapter heading!).

Special (post-midterm) features of analysis

- Random effects and mixed effects models. Here some or all of the effects are random. The major changes in the analysis are:

  - The F-ratios no longer all have MS_E in the denominator. The correct choice will depend on the expected value of the numerator mean square, when the null hypothesis is true - the denominator must be a mean square with this same expectation. (Why?)

  - We can still do multiple comparisons among the levels of the fixed factors (Tukey intervals, etc.) but now the standard error of a difference in means employs the mean square, and d.f., from the previous point.

  - It doesn't make sense to compare levels of random factors (why not?), but we can estimate the variation between them. Again the expected mean squares are used to determine

appropriate estimates of these *variance components*.

- Regression. Possible uses in our analyses:
  - ANOVA can always be thought of as a special case of regression, although the formulas are much more transparent in ANOVA (e.g. using obvious averages to estimate effects) as long as the design is balanced.
  - When the levels are quantitative (e.g. length of time for which a treatment is applied) but not exact (e.g. 57, 61, 63 minutes when 1 hour is intended) then one might regress on the numerical values rather than the level labels.
  - ANCOVA: we discussed (only) the case of a CRD in which a significant amount of the variation might be a result of changes in a covariate (e.g. responses to a drug might depend

on the age of the patient as well as on the type of drug). One of the basic techniques is that of estimating and comparing the treatment means by setting the covariate equal to its overall average (so $X = x - \bar{x}_{\cdot\cdot} = 0$); this yields the adjusted treatment means.

* Here classical design theory and regression are combined to make for a much simpler analysis (numerically) than otherwise.