UNBALANCED DESIGNS

Recall that an experimental design is called *unbalanced* if the sample sizes for the treatment combinations are not all equal.

Reasons why balanced designs are better:

- The test statistic is less sensitive to small departures from the equal variance assumption.
- The power of the test is largest when sample sizes are equal.

Reasons why you need to be able to work with unbalanced designs:

- Balanced designs produce unbalanced data when something goes wrong. (e.g., plants die, machinery breaks down, shipments of raw materials don't come in on time, subjects get sick, etc.)
- Some treatments may be more expensive or more difficult to run than others.
- Some treatment combinations may be of particular interest, so the experimenter chooses to sample more heavily from them.

Ways of analyzing unbalanced designs:

- If the data are "proportional" (meaning that $r_{ij} = (r_{i\cdot}r_{\cdot j})/r_{\cdot\cdot}$), there is a minor variation to the usual analysis that works. (See Montgomery, p. 601 for details.)
- If the data are only slightly unbalanced, there are several approximate procedures that might be used (e.g., estimating missing observations, omitting observations from cells with larger numbers, various methods adjusting weights). (See Montgomery, pp. 601-603 for details.)
- The "Exact Method, " representing the analysis of variance model as a regression model. This is the only method we will discuss for unbalanced factorial designs. It requires some caveats:
  - The same problem might be done in more than one way, resulting in different sums of squares.
  - The hypotheses tested might be different from those tested in balanced ANOVA.
  - The tests sometimes create their own problems in interpretation.