Higher-order Factorial Experiments

Suppose we have

- three factors A, B, and C
- with levels l_1, l_2 and l_3.

A full factorial experiment has at least one run for each combination of the factors, i.e., we need to have at least $n = l_1 \times l_2 \times l_3$ observations. The analysis is very similar to the one for two-way ANOVA models.

Factional Factorials

- Disadvantage of factorial designs: the total number of runs can be very large when the number of factors is large.
- Factional factorials use only a fraction of the number of runs.
  - Save the cost of the full experiment;
  - It is still possible to estimate the lower order effects with just a fraction; we have to assume the high-order interactions are negligible.
  - Consider an example with 3 factors (each with 2 levels).

Blocking Designs

First, let's refresh our memory on two-sample t-tests.

- The `shoe` data.
- A simple two-sample t-test: a completely randomized design (CRD)
- A paired two-sample t-test: a randomized block design (RCBD)
- Why blocking? Reduce variance; of course, meanwhile we lose df.

Randomized Block Design

- Two factors: one is of primary interest, and the other is the blocking variable.
- Block size is the same as the number of levels of the primary factor: we assign the primary factor randomly within each block.
- The analysis is similar to the two-way ANOVA with one observation per cell.

- What does it mean – “Blocking is a feature of the experimental units and restricts the randomized assignment of the treatments. This means that we cannot regain the df devoted to blocking even if the blocking effect turns out not to be significant” ?

- Relative advantage of RCBD over CRD: $\hat{\sigma}^2_{\mathsf{CRD}}/\hat{\sigma}^2_{\mathsf{RCBD}}$.

Latin Squares

- Three factors: one is of primary interest (with m levels), and the other two are blocking variables.
- Latin square design: only m^2 runs are needed; cannot estimate all pairwise interactions, but can estimate the main effects (i.e., the additive model).
- When there are three block variables, a Graeco-Latin square may be used.

- The analysis is the same as an additive model with three factors: the RSS has df $(m-1)(m-2)$ where m is the number of levels for the primary factor. For example, the Tukey pairwise CIs for the primary factor are

$$\bar{y}_{i\cdot\cdot} - \bar{y}_{j\cdot\cdot} \pm q^{\alpha}_{m,(m-1)(m-2)}\hat{\sigma}\sqrt{\frac{1}{m}}.$$