2^{nd} practice sheet Experimental Design

Werner G. Müller

submit and set checkmarks on Moodle until April 9^{th} 2024

5. Consider a CRD with five treatment groups, in which a total of $n = 50$ units are to be used. Although it won't be explicitly used in the analysis model, treatments 1 through 5 actually represent increasing concentrations of one component in an otherwise standard chemical compound, and the primary purpose of the experiment is to understand whether certain measurable properties of the compound change with this concentration. The investigator decides to address these questions by estimating four quantities:

$$\tau_2 - \tau_1, \tau_3 - \tau_2, \tau_4 - \tau_3, \tau_5 - \tau_4$$

   where each τ_i is a parameter in the standard effects model. Find the optimal allocation for the 50 available units (i.e., values for $n_1, \ldots, n_5$) that minimizes the average variance of estimates of the four contrasts of interest. Do this as a constrained, continuous optimization problem, then round the solution to integer values that are consistent with the required constraint.

6. Continue working with the experimental design described in exercise 5. Suppose the experiment-specific treatment means in this problem, as would be expressed in the cell means model, are actually:

μ_1	μ_2	μ_3	μ_4	μ_5
10	11	12	12	12

   and $\sigma = 2$. What is the power of the standard F-test for the hypothesis $\tau_1 = \tau_2 = \tau_3 = \tau_4 = \tau_5$, at $\alpha = 0.05$:

   (a) if all $n_i = 10$?

   (b) under the optimal sample allocation you found in exercise 5?

   (c) Derive an optimal allocation for the F-test of equal treatment effects, i.e. the sample sizes (totaling 50) that would result in the greatest power, if in reality the experiment-specific means are $\mu_1 = 10$ and $\mu_2 = \mu_3 = \mu_4 = \mu_5 = 8$.

7. Consider a modification of the usual randomized complete block design. Suppose there are b blocks, and that each block contains units allocated to each of t treatments, as in the CBD. However, while each treatment is assigned a unit in each block (as with a CBD), $r > 0$ additional units are allocated to treatment 1 in each block (this is an augmented CBD). I.e., each block contains $t+r$ units, $r+1$ of these are allocated to treatment 1, and one of these is allocated to each of treatments 2 through t. There are thus $n = b(t+r)$ units used in the entire experiment. Answer the following questions, using the usual notation for an effects model parameterization and assuming that blocks and treatment effects do not interact.

(a) Write expressions for the partitioned model matrices, $\mathbf{X}_1$ and $\mathbf{X}_2$. (Use general symbolic expressions like $\mathbf{1}_t$ and $\mathbf{J}_{n\times b}$ to do this; segments of the matrices that are filled with zeros can be left blank.)

(b) Write expressions for the matrices $\mathbf{H}_1$, $\mathbf{X}_{2|1}$, and $\mathcal{I}$.

(c) Give a scalar expression (i.e., not in terms of matrices) for the variance of $\widehat{\tau_1 - \tau_2}$

(d) Similarly, give a scalar expression for the variance of $\widehat{\tau_2 - \tau_3}$

(e) Does this design and a CRD with $n_1 = b(r+1)$ and $n_j = b$, $j = 2, \ldots t$ satisfy Condition E?

(f) For this design, σ^2 can be unbiasedly estimated even if treatments and blocks interact. Write this estimate, a quadratic form in the data for which the rank of the central matrix is rb, and show that it is statistically independent of the least-squares estimate of any estimable $\mathbf{c}^T\boldsymbol{\tau}$.

8. An investigator wishes to design an experiment to compare three treatments using a CBD (of three units per block). From long experience with similar experiments, he knows that σ^2 will be very close to 2.4. He believes that there really is no difference between treatments 1 and 2, but that treatment 3 produces responses about 0.6 larger, on average, than these. Assuming this is true:

   (a) How many blocks should be included in the design to provide power of 0.8 for testing $H_0: \tau_1 = \tau_2 = \tau_3$ with type I error probability of 0.05?

   (b) If $b = 10$ blocks are used, what will be the expected squared length of a 95% confidence interval for $\frac{1}{2}(\tau_1 + \tau_2) - \tau_3$?