2^{nd} practice sheet Experimental Design

Werner Müller

submit and set checkmarks on Moodle until March 18th

2. Consider the lady tasting tea problem: The experiment provides a subject with 8 randomly ordered cups of tea - 4 prepared by first adding milk, 4 prepared by first adding the tea. The subject knows that 4 of the 8 cups are prepared by one method and the other 4 by the other method, and has to select 4 cups prepared by one method.

   (a) Carefully identify the units in this experiment.

   (b) Carefully define the treatments in this experiment.

   (c) Fisher mentioned several physical devices that might be used to determine a random temporal order of treatments to the available units. Carefully describe exactly how this might be done using any of these devices, while honoring the sample size constraints of the problem.

   (d) Suppose eight (physical cups) are available for the execution of this experiment, but they are from two sets. Four are made from heavy, thick porcelain, while the other four are made from much lighter china. If operational restrictions are such that each cup can only be used once, how might this fact be incorporated into the experimental design?

3. Consider the lady tasting tea problem:

   (a) what is a reasonable null hypothesis?

   (b) how many cups does the lady have to identify if we want to reject the null hypothesis at a significance level of 5%?

   (c) what is the exact p-value if we can reject H_0 at a level of 5%?

4. Recall R.A. Fisher's tea-tasting experiment (exercise 2(d)). Suppose that, rather than stating her guess as to which ingredient was put into the cup first, the lady responded to each cup with a real-valued number (a response) reflecting her judgment of the tea's taste, and that we are willing to adopt the model:

$$y_{ijk} = \alpha_i + \beta_j + \varepsilon_{ijk}$$

   where $i = 1$ for porcelain cups, $i = 2$ for china cups, $j = 1$ for milk-before-tea preparations, $j = 2$ for tea-before-milk preparations, and $k \in \{1, 2\}$ indexes the first or second beverage made for the indicated values of i and j. Assuming that inferences about β_1 and β_2 are of greatest interest:

   (a) construct a partitioned model for the experiment,

(b) compute $\mathbf{H}_1$, the hat matrix for the model including only cup-type parameters and random noise, and

(c) compute $\mathbf{X}_{2|1}$ and $\mathcal{I}_{2|1}$.

(d) the expected squared length of a 90% confidence interval for $\beta_1 - \beta_2$ (in terms of σ^2).

5. Continuing exercise 3, suppose now that one of the china cups was broken, and one of the tea-before-milk preparations was not made. For the resulting seven-trial experiment, find:

   (a) $\mathbf{H}_1$

   (b) $\mathbf{X}_{2|1}$ and $\mathcal{I}_{2|1}$

   (c) the expected squared length of a 90% confidence interval for $\beta_1 - \beta_2$ (in terms of σ^2).

6. Once again, continuing exercise 3, suppose only three beverages are prepared:

   - china cup, with milk-before-tea
   - porcelain cup, with milk-before-tea
   - porcelain cup, with tea-before-milk

   Show that the data value collected from the beverage prepared in the china cup is not used in estimating $\beta_1 - \beta_2$. Explain clearly, and without using mathematics, why this is reasonable.