STAT 514 Homework#2
(Due Thursday 9/10/2015 BEFORE CLASS)

1. Two types of fertilizer mixture A and B are used to feed six bean plants with each randomly applied to three plants. The yields of the plants (in pounds) are given below. We want to test whether A and B have different effects on the yield of bean plants.

A	B	B	A	B	A
3	5	4	3	7	6

(a) Construct the randomization distribution;

(b) Calculate the p-value based on the randomization test;

(c) Calculate the p-value based on two-sample t-test;

(d) Compare the results from (b) and (c).

2. You are requested to design an experiment to compare the typing efficiency of three different types of keyboards denoted by A, B and C. Two typists, denoted by T_1 and T_2, are employed and six standard manuscripts m_1, m_2, m_3, m_4, m_5 and m_6 are used.

(a) Give your experimental plan.

(b) Which principles you have used in designing the experiment? Comment on their advantages in this particular experiment.

3. We introduce a simple statistical model for the keyboard experiment discussed in class. Let y be the amount of time used to type up a manuscript. Note that y depends on keyboard, manuscript, whether the manuscript has already been typed, and experimental error. Let τ_A and τ_B denote the effects of keyboard A and B respectively, β_i the effect of manuscript i for $i = 1, 2, 3, 4, 5, 6$, and ϵ the experimental error. Let η_l

1

denote the learning effect. We are interested in estimating the difference between τ_B and τ_A. Suppose Design 2 from the lecture notes is used in the experiment, which is

$$1.A - B;\, 2.B - A;\, 3.A - B;\, 4.B - A;\, 5.A - B;\, 6.A - B$$

The statistical model for the amount of time used in $1.A$, denoted by y_{1A}, is

$$y_{1A} = \mu + \tau_A + \beta_1 + \epsilon_{1A}, \tag{1}$$

and the model for the amount of time used in $1.B$ is

$$y_{1B} = \mu + \tau_B + \beta_1 + \eta_l + \epsilon_{1B}, \tag{2}$$

where μ is some constant.

(a) Is η_l positive or negative? Why it is included in (2) but not in (1)?

(b) Write down the statistical models for the other runs.

(c) How to estimate $\tau_B - \tau_A$, using $y_{1A}, y_{1B}, \ldots, y_{6A}, y_{6B}$.

(d) Compared to Design 1 in the lecture, what is the gain (in terms of estimating $\tau_B - \tau_A$) from using randomization? Recall that the typing order of the manuscripts is randomized in Design 2.

(e) Derive an experimental plan using balanced randomization as discussed in class. Is there any further gain of your plan over Design 2?