HOMEWORK#2

Due on Midnight Dec 31

Please email your homework (scanned handwritten solution or typed solution) to my email address with subject "HW 2 of NCKU course"

1. An experiment tries to compare the typing efficiency of two keyboards denoted by A and B. One typist uses the keyboards on six different manuscripts, denoted by 1-6.

   Let y be the amout 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 μ_A and μ_B denote the effects of keyboard A and B respectively, τ_i the effect of manuscript i for $i = 1, 2, 3, 4, 5, 6$ and ϵ denotes the experimental error. Let α_l denote the learning effect. We are interested in estimating the difference between μ_B and μ_A. Consider a design as follows:

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

   (For the 1st manuscript, the typist tries keyboard A first, then tries keyboard B.) The statistical model for the amount of time for 1st manuscript with keyboard A, denotes by y_{1A} is

   $$y_{1A} = \gamma + \mu_A + \tau_1 + \epsilon_{1A},$$

   and the model for the amount of time used for 1st manuscript with keyboard B is

   $$y_{1B} = \gamma + \mu_B + \tau_1 + \alpha_l + \epsilon_{1B},$$

   where γ is some constant value.

   - Is α_l positive or negative? Why it is not included in first model?
   - Write down the statistical models for the other runs.
   - Is $(\sum_{i=1}^{6} y_{iA} - \sum_{i=1}^{6} y_{iB})/6$ a good estimate for $\mu_A - \mu_B$
   - Please propose a better design.

2. A 2^{5-2} design is defined by $\mathbf{D} = \mathbf{AC},\ \mathbf{E} = \mathbf{BC}$.

   - Find its defining words and resolution.
   - In the course of the analysis of this experiment, it is thought that factor **E** and all interactions involving **E** are negligible. In addition to estimating the four main effects, there are still three degrees of freedom left. What two-factor interactions can be estimated with these three degrees of freedom? Here an effect is considered to be estimatible, if it does aliased with any same-order or lower-order interaction effects.

3. Consider a linear regression mode with one x variable and no intercept, i.e., $Y = \beta X + \epsilon$, where the noise satisfying $var(\epsilon) = \sigma^2$. Given observations (x_1, y_1), ..., (x_n, y_n)

(a) What is the least square estimation of β? That is to solve the minimization of $\sum(y_i - x_i\beta)^2$ w.r.t. β.

(b) What is mean and variance of your estimation?

(c) Given a new $x_{n+1} = 10$. What is your prediction for y_{n+1}? what is the mean and variance of the prediction?