HOMEWORK#2

Due on 12PM (noon) May 29

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

1. A clay tile company is interested in studying the effect of cooling temperature on strength. The company has five ovens which produce the tiles, four tiles were baked in each oven and then randomly assigned to one of the four cooling temperatures. The data are shown below.

Cooling	Oven					
Temp	1	2	3	4	5	mean
$5°$	3	10	7	4	3	5.4
$10°$	3	8	12	2	4	5.8
$15°$	9	13	15	3	10	10
$20°$	7	12	9	8	13	9.8
Mean	5.50	10.75	10.75	4.25	7.50	7.75

   (a) Which type of design was employed? Describe how the fundamental principles of experimental design were followed in this design.

   (b) If $MS_E = 6.275$, compute the F-statistic to determine if there is a difference among the four cooling temperatures (use α= 5%).

   (c) Estimate the relative effiency, and interpret your result.

   (d) Suppose the company believes there is a jump in the strength at 12.5° but otherwise cooling temperature has no effect, that is, 5° and 10° are not different, neither are 15° and 20° , but these two groups of temperatures have different effects. Find a set of orthogonal contrasts that would allow you to test this.

   (e) Test these contrasts.

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 regression model $y_i \sim N(\beta_2 + \beta_1 x_i, \sigma^2)$ for $i = 1, ..., n_1$ and $y_i \sim N(\beta_1 + \beta_2 x_i, \sigma^2)$ for $i = n_1 + 1, ..., n_1 + n_2$, and all y_i's are independent. Furthermore, we assume that $\sum_{i=1}^{n_1+n_2} x_i = 0$.

- Express the model in a matrix form, i.e., rewritten the model as $Y = X\beta + \epsilon$, where $Y = (y_1, \dots, y_{n_1+n_2})'$, and β is the parameter vector.
- Find out the OLS estimation (b_1, b_2) for β_1 and β_2. (Please provide an answer which no longer involves matrix operation, otherwise only partial credits will be granted.)
- What is the variance of $b_1 + b_2$?

4. You are interested in determining the linear relationship between the flower opening size (nearest tenth of a mm) and the incidence of Fusarium head blight (% of infected spikes). You hire an undergraduate to collect the data and she comes back from the field with the following table based on a collection of 29 wheat plants.

Opening (mm)	1.2	1.4	1.7	1.8	1.9	2.0	2.2	2.4
Mean incidence (%)	5.3	8.6	10.2	12.5	14.4	17.1	18.2	22.3
n	3	5	2	5	5	6	5	3

- The individual measurements used to generate this table were thrown away. Can you still do the inference here? If so, what would you do? Will you get the same slope, intercept, and variance estimates? Explain.