實驗設計與分析 HW1 solution

Problem 1.

Do Problem 6 in Chapter 1 of textbook.

(a) For the typing experiment considered in Section 1.3, use a statistical model to quantify the gains from using randomization (as illustrated in the second sequence) and from using balance in addition to randomization.

(b) Suppose that the following sequence is obtained from using balanced randomization:

1.A,B, 2.A,B, 3.A,B, 4.B,A, 5.B,A, 6.B,A.

Would you use it for the study? If not, what would you do? What aspect of the sequence makes you uneasy? Can you relate it to the possibility that the advantage of the learning effect may diminish over time and express it in more rigorous terms? (Hint: The terms in the model should represent the effects you identified as potentially influencing the comparison.)

(a) For the typing experiment considered in Section 1.3, use a statistical model to quantify the gains from using randomization (as illustrated in the second sequence) and from using balance in addition to randomization.

根據題意我們可以定義以下模型：

$$y_{ij} = \mu_y + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ij}$$

其中

- μ_y: 整體打字效率的平均值
- α_i: 為 treatment effect(不同鍵盤帶來的效應),i=1,2
- β_j 為 blocking effect(不同份手稿帶來的效應),j=1,2,...,6
- γ_{ij}: 為 learning effect(此效應是假設打第二次稿的熟練度會比第一次高，且第二次打稿提升的效率皆相同),
  若 order 為 A→B, 則 $\gamma_{1j} = 0, \gamma_{2j} = c$
  若 order 為 B→A, 則 $\gamma_{1j} = c, \gamma_{2j} = 0$
  其中 c>0
- ϵ_{ij}: 誤差項

在這個實驗中我們有興趣的是 $\alpha_1 - \alpha_2$(A&B 兩種鍵盤打字效率的差異)
為了得到 $\alpha_1 - \alpha_2$，我們需要用 $\bar{y}_1 - \bar{y}_2$ 來估計
其中 $\bar{y}_i = \frac{1}{6}\sum_{j=1}^{6} y_{ij}$

$$E(\bar{y}_1 - \bar{y}_2) = \alpha_1 - \alpha_2 + \frac{1}{6}\sum_{j=1}^{6}(\gamma_{1j} - \gamma_{2j})$$

從這個式子中我們可以看出只有當 $\frac{1}{6}\sum_{j=1}^{6}(\gamma_{1j}-\gamma_{2j})=0$ 時
我們才得以利用 $\bar{y}_1-\bar{y}_2$ 得到 $\alpha_1-\alpha_2$ 的不偏估計量

針對這個問題我們可以使用三種實驗設計的方法

case 1. Block design

Keyboard	Manu	Order	Model
A	1	I	$y_{11}=\mu_y+\alpha_1+\beta_1$
B	1	II	$y_{21}=\mu_y+\alpha_2+\beta_1+c$
A	2	I	$y_{12}=\mu_y+\alpha_1+\beta_2$
B	2	II	$y_{22}=\mu_y+\alpha_2+\beta_2+c$
A	3	I	$y_{13}=\mu_y+\alpha_1+\beta_3$
B	3	II	$y_{23}=\mu_y+\alpha_2+\beta_3+c$
A	4	I	$y_{14}=\mu_y+\alpha_1+\beta_4$
B	4	II	$y_{24}=\mu_y+\alpha_2+\beta_4+c$
A	5	I	$y_{15}=\mu_y+\alpha_1+\beta_5$
B	5	II	$y_{25}=\mu_y+\alpha_2+\beta_5+c$
A	6	I	$y_{16}=\mu_y+\alpha_1+\beta_6$
B	6	II	$y_{26}=\mu_y+\alpha_2+\beta_6+c$

這個 case 中

$$E(\bar{y}_1-\bar{y}_2)=\alpha_1-\alpha_2-c$$

case 2. Block design + randomized

Keyboard	Manu	Order	Model
A	1	I	$y_{11}=\mu_y+\alpha_1+\beta_1$
B	1	II	$y_{21}=\mu_y+\alpha_2+\beta_1+c$
A	2	I	$y_{12}=\mu_y+\alpha_1+\beta_2$
B	2	II	$y_{22}=\mu_y+\alpha_2+\beta_2+c$
B	3	I	$y_{23}=\mu_y+\alpha_2+\beta_3$
A	3	II	$y_{13}=\mu_y+\alpha_1+\beta_3+c$
A	4	I	$y_{14}=\mu_y+\alpha_1+\beta_4$
B	4	II	$y_{24}=\mu_y+\alpha_2+\beta_4+c$
A	5	I	$y_{15}=\mu_y+\alpha_1+\beta_5$
B	5	II	$y_{25}=\mu_y+\alpha_2+\beta_5+c$
B	6	I	$y_{26}=\mu_y+\alpha_2+\beta_6$
A	6	II	$y_{16}=\mu_y+\alpha_1+\beta_6+c$

這個 case 中

$$E(\bar{y}_1 - \bar{y}_2) = \alpha_1 - \alpha_2 - \frac{1}{3}c$$

case 3. Block design + Balance randomized

Keyboard	Manu	Order	Model
A	1	I	$y_{11} = \mu_y + \alpha_1 + \beta_1$
B	1	II	$y_{21} = \mu_y + \alpha_2 + \beta_1 + c$
A	2	I	$y_{12} = \mu_y + \alpha_1 + \beta_2$
B	2	II	$y_{22} = \mu_y + \alpha_2 + \beta_2 + c$
B	3	I	$y_{23} = \mu_y + \alpha_2 + \beta_3$
A	3	II	$y_{13} = \mu_y + \alpha_1 + \beta_3 + c$
A	4	I	$y_{14} = \mu_y + \alpha_1 + \beta_4$
B	4	II	$y_{24} = \mu_y + \alpha_2 + \beta_4 + c$
B	5	I	$y_{25} = \mu_y + \alpha_2 + \beta_5$
A	5	II	$y_{15} = \mu_y + \alpha_1 + \beta_5 + c$
B	6	I	$y_{26} = \mu_y + \alpha_2 + \beta_6$
A	6	II	$y_{16} = \mu_y + \alpha_1 + \beta_6 + c$

這個 case 中

$$E(\bar{y}_1 - \bar{y}_2) = \alpha_1 - \alpha_2$$

從上面三個 case 中我們可以發現，
1. 使用 Block design 首先消除了各個手稿不同造成的影響
2. 執行 randomized 稍稍減少了 learning effect 的影響
3. 接著最後執行 balance randomized 則是消除了 $E(\bar{y}_1 - \bar{y}_2)$ 的 bias term，使得估計 treatment effect 時可以更加精準。

(b) Would you use it for the study? If not, what would you do? What aspect of the sequence makes you uneasy? Can you relate it to the possibility that the advantage of the learning effect may diminish over time and express it in more rigorous terms?

上述題目中的設計並不好，因為這個順序看起來並不是那麼 random，可以重新執行 randomized 一次盡量讓兩種次序交錯，這種設計可能會導致一些隱藏的效應影響到模型的估計，例如打前三份稿與後三份稿的人剛好不同的話，會與鍵盤打字順序的 effect 產生很強的共線性或者是有 unidentifiable 的狀況發生，導致模型沒辦法正確估計。
就 learning effect 而言，第二次打同一份稿會相較第一次更為熟練，打字效率更好，但如果 learning effect 並不是一個固定常數的話，這個順序即使是使用 balance design 也可能在估計 treatment effect 時也會產生偏差，故若 learning effect 是固定的話，則 treatment effect 才會有不偏的性質。

Problem 2.

Two groups of students are conducting the same experiment on some sample units. There are two factors which can be adjusted to affect the outcome. Each factor has only two settings. There are only two students in each group. Group 1 does their experiment in the following way: every treatment is only applied to one unit, then both students read the measurements and record their readings. Group 2 does their experiment in another way: each treatment is applied to two units, and each student reads only one of the measurements. The two groups have the same number of observations. What are the differences between their observations? In what situation would one group obtain more accurate results than the other? Why? Is there a block factor?

What are the differences between their observations?

此實驗包括 2 個 factors 及 2 個 settings，因此實驗中的 treatment 個數為 2 (factors) $\times$ 2 (settings) = 4 (treatments)，假設 2 個 factors 分別為 Factor One 及 Factor Two，2 個 settings 分別為 + 及 - ，treatment 可以分為以下四種

Treatment	Factor One	Factor Two
1	+	+
2	+	-
3	-	+
4	-	-

- Group1 的分法為將每種 treatment 個別作用在一個 experimental unit(EU)，並讓 student1 及 student2 都對其觀察，因此 Group1 共使用了 4 個 experimental units，並得到 8 筆觀測值，且因為此實驗將一種 treatment 對應一個 EU，因此為 repetition，實驗誤差 (experimental error) 僅包含測量誤差 (measurement error)。
- Group2 的分法為將每種 treatment 個別作用在兩個 EU，並讓 student1 或 student2 其一觀察，因此 Group2 共使用了 8 個 experimental units，並得到 8 筆觀測值，且因為此實驗將同種 treatment 作用在不同的 EU 上，因此為 replication，實驗誤差包含測量誤差及 EU 間的誤差。
- Group1 因為是 repetition，沒有將同種 treatment 作用在不同 EU，因此缺乏 unit-unit variation，treatment effect 的估計誤差較大，但因為讓不同學生觀察，因此 Group1 實驗結果包含 student-student variation 及 measurement error，但通常對 student-student variation 不感興趣。
- Group2 因為是 replication，將同種 treatment 作用在不同 EU，因此包含 unit-unit variation，treatment effect 的估計誤差較小，但因為擇一學生觀察，因此 Group2 實驗結果缺乏 student-student variation，只有 unit-unit variation 及 measurement error。

In what situation would one group obtain more accurate results than the other? Why?

此實驗中可能造成的變異有以下幾種

1. student-student variation: 不同學生觀測造成差異
2. unit-unit variation: 同種 treatment 作用於不同 EU 造成的差異
3. measurement error: 測量時造成的誤差

觀察以下兩個 case:

- Group1 因為是 repetition，沒有將同種 treatment 作用在不同 EU，因此缺乏 unit-unit variation，但因為讓不同學生觀察，因此 Group1 實驗結果包含 student-student variation 及 measurement error，另外，比較 Group1 和 Group2，因為 Group1 的 EU 較少，因此我們得到關於 unit-unit variation 的資訊也會比較少，treatment 估計的 variance 也會比較大。

- Group2 因為是 replication，將同種 treatment 作用在不同 EU，因此包含 unit-unit variation，treatment effect 的估計誤差較小，但因為擇一學生觀察，因此 Group2 實驗結果缺乏 student-student variation。

討論以下兩個 case:

1. 當 student-student variation≫unit-unit variation，我們需要利用 repetition 來降低 student-student variation 所造成的誤差，因此 Group1 較適合。
2. 當 student-student variation≪unit-unit variation，我們需要利用 replication 來降低 unit-unit variation 所造成的誤差，因此 Group2 較適合。

Is there a block factor?

通常會將可控且會影響實驗結果但並非實驗所重視的因子設定為 block factor，根據此題可以發現 student 並非此實驗重視的變數，且隨著更多不同的觀察者會帶來與先前完全不同的影響，因此為了著重在 treatment effect，須將 student 設定為 block factor。

Problem 3.

Suppose in Problem 2, the settings of the two factors require highly accurate measurements. Minor deviation from the nominal value of the settings could lead to the failure of the experiment, which is judged by the reading of the outcome. If the measurement is less than a certain value, this run of the experiment fails. The following table shows how Group 2 conducts the experiment. For example, in the first run, setting of the two factors is completed by student A and the measurement is read by the same student. Note that in this table, each treatment has been applied to two units.

Run	Factor One	Factor Two	Setting	Reading
1	+	+	A	A
2	+	+	A	A
3	+	-	A	A
4	+	-	B	B
5	-	+	B	B
6	-	+	A	A
7	-	-	A	A
8	-	-	B	B

The higher the reading, the better is the experimental run. If the reading shows that this run fails, it has to be repeated again by the same student. Is there any problem with this design? What suggestion can you provide to improve the design?

本題實驗設計存在以下幾個問題：

1. student A 的實驗次數 (setting) 及實驗紀錄 (reading) 較多，並不 balanced，A 與 B 的實驗及記錄次相同較能使 treatment effect 的估計更加精準，且通常 student 並非實驗所重視的因子，因此會設定為 block factor
2. 能發現當 Factor1 和 Factor2 同設定時 (Factor1,Factor2)=(+,+)/(-,-) 時多為 student A 實驗，這可能造成 slightly confounded，此外，run1 和 run2 實驗設定相同，也都由 student A 實驗，仍可能會有 measurement error，會對 learning effect 的估計準確度造成影響
3. 所有實驗皆由同一人 setting 和 reading，若將兩者 student 錯開，也許能減少實驗失誤發生

從以上三點來做改善，我們在以下提出新的設計

Run	Factor One	Factor Two	Setting	Reading
1	+	+	A	B
2	-	+	A	B
3	+	-	A	B
4	-	-	A	B
5	+	+	B	A
6	-	+	B	A
7	+	-	B	A
8	-	-	B	A

1. 為改善 1，我們讓 student A 和 student B 分別做四次實驗及記錄，其中 run1~run4 為 student A 實驗，run5~run8 為 student B 實驗
2. 為改善 2，我們讓 student A 和 student B 各執行 (+,+),(+,-),(-,+),(+,-) 一次
3. 為改善 3，我們讓每次實驗及記錄人不同