實驗設計與分析 HW4 solution

Problem 1.

Do Problem 1 in Chapter 3 of textbook.

The following are the weights (in grams) of six rock samples (from different lakes) measured on two different scales. The purpose of this data collection is to test whether the two scales are different.

	Weight in Grams by Sample Number					
Scale	1	2	3	4	5	6
I	8	14	16	19	18	12
II	11	16	20	18	20	15

Figure 1: Table

(a) Perform an appropriate test at the 0.05 level to test the null hypothesis that scale I and scale II give the same measurement. Formulate this hypothesis and give details of your computations.

Set model:

$$y_{ij} = \eta + \alpha_i + \tau_j + \epsilon_{ij} \quad \text{for } i = 1, \cdots, 6 \quad j = 1, 2$$

其中

α_i 為不同石頭間的 block effect，τ_j 則是兩不同測量方式帶來的 treatment effect，而我們欲檢定的是兩種測量方式有沒有顯著差異，對於這個議題我們設定假設如下

$$H_0 : \tau_1 = \tau_2 \quad v.s. \quad H_1 : \tau_1 \neq \tau_2$$

則我們可以使用 paired t test 或是 F test 來做檢定。首先使用 paired t test

```
kable(tidy(t.test(weight~scale,data = dat1,paired = T)),
      format='markdown',caption = "paired t test table")
```

Table 1: paired t test table

estimate	statistic	p.value	parameter	conf.low	conf.high	method	alternative
-2.166667	-3.081295	0.0274292	5	-3.974216	-0.3591172	Paired t-test	two.sided

在 0.05 的信心水準底下，由 paired t test 檢定結果我們有足夠證據說明兩種測量方法存在明顯差距。接著使用 F test

```
options(knitr.kable.NA = '')
kable(tidy(aov(weight~rock + scale,data = dat1)),
      format='markdown',caption = "ANOVA table")
```

Table 2: ANOVA table

term	df	sumsq	meansq	statistic	p.value
rock	5	135.416667	27.083333	18.258427	0.0031539
scale	1	14.083333	14.083333	9.494382	0.0274292
Residuals	5	7.416667	1.483333		

可得相同檢定結果。

(b) Compute the p value of the test statistic you obtained in (a).
根據 (a) 小題我們可以得到 p value = 0.0274292

Problem 2.

Do Problem 9 in Chapter 3 of textbook.

An experimenter believes that drug A lowers the diastolic blood pressure in people with high blood pressure and that it takes effect in two hours. Design a simple comparative experiment with a control group to assess the effect of the drug A. Tell the experimenter exactly what steps to take to carry out the experiment.
為比較 A 藥物是否有實質藥效，我們可以使用 paired comparison design。實驗步驟如下：

1. 將具有相似身體數據的受測者兩兩一組。
2. 隨機將每組受測者的其中一人分配 drug A，另一人則分配安慰劑。
3. 實驗前先進行一次舒張壓的測量，服用藥物後的兩個小時再進行一次測量。
4. 針對最後測量值進行 paired t test。

Problem 3.

Do Problem 13 in Chapter 3 of textbook.

For the composite experiment of Section 2.3, theoriginal paper by Mazumdar and Hoa (1995) reported a second factor, tape speed. Table 3.44 shows the three replicates for each level of laser power corresponding to tape speed of 6.42, 13.0, and 27.0 m/s, respectively. The levels of tape speed are roughly evenly spaced on the log scale, so that linear and quadratic effects can be entertained for the second quantitative factor. Analyze the experiment as a two-way layout with a single replicate, including model building, parameter estimation, ANOVA and residual analysis.

Table 3.44 Strength Data, Revised Composite Experiment

	Laser Power		
Tape Speed	40W	50W	60W
6.42	25.66	29.15	35.73
13.00	28.00	35.09	39.56
27.00	20.65	29.79	35.66

Figure 2: Table 3.44

根據題目所述，此實驗為一 two-way layout with a single replicate，且須考慮 linear 及 quadratic effects，首先先在下方建立模型

$$y_{ijl} = \eta + \alpha_i^{(l)} + \beta_j^{(l)} + \alpha_i^{(q)} + \beta_j^{(q)} + \omega_{ij}^{(l\times l)} + \omega_{ij}^{(q\times q)} + \epsilon_{ijl} \quad \text{for } i = 1,2,3 \quad j = 1,2,3 \quad l = 1$$

其中 η 代表 overall mean，$\alpha_i^{(l)}$ 代表 laser power 的 linear effect，$\alpha_i^{(q)}$ 代表 laser power 的 quadratic effect，$\beta_j^{(l)}$ 代表 log(tape speed) 的 linear effect，$\beta_j^{(q)}$ 代表 log(tape speed) 的 quadratic effect，$\omega_{ij}^{(l\times l)}$ 代表 linear effect 間的交互作用項，$\omega_{ij}^{(q\times q)}$ 代表 quadratic effect 間的交互作用項，$\epsilon_{ijl} \sim N(0, \sigma^2)$ 代表誤差項，另外 l 為 replicate 的次數，因為此題為 single replicate，l 皆為 1，因此省略，隨後觀察到因為實驗次數為 9 次，而但考慮到 laser power 和 log(tape speed) 的 d.f. 皆為 2，交互作用項的 d.f. 為 4，若同時加入所有交互作用會導致模型變為 saturated model 而 residual 的 d.f. 變為 0，導致無法估計，因此這裡先不考慮交互作用進行模型配適，重新得到以下模型

$$y_{ij} = \eta + \alpha_i^{(l)} + \beta_j^{(l)} + \alpha_i^{(q)} + \beta_j^{(q)} + \epsilon_{ij} \quad \text{for } i = 1,2,3 \quad j = 1,2,3$$

另外，我們可以觀察到因為 laser power 及 log(tape speed) 皆為 equally spaced with order 的 factors，因此我們增加 normalized 的 linear contrast 及 quadratic contrast，並稍加更改模型

$$y_i = \eta + \alpha^{(l)} P^{(l)}(\text{power}_i) + \beta^{(l)} P^{(l)}(\log(\text{speed})_i) + \alpha^{(q)} P^{(q)}(\text{power}_i) + \beta^{(q)} P^{(q)}(\log(\text{speed})_i) + \epsilon_i \quad \text{for } i = 1, \cdots, 9$$

其中 normalized contrast vectors 為

$$\text{linear}: \ (-1, 0, 1)/\sqrt{2}$$
$$\text{quadratic}: \ (1, -2, 1)/\sqrt{6}$$

接著在下方進行參數估計

```
dat3$power.L <- rep(c(-1,0,1),each=3)/sqrt(2)
dat3$logspeed.L <- rep(c(-1,0,1),times=3)/sqrt(2)
dat3$power.Q <- rep(c(1,-2,1),each=3)/sqrt(6)
```

```
dat3$logspeed.Q <- rep(c(1,-2,1),times=3)/sqrt(6)
lm.fit3 <- lm(strength~power.L+power.Q+logspeed.L+logspeed.Q,data = dat3)
sumtab3 <- summary.lm(lm.fit3)
kable(sumtab3$coefficients,format='markdown',caption = "model fitting table")
```

Table 3: model fitting table

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	31.0322222	0.5401400	57.452182	0.0000005
power.L	8.6361308	0.9355499	9.231074	0.0007654
power.Q	-0.3810317	0.9355499	-0.407281	0.7046552
logspeed.L	-1.0465180	0.9355499	-1.118613	0.3259444
logspeed.Q	-3.9001320	0.9355499	-4.168812	0.0140447

可以發現 power 的 quadratic effect 及 log(tape speed) 的 linear effect 並不顯著，因此我們可以考慮移除，然而因為 log(tape speed) 的 quadratic effect 顯著，若只保留 quadratic effect 會失去模型的解釋意義，因此僅移除 power 的 quadratic effect 重新配適以下模型

$$y_i = \eta + \alpha^{(l)} P^{(l)}(\text{power}_i) + \beta^{(l)} P^{(l)}(\log(\text{speed})_i) + \beta^{(q)} P^{(q)}(\log(\text{speed})_i) + \epsilon_i \quad \text{for } i = 1, \cdots, 9$$

接著進行參數估計

```
lm.fit3_r <- lm(strength~power.L+logspeed.L+logspeed.Q,data = dat3)
sumtab3_r <- summary.lm(lm.fit3_r)
kable(sumtab3_r$coefficients,format='markdown',caption = "model fitting table")
```

Table 4: model fitting table

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	31.032222	0.4930314	62.941670	0.0000000
power.L	8.636131	0.8539555	10.113092	0.0001620
logspeed.L	-1.046518	0.8539555	-1.225495	0.2749607
logspeed.Q	-3.900132	0.8539555	-4.567137	0.0060176

透過上方的 model fitting table，我們得到的參數估計為

$$\hat{\eta} = 31.03222, \ \hat{\alpha}^{(l)} = 8.636131, \ \hat{\beta}^{(l)} = -1.046518, \ \hat{\beta}^{(q)} = -3.900132$$

接著使用 sequential F test 建立下方的 ANOVA table

```
options(knitr.kable.NA = '')
kable(anova(lm.fit3_r),format='markdown',caption = "ANOVA table")
```

Table 5: ANOVA table

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
power.L	1	223.74827	223.74827	102.274636	0.0001620
logspeed.L	1	3.28560	3.28560	1.501837	0.2749607
logspeed.Q	1	45.63309	45.63309	20.858743	0.0060176
Residuals	5	10.93860	2.18772		

可以發現將 power 及 log(speed) 視為連續型變數時有和 drop-one t test 有一樣的結果，接著將兩變數以沒有 order 的類別型變數建立 ANOVA table 觀察

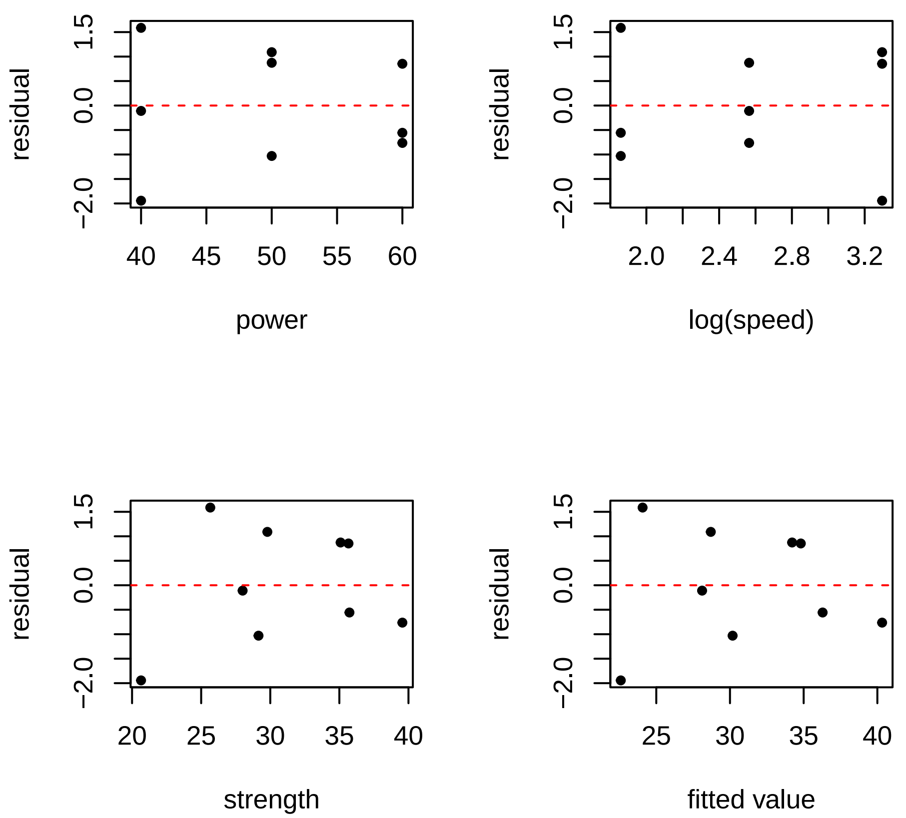
```r
lm.fit3_f <- lm(strength~factor(power)+factor(speed),data = dat3)
options(knitr.kable.NA = '')
kable(anova(lm.fit3_f),format='markdown',caption = "ANOVA table")
```

Table 6: ANOVA table

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
factor(power)	2	224.18382	112.091911	42.689303	0.0020029
factor(speed)	2	48.91869	24.459344	9.315145	0.0312421
Residuals	4	10.50304	2.625761		

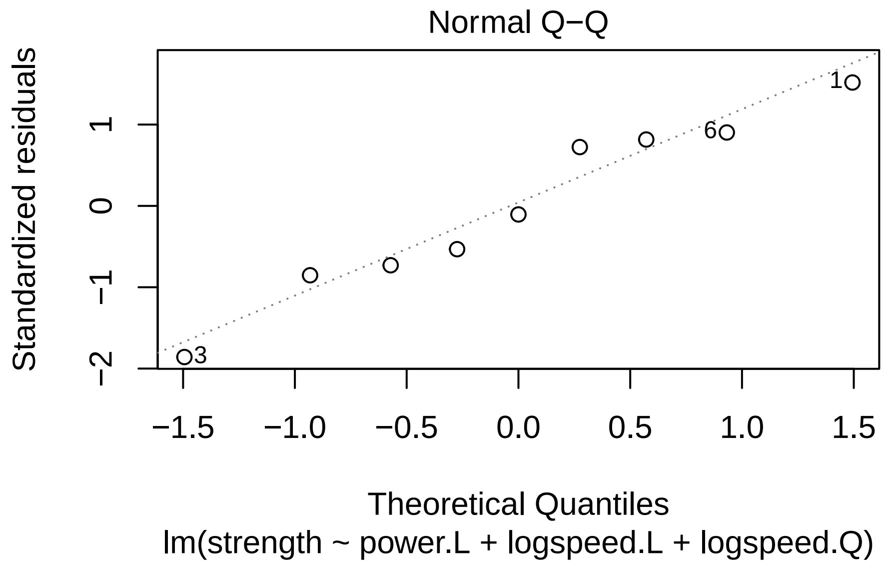
可以發現 power 彼此之間及 speed 彼此之間皆有顯著差異，最後進行 residual analysis

```r
par(mfrow=c(2,2))
plot(dat3$power,lm.fit3_r$residuals,xlab='power',ylab='residual',pch=20)
abline(h=0,lty=2,col="red")
plot(log(dat3$speed),lm.fit3_r$residuals,
     xlab='log(speed)',ylab='residual',pch=20)
abline(h=0,lty=2,col="red")
plot(dat3$strength,lm.fit3_r$residuals,xlab='strength',ylab='residual',pch=20)
abline(h=0,lty=2,col="red")
plot(lm.fit3_r$fitted.values,lm.fit3_r$residuals,
     xlab='fitted value',ylab='residual',pch=20)
abline(h=0,lty=2,col="red")
```



透過上方四張 residual plot 可以發現都沒有太過明顯的 pattern 且 repicates 的平均皆位於 0 左右，接著觀察 Normal-QQ

```
plot(lm.fit3_r,2)
```



可以發現 residul 整體還算服從 normal，因此模型配適還算可行。

Problem 4.

Do Problem 28 in Chapter 3 of textbook.

Natrella (1963, pp. 13–14) described an experiment on a resistor mounted on a ceramic plate in which the impact of four geometrical shapes of the resistors on the current noise of the resistor is studied. Only three resistors can be mounted on one plate. The design and data for the resistor experiment are given in Table 3.48.

Table 3.48 Data, Resistor Experiment

	Shape			
Plate	*A*	*B*	*C*	*D*
1	1.11		0.95	0.82
2	1.70	1.22		0.97
3	1.60	1.11	1.52	
4		1.22	1.54	1.18

Figure 3: Table 3.48

(a) Describe and justify the design.
根據題目所述及 Table3.48 可以發現 shape 為 treatment factor，plate 則為 block factor，且根據表格可以發現並非所有 pairs 都執行實驗，而是在每個 block 下執行三種不同 treatment 的實驗，且滿足每兩種 treatment 間皆存在兩組 WBC 及一組 BBC，因此得出此為一 Balanced Imcomplete Block Design(BIBD)，且各參數如下

$$\text{number of treatments}: t = 4$$
$$\text{number of blocks}: b = 4$$
$$\text{block size}: k = 3$$
$$\text{each treatment repicated times}: r = 3$$
$$\text{number of within blocks comparison}: \lambda = \frac{r(k-1)}{t-1} = 2$$

(b) Analyze the experiment.
首先配適以下模型

$$y_{ij} = \eta + \alpha_i + \tau_j + \epsilon_{ij} \quad \text{for } i = 1,2,3,4, \quad j = A, B, C, D$$

其中 η 代表 overall mean，α_i 代表 block effect(plate)，τ_j 代表 treatment effect(shape)，$\epsilon_{ij} \sim N(0, \sigma^2)$ 代表誤差項，接著在下方建立 ANOVA table

```r
y <- c(1.11,1.70,1.60,1.22,1.11,1.22,0.95,1.52,1.54,0.82,0.97,1.18)
blk <- factor(c(1,2,3,2,3,4,1,3,4,1,2,4))
trt <- gl(4,3,labels = c('A','B','C','D'))
dat4 <- as.data.frame(t(rbind(y,blk,trt)))
aovibd <- aov.ibd(y~factor(blk)+factor(trt),data = dat4,
                  specs = "trt",alpha=0.05,add.intercept = TRUE)$ANOVA.table
options(knitr.kable.NA = '')
kable(aovibd,format='markdown',caption = "ANOVA table")
```

Table 7: ANOVA table

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	2.3869714	1	174.29510	0.0000445
factor(blk)	0.4288583	3	10.43832	0.0136052
factor(trt)	0.4650583	3	11.31942	0.0114559
Residuals	0.0684750	5		

根據上方 ANOVA table 的結果可以發現 treatment(shape) 的 p-value=0.01146<0.05，在顯著水準 $\alpha = 0.05$ 下拒絕 H_0，因此不同 shape 之間具有顯著差異，接著使用 Tukey method 進行 multiple comparison testing，首先建立以下假設檢定

$$H_0^{ij} : \tau_i = \tau_j \quad v.s. \quad H_1^{ij} : \tau_i \neq \tau_j \quad \text{for } i, j = 1, \cdots, 4, i \neq j$$

其中 multiple comparison test statistic 為

$$t_{ij} = \frac{\hat{\tau}_i - \hat{\tau}_j}{se(\hat{\tau}_i - \hat{\tau}_j)}$$

若

$$|t_{ij}| > \frac{1}{\sqrt{2}} q_{t,bk-b-t+1,\alpha} = \frac{1}{\sqrt{2}} q_{4,5,0.05} = \frac{5.218325}{\sqrt{2}} = 3.689913$$

則拒絕 H_0

```r
y <- c(1.11,1.70,1.60,1.22,1.11,1.22,0.95,1.52,1.54,0.82,0.97,1.18)
blk <- factor(c(1,2,3,2,3,4,1,3,4,1,2,4))
trt <- gl(4,3,labels = c('A','B','C','D'))
aovtukey <- aov(y~blk+trt)
library(multcomp)
fit <- glht(aovtukey,linfct=mcp(trt="Tukey"))
aovcom <- summary(fit)
aovcom
```

```
##
##   Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: aov(formula = y ~ blk + trt)
##
## Linear Hypotheses:
##              Estimate Std. Error t value Pr(>|t|)
## B - A == 0 -0.45500    0.10135  -4.490   0.0235 *
## C - A == 0 -0.15625    0.10135  -1.542   0.4813
## D - A == 0 -0.50375    0.10135  -4.971   0.0154 *
## C - B == 0  0.29875    0.10135   2.948   0.1073
## D - B == 0 -0.04875    0.10135  -0.481   0.9602
## D - C == 0 -0.34750    0.10135  -3.429   0.0652 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)
```

根據上方結果可以發現 (A,B)、(A,D) 組合間的 treatment effect 具有顯著的差異。

Problem 5.

Do Problem 35 in Chapter 3 of textbook.

To compare the effects of five different assembly methods (denoted by the Latin letters A, B, C, D, and E) on the throughput, an experiment based on a Graeco-Latin square was conducted which involved three blocking variables: day, operator, and machine type. The data are given in Table 3.50, where the machine type is denoted by the five Greek letters. The response, throughput, is the number of completed pieces per day and is given in the parentheses in the table. Analyze the data and compare the five assembly methods. Which of the methods are significantly better than the others?

Table 3.50 Throughput Data

Day	Operator 1		2		3		4		5	
1	$A\alpha$	(102)	$B\beta$	(105)	$C\gamma$	(82)	$D\delta$	(141)	$E\epsilon$	(132)
2	$B\gamma$	(92)	$C\delta$	(112)	$D\epsilon$	(131)	$E\alpha$	(112)	$A\beta$	(99)
3	$C\epsilon$	(96)	$D\alpha$	(130)	$E\beta$	(108)	$A\gamma$	(73)	$B\delta$	(129)
4	$D\beta$	(120)	$E\gamma$	(100)	$A\delta$	(111)	$B\epsilon$	(116)	$C\alpha$	(100)
5	$E\delta$	(123)	$A\epsilon$	(110)	$B\alpha$	(111)	$C\beta$	(85)	$D\gamma$	(100)

Figure 4: Table 3.50

首先配適以下模型

$$y_{ij} = \eta + \alpha_i + \beta_j + \tau_l + \zeta_m + \epsilon_{ijlm} \quad \text{for } i = 1, \cdots, 5, \quad j = 1, \cdots, 5, \quad l = A, \cdots, E, \quad m = \alpha, \cdots, \epsilon$$

其中 η 代表 overall mean，α_i 代表 Day，β_j 代表 Operator，τ_l 代表 different assembly methods(Latin letter)，ζ_m 代表 machine type(Greek letter)，$\epsilon_{ijlm} \sim N(0, \sigma^2)$ 代表誤差項，接著在下方建立 ANOVA table

```
aovglsd <- tidy(aov(Throughput~Day+Operator+Method+Machine,data = dat5))
options(knitr.kable.NA = '')
kable(aovglsd,format='markdown',caption = "ANOVA table")
```

Table 8: ANOVA table

term	df	sumsq	meansq	statistic	p.value
Day	4	125.2	31.3	1.534314	0.2806024
Operator	4	167.2	41.8	2.049020	0.1800250
Method	4	2857.6	714.4	35.019608	0.0000408
Machine	4	3424.8	856.2	41.970588	0.0000206
Residuals	8	163.2	20.4		

根據上方 ANOVA table 的結果可以發現 treatment effect(Method) 及 block effect(Machine) 的 p-value<0.05，在顯著水準 $\alpha = 0.05$ 下拒絕 H_0，因此不同 assembly methods 及 machine type 各自間具有顯著差異，接著使用 Tukey method 進行 multiple comparison testing

```
aovtukey <- aov(Throughput~Day+Operator+Method+Machine,data = dat5)
fit <- glht(aovtukey,linfct=mcp(Method="Tukey"))
aovcom <- summary(fit)
aovcom
```

```
##
##   Simultaneous Tests for General Linear Hypotheses
##
```

```
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: aov(formula = Throughput ~ Day + Operator + Method + Machine,
##     data = dat5)
##
## Linear Hypotheses:
##            Estimate Std. Error t value Pr(>|t|)
## B - A == 0   11.600      2.857   4.061  0.02222 *
## C - A == 0   -4.000      2.857  -1.400  0.64405
## D - A == 0   25.400      2.857   8.892  < 0.001 ***
## E - A == 0   16.000      2.857   5.601  0.00338 **
## C - B == 0  -15.600      2.857  -5.461  0.00395 **
## D - B == 0   13.800      2.857   4.831  0.00827 **
## E - B == 0    4.400      2.857   1.540  0.56752
## D - C == 0   29.400      2.857  10.292  < 0.001 ***
## E - C == 0   20.000      2.857   7.001  < 0.001 ***
## E - D == 0   -9.400      2.857  -3.291  0.06261 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)
```

根據上方結果可以發現 (A,B)、(A,D)、(A,E)、(B,C)、(B,D)、(C,D)、(C,E) 組合具有顯著的差異，且根據估計結果可以發現 Method D 最好。