實驗設計 HW3
統碩一 106354012 曹立諭

2018/04/03

1. Problem 3.1

Cardiac pacemakers contain electrical connections that are platinum pins soldered onto a substrate. The question of interest is whether different operators produce solder joints with the same strength. Twelve substrates are randomly assigned to four operators. Each operator solders four pins on each substrate, and then these solder joints are assessed by measuring the shear strength of the pins. Data from T. Kerkow.

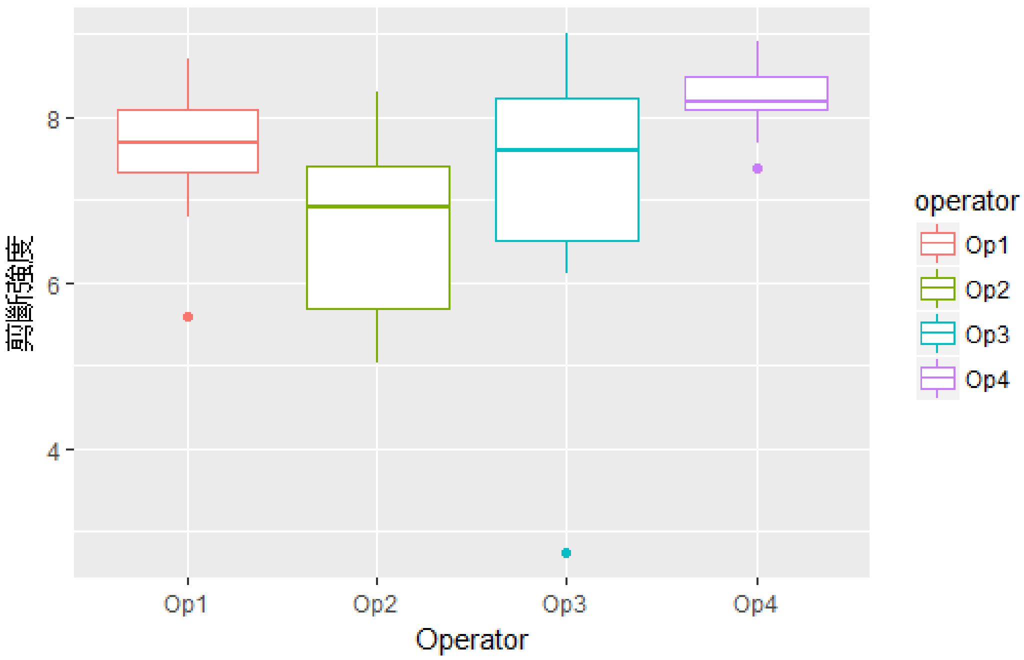
	Strength (lb)		
Operator	Substrate 1	Substrate 2	Substrate 3
1	5.60 6.80 8.32 8.70	7.64 7.44 7.48 7.80	7.72 8.40 6.98 8.00
2	5.04 7.38 5.56 6.96	8.30 6.86 5.62 7.22	5.72 6.40 7.54 7.50
3	8.36 7.04 6.92 8.18	6.20 6.10 2.75 8.14	9.00 8.64 6.60 8.18
4	8.30 8.54 7.68 8.92	8.46 7.38 8.08 8.12	8.68 8.24 8.09 8.06

Analyze these data to determine if there is any evidence that the operators produce different mean shear strengths. (Hint: what are the experimental units?)

$$H_0: 工人的剪斷強度沒有差別$$

$$H_1: 工人的剪斷強度有差別$$

畫個合鬚圖比較一下



這是一個 ANOVA 表格

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Operators	3	15.1888	5.062961	4.24272	0.0101966
Residuals	44	52.50648	1.193329		

給定顯著水準 $\alpha = 0.05$ ，因為p-value小於α

所以傾向拒絕虛無假設

結論 ： 四位工人的剪斷強度有顯著的差別

2. Exercise 4.3

Refer to the data in Problem 3.1. Workers 1 and 2 were experienced,whereas workers 3 and 4 were novices. Find a contrast to compare the experienced and novice workers and test the null hypothesis that experienced and novice works produce the same average shear strength.

H_0: experienced and novice works produce the same average shear strength
H_1: experienced and novice works does not produce the same average shear strength

假設也可以寫成
H_0: 有經驗的工人剪斷強度平均與沒經驗的工人剪斷強度平均相同
H_1: 有經驗的工人剪斷強度平均與沒經驗的工人剪斷強度平均不同

Let contrast $w = (\ 0.5\ ,\ 0.5\ , -0.5\ , -0.5\ )$

$$\sum_{i=1}^{t} w_i \bar{y}_{i.} \sim N\left(\sum_{i=1}^{g} w_i \mu_i , \sigma^2 \sum_{i=1}^{g} \frac{w_i^2}{n_i}\right)$$

$$\frac{\sum_{i=1}^{g} w_i \bar{y}_{i.} - \sum_{i=1}^{g} w_i \mu_i}{\sqrt{\sigma^2 \sum_{i=1}^{g} \frac{w_i^2}{n_i}}} \sim N(0,1) \quad => \quad t = \frac{\sum_{i=1}^{g} w_i\ \bar{y}_i}{\sqrt{MS_E}\sqrt{\sum_{i=1}^{g} \frac{w_i^2}{n_i}}} \sim t_{N-g}$$

給定顯著水準 $\alpha = 0.05$ 下

$|t_0| = 1.807751 < t_{0.025}(44) = 2.015368$ (雙尾檢定)

$\therefore |t_0| < t_{0.025}(44)$ 所以我們傾向不拒絕虛無假設

結論 : 我們認為有經驗的工人與沒經驗的工人剪斷強度平均相同

實驗設計 HW3
統碩一 106354012 曹立諭

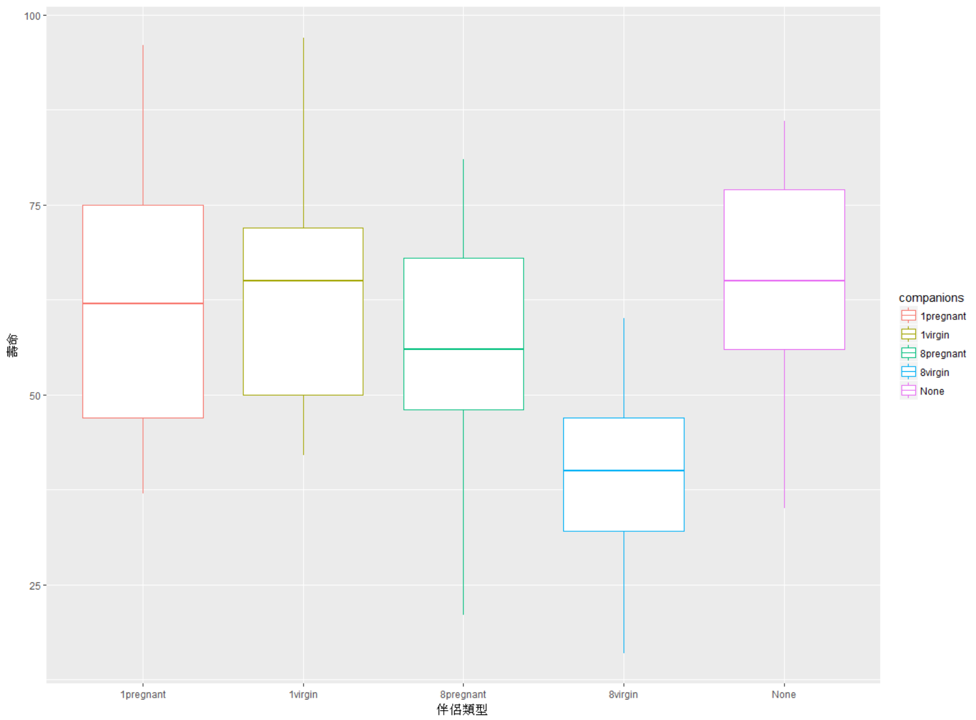
2018/04/03

3. Problem 3.2 (Descriptive analysis)

Scientists are interested in whether the energy costs involved in reproduction affect longevity. In this experiment, 125 male fruit flies were divided at random into five sets of 25. In one group, the males were kept by themselves.In two groups, the males were supplied with one or eight receptive virgin female fruit flies per day. In the final two groups, the males were supplied with one or eight unreceptive (pregnant) female fruit flies per day. Other than the number and type of companions, the males were treated identically. The longevity of the flies was observed. Data from Hanley and Shapiro (1994).

Companions	Longevity (days)												
None	35	37	49	46	63	39	46	56	63	65	56	65	70
	63	65	70	77	81	86	70	70	77	77	81	77	
1 pregnant	40	37	44	47	47	47	68	47	54	61	71	75	89
	58	59	62	79	96	58	62	70	72	75	96	75	
1 virgin	46	42	65	46	58	42	48	58	50	80	63	65	70
	70	72	97	46	56	70	70	72	76	90	76	92	
8 pregnant	21	40	44	54	36	40	56	60	48	53	60	60	65
	68	60	81	81	48	48	56	68	75	81	48	68	
8 virgin	16	19	19	32	33	33	30	42	42	33	26	30	40
	54	34	34	47	47	42	47	54	54	56	60	44	

Analyze these data to test the null hypothesis that reproductive activity does not affect longevity. Write a report on your analysis. Be sure to describe the experiment as well as your results.



$$H_0: Reproductive\ \ activity\ \ does\ \ not\ \ affect\ \ longevity.$$

$$H_1: Reproductive\ \ activity\ \ affect\ \ longevity.$$

ANOVA TABLE

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Flys	4	11939	2984.82	13.612	0.000000003516
Residuals	120	26314	219.28		

給定顯著水準 $\alpha = 0.05$ 下 ， 因為p-value遠小於α，我們傾向拒絕虛無假設

結論 ：即繁殖行為會影響壽命長短

4. Problem 4.2

Consider the data in Problem 3.2. Design a set of contrasts that seem meaningful. For each contrast, outline its purpose and test the null hypothesis that the contrast has expected value zero.

(a) 懷孕 v.s. 處女

H_0: Virgin fly and Prgnant fly reproductive activity does not affect longevity.

H_1: Virgin fly and Prgnant fly reproductive activity affect longevity.

Let contrasts w=(1,-1,1,-1),

$$\because t = \frac{\sum_{i=1}^{g} w_i \ \bar{y}_i}{\sqrt{MS_E}\sqrt{\sum_{i=1}^{g}\frac{w_i^2}{n_i}}} \ , |t_0| = 2.823645 \ \ \geq \ \ t_{N-g,\alpha/2} = t_{0.025}(96) = 1.985$$

我們傾向拒絕虛無假設，即伴侶類型會影響壽命長短。

(b) 沒放伴侶 v.s. 有放伴侶

H_0: fly with no companion and fly with companions reproductive activity does not affect longevity.

H_1: fly with no companion and fly with companions reproductive activity affect longevity.

Let contrasts w=(1,-1/4,-1/4,-1/4,-1/4),

$$\because t = \frac{\sum_{i=1}^{g} w_i \ \bar{y}_i}{\sqrt{MS_E}\sqrt{\sum_{i=1}^{g}\frac{w_i^2}{n_i}}} \ , |t_0| = 2.224885 \ \ \geq \ \ t_{N-g,\alpha/2} = t_{0.025}(120) = 1.9799$$

我們傾向拒絕虛無假設，即伴侶有無會影響壽命長短。

5. Exercise 4.4

Consider an experiment taste-testing six types of chocolate chip cookies:1 (brand A, chewy, expensive), 2 (brand A, crispy, expensive), 3 (brand B,chewy, inexpensive), 4 (brand B, crispy, inexpensive), 5 (brand C, chewy,expensive), and 6 (brandD, crispy, inexpensive). We will use twenty different raters randomly assigned to each type (120 total raters).

(a) Design contrasts to compare chewy with crispy, and expensive with inexpensive.

- ✓ chewy with crispy

  Let contrasts w=(1,-1,1,-1,1,-1)

  H_0: 1、3、5組的分數平均與2、4、6組的分數平均相同

  H_1: 1、3、5組的分數平均與2、4、6組的分數平均不同

- ✓ expensive with inexpensive

  Let contrast w=(1,1,-1,-1,1,-1)

  H_0: 1、2、5組的分數總和平均與3、4、6組的分數總和平均相同

  H_1: 1、2、5組的分數總和平均與3、4、6組的分數總和平均不同

(b) Are your contrasts in part (a) orthogonal? Why or why not?

$$\mathrm{W} = \sum_{i=1}^{6} \frac{w_i * t(w_i)}{n_i} = 0.1 \neq 0$$

So , it’s not orthogonal.

6. Question 3.2

Proof:

$$\sum_{i=1}^{g}\sum_{j=1}^{n_i}\hat{\alpha}_i r_{ij} = 0$$

Solve:

$$\sum_{i=1}^{g}\sum_{j=1}^{n_i}\hat{\alpha}_i r_{ij}$$

$$= \sum_{i=1}^{g}\sum_{j=1}^{n_i}(\bar{y}_i - \bar{y}_{..})\left(y_{ij} - \bar{y}_{i.}\right)\ \left(\because r_{ij} = y_{ij} - \bar{y}_{i.}\right)$$

$$= \sum_{i=1}^{g}(\bar{y}_i - \bar{y}_{..})\sum_{j=1}^{n_i}\left(y_{ij} - \bar{y}_{i.}\right)\left(\because \sum_{j=1}^{n_i}\left(y_{ij} - \bar{y}_i\right) = 0\right)$$

$$= 0$$

7. Question 4.1

Show that orthogonal contrasts in the observed treatment means are uncorrelated random variables.

Let w_1 and w_2 be the orthogonal contrasts.

$\bar{Y}_{i.} \sim N(\mu\ , \frac{\sigma^2}{n_i})$

$$u = \sum_{i=1}^{n} w_{1i}\bar{Y}_{i.} \quad , \qquad v = \sum_{j=1}^{n} w_{2j}\bar{Y}_{j.} \ ,$$

$\text{cov}(\mu, v)$

$$= \text{cov}\left(\sum_{i=1}^{n} w_{1i}\bar{Y}_{i.}\ , \sum_{j=1}^{n} w_{2j}\bar{Y}_{j.}\right)$$

$$= \sum_{i=j}^{n} w_{1i}w_{2j}\,\text{cov}(\bar{Y}_{i.}, \bar{Y}_{j.}) + \sum_{i\neq j}^{n} w_{1i}w_{2j}\,\text{cov}(\bar{Y}_{i.}, \bar{Y}_{j.})$$

$$(\because \text{cov}(\bar{Y}_{i.}, \bar{Y}_{j.}) = 0\ , when\ \ i \neq j\ \ and\ \ var(\bar{Y}_{i.}) = var(\bar{Y}_{j.}) = \frac{\sigma^2}{n_i}\ )$$

$$= \sum_{i=1}^{n} w_{1i}w_{2i}var(\bar{Y}_{i.}) + 0$$

$$= \sigma^2 \sum_{i=1}^{n} \frac{w_{1i}w_{2i}}{n_i} \quad (\because w_1 \text{ and } w_2 \text{ are orthogonal})$$

$$= 0$$

So,the orthogonal contrasts in the observed treatment means are uncorrelated.