Design of experiment - Homework 03

統碩一 106354003 林健宏

1. Question 3.2

Prove

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

Solve

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

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

$$\sum_{i=1}^{g}\sum_{j=1}^{n_i}\hat{\alpha}_i r_{ij} = \sum_{i=1}^{g}\sum_{j=1}^{n_i}(\bar{y}_{i.} - \bar{y}_{..})(y_{ij} - \bar{y}_{i.}) = \sum_{i=1}^{g}(\bar{y}_{i.} - \bar{y}_{..})\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.}) = 0$$

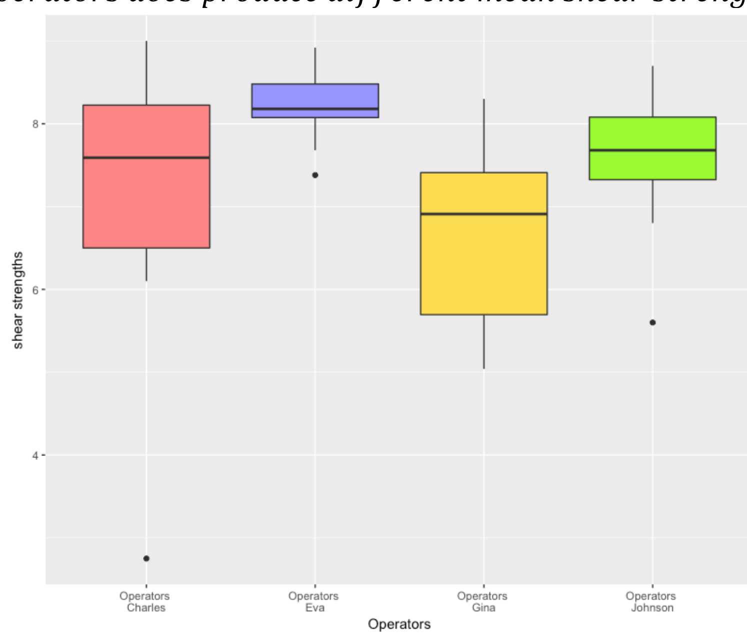
2. Problem 3.1

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.

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?)

$$H_0: Operators\ doesn't\ produce\ different\ mean\ shear\ strengths.$$
$$H_1: Operators\ does\ produce\ different\ mean\ shear\ strengths.$$



This is one way ANOVA table test

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
operators	3	15.18888	5.062961	4.24272	0.0101966
Residuals	44	52.50648	1.193329	NA	NA

Under the level of significant α = 0.05, the p-value = 0.010966 is smaller than α. Therefore, We reject H_0 hypothesis. There is difference between operators.

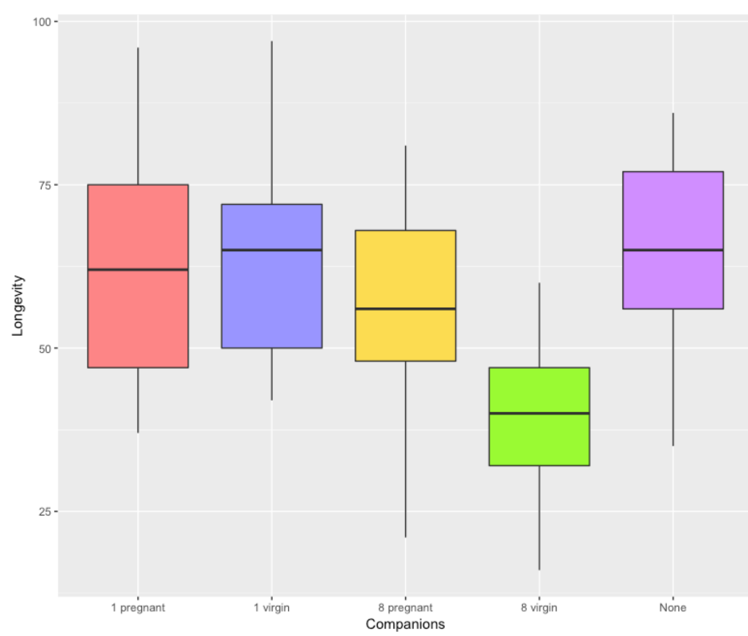
3. Problem 3.2

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.

$$H_0: \textit{Reproductive activity does not affect longevity.}$$
$$H_1: \textit{Reproductive activity does affect longevity.}$$



This is one way ANOVA table test:

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
flys	4	11939.28	2984.8200	13.61195	0
Residuals	120	26313.52	219.2793	NA	NA

Under the level of significant α = 0.05, the p-value = 0.000000003516 is smaller than α.
Therefore, We reject H_0 hypothesis.

4. 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 do not produce the same average shear strength*.

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

H_0: operators1's mean + operators2's mean – operators3's mean - operators4's mean = 0
H_1: operators1's mean + operators2's mean – operators3's mean - operators4's mean $\neq$ 0

We use statistic $$|t| = \left| \frac{\sum_{i=1}^{g} w_i \bar{y}_{i\cdot}}{\sqrt{MSE \sum_{i=1}^{g} w_i^2 / n_i}} \right| \geq t_{N-g,\alpha/2}$$ to test data.

Then, data's t_0 is -1.807529 and $t_{0.025}(44)$ is 2.015368.
Therefore, $|t_0| < t_{0.025}(44)$ means that we don't reject H_0 hypothesis.

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 (brand D, 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: We use contrast as w = (1 ,-1, 1, -1, 1, -1), means

$$\begin{aligned} \mathrm{H_0}&: (\text{brand A, chewy, expensive}) - (\text{brand A, crispy, expensive}) + (\text{brand B, chewy, inexpensive}) \\ &- (\text{brand B, crispy, inexpensive}) + (\text{brand C, chewy, expensive}) \\ &- (\text{brand D, crispy, inexpensive}) = 0 \end{aligned}$$

$\mathrm{H_1}$: Not equals to zero.

- expensive with inexpensive: We use contrast as w* = (1 ,1, -1, -1, 1, -1), means

$$\begin{aligned} \mathrm{H_0}&: (\text{brand A, chewy, expensive}) + (\text{brand A, crispy, expensive}) - (\text{brand B, chewy, inexpensive}) \\ &- (\text{brand B, crispy, inexpensive}) + (\text{brand C, chewy, expensive}) \\ &- (\text{brand D, crispy, inexpensive}) = 0 \end{aligned}$$

$\mathrm{H_1}$: Not equals to zero.

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

$\because$ W = $\sum_{i=1}^{6} \frac{w_i w_i^*}{n_i}$ = 0.1 $\neq$ 0, $\mathrm{n_i} = 20$ $\therefore$ It is not orthogonal.

6. Question 4.1

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

Let $\boldsymbol{w_1}$ *and* $\boldsymbol{w_2}$ be the orthogonal contrasts.

$u = \sum_{i=1}^{n} w_{1i}\,\bar{Y}_{i\cdot}$, $v = \sum_{j=1}^{n} w_{2j}\,\bar{Y}_{j\cdot}$, and every number of treatment's sample is n. s.t, $\bar{Y}_{i\cdot} \sim N(\mu, \frac{\sigma^2}{n_i})$

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

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

$$\therefore \sum_{i=j}^{n} w_{1i}w_{2j}cov(\bar{Y}_{i\cdot}, \bar{Y}_{j\cdot}) + \sum_{i\neq j}^{n} w_{1i}w_{2j}cov(\bar{Y}_{i\cdot}, \bar{Y}_{j\cdot}) = \sum_{i=1}^{n} w_{1i}w_{2i}var(\bar{Y}_{i\cdot}) + 0 = \sigma^2 \sum_{i=1}^{n} w_{1i}w_{2i}/n_i$$

$$\because \boldsymbol{w_1}\ and\ \boldsymbol{w_2}\ are\ orthogonal$$

$\therefore \sum_{i=1}^{n} w_{1i}w_{2i}/n_i = 0,\ \sigma^2 \sum_{i=1}^{n} w_{1i}w_{2i}/n_i = 0$ $\rightarrow$ cov(u, v) = 0 $\rightarrow$ correlation of u and v is 0

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

7. 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) 1 fly vs 8 flies

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

$$\mathrm{H_0}: One\ female\ fly\ and\ eight\ flies\ reproductive\ activity\ does\ not\ affect\ longevity.$$

$$\mathrm{H_1}: One\ female\ fly\ and\ eight\ flies\ reproductive\ activity\ does\ affect\ longevity.$$

We use statistic $|t| = \left| \frac{\sum_{i=1}^{g} w_i \bar{y}_{i\cdot}}{\sqrt{MSE \sum_{i=1}^{g} w_i^2/n_i}} \right| \geq t_{N-g,\alpha/2}$ to test data.

Then, data's t_0 is 5.526 and $t_{0.025}(96)$ is 1.985.

Therefore, $|t_0| \geq t_{0.025}(96)$ means that we reject $\mathrm{H_0}$ hypothesis.

(b) Virgin vs pregnant

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

$$\mathrm{H_0}: Virgin\ fly\ and\ prgnant\ fly\ reproductive\ activity\ does\ not\ affect\ longevity.$$

$$\mathrm{H_1}: Virgin\ fly\ and\ prgnant\ fly\ reproductive\ activity\ does\ affect\ longevity.$$

We use statistic $|t| = \left| \frac{\sum_{i=1}^{g} w_i \bar{y}_{i\cdot}}{\sqrt{MSE \sum_{i=1}^{g} w_i^2/n_i}} \right| \geq t_{N-g,\alpha/2}$ to test data.

Then, data's t_0 is 2.824 and $t_{0.025}(96)$ is 1.985.

Therefore, $|t_0| \geq t_{0.025}(96)$ means that we reject $\mathrm{H_0}$ hypothesis.

(c) Non vs …

Let contrast = (1,-1/4,-1/4,-1/4,-1/4) , means

$$\mathrm{H_0}: Non\ of\ fly\ and\ fly\ reproductive\ activity\ does\ not\ affect\ longevity.$$

$$\mathrm{H_1}: Non\ of\ fly\ and\ fly\ reproductive\ activity\ does\ affect\ longevity.$$

We use statistic $|t| = \left| \frac{\sum_{i=1}^{g} w_i \bar{y}_{i\cdot}}{\sqrt{MSE \sum_{i=1}^{g} w_i^2/n_i}} \right| \geq t_{N-g,\alpha/2}$ to test data.

Then, data's t_0 is 2.23 and $t_{0.025}(120)$ is 1.9799.

Therefore, $|t_0| \geq t_{0.025}(120)$ means that we reject $\mathrm{H_0}$ hypothesis.

8. Appendix

code: http://rpubs.com/YaPi/374816

```
library(dplyr)
library(knitr)
library(kableExtra)
library(ggplot2)
# Problem 3.1
x1 <- c(5.60, 6.80, 8.32, 8.70, 7.64, 7.44, 7.48, 7.80, 7.72, 8.40, 6.98, 8.00)
x2 <- c(5.04, 7.38, 5.56, 6.96, 8.30, 6.86, 5.62, 7.22, 5.72, 6.40, 7.54, 7.50)
x3 <- c(8.36, 7.04, 6.92, 8.18, 6.20, 6.10, 2.75, 8.14, 9.00, 8.64, 6.60, 8.18)
x4 <- c(8.30, 8.54, 7.68, 8.92, 8.46, 7.38, 8.08, 8.12, 8.68, 8.24, 8.09, 8.06)
x <- c(x1, x2, x3, x4)
operators <- as.factor(rep(c("Johnson", "Gina", "Charles", "Eva"), each = 12))
y <- data.frame(ww = x, op = operators)
ggplot(y) + geom_boxplot(aes(x = op,y = ww), fill = c("#FF8888","#9999FF","#FFDD55","#99FF
33"))+
  xlab("Operators")+
  ylab("shear strengths")

lm(x ~ operators)

##
## Call:
## lm(formula = x ~ operators)
##
## Coefficients:
##     (Intercept)      operatorsEva     operatorsGina  operatorsJohnson
##          7.1758            1.0367           -0.5008            0.3975

summary(lm(x ~ operators))

##
## Call:
## lm(formula = x ~ operators)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.4258 -0.5433  0.0771  0.7173  1.8242
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)        7.1758     0.3153  22.755   <2e-16 ***
## operatorsEva       1.0367     0.4460   2.325   0.0248 *
## operatorsGina     -0.5008     0.4460  -1.123   0.2675
## operatorsJohnson   0.3975     0.4460   0.891   0.3776
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.092 on 44 degrees of freedom
```

```
## Multiple R-squared:  0.2244, Adjusted R-squared:  0.1715
## F-statistic: 4.243 on 3 and 44 DF,  p-value: 0.0102

p31 <- anova(lm(x ~ operators))
p31 %>% kable(., "html") %>%
  kable_styling(bootstrap_options = "striped", full_width = F)

# Problem 3.2
N <- c(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)
p1 <- c(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)
v1 <- c(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)
p8 <- c(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)
v8 <- c(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)
x <- c(N,p1,v1,p8,v8)
flys <- as.factor(rep(c("None", "1 pregnant", "1 virgin", "8 pregnant", "8 virgin"), each
= 25))
y <- data.frame(ww = x, op = flys)
ggplot(y) + geom_boxplot(aes(x = op,y = ww), fill = c("#FF8888","#9999FF","#FFDD55","#99FF
33","#D28EFF"))+
  xlab("Companions")+
  ylab("Longevity")

lm(x ~ flys)

##
## Call:
## lm(formula = x ~ flys)
##
## Coefficients:
##    (Intercept)    flys1 virgin  flys8 pregnant    flys8 virgin
##          63.56            1.24           -6.80          -24.84
##       flysNone
##          -0.20

summary(lm(x ~ flys))

##
## Call:
## lm(formula = x ~ flys)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -35.76  -8.76   0.20  11.20  32.44
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)      63.560      2.962  21.461  < 2e-16 ***
## flys1 virgin      1.240      4.188   0.296    0.768
## flys8 pregnant   -6.800      4.188  -1.624    0.107
## flys8 virgin    -24.840      4.188  -5.931 2.98e-08 ***
## flysNone         -0.200      4.188  -0.048    0.962
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 14.81 on 120 degrees of freedom
## Multiple R-squared:  0.3121, Adjusted R-squared:  0.2892
## F-statistic: 13.61 on 4 and 120 DF,  p-value: 3.516e-09
```

```
anova(lm(x ~ flys)) %>% kable(., "html") %>%
  kable_styling(bootstrap_options = "striped", full_width = F)

# Exercise 4.3
# Refer to the data in Problem 3.1. Workers 1 and 2 were experienced, whereas workers 3 an
d 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.
x1 <- c(5.60, 6.80, 8.32, 8.70, 7.64, 7.44, 7.48, 7.80, 7.72, 8.40, 6.98, 8.00)
x2 <- c(5.04, 7.38, 5.56, 6.96, 8.30, 6.86, 5.62, 7.22, 5.72, 6.40, 7.54, 7.50)
x3 <- c(8.36, 7.04, 6.92, 8.18, 6.20, 6.10, 2.75, 8.14, 9.00, 8.64, 6.60, 8.18)
x4 <- c(8.30, 8.54, 7.68, 8.92, 8.46, 7.38, 8.08, 8.12, 8.68, 8.24, 8.09, 8.06)
x <- c(x1, x2, x3, x4)
operators <- as.factor(rep(c("Johnson", "Gina", "Charles", "Eva"), each = 12))
boxplot(x ~ operators, ylab = "Strength of the pins", xlab = "Operators")

lm(x ~ operators)

##
## Call:
## lm(formula = x ~ operators)
##
## Coefficients:
##     (Intercept)      operatorsEva     operatorsGina  operatorsJohnson
##          7.1758            1.0367           -0.5008            0.3975

summary(lm(x ~ operators))

##
## Call:
## lm(formula = x ~ operators)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.4258 -0.5433  0.0771  0.7173  1.8242
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)        7.1758     0.3153  22.755   <2e-16 ***
## operatorsEva       1.0367     0.4460   2.325   0.0248 *
## operatorsGina     -0.5008     0.4460  -1.123   0.2675
## operatorsJohnson   0.3975     0.4460   0.891   0.3776
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.092 on 44 degrees of freedom
## Multiple R-squared:  0.2244, Adjusted R-squared:  0.1715
## F-statistic: 4.243 on 3 and 44 DF,  p-value: 0.0102

p31 <- anova(lm(x ~ operators))
w <- c(1, 1, -1, -1)
mse <- p31$`Mean Sq`[2]
x <- matrix(x,ncol=12,byrow = T) %>% apply(.,1,mean)
t0 <- (w %*% x)/sqrt(mse*sum(w^2/12))
t0
```

```
##           [,1]
## [1,] -1.807529
```

```
quantt <- qt(0.975,48-4)
quantt
```

```
## [1] 2.015368
```

```
# test
abs(t0) >= quantt
```

```
##       [,1]
## [1,] FALSE
```

```
# do not reject

# Problem 4.2
## (a) 1 v 8 female fly
p1 <- c(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)
v1 <- c(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)
p8 <- c(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)
v8 <- c(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)
x <- c(p1,v1,p8,v8)
flys <- as.factor(rep(c("1 pregnant", "1 virgin", "8 pregnant", "8 virgin"), each = 25))
boxplot(x ~ flys, ylab = "Longevity (days)", xlab = "Companions")
lm(x ~ flys)
```

```
##
## Call:
## lm(formula = x ~ flys)
##
## Coefficients:
##    (Intercept)    flys1 virgin  flys8 pregnant    flys8 virgin
##          63.56            1.24           -6.80          -24.84
```

```
summary(lm(x ~ flys))
```

```
##
## Call:
## lm(formula = x ~ flys)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -35.76  -8.77  -0.28   9.13  32.44
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)       63.560      2.975  21.366  < 2e-16 ***
## flys1 virgin       1.240      4.207   0.295    0.769
## flys8 pregnant    -6.800      4.207  -1.616    0.109
## flys8 virgin     -24.840      4.207  -5.904 5.34e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 14.87 on 96 degrees of freedom
## Multiple R-squared:  0.338,	Adjusted R-squared:  0.3173
## F-statistic: 16.34 on 3 and 96 DF,  p-value: 1.177e-08
```

```r
pfly <- anova(lm(x ~ flys))
pfly%>% kable(., "html") %>%
  kable_styling(bootstrap_options = "striped", full_width = F)

w <- c(1, 1, -1, -1)
mse <- pfly$`Mean Sq`[2]
x <- matrix(x,ncol=25,byrow = T) %>% apply(.,1,mean)
t0 <- (w %*% x)/sqrt(mse*sum(w^2/25))
t0

##          [,1]
## [1,] 5.526277

quantt <- qt(0.975,100-4)
quantt

## [1] 1.984984

# test
abs(t0) >= quantt

##      [,1]
## [1,] TRUE

# reject

## (b) virgin v pregant
p1 <- c(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)
v1 <- c(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)
p8 <- c(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)
v8 <- c(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)
x <- c(p1,v1,p8,v8)
flys <- as.factor(rep(c("1 pregnant", "1 virgin", "8 pregnant", "8 virgin"), each = 25))
boxplot(x ~ flys, ylab = "Longevity (days)", xlab = "Companions")

lm(x ~ flys)

##
## Call:
## lm(formula = x ~ flys)
##
## Coefficients:
##    (Intercept)    flys1 virgin  flys8 pregnant    flys8 virgin
##          63.56            1.24           -6.80          -24.84

summary(lm(x ~ flys))

##
## Call:
## lm(formula = x ~ flys)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -35.76  -8.77  -0.28   9.13  32.44
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)      63.560      2.975  21.366  < 2e-16 ***
```

```
## flys1 virgin      1.240      4.207   0.295    0.769
## flys8 pregnant   -6.800      4.207  -1.616    0.109
## flys8 virgin    -24.840      4.207  -5.904 5.34e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 14.87 on 96 degrees of freedom
## Multiple R-squared:  0.338,  Adjusted R-squared:  0.3173
## F-statistic: 16.34 on 3 and 96 DF,  p-value: 1.177e-08
```

```
pfly <- anova(lm(x ~ flys))
pfly%>% kable(., "html") %>%
  kable_styling(bootstrap_options = "striped", full_width = F)
```

```
w <- c(1, -1, 1, -1)
mse <- pfly$`Mean Sq`[2]
x <- matrix(x,ncol=25,byrow = T) %>% apply(.,1,mean)
t0 <- (w %*% x)/sqrt(mse*sum(w^2/25))
t0
```

```
##          [,1]
## [1,] 2.823645
```

```
quantt <- qt(0.975,100-4)
quantt
```

```
## [1] 1.984984
```

```
# test
abs(t0) >= quantt
```

```
##      [,1]
## [1,] TRUE
```

```
# reject
```

```
## (c) non v more
N <- c(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)
p1 <- c(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)
v1 <- c(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)
p8 <- c(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)
v8 <- c(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)
x <- c(N,p1,v1,p8,v8)
flys <- as.factor(rep(c("None", "1 pregnant", "1 virgin", "8 pregnant", "8 virgin"), each
= 25))
boxplot(x ~ flys, ylab = "Longevity (days)", xlab = "Companions")
```

```
lm(x ~ flys)
```

```
##
## Call:
## lm(formula = x ~ flys)
##
## Coefficients:
##    (Intercept)    flys1 virgin  flys8 pregnant    flys8 virgin
##          63.56            1.24           -6.80          -24.84
##       flysNone
##          -0.20
```

```r
summary(lm(x ~ flys))

##
## Call:
## lm(formula = x ~ flys)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -35.76  -8.76   0.20  11.20  32.44
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)      63.560      2.962  21.461  < 2e-16 ***
## flys1 virgin      1.240      4.188   0.296    0.768
## flys8 pregnant   -6.800      4.188  -1.624    0.107
## flys8 virgin    -24.840      4.188  -5.931 2.98e-08 ***
## flysNone         -0.200      4.188  -0.048    0.962
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 14.81 on 120 degrees of freedom
## Multiple R-squared:  0.3121, Adjusted R-squared:  0.2892
## F-statistic: 13.61 on 4 and 120 DF,  p-value: 3.516e-09

pfly <- anova(lm(x ~ flys))
pfly%>% kable(., "html") %>%
  kable_styling(bootstrap_options = "striped", full_width = F)

w <- c(1,-1/4,-1/4,-1/4,-1/4)
mse <- pfly$`Mean Sq`[2]
x <- matrix(x,ncol=25,byrow = T) %>% apply(.,1,mean)
t0 <- (w %*% x)/sqrt(mse*sum(w^2/25))
t0

##          [,1]
## [1,] 2.234847

quantt <- qt(0.975,125-5)
quantt

## [1] 1.97993

# test
abs(t0) >= quantt

##      [,1]
## [1,] TRUE

# reject
```