STA305 Homework 3

Rui Qiu #999292509

2016-03-20

Problem 1

(a) Solution:

```
 1 > mu_bar <- (15+14.5+12.5+15.3)/4
 2 > eff.size1 <- sqrt(((15-mu_bar)^2+(14.5-mu_bar)^2+(12.5-mu_bar)^2+(15.3-mu_bar)^2)/4/36)
 3 > eff.size2 <- sqrt(((15-mu_bar)^2+(14.5-mu_bar)^2+(12.5-mu_bar)^2+(15.3-mu_bar)^2)/4/16)
 4 > eff.size3 <- sqrt(((15-mu_bar)^2+(14.5-mu_bar)^2+(12.5-mu_bar)^2+(15.3-
   mu_bar)^2)/4/3.5^2)
 5 > eff.size4 <- sqrt(((15-mu_bar)^2+(14.5-mu_bar)^2+(12.5-mu_bar)^2+(15.3-mu_bar)^2)/4/9)
 6 > library(pwr)
 7 > pwr.anova.test(k=4,n=15,f=eff.size1,sig.level=0.05)
 8
 9      Balanced one-way analysis of variance power calculation
10
11               k = 4
12               n = 15
13               f = 0.181955
14       sig.level = 0.05
15           power = 0.1805942
16
17 NOTE: n is number in each group
18
19 > pwr.anova.test(k=4,n=15,f=eff.size2,sig.level=0.05)
20
21      Balanced one-way analysis of variance power calculation
22
23               k = 4
24               n = 15
25               f = 0.2729326
26       sig.level = 0.05
27           power = 0.3727171
28
29 NOTE: n is number in each group
30
31 > pwr.anova.test(k=4,n=15,f=eff.size3,sig.level=0.05)
32
33      Balanced one-way analysis of variance power calculation
34
35               k = 4
36               n = 15
```

```
37                    f = 0.3119229
38         sig.level = 0.05
39            power = 0.475779
40
41 NOTE: n is number in each group
42
43 > pwr.anova.test(k=4,n=15,f=eff.size4,sig.level=0.05)
44
45        Balanced one-way analysis of variance power calculation
46
47               k = 4
48               n = 15
49               f = 0.3639101
50       sig.level = 0.05
51           power = 0.6170057
52
53 NOTE: n is number in each group
```

The effective sizes assuming 4 different within group variances are:

- $f_1 = 0.181955$,
- $f_2 = 0.2729326$,
- $f_3 = 0.3119229$,
- $f_4 = 0.3639101$.

The power to detect each effect size at the 5% level, are

- power 1 $= 0.1805942$,
- power 2 $= 0.3727171$,
- power 3 $= 0.475779$,
- power 4 $= 0.6170057$.

(b) Solution

```
 1 > NSIM <- 10000
 2 > res <- numeric(NSIM)
 3 > mu1 <- 15; mu2 <- 14.5; mu3 <- 12.5; mu4 <- 15.3
 4 > sigma1 <- 6; sigma2 <- 4; sigma3 <- 3.5; sigma4 <- 3
 5 > n1 <- 15; n2 <- 15; n3 <- 15; n4 <- 15
 6
 7 > for (i in 1:NSIM) {
 8     y1 <- rnorm(n=n1,mean=mu1,sd=sigma1);
 9     y2 <- rnorm(n=n2,mean=mu2,sd=sigma2);
10     y3 <- rnorm(n=n3,mean=mu3,sd=sigma3);
11     y4 <- rnorm(n=n4,mean=mu4,sd=sigma4);
12     y <- c(y1,y2,y3,y4);
13     trt <- as.factor(c(rep(1,n1),rep(2,n2),rep(3,n3),rep(4,n4)));
14     m <- lm(y~trt);
15     res[i] <- anova(m)[1,5]
16 }
17 > sum(res<=0.05)/NSIM
18 [1] 0.3273
```

The power of the study using 15 subjects per group assuming that the standard deviations for the four methods are not equal is 0.3273.

(c) Solution:

By comparison between this power 0.3273 and powers calculated in part(a), it is lower than 3 out of 4 part(a) powers.

We know that the power stands for the probability of rejecting H_0 when H_0 is false, so higher power is always welcomed.

Hence we would like to conclude that the assumption of a common within group variance in calculating power for an ANOVA experiment is not a good choice.

Problem 2

(a) Solution:

```
 1 > q2data <- read.csv("q2data.csv")
 2 > treatment <- as.factor(q2data$trt)
 3 > y <- q2data$y
 4 > sapply(split(y,treatment),mean)
 5         1        2        3        4
 6  2.206182 2.290470 2.320007 2.855205
 7 > sapply(split(y,treatment),sd)
 8         1        2        3        4
 9  1.799560 1.774576 1.812561 1.763111
10 > boxplot(y~treatment)
```

$\mu_1 = 2.206182, \mu_2 = 2.290470, \mu_3 = 2.320007, \mu_4 = 2.855205$
$\sigma_1 = 1.799560, \sigma_2 = 1.774576, \sigma_3 = 1.812561, \sigma_4 = 1.763111$

And the boxplot indicates that the 4 distributions look pretty similar.
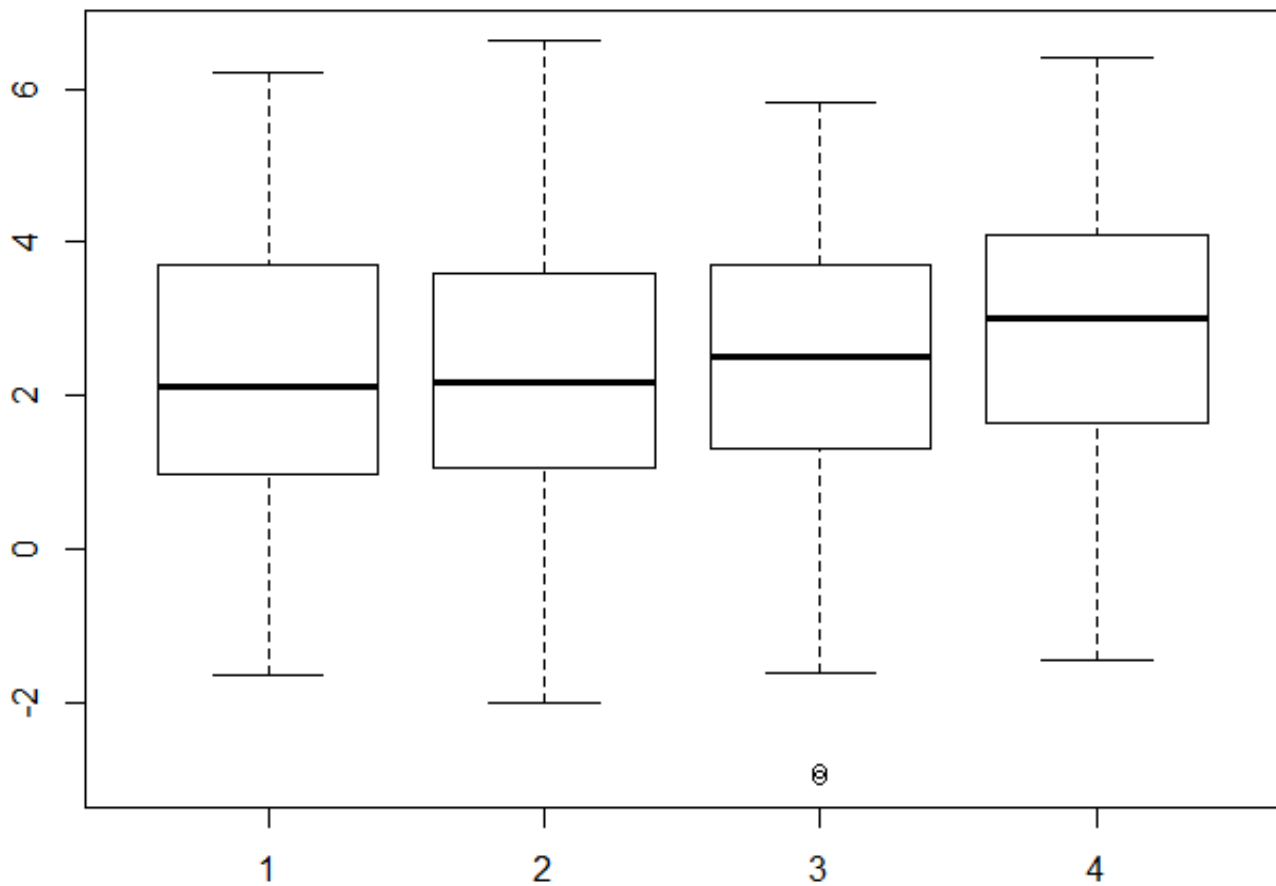
(b) Solution:

```
1 > anova(lm(y~treatment))
2 Analysis of Variance Table
3
4 Response: y
5           Df  Sum Sq Mean Sq F value  Pr(>F)
6 treatment   3   26.29  8.7639  2.7429 0.04294 *
7 Residuals 391 1249.31  3.1952
8 ---
9 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

As p-value is less than 0.05, reject null hypothesis, i.e. means are different.

(c) Solution:

```
1 > contrasts(treatment) <-contr.treatment(4, base = 3)
2 > contrasts(treatment)
3    1 2 4
4 1 1 0 0
5 2 0 1 0
6 3 0 0 0
7 4 0 0 1
```

Model: $y_{ij} = \mu + \tau_1 X_1 + \tau_2 X_2 + \tau_4 X_4 + \epsilon_{ij}$, where dummy variables in this case are:

$$X_1 = \begin{cases} 1, \text{ if trt 1,} \\ 0, \text{ otherwise.} \end{cases}$$

$$X_2 = \begin{cases} 1, \text{ if trt 2,} \\ 0, \text{ otherwise.} \end{cases}$$

$$X_4 = \begin{cases} 1, \text{ if trt 4,} \\ 0, \text{ otherwise.} \end{cases}$$

If $X_1 = X_2 = X_4 = 0$, then the treatment is trt3, which is the control treatment.

```
1 > summary(lm(y~treatment),data=q2data)
2
3 Call:
4 lm(formula = y ~ treatment)
5
6 Residuals:
7    Min     1Q Median     3Q    Max
8 -5.295 -1.180 -0.048  1.391  4.336
9
10 Coefficients:
11            Estimate Std. Error t value Pr(>|t|)
12 (Intercept)  2.32001    0.18056  12.849   <2e-16 ***
13 treatment1  -0.11382    0.25408  -0.448   0.6544
14 treatment2  -0.02954    0.25668  -0.115   0.9084
15 treatment4   0.53520    0.25345   2.112   0.0354 *
16 ---
17 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
18
19 Residual standard error: 1.788 on 391 degrees of freedom
20 Multiple R-squared:  0.02061,   Adjusted R-squared:  0.0131
21 F-statistic: 2.743 on 3 and 391 DF,  p-value: 0.04294
```

The design matrix can be shown by `model.matrix(lm(y~treatment))`.

$\bar{y}_{3.} = 2.32001$
$\hat{\tau}_1 = \bar{y}_{1.} - \bar{y}_{3.} = -0.11382$
$\hat{\tau}_2 = \bar{y}_{2.} - \bar{y}_{3.} = -0.02954$
$\hat{\tau}_4 = \bar{y}_{4.} - \bar{y}_{3.} = 0.53520$

So the means are:
$\bar{y}_{1.} = 2.32001 - 0.11382 = 2.20619$

$\bar{y}_{2.} = 2.32001 - 0.02954 = 2.29047$
$\bar{y}_{3.} = 2.32001$
$\bar{y}_{4.} = 2.32001 + 0.53520 = 2.85521$
Thus, they verified the results from part(a).

(d) Solution:

```
1 > contrasts(treatment) <- contr.helmert(4) # no argument in `contr.helmert()` to set up bas
  eline
2 > contrasts(treatment)
3    [,1] [,2] [,3]
4 1   -1   -1   -1
5 2    1   -1   -1
6 3    0    2   -1
7 4    0    0    3
```

Model: $y_{ij} = \mu + \tau_1 X_{1j} + \tau_2 X_{2j} + \tau_3 X_{3j0} + \epsilon_{ij}$ with dummy variables:

$$X_{1j} = \begin{cases} -1, \text{ if jth unit recieves trt 1,} \\ 1, \text{ if trt 2,} \\ 0, \text{ if trt 3 or trt 4.} \end{cases}$$

$$X_{2j} = \begin{cases} -1, \text{ if trt1 or trt 2,} \\ 2, \text{ if trt 3,} \\ 0, \text{ if trt 4.} \end{cases}$$

$$X_{3j} = \begin{cases} -1, \text{ if trt 1 or trt 2 or trt 3,} \\ 3, \text{ if trt 4.} \end{cases}$$

```
1 > summary(lm(y~treatment),data=q2data)
2
3 Call:
4 lm(formula = y ~ treatment)
5
6 Residuals:
7    Min      1Q Median     3Q    Max
8 -5.295 -1.180 -0.048  1.391  4.336
9
10 Coefficients:
11           Estimate Std. Error t value Pr(>|t|)
12 (Intercept)  2.41797    0.08996   26.879   < 2e-16 ***
13 treatment1   0.04214    0.12771    0.330   0.74157
14 treatment2   0.02389    0.07372    0.324   0.74603
15 treatment3   0.14575    0.05154    2.828   0.00493 **
16 ---
17 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
18
19 Residual standard error: 1.788 on 391 degrees of freedom
20 Multiple R-squared:  0.02061,   Adjusted R-squared:  0.0131
21 F-statistic: 2.743 on 3 and 391 DF,  p-value: 0.04294
```

Means of each treatment are:
$$\bar{y}_{1.} = \hat{\mu} - \hat{\tau}_1 - \hat{\tau}_2 - \hat{\tau}_3 = 2.41797 - 0.04214 - 0.02389 - 0.14575 = 2.20619,$$
$$\bar{y}_{2.} = \hat{\mu} + \hat{\tau}_1 - \hat{\tau}_2 - \hat{\tau}_3 = 2.41797 + 0.04214 - 0.02389 - 0.14575 = 2.29047,$$
$$\bar{y}_{3.} = \hat{\mu} + 2\hat{\tau}_2 - \hat{\tau}_3 = 2.41797 + 2 \times 0.02389 - 0.14575 = 2.32,$$
$$\bar{y}_{4.} = \hat{\mu} + 3\hat{\tau}_3 = 2.41797 + 3 \times 0.14575 = 2.85522$$
This set of means is also very consistent with the result of part (a).

(e) Solution:
The Helmert contrast makes more sense because it has more "involvement" of the different treatments. Not like the original treatment contrast (dummy coding), each linear regression model only contains one treatment. On the other hand, this type of "simplification" could possibly ignore some significant difference between treatment and placebo. But the Helmert contrast could contain multiple treatment in the model, so it can detect significant difference more easily.

(f) Solution:

- unadjusted:

```
1  > pairwise.t.test(q2data$y, q2data$trt, p.adjust.method = "none")
2
3      Pairwise comparisons using t tests with pooled SD
4
5  data:  q2data$y and q2data$trt
6
7    1     2     3
8  2 0.742 -     -
9  3 0.654 0.908 -
10 4 0.010 0.027 0.035
11
12 P value adjustment method: none
```

There are significant differences at the 5% level between treatment 1 & 4, 2 & 4, 3 & 4.

- Bonferroni:

```
1  > pairwise.t.test(q2data$y, q2data$trt, p.adjust.method = "bonferroni")
2
3      Pairwise comparisons using t tests with pooled SD
4
5  data:  q2data$y and q2data$trt
6
7    1     2     3
8  2 1.000 -     -
9  3 1.000 1.000 -
10 4 0.063 0.163 0.212
11
12 P value adjustment method: bonferroni
```

There are significant differences at the 5% level between none of 2 treatments.

- Tukey:

```
 1 > library(multcomp)
 2 > TukeyHSD(aov(y~treatment,data=q2data))
 3   Tukey multiple comparisons of means
 4     95% family-wise confidence level
 5
 6 Fit: aov(formula = y ~ treatment, data = q2data)
 7
 8 $treatment
 9          diff          lwr       upr     p adj
10 2-1 0.08428777 -0.574699935 0.7432755 0.9875857
11 3-1 0.11382450 -0.541723430 0.7693724 0.9699886
12 4-1 0.64902293 -0.001589317 1.2996352 0.0508310
13 3-2 0.02953673 -0.632736356 0.6918098 0.9994544
14 4-2 0.56473516 -0.092652737 1.2221231 0.1205489
15 4-3 0.53519843 -0.118741276 1.1891381 0.1511693
```

Again, none of two pairs have significant differences.

Thus, the 3 methods have different results. (In fact, Bonferroni and Tukey have the same result.)
Why?

- The family-wise type I error rate of unadjusted at 5% is $1 - (1 - 0.05)^6 = 0.2649081$, which is greater than 0.05.
- Actually, testing all possible pairs increase the type I error rate. This means the chance that there is a higher probability that a significant difference is detected when the truth is that no difference exists.
- For Bonferroni and Tukey's HSD, family-wise type I error rates are controlled.

So different results occured.
It is not necessary to consider all pairs of treatment means, since the objective of our study is to compare new treatments to the control treatment.
So we only need to compare treatment **1 & 3, 2 & 3, 4 & 3.**

Problem 3

(a) Solution

```
 1 > q3data <- read.csv("q3data.csv")
 2 > q3data
 3    run x1 x2 x3  y
 4 1    1 -1 -1 -1 59
 5 2    2  1 -1 -1 60
 6 3    3 -1  1 -1 63
 7 4    4  1  1 -1 62
 8 5    5 -1 -1  1 63
 9 6    6  1 -1  1 64
10 7    7 -1  1  1 54
11 8    8  1  1  1 59
12 9    9 -1 -1 -1 55
13 10  10  1 -1 -1 56
14 11  11 -1  1 -1 61
15 12  12  1  1 -1 58
16 13  13 -1 -1  1 63
17 14  14  1 -1  1 53
18 15  15 -1  1  1 56
19 16  16  1  1  1 56
```

Yes, because of two points:

- The design has a fixed number of levels of each factor.
- Also, it is run in all possible combinations which is 8 times and repeated twice, so 16 runs in total.

So it is a 2^3 factorial experiment, where $k = 3$.
The experiment has three quantitative factors, temperature(x_1), pH(x_2) and agitation rate(x_3).
Each of variables has two levels, coded as -1 and +1 respectively.

(b) Solution:

```
 1 > fact.mod <- lm(y~x1*x2*x3,data=q3data)
 2 > fact.mod
 3
 4 Call:
 5 lm(formula = y ~ x1 * x2 * x3, data = q3data)
 6
 7 Coefficients:
 8 (Intercept)           x1             x2             x3          x1:x2
 9      58.875        -0.375         -0.250         -0.375          0.500
10       x1:x3         x2:x3       x1:x2:x3
```

```
11      -0.125           -2.000          1.250
12
13 > round(summary(fact.mod)$coefficients,2)
14           Estimate Std. Error t value Pr(>|t|)
15 (Intercept)    58.88       0.85   69.07     0.00
16 x1             -0.37       0.85   -0.44     0.67
17 x2             -0.25       0.85   -0.29     0.78
18 x3             -0.37       0.85   -0.44     0.67
19 x1:x2           0.50       0.85    0.59     0.57
20 x1:x3          -0.12       0.85   -0.15     0.89
21 x2:x3          -2.00       0.85   -2.35     0.05
22 x1:x2:x3        1.25       0.85    1.47     0.18
23
24 > round(2*fact.mod$coefficients,2)
25 (Intercept)          x1          x2          x3       x1:x2       x1:x3       x2:x3      x1:x
   2:x3
26      117.75       -0.75       -0.50       -0.75        1.00       -0.25       -4.00
   2.50
27
28 > 2*confint.lm(fact.mod)
29                      2.5 %         97.5 %
30 (Intercept) 113.818787 121.68121292
31 x1            -4.681213   3.18121292
32 x2            -4.431213   3.43121292
33 x3            -4.681213   3.18121292
34 x1:x2         -2.931213   4.93121292
35 x1:x3         -4.181213   3.68121292
36 x2:x3         -7.931213  -0.06878708
37 x1:x2:x3      -1.431213   6.43121292
```

A linear model for a 2^3 factorial design is:
$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i1} x_{i2} + \beta_5 x_{i1} x_{i3} + \beta_6 x_{i2} x_{i3} + \beta_7 x_{i1} x_{i2} x_{i3} + \epsilon_i$$

Note that the estimated least squares coefficients are one-half the factorial estimates, and the intercept β_0 is the sample mean. Therefore, the factorial estimates are twice the least squares coefficients.
The standard error for all factorial effects is the same, which is can be calculated as below:

The pooled estimate of σ^2 : $s^2 = \frac{\sum \text{diff}_i^2/2}{8} = \frac{93}{8} = 11.625$

$\text{Var(effect)} = (\frac{1}{8} + \frac{1}{8})s^2 = 2.90625$

$\text{sd(effect)} = \sqrt{\text{Var(effect)}} = 1.705$

The factorial effect of x_1 is -0.75,
the factorial effect of x_2 is -0.50,
the factorial effect of x_3 is -0.75,
the interaction effect of $x_1 : x_2$ is 1.00,
the interaction effect of $x_1 : x_3$ is -0.25,
the interaction effect of $x_2 : x_3$ is -4.00,
the interaction effect of $x_1 : x_2 : x_3$ is 2.50.

(c) Solution:

```
1 > library("FrF2")
2 > cubePlot(fact.mod,"x1","x2","x3",main="Cube Plot for q3data")
3 > interaction.plot(q3data$x1,q3data$x2,q3data$y,type="l",xlab="temperature",trace.label="ag
  itation rate",ylab="pH")
4 > interaction.plot(q3data$x1,q3data$x3,q3data$y,type="l",xlab="temperature",trace.label="pH
  ",ylab="agitation rate")
5 > interaction.plot(q3data$x3,q3data$x2,q3data$y,type="l",xlab="agitation rate",trace.label=
  "temperature",ylab="pH")
```
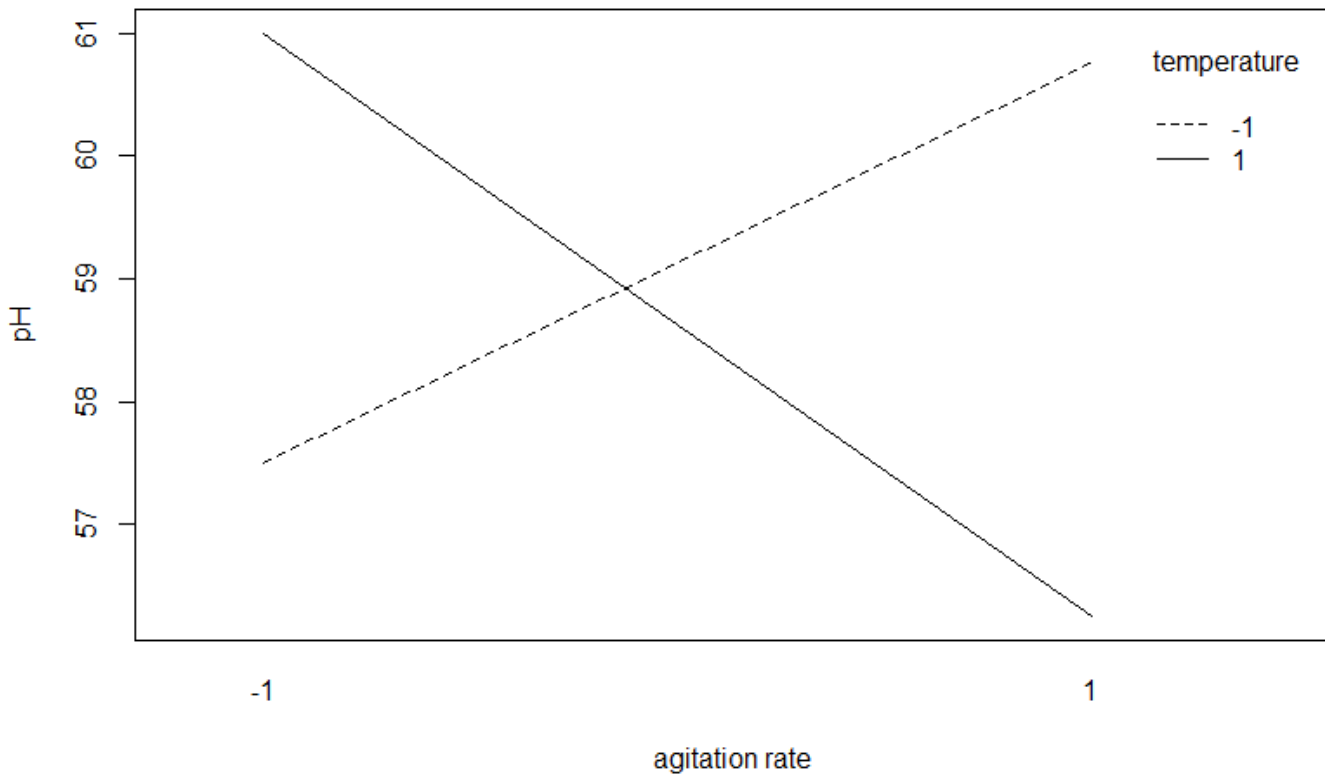
Cube Plot for q3data
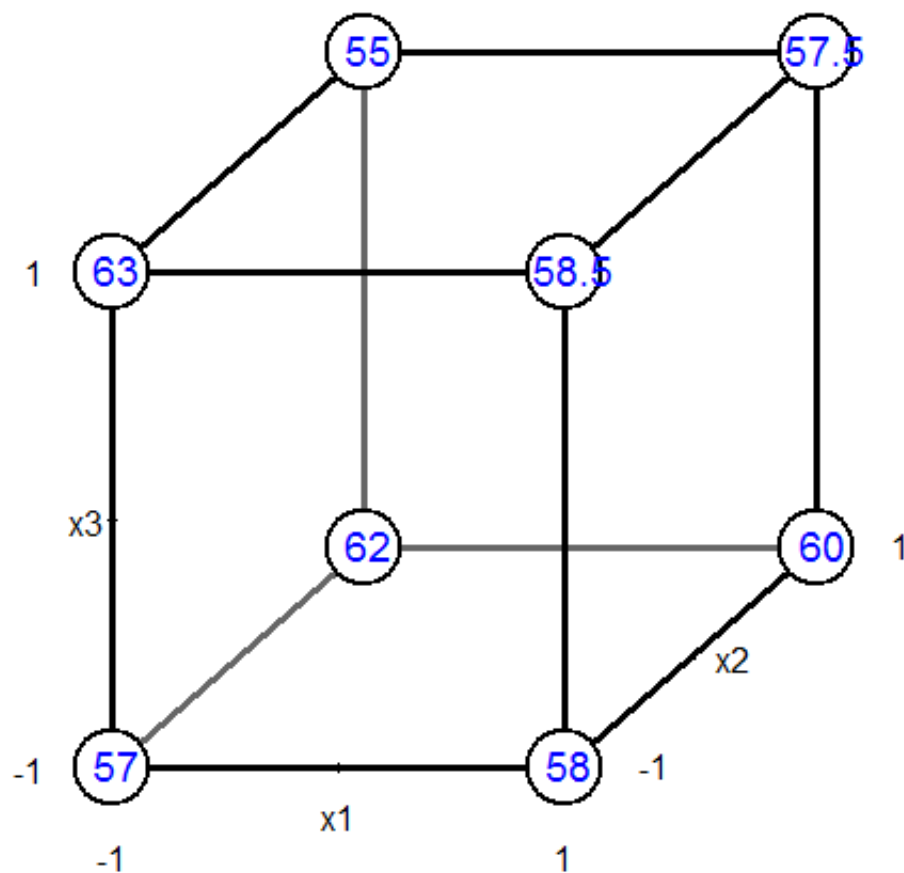
modeled = TRUE
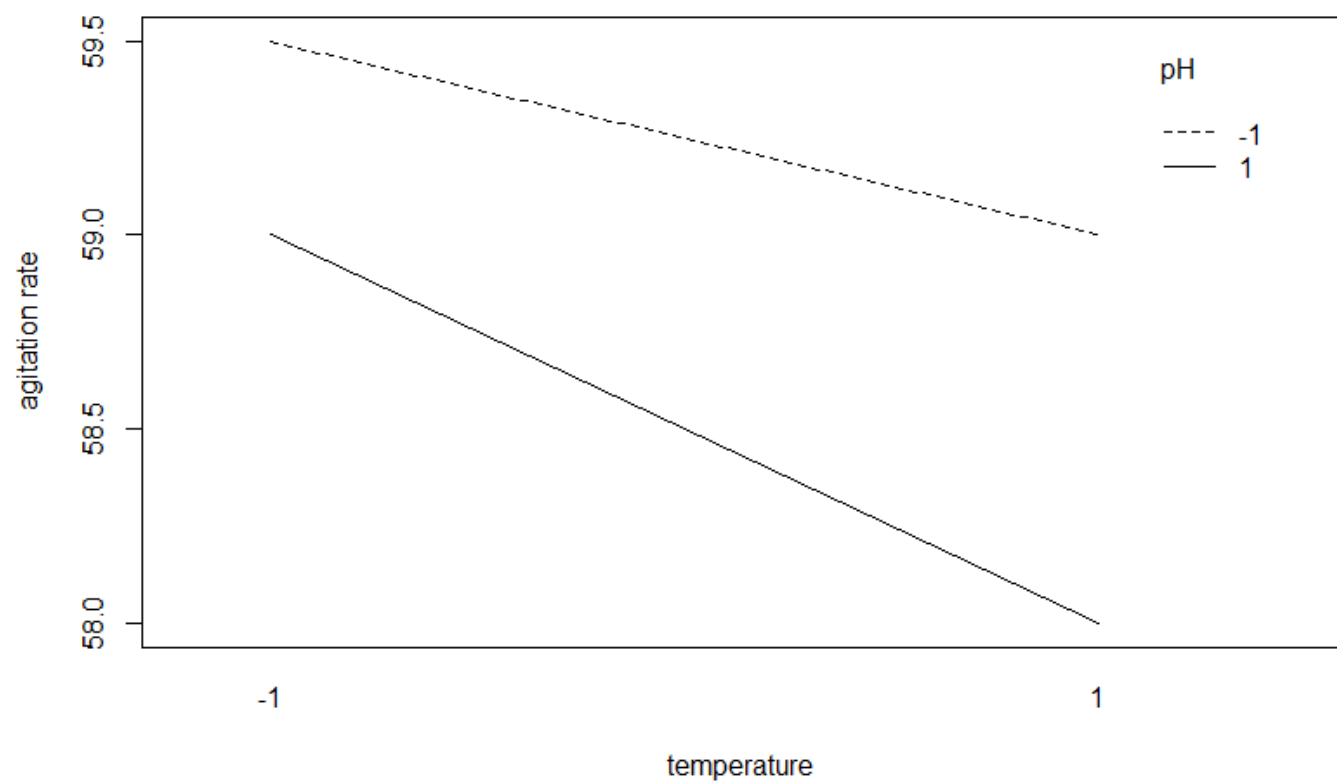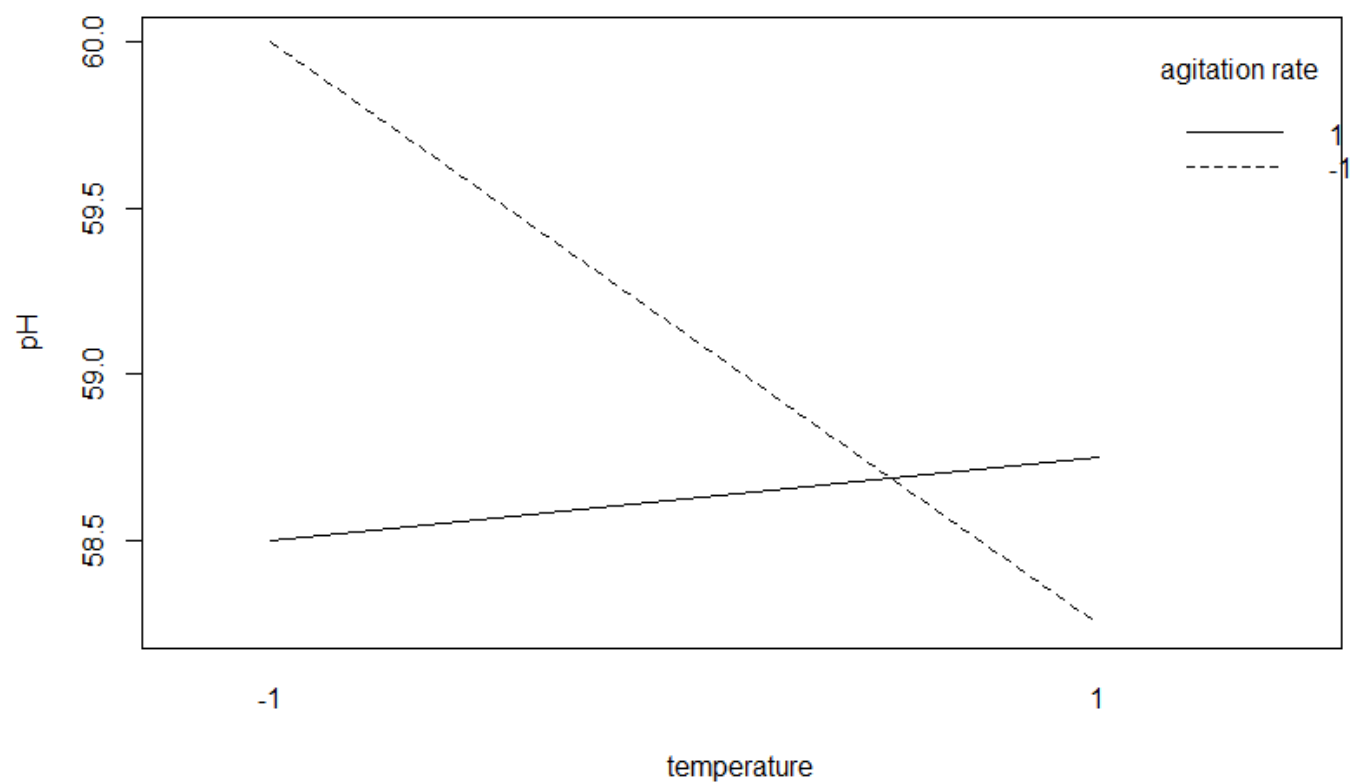
- The main effect for x_1 is interpreted as the change in yield of a chemical reaction after temperature change from level -1 to level 1.
- The main effect for x_2 is interpreted as the change in yield of a chemical reaction after pH change from level -1 to level 1.
- The main effect for x_3 is interpreted as the change in yield of a chemical reaction after agitation rate change from level -1 to level 1.
- The interaction effect for x_1 and x_2 is interpreted as the difference in the change in yield of a chemical reaction for 2 levels of temperature for 2 levels of pH.
- The interaction effect for x_1 and x_3 is interpreted as the difference in the change in yield of a chemical reaction for 2 levels of temperature for 2 levels of agitation rate.
- The interaction effect for x_2 and x_3 is interpreted as the difference in the change in yield of a chemical reaction for 2 levels of pH for 2 levels of agitation rate.
- The interaction effect for x_1 and x_2 and x_3 is interpreted as the difference in the change in yield of a chemical reaction for 2 levels of temperature for 2 levels of pH and for 2 levels of agitation rate.

The interaction plot for x_2 and x_3 indicates that there is an interaction, also the p-value is rounded to 0.05, but its 95% CI does not contain a zero.
The rest interaction plots are consistent with the hypothesis tests.
So, we reject null hypothesis, and we draw the conclusion that **there is significant interaction effect between pH and agitation rate.**

Problem 4

(a) Solution:

Let the 4 observations for treatment $a_1 b_1$ be x_1, x_2, x_3, x_4, we claim the mean to be $\bar{x} = \frac{1}{4} \sum_{i=1}^{4} x_i$.
Similarly, the 4 observations for treatment $a_2 b_1$ are y_1, y_2, y_3, y_4, the mean is $\bar{y} = \frac{1}{4} \sum_{i=4}^{4} y_i$.
Since the variance of output is $\sigma^2 = \mathrm{Var}(x) = \mathrm{Var}(y)$,

$$
\begin{aligned}
\mathrm{Var}(\bar{x} - \bar{y}) &= \mathrm{Var}(\bar{x}) + \mathrm{Var}(\bar{y}) \\
&= \frac{1}{16} \cdot 4\sigma^2 + \frac{1}{16} \cdot 4\sigma^2 \\
&= \frac{1}{2} \sigma^2
\end{aligned}
$$

So the variance of the mean difference of the effect of flour is $\frac{1}{2} \sigma^2$.

(b) Solution:
This is a 2^2 factorial desig with A, B replicated twice.
Let y_{i1} be the first outcome from ith run, then $\mathrm{diff}_i = (y_{i1} - y_{i2})$, $s_i^2 = \frac{(y_{i1} - y_{i2})^2}{2}$.
So the pooled estimate of σ^2 is

$$
s^2 = \frac{\sum_{i=1}^{4} s_i^2}{4} = \frac{\sum_{i=1}^{4} \mathrm{diff}_i^2}{8}
$$

$$
\mathrm{Var}(\mathrm{effect}) = \mathrm{Var}((\frac{1}{4} + \frac{1}{4})s^2) = \mathrm{Var}(\frac{\sum_{i=1}^{4} \mathrm{diff}_i^2}{16}) = \frac{4}{16} \mathrm{Var}(\mathrm{diff}^2)
$$

$$
= \frac{1}{4} (\mathrm{Var}(y_{i1}^2) + \mathrm{Var}(y_{i2}^2) + 2\mathrm{Cov}(y_{i1}, y_{i2})) = \frac{1}{4} (\frac{\sigma^2}{2} + \frac{\sigma^2}{2} + 2 \cdot \frac{\sigma^2}{2}) = \frac{1}{2} \sigma^2.
$$

(c) Solution:
The experiment in (b) would be a better choice since it is a factorial design with two levels of flour and two levels of flavouring (high and low), so that we can observe the effect of interactions in order to know how the effect of A changes when we change B from low to high level.

We can fit a linear regression model to find the estimated coefficients of effects. And we multiple the result by 2 to get the main and interaction effects. And among all, we are mostly interested in the interaction effect for a and b. We want to check the p-value of such interaction effect in hypothesis test:

- if the p-value is small enough, then reject null hypothesis, there is interaction between A and B, so the effect of flour is different when the amount flavouring is changed.
- if the p-value is large, then fail to rejct, there is no interaction between A and B, so the effect of flour is the same when the amount flavoring is changed.