

STA305/1004 - Homework #3 - Solutions

Due Date: March 22, 2016

Due date: Electronic submission on UofT Learning Portal course page by Tuesday, March 22, 2016 at 22:00. NB: e-mail submissions will NOT be accepted.

If you work with other students on this assignment then:

- indicate the names of the students on your solutions;
- your solutions must be written up independently (i.e., your solutions should not be the same as another students solutions).

1. A psychologist is designing an experiment to investigate the effects of four different learning methods on short term memory. Subjects will be shown a series of 20 words after undergoing some training in the learning method that they were assigned. The outcome of the experiment is the total number of words that a subject is able to recall after being trained in one of the learning methods. An equal number of subjects will be randomly assigned to each learning method.

Based on previous research the psychologist estimates that the mean and standard deviation for each method are:

Learning Method	Mean	Standard deviation
1	15	6
2	14.5	4
3	12.5	3.5
4	15.3	3

The psychologist would like to know how many subjects she will require so that her study has 80% power at the 5% significance level.

- (a) Use R to calculate the effect sizes that the psychologist can detect if she uses the different variances of the different learning methods. The formula for effect size is

$$f = \sqrt{\frac{\sum_{i=1}^k (\mu_i - \bar{\mu})^2 / k}{\sigma^2}}.$$

$\bar{\mu} = \sum_{i=1}^k \mu_i / k$, and σ^2 is the within group error variance.

For example, for the first effect size assume that the within group variance is 36, for the second effect size assume that the within group variance is 16, and so on. Now, assuming that she can enrol 15 subjects per group, what is the power to detect each effect size at the 5% level? Use `pwr.anova.test()` to calculate power. (Hand in your R code and output)

The power is calculated below for when $\sigma = 6, 4, 3.5, 3$.

```
library(pwr)
mu1 <- 15; mu2 <- 14.5; mu3 <- 12.5; mu4 <- 15.3
sigma1 <- 6; sigma2 <- 4; sigma3 <- 3.5; sigma4 <- 3;
```

```

mug <- sum(mu1,mu2,mu3,mu4)/4
mui <- c(mu1,mu2,mu3,mu4)

f1 <- sqrt(sum((mui-mug)^2)/4)/sigma1
pwr.anova.test(k = 4,f = f1,n=15,sig.level = 0.05)

##
##      Balanced one-way analysis of variance power calculation
##
##              k = 4
##              n = 15
##              f = 0.181955
##      sig.level = 0.05
##      power = 0.1805942
##
## NOTE: n is number in each group

f2 <- sqrt(sum((mui-mug)^2)/4)/sigma2
pwr.anova.test(k = 4,f = f2,n=15,sig.level = 0.05)

##
##      Balanced one-way analysis of variance power calculation
##
##              k = 4
##              n = 15
##              f = 0.2729326
##      sig.level = 0.05
##      power = 0.3727171
##
## NOTE: n is number in each group

f3 <- sqrt(sum((mui-mug)^2)/4)/sigma3
pwr.anova.test(k = 4,f = f3,n=15,sig.level = 0.05)

##
##      Balanced one-way analysis of variance power calculation
##
##              k = 4
##              n = 15
##              f = 0.3119229
##      sig.level = 0.05
##      power = 0.475779
##
## NOTE: n is number in each group

f4 <- sqrt(sum((mui-mug)^2)/4)/sigma4
pwr.anova.test(k = 4,f = f4,n=15,sig.level = 0.05)

##
##      Balanced one-way analysis of variance power calculation
##

```

```
##          k = 4
##          n = 15
##          f = 0.3639101
##      sig.level = 0.05
##          power = 0.6170057
##
## NOTE: n is number in each group
```

- (b) Use simulation to calculate the power of the study using 15 subjects per group assuming that the standard deviations for the four methods are not equal, but are as shown in the table above, and that the distribution of observations in each group is normal. A random sample of size n from a $N(\mu, \sigma^2)$ can be generated in R using the function `rnorm(n,mu,sigma)`. (Hand in your R output and R code)

The power is approximately 33% (results will vary).

```
NSIM <- 10000
res <- numeric(NSIM)

mu1 <- 15; mu2 <- 14.5; mu3 <- 12.5; mu4 <- 15.3
sigma1 <- 6; sigma2 <- 4; sigma3 <- 3.5; sigma4 <- 3;
n <- 15

for (i in 1:NSIM){

y1 <- rnorm(n,mu1,sigma1)
y2 <- rnorm(n,mu2,sigma2)
y3 <- rnorm(n,mu3,sigma3)
y4 <- rnorm(n,mu4,sigma4)

y <- c(y1,y2,y3,y4)
trt <- as.factor(c(rep(1,n),rep(2,n),rep(3,n),rep(4,n)))
m <- lm(y~trt)
res[i] <- anova(m)[1,5] # p-value of F test
}

sum(res<=0.05)/NSIM
```

- (c) What does part (b) tell you about the assumption of a common within group variance in calculating power for an ANOVA experiment? Explain.

The power in part (b) depends on the within group σ . Smaller values correspond to larger power. But none of the power values are close to the power when the groups have different within group variances. So, if the within group variances are not approximately equal then the power will not be accurate.

2. A clinical trial was conducted where patients were randomized to four different treatments. The data is available in the file `q2data.csv`. The outcome is a continuous response y_{ij} the response for the i th subject in the j th treatment group. There are three new treatments in this study and one control treatment. The control treatment is the third treatment ($j = 3$). The main objective of the study is to compare the three new treatments to the control treatment.

NB: The file can be read into R and put into a data.frame using the command

```
q2data <- read.csv("q2data.csv")
```

In this question use the 5% significance level.

- (a) What are the averages and standard deviations of each treatment? Plot the distributions of the four treatment groups. Do the distributions look similar or different? (Hand in your R code and output)

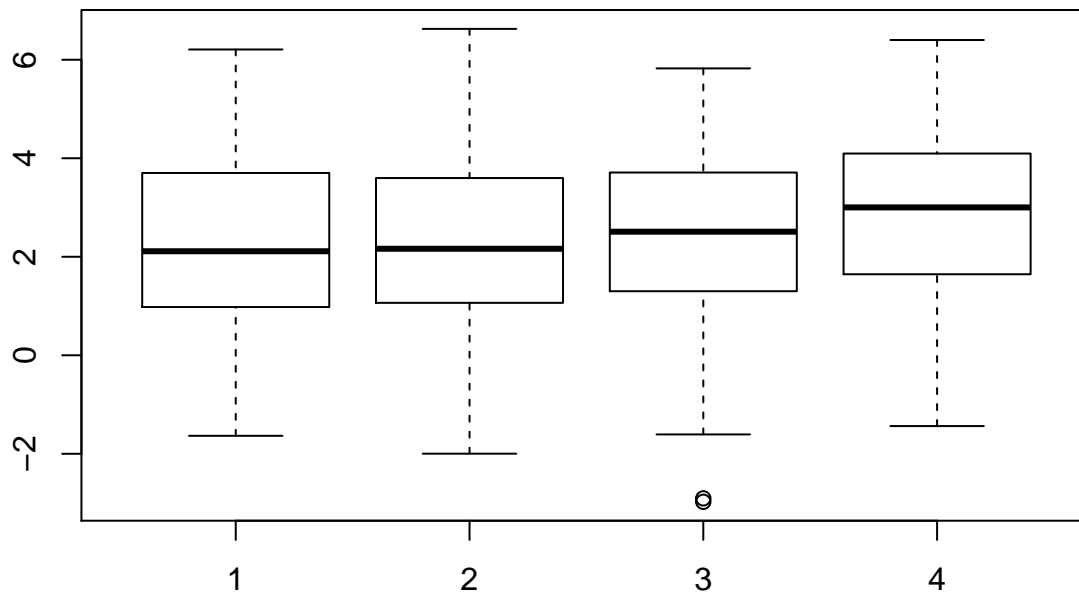
```
q2data <- read.csv("q2data.csv")
sapply(split(q2data$y,q2data$trt),mean) # treatment averages
```

```
##          1          2          3          4
## 2.206182 2.290470 2.320007 2.855205
```

```
sapply(split(q2data$y,q2data$trt),sd) # treatment SDs
```

```
##          1          2          3          4
## 1.799560 1.774576 1.812561 1.763111
```

```
boxplot(y~trt,data=q2data) # plot of distributions - other plots could also be used
```



- (b) Use linear regression to calculate the ANOVA table. What do you conclude from the ANOVA table? (NB: when using linear regression to calculate the effects the treatment variable should be specified as a factor as `as.factor(trt)`.) (Hand in your R code and output)

The ANOVA table indicates that we would reject $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ at the 5% level.

```
anova(lm(y~as.factor(trt),data = q2data))
```

```
## Analysis of Variance Table
##
## Response: y
##          Df Sum Sq Mean Sq F value Pr(>F)
```

```
## as.factor(trt)    3    26.29  8.7639  2.7429 0.04294 *
## Residuals        391 1249.31  3.1952
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- (c) Use the model you obtained in part (b) to obtain the appropriate parameter estimates using the treatment contrast (dummy coding) to answer the main objective. In R this can be done using the `contr.treatment()` function. Define the underlying statistical model in terms of dummy variables. Explicitly state the dummy variables. Interpret the parameter estimates. Verify the parameter estimates using the table of means that you obtained in part (a). (Hand in your R code and output)

The statistical model is

y_{ij} is the j^{th} observation under the i^{th} treatment. Let μ be the overall mean. The one-way model $y_{ij} = \mu + \tau_i + \epsilon_{ij}$, $\epsilon_{ij} \sim N(0, \sigma^2)$ can be written in terms of the dummy variables X_1, X_2, X_3 as:

$$y_{ij} = \mu + \tau_1 X_{i1} + \tau_2 X_{i2} + \tau_3 X_{i3} + \epsilon_{ij},$$

where,

$$X_{1j} = \begin{cases} 1 & \text{if } j\text{th unit receives treatment 1} \\ 0 & \text{otherwise} \end{cases}$$

$$X_{2j} = \begin{cases} 1 & \text{if } j\text{th unit receives treatment 2} \\ 0 & \text{otherwise} \end{cases}$$

$$X_{3j} = \begin{cases} 1 & \text{if } j\text{th unit receives treatment 4} \\ 0 & \text{otherwise} \end{cases}$$

Let $\mu_1, \mu_2, \mu_3, \mu_4$ be the treatment means. It follows that $E(y_{3j}) = \mu_3$ is the mean of the control treatment (treatment 3) so

$$E(y_{1j}) = \mu_1 = \mu_3 + \tau_1$$

$$E(y_{2j}) = \mu_2 = \mu_3 + \tau_2$$

$$E(y_{4j}) = \mu_4 = \mu_3 + \tau_3$$

The least squares estimates are:

$$\hat{\mu} = \bar{y}_{1.},$$

$$\hat{\tau}_1 = \bar{y}_{1.} - \bar{y}_{3.},$$

$$\hat{\tau}_2 = \bar{y}_{2.} - \bar{y}_{3.},$$

$$\hat{\tau}_3 = \bar{y}_{4.} - \bar{y}_{3.}.$$

```
q2data$trt <- as.factor(q2data$trt)
contrasts(q2data$trt) <- contr.treatment(n = 4, base = 3)
summary(lm(y~trt, data = q2data))
```

```
##
## Call:
## lm(formula = y ~ trt, data = q2data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.295 -1.180 -0.048  1.391  4.336
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.32001    0.18056  12.849  <2e-16 ***
## trt1         -0.11382    0.25408  -0.448   0.6544
## trt2         -0.02954    0.25668  -0.115   0.9084
## trt4          0.53520    0.25345   2.112   0.0354 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.788 on 391 degrees of freedom
## Multiple R-squared:  0.02061,    Adjusted R-squared:  0.0131
## F-statistic: 2.743 on 3 and 391 DF,  p-value: 0.04294
```

- (d) Obtain the parameter estimates using the Helmert contrast. In R this can be done using the `contr.helmert(4)` function. Explicitly state the dummy variables. Define the underlying statistical model in terms of dummy variables. Interpret the parameter estimates. Verify the parameter estimates using the table of means that you obtained in part (a). (Hand in your R code and output)

The Helmert contrast is:

```
contr.helmert(4)
```

```
##      [,1] [,2] [,3]
## 1     -1    -1    -1
## 2      1    -1    -1
## 3      0     2    -1
## 4      0     0     3
```

$$X_{1j} = \begin{cases} -1 & \text{if } j\text{th unit receives treatment 1} \\ 1 & \text{if } j\text{th unit receives treatment 2} \\ 0 & \text{if } j\text{th unit receives treatment 3 or 4} \end{cases}$$

$$X_{2j} = \begin{cases} -1 & \text{if } j\text{th unit receives treatment 1} \\ -1 & \text{if } j\text{th unit receives treatment 2} \\ 2 & \text{if } j\text{th unit receives treatment 3} \\ 0 & \text{if } j\text{th unit receives treatment 4} \end{cases}$$

$$X_{3j} = \begin{cases} -1 & \text{if } j\text{th unit receives treatment 1} \\ -1 & \text{if } j\text{th unit receives treatment 2} \\ -1 & \text{if } j\text{th unit receives treatment 3} \\ 3 & \text{if } j\text{th unit receives treatment 4} \end{cases}$$

Let $\mu_1, \mu_2, \mu_3, \mu_4$ be the treatment means.

$$\begin{aligned}
E(y_{1j}) &= \mu_1 = \mu - \tau_1 - \tau_2 - \tau_3 \\
E(y_{2j}) &= \mu_2 = \mu + \tau_1 - \tau_2 - \tau_3 \\
E(y_{3j}) &= \mu_3 = \mu + 2\tau_2 - \tau_3 \\
E(y_{4j}) &= \mu_4 = 3\tau_3
\end{aligned}$$

Use all four equations to solve for μ .

$$\mu = (\mu_1 + \mu_2 + \mu_3 + \mu_4)/4.$$

Use the first two equations to solve for τ_1 .

$$\mu_2 - \mu_1 = 2\tau_1 \Rightarrow \tau_1 = (1/2)(\mu_2 - \mu_1).$$

Use the first three equations to solve for τ_2 .

$$\mu_3 - (\mu_1 + \mu_2)/2 = 3\tau_2 \Rightarrow \tau_2 = (1/3)(\mu_3 - (\mu_1 + \mu_2)/2).$$

Use the all four equations to solve for τ_3 .

$$\mu_4 - (\mu_1 + \mu_2 + \mu_3)/3 = 4\tau_3 \Rightarrow \tau_3 = (1/4)(\mu_4 - (\mu_1 + \mu_2 + \mu_3)/3).$$

The interpretation of the parameter estimates: the intercept μ is the grand average; $\tau_i, i = 1, 2, 3$ is the average difference of the $(i + 1)th$ mean and the average of the subsequent means.

```

q2data$trt <- as.factor(q2data$trt)
contrasts(q2data$trt) <- contr.helmert(4)
summary(lm(y~as.factor(trt),data = q2data))

##
## Call:
## lm(formula = y ~ as.factor(trt), data = q2data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.295 -1.180 -0.048  1.391  4.336
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.41797    0.08996  26.879 < 2e-16 ***
## as.factor(trt)1  0.04214    0.12771   0.330  0.74157
## as.factor(trt)2  0.02389    0.07372   0.324  0.74603
## as.factor(trt)3  0.14575    0.05154   2.828  0.00493 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.788 on 391 degrees of freedom
## Multiple R-squared:  0.02061,    Adjusted R-squared:  0.0131
## F-statistic: 2.743 on 3 and 391 DF,  p-value: 0.04294

```

Verify the parameter estimates using the table of means.

```
xbar <- sapply(split(q2data$y,q2data$trt),mean) # treatment averages
sum(xbar)/4 # intercept mu
```

```
## [1] 2.417966
```

```
(xbar[2]-xbar[1])/2 #beta_1_hat
```

```
##          2
## 0.04214389
```

```
(xbar[3]-(xbar[1]+xbar[2])/2)/3 #beta_2_hat
```

```
##          3
## 0.02389354
```

```
(xbar[4]-(xbar[1]+xbar[2]+xbar[3])/3)/4 #beta_3_hat
```

```
##          4
## 0.1457464
```

- (e) Which coding scheme do you think makes more sense for evaluating if there is a significant difference between any of the new treatments and placebo.

The coding scheme in part (b) is more appropriate since the Helmert coding scheme does not compare any of the new treatments to the placebo.

- (f) Which pairs of treatments have a statistically significant difference? Do your results change if you adjust for multiple comparisons using either the Bonferroni or Tukey method? Compare all pairs of treatment means using no adjustment, Bonferroni, and Tukey. If the unadjusted, Bonferroni, and Tukey lead to different conclusions then explain why these methods give different results. Does it make sense to consider all pairs of treatment means given the main objective of this study? (Hand in your R code and output)

It doesn't really make sense to consider all possible treatment pairs since the primary question is to compare placebo to all the new treatments. Nevertheless, both Bonferroni and Tukey make these comparisons plus other comparisons.

Treatment 4 is significantly different from placebo in an unadjusted comparison, but this difference becomes non-significant after adjusting for multiple comparisons using both Tukey and Bonferroni. The reason that Tukey and Bonferroni give different p-values compared to the unadjusted is because both Tukey and Bonferroni ensure that the family-wise type I error rate remains at 5%, and the unadjusted family-wise type I error rate is $1 - (1 - .05)^6 = 0.26$.

```
TukeyHSD(aov(y~as.factor(trt),data = q2data))
```

```
##    Tukey multiple comparisons of means
##      95% family-wise confidence level
##
## Fit: aov(formula = y ~ as.factor(trt), data = q2data)
##
```



```
## $`as.factor(trt)`
##      diff      lwr      upr      p adj
## 2-1 0.08428777 -0.574699935 0.7432755 0.9875857
## 3-1 0.11382450 -0.541723430 0.7693724 0.9699886
## 4-1 0.64902293 -0.001589317 1.2996352 0.0508310
## 3-2 0.02953673 -0.632736356 0.6918098 0.9994544
## 4-2 0.56473516 -0.092652737 1.2221231 0.1205489
## 4-3 0.53519843 -0.118741276 1.1891381 0.1511693
```

```
pairwise.t.test(q2data$y,as.factor(q2data$trt),p.adjust.method = "bonferroni")
```

```
##
## Pairwise comparisons using t tests with pooled SD
##
## data: q2data$y and as.factor(q2data$trt)
##
##      1      2      3
## 2 1.000 -      -
## 3 1.000 1.000 -
## 4 0.063 0.163 0.212
##
## P value adjustment method: bonferroni
```

```
pairwise.t.test(q2data$y,as.factor(q2data$trt),p.adjust.method = "none")
```

```
##
## Pairwise comparisons using t tests with pooled SD
##
## data: q2data$y and as.factor(q2data$trt)
##
##      1      2      3
## 2 0.742 -      -
## 3 0.654 0.908 -
## 4 0.010 0.027 0.035
##
## P value adjustment method: none
```

3. A chemical engineer studied three variables: temperature (x_1), pH (x_2), and agitation rate (x_3). Each variable has two levels. The dependent variable is the yield of a chemical reaction. The data is shown in the table below and is available in the file `q3data.csv`.

X	run	x1	x2	x3	y
1	1	-1	-1	-1	59
2	2	1	-1	-1	60
3	3	-1	1	-1	63
4	4	1	1	-1	62
5	5	-1	-1	1	63
6	6	1	-1	1	64
7	7	-1	1	1	54
8	8	1	1	1	59
9	9	-1	-1	-1	55

X	run	x1	x2	x3	y
10	10	1	-1	-1	56
11	11	-1	1	-1	61
12	12	1	1	-1	58
13	13	-1	-1	1	63
14	14	1	-1	1	53
15	15	-1	1	1	56
16	16	1	1	1	56

- (a) Explain why this a 2^k factorial experiment? What is the value of k . How many replications of each run did the engineer conduct? Explain.

$k = 3$ since there are three factors at two levels. All factor-level combinations are studied therefore it's a factorial design.

- (b) Estimate all the factorial effects and their standard errors by fitting a linear regression model in R. (Hand in your R code and output)

```
attach(q3data)
x1 <- as.factor(x1)
x2 <- as.factor(x2)
x3 <- as.factor(x3)
fact.mod <- lm(y~x1*x2*x3,data=q3data)
summary(fact.mod)

##
## Call:
## lm(formula = y ~ x1 * x2 * x3, data = q3data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.500 -1.625  0.000  1.625  5.500
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   58.8750     0.8524   69.071 2.15e-12 ***
## x1             -0.3750     0.8524   -0.440  0.6716
## x2             -0.2500     0.8524   -0.293  0.7768
## x3             -0.3750     0.8524   -0.440  0.6716
## x1:x2           0.5000     0.8524    0.587  0.5737
## x1:x3          -0.1250     0.8524   -0.147  0.8870
## x2:x3          -2.0000     0.8524   -2.346  0.0469 *
## x1:x2:x3        1.2500     0.8524    1.466  0.1807
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.41 on 8 degrees of freedom
## Multiple R-squared:  0.515, Adjusted R-squared:  0.09061
## F-statistic: 1.214 on 7 and 8 DF, p-value: 0.3926
```

```
2*fact.mod$coefficients # factorial effects
```

```
## (Intercept)      x1      x2      x3      x1:x2      x1:x3
##      117.75     -0.75     -0.50     -0.75      1.00     -0.25
##      x2:x3      x1:x2:x3
##      -4.00      2.50
```

The standard error of the effects is $2 \times 0.8524 = 1.7048$.

(c) Interpret all the main and interaction effects in the context of this study.

The main effect of, say, temperature is the average difference between yield when temperature is at the level coded as +1 versus yield when temperature is at the level coded -1.

A two-way interaction effect between, say, temperature and pH is the average difference between the effect of temperature when pH is coded as +1 and the effect of temperature when pH is coded as -1.

The other main effects and two-factor interaction effects have similar interpretations.

The three-way interaction effect between temperature, pH, agitation rate is the average difference between the interaction effect of temperature and pH when agitation rate is coded as +1 and the interaction effect of temperature and pH when agitation rate is coded as -1.

Side note

The following calculations show why the factorial effects are interpreted as above (the question did not ask for this, but it's provided so that you can verify the interpretations above.)

The R code

- `q3data$y[q3data$x1==1]` extracts the yield values when `x1=1`.
- `q3data$y[q3data$x1==-1]` extracts the yield values when `x1=-1`.
- `q3data$y[q3data$x1==1&q3data$x2==1]` extracts the yield values when `x1=1` and `x2=1`.
- `q3data$y[q3data$x1==1&q3data$x2==1&q3data$x3==1]` extracts the yield values when `x1=1` and `x2=1` and `x3=1`.
- Other data vectors are defined in a similar way below.

The average yield when temperature is at +1 level is:

```
mean(q3data$y[q3data$x1==1])
```

```
## [1] 58.5
```

The average yield when temperature is at -1 level is:

```
mean(q3data$y[q3data$x1==-1])
```

```
## [1] 59.25
```

The average difference is the main effect of temperature:

```
58.5-59.25
```

```
## [1] -0.75
```

The average effect of temperature when pH=+1 is:

```
mean(q3data$y[q3data$x1==1&q3data$x2==1])-mean(q3data$y[q3data$x1==1&q3data$x2==1])
```

```
## [1] 0.25
```

The average effect of temperature when pH=-1 is:

```
mean(q3data$y[q3data$x1==1&q3data$x2==1])-mean(q3data$y[q3data$x1==1&q3data$x2==1])
```

```
## [1] -1.75
```

The interaction between temperature and pH is the average difference between these two effects.

```
(0.25-(-1.75))/2
```

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

When agitation rate is at it's +1 level the interaction between temperature and pH is

```
mean(q3data$y[q3data$x1==1&q3data$x2==1&q3data$x3==1])
```

```
## [1] 57.5
```

```
mean(q3data$y[q3data$x1==1&q3data$x2==1&q3data$x3==1])
```

```
## [1] 55
```

```
mean(q3data$y[q3data$x1==1&q3data$x2==1&q3data$x3==1])
```

```
## [1] 58.5
```

```
mean(q3data$y[q3data$x1==1&q3data$x2==1&q3data$x3==1])
```

```
## [1] 63
```

```
((57.5-55)-(58.5-63))/2
```

```
## [1] 3.5
```

When agitation rate is at it's -1 level the interaction between temperature and pH is

```
mean(q3data$y[q3data$x1==1&q3data$x2==1&q3data$x3==1])
```

```
## [1] 60
```

```
mean(q3data$y[q3data$x1==1&q3data$x2==1&q3data$x3==1])
```

```
## [1] 62
```

```
mean(q3data$y[q3data$x1==1&q3data$x2==1&q3data$x3==1])
```

```
## [1] 58
```

```
mean(q3data$y[q3data$x1==1&q3data$x2==1&q3data$x3==1])
```

```
## [1] 57
```

```
((60-62)-(58-57))/2
```

```
## [1] -1.5
```

The three-way interaction between temperature, pH, and agitation rate is

```
(3.5-(-1.5))/2
```

```
## [1] 2.5
```

All interactions are symmetric in their interpretation.

4. A food scientist is interested in studying the effect of two types of flour (call this factor A with levels a_1, a_2) and two amounts of flavouring (call this factor B with levels b_1, b_2) on taste (y). y is measured on scale of 0 to 10 with 0 meaning poor taste and 10 meaning excellent taste. The variance of y is σ^2 .

- (a) The scientist wants to compare the two types of flour so decides to keep factor B at the low level since he almost always uses B at the low level in most of the products he works on. So there are two treatments: a_1b_1 and a_2b_1 . If the scientist obtains four observations for each treatment then what is the variance of the mean difference of the effect of flour (i.e., factor A)? Show your work.

Suppose that we have four independent observations for each treatment.

a_1b_1	a_2b_1
y_{11}	y_{12}
y_{21}	y_{22}
y_{31}	y_{32}
y_{41}	y_{42}

Let \bar{y}_1 be the mean of the observations under treatment a_1b_1 and \bar{y}_2 be the mean of the observations under

treatment a_2b_1 . The $Var(y_{ij}) = \sigma^2$. Therefore,

$$Var(\bar{y}_1 - \bar{y}_2) = 4\sigma^2/16 + 4\sigma^2/16 = \sigma^2/2$$

- (b) A 2^2 factorial design with factors A and B replicated twice is also conducted. What are the variances of the main effect of A? Show your work.

A replicated 2^2 design looks like

A	B	AB	Rep 1	Rep 2	Average
-1	-1	+1	y_{11}	y_{12}	\bar{y}_1
+1	-1	-1	y_{21}	y_{22}	\bar{y}_2
-1	+1	-1	y_{31}	y_{32}	\bar{y}_3
+1	+1	+1	y_{41}	y_{42}	\bar{y}_4

The main effect of A is

$$A = (-\bar{y}_1 + \bar{y}_2 - \bar{y}_3 + \bar{y}_4)/2$$

So the variance is:

$$\begin{aligned}
 Var(A) &= Var((-\bar{y}_1 + \bar{y}_2 - \bar{y}_3 + \bar{y}_4)/2) \\
 &= (1/4)(Var(\bar{y}_1) + Var(\bar{y}_2) + Var(\bar{y}_3) + Var(\bar{y}_4)) \\
 &= (1/4)(4\sigma^2/2) \\
 &= \sigma^2/2
 \end{aligned}$$

Side note:

The replicated factorial experiment used 8 observations to estimate all the factorial effects (2 main effects and the interaction) with the same precision as the experiment in part (a) which is only able to estimate the main effect of A at a fixed level of B. This is why factorial experiments are considered more efficient.

- (c) Another scientist wants to know if the effect of flour is the same when the amount flavouring used is low and high? Which experiment can answer this question: the experiment in (a) or the experiment in (b)? Explain which estimate you would report to the scientist. Explain your reasoning.

The experiment in (b) can answer this question since a factorial experiment is able to estimate interaction effects.