A3

Yue Han

```
library(tidyverse)
library(biotools)
```

1a.)

```
x = read.table("profilePotato.txt", header = TRUE)
head(x)
```

```
##   texture flavor moist cookmeth
## 1     2.9    3.2   3.0        1
## 2     1.8    3.0   1.7        1
## 3     1.8    2.6   1.5        1
## 4     2.6    3.1   2.4        1
## 5     3.1    3.0   2.8        1
## 6     1.8    2.6   1.8        1
```

```
xm1 = subset(x, cookmeth== 1, select=c(texture, flavor, moist))
xm2 = subset(x, cookmeth== 2, select=c(texture, flavor, moist))
head(xm1) #cooking method 1
```

```
##   texture flavor moist
## 1     2.9    3.2   3.0
## 2     1.8    3.0   1.7
## 3     1.8    2.6   1.5
## 4     2.6    3.1   2.4
## 5     3.1    3.0   2.8
## 6     1.8    2.6   1.8
```

```
head(xm2) #cooking method 2
```

```
##    texture flavor moist
## 33     1.9    2.8   2.2
## 34     2.0    2.8   1.9
## 35     2.5    3.0   2.1
## 36     3.0    2.6   2.7
## 37     1.9    2.9   2.0
## 38     2.0    3.0   2.0
```

```
p = ncol(xm1)
n1 = nrow(xm1)
n2 = nrow(xm2)

xbar1 = colMeans(xm1)
xbar1 #mean cooking quality measures for method 1
```

```
## texture    flavor     moist
## 2.400000 2.937500 2.240625
```

```
xbar2 = colMeans(xm2)
xbar2 #mean cooking quality measures for method 2
```

```
## texture    flavor     moist
## 2.465625 2.806250 2.325000
```

Below is the plot with method 1 in blue and method 2 in red. We see that the means between method 1 and 2 are rather close and follow the same pattern.

```
variables = c(1,2,3)
plot(variables, xbar1, ylim=range(c(1,4)),
type="l", col="blue",
ylab = "Mean Score",
axes = FALSE,)
axis(side = 1, at = variables)
axis(side = 2)
lines(variables, xbar2, col="red")
```
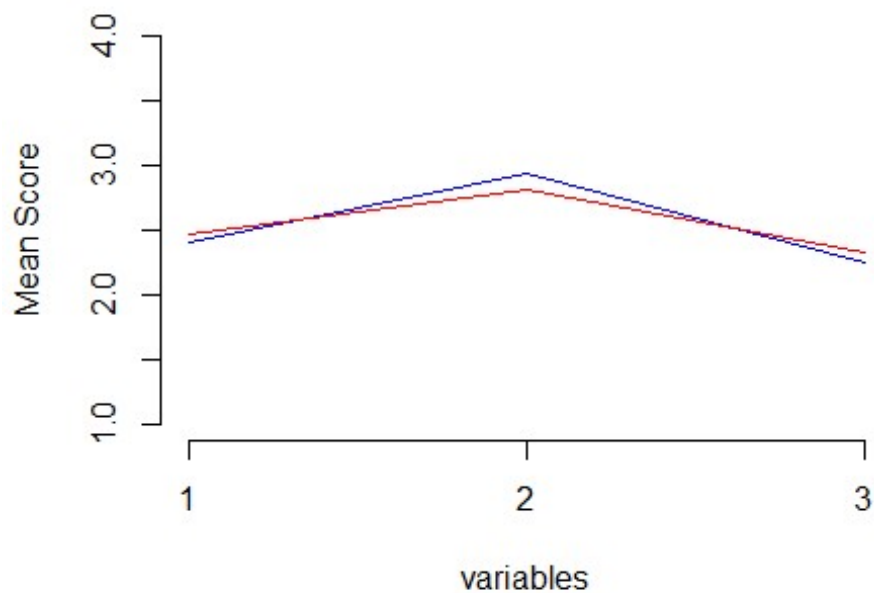

1b.) # test for parallel profiles

```
S1 = cov(xm1)
S2 = cov(xm2)
Sp = (1/(n1+n2-2))*((n1-1)*S1+(n2-1)*S2)
S1
```

```
##            texture      flavor       moist
## texture 0.24774194 0.03354839 0.19612903
## flavor  0.03354839 0.08435484 0.02552419
## moist   0.19612903 0.02552419 0.21023185

S2

##           texture      flavor       moist
## texture 0.2565222 0.03312500 0.15830645
## flavor  0.0331250 0.09415323 0.02758065
## moist   0.1583065 0.02758065 0.14322581

Sp

##            texture      flavor       moist
## texture 0.25213206 0.03333669 0.17721774
## flavor  0.03333669 0.08925403 0.02655242
## moist   0.17721774 0.02655242 0.17672883
```

```r
C = matrix(c(-1, 1, 0, 0, -1, 1), byrow = TRUE, 2, 3)
C
```

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

```r
C%*%(xbar1-xbar2)
```

```
##            [,1]
## [1,]  0.196875
## [2,] -0.215625
```

Testing profile parallelism $H_0: C\mu_1 = C\mu_2$ based on page 35 lec chap6_2way_manova_profile we get 3.527515 < 6.398789 so do not reject hypothesis of parallel profiles and p-value greater than 0.05 so do not reject could be parallel.

```r
T_sq_parallel = ((1/n1+1/n2)^(-1))*t(xbar1-
xbar2)%*%t(C)%*%solve(C%*%Sp%*%t(C))%*%C%*%(xbar1-xbar2)
T_sq_parallel
```

```
##          [,1]
## [1,] 3.527515
```

```r
alpha =0.05
CritValue_parallel = (((n1+n2-2)*(p-1))/(n1+n2-p))*qf(1-alpha, p-1, n1+n2-p)
CritValue_parallel
```

```
## [1] 6.398789
```

```r
F_parallel = T_sq_parallel/(((((n1+n2-2)*(p-1))/(n1+n2-p)))
F_parallel
```

```
##         [,1]
## [1,] 1.73531
```

```
p_parallel = 1-pf(F_parallel, p-1 , n1+n2-p)
p_parallel
```

```
##          [,1]
## [1,] 0.184937
```

Test for coincident profiles $H_0 : 1'(\mu_1-\mu_2)=0$ based on pg 36 lec chap6_2way_manova_profile we get 0.00567 < 3.996 so do not reject. We also see that the p-value is much larger than 0.05 so do not reject, could have coincident.

```
one <- as.matrix(c(rep(1,p)))
T_sq_coincident <- ((t(one)%*%(xbar1-
xbar2))/sqrt((1/n1+1/n2)*(t(one)%*%Sp%*%one)))^2
T_sq_coincident
```

```
##             [,1]
## [1,] 0.005668485
```

```
alpha <- 0.05
CritValue_coincident <- qf(1-alpha, 1, n1+n2-2)
CritValue_coincident
```

```
## [1] 3.995887
```

```
p_coincident <- 1-pf(T_sq_coincident, 1, n1+n2-2 )
p_coincident
```

```
##           [,1]
## [1,] 0.9402269
```

Test for level profiles $H_0: C\mu$ based on page 38 lec chap6_2way_manova_profile we get 115.3216 > 6.391977 so reject hypothesis also p-value is way smaller than 0.05 so reject. Cannot be level

```
S <- cov(x[,1:3])
S
```

```
##            texture     flavor      moist
## texture 0.24922371 0.03062004 0.17581101
## flavor  0.03062004 0.09221230 0.02331845
## moist   0.17581101 0.02331845 0.17573165
```

```
xbar <- colMeans(x[,1:3])
xbar
```

```
##  texture    flavor     moist
## 2.432812 2.871875 2.282812
```

```
C = matrix(c(-1, 1, 0, 0, -1, 1), byrow = TRUE, 2, 3)
C
```

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

T_sq_level <- (n1+n2)*t(C%*%xbar)%*%solve(C%*%S%*%t(C))%*%(C%*%xbar)
T_sq_level

##          [,1]
## [1,] 115.3216

CritValue_level <- (((n1+n2-1)*(p-1))/(n1+n2-p+1))*qf(1-alpha, p-1, n1+n2-
p+1)
CritValue_level

## [1] 6.391977

F_level <- T_sq_level/(((n1+n2-1)*(p-1))/(n1+n2-p+1)))
F_level

##          [,1]
## [1,] 56.74555

p_level <- 1-pf(F_level, p-1 , n1+n2-p+1)
p_level

##              [,1]
## [1,] 9.769963e-15

y = read.table("2sample_data.txt", header = TRUE)
head(y)

##     x1  x2  x3 x4 Period
## 1 131 138  89 49      1
## 2 125 131  92 48      1
## 3 131 132  99 50      1
## 4 119 132  96 44      1
## 5 136 143 100 54      1
## 6 138 137  89 56      1

yp1 = subset(y, Period== 1, select=c(x1, x2, x3, x4))
yp2 = subset(y, Period== 2, select=c(x1, x2, x3, x4))
head(yp1) #period 1

##     x1  x2  x3 x4
## 1 131 138  89 49
## 2 125 131  92 48
## 3 131 132  99 50
## 4 119 132  96 44
## 5 136 143 100 54
## 6 138 137  89 56

head(yp2) #period 2
```

```
##      x1   x2   x3 x4
## 31 125 136   99 46
## 32 134 132   95 46
## 33 139 132   96 43
## 34 149 127 102 49
## 35 127 122   93 43
## 36 136 134   96 50

p = ncol(yp1)
n1 = nrow(yp1)
n2 = nrow(yp2)
```

2a.) observed sample means for period 1 and 2 respectively

```
ybar1 = colMeans(yp1)
ybar1 #sample means for period 1

##         x1         x2         x3         x4
## 131.36667 133.60000  99.16667  50.53333

ybar2 = colMeans(yp2)
ybar2 #sample means for period 2

##         x1         x2         x3         x4
## 133.36667 130.70000  97.06667  48.23333
```

Observed sample covariance matrix of period 1, period 2, and pooled respectively

```
S1 = cov(yp1)
S2 = cov(yp2)
Sp = (1/(n1+n2-2))*((n1-1)*S1+(n2-1)*S2)
S1

##            x1         x2         x3         x4
## x1 26.309195   4.1517241   0.4540230   7.2459770
## x2  4.151724  19.9724138  -0.7931034   0.3931034
## x3  0.454023  -0.7931034  34.6264368  -1.9195402
## x4  7.245977   0.3931034  -1.9195402   7.6367816

S2

##            x1         x2         x3         x4
## x1 23.136782   1.010345   4.7678161 1.8425287
## x2  1.010345  21.596552   3.3655172 5.6241379
## x3  4.767816   3.365517  18.8919540 0.1908046
## x4  1.842529   5.624138   0.1908046 8.7367816

Sp

##            x1         x2         x3         x4
## x1 24.722989   2.581034   2.6109195   4.5442529
## x2  2.581034  20.784483   1.2862069   3.0086207
```

```
## x3   2.610920   1.286207 26.7591954 -0.8643678
## x4   4.544253   3.008621 -0.8643678  8.1867816
```

2c.) $H_0: \Sigma_1 =\ldots= \Sigma_g = \Sigma$ $\alpha = 0.05 < p - value = 0.428$ fail to reject hypothesis. So the sigmas could be equal

```
boxM(y[, 1:4], y[, "Period"])

##
##  Box's M-test for Homogeneity of Covariance Matrices
##
## data:  y[, 1:4]
## Chi-Sq (approx.) = 10.143, df = 10, p-value = 0.428
```

2d.) $H_0: \mu_1 - \mu_2 = \delta_0$ from page 37 Chap6_Comp_of_several_means 24.59121 > 10.71287 so reject that hypothesis. So the covariance matrices are not equal. The p-value is 1.951772e-13 which is <0.05 so we reject hypothesis.

```
delta0 <- c(0,0)
T_sq <- t(ybar1-ybar2-delta0)%*%solve((1/n1+1/n2)*Sp)%*%(ybar1-ybar2-delta0)
T_sq

##           [,1]
## [1,] 24.59121

criticalValue <- (((n1+n2-2)*p)/(n1+n2-p-1))*qf(1-0.05,p, n1+n2-p-1 )
criticalValue

## [1] 10.71287

F_level <- T_sq_level/(((((n1+n2-1)*(p-1))/(n1+n2-p+1)))
F_level

##           [,1]
## [1,] 37.13747

p_level <- 1-pf(F_level, p-1 , n1+n2-p+1)
p_level

##                [,1]
## [1,] 1.951772e-13
```

Most critical linear combinination leading to rejection of $H_0: \mu_1 - \mu_2 = \delta_0$ is

```
solve(S1*1/n1+S2*1/n2)%*%(ybar1-ybar2)

##          [,1]
## x1 -2.488406
## x2  1.558151
## x3  1.512469
## x4  5.182426
```

e.)

```
dybar <- ybar1-ybar2
Spd <- diag(Sp)
c <- sqrt(criticalValue)
SE <- sqrt(Spd*(1/n1+1/n2))
CI_L_sim <- dybar-c*SE
CI_U_sim <- dybar+c*SE
CI_L_Bon <- dybar-qt(1-0.05/(2*p),n1+n2-2)*SE
CI_U_Bon <- dybar+qt(1-0.05/(2*p),n1+n2-2)*SE
CIs <- cbind(dybar, CI_L_sim, CI_U_sim, CI_L_Bon, CI_U_Bon)
colnames(CIs) <- c("ybar period 1-2", "sim Lower", "sim Upper", "Bon Lower",
"Bon Upper")
CIs

##    ybar period 1-2  sim Lower sim Upper  Bon Lower Bon Upper
## x1           -2.0 -6.2020164  2.202016 -5.3096785  1.309679
## x2            2.9 -0.9528036  6.752804 -0.1346243  5.934624
## x3            2.1 -2.2716339  6.471634 -1.3432761  5.543276
## x4            2.3 -0.1180421  4.718042  0.3954520  4.204548
```

f.)

```
n <- n1+n2
g = 2

CritValue <- qf(1-0.01, 2*(g-1), 2*(n-g-1))
CritValue

## [1] 4.796315
```

We are hypothesizing that the means are equal $5.8298 > 4.786$ so reject hypothesis p-value is larger than the one in part d but still < 0.05

```
y14<- as.matrix(y[,1:4])
per = factor(y$Period)
fit <- manova(y14~per)
summary(fit, test = "Wilks")

##           Df   Wilks approx F num Df den Df    Pr(>F)
## per        1 0.70225   5.8298      4     55 0.0005518 ***
## Residuals 58
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

3.)

```
z = read.table("t6_17.txt", header = TRUE)
head(z)

##   Fac1 Fac2    x1    x2   x3
## 1    1    5 195.3 153.1 51.4
## 2    1    5 194.3 167.7 53.7
## 3    2    5 189.7 139.5 55.5
```

```
## 4    2    5 180.4 121.1 44.4
## 5    1    6 203.0 156.8 49.8
## 6    1    6 195.9 166.0 45.8

fac1 <- factor(z$Fac1)
fac2 <- factor(z$Fac2)
z345 <- as.matrix(z[,3:5])
fit <- manova(z345~fac1*fac2)
moreinfo <- summary(fit)
summary(fit)

##            Df  Pillai approx F num Df den Df   Pr(>F)
## fac1        1 0.89348  11.1843      3      4 0.020502 *
## fac2        2 1.70911   9.7924      6     10 0.001056 **
## fac1:fac2   2 1.29086   3.0339      6     10 0.058708 .
## Residuals   6
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

a.)

```
SSPfac1 <- moreinfo$SS[1]
SSPfac1

## $fac1
##             x1         x2          x3
## x1   0.7008333   -10.6575    7.129167
## x2 -10.6575000   162.0675 -108.412500
## x3   7.1291667  -108.4125   72.520833

SSPfac2 <- moreinfo$SS[2]
SSPfac2

## $fac2
##           x1        x2        x3
## x1 196.1150   365.1825   42.6275
## x2 365.1825  1089.0150  414.6550
## x3  42.6275   414.6550  284.1017

SSPint = moreinfo$SS[3]
SSPint

## $`fac1:fac2`
##           x1        x2         x3
## x1 205.1017  363.6675  107.78583
## x2 363.6675  780.6950  254.22000
## x3 107.7858  254.2200   85.95167

SSPres = moreinfo$SS[4]
SSPres
```

```
## $Residuals
##          x1      x2       x3
## x1 104.205  49.365  76.480
## x2  49.365 352.105 121.995
## x3  76.480 121.995  94.835
```

b.)

```
fitb <- manova(z345~fac1*fac2-fac1:fac2)
moreinfob <- summary(fitb)
summary(fitb)

##            Df  Pillai approx F num Df den Df    Pr(>F)
## fac1        1 0.76661   6.5693      3      6 0.025249 *
## fac2        2 1.32107   4.5402      6     14 0.009268 **
## Residuals   8
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

SSPfac1b <- moreinfob$SS[1]
SSPfac1b

## $fac1
##            x1        x2         x3
## x1  0.7008333 -10.6575   7.129167
## x2 -10.6575000 162.0675 -108.412500
## x3  7.1291667 -108.4125  72.520833

SSPfac2b <- moreinfob$SS[2]
SSPfac2

## $fac2
##          x1       x2       x3
## x1 196.1150  365.1825  42.6275
## x2 365.1825 1089.0150 414.6550
## x3  42.6275  414.6550 284.1017

SSPresb = moreinfob$SS[3]
SSPresb

## $Residuals
##          x1       x2       x3
## x1 309.3067  413.0325 184.2658
## x2 413.0325 1132.8000 376.2150
## x3 184.2658  376.2150 180.7867
```

$SSP_{int} + SSP_{res}$ for the model with interaction $= SSP_{res}$ of model without interaction

```
SSPintm <- SSPint$`fac1:fac2`
SSPintm
```

```
##            x1        x2        x3
## x1 205.1017 363.6675 107.78583
## x2 363.6675 780.6950 254.22000
## x3 107.7858 254.2200  85.95167

SSPresm <- SSPres$Residuals
SSPresm

##           x1      x2      x3
## x1 104.205  49.365  76.480
## x2  49.365 352.105 121.995
## x3  76.480 121.995  94.835

SSPintm+SSPresm

##            x1         x2       x3
## x1 309.3067   413.0325 184.2658
## x2 413.0325 1132.8000 376.2150
## x3 184.2658   376.2150 180.7867
```