Homework 3 (30 points)

1. **Chapter 6 Page 245-246: #6.27 (a) (b) (c)**
2. **Chapter 6 Page 246-247: #6.28 (a)**

HW 3 HELP FILE

6.27 (a)

Input the data into R

```
> method1
   y1  y2  y3  y4
1  5.4 6.0 6.3 6.7
2  5.2 6.2 6.0 5.8
3  6.1 5.9 6.0 7.0
4  4.8 5.0 4.9 5.0
5  5.0 5.7 5.0 6.5
6  5.7 6.1 6.0 6.6
7  6.0 6.0 5.8 6.0
8  4.0 5.0 4.0 5.0
9  5.7 5.4 4.9 5.0
10 5.6 5.2 5.4 5.8
11 5.8 6.1 5.2 6.4
12 5.3 5.9 5.8 6.0
```

```
> method2
   y1  y2  y3  y4
13 5.0 5.3 5.3 6.5
14 4.8 4.9 4.2 5.6
15 3.9 4.0 4.4 5.0
16 4.0 5.1 4.8 5.8
17 5.6 5.4 5.1 6.2
18 6.0 5.5 5.7 6.0
19 5.2 4.8 5.4 6.0
20 5.3 5.1 5.8 6.4
21 5.9 6.1 5.7 6.0
22 6.1 6.0 6.1 6.2
23 6.2 5.7 5.9 6.0
24 5.1 4.9 5.3 4.8
```

```
> method3
   y1  y2  y3  y4
25 4.8 5.0 6.5 7.0
26 5.4 5.0 6.0 6.4
27 4.9 5.1 5.9 6.5
28 5.7 5.2 6.4 6.4
29 4.2 4.6 5.3 6.3
30 6.0 5.3 5.8 6.4
31 5.1 5.2 6.2 6.5
32 4.8 4.6 5.7 5.7
33 5.3 5.4 6.8 6.6
34 4.6 4.4 5.7 5.6
35 4.5 4.0 5.0 5.9
36 4.4 4.2 5.6 5.5
```

Calculate "between" matrices

$$\boldsymbol{H} = n\sum_{i=1}^{k}(\bar{\boldsymbol{y}}_{i.}-\bar{\boldsymbol{y}}_{..})(\bar{\boldsymbol{y}}_{i.}-\bar{\boldsymbol{y}}_{..})'$$

$\bar{\boldsymbol{y}}_{i.}$ is the column mean of method i, $\bar{\boldsymbol{y}}_{..}$ is the column mean of all samples. n is the number of samples.

```
method1.bar <- colMeans(method1)
method2.bar <- colMeans(method2)
method3.bar <- colMeans(method3)
```
$\boxed{\leftarrow \bar{\boldsymbol{y}}_{i.}}$

```
method.all.bar <- (method1.bar+method2.bar+method3.bar)/3
```
$\boxed{\leftarrow \bar{\boldsymbol{y}}_{..}}$

```
method1.bar.diff <- method1.bar - method.all.bar
method2.bar.diff <- method2.bar - method.all.bar
method3.bar.diff <- method3.bar - method.all.bar
```
$\boxed{\leftarrow \bar{\boldsymbol{y}}_{i.}-\bar{\boldsymbol{y}}_{..}}$

```
H <- 12 * unname(method1.bar.diff %*% t(method1.bar.diff) + method2.bar.diff %*%
t(method2.bar.diff) + method3.bar.diff %*% t(method3.bar.diff))
```

```
> H
           [,1]    [,2]      [,3]       [,4]
[1,]  1.0505556  2.1200 -1.375556 -0.7602778
[2,]  2.1200000  4.6050 -2.340000 -1.2425000
[3,] -1.3755556 -2.3400  2.382222  1.3844444
[4,] -0.7602778 -1.2425  1.384444  0.8105556
```

Calculate "within" matrices

$$\boldsymbol{E} = \sum_{i=1}^{k}\sum_{j=1}^{n}(\boldsymbol{y}_{ij} - \bar{\boldsymbol{y}})(\boldsymbol{y}_{ij} - \bar{\boldsymbol{y}})'$$

```
"compute.within.matrix" <- function(data, mean) {
  ret <- matrix(as.numeric(0), nrow=4, ncol=4)
  for (i in 1:12) {
     diff <- as.numeric(unname(data[i,] - mean))
     ret <- ret + diff %*% t(diff)
  }
  return(ret)
}


E <- compute.within.matrix(method1, method1.bar) + compute.within.matrix(method2, method2.bar) +
compute.within.matrix(method3, method3.bar)
```

```
> E
          [,1]     [,2]      [,3]      [,4]
[1,] 13.408333 7.778333  8.675000  5.864167
[2,]  7.778333 8.102500  7.359167  6.268333
[3,]  8.675000 7.359167 11.607500  7.037500
[4,]  5.864167 6.268333  7.037500 10.565833
```

#Four MANOVA

Wilks' Test Statistic $\Lambda = \frac{|\boldsymbol{E}|}{|\boldsymbol{E}+\boldsymbol{H}|}$

← Determinant of E/Determinant of E+H

```
Lambda <- det(E) / det(E + H)
```

```
> Lambda
[1] 0.2200479
```

Pillai Statistic $V^{(s)} = tr[(E + H)^{-1}H]$

To calculate trace of the matrix in R, you will need to install "psych" package.

```
> install.packages("psych")
```

```
> library(psych)

V.s <- tr(solve(E + H) %*% H)
```

```
> V.s
[1] 0.8641859
```

Lawley-Hotelling statistic is defined as $U^{(s)} = tr(\boldsymbol{E}^{-1}\boldsymbol{H})$

```
U.s <- tr(solve(E) %*% H)
```

```
> U.s
[1] 3.161668
```

Roy's largest root test $\Theta = \frac{\lambda_1}{1+\lambda_1}$

λ_1 is the largest eigenvalue of matrix E

```
lambda.1 <- eigen(solve(E) %*% H)$values[1]

theta <- lambda.1 / (1 + lambda.1)
```

```
> theta
[1] 0.7522032
```

6.27 (b)

See textbook page 188 example 6.1.8.

6.27 (c)

Read textbook page 189-190 and see example 6.2