DA410_Assignment3_MattGraham

```
library(psych)  # used for Pillai Statistic
```

6.27

Baten, Tack, and Baeder (1958) compared judges' scores on fish prepared by three methods. Twelve fish were cooked by each method, and several judges tasted fish samples and rated each on four variables: y1 — aroma, y2 = flavor, y3 = texture, and y4 — moisture. The data are in Table 6.17. Each entry is an average score for the judges on that fish.

```
t_6.17 <- read.table("C:/mattgraham93.github.io/school/22_3_DA410/data/T6_17_FISH.DAT", heade
r=FALSE)
colnames(t_6.17) <- c("method", paste("Y", sep = '.', 1:4))
t_6.17$method <- as.factor(t_6.17$method)

method1 <- t_6.17[t_6.17$method==1, -1]
method2 <- t_6.17[t_6.17$method==2, -1]
method3 <- t_6.17[t_6.17$method==3, -1]

as.data.frame(t_6.17)
```

method <fct>	Y.1 <dbl>	Y.2 <dbl>	Y.3 <dbl>	Y.4 <dbl>
1	5.4	6.0	6.3	6.7
1	5.2	6.5	6.0	5.8
1	6.1	5.9	6.0	7.0
1	4.8	5.0	4.9	5.0
1	5.0	5.7	5.0	6.5
1	5.7	6.1	6.0	6.6
1	6.0	6.0	5.8	6.0
1	4.0	5.0	4.0	5.0
1	5.7	5.4	4.9	5.0
1	5.6	5.2	5.4	5.8

1-10 of 36 rows Previous **1** 2 3 4 Next

```
# calculating between matricies

method1.bar <- colMeans(method1)
method2.bar <- colMeans(method2)
method3.bar <- colMeans(method3)

method.all.bar <- (method1.bar + method2.bar + method3.bar) / 3

method1.bar.diff <- method1.bar - method.all.bar
method2.bar.diff <- method2.bar - method.all.bar
method3.bar.diff <- method3.bar - method.all.bar

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



"compute.within.matrix" <- function(data, mean) {
  ret <- matrix(as.numeric(0), nrow=4, ncol=4)

  for (i in 1:12) {
      diff <- as.numeric(data[i,] - mean)
      ret <- ret + diff %*% t(diff)
  }
  return(ret)
}

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

(a) Compare the three methods using all four MANOVA tests.

```
# Wilks' Test
print("Wilks' test")
```

```
## [1] "Wilks' test"
```

```
summary(manova(cbind(Y.1, Y.2, Y.3, Y.4) ~ method, data=t_6.17), test="Wilks")
```

```
##           Df   Wilks approx F num Df den Df    Pr(>F)
## method     2 0.22449   8.3294      8     60 1.609e-07 ***
## Residuals 33
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
print("")
```

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

```
# Roy's test
print("Wilks' test")
```

```
## [1] "Wilks' test"
```

```
summary(manova(cbind(Y.1, Y.2, Y.3, Y.4) ~ method, data=t_6.17), test="Roy")
```

```
##           Df    Roy approx F num Df den Df    Pr(>F)    
## method     2 2.9515   22.874      4     31 7.077e-09 ***
## Residuals 33                                            
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
print("")
```

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

```
# Pillai test
print("Pillai test")
```

```
## [1] "Pillai test"
```

```
summary(manova(cbind(Y.1, Y.2, Y.3, Y.4) ~ method, data=t_6.17), test="Pillai")
```

```
##           Df  Pillai approx F num Df den Df    Pr(>F)    
## method     2 0.85987    5.845      8     62 1.465e-05 ***
## Residuals 33                                             
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
print("")
```

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

```
# Lawley-Hotelling Test
print("Lawley-Hotelling' test")
```

```
## [1] "Lawley-Hotelling' test"
```

```
summary(manova(cbind(Y.1, Y.2, Y.3, Y.4) ~ method, data=t_6.17), test="Hotelling-Lawley")
```

```
##           Df Hotelling-Lawley approx F num Df den Df    Pr(>F)
## method     2           3.0788   11.161      8     58 2.161e-09 ***
## Residuals 33
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
print("")
```

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

```
# F-stat @ 95%
```

Analysis

According to all four of our tests, we can conclude there is significant evidence there are differences among the methods and how it impacts the overall score on that fish.

(b) Compute the following measures of multivariate association from Section 6.1.8.

```
k = 3
s = k - 1

# eta.lambda.sq
lambda <- det(E) / det(E+H)
eta.lambda <- 1 - lambda
sprintf("eta.lambda: %s", eta.lambda)
```

```
## [1] "eta.lambda: 0.775512677874853"
```

```
# eta.theta.sq:
lambda.1 <- eigen(solve(E) %*% H)$values[1]
theta <- lambda.1 / (1 + lambda.1)
sprintf("eta.theta.sq: %s", theta)
```

```
## [1] "eta.theta.sq: 0.746929971701092"
```

```
# a.lambda
a.lambda <- 1 - lambda ** (1 / s)
sprintf("a.lambda: %s", a.lambda)
```

```
## [1] "a.lambda: 0.526199069096369"
```

```
# a.rho
V.s <- tr(solve(E + H) %*% H)
a.rho <- V.s / s
sprintf("a.rho: %s", a.rho)
```

```
## [1] "a.rho: 0.42993691649735"
```

```
# a.lh
U.s <- tr(solve(E) %*% H)
a.lh <- (U.s / s) / (1 + (U.s / s))
sprintf("a.lh: %s", a.lh)
```

```
## [1] "a.lh: 0.606206175032727"
```

(c) Based on the eigenvalues, is the essential dimensionality of the space

containing the mean vectors equal to 1 or 2

```
# See example 6.2

vals <- eigen(solve(E) %*% H)[1]
eigen_mean <- sapply(vals, mean)
sprintf("Eigenvalue mean: %s", eigen_mean)
```

```
## [1] "Eigenvalue mean: 0.769699950326934"
```

—- Analysis Given our eigenvalues and averaging them out, we see that our mean vectors are more equal to 1 than 2.

6.28

Table 6.18 from Keuls et al. (1984) gives data from a two-way (fixed-effects) MANOVA on snap beans showing the results of four variables: yi = yield earliness, y2 = specific leaf area (SLA) earliness, y3 = total yield, and y4 = average SLA. The factors are sowing date (S) and variety (V).

```
t_6.18 <- read.table("C:/mattgraham93.github.io/school/22_3_DA410/data/T6_18_SNAPBEAN.DAT")
t_6.18 <- t_6.18[-3]  # remove sample id
colnames(t_6.18) <- c('S', 'V', 'Y1','Y2','Y3','Y4')  # label columns
as.data.frame(t_6.18)
```

S	V	Y1	Y2	Y3	Y4
<int>	<int>	<dbl>	<dbl>	<dbl>	<int>
1	1	59.3	4.5	38.4	295

S	V	Y1	Y2	Y3	Y4
<int>	<int>	<dbl>	<dbl>	<dbl>	<int>
1	1	60.3	3.5	38.6	302
1	1	60.9	5.3	37.2	318
1	1	60.6	5.8	38.1	345
1	1	60.4	6.0	38.8	325
1	2	59.3	6.7	37.9	275
1	2	59.4	4.8	36.6	290
1	2	60.0	5.1	38.7	295
1	2	58.9	5.8	37.5	296
1	2	59.5	4.8	37.0	330

1-10 of 60 rows Previous **1** 2 3 4 5 6 Next

a. Test for main effects and interaction using all four MANOVA statistics

```
# Wilks' Test
print("Wilks' test")
```

```
## [1] "Wilks' test"
```

```
summary(manova(cbind(Y1, Y2, Y3, Y4) ~ S*V, data=t_6.18), test="Wilks")
```

```
##           Df   Wilks approx F num Df den Df Pr(>F)
## S          1 0.03384   378.32      4     53 <2e-16 ***
## V          1 0.90842     1.34      4     53 0.2689
## S:V        1 0.94029     0.84      4     53 0.5052
## Residuals 56
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
print("")
```

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

```
# Roy's test
print("Wilks' test")
```

```
## [1] "Wilks' test"
```

```
summary(manova(cbind(Y1, Y2, Y3, Y4) ~ S*V, data=t_6.18), test="Roy")
```

```
##           Df     Roy approx F num Df den Df Pr(>F)
## S          1 28.5524   378.32      4     53 <2e-16 ***
## V          1  0.1008     1.34      4     53 0.2689
## S:V        1  0.0635     0.84      4     53 0.5052
## Residuals 56
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
print("")
```

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

```
# Pillai test
print("Pillai test")
```

```
## [1] "Pillai test"
```

```
summary(manova(cbind(Y1, Y2, Y3, Y4) ~ S*V, data=t_6.18), test="Pillai")
```

```
##           Df  Pillai approx F num Df den Df Pr(>F)
## S          1 0.96616   378.32      4     53 <2e-16 ***
## V          1 0.09158     1.34      4     53 0.2689
## S:V        1 0.05971     0.84      4     53 0.5052
## Residuals 56
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
print("")
```

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

```
# Lawley-Hotelling Test
print("Lawley-Hotelling' test")
```

```
## [1] "Lawley-Hotelling' test"
```

```
summary(manova(cbind(Y1, Y2, Y3, Y4) ~ S*V, data=t_6.18), test="Hotelling-Lawley")
```

```
##           Df Hotelling-Lawley approx F num Df den Df Pr(>F)
## S          1          28.5524   378.32      4     53 <2e-16 ***
## V          1           0.1008     1.34      4     53 0.2689
## S:V        1           0.0635     0.84      4     53 0.5052
## Residuals 56
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
print("")
```

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

Analysis

According to our 4 tests, we can conclude there are differences among our sowing dates across all measured variables. Variety, however, does not have significance, which implies and is determined to be ineffective at their cross product. This means sowing date and variety have no difference among any variable. Extra simply, varieties are independent of each other.