Chapter 6: Multivariate Analysis of Variance

DA 410

Marjorie Blanco

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: y_1 = aroma, y_2 = flavor, y_3 = texture, and y_4 = moisture. The data are in Table 6.17. Each entry is an average score for the judges on that fish.

```r
fish <- read_table2("Software-Files/T6_17_FISH.CSV",
                    col_names = c("Method", "y1", "y2", "y3", "y4", "y5"),
                    col_types = cols(y5 = col_skip()))
```

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

	x
y1	5.205556
y2	5.266667
y3	5.552778
y4	6.030556

	1	2	3
y1	5.383333	5.258333	4.975000
y2	5.733333	5.233333	4.833333
y3	5.441667	5.308333	5.908333
y4	5.983333	5.875000	6.233333

H =

1.0505556	2.173333	-1.375556	-0.7602778
2.1733333	4.880000	-2.373333	-1.2566667
-1.3755556	-2.373333	2.382222	1.3844444
-0.7602778	-1.256667	1.384444	0.8105556

```r
# E = matrix(data = 0, nrow = 4, ncol = 4)
# for (i in 1:dim(E)[1]) {
#   for (j in 1:i) {
#     b <- c()
#     for (k in fish.group) {
#       a <- sum((k[,i] - mean(k[,i])) * (k[,j] - mean(k[,j])))
#       print(a)
#       b <- append(b, a)
#
#     }
#     E[i,j] <- sum(b)
```

```r
#     E[j,i] <- sum(b)
#   }
# }
# E
```

```r
fish.manova <- manova(cbind(fish$y1, fish$y2, fish$y3, fish$y4) ~ Method, data = fish)
fish.summary <- summary(fish.manova)
fish.summary
```

```
##           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
```

```r
fish.summary$SS
```

```
## $Method
##            [,1]      [,2]      [,3]       [,4]
## [1,]  1.0505556  2.173333 -1.375556 -0.7602778
## [2,]  2.1733333  4.880000 -2.373333 -1.2566667
## [3,] -1.3755556 -2.373333  2.382222  1.3844444
## [4,] -0.7602778 -1.256667  1.384444  0.8105556
## 
## $Residuals
##           [,1]     [,2]      [,3]      [,4]
## [1,] 13.408333 7.723333  8.675000  5.864167
## [2,]  7.723333 8.480000  7.526667  6.213333
## [3,]  8.675000 7.526667 11.607500  7.037500
## [4,]  5.864167 6.213333  7.037500 10.565833
```

We would then like to test if the properties are the same across the three sports.

$H_0 : \mu_1 = \mu_2$

H_1 : The two $\mu's$ are unequal

```r
summary(manova(cbind(fish$y1, fish$y2, fish$y3, fish$y4) ~ fish$Method), test = "Wilks")
```

```
##             Df   Wilks approx F num Df den Df    Pr(>F)    
## fish$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
```

The Wilks's test rejects the hypothesis H_0 that the mean vector for the three sports are equal.

```r
summary(manova(cbind(fish$y1, fish$y2, fish$y3, fish$y4) ~ fish$Method), test = "Roy")
```

```
##             Df    Roy approx F num Df den Df    Pr(>F)    
## fish$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
```

The Roy's test also rejects the hypothesis H_0 that the mean vector for the three sports are equal.

```r
summary(manova(cbind(fish$y1, fish$y2, fish$y3, fish$y4) ~ fish$Method), test = "Hotelling-Lawley")
```

```
##             Df Hotelling-Lawley approx F num Df den Df    Pr(>F)
## fish$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
```

The Hotelling-Lawley's test also rejects the hypothesis H_0 that the mean vector for the three sports are equal.

```r
summary(manova(cbind(fish$y1, fish$y2, fish$y3, fish$y4) ~ fish$Method), test = "Pillai")
```

```
##             Df  Pillai approx F num Df den Df    Pr(>F)
## fish$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
```

(b) Compute the following measures of multivariate association from Section 6.1.8 : r), rfg, A_A, ALH, A_P.

(c) Based on the eigenvalues, is the essential dimensionality of the space containing the mean vectors equal to 1 or 2?