Math 760

# Chapter 6 HW

Gabrielle Salamanca

April 24, 2024

## 

## 2. Using the information in Example 6.1, construct the 95% Bonferroni simultaneous intervals for the components of the mean difference vector . Compare the lengths of these intervals with those of the simultaneous intervals constructed in the example.

The 95% Bonferroni simultaneous interval formula is:

## Commercial Lab: ( -22.29112 , 3.571124 )

## State Lab of Hygiene: ( -5.472651 , 32.01265 )

From Example 6.1, the simultaneous intervals are:

We see that the simultaneous intervals from Example 6.1 are larger than the Bonferroni ones, but not by much.

## Simultaneous - Bonferroni, d1 lower: -0.1688761

## Simultaneous - Bonferroni, d1 upper: 0.1688761

## Simultaneous - Bonferroni, d2 lower: -0.2373493

## Simultaneous - Bonferroni, d2 upper: 0.2373493

## 

## 5. A researcher considered three indices measuring the severity of heart attacks. The values of these indices for n = 40 heart-attack patients arriving at a hospital emergency room produced the summary statistics

## xbar matrix:

## [,1]  
## [1,] 46.1  
## [2,] 57.3  
## [3,] 50.4

##   
## S matrix:

## [,1] [,2] [,3]  
## [1,] 101.3 63.0 71.0  
## [2,] 63.0 80.2 55.6  
## [3,] 71.0 55.6 97.4

### 

### All three indices are evaluated patient. Test the equality of mean indices using (6-16) with .

**vs**

We will be using (6-16) to decide if we accept or reject .

The formula is:

If our value is greater than the critical value, we will reject . That said, our critical value is:

## [1] 6.660417

Our contrast matrix will be:

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

Thus, our value is:

## [,1]  
## [1,] 90.49458

Because , we reject .

### 

### (b) Judge the differences in pairs of mean indices using 95% simultaneous confidence intervals. [See (6-18)]

The (6-18) formula is:

We’ll be finding these pairs: , , and .

## 95% simultaneous confidence intervals

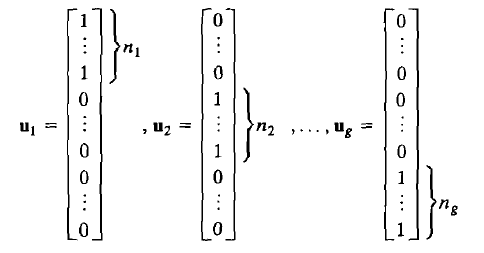
##   
## mu1 - mu2: ( -14.23996 , -8.160045 )

## mu1 - mu3: ( -7.6251 , -0.9748998 )

## mu2 - mu3: ( 3.5749 , 10.2251 )

The differences in pairs of mean indices are all quite different from each other when looking at these intervals.

**10. Consider the univariate one-way decomposition of the observation given by (6-34). Show that the mean vector is always perpendicular to the treatment effect vector where**



The formula of (6-34) is: . In writing, this is: observation = overall sample mean + estimated treatment effect + residual.

## 

## 16. Four measures of the response stiffness on each of 30 boards are listed in Table 4.3 (see Example 4.14). The measures, on a given board, are repeated int he sense that they were made one after another. Assuming that the measures of stiffness arise from four treatments, test for the equality of treatments in a repeated measures design context. Set . Construct a 95% (simultaneous) confidence interval for a contrast in the mean levels representing a comparison of the dynamic measurements with the static measurements.

**va**

Our contrast, C, matrix is:

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

We’ll be once again using (6-16) for this problem.

## The critical value is 9.53891

Now, we find our value and see if it’s greater than our critical value

## Our T-squared value is 254.7158

## Is it greater than 9.53891 ?: TRUE

Because , we reject at level.

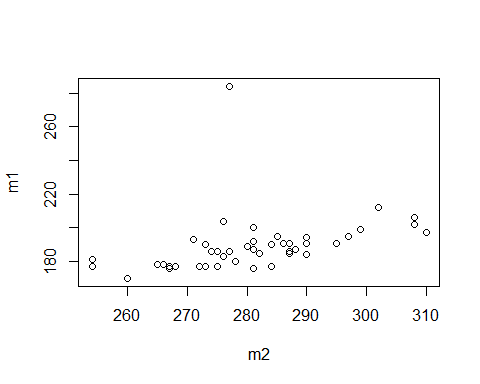
Now, we will construct our 95% simultaneous confidence interval for dynamic and static measurements.

## The 95% simultaneous confidence interval is ( 247.0397 , 596.027 )

## 

## 20. The tail lengths in measurements () and wing lengths in millimeters (x\_2) for 45 male hook-billed kites are given in Table 6.11 on page 346. Similar measurements for female hook-billed kites were given in Table 5.12.

### (a) Plot the male hook-billed kite data as a scatter diagram, and (visually) check for outliers. (Note, in particular, observation 31 with )



Only one observation stands out:

## V1 V2  
## 31 284 277

### (b) Test for equality of mean vectors for the populations of male and female hook-billed kites. Set . If is rejected, find the linear combination most responsible for the rejection of . (You may want to eliminate any outliers found in (a) for the male hook-billed data before conducting this test. Alternatively, you may want to interperet for observation 31 as a misprint and conduct the test with for this observation. Does it make any difference in this case how observation 31 for the male hook-billed kite data is treated?)

We will have two male datasets: one with observation 31 as a misprint, and one without observation 31.

**vs**

##   
## Hotelling's two sample T2-test  
##   
## data: male1 and female  
## T.2 = 6.713, df1 = 2, df2 = 87, p-value = 0.001944  
## alternative hypothesis: true location difference is not equal to c(0,0)

##   
## Hotelling's two sample T2-test  
##   
## data: male2 and female  
## T.2 = 12.339, df1 = 2, df2 = 86, p-value = 1.944e-05  
## alternative hypothesis: true location difference is not equal to c(0,0)

We see that with either male dataset, we will reject because .

We now will find the linear combination most responsible for this rejection. It is said that the coefficient vector for the linear combination most responsible for rejection is proportional to . Now, if we used the misprint dataset, there is a remark about datasets with the same sample size:

**Remark:** If , then

, so

With this in mind, our linear combination is:

## [,1]  
## [1,] -2.1886841  
## [2,] 0.7353992

### 

### (c) Determine the 95% confidence region for and 95% simultaneous confidence intervals for the components of .

The 95% confidence region formula for is from **(6-30)**:

## The left side of the inequality is 4.142067

## The right side of the inequality is 2.888072

The 95% confidence region is

For the 95% simultaneous intervals for the components of , we get:

Tail length (): (-10.800065, -2.266602)

Wing length (): (-6.204043, 8.159598)

##   
## Welch Two Sample t-test  
##   
## data: male1$V1 and female$V1  
## t = -3.0441, df = 85.679, p-value = 0.003099  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -10.800065 -2.266602  
## sample estimates:  
## mean of x mean of y   
## 187.0889 193.6222

##   
## Welch Two Sample t-test  
##   
## data: female$V2 and male1$V2  
## t = 0.27088, df = 81.231, p-value = 0.7872  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -6.204043 8.159598  
## sample estimates:  
## mean of x mean of y   
## 279.7778 278.8000

### 

### (d) Are male or female birds generally larger.

Female birds are generally larger, though more particularly with the tails as shown with the confidence intervals:

Tail length (): (-10.800065, -2.266602)

Wing length (): (-6.204043, 8.159598)

Wing length does include 0, which means no significance difference; but tail length is all negative.

## 25. Construct a one-way MANOVA of the crude-oil data listed in Table 11.7 on page 662. Construct 95% simultaneous confidence intervals to determine which mean components differ among the populations. (You may want to consider transformations of the data to make them more closely conform to the usual MANOVA assumptions.)

Without transforming the data, here are our results:

**The mean components do not differ among populations**

**The mean components do differ among populations**

## Df Wilks approx F num Df den Df Pr(>F)   
## pie 2 0.11591 18.985 10 98 < 2.2e-16 \*\*\*  
## Residuals 53   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Because the p-value is extremely small, we reject , which means the mean components do differ among populations.

The 95% simultaneous confidence intervals will be based on groups in V6: Wilhelm, SubMuli, and Upper. The intervals are:

## tau1 - tau2

## tau1[1]-tau2[1] belongs to (-2.845798,0.412032)  
## tau1[2]-tau2[2] belongs to (2.040458,18.92058)  
## tau1[3]-tau2[3] belongs to (-0.2759328,0.1684003)  
## tau1[4]-tau2[4] belongs to (-0.7699259,1.239536)  
## tau1[5]-tau2[5] belongs to (3.570383,8.542344)

##   
## tau1 - tau3

## tau1[1]-tau3[1] belongs to (-5.383456,-2.612032)  
## tau1[2]-tau3[2] belongs to (14.13887,28.49873)  
## tau1[3]-tau3[3] belongs to (-0.5039587,-0.1259662)  
## tau1[4]-tau3[4] belongs to (1.282835,2.992277)  
## tau1[5]-tau3[5] belongs to (3.657291,7.88692)

##   
## tau2 - tau3

## tau2[1]-tau3[1] belongs to (-3.934361,-1.627362)  
## tau2[2]-tau3[2] belongs to (4.861534,16.81502)  
## tau2[3]-tau3[3] belongs to (-0.4185211,-0.1038712)  
## tau2[4]-tau3[4] belongs to (1.191261,2.614241)  
## tau2[5]-tau3[5] belongs to (-2.04468,1.476164)

Of all the simultaneous confidence intervals, the ones that don’t cover 0 are:

This means there are significant differences between the three different groups, especially between Wilhelm-Upper and SubMuli-Upper.

## 29. Using the data on bone mineral content in Table 1.8, investigate equality between the dominant and nondominant bones.

### (a) Test using

**vs**

##   
## Hotelling's two sample T2-test  
##   
## data: dom and weak  
## T.2 = 0.29523, df1 = 3, df2 = 46, p-value = 0.8286  
## alternative hypothesis: true location difference is not equal to c(0,0,0)

Because , we fail to reject at 95% significance.

### Construct 95% simultaneous confidence intervals for the mean differences.

## Radius: ( -0.01458661 , 0.06554661 )

## Humerus Height: ( -0.0237914 , 0.1394714 )

## Ulna: ( -0.04262241 , 0.06374241 )

### 

### Contruct the Bonferroni 95% simultaneous intervals, and compare these with the intervals in (b).

## Radius: ( -0.008338575 , 0.05929857 )

## Humerus Height: ( -0.01106171 , 0.1267417 )

## Ulna: ( -0.03432909 , 0.05544909 )

## 33. Refer to Exercise 6.32. The data in Table 6.18 are measurements on the variables [percent spectral reflectance at wavelength 560 nm (green)], [percent spectral reflectance at wavelength 720 nm (near infrared)] for three species (sitka spruce [SS], Japanese larch [JL], and lodgepole pine [LP]) of 1-year-old seedlings taken at three different times (Julian day 150 [1], Julian day 235 [2], and Julian day 320 [3]) during the growing season. The seedlings were all grown with the optimal level of nutrient.

### 

### Perform a two-factor MANOVA using the data in Table 6.18. Test for a species effect, a time effect, and species-time interaction. Use .

**There is a species effect vs There is no species effect.**

## Df Wilks approx F num Df den Df Pr(>F)   
## x3 2 0.67704 3.4452 4 64 0.013 \*  
## Residuals 33   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Because , we reject ; there is no species effect.

**There is a time effect vs There is no time effect.**

## Df Wilks approx F num Df den Df Pr(>F)   
## x4 1 0.48184 17.744 2 33 5.86e-06 \*\*\*  
## Residuals 34   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Because , we reject ; there is no time effect.

**There is a species-time effect vs There is no species-time effect.**

## Df Wilks approx F num Df den Df Pr(>F)   
## x3 2 0.32499 10.935 4 58 1.125e-06 \*\*\*  
## x4 1 0.26621 39.968 2 29 4.632e-09 \*\*\*  
## x3:x4 2 0.42843 7.653 4 58 5.063e-05 \*\*\*  
## Residuals 30   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Because for all variables, we reject ; there is no species-time effect.

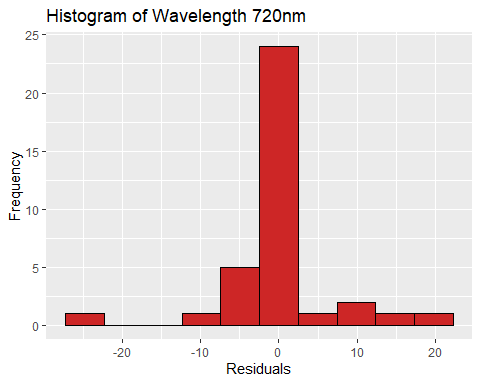
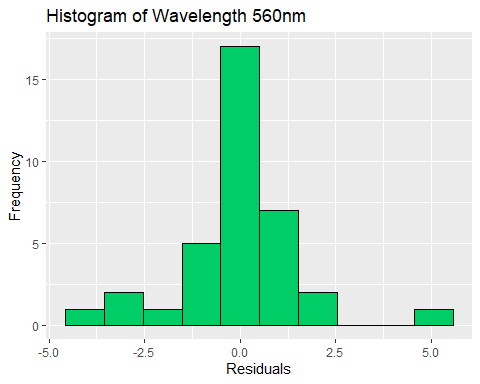
### 

### (b) Do you think the usual MANOVA assumptions are satisfied for these data? Discuss with reference to a residual analysis, and the possibility of correlated observations over time.

The MANOVA assumptions are:

1. We assume that the observations are independent of one another, that the sample is completely random.
2. We assume that the independent variables are categorical and dependent variables are continuous or scale variables.
3. There is an absence of multicollinearity between the dependent variables.
4. The data follows a multivariate normal.
5. The variance between groups is equal.

We can check the 4th assumption with the histograms of Wavelength 560nm and Wavelength 720nm below. There are a few outliers in the data, but overall, they approximately follow a normal distribution.



Then, we can check with the rest of the assumptions. The test is used for independence, and we see that the p-value is large, which means the 1st assumption does not hold. The correlation test is used to check for linearity, which can in turn tell us about the variance, because variance is part of the correlation equation. We see that p-value is small, which means there is multicollinearity between the dependent variables, which means variance is not equal between the groups. The 3rd and 5th assumptions do not hold.

## Warning in chisq.test(table(x1, x2)): Chi-squared approximation may be  
## incorrect

##   
## Pearson's Chi-squared test  
##   
## data: table(x1, x2)  
## X-squared = 1188, df = 1156, p-value = 0.2504

##   
## Pearson's product-moment correlation  
##   
## data: x1 and x2  
## t = 8.144, df = 34, p-value = 1.691e-09  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## 0.6611588 0.9009499  
## sample estimates:  
## cor   
## 0.8130816

### 

### Foresters are particularly interested in the interaction of species and time. Does interaction show up for one variable but not the other? Check by running a univariate two-factor ANOVA for each of the two responses.

## Analysis of Variance Table  
##   
## Response: x1  
## Df Sum Sq Mean Sq F value Pr(>F)   
## x3 2 965.18 482.59 169.973 5.027e-16 \*\*\*  
## factor(x4) 2 1275.25 637.62 224.578 < 2.2e-16 \*\*\*  
## x3:factor(x4) 4 795.81 198.95 70.073 7.341e-14 \*\*\*  
## Residuals 27 76.66 2.84   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## -------------------------------------------------------

## Analysis of Variance Table  
##   
## Response: x2  
## Df Sum Sq Mean Sq F value Pr(>F)   
## x3 2 2026.9 1013.43 15.4622 3.348e-05 \*\*\*  
## factor(x4) 2 5573.8 2786.90 42.5207 4.537e-09 \*\*\*  
## x3:factor(x4) 4 193.5 48.39 0.7383 0.5741   
## Residuals 27 1769.6 65.54   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

After running a univariate two-factor ANOVA for each response, the interaction between predictor variables occurs for wavelength 560nm (green), but not for wavelength 720nm (near infrared).

### 

### Can you think of another method of analyzing these data (or a different experimental design) that would allow for a potential time trend in the spectral reflectant numbers?

We could analyze the data by date. The time column goes from 1 to 3, so we can have 3 groups. Then, we can see if the spectral reflectance is either the same or not for all the species within that certain time.

**Code Appendix**

knitr::opts\_chunk$set(echo = FALSE)  
library(car)  
library(DescTools)  
library(dplyr)  
library(GGally)  
library(ggplot2)  
library(graphics)  
library(gridExtra)  
library(gtools)  
library(heplots)  
library(investr)  
library(matlib)  
library(MVN)  
library(mvnormtest)  
library(robustbase)  
library(SIBER)  
effluent <- stiff <- read.table("D:/Coding/R Storage/T6-1.dat", header = FALSE, sep = "")  
# mats  
dMat <- c(-9.36, 13.27)  
Smat <- c(199.26, 88.38, 88.38, 418.61)  
d <- matrix(dMat, nrow = 2, ncol = 1, byrow = TRUE)  
S <- matrix(Smat, nrow = 2, ncol = 2, byrow = TRUE)  
# vars  
n <- dim(effluent)[1]  
p <- dim(effluent)[2]  
crit <- -qt(0.05/(2\*p), df = n-1)  
# comm  
neg1 <- d[1] - crit\*sqrt((S[1,1]/n))  
pos1 <- d[1] + crit\*sqrt((S[1,1]/n))  
cat("Commercial Lab: (", neg1, ",", pos1, ") \n")  
# state  
neg2 <- d[2] - crit\*sqrt((S[2,2]/n))  
pos2 <- d[2] + crit\*sqrt((S[2,2]/n))  
cat("State Lab of Hygiene: (",neg2, ",", pos2, ")")  
eneg1 <- -22.46  
epos1 <- 3.74  
eneg2 <- -5.71  
epos2 <- 32.25  
cat("Simultaneous - Bonferroni, d1 lower: ", eneg1 - neg1, "\n")  
cat("Simultaneous - Bonferroni, d1 upper: ", epos1 - pos1, "\n")  
cat("Simultaneous - Bonferroni, d2 lower: ", eneg2 - neg2, "\n")  
cat("Simultaneous - Bonferroni, d2 upper: ", epos2 - pos2)  
n <- 40  
q <- 3  
# mats  
xbarMat <- c(46.1, 57.3, 50.4)  
Smat <- c(101.3,63.0,71.0,63.0,80.2,55.6,71.0,55.6,97.4)  
xbar <- matrix(xbarMat, nrow = 3, ncol = 1, byrow = TRUE)  
S <- matrix(Smat, nrow = 3, ncol = 3, byrow = TRUE)  
# print  
cat("xbar matrix: \n")  
xbar  
cat("\n S matrix: \n")  
S  
crit <- qf(0.05, df1 = q-1, df2 = n-q+1, lower.tail = FALSE)  
frac <- ((n-1)\*(q-1))/(n-q+1)  
right <- frac\*crit  
right  
Cmat <- c(1,-1,0,0,1,-1)  
C <- matrix(Cmat, nrow = 2, ncol = 3, byrow = TRUE)  
C  
Cx <- C %\*% xbar  
CSC <- C %\*% S %\*% t(C)  
T2 <- n %\*% t(Cx) %\*% inv(CSC) %\*% Cx  
T2  
m12 <- xbar[1] - xbar[2]  
m13 <- xbar[1] - xbar[3]  
m23 <- xbar[2] - xbar[3]  
crit1 <- sqrt(right)  
S1 <- (CSC/n)  
# intervals  
neg12 <- m12 - crit1\*sqrt(S1[1,1])   
neg13 <- m13 - crit1\*sqrt(S1[2,2])  
neg23 <- m23 - crit1\*sqrt(S1[2,2])  
pos12 <- m12 + crit1\*sqrt(S1[1,1])   
pos13 <- m13 + crit1\*sqrt(S1[2,2])  
pos23 <- m23 + crit1\*sqrt(S1[2,2])  
# print  
cat("95% simultaneous confidence intervals \n")  
cat("\n mu1 - mu2: (", neg12, ",", pos12, ") \n")  
cat("mu1 - mu3: (", neg13, ",", pos13, ") \n")  
cat("mu2 - mu3: (", neg23, ",", pos23, ") \n")  
stiff <- read.table("D:/Coding/R Storage/T4-3.dat", header = FALSE, sep = "")  
newStiff <- stiff[,1:4]  
# cols  
x1 <- newStiff$V1  
x2 <- newStiff$V2  
x3 <- newStiff$V3  
x4 <- newStiff$V4  
# vars  
n <- dim(newStiff)[1]  
q <- dim(newStiff)[2]   
Cmat <- c(1,-1,0,0,0,1,-1,0,0,0,1,-1)  
C <- matrix(Cmat, nrow = 3, ncol = 4, byrow = TRUE)  
C  
crit <- qf(0.05, df1 = q-1, df2 = n-q+1, lower.tail = FALSE)  
frac <- ((n-1)\*(q-1))/(n-q+1)  
right <- frac\*crit  
cat("The critical value is ", right)  
Stiff <- as.matrix(newStiff)  
# xbar  
one <- as.matrix(rep(1, n))  
xbar <- 1/n\*t(Stiff)%\*%one  
# S  
mean\_matrix <- matrix(data = 1, nrow = n)%\*%cbind(xbar[[1]], xbar[[2]], xbar[[3]], xbar[[4]])  
xstar <- Stiff - mean\_matrix  
S <- 1/(n-1) \* t(xstar) %\*% xstar  
# T2  
T2 <- n %\*% t(C %\*% xbar) %\*% inv(C %\*% S %\*% t(C)) %\*% (C %\*% xbar)  
ans <- T2 > right  
# print  
cat("Our T-squared value is ", T2, "\n")  
cat("Is it greater than ", right, "?: ", ans)  
cMat <- c(1,1,-1,-1)  
c <- matrix(cMat, nrow = 4, ncol = 1, byrow = TRUE)  
# ops  
before <- t(c) %\*% xbar   
after <- sqrt(right) \* sqrt((t(c) %\*% S %\*% c)/n)  
cat("The 95% simultaneous confidence interval is (", before - after, ",", before + after, ")")  
male <- read.table("D:/Coding/R Storage/T6-11.dat", header = FALSE, sep = "")  
female <- read.table("D:/Coding/R Storage/T5-12.dat", header = FALSE, sep = "")  
# cols  
m1 <- male$V1 # tail length  
m2 <- male$V2 # wing length  
f1 <- female$V1 # tail length  
f2 <- female$V2 # wing length  
plot(m2,m1)  
male[31,]  
male1 <- male  
male2 <- male  
male1[31,] <- 184  
male2 <- male2[-31,]  
HotellingsT2Test(x = male1, y = female)  
HotellingsT2Test(x = male2, y = female)  
n <- 45  
# xbar  
onem1 <- as.matrix(rep(1, n))  
onef <- as.matrix(rep(1, n))  
xbarm1 <- 1/n\*t(male1)%\*%onem1  
xbarf <- 1/n\*t(female)%\*%onef  
# S  
mean\_matrixm1 <- matrix(data = 1, nrow = n)%\*%cbind(xbarm1[[1]], xbarm1[[2]])  
mean\_matrixf <- matrix(data = 1, nrow = n)%\*%cbind(xbarf[[1]], xbarf[[2]])  
xstarm1 <- male1 - mean\_matrixm1  
xstarf <- female - mean\_matrixf  
Sm1 <- 1/(n-1) \* t(xstarm1) %\*% as.matrix(xstarm1)  
Sf <- 1/(n-1) \* t(xstarf) %\*% as.matrix(xstarf)  
# combo  
fracnum1 <- (n-1)\*Sm1 + (n-1)\*Sf  
fracdenom1 <- n + n - 2  
frac1 <- fracnum1/fracdenom1  
frac2 <- (1/n)+(1/n)  
Sp <- frac1\*frac2  
xbar <- xbarm1 - xbarf  
combo <- inv(Sp)%\*%xbar  
combo  
v <- 87  
p <- 2  
# crit  
crit <- qf(0.05, df1 = v, df2 = v-p+1, lower.tail = FALSE)  
frac <- (v\*p)/(v-p+1)  
right <- crit \* frac  
# mu  
mumV1 <- mean(male1$V1)  
mufV1 <- mean(female$V1)  
mumV2 <- mean(male1$V2)  
mufV2 <- mean(female$V2)  
mu1 <- mumV1 - mufV1  
mu2 <- mumV2 - mufV2  
mu <- mu1 - mu2  
# left  
mux <- xbar - mu  
left <- t(mux) %\*% inv(Sp) %\*% mux  
# print  
cat("The left side of the inequality is ", left, "\n")  
cat("The right side of the inequality is ", right)  
# t-test  
t.test(x = male1$V1, y = female$V1)  
t.test(x = female$V2, y = male1$V2)  
oil <- read.table("D:/Coding/R Storage/T11-7.dat", header = FALSE, sep = "")  
# cols  
x1 <- oil$V1  
x2 <- oil$V2  
x3 <- oil$V3  
x4 <- oil$V4  
x5 <- oil$V5  
pie <- oil$V6  
# model  
depend <- cbind(x1,x2,x3,x4,x5)  
model1 <- manova(depend ~ pie, data = oil)  
summary(model1, test = "Wilks")  
g <- 3  
n1 <- length(which(oil$V6 == "Wilhelm"))  
n2 <- length(which(oil$V6 == "SubMuli"))  
n3 <- length(which(oil$V6 == "Upper"))  
n <- n1+n2+n3  
alpha <- 0.05  
# subset  
Wilhelm <- oil[1:7,]  
Wilhelm <- Wilhelm[,-6]  
SubMuli <- oil[8:18,]  
SubMuli <- SubMuli[,-6]  
Upper <- oil[19:56,]  
Upper <- Upper[,-6]  
# xbar  
xbar1 <- colMeans(Wilhelm)  
xbar2 <- colMeans(SubMuli)  
xbar3 <- colMeans(Upper)  
xbar <- (n1\*xbar1 + n2\*xbar2 + n3\*xbar3)/n  
# cov  
S1 <- cov(Wilhelm)  
S2 <- cov(SubMuli)  
S3 <- cov(Upper)  
W <- (n1-1)\*S1 + (n2-1)\*S2 + (n2-1)\*S3   
# crit  
qtlevel <- qt(1-alpha/(p\*g\*(g-1)), df = n-g)  
# loop  
cat("tau1 - tau2 \n")  
for ( i in 1:5 ){  
# \tau\_{11}-\tau\_{21}  
LCI12 <- (xbar1[i]-xbar2[i])-qtlevel\*sqrt(W[i,i]/(n-g)\*(1/n1+1/n2))  
UCI12 <- (xbar1[i]-xbar2[i])+qtlevel\*sqrt(W[i,i]/(n-g)\*(1/n1+1/n2))  
cat("tau1[",i,"]-tau2[",i,"] belongs to (",LCI12,",",UCI12,")\n",sep="")  
}  
cat("\n tau1 - tau3 \n")  
for ( i in 1:5 ){  
# \tau\_{11}-\tau\_{31}  
LCI13 <- (xbar1[i]-xbar3[i])-qtlevel\*sqrt(W[i,i]/(n-g)\*(1/n1+1/n3))  
UCI13 <- (xbar1[i]-xbar3[i])+qtlevel\*sqrt(W[i,i]/(n-g)\*(1/n1+1/n3))  
cat("tau1[",i,"]-tau3[",i,"] belongs to (",LCI13,",",UCI13,")\n",sep="")  
}  
cat("\n tau2 - tau3 \n")  
for ( i in 1:5 ){  
# \tau\_{21}-\tau\_{31}  
LCI23 <- (xbar2[i]-xbar3[i])-qtlevel\*sqrt(W[i,i]/(n-g)\*(1/n2+1/n3))  
UCI23 <- (xbar2[i]-xbar3[i])+qtlevel\*sqrt(W[i,i]/(n-g)\*(1/n2+1/n3))  
cat("tau2[",i,"]-tau3[",i,"] belongs to (",LCI23,",",UCI23,")\n",sep="")  
}  
bone <- read.table("D:/Coding/R Storage/T1-8.dat", header = FALSE, sep = "")  
# vars  
x1 <- bone$V1 # dom radius  
x2 <- bone$V2 # radius  
x3 <- bone$V3 # dom humerus  
x4 <- bone$V4 # humerus  
x5 <- bone$V5 # dom ulna  
x6 <- bone$V6 # ulna  
# subset  
dom <- bone %>%  
 select(V1, V3, V5)  
weak <- bone %>%  
 select(V2, V4, V6)  
dom <- as.matrix(dom)  
weak <- as.matrix(weak)  
# new data  
radius <- x1-x2  
humerus <- x3-x4  
ulna <- x5-x6  
distance <- data.frame(radius, humerus, ulna)  
# dim  
n <- dim(distance)[1]  
p <- dim(distance)[2]  
# xbar  
one <- as.matrix(rep(1, n))  
xbar <- 1/n \* t(distance) %\*% one  
# S  
mean\_matrix <- matrix(data = 1, nrow = n1) %\*% cbind(xbar[[1]], xbar[[2]], xbar[[3]])  
xstar <- distance - mean\_matrix  
S <- 1/(n-1) \* t(xstar) %\*% as.matrix(xstar)  
HotellingsT2Test(x = dom, y = weak)  
# vars  
crit1 <- qf(0.05, df1 = p, df2 = n-p, lower.tail = FALSE)  
frac <- ((n-1)\*p)/(n-p)  
right <- sqrt(frac\*crit1)  
# Radius  
bneg1 <- xbar[1] - crit1\*sqrt((S[1,1]/n))  
bpos1 <- xbar[1] + crit1\*sqrt((S[1,1]/n))  
cat("Radius: (", bneg1, ",", bpos1, ") \n")  
# Humerus  
bneg2 <- xbar[2] - crit1\*sqrt((S[2,2]/n))  
bpos2 <- xbar[2] + crit1\*sqrt((S[2,2]/n))  
cat("Humerus Height: (", bneg2, ",", bpos2, ") \n")  
# Ulna  
bneg3 <- xbar[3] - crit1\*sqrt((S[3,3]/n))  
bpos3 <- xbar[3] + crit1\*sqrt((S[3,3]/n))  
cat("Ulna: (", bneg3, ",", bpos3, ") \n")  
# vars  
crit2 <- -qt(0.05/(2\*p), df = n-1)  
# Radius  
bneg1 <- xbar[1] - crit2\*sqrt((S[1,1]/n))  
bpos1 <- xbar[1] + crit2\*sqrt((S[1,1]/n))  
cat("Radius: (", bneg1, ",", bpos1, ") \n")  
# Humerus  
bneg2 <- xbar[2] - crit2\*sqrt((S[2,2]/n))  
bpos2 <- xbar[2] + crit2\*sqrt((S[2,2]/n))  
cat("Humerus Height: (", bneg2, ",", bpos2, ") \n")  
# Ulna  
bneg3 <- xbar[3] - crit2\*sqrt((S[3,3]/n))  
bpos3 <- xbar[3] + crit2\*sqrt((S[3,3]/n))  
cat("Ulna: (", bneg3, ",", bpos3, ") \n")  
specter <- read.table("D:/Coding/R Storage/T6-18.dat", header = FALSE, sep = "")  
# vars  
x1 <- specter$V1 # percent spectral reflectance at wavelength 560nm (green)  
x2 <- specter$V2 # percent spectral reflectance at wavelength 720nm (near infrared)  
x3 <- specter$V3 # species  
x4 <- specter$V4 # time  
x5 <- specter$V5 # replication  
test1 <- manova(cbind(x1, x2)~ x3, data = specter)  
summary(test1, test = 'Wilks')  
test2 <- manova(cbind(x1, x2)~ x4, data = specter)  
summary(test2, test = 'Wilks')  
test3 <- manova(cbind(x1, x2)~ x3\*x4, data = specter)  
summary(test3, test = 'Wilks')  
fit1 <- lm(x1 ~ x3\*factor(x4), data = specter)  
ggplot(data = specter, aes(x = fit1$residuals)) +  
 geom\_histogram(bins = 10, fill = 'springgreen3', color = 'black') +  
 labs(title = 'Histogram of Wavelength 560nm',   
 x = 'Residuals', y = 'Frequency')  
# fit 2  
fit2 <- lm(x2 ~ x3\*factor(x4), data = specter)  
ggplot(data = specter, aes(x = fit2$residuals)) +  
 geom\_histogram(bins = 10, fill = 'firebrick3', color = 'black') +  
 labs(title = 'Histogram of Wavelength 720nm',   
 x = 'Residuals', y = 'Frequency')  
chisq.test(table(x1,x2))  
cor.test(x1,x2)  
anova(fit1)  
cat("------------------------------------------------------- \n")  
anova(fit2)