

Homework 7

Chris Cameron

11/10/2021

A)

```
bluecrab <- read.csv("data/bluecrab.dat", sep = ' ', header = FALSE)
orangecrab <- read.csv("data/orangecrab.dat", sep = ' ', header = FALSE)
```

```
#Set Prior Parameters
mu_blue <- apply(bluecrab,2,mean)
mu_orange <- apply(orangecrab,2,mean)
L_blue <- cov(bluecrab)
L_orange <- cov(orangecrab)
S_blue = L_blue
S_orange = L_orange
nu_blue <- 7
nu_orange <- 7
iterations <- 10000
```

```
library(mvtnorm)
library(MCMCpack)
```

```
## Loading required package: coda
```

```
## Loading required package: MASS
```

```
## ##
```

```
## ## Markov Chain Monte Carlo Package (MCMCpack)
```

```
## ## Copyright (C) 2003-2021 Andrew D. Martin, Kevin M. Quinn, and Jong Hee Park
```

```
## ##
```

```
## ## Support provided by the U.S. National Science Foundation
```

```
## ## (Grants SES-0350646 and SES-0350613)
```

```
## ##
```

```

#rbind does not seem to be (properly) returnable in a list,
#which is why I am not using a function

mu <- mu_blue
L <- L_blue
S <- S_blue
nu <- nu_blue
Y <- bluecrab

n <- dim (Y)[1]
Sigma <- S #initialize to prior?

THETA_blue <- SIGMA_blue <- NULL
set.seed(6)

for(s in 1:iterations){

  #update theta
  Ln <- solve(solve(L) + n*solve(Sigma))
  mun <- Ln %*% (solve(L) %*% mu + n*solve(Sigma) %*% mu) #mu is equal to mean
  theta <- rmvnorm(1, mu, L)

  #update sigma
  Sn <- S + (t(Y) - c(theta)) %*% t(t(Y)-c(theta))

  Sigma <- solve(rwish(nu + n, solve(Sn)))

  #Save Results
  THETA_blue <- rbind(THETA_blue, theta)
  SIGMA_blue <- rbind(SIGMA_blue, c(Sigma))
}

```

```

mu <- mu_orange
L <- L_orange
S <- S_orange
nu <- nu_orange
Y <- orangecrab

n <- dim (Y)[1]
Sigma <- S #initialize to prior?

THETA_orange <- SIGMA_orange <- NULL

for(s in 1:iterations){

```

```

#update theta
Ln <- solve(solve(L) + n*solve(Sigma))
mun <- Ln %*% (solve(L) %*% mu + n*solve(Sigma) %*% mu) #mu is equal to mean
theta <- rmvnorm(1, mu, L)

#update sigma
Sn <- S + (t(Y) - c(theta)) %*% t(t(Y)-c(theta))

Sigma <- solve(rwish(nu + n, solve(Sn)))

#Save Results
THETA_orange <- rbind(THETA_orange, theta)
SIGMA_orange <- rbind(SIGMA_orange, c(Sigma))
}

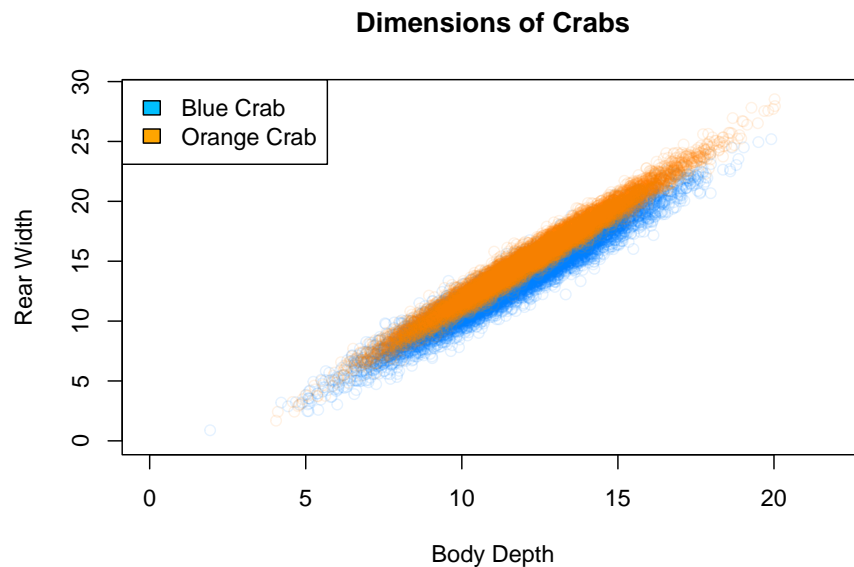
```

B)

```

plot(THETA_blue, col = rgb(red = 0, green = 0.5, blue = 1, alpha = 0.1),
     main = "Dimensions of Crabs", xlab = "Body Depth", ylab = "Rear Width",
     xlim = c(0, 22), ylim = c(0, 29))
points(THETA_orange, col = rgb(red = 1, green = 0.5, blue = 0, alpha = 0.1))
legend("topleft", c("Blue Crab", "Orange Crab"), fill = c("deepskyblue", "orange"))

```



#Citation: <https://stackoverflow.com/questions/14683766/adding-key-legend-to-multi-histogram-plot-in-r>

```
## [1] "Average Blue Crab Body Depth"
## [1] 11.13554
## [1] "Average Blue Crab Rear Width"
## [1] 10.2302
## [1] "Average Orange Crab Body Depth"
## [1] 11.36395
## [1] "Average Orange Crab Rear Width"
## [1] 12.10057
```

The above plot demonstrates that both orange and blue crabs demonstrate a positive linear relationship between their rear widths and body depths. The relationship between width and depth appears to have a similar slope for both species, with an orange crab having a larger rear width than a blue crab of equal body depth. Furthermore, orange crabs have rear widths that are 2.1 cm larger than that of blue crabs, and body depths that are 0.2 cm larger, demonstrating that orange crabs are larger on average.

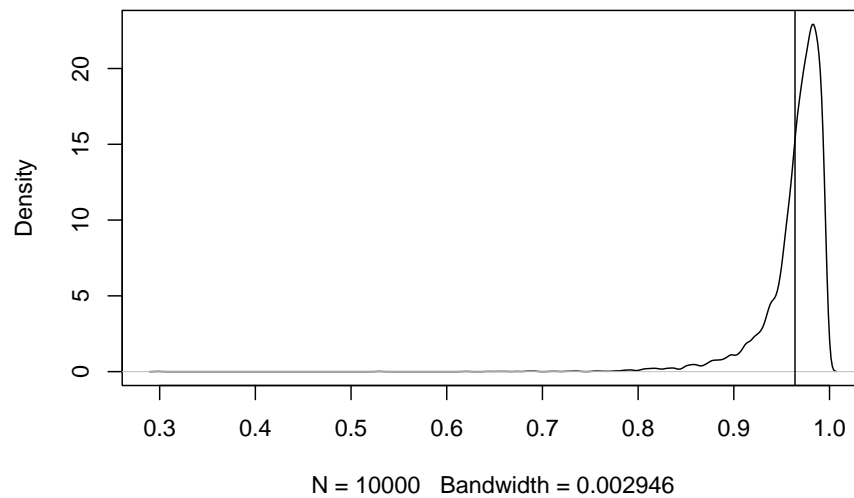
C)

```
blue_cor <- list()
for (x in 1:iterations){
  blue_cor[x] <- cov2cor(matrix(SIGMA_blue[x,1:4], nrow=2, ncol=2))[2]
}

orange_cor <- list()
for (x in 1:iterations){
  orange_cor[x] <- cov2cor(matrix(SIGMA_orange[x,1:4], nrow=2, ncol=2))[2]
}

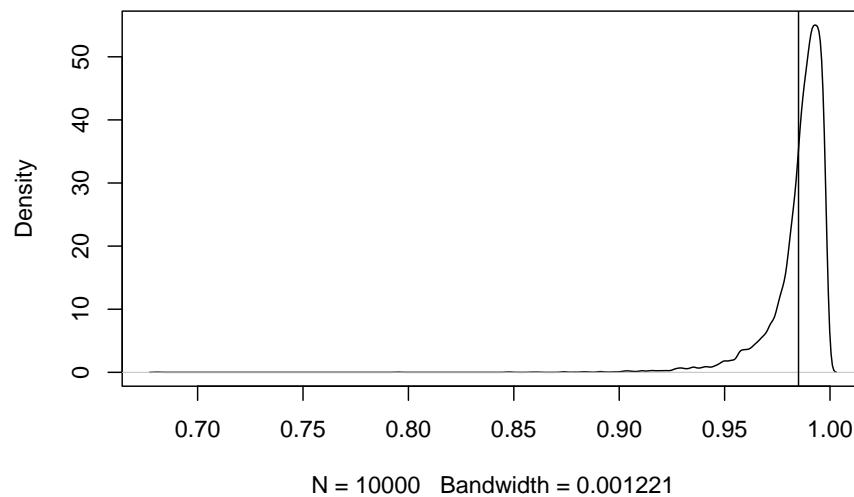
plot(density(unlist(blue_cor)),
     main = "Density of Correlation Between Width & Depth For Blue Crabs")
abline(v = mean(unlist(blue_cor)))
```

Density of Correlation Between Width & Depth For Blue Crabs



```
plot(density(unlist(orange_cor)),  
      main = "Density of Correlation Between Width & Depth For Orange Crabs")  
abline(v = mean(unlist(orange_cor)))
```

Density of Correlation Between Width & Depth For Orange Crabs



```

a = 0
for(i in 1:iterations){
  if(unlist(orange_cor)[i] > unlist(blue_cor)[i]){
    a <- a+1
  }
}
print(a/iterations)

## [1] 0.789

t.test(unlist(orange_cor), unlist(blue_cor), alternative = "less")

##
## Welch Two Sample t-test
##
## data: unlist(orange_cor) and unlist(blue_cor)
## t = 58.278, df = 13460, p-value = 1
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf 0.02170674
## sample estimates:
## mean of x mean of y
## 0.9850548 0.9639440

```

Both types of crabs have correlations between their body depths and rear widths that are very high, with the density of this correlation being skewed left for both. The average correlation for orange crabs is .985, while that for blue crabs is .964. The average correlation for any one sample of orange crabs has a probability of .789 of being higher than that of any one sample of blue crabs. The probability that the overall mean correlation of orange crabs is higher than that of blue crabs, using a t-test with 13460 degrees of freedom, is nearly 1. These values suggest that the correlation between correlation between body depth and rear width is on average higher in orange crabs than in blue crabs.