

Homework 7

STA 360: Assignment 7, Fall 2020

Due Monday November 2nd, 5 PM Standard Eastern Time

Question 1

Hoff 7. 3

Part A.

```
#Read in the data sets from github
blue_crab_data = as.matrix(read.table(url('https://raw.githubusercontent.com/resteorts/modern-bayes/master/data/BlueCrabData.csv')))
orange_crab_data = as.matrix(read.table(url('https://raw.githubusercontent.com/resteorts/modern-bayes/master/data/OrangeCrabData.csv')))

#We need to sample from the posterior for both data sets, so let's use lapply with a function
#The function is heavily based on Rebecca Steort's code in Module 8
crab.mcmc = lapply(list('blue_crab' = blue_crab_data, 'orange_crab' = orange_crab_data), function(crab) {
  n = nrow(crab)
  # Set prior parameters according to problem
  mu0 = colMeans(crab)
  lambda0 = cov(crab)
  s0 = cov(crab)
  nu0 = 4
  THETA <- SIGMA <- CORR <- NULL
  # Start with sigma sample
  sigma = s0
  for (s in 1:10000) {
    # Update theta
    Ln <- solve(solve(lambda0) + n * solve(sigma))
    mun <- Ln %*% (solve(lambda0) %*% mu0 + n * solve(sigma) %*% mu0)
    theta <- rmvnorm(1, mun, Ln)

    ## update sigma
    Sn <- s0 + (t(crab) - c(theta)) %*% t(t(crab) - c(theta))
    sigma <- solve(rwish(nu0 + n, solve(Sn)))

    #Correlation Coefficient
    corr <- sigma[2]/(sqrt(sigma[1] * sigma[4]))

    ## save results
    THETA <- rbind(THETA, theta)
    SIGMA <- rbind(SIGMA, c(sigma))
    CORR <- rbind(CORR, corr)
  }

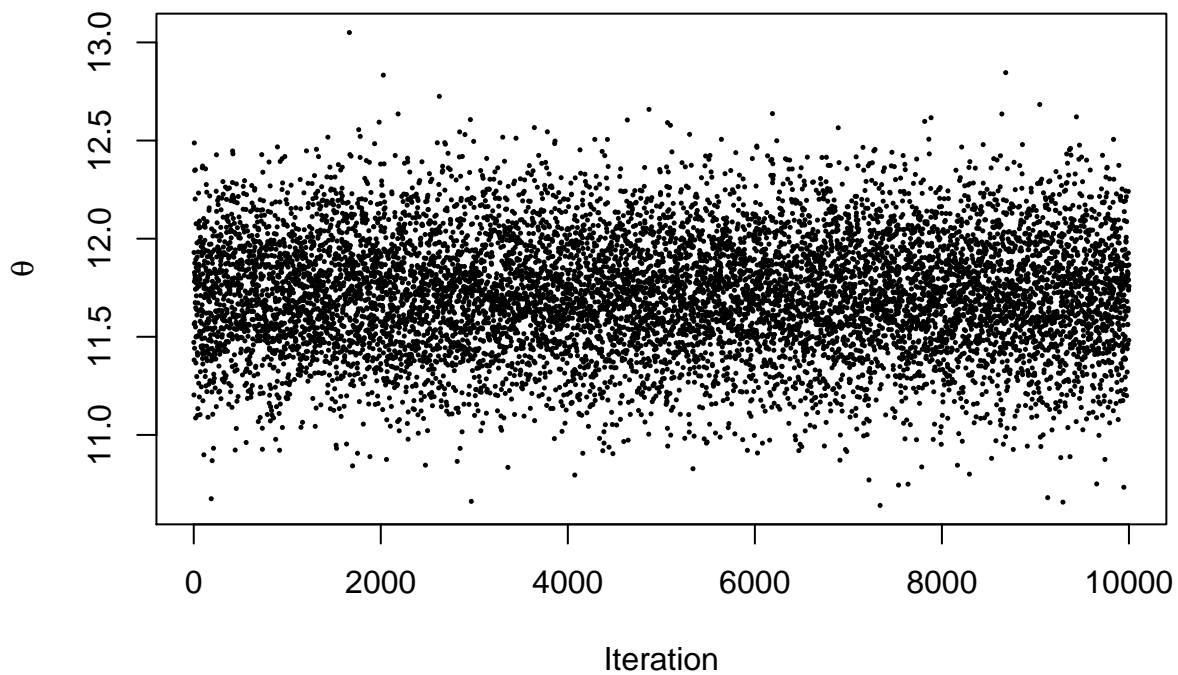
  list(theta = THETA, sigma = SIGMA, corr = CORR)
})
```

```
})
```

Part B.

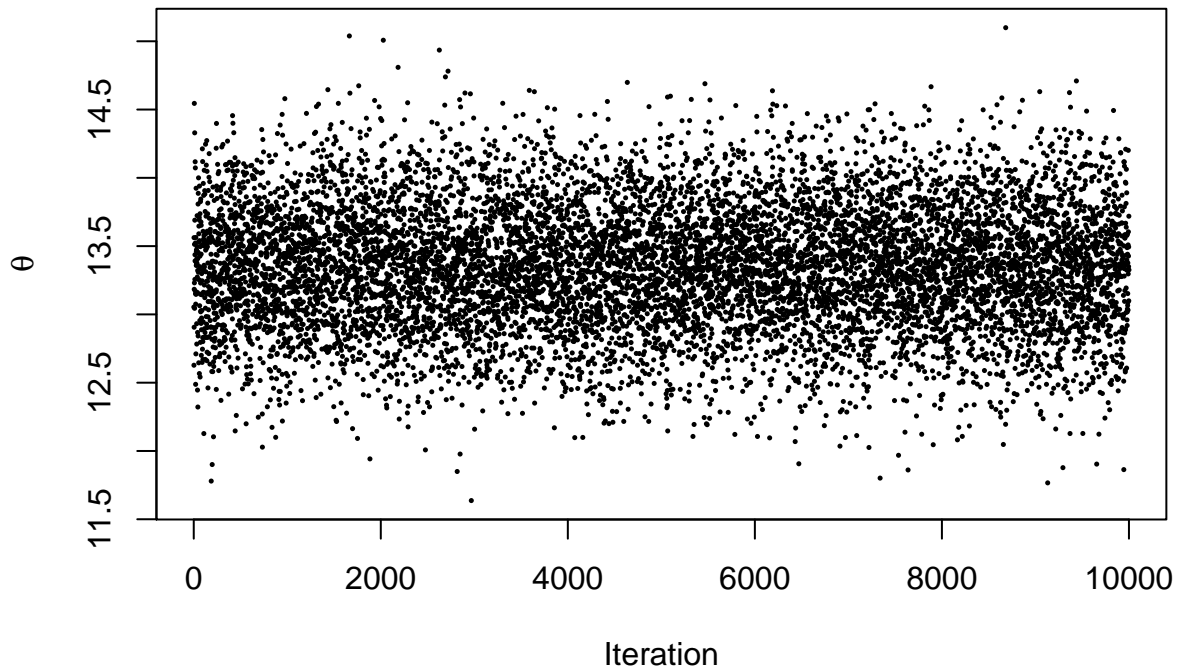
```
#Create Traceplots for Theta 1 and 2 for Orange and Blue
n.iter<-10000
plot(1:n.iter, crab.mcmc$blue_crab$theta[,1], pch = 16, cex = .35,
     xlab = "Iteration", ylab = expression(theta),
     main = expression(paste("Traceplot of ", theta, "1 for Blue Crab")))
```

Traceplot of θ_1 for Blue Crab



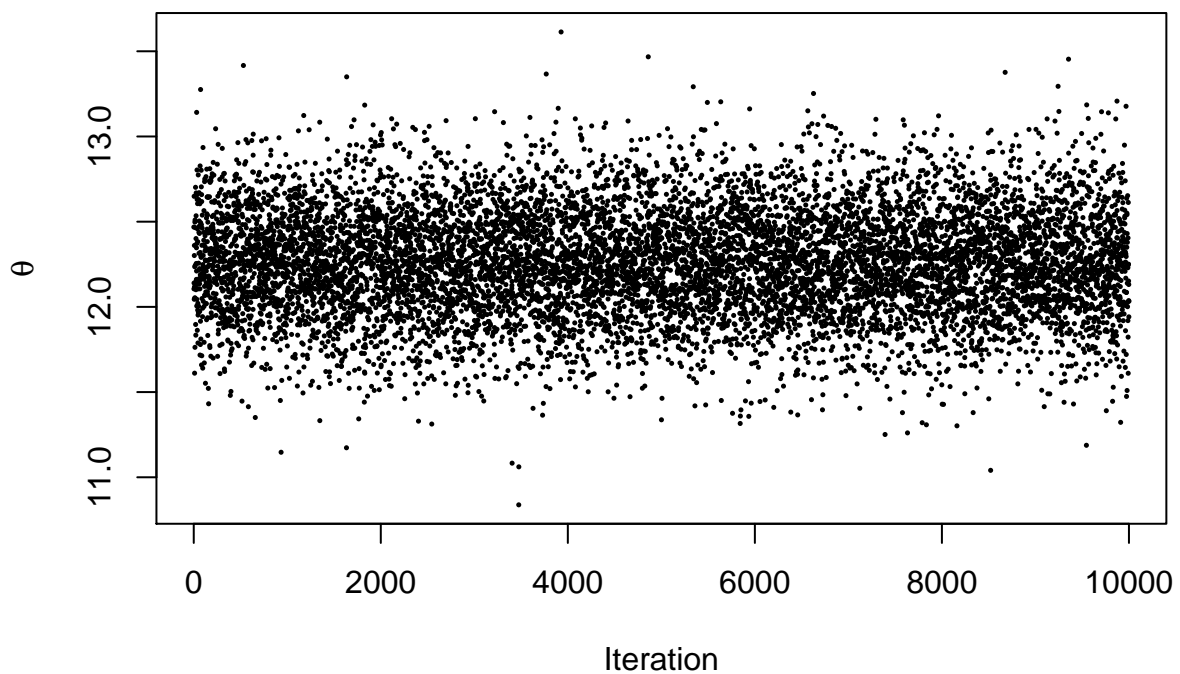
```
n.iter<-10000
plot(1:n.iter, crab.mcmc$blue_crab$theta[,2], pch = 16, cex = .35,
     xlab = "Iteration", ylab = expression(theta),
     main = expression(paste("Traceplot of ", theta, "2 for Blue Crab")))
```

Traceplot of θ_2 for Blue Crab



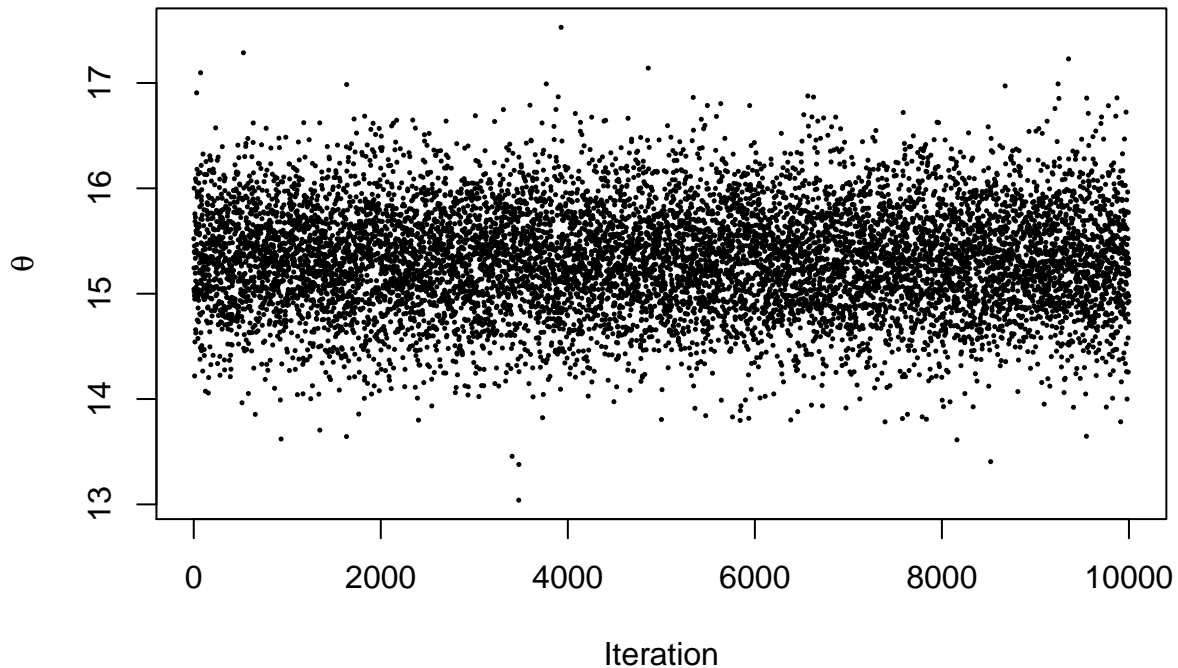
```
n.iter<-10000
plot(1:n.iter, crab.mcmc$orange_crab$theta[,1], pch = 16, cex = .35,
     xlab = "Iteration", ylab = expression(theta),
     main = expression(paste("Traceplot of ", theta, "1 for Orange Crab")))
```

Traceplot of θ_1 for Orange Crab



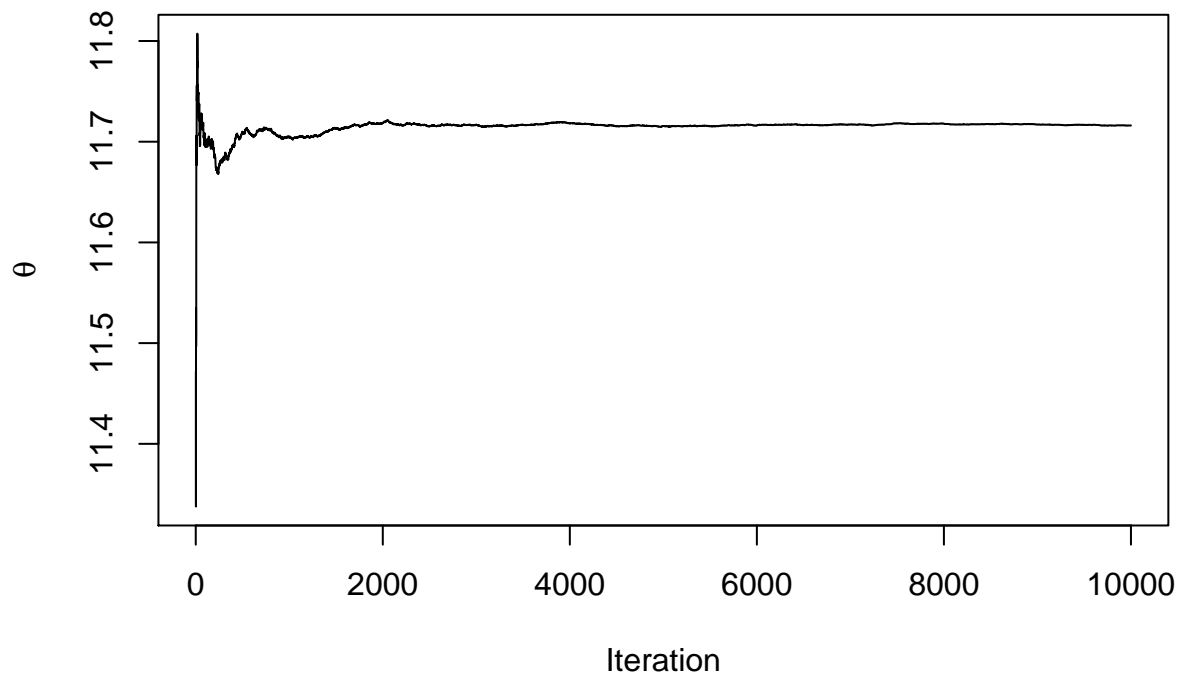
```
n.iter<-10000
plot(1:n.iter, crab.mcmc$orange_crab$theta[,2], pch = 16, cex = .35,
     xlab = "Iteration", ylab = expression(theta),
     main = expression(paste("Traceplot of ", theta, "2 for Orange Crab")))
```

Traceplot of θ_2 for Orange Crab



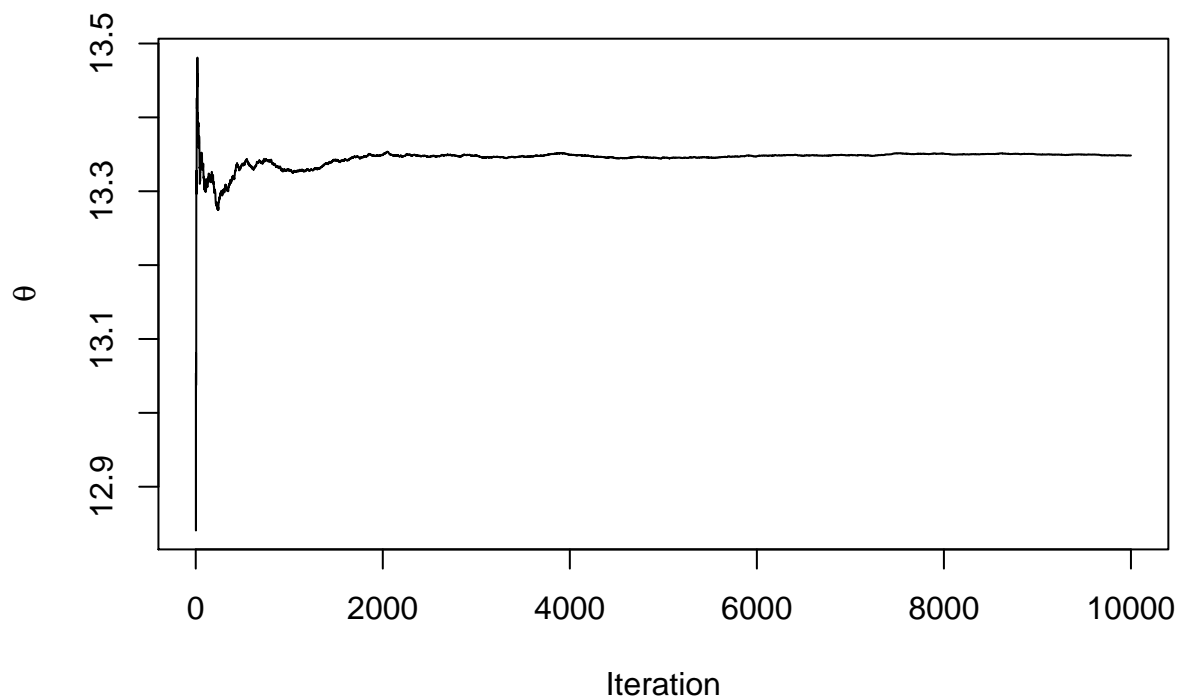
```
run.avg <- apply(crab.mcmc$blue_crab$theta, 2, cumsum)/(1:n.iter)
plot(1:n.iter, run.avg[,1], type = "l",
     xlab = "Iteration", ylab = expression(theta),
     main = expression(paste("Running Average Plot of ", theta, "1 for Blue Crab")))
```

Running Average Plot of θ_1 for Blue Crab



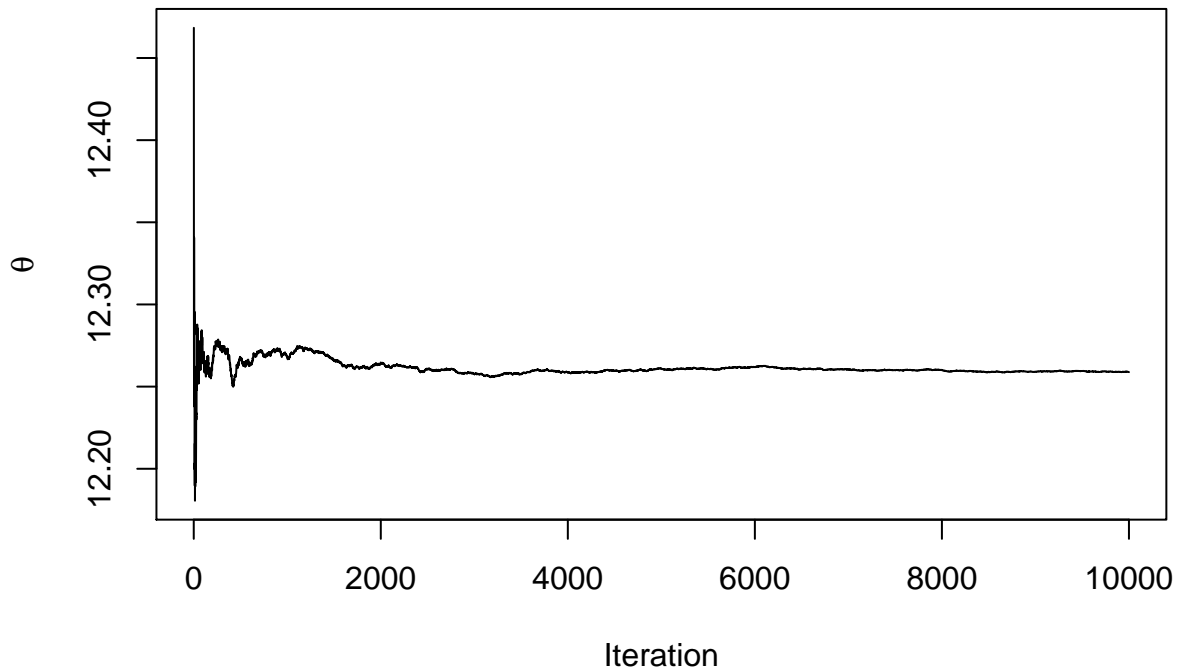
```
plot(1:n.iter, run.avg[,2], type = "l",  
     xlab = "Iteration", ylab = expression(theta),  
     main = expression(paste("Running Average Plot of ", theta, "2 for Blue Crab")))
```

Running Average Plot of θ_2 for Blue Crab



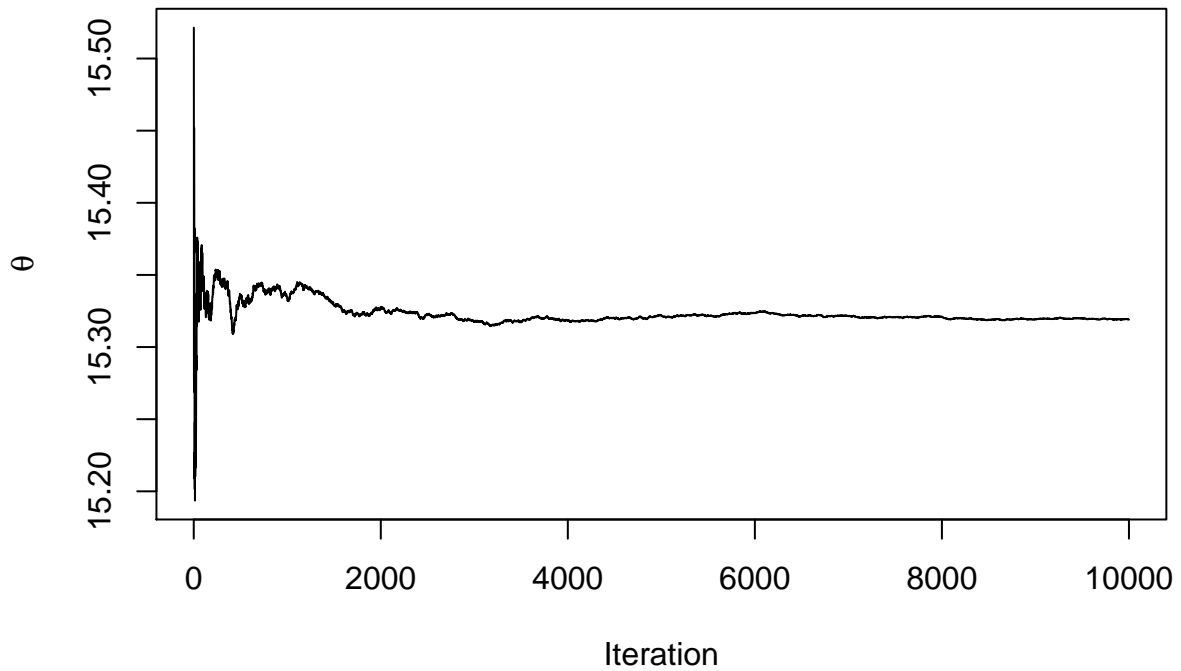
```
run.avg2 <- apply(crab.mcmc$orange_crab$theta, 2, cumsum)/(1:n.iter)
plot(1:n.iter, run.avg2[,1], type = "l",
     xlab = "Iteration", ylab = expression(theta),
     main = expression(paste("Running Average Plot of ", theta, "1 for Orange Crab")))
```

Running Average Plot of θ_1 for Orange Crab



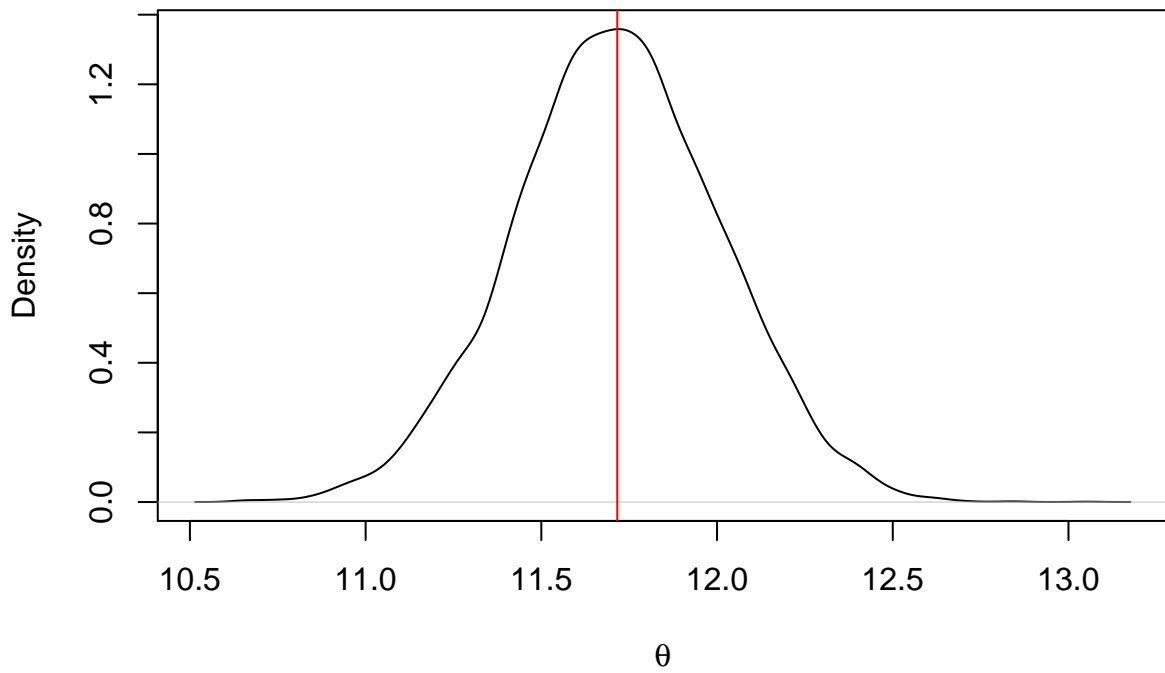
```
plot(1:n.iter, run.avg2[,2], type = "l",
     xlab = "Iteration", ylab = expression(theta),
     main = expression(paste("Running Average Plot of ", theta, "2 for Orange Crab")))
```

Running Average Plot of θ_2 for Orange Crab



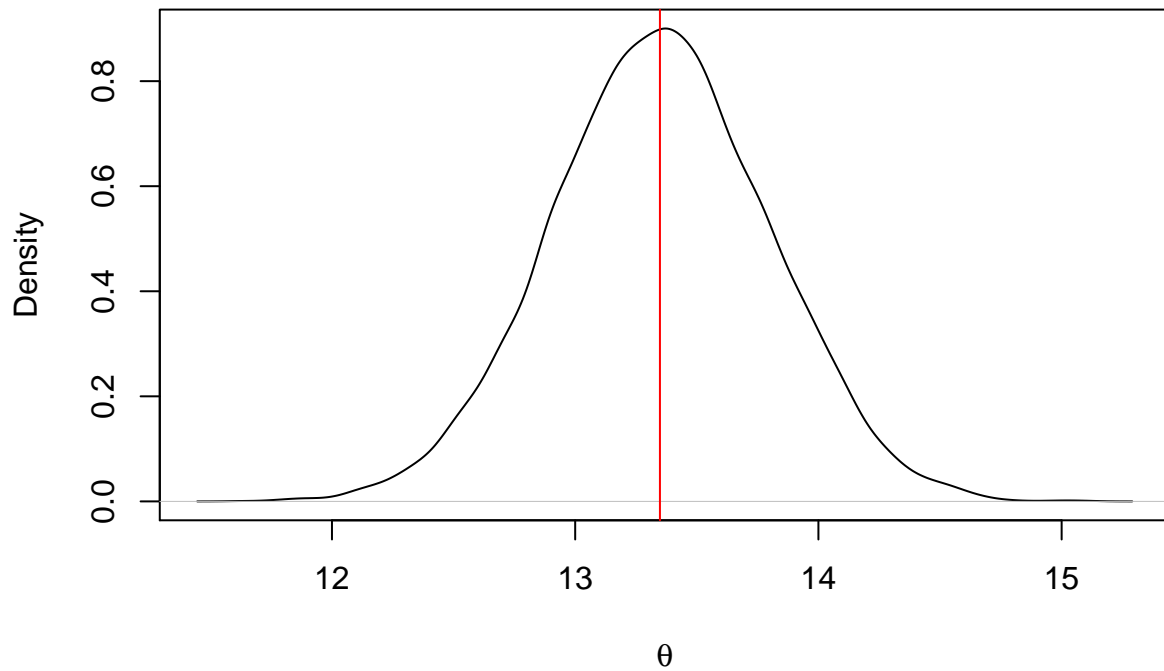
```
plot(density(crab.mcmc$blue_crab$theta[,1]), xlab = expression(theta),  
     main = expression(paste("Density of ", theta, "1 for Blue Crab"))  
abline(v = mean(crab.mcmc$blue_crab$theta[,1]), col = "red")
```

Density of θ_1 for Blue Crab



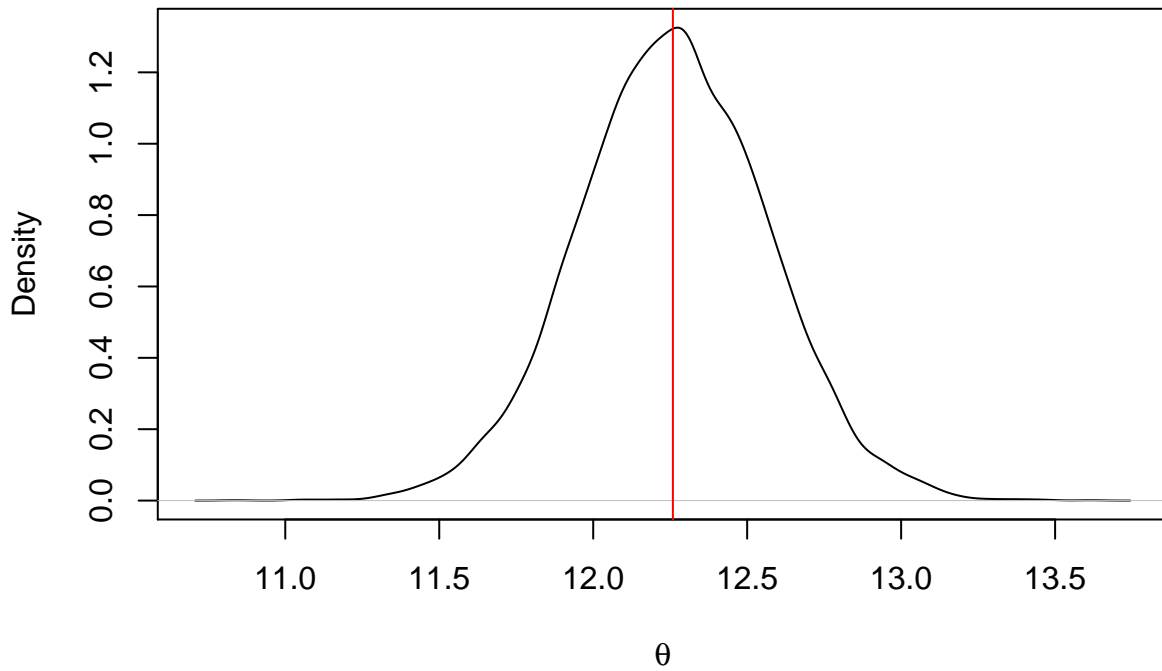
```
plot(density(crab.mcmc$blue_crab$theta[,2]), xlab = expression(theta),
     main = expression(paste("Density of ", theta, "2 for Blue Crab")))
abline(v = mean(crab.mcmc$blue_crab$theta[,2]), col = "red")
```

Density of θ_2 for Blue Crab



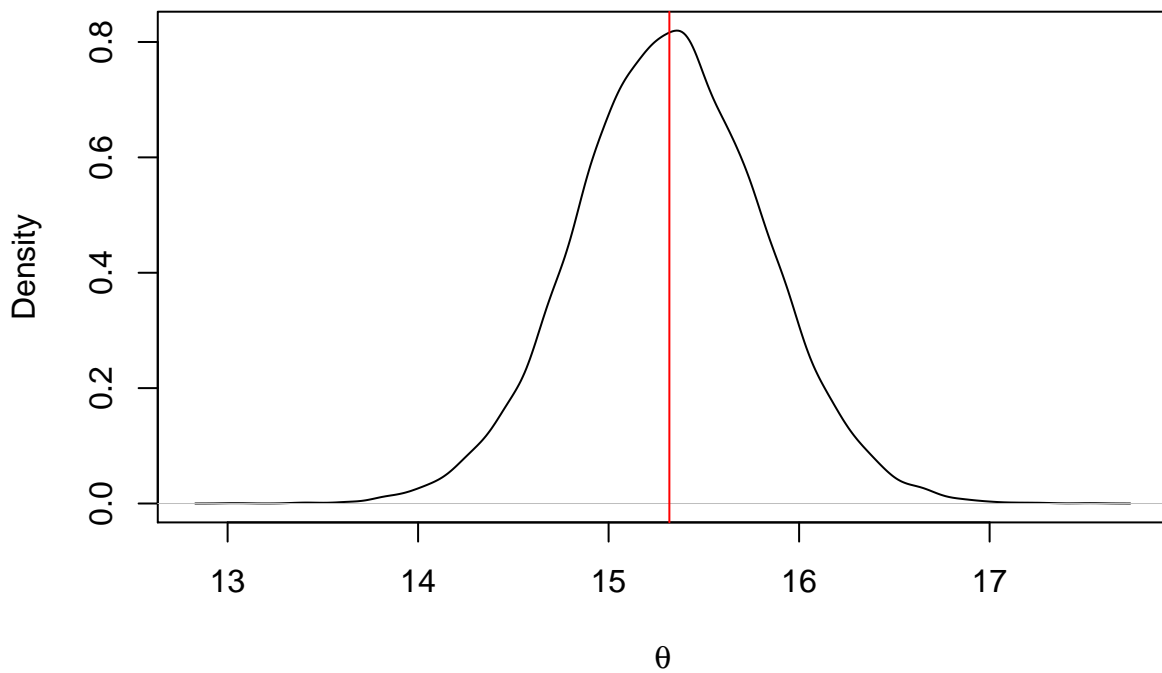
```
plot(density(crab.mcmc$orange_crab$theta[,1]), xlab = expression(theta),
     main = expression(paste("Density of ", theta, "1 for Orange Crab")))
abline(v = mean(crab.mcmc$orange_crab$theta[,1]), col = "red")
```


Density of θ_1 for Orange Crab



```
plot(density(crab.mcmc$orange_crab$theta[,2]), xlab = expression(theta),  
     main = expression(paste("Density of ", theta, "2 for Orange Crab"))  
abline(v = mean(crab.mcmc$orange_crab$theta[,2]), col = "red")
```

Density of θ_2 for Orange Crab



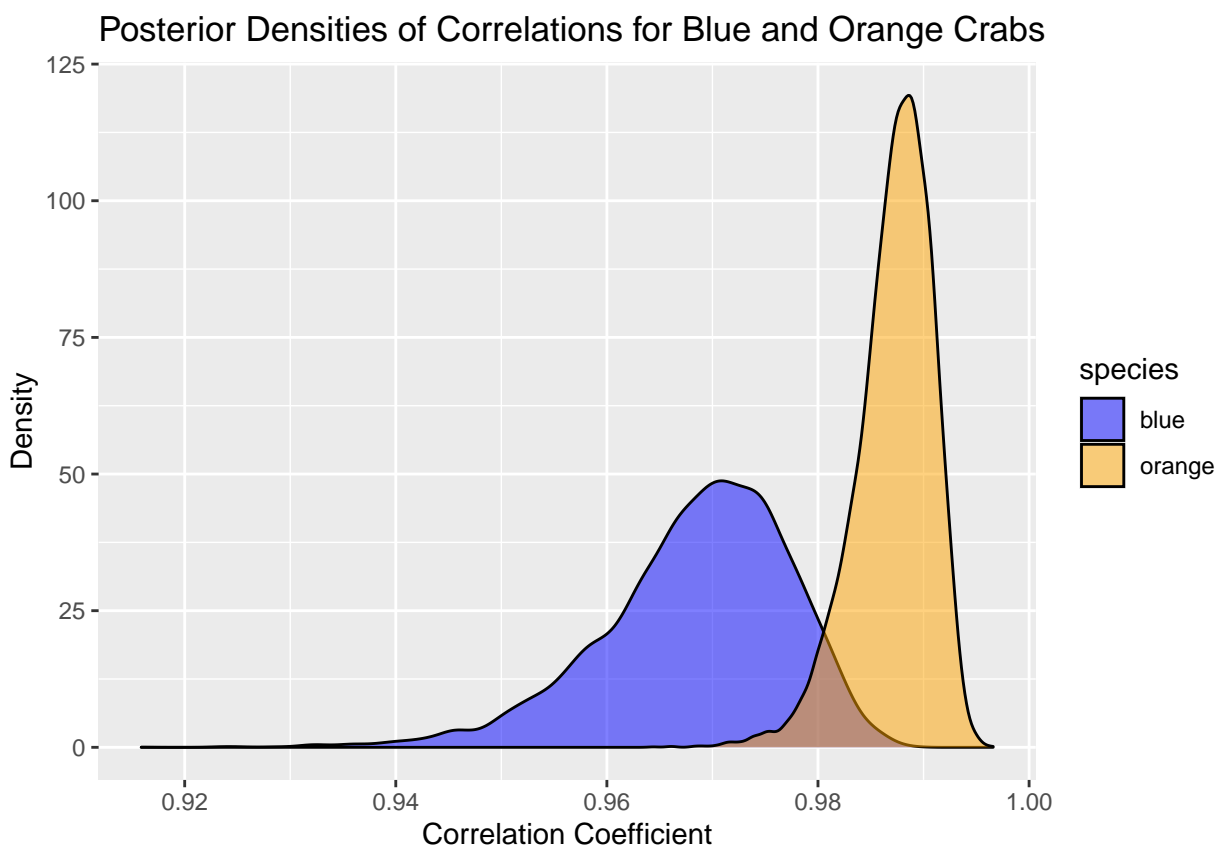
From the plots above, we can see that the chain has seemingly not failed to converge. And we can see that

for body width and for rear width, the orange crabs seem to be bigger. The density and running average plots both reveal, that the orange crabs have larger values for body with and rear width.

Part C.

```
#Creating Data Frame for the plot
cor.df = data.frame(species = c(rep('blue', length(crab.mcmc$blue_crab$corr)), rep('orange', length(crab.mcmc$orange_crab$corr))),
  cor = c(crab.mcmc$blue_crab$corr, crab.mcmc$orange_crab$corr))

#Plotting the values
ggplot(cor.df, aes(x = cor, fill = species)) +
  geom_density(alpha = 0.5) +
  labs(x = "Correlation Coefficient",
    y = "Density") +
  ggtitle("Posterior Densities of Correlations for Blue and Orange Crabs") +
  scale_fill_manual(values = c('blue', 'orange'))
```



```
#Finding the posterior probability
p = mean(crab.mcmc$blue_crab$corr < crab.mcmc$orange_crab$corr)
p
```

```
## [1] 0.9871
```

With our approximated posterior probability found, we can say that the posterior probability that the correlation between body depth and rear width for blue crabs is smaller than that for orange crabs is 0.9871. Which in general means, that for both measurements, orange crabs are most likely to be bigger than blue crabs.