hw3

Sharon Zhao

1/27/2022

Q2

The null hypothesis is that the population means of all six variables are the same for the genuine and counterfeit bank notes. The alternative hypothesis is that they are different. We can assume that they have the same population variance-covariance matrices.

```
library(glue)
banknotes <- read.delim("SwissBankNotes.txt", sep = "")
df <- banknotes
genuine <- df[0:100,]
counterfeit <- df[101:200,]
n <- 100
delta <- colMeans(genuine) - colMeans(counterfeit)
p <- 6
Sg <- cov(genuine)
Sc <- cov(counterfeit)
S_pooled <- ((n-1)*Sg + (n-1)*Sc) / (n+n-2)
t_squared <- (n*n)/(n+n) * t(delta) %*% solve(S_pooled) %*% (delta)
f_statistic <- t_squared * (n+n-p-1)/(p*(n+n-2))
cv <- qf(0.05, p, n+n-p-1, lower.tail = FALSE)
print(glue("The F statistic is {f_statistic} and the critical value is {cv}"))
```

```
## The F statistic is 391.921702277771 and the critical value is 2.14580146767029
```

Since the F statistic is larger than the critical value, we are able to reject the null hypothesis. We have evidence that the counterfeits are distinguishable from the genuine bank notes on at least one of the variables.