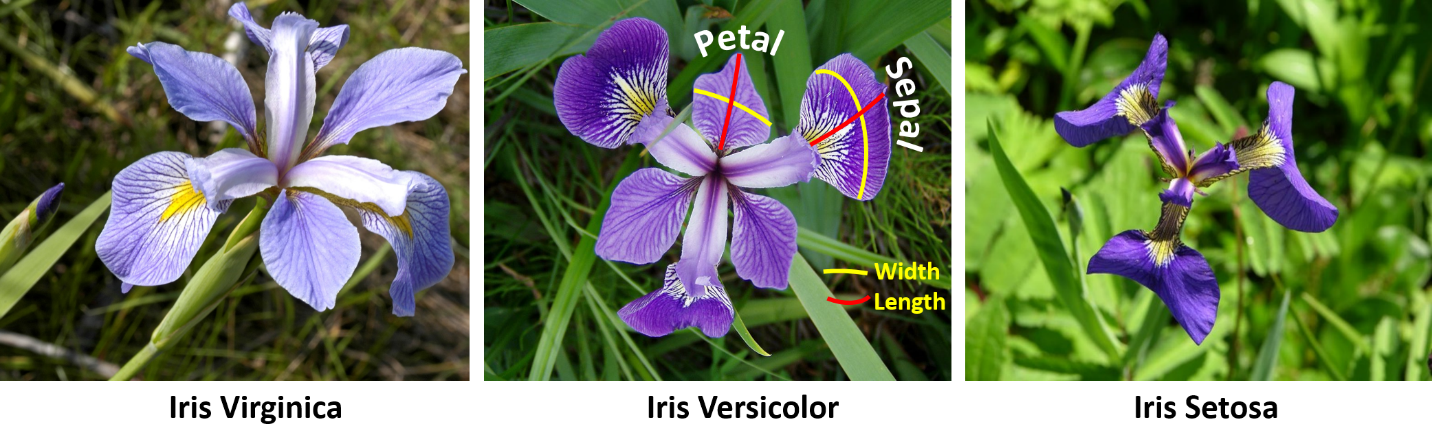
Chapter 14 - MANOVA

The Iris dataset quantifies the morphological variation of Iris flowers of three related species. It consists of 50 samples from each of three species of Iris—Iris setosa, Iris virginica, and Iris versicolor. The length and the width of the sepals and petals in centimeters were measured for each sample to compare the features per species.



The data set Iris\_data.csv contains this information. Use MANOVA to find out if petal width and sepal length differ significantly between species.

Which variables are dependent and independent variables?

Petal width and sepal width are both dependent variables. The independent variable (or factor) consists of the species.

For speed, ignore the usual assumption checks for parametric tests. Check the assumptions specific to MANOVA before reporting the MANOVA results but for simplicity, report the results 'as is' even though some assumptions may be violated. No data transformation will be conducted. No data transformation will be conducted. For the MANOVA, just perform the basic version as shown first in the book, followed by the univariate test procedure based on the initial model.

First, some file preparation.

file = read.csv("Iris\_data.csv")

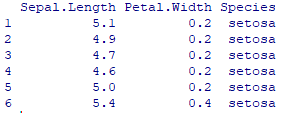
flowerStudy = with(file,

data.frame(Sepal.Length, Petal.Width, Species))

flowerStudy = na.omit(flowerStudy)

data = flowerStudy # For a variable that is easier to type for reuse

head(data)



sepL = data$Sepal.Length

petW = data$Petal.Width

library(mvnormtest) # For multivariate normality

library(biotools) # For variance/covariance matrices

iv = data$Species # The independent variable

dv = cbind(sepL, petW) # A matrix of the current dependent variables

Run the **multivariate Shapiro test**, as shown in the book. This should give the following result:



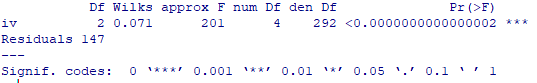
Run Box's M test, as shown in the book. This should give the following result:



Run MANOVA, ignoring the 'as.factor' as given in the book. The Species variable is clearly a factor, not a numerical variable.

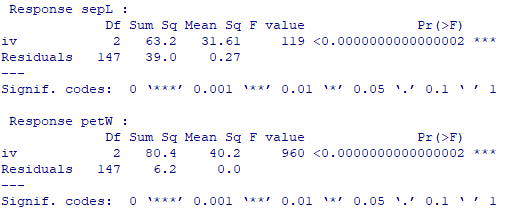
manova1 = manova(dv ~ iv, data=data)

There is no great difference between the different tests. In this case, as shown in the book, show a basic summary using Wilks' Lambda.



All of the variants show the same overall result.

To run a univariate test, as shown in the book, use summary.aov() with the model.



What do the assumption tests show?

Both homogeneity and normality assumptions were not met.

What does the MANOVA show?

The multivariate analysis is significant (*p* <.001) showing that at least one of the two variables differs between the species. Looking at the univariate results, both sepal length and petal width have significantly different mean values among the iris species.