Chapter 15

Multivariate Analysis of Variance

Contents

15.1 Example: Painted turtle shells ................................. 495
15.2 MANOVA model, inference, and tests ......................... 499
   15.2.1 Multivariate test statistics ............................... 507

15.1 Example: Painted turtle shells

Jolicouer and Mosimann studied the relationship between the size and shape of painted turtles. The table below gives the length, width, and height (all in mm) for 24 males and 24 females.

library(tidyverse)

# load ada functions
source("ada_functions.R")

### Example: Painted turtle shells
dat_shells <- read_table2("http://statacumen.com/teach/ADA2/notes/ADA2_notes_Ch15_shells_mf.dat") %>%
  mutate(UNM, Stat 428/528 ADA2)
`sex = factor(sex)`

### Parsed with column specification:
### cols(
### `sex = col_character()`,
### `length = col_double()`,
### `width = col_double()`,
### `height = col_double()`
### )

`str(dat_shells)`

### Classes 'spec_tbl_df', 'tbl_df', 'tbl' and 'data.frame': 48 obs. of 4 variables:
### $ sex : Factor w/ 2 levels "F","M": 1 1 1 1 1 1 1 1 ...
### $ length: num 98 103 105 109 123 133 133 133 ...
### $ width : num 81 84 86 86 92 95 99 102 102 ...
### $ height: num 38 38 42 42 44 50 46 51 51 ...

`#head(dat_shells)`

<table>
<thead>
<tr>
<th>sex</th>
<th>length</th>
<th>width</th>
<th>height</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>98</td>
<td>81</td>
<td>38</td>
</tr>
<tr>
<td>F</td>
<td>103</td>
<td>84</td>
<td>38</td>
</tr>
<tr>
<td>F</td>
<td>103</td>
<td>86</td>
<td>42</td>
</tr>
<tr>
<td>F</td>
<td>105</td>
<td>86</td>
<td>42</td>
</tr>
<tr>
<td>F</td>
<td>109</td>
<td>88</td>
<td>44</td>
</tr>
<tr>
<td>F</td>
<td>123</td>
<td>92</td>
<td>50</td>
</tr>
<tr>
<td>F</td>
<td>123</td>
<td>95</td>
<td>46</td>
</tr>
<tr>
<td>F</td>
<td>133</td>
<td>99</td>
<td>51</td>
</tr>
<tr>
<td>F</td>
<td>133</td>
<td>102</td>
<td>51</td>
</tr>
<tr>
<td>F</td>
<td>133</td>
<td>102</td>
<td>51</td>
</tr>
<tr>
<td>F</td>
<td>134</td>
<td>100</td>
<td>48</td>
</tr>
<tr>
<td>F</td>
<td>136</td>
<td>102</td>
<td>49</td>
</tr>
<tr>
<td>F</td>
<td>138</td>
<td>98</td>
<td>51</td>
</tr>
<tr>
<td>F</td>
<td>138</td>
<td>99</td>
<td>51</td>
</tr>
<tr>
<td>F</td>
<td>141</td>
<td>105</td>
<td>53</td>
</tr>
<tr>
<td>F</td>
<td>147</td>
<td>108</td>
<td>57</td>
</tr>
<tr>
<td>F</td>
<td>149</td>
<td>107</td>
<td>55</td>
</tr>
<tr>
<td>F</td>
<td>153</td>
<td>107</td>
<td>56</td>
</tr>
<tr>
<td>F</td>
<td>155</td>
<td>115</td>
<td>63</td>
</tr>
<tr>
<td>F</td>
<td>155</td>
<td>117</td>
<td>60</td>
</tr>
<tr>
<td>F</td>
<td>158</td>
<td>115</td>
<td>62</td>
</tr>
<tr>
<td>F</td>
<td>159</td>
<td>118</td>
<td>63</td>
</tr>
<tr>
<td>F</td>
<td>162</td>
<td>124</td>
<td>61</td>
</tr>
<tr>
<td>F</td>
<td>177</td>
<td>132</td>
<td>67</td>
</tr>
</tbody>
</table>

### Scatterplot matrix

```r
library(ggplot2)
library(GGally)
# color by sex
p <- ggpairs(
  dat_shells
, mapping = ggplot2::aes(colour = sex, alpha = 0.5)
, title = "Painted turtle shells"
, progress = FALSE
)```

---

Prof. Erik B. Erhardt
print(p)

```r
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
##  3D scatterplot
library(scatterplot3d)
par(mfrow=c(1,1))
scatterplot3d(
  x = dat_shells$length,
  y = dat_shells$width,
  z = dat_shells$height,
  main = "Shells 3D Scatterplot",
  type = "h", # lines to the horizontal xy-plane
  color = as.integer(dat_shells$sex), # color by group
  pch = as.integer(dat_shells$sex) + 19, # plotting character by group
  angle = 40, # perspective
  #, highlight.3d = TRUE # makes color change with z-axis value
)

### Try this!
### For a rotatable 3D plot, use plot3d() from the rgl library
# ## This uses the R version of the OpenGL (Open Graphics Library)
# library(rgl)
# rgl::plot3d(
#   x = dat_shells$length,
#   y = dat_shells$width,
#   z = dat_shells$height,
#   col = dat_shells$sex %>% as.integer()
# )
```
Painted turtle shells

Shells 3D Scatterplot

Cor : 0.978
F: 0.973
M: 0.95
Cor : 0.963
F: 0.971
M: 0.947
Cor : 0.96
F: 0.966
M: 0.912
15.2 MANOVA model, inference, and tests

Multivariate Analysis of Variance (MANOVA) considers the following two questions:

- Are the population mean length, width, and height the same for males and females?
- If not, then what combination of features is most responsible for the differences?

To describe MANOVA, suppose you measure $p$ features on independent random samples from $k$ strata, groups, or populations. Let

$$\mu_i' = [\mu_{i1} \ \mu_{i2} \ \cdots \ \mu_{ip}]'$$

be the vector of population means for the $i^{th}$ population, where $\mu_{ij}$ is the $i^{th}$ population mean on the $j^{th}$ feature. For the turtles, $p = 3$ features and $k = 2$ strata (sexes).

A one-way MANOVA tests the hypothesis that the population mean vectors are identical: $H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$ against $H_A : \text{not } H_0$. For the carapace data, you are simultaneously testing that the sexes have equal population mean lengths, equal population mean widths, and equal population mean heights.

Assume that the sample sizes from the different groups are $n_1, n_2, \ldots, n_k$. The total sample size is $n = n_1 + n_2 + \cdots + n_k$. Let

$$X_{ij}' = [X_{ij1} \ X_{ij2} \ \cdots \ X_{ijp}]'$$

be the vector of responses for the $j^{th}$ individual from the $i^{th}$ sample. Let

$$\bar{X}_i' = [\bar{X}_{i1} \ \bar{X}_{i2} \ \cdots \ \bar{X}_{ip}]'$$

and $S_i$ be the mean vector and variance-covariance matrix for the $i^{th}$ sample. Finally, let

$$\bar{X}' = [\bar{X}_1 \ \bar{X}_2 \ \cdots \ \bar{X}_p]'$$
be the vector of means ignoring samples (combine all the data across samples and compute the average on each feature), and let

\[ S = \frac{\sum_i (n_i - 1) S_i}{n - k} \]

be the pooled variance-covariance matrix. The pooled variance-covariance matrix is a weighted average of the variance-covariance matrices from each group.

To test \( H_0 \), construct the following MANOVA table, which is the multivariate analog of the ANOVA table:

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between</td>
<td>( k - 1 )</td>
<td>( \sum_i n_i (\bar{X}_i - \bar{X})(\bar{X}_i - \bar{X})' )</td>
<td>( \frac{\text{SS}}{\text{df}} )</td>
</tr>
<tr>
<td>Within</td>
<td>( n - k )</td>
<td>( \sum_i (n_i - 1) S_i )</td>
<td>( \frac{\text{SS}}{\text{df}} )</td>
</tr>
<tr>
<td>Total</td>
<td>( n - 1 )</td>
<td>( \sum_{ij} (X_{ij} - \bar{X})(X_{ij} - \bar{X})' )</td>
<td>( \frac{\text{SS}}{\text{df}} )</td>
</tr>
</tbody>
</table>

where all the MSs are SS/df.

The expressions for the SS have the same form as SS in univariate analysis of variance, except that each SS is a \( p \times p \) symmetric matrix. The diagonal elements of the SS matrices are the SS for one-way ANOVAs on the individual features. The off-diagonal elements are SS between features. The Error MS matrix is the pooled variance-covariance matrix \( S \).

The standard MANOVA assumes that you have independent samples from multivariate normal populations with identical variance-covariance matrices. This implies that each feature is normally distributed in each population, that a feature has the same variability across populations, and that the correlation (or covariance) between two features is identical across populations. The Error MS matrix estimates the common population variance-covariance matrix when the population variance-covariance matrices are identical.

The \( H_0 \) of equal population mean vectors should be rejected when the difference among mean vectors, as measured by the Between MS matrix, is large relative to the variability within groups, as measured by the Error MS matrix. Equivalently, \( H_0 \) is implausible if a significant portion of the total variation in the data, as measured by the Total SS matrix, is due to differences among the groups. The same idea is used in a one-way ANOVA to motivate the \( F \)-test of
no differences in population means. However, some care is needed to quantify these ideas in a MANOVA because there are several natural matrix definitions for comparing the Between MS matrix to the Error MS matrix. As a result, several MANOVA tests of $H_0$ have been proposed.

Graphical summaries for the carapace data are given above, with numerical summaries below.

```r
# summary statistics for each sex
by(dat_shells, dat_shells$sex, summary)

## dat_shells$sex: F
## sex length width height
## F:24 Min. : 98.0 Min. : 81.00 Min. :38.00
## M: 0 1st Qu.:123.0 1st Qu.: 94.25 1st Qu.:47.50
## Median :137.0 Median :102.00 Median :51.00
## Mean :136.0 Mean :102.58 Mean :52.04
## 3rd Qu.:153.5 3rd Qu.:109.75 3rd Qu.:57.75
## Max. :177.0 Max. :132.00 Max. :67.00
```

```r
## dat_shells$sex: M
## sex length width height
## F: 0 Min. : 93.0 Min. : 74.00 Min. :35.00
## M:24 1st Qu.:103.8 1st Qu.: 83.00 1st Qu.:38.75
## Median :115.0 Median : 89.00 Median :40.00
## Mean :113.4 Mean : 88.29 Mean :40.71
## 3rd Qu.:121.0 3rd Qu.: 93.00 3rd Qu.:43.25
## Max. :135.0 Max. :106.00 Max. :47.00

# standard deviations
by(dat_shells[, 2:4], dat_shells$sex, apply, 2, sd)

## dat_shells$sex: F
## length width height
## 21.24900 13.10465 8.04595

## dat_shells$sex: M
## length width height
## 11.806103 7.074013 3.355452

# correlation matrix (excluding associated p-values testing "H0: rho == 0")
#library(Hmisc)
Hmisc::rcorr(as.matrix(dat_shells[, 2:4]))[[1]] # all

## length width height
## length 1.0000000 0.9778962 0.9628010
## width 0.9778962 1.0000000 0.9599055
## height 0.9628010 0.9599055 1.0000000
Hmisc::rcorr(as.matrix(dat_shells[dat_shells$sex == "F", 2:4]))[[1]] # females

## length width height
```
The features are positively correlated within each sex. The correlations between pairs of features are similar for males and females. Females tend to be larger on each feature. The distributions for length, width, and height are fairly symmetric within sexes. No outliers are present. Although females are more variable on each feature than males, the MANOVA assumptions do not appear to be grossly violated here. (Additionally, you could consider transforming the each dimension of the data in hopes to make the covariances between sexes more similar, though it may not be easy to find a good transformation to use.)

```r
pca_sh <-
  princomp( 
    dat_shells[, 2:4] 
  )

df_pca_sh <-
  data.frame( 
    sex = dat_shells$sex
  , pca_sh$scores 
  )

str(df_pca_sh)
## 'data.frame': 48 obs. of  4 variables:
## $ sex : Factor w/ 2 levels "F","M": 1 1 1 1 1 1 1 1 1 1 ...
## $ Comp.1: num -31.4 -25.9 -23.6 -22 -17.1 ...
## $ Comp.2: num -2.27 -1.43 -4.44 -3.26 -3.11 ...
## $ Comp.3: num -0.943 0.73 -1.671 -1.618 -2.2 ...

# Scatterplot matrix
library(ggplot2)
library(GGally)
# put scatterplots on top so y axis is vertical
p <-
  ggpairs( 
    df_pca_sh
  , mapping = ggplot2::aes(colour = sex, alpha = 0.5)
  , title = "Principal components of Shells"
  , progress = FALSE
  )
```
For comparison with the MANOVA below, here are the univariate ANOVAs for each feature. For the carapace data, the univariate ANOVAs indicate significant differences between sexes on length, width, and height. Females are larger on average than males on each feature.

```
# Univariate ANOVA tests, by each response variable
lm_sh <-
  lm(
    cbind(length, width, height) ~ sex
  , data = dat_shells
  )
summary(lm_sh)
## Response length :
```
## Call:

```
lm(formula = length ~ sex, data = dat_shells)
```

## Residuals:

```
Min 1Q Median 3Q Max
-38.042 -10.667 1.271 11.927 40.958
```

## Coefficients:

```
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)        136.042    3.509   38.77   < 2e-16 ***
sexM             -22.625     4.962   -4.56    3.79e-05 ***
```

---

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Residual standard error: 17.19 on 46 degrees of freedom

## Multiple R-squared:  0.3113, Adjusted R-squared:  0.2963

## F-statistic: 20.79 on 1 and 46 DF,  p-value: 3.788e-05

## Response width:

## Call:

```
lm(formula = width ~ sex, data = dat_shells)
```

## Residuals:

```
Min 1Q Median 3Q Max
-21.5833 -5.5417 -0.4375 4.8854 29.4167
```

## Coefficients:

```
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)       102.583     2.149  47.725   < 2e-16 ***
sexM              -14.292     3.040  -4.701    2.38e-05 ***
```

---

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Residual standard error: 10.53 on 46 degrees of freedom

## Multiple R-squared:  0.3246, Adjusted R-squared:  0.3099

## F-statistic: 22.1 on 1 and 46 DF,  p-value: 2.376e-05

## Response height:

## Call:

```
lm(formula = height ~ sex, data = dat_shells)
```

## Residuals:

```
Min 1Q Median 3Q Max
-10.667 1.271 11.927 40.958
```

## Coefficients:

```
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)       136.042     3.509   38.77   < 2e-16 ***
sexM              -22.625     4.962   -4.56    3.79e-05 ***
```

---

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Residual standard error: 17.19 on 46 degrees of freedom

## Multiple R-squared:  0.3113, Adjusted R-squared:  0.2963

## F-statistic: 20.79 on 1 and 46 DF,  p-value: 3.788e-05

---

Prof. Erik B. Erhardt
A few procedures can be used for one-way MANOVA; two are `manova()` and the `car` package’s `Manova()`. First we check the assumption of multivariate normality.

```r
# Test multivariate normality using the Shapiro-Wilk test for multivariate normality
library(mvnormtest)
# The data needs to be transposed t() so each variable is a row
# with observations as columns.

mshapiro.test(t(dat_shells[dat_shells$sex == "F", 2:4]))
# Shapiro-Wilk normality test
#
# data: Z
# W = 0.89324, p-value = 0.01551

mshapiro.test(t(dat_shells[dat_shells$sex == "M", 2:4]))
# Shapiro-Wilk normality test
#
# data: Z
# W = 0.93602, p-value = 0.1329

f_mnv_norm_qqplot(dat_shells[dat_shells$sex == "F", 2:4], "Female")
f_mnv_norm_qqplot(dat_shells[dat_shells$sex == "M", 2:4], "Male")
```

UNM, Stat 428/528 ADA2
The curvature in the Female sample cause us to reject normality, while the males do not deviate from normality. We’ll proceed anyway since this deviation from normality in the female sample will largely increase the variability of the sample and not displace the mean greatly, and the sample sizes are somewhat large.
15.2.1 Multivariate test statistics

These four multivariate test statistics are among the most common to assess differences across the levels of the categorical variables for a linear combination of responses. In general Wilks’ lambda is recommended unless there are problems with small total sample size, unequal sample sizes between groups, violations of assumptions, etc., in which case Pillai’s trace is more robust.

Wilks’ lambda, (λ)

- Most commonly used statistic for overall significance
- Considers differences over all the characteristic roots
- The smaller the value of Wilks’ lambda, the larger the between-groups dispersion

Pillai’s trace

- Considers differences over all the characteristic roots
- More robust than Wilks’; should be used when sample size decreases, unequal cell sizes or homogeneity of covariances is violated

Hotelling’s trace

- Considers differences over all the characteristic roots

Roy’s greatest characteristic root

- Tests for differences on only the first discriminant function (Chapter 16)
- Most appropriate when responses are strongly interrelated on a single dimension
- Highly sensitive to violation of assumptions, but most powerful when all assumptions are met.

```r
# Multivariate MANOVA test
# the specific test is specified in summary()
# test = c("Pillai", "Wilks", "Hotelling-Lawley", "Roy")
man_sh <-
  manova(
    cbind(length, width, height) ~ sex,
    data = dat_shells
  )
summary(man_sh, test = "Wilks")
```

## Df Wilks approx F num Df den Df Pr(>F)
## sex 1 0.38695 23.237 3 44 3.622e-09 ***
## Residuals 46
The four MANOVA tests of no differences between sexes are all highly significant. These tests reinforce the univariate analyses. Of the four tests, I prefer Roy’s test because it has an intuitive interpretation. I will mostly ignore the other three tests for discussion.

**Roy’s test** locates the linear combination of the features that produces the most significant one-way ANOVA test for no differences among groups. If
the groups are not significantly different on the linear combination that best separates the groups in a one-way ANOVA sense, then there is no evidence that the population mean vectors are different. The critical value for Roy’s test accounts for the linear combination being suggested by the data. That is, the critical value for Roy’s test is not the same critical value that is used in a one-way ANOVA. The idea is similar to a Bonferroni-type correction with multiple comparisons.

Roy’s method has the ability to locate linear combinations of the features on which the groups differ, even when the differences across groups are not significant on any feature. This is a reason for treating multivariate problems using multivariate methods rather than through individual univariate analyses on each feature.

```r
## For Roy's characteristic Root and vector
str(man_sh)
# H = hypothesis matrix
# man_sh$df # hypothesis df
# E = error matrix
# man_sh$error.df # error df
# characteristic roots of (E inverse * H)
EinvH <- solve(E) %*% H
# solve() computes the matrix inverse
ev <- eigen(EinvH)
# eigenvalue/eigenvectors

ev
```

The **characteristic vector** (eigenvector) output gives one linear combination of the features for each variable in the data set. By construction, the linear combinations are uncorrelated (adjusting for groups). In general, the first $a = \min(p, k - 1)$ linear combinations contain information in decreasing amounts for distinguishing among the groups. The first linear combination is used by Roy’s test.
The three linear combinations for the carapace data are (reading down the columns in the matrix of eigenvectors)

\[
\begin{align*}
D1 &= 0.2152 \text{ Length} + 0.1014 \text{ Width} - 0.9713 \text{ Height} \\
D2 &= 0.3564 \text{ Length} + 0.1651 \text{ Width} - 0.9196 \text{ Height} \\
D3 &= 0.02975 \text{ Length} + 0.5917 \text{ Width} - 0.8056 \text{ Height}.
\end{align*}
\]

Here \( p = 3 \) and \( k = 2 \) gives \( a = \min(p, k - 1) = \min(3, 2 - 1) = 1 \) so only \( D1 \) from \( (D1, D2, \text{ and } D3) \) contains information for distinguishing between male and female painted turtles.

As in PCA, the linear combinations should be interpreted. However, do not discount the contribution of a feature with a small loading (though, in the shells example, \( D2 \) and \( D3 \) have 0 for loadings). In particular, the most important feature for distinguishing among the groups might have a small loading because of the measurement scale.

The following output shows the differences between sexes on \( D1 \). The separation between sexes on \( D1 \) is greater than on any single feature. Females typically have much larger \( D1 \) scores than males.

Since the matrix of eigenvectors are a rotation matrix\(^1\) we can create the \( D \) linear combinations by matrix multiplication of the eigenvector (rotation) matrix with the original data (being careful about dimensions).

\[
D \leftarrow \text{as.matrix}(\text{dat\_shells[,2:4]}) \times \text{ev$vectors}
\]

```r
colnames(D) \leftarrow \text{c("D1", "D2", "D3")}
```

```r
df\_D \leftarrow \text{data.frame}(\text{sex = dat\_shells$sex}, D)
```

```r
\text{str(df\_D)}
```

\( \text{## 'data.frame': 48 obs. of 4 variables:} \)
\( \text{##$ sex: Factor w/ 2 levels "F","M": 1 1 1 1 1 1 1 1 1 1 ...} \)
\( \text{##$ D1 : num -7.6 -6.22 -9.91 -9.48 -10.36 ...} \)
\( \text{##$ D2 : num 13.4 15.6 12.3 13 12.9 ...} \)
\( \text{##$ D3 : num 20.2 22.2 20.1 20.2 19.9 ...} \)

\( \text{## Scatterplot matrix} \)

\( \text{library(ggplot2)} \)

\( \text{library(GGally)} \)

\( \text{\texttt{\textbackslash http://en.wikipedia.org/wiki/Rotation_matrix}} \)

---

\(^1\text{http://en.wikipedia.org/wiki/Rotation_matrix}\)}
# put scatterplots on top so y axis is vertical

```r
p <-
ggpairs(
df_D
, mapping = ggplot2::aes(colour = sex, alpha = 0.5)
, title = "D1 is the linear combination that best distinguishes the sexes"
, progress = FALSE
)
print(p)
```

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```

```
D1 is the linear combination that best distinguishes the sexes
```

```
<table>
<thead>
<tr>
<th>sex</th>
<th>D1</th>
<th>D2</th>
<th>D3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

```
0  5  10  15  20  25  30
```

```
Cor : 1.56e−15
F: −0.184
M: 0.501
Cor : 1.47e−16
F: −0.19
M: 0.515
Cor : 0.877
F: 0.863
M: 0.897
```

```
# Univariate ANOVA tests, by D1 linear combination variable

```r
lm_D_sh <-
  lm(D1 ~ sex
  , data = df_D
  )
```

UNM, Stat 428/528 ADA2
summary(lm_D_sh)

## Call:
## lm(formula = D1 ~ sex, data = df_D)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.3047  -0.9791   0.3260   0.9570   4.6434
##
## Coefficients:
##             Estimate Std. Error t value  Pr(>|t|)
## (Intercept) -10.8679    0.3884  -27.98   < 2e-16  ***
## sexM         4.6897     0.5493   8.54   4.84e-11  ***
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 1.903 on 46 degrees of freedom
## Multiple R-squared:  0.6131, Adjusted R-squared:  0.6046
## F-statistic: 72.88 on 1 and 46 DF,  p-value: 4.841e-11