

## exercise5

Lee Rui (BA19008042)

2020/3/29

```
data(iris)
head(iris)

##      Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 1           5.1           3.5           1.4           0.2  setosa
## 2           4.9           3.0           1.4           0.2  setosa
## 3           4.7           3.2           1.3           0.2  setosa
## 4           4.6           3.1           1.5           0.2  setosa
## 5           5.0           3.6           1.4           0.2  setosa
## 6           5.4           3.9           1.7           0.4  setosa

colnames(iris)[1:5]=c("sepal_length","sepal_width","petal_length","petal_width","class")

iris$class=factor(iris$class)
str(iris)

## 'data.frame':   150 obs. of  5 variables:
##  $ sepal_length: num  5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...
##  $ sepal_width : num  3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...
##  $ petal_length: num  1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
##  $ petal_width : num  0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
##  $ class       : Factor w/ 3 levels "setosa","versicolor",...: 1 1 1 1 1 1 1 1 1 1 ...

sample_iris=sample(150,110,replace = FALSE)
sample_iris

##      [1] 132  48  44  80  29  97  60 125 104 123  23  65  61 112  13 126  68 121
##     [19]  94 147 117  59  16  19 149  51  39  3  95 128  70  69 111  57  37  21
##     [37] 102  35  31 120  88  81  50  96  1  10 119  56  54 101 135 133 131  63
##     [55]  18  24  78  20  93  43 115  17  82 134 141 150  87  73  86  6 145 109
##     [73] 114 142 107  55 140  99  64  90 146  36  77  2  38 106 118  14  89 105
##     [91]  79  11  41  5  72  8 103  53  26 110  58 138 136  7  42  98  40  9
##    [109]  91  33

iris_training=iris[sample_iris,]
iris_test=iris[-sample_iris,]
iris_training_labels=iris[sample_iris,]$class
iris_test_labels=iris[-sample_iris,]$class
table(iris_training$class)

##
##      setosa versicolor  virginica
##           36           37           37

table(iris_test$class)
```

```
##
##      setosa versicolor  virginica
##      14         13         13
```

```
library(e1071)
```

```
## Warning: package 'e1071' was built under R version 3.6.3
```

```
iris_classifier=naiveBayes(iris_training,iris_training_labels)
iris_test_pred=predict(iris_classifier,iris_test)
iris_test_pred
```

```
## [1] setosa      setosa      setosa      setosa      setosa      setosa
## [7] setosa      setosa      setosa      setosa      setosa      setosa
## [13] setosa      setosa      versicolor  versicolor  versicolor  versicolor
## [19] versicolor  versicolor  versicolor  versicolor  versicolor  versicolor
## [25] versicolor  versicolor  versicolor  virginica   virginica   virginica
## [31] virginica   virginica   virginica   virginica   virginica   virginica
## [37] virginica   virginica   virginica   virginica
## Levels: setosa versicolor virginica
```

```
library(gmodels)
```

```
## Warning: package 'gmodels' was built under R version 3.6.3
```

```
CrossTable(iris_test_pred,iris_test_labels,prop.chisq = FALSE, prop.t = FALSE, prop.r = FALSE, dnn = c(
```

```
##
##
##      Cell Contents
## |-----|
## |                      N |
## |          N / Col Total |
## |-----|
##
##
## Total Observations in Table:  40
##
##
```

	actual			
predicted	setosa	versicolor	virginica	Row Total
setosa	14	0	0	14
	1.000	0.000	0.000	
versicolor	0	13	0	13
	0.000	1.000	0.000	
virginica	0	0	13	13
	0.000	0.000	1.000	
Column Total	14	13	13	40
	0.350	0.325	0.325	

```
##
##
```

```
iris_classifier2=naiveBayes(iris_training,iris_training_labels,laplace = 1)
iris_test_pred2=predict(iris_classifier2,iris_test)
iris_test_pred2
```

```
## [1] setosa      setosa      setosa      setosa      setosa      setosa
## [7] setosa      setosa      setosa      setosa      setosa      setosa
## [13] setosa      setosa      versicolor  versicolor  versicolor  versicolor
## [19] versicolor  versicolor  versicolor  versicolor  versicolor  versicolor
## [25] versicolor  versicolor  versicolor  virginica   virginica   virginica
## [31] virginica   virginica   virginica   virginica   virginica   virginica
## [37] virginica   virginica   virginica   virginica
## Levels: setosa versicolor virginica
```

```
CrossTable(iris_test_pred2,iris_test_labels,prop.chisq = FALSE, prop.t = FALSE, prop.r = FALSE, dnn = c
```

```
##
##
##      Cell Contents
## |-----|
## |              N |
## |      N / Col Total |
## |-----|
##
##
## Total Observations in Table:  40
##
##
##      | actual
## predicted |      setosa | versicolor |  virginica | Row Total |
## -----|-----|-----|-----|-----|
##      setosa |          14 |           0 |           0 |          14 |
##      |          1.000 |          0.000 |          0.000 |
## -----|-----|-----|-----|
##      versicolor |           0 |          13 |           0 |          13 |
##      |          0.000 |          1.000 |          0.000 |
## -----|-----|-----|-----|
##      virginica |           0 |           0 |          13 |          13 |
##      |          0.000 |          0.000 |          1.000 |
## -----|-----|-----|-----|
## Column Total |          14 |          13 |          13 |          40 |
##      |          0.350 |          0.325 |          0.325 |
## -----|-----|-----|-----|
##
##
##
```