

Home Study 3: To be completed by Monday 2nd November

MAS2905: Probability and Bayesian Methods with R

Newcastle University
`lee.fawcett@ncl.ac.uk`

Semester 1, 2020/21

What you should have done so far...

- Completed everything in the first two chapters of the course notes!
 - General course structure (Chapter 1)
 - R Refresher – data types, vectors and dataframes (Chapter 2 – up to Section 2.4): In our first synchronous online lecture and through home study exercises
 - R Refresher – functions, for loops and if statements (Section 2.5): In our first computer practical (Tuesday 27th October)
- You should now be ready to see some solutions to **Example 2.4**

What you should have done so far...

- Completed everything in the first two chapters of the course notes!
 - General course structure (Chapter 1)
 - R Refresher – data types, vectors and dataframes (Chapter 2 – up to Section 2.4): In our first synchronous online lecture and through home study exercises
 - R Refresher – functions, for loops and if statements (Section 2.5): In our first computer practical (Tuesday 27th October)
- You should now be ready to see some solutions to **Example 2.4**

What you should have done so far...

- Completed everything in the first two chapters of the course notes!
 - General course structure (Chapter 1)
 - R Refresher – data types, vectors and dataframes (Chapter 2 – up to Section 2.4): In our first synchronous online lecture and through home study exercises
 - R Refresher – functions, for loops and if statements (Section 2.5): In our first computer practical (Tuesday 27th October)
- You should now be ready to see some solutions to Example 2.4

What you should have done so far...

- Completed everything in the first two chapters of the course notes!
 - General course structure (Chapter 1)
 - R Refresher – data types, vectors and dataframes (Chapter 2 – up to Section 2.4): In our first synchronous online lecture and through home study exercises
 - R Refresher – functions, for loops and if statements (Section 2.5): In our first computer practical (Tuesday 27th October)
- You should now be ready to see some solutions to Example 2.4

What you should have done so far...

- Completed everything in the first two chapters of the course notes!
 - General course structure (Chapter 1)
 - R Refresher – data types, vectors and dataframes (Chapter 2 – up to Section 2.4): In our first synchronous online lecture and through home study exercises
 - R Refresher – functions, for loops and if statements (Section 2.5): In our first computer practical (Tuesday 27th October)
- You should now be ready to see some solutions to **Example 2.4**

Example 2.4: Solutions



Load the music dataset into RStudio, just as you did in Example 2.3.

```
1 library(ncldata)
2 data(music)
3 attach(music)
```

Example 2.4: Solutions

- (a) Find the mean and standard deviation for the duration of tracks in this dataset.

```
1 mean(duration); sd(duration)
2 [1] 214.8182
3 [1] 83.53777
```

So mean duration is 214.8s, with a standard deviation of 83.54s.

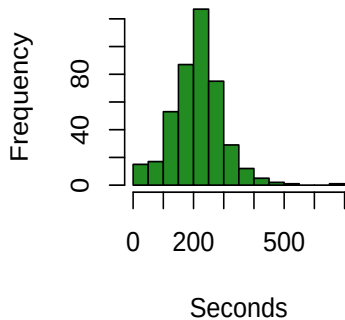
Example 2.4: Solutions

- (b) Produce histograms for track duration and danceability; put these histograms on the same panel of plots, use different colours for the histograms (you can choose whichever colours you like!) and make sure you provide suitable titles and labels on the axes.

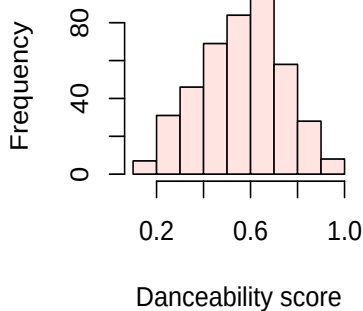
```
1 par(mfrow=c(1,2))
2
3 hist(duration , main="Track duration" , xlab="Seconds" ,
4       col="forest green")
5
6 hist(danceability , main="Danceability" , xlab="
7       Danceability score" , col="misty rose")
```

Example 2.4: Solutions

Track duration



Danceability

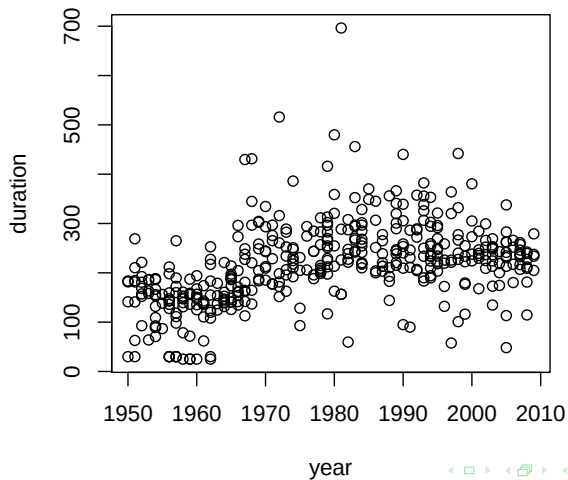


Example 2.4

- (c) You are asked to investigate whether or not the duration of tracks has changed over time.
 - (i) Use the `plot` command in R to produce a scatterplot of the year each track was released against track duration.

```
1 plot(year, duration)
```

Example 2.4

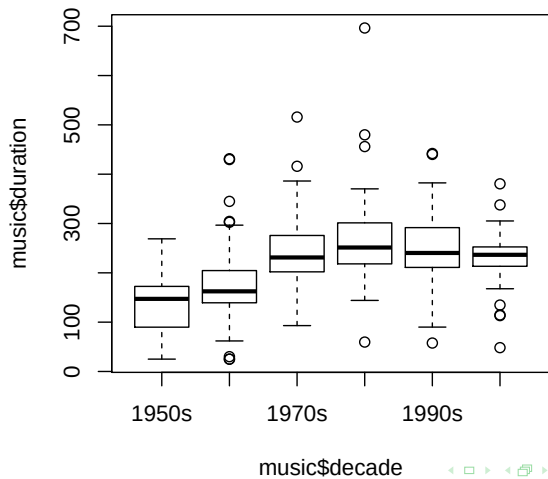


Example 2.4

- (c) You are asked to investigate whether or not the duration of tracks has changed over time.
 - (ii) Try out the following code, and comment on what you see in the plots in part (c):

```
1 boxplot(music$duration ~ music$decade)
```

Example 2.4



Example 2.4

- Track duration seems to increase through time, from the 1950s through to the 1980s
- Since the 1980s track duration seems to have decreased
- There is a very long track in the 1980s of over 700s (*sim11* minutes)

Example 2.4

- (d) On average, how do energy scores for “Hip Hop” tracks compare to “Electronic” tracks?

```
1 mean(energy[genre=="Hip Hop"])  
2 [1] 0.6315726  
3 mean(energy[genre=="Electronic"])  
4 [1] 0.7280796
```

Electronic tracks seem more energetic!

Chapter 2: Summary of learning outcomes

We've finished Chapter 2, and the R Refresher!

Hopefully you now feel a bit more comfortable with:

- **Data types** in R: Doubles, logicals, Characters, NA, NaN, Inf
- **Vectors** in R: Operations and extraction
- **Dataframes** in R: Creating and querying
- **Graphical** and **numerical summaries** in R: Measures of average and spread; simple plots
- **Control flow** and **functions** in R: for loops, if statements and functions

Chapter 2: Summary of learning outcomes

We've finished Chapter 2, and the R Refresher!

Hopefully you now feel a bit more comfortable with:

- **Data types** in R: Doubles, logicals, Characters, NA, NaN, Inf
- **Vectors** in R: Operations and extraction
- **Dataframes** in R: Creating and querying
- **Graphical** and **numerical summaries** in R: Measures of average and spread; simple plots
- **Control flow** and **functions** in R: for loops, if statements and functions

Chapter 2: Summary of learning outcomes

We've finished Chapter 2, and the R Refresher!

Hopefully you now feel a bit more comfortable with:

- **Data types** in R: Doubles, logicals, Characters, NA, NaN, Inf
- **Vectors** in R: Operations and extraction
- **Dataframes** in R: Creating and querying
- **Graphical** and **numerical summaries** in R: Measures of average and spread; simple plots
- **Control flow** and **functions** in R: for loops, if statements and functions

Chapter 2: Summary of learning outcomes

We've finished Chapter 2, and the R Refresher!

Hopefully you now feel a bit more comfortable with:

- **Data types** in R: Doubles, logicals, Characters, NA, NaN, Inf
- **Vectors** in R: Operations and extraction
- **Dataframes** in R: Creating and querying
- **Graphical** and **numerical summaries** in R: Measures of average and spread; simple plots
- **Control flow** and **functions** in R: for loops, if statements and functions

Chapter 2: Summary of learning outcomes

We've finished Chapter 2, and the R Refresher!

Hopefully you now feel a bit more comfortable with:

- **Data types** in R: Doubles, logicals, Characters, NA, NaN, Inf
- **Vectors** in R: Operations and extraction
- **Dataframes** in R: Creating and querying
- **Graphical** and **numerical summaries** in R: Measures of average and spread; simple plots
- **Control flow** and **functions** in R: for loops, if statements and functions

Chapter 2: Summary of learning outcomes

We've finished Chapter 2, and the R Refresher!

Hopefully you now feel a bit more comfortable with:

- **Data types** in R: Doubles, logicals, Characters, NA, NaN, Inf
- **Vectors** in R: Operations and extraction
- **Dataframes** in R: Creating and querying
- **Graphical** and **numerical summaries** in R: Measures of average and spread; simple plots
- **Control flow** and **functions** in R: for loops, if statements and functions

Chapter 2: Summary of learning outcomes

We've finished Chapter 2, and the R Refresher!

Hopefully you now feel a bit more comfortable with:

- **Data types** in R: Doubles, logicals, Characters, NA, NaN, Inf
- **Vectors** in R: Operations and extraction
- **Dataframes** in R: Creating and querying
- **Graphical** and **numerical summaries** in R: Measures of average and spread; simple plots
- **Control flow** and **functions** in R: for loops, if statements and functions

Prize question



Suppose we randomly permute the letters of the word Mississippi.

- (a) What is the expected number of times we have consecutive S's in the jumbled up word?
- (b) What is the probability the word contains no consecutive S's?

Home Study 3: To be completed by Monday 2nd November

- That's this Home Study activity completed. Well done!
- Before the end of **Monday 2nd November**, you should:
 - Make sure you're on top of EVERYTHING in Chapter 2 – take some time to read over the notes again, and complete ALL exercises if you've not already done so
 - Complete the prize question and submit your answers to me, via email
- See you next week, in our synchronous online lecture:
12 noon, Wednesday 4th November
 - Pre-lecture reading, available from Monday 2nd November: Chapter 3 of the course notes, Sections 3.1–3.2 (look in the week 3 materials in Canvas)
 - Try to fill in the gaps, and try as much of **Examples 3.1–3.4** as you can – answers will be revealed on Wednesday!

Home Study 3: To be completed by Monday 2nd November

- That's this Home Study activity completed. Well done!
- Before the end of **Monday 2nd November**, you should:
 - Make sure you're on top of EVERYTHING in Chapter 2 – take some time to read over the notes again, and complete ALL exercises if you've not already done so
 - Complete the prize question and submit your answers to me, via email
- See you next week, in our synchronous online lecture:
12 noon, Wednesday 4th November
 - Pre-lecture reading, available from Monday 2nd November: Chapter 3 of the course notes, Sections 3.1–3.2 (look in the week 3 materials in Canvas)
 - Try to fill in the gaps, and try as much of Examples 3.1–3.4 as you can – answers will be revealed on Wednesday!

Home Study 3: To be completed by Monday 2nd November

- That's this Home Study activity completed. Well done!
- Before the end of **Monday 2nd November**, you should:
 - Make sure you're on top of EVERYTHING in Chapter 2 – take some time to read over the notes again, and complete ALL exercises if you've not already done so
 - Complete the prize question and submit your answers to me, via email
- See you next week, in our synchronous online lecture:
12 noon, Wednesday 4th November
 - Pre-lecture reading, available from Monday 2nd November: Chapter 3 of the course notes, Sections 3.1–3.2 (look in the week 3 materials in Canvas)
 - Try to fill in the gaps, and try as much of Examples 3.1–3.4 as you can – answers will be revealed on Wednesday!

Home Study 3: To be completed by Monday 2nd November

- That's this Home Study activity completed. Well done!
- Before the end of **Monday 2nd November**, you should:
 - Make sure you're on top of EVERYTHING in Chapter 2 – take some time to read over the notes again, and complete ALL exercises if you've not already done so
 - Complete the prize question and submit your answers to me, via email
- See you next week, in our synchronous online lecture:
12 noon, Wednesday 4th November
 - Pre-lecture reading, available from Monday 2nd November: Chapter 3 of the course notes, Sections 3.1–3.2 (look in the week 3 materials in Canvas)
 - Try to fill in the gaps, and try as much of **Examples 3.1–3.4** as you can – answers will be revealed on Wednesday!