Homework 2

Write your name here

31 March, 2022

Section 1

Background: For section we will work with our coffee data set again. This data set comes from an experiment designed to test whether drinking coffee makes people smarter. Participants were divided into two groups. All participants took an IQ test, but participants in **group 1** drank water 5 minutes before the test, while participants in **group 2** drank coffee.

We have two competing hypothesis (models), the first one is that drinking coffee doesn't make a difference in IQ test scores between people who drank water (group 1) and people who drank coffee (group 2). The second one states that people who drank coffee are smarter, in other words, that the IQ test scores should be different between groups. For both hypothesis we will assume that variances are equal across observations and groups.

```
# In this chunk you should write (a.k.a. copy from the slides) the code needed
# to load the data for section 1. IMPORTANT: Remember to delete the comments
# inside the code chunks BEFORE submitting your homework.
```

1. Write the first model (**no differences between groups**) in mathematical notation

ANS: [WRITE YOUR ANSWER HERE]

2. Make a density graph of model 1 (**no differences between groups**). Remember that at this point the units (values of x) don't matter, we are only interested in what the model expects the data to look like. Assume that the model variance is equal to 3.

```
# In this code chunk you can write the code used to generate the density plot
# for model 1. remember that the values on the x axis are not important, what we
# look at is what the model expects the data to look like.
```

3. Calculate the prediction of the first model and add it as a new variable (column) to the data set (remember to give the variable a name!). What are the prediction of the first model?

```
# In this code chunk you can write the functions that you need to use in order
# to calculate the predictions of model one and add it to the data set as a new
# column. Report the value(s) as text text below.
```

ANS: [WRITE YOUR ANSWER HERE]

4. Calculate the squared error (SE) of each observation for the first model and add it to the data set as a new variable. What is the mean Sum of Squared Errors (m-SSE) of the no group differences model (model 1)?

```
# In this code chunk you can write the functions that you need to use in order
# to calculate the squared error of each observation and the functions you need
# to calculate the sum of squared errors of the no differences model.
```

5. What is the value of the Bayesian Information Criterion for model 1?

```
# In this code chunk you can write the functions that you need in order to
# calculate the BIC associated to model 1. Save it as a new variable and
# report the value as text below.
```

ANS: [WRITE YOUR ANSWER HERE]

6. Write the second model (**groups are different**) in mathematical notation.

ANS: [WRITE YOUR ANSWER HERE]

7. Make a density graph of model 2 (**groups are different**). Remember that at this point the units (values of x) don't matter, we are only interested in what the model expects the data to look like. Assume that the model variance is equal to 3.

```
# In this code chunk you can write the code used to generate the density plot
# for model 2. remember that the values on the x axis are not important, what we
# look at is what the model expects the data to look like.
```

8. Calculate the prediction of the second model and add it as a new variable (column) to the data set (remember to give the variable a different name!). What are the prediction of the second model?

```
# In this code chunk you can write the functions that you need to use in order
# to calculate the predictions of model two and add it to the data set as a new
# column. Report the value(s) as text below
```

ANS: [WRITE YOUR ANSWER HERE]

9. Calculate the squared error (SE) of each observation for the second model and add it to the data set as a new variable. What is the mean Sum of Squared Errors (m-SSE) of the group differences model (model 2)?

```
# In this code chunk you can write the functions that you need to use in order
# to calculate the squared error of each observation and the functions you need
# to calculate the mean sum of squared errors of the group differences model.
```

ANS: [WRITE YOUR ANSWER HERE]

10. What is the value of the Bayesian Information Criterion for model 2?

```
# In this code chunk you can write the functions that you need in order to
# calculate the BIC associated to model 2. Save it as a new variable and
# report the value as text below.
```

ANS: [WRITE YOUR ANSWER HERE]

11. Which of the two models would you select and why?

ANS: [WRITE YOUR ANSWER HERE]

Section 2

Background: For section we will work with our response time data set again. This data set comes from an experiment designed to test whether a new training technique can improve participants' ability to navigate a maze. Participants first completed a maze after playing a game that lasted as long as the new training approach. Then they underwent the new training and completed a second maze of the same difficulty (we will assume that, if there is a difference in response times it is only because of the training and not because of experience with maze solving). The data provides information about the time (in seconds) in which each participant completed each maze.

We have two competing models (hypothesis) about the performance of participants on the maze before and after training, the first one (**no improvement**) states that the new training method does not make any

difference, in other words that there will be no improvement on time to solve the maze.

The second model (**improvement**) states that the new training method is effective, and so the time to solve a maze should be shorter after training than before the training session.

```
# In this section you should write the code needed to load the data for
# section 2. IMPORTANT: Remember to delete the comments inside the code
# chunks BEFORE submitting your homework.
```

12. Write the first model (**no improvement**) in mathematical notation

ANS: [WRITE YOUR ANSWER HERE]

13. Make a density graph of model 1 (**no improvement**). To make the plot assume that the model variance is equal to 2.

```
# In this code chunk you can write the code used to generate the density plot
# for model 1.
```

14. Write the predictions of the first model (**no improvement**)?

ANS: [WRITE YOUR ANSWER HERE]

15. Calculate the squared error (SE) of each observation for the first model and add it to the data set as a new variable. What is the mean Sum of Squared Errors (m-SSE) of the no group differences model (model 1)?

```
# In this code chunk you can write the functions that you need to use in order
# to calculate the squared error of each observation and the functions you need
# to calculate the sum of squared errors of the no improvement model.
```

16. What is the value of the Bayesian Information Criterion for model 1?

```
# In this code chunk you can write the functions that you need in order to
# calculate the BIC associated to model 1. Save it as a new variable and
# report the value as text below.
```

ANS: [WRITE YOUR ANSWER HERE]

17. Write the second model (**improvement**) in mathematical notation.

ANS: [WRITE YOUR ANSWER HERE]

18. Make a density graph of model 2 (**improvement**). Remember that at this point the units (values of x) don't matter, we are only interested in what the model expects the data to look like.

```
# In this code chunk you can write the code used to generate the density plot
# for model 2. Remember that the values on the x axis are not important, what we
# look at is what the model expects the data to look like.
```

19. Calculate the prediction of the second model and add it as a new variable (column) to the data set (remember to give the variable a different name!). What are the prediction of the second model?

```
# In this code chunk you can write the functions that you need to use in order
# to calculate the predictions of model two and add it to the data set as a new
# column. Report the value(s) as text below
```

ANS: [WRITE YOUR ANSWER HERE]

20. Calculate the squared error (SE) of each observation for the second model and add it to the data set as a new variable. What is the mean Sum of Squared Errors (m-SSE) of the group differences model (model 2)?

```
# In this code chunk you can write the functions that you need to use in order
# to calculate the squared error of each observation and the functions you need
# to calculate the mean sum of squared errors of the group differences model.
```

ANS: [WRITE YOUR ANSWER HERE]

21. What is the value of the Bayesian Information Criterion for model 2?

```
# In this code chunk you can write the functions that you need in order to
# calculate the BIC associated to model 2. Save it as a new variable and
# report the value as text below.
```

ANS: [WRITE YOUR ANSWER HERE]

22. Which of the two models would you select and why?

ANS: [WRITE YOUR ANSWER HERE]