Homework 2

Write your name here

01 April, 2022

Section 1

Background: For this 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 hypotheses (models). Model 1 states 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). Model 2 predicts that people who drank coffee will be smarter; in other words, the IQ test scores should be different between groups (with group 2 being better than group 1). Both hypotheses assume that variances are equal across observations and groups.

```
# Code 1: Load the data (Hint: Go back to Lec-1 (or homework-1!))
#
# In this chunk you should write (a.k.a. copy from the slides) the code needed
# to load the data for section 1.
```

1. Write the Model 1 (**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 actual values of x don't matter. We are interested in **what the model expects the data to look like**. Do it with a variance equal to 3.

```
# Code 2: Density plot for Model 1
#
# 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. We
# just want to visualize what the model expects the data to look like.
```

3. Test Model 1 by adding the required calculations as a new variable (column) to the data set (remember to give this new variable a name!). What are the predictions that you need to test for the first model?

```
# Code 3: Testing the predictions made by Model 1
#
# In this code chunk you can write the functions that you need to
# calculate the predictions of model 1 and add it to the data set as a new
# column. Report the value(s) as text in the space below.
```

ANS: [WRITE YOUR ANSWER HERE]

4. Calculate the squared error (SE) of each observation under Model 1 by adding a new column to our data set. What is the mean Sum of Squared Errors (m-SSE) of the "no group differences" model (model 1)?

```
# Code 4: Add a Squared Error column and compute the m-SSE for 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 mean sum of squared errors of the "no differences" model.
```

5. What is the value of the Bayesian Information Criterion (BIC) for Model 1?

```
# Code 5: Calculate the BIC for Model 1
#
# In this code chunk you can write the functions that you need in order to
# calculate the BIC associated with Model 1. Save it as a new variable and
# report the value as text below.
```

ANS: [WRITE YOUR ANSWER HERE]

6. Write Model 2 (**groups are different**) in mathematical notation.

ANS: [WRITE YOUR ANSWER HERE]

7. Make a density graph of Model 2 (**coffee group is better**). Remember that the actual values of x don't matter. Right now, we only want to visualize what the model expects the data to look like. Assume that the model variance is equal to 3.

```
# Code 6: Density plot for Model 2
#
# 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
# are looking at is what the model expects the data to look like.
```

8. Test Model 2 by adding the required calculations as a new variable (column) to the data set (remember to give this new variable a DIFFERENT name!). What are the predictions that you need to test for this second model?

```
# Code 7: Testing the predictions made by Model 2
#
# In this code chunk you can write the functions that you need to use in order
# to calculate the predictions of Model 2 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) for each observation under Model 2 by adding a new column to the data set (again, watch out the names you give to the new column!). What is the mean Sum of Squared Errors (m-SSE) of this "group differences" model (Model 2)?

```
# Code 8: Add a Squared Error column and compute the m-SSE for 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 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 (BIC) for Model 2?

```
# Code 9: Calculate the BIC for Model 1
#
# In this code chunk you can write the functions that you need in order to
```

```
# calculate the BIC associated with 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]

12. What does the Model that you selected tell us about our research question?

ANS: [WRITE YOUR ANSWER HERE]

Section 2

Background: For this 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.

Once again, we have two competing models (hypotheses) about the performance of participants on the maze before and after training. Model 1 (**no improvement**) states that the new training method wouldn't make any difference. In other words, Model 1 predicts that there will be no improvement in terms of how quickly people solve the maze after being trained.

Model 2 (**improvement**) states that the new training method should be effective, and so the time people need to solve the maze should be shorter the second time (after participants have undergone the training).

```
# Code 10: Load the data (Hint: Go back to Lec-1 (or homework-1!))
#
# 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.
```

13. Write Model 1 (**no improvement**) in mathematical notation.

ANS: [WRITE YOUR ANSWER HERE]

14. Make a density graph for Model 1 (**no improvement**). To make the plot, assume that the model variance is equal to 2.

```
# Code 11: Density plot for Model 1
#
# In this code chunk you can write the code used to generate the density plot
# for model 1.
```

15. What are the predictions made by Model 1 (**no improvement**)? Again, make sure to test them by adding the corresponding variable (column) to our data base.

```
# Code 12: Testing the predictions made by Model 1
#
# In this code chunk you can write the functions that you need to use in order
# to calculate the predictions of model 1 and add it to the data set as a new
# column. Report the value(s) as text below.
```

ANS: [WRITE YOUR ANSWER HERE]

16. Calculate the squared error (SE) of each observation for Model 1 and add them to the data set as a new variable. What is the mean Sum of Squared Errors (m-SSE) of the "no improvement" model (Model 1)?

```
# Code 13: Adding a column for the SE and calculating the m-SSE
#
# 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.
```

ANS: [WRITE YOUR ANSWER HERE]

17. What is the value of the Bayesian Information Criterion for Model 1?

```
# Code 14: Calculating the BIC 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]

18. Write Model 2 (**improvement**) in mathematical notation.

ANS: [WRITE YOUR ANSWER HERE]

19. Make a density graph for 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.

```
# Code 15: Density plot for Model 2
#
# 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.
```

20. Calculate the predictions of the second model and add them as a new variable (column) to the data set (remember to give the variable a DIFFERENT name!). What are the predictions of the second model?

```
# Code 16: Testing the predictions of Model 2
#
# In this code chunk you can write the functions that you need to use in order
# to calculate the predictions of model 2 and add it to the data set as a new
# column. Report the value(s) as text below
```

ANS: [WRITE YOUR ANSWER HERE]

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

```
# Code 17: Adding a column for the SE and calculating the m-SSE
#
# 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 "improvement" model.
```

ANS: [WRITE YOUR ANSWER HERE]

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

```
# Code 18: Calculating the BIC for Model 2
#
```

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

ANS: [WRITE YOUR ANSWER HERE]

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

ANS: [WRITE YOUR ANSWER HERE]

24. What does the Model that you selected tell us about our research question?

ANS: [WRITE YOUR ANSWER HERE]