**Dear all students,**

**Follow the instructions carefully, please:**

* You can find the following three questions in the a and b parts. You have **120 minutes** to answer the questions and **10 minutes of upload** time (not extendable).
* For the questions that you need to answer in the text, add your answer in this world file and upload it into Moodle. For the questions that you need to write code, upload your Python file into Moodle. Rename both files using your English name or student ID.
* You can use Slides and class code examples only.
* No need to mention using the Internet or communicating during the exam is prohibited.
* Cell phones are only in your pocket at all.
* **Important:** Attach the Word file and Python code (\*.py) to Moodle (do not compress files as .zip or .rar files).

**Good Luck.**

--------------------------------------------------------------------------------------------------

English Name:

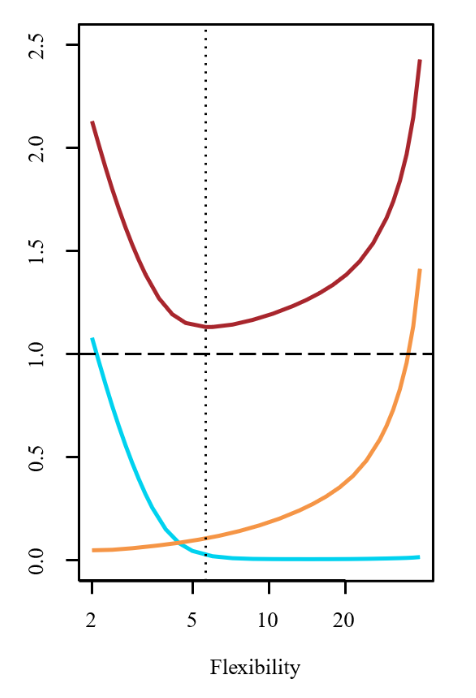
Chinese name:

Student ID:

--------------------------------------------------------------------------------------------------

**Questions:**

1. (a) Why is the error in regression model predictions not reproducible completely (Irreducible)? Explain the irreducible error by example. (b) Explain clearly what we can understand in the following Figure (what are two dash lines), and answer why we need to plot it.



Answer:

1. (a) Consider advertising on Facebook, YouTube, and Instagram for a product, then write the complete linear regression model ( for it by considering Interactions Between independent variables. (b) add a categorical independent variable for the previous model as an advertisement for young (under 18 years old) and Adults (equal or over 18 years old).

Answer:

1. Implement the SGD based on a regression model on the following data report: a) Convergence plot, b) the LOOCV cross-validation results, c) P-value, and null hypotheses after convergence.

import pandas as pd

data = pd.DataFrame({

    'FaceBook': [340.1, 154.5, 127.2, 261.5, 290.8, 115.7, 167.5, 230.2, 115.6, 309.8, 176.1, 324.7, 130.8, 207.5, 314.1, 305.4, 177.8, 391.4, 179.2, 257.3],

    'Youtube': [147.8, 150.3, 157.9, 151.3, 133.8, 156.9, 140.8, 127.6, 109.1, 109.6, 112.8, 131.0, 142.1, 130.6, 140.9, 170.7, 149.6, 153.6, 130.5, 130.9],

    'Instagram': [169.2, 145.1, 169.3, 157.5, 157.4, 182.0, 130.5, 118.6, 108.0, 128.2, 123.2, 103.0, 164.9, 106.2, 145.0, 151.9, 213.0, 154.8, 117.3, 126.1],

    'Sales': [29.1, 17.4, 16.3, 25.5, 19.9, 14.2, 18.8, 20.2, 11.8, 18.6, 16.6, 23.4, 15.2, 15.7, 26.0, 29.4, 19.5, 31.4, 18.3, 21.6]})

**Note:** You can use the homework that you already extended based on the given class example as homework (Chapter 5). You are allowed to use the same codes to answer this question, but do not use libraries to solve the SGD.