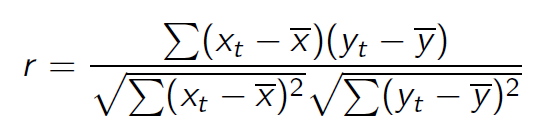
Multivariate Modeling

DATS 6450

LAB#2- Correlation coefficient and Scatter plot

Using the Python program and “pandas”, “matplotlib.pyplot” and “numpy” library perform the following tasks:

1. Write a python function called “ correlation\_coefficent\_cal(x,y)” that implement the correlation coefficient. The formula for correlation coefficient is given below. The function should be written in a general form than can work for any dataset x and dataset y. The return value for this function is r.



1. Test the “ correlation\_coefficent\_cal(x,y)” function with the following simple dataset. The x and y here are dummy variable and should be replaced by any other dataset.

X = [1, 2, 3, 4, 5]

Y = [1, 2, 3, 4, 5]

Z = [-1, -2, -3, -4,-5]

G = [1,1,0,-1,-1,0,1]

H = [0,1,1,1,-1,-1,-1]

* 1. Plot the scatter plot between X, Y
  2. Plot the scatter plot between X, Z
  3. Plot the scatter plot between G, H
  4. Without using Python program, implement the above formula to derive the r\_xy, r\_xz, r\_gh. You should NOT use computer to answer this section. You need to show all your work for this section on the paper.
  5. Calculate r\_xy , r\_xz and r\_gh using the written python function “correlation\_coefficent\_cal(x,y)”.
  6. Compare the answer in section d and e. Any difference in value?
  7. Display the message as:
     1. The correlation coefficient between x and y is \_\_\_\_\_\_
     2. The correlation coefficient between x and z is \_\_\_\_\_\_
     3. The correlation coefficient between g and h is \_\_\_\_\_\_
  8. Add an appropriate x-axis label and y-axis label to all your scatter graphs.
  9. Include the r\_xy , r\_xz, r\_gh as a variable on the scatter plot title in part a and part b. The code should be written in a way that the r value changes on the figure title automatically. Hint: You can use the following command:

*plt.title("Scatter plot of X and Y with r ={}".format(r\_xy))*

* 1. Does the calculated r\_xy, r\_xz and r\_gh make sense with respect to the scatter plots? Explain why?

**Submission guidelines:**

1. Write a formal report for this LAB and include graphs (add appropriate title, x-label, y-label, and legend if applies) and the developed python code.
2. The developed python must be included at the end of your report as an appendix.
3. The softcopy of the developed Python code .py must also be submitted separately. Please make sure the developed python code runs without any error by testing it through Pycharm software. **The developed python code with any error will subject to 50% points penalty.**
4. Make sure to answer all the questions asked. Your report should be written in a formal format. A formal report includes a cover page, abstract, introduction, answering questions, conclusion, appendix, and references. Upload a formal **report (as a single pdf**) plus **the .py file** through BB by the due date.

1- Cover Page:   A formal report must have a cover page  which includes :

|  |
| --- |
| * Course number and name. * Instructor(s) name(s). * Lab number or title * Your name * Initials and Date. |

2- Abstract :

|  |
| --- |
| A brief description of the LAB. The abstract should not exceed four or five sentences. The abstract should be self contained and meaningful on its own with the intent that any reader can understand the primary purpose and findings of the lab at a very high summary level. |

3- Introduction : In your own words, explain the reason for performing the experiment and give a concise summary of the theory involved, including any mathematical detail relevant to later discussion in the report.

4- Method, theory, and procedures:  A brief explanation of the theory and method behind the LAB.

5- Answer to asked questions:     Make sure to answer all the questions asked in the LAB. You may need to add figures, table or chart in this section. All figures should have correct x-label, y-label, title and legend. This also applies to all charts and tables.

6- Conclusion :  This section should reflect your understanding of the LAB. Important points to include are a brief discussion of your final results, an interpretation of the actual experimental results as they apply to the objectives of the LAB set out in the introduction should be given. Also discuss any problems encountered and how they were resolved.

7- Appendix:  A copy of the developed python code to be added in this section. Make sure to run your Python code in Pycharm before submission and make sure the code runs without an error. **50%-point reduction for a LAB or homework submission with an error in Python code**.

8- References (if applies)