23MAT204 – Mathematics for Intelligent Systems - 3
Practise Sheet-8
(Covariance, Correlation)

- **Evaluation of Covariance and Correlation coefficient**

$$\mathrm{Cov}(X_1,X_2) = \frac{1}{N-1}\sum_{i=1}^{N}(x_{1i}-\bar{x}_1)(x_{2i}-\bar{x}_2),$$

$$\text{where } \bar{x}_1 = \frac{1}{N}\sum_{i=1}^{N} x_{1i}\ , \bar{x}_2 = \frac{1}{N}\sum_{i=1}^{N} x_{2i}$$

$$\boldsymbol{\rho = \frac{Cov(X_1,X_2)}{\sigma_{X_1}\sigma_{X_2}} = \frac{\frac{1}{N-1}\sum_{i=1}^{N}(x_{1i}-\bar{x}_1)(x_{2i}-\bar{x}_2)}{\left(\frac{1}{N-1}\sum_{i=1}^{N}(x_{1i}-\bar{x}_1)^2\right)^{1/2}\left(\frac{1}{N-1}\sum_{i=1}^{N}(x_{2i}-\bar{x}_2)^2\right)^{1/2}}}$$

1. Ten students got the following marks in Mathematics in XII board exam (X) and in the Engineering entrance exam(Y). Calculate the coefficient of correlation:

Student:	1	2	3	4	5	6	7	8	9	10
XII marks:	78	40	94	22	76	84	90	62	65	39
Engg marks:	84	51	91	60	68	62	86	58	53	47

Correlation using excel

	A	B	C	D	E	F	G	H	I
1									
2		**X**	**Y**	**X-Xbar**	**Y-Ybar**	**(X-Xbar)(Y-Ybar)**	**(X-Xbar)^2**	**(Y-Ybar_^2)**	
3		78	84	13	18	234	169	324	
4		40	51	-25	-15	375	625	225	
5		94	91	29	25	725	841	625	
6		22	60	-43	-6	258	1849	36	
7		76	68	11	2	22	121	4	
8		84	62	19	-4	-76	361	16	
9		90	86	25	20	500	625	400	
10		62	58	-3	-8	24	9	64	
11		65	53	0	-13	0	0	169	
12		39	47	-26	-19	494	676	361	
13									
14	mean	**65**	**66**			**284**	**586.22222**	**247.1111111**	
15									
16									
17						Correlation=	0.7461756		
18							=F14/(SQRT(G14)*SQRT(H14))		

Correlation using MATLAB

```
X=[78 40 94 22 76 84 90 62 65 39];
Y=[84 51 91 60 68 62 86 58 53 47];
C=cov(X,Y) % Covariance matrix
R=C(1,2)/(sqrt(C(1,1)).*sqrt(C(2,2))) % Correlation coefficient
```

Practise Questions:

1. Find the correlation coefficient for the following data and explain about the strength of correlation.

X	2	4	5	6	8	11
Y	18	12	10	8	7	5

2. Janice and Paul did a study on feelings of stress and life satisfaction. Participants completed a measure on how stressed they were feeling (on a 1 to 30 scale) and a measure of how satisfied they felt with their lives (measured on a 1 to 10 scale). The table below indicates the participants' scores. Draw the scatter plot of the data and evaluate the correlation coefficient for the case.

Participant #	Stress score (X)	Life Satisfaction (Y)
1	11	7
2	25	1
3	19	4
4	7	9
5	23	2
6	6	8
7	11	8
8	22	3
9	25	3
10	10	6

3. At Hogwarts School of Witchcraft and Wizardry, students often have a lot of homework. The table below indicates the number of hours students studied, and how they performed on an exam in two of their classes.

Student	Potions		Defense against the dark arts	
	study hours	exam score	study hours	exam score
1	3	75	4	70
2	15	95	12	98
3	6	65	9	85
4	8	70	6	80
5	4	85	2	65
6	2	80	3	75
7	10	65	10	92
Mean	6.86	76.43	6.57	80.71
SD	4.22	10.25	3.54	10.95

   a. Find the correlations between hours spent studying and how students performed in their potions and defense against the dark arts classes.

   b. Which class was more strongly correlated with studying?

4. The marks of 10 engineering students in Engineering Mechanics(X) and Engineering Mathematics(Y) are given below. Find the correlation coefficient for this data and explain about the strength of correlation.

X:	78	40	94	22	76	84	90	62	65	39
Y:	84	51	91	60	68	62	86	58	53	47

5. Find the correlation coefficient for the following data and explain about the strength of correlation.

X:	1	4	7	9	12	15	21	13
Y:	90	99	85	76	70	65	55	56

6. Find the correlation coefficient for the following data and explain about the strength of correlation.

X:	5	8	12	3	7	9	15	1
Y:	21	28	30	14	23	31	40	9