

Department of Mathematics
MTL 390 (Regression)
Tutorial Sheet No. 6

- Show that for any collection (X_1, X_2, \dots, X_n) of random variables the covariance matrix $\Sigma = (Cov(X_i, X_j))$ which is symmetric, is always positive definite.
- A sample of size 27 from a bivariate normal population had an observed correlation of 0.2. Can you discard the claim that components are independent? Use 5% level of significance.
- A sample of size 100 from a normal population had an observed correlation of 0.6. Is the shortfall from the claimed correlation of at least 0.75 significant at 5% level of significance? What would a confidence interval for the correlation coefficient be at 95% level of significance?
- A random sample of size 8 from a bivariate normal distribution yields a value of the correlation coefficient of 0.75. Would we accept or reject at the 5% significance level, the hypothesis that $\rho = 0$.
- A sample of 10 students training as analysts are ranked by a tutor at the end of the course according to (a) suitability to their career (b) knowledge. are given in Table 1.

Table 1: Ranks given by tutor

Student	A	B	C	D	E	G	H	I	J	
Ranks on (a)	4	10	3	1	9	2	6	7	8	5
Ranks on (b)	5	8	6	2	10	3	9	4	7	1

Find Kendall rank correlation coefficient.

- Ten competitors in a dance competition were ranked by three judges A, B and C in the following order 2:

Table 2: Ranks given by judges

Ranks by A	1	6	5	10	3	2	4	9	7	8
Ranks by B	3	5	8	4	7	10	2	1	6	9
Ranks by C	6	4	9	8	1	2	3	10	5	7

Using Spearman rank correlation method, discuss which pair of judges has the nearest approach to common likings in dance.

- Consider the bivariate data $(x_i, y_i), i = 1, 2, \dots, 12$ given in Table 3

Table 3: Data

x	32	41	31	38	21	13	17	22	24	11	17	20
y	6	5	3	3	7	9	9	8	6	9	7	8

Find the Spearman rank correlation coefficient. Test the null hypothesis that the variables do not have a rank order relationship in the population represented by the sample.

- For a bivariate data set $(x_i, y_i), i = 1, 2, \dots, n$, suppose the least square regression lines are $5x - 8y + 14 = 0$ and $2x - 5y + 11 = 0$. Show that
 - Standard deviation of y is less than standard deviation of x .
 - correlation coefficient between x and y is 0.80.

- (c) mean of x and y are 2 and 3 respectively.
9. Consider the following five observations on $(X, Y) : (0, 1), (1, 2), (2, 3), (3, 2), (4, 1)$. Find the correlation coefficient between X, Y . Also, show that the two lines of regression are perpendicular to each other.
10. Students' scores in the Probability course examination, x , and on the semester CGPA, y are given in Table 4:

Table 4: Scores in Probability course

x	4	2	5	3	2	3	4	3	5	2
y	3.12	3.00	4.5	4.75	3	3.5	3.75	4.12	4.54	3.1

- (i) Calculate the least square regression line for the data.
- (ii) Plot the points and the least square regression line on the same graph.
- (iii) Find point estimates for α, β and σ^2 .
- (iv) Find 95% confidence interval for the α and β under the usual assumptions.
11. Can $Y = 9 + 2.9X$ and $X = 3 - 0.4Y$ be the estimated regression equations of Y on X and X on Y respectively? Explain?