Statistics for Decision Making – I

End Semester Examination, 2025

Full Marks: 50 Time : 2 hr 30 minutes

Answer 1, 2 and 3 and any two from the rest.

1. Following are the scores of 8 students in pre-test (x3), test (x2) and final examination (x1). If a multiple linear regression equation is to be set,
   i. Check how good the regression equation will perform?
   ii. Among the pre-test and test scores, which one has closer alignment with the final scores? Give your remark. (3+4)

Student	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Pre-test	43	38	27	28	35	21	19	13
Test	22	29	23	33	20	8	17	19
Final	66	38	55	63	25	17	33	18

2. i. The following data shows detection cases of Tuberculosis of different types among male and female residents of a city. Check whether gender has any impact on type of tuberculosis. (4)

Type of Tuberculosis	Gender	
	Male	Female
Tuberculosis of respiratory system	3534	1319
Other forms of Tuberculosis	270	252

ii. The following data give size of each household in a block of 40 households. (9)

6	5	4	7	4	5	3
4	4	2	5	5	8	2
4	5	3	5	2	3	2
3	2	6	2	2	2	4
2	3	1	3	3	4	5

Draw a SRSWR and SRSWOR both of size 10. In each case, estimate the average marks and also estimate the corresponding standard error.

3. Consider the following bivariate data (in ‘000 $) on return-on-investment in a certain market. Is there any outlier and / or leverage point in the data? If yes, can it/they be considered influential? Justify. (10)

Investment	Return
12	45
15	54
17	61

13	51
18	70
20	73
22	71
18	66
25	106

4. i. A total of 1500 children have a rapid strep test (RST) done by a standardized culture technique. Among them, 1338 have a negative RST and 162 have a positive RST. In addition, a backup throat culture (gold standard) was done on all children. Of those children with a negative RST, 1302 have a negative throat culture while in the group with a positive RST, 159 have a positive throat culture. Calculate the sensitivity and specificity of the RST. (4)

ii. If X_1 and X_2 are independently distributed random variables each following R(0, 1), then show that $U_1 = \sqrt{-2\ln X_1}\ Cos\ 2\pi\ X_2$ and $U_2 = \sqrt{-2\ln X_1}\ Sin\ 2\pi\ X_2$ are independently distributed. Can you identify their distributions? (6)

5. i. Show that, under the set-up of multiple linear regression, the regressed values are independent of the corresponding errors. (5)
   ii. Derive the expression for the sum of squares of errors under the same set-up. (5)
6. i. Statin is a regular drug used to lower the cholesterol level in blood. However, some Statin users find it causing sleep disorder. Based on the following data compute odds ratio and make your comment on the impact of Statin in sleep disorder. (4)

	Sleep disorder	Sleep ok
Statin User	14	91
Non-statin User	28	369

   ii. For proportional allocation, define the stratified random sampling estimator of the population mean. Show that it is an unbiased estimator. Derive the expression for its variance under SRSWOR. (6)