

ISLAMIC UNIVERSITY OF TECHNOLOGY (IUT)
ORGANISATION OF ISLAMIC COOPERATION (OIC)
DEPARTMENT OF CIVIL AND ENVIRONMENTAL ENGINEERING

Semester Final Examination
Course No.: CEE 4655
Course Title: Civil Engineering Data Analysis

Summer Semester: 2021-2022
Full Marks: 150
Time: 3.0 Hours

There are 8 (Eight) questions. Question No. 1, 2, 3, 4 are compulsory. Answer any 2 (two) questions from Question No. 5, 6, 7 and 8. Programmable calculators are not allowed. It's an OPEN BOOK exam. Only ONE TEXT BOOK is allowed in the exam. Do not write on this question paper. The figures in the right margin indicate full marks. The Symbols have their usual meaning.

- 1(a) If the probability density function of a random variable is given by (10)
 $f(x) = k(1 - x^2) \quad 0 < x < 1$ (CO1-PO1)
 $= 0 \quad \text{otherwise}$

Find the value of k and the probabilities that a random variable having this probability density will take on a value(i) between 0.1 and 0.2, and (ii) greater than 0.5.

- (b) An insurance company insured 4000 people against loss of both eyes in a car accident. Based on previous data, the rates were computed on the assumption that on the average, 10 persons in 100000 will have car accidents each year that result in this type of injury. What is the probability that more than 3 of the insured will collect on their policy in a given year? (05)
(CO1-PO1)

- (c) The incidence of an occupational disease in an industry is such that the workers have a 20% chance of suffering from it. What is the probability that out of 6 workers chosen at random, four or more will suffer from the disease? (05)
(CO1-PO1)

- (d) A discrete random variable X has the following distribution function: (10)
(CO1-PO1)

$$F(x) = \begin{cases} 0 & x < 1 \\ \frac{1}{3} & 1 \leq x < 4 \\ \frac{1}{2} & 4 \leq x < 6 \\ \frac{5}{6} & 6 \leq x < 10 \\ 1 & x \geq 10 \end{cases}$$

Find (i) $P(2 < X \leq 6)$ (ii) $P(X = 5)$ (iii) $P(X = 4)$ (iv) $P(X \leq 6)$ and (v) $P(X = 6)$

2. An article in *Wear* (Vol. 152, 1992, pp. 171- 181) presents data on the fretting wear of mild steel and oil viscosity. Representative data follow, with x =oil viscosity and y = wear volume (10^{-4} mm³). (30)
(CO3-PO4)
- Using a graph paper construct a scatter plot of data. Does a simple linear regression model appear to be plausible?
 - Fit the simple liner regression model using least squares.
 - Obtain the fitted value of y when $x=22.0$ and calculate the corresponding residual.
 - Test at the 5% level of significance for the statistical significance of the parameter x .
 - Determine and interpret the coefficient of determination.
3. Given the following regression output: (30)
(CO3-PO4)

Predictor	Coefficient	Standard Error	T	p-value
X1	2.391	1.200	1.99	0.051
X2	-0.4086	0.1717	-2.38	0.020
Constant	84.998	1.863	45.61	0.000

Analysis of Variance					
Source	Degrees of Freedom	SS	MS	F	P
Regression	2	77.907	38.954	4.14	0.021
Residual Error	62	583.693	9.414	-	-
Total	64	661.600	-	-	-

Answer the following:

- Write the regression equation
 - If X_1 is 4 and X_2 is 11, what is the value of the dependent variable?
 - How large is sample? How many independent variables are there?
 - Conduct F-test of hypothesis to see if any of the set of regression coefficients could be different from 0. Use the 0.05 significance level? What is your conclusion?
 - Conduct a test of hypothesis for each independent variable. Use the 0.05 significance level. Which variable would you consider eliminating?
- 4(a) Before an increase in excise duty on tea, 800 people out of a sample of 1000 were consumers of tea. After an increase in exercise duty, 800 people were consumers of tea in a sample of 1200 persons. Find whether there is significant decrease in the consumption of tea after the increase in duty. ($\alpha = 0.05$) (10)
(CO2-PO2)
- (b) The breaking strengths of cables produced by a manufacturer have mean 1800 lb and standard deviation 100lb. By a new technique in the manufacturing process it is claimed that the breaking strength can be increased. To test this claim, a sample of 50 cables is tested, and it is found that the mean breaking strength is 1850 lb. (i) (10)
(CO2-PO2)

Can we support the claim at a 0.01 level of significance? (b) What is the P value of the test?

5. Two researchers adopted different sampling techniques while investigating some group of students to find the number of students falling into different intelligence level. The results are as follows: (20)
(CO2-PO2)

Researchers	Below Average	Average	Above average	Genius
X	86	60	44	10
Y	40	33	25	2

Would you say that the sampling techniques adopted by the two researchers are significantly different? ($\alpha=0.1$)

6. Civil Engineering students were given different treatments before revising for their exams. Some were given a memory drug, some a placebo drug and some no treatment. The exam scores(%) are shown below for the three different groups: (20)
(CO2-PO2)

Memory Drug	Placebo	No treatment
70	37	3
77	43	10
83	50	17
90	57	23
97	63	30

Carry out an ANOVA to test the hypothesis that the treatments will have different. ($\alpha=0.05$).

- 7(a) A test of the breaking strengths of 6 ropes manufactured by a company showed a mean breaking strength of 7750 lb. and a standard deviation of 145 lb., whereas the manufacturer claimed a mean breaking strength of 8000 lb. Can we support the manufacturer's claim at a level of significance (i) 0.05 (ii) 0.01? (iii) What is the P value of the test? (10)
(CO2-PO2)

- (b) Lester Hollar is Vice President for human resources for a large manufacturing company. In recent years he has noticed an increase in absenteeism that he thinks is related to the general health of the employees. Four years ago, in an attempt to improve the situation, he began a fitness program in which employees exercise during their lunch hour. To evaluate the program, he selected a random sample of eight participants and found the number of days each was absent in the six months before the exercise program began and in the last six months. Below are the results. At the 0.05 significance level, can he conclude that the number of absences has declined? Estimate the p -value. (10)
(CO2-PO2)

Employee	Before	After
1	6	5
2	6	2
3	7	1
4	7	3
5	4	3
6	3	6
7	5	3
8	6	7

- 8(a) At an agricultural station it was desired to test the effect of a given fertilizer on wheat production. To accomplish this, 24 plots of land having equal areas were chosen: half of these were treated with the fertilizer and the other half were untreated (control group). Otherwise the conditions were the same. The mean yield of wheat on the untreated plots was 4.8 bushels with a standard deviation of 0.40 bushels, while the mean yield on the treated plots was 5.1 bushels with a standard deviation of 0.36 bushels. Can we conclude that there is a significant improvement in wheat production because of the fertilizer if a significance level of (i) 1% (ii) 5% is used?(iii) What is the P value of the test? (10)
(CO2-PO2)
- (b) The assembly area of Gotrac Products was recently redesigned. Installing a new lighting system and purchasing new workbenches were two features of the redesign. The production supervisor would like to know if the changes resulted in an improvement in worker productivity. To investigate she selected a sample of 11 workers and determined the production rate before and after the changes. The sample information is reported below. (10)
(CO2-PO2)

Operator	Production Before	Production
S. M.	17	18
D. J.	21	23
M. D.	25	22
B. B.	15	25
M. F.	10	28
A.A.	16	16
U.Z.	10	22
Y. U.	20	19
U.T.	17	20
Y.H.	24	30
Y.Y.	23	26

Use the Wilcoxon signed rank test to determine whether the new procedures actually increased production. Use the 0.05 level.