

GUJARAT TECHNOLOGICAL UNIVERSITY
BACHELOR OF ENGINEERING – SEMESTER 3 – EXAMINATION – SUMMER 2026

Subject Code: BE03013041

Date: 02/07/2026

Subject Name: STATISTICAL TECHNIQUES FOR ENVIRONMENTAL ENGINEERING

Time: 02:30 PM TO 05:00 PM

Total Marks: 70

Instructions:

- 1. Attempt all questions.**
- 2. Make suitable assumptions wherever necessary.**
- 3. Figures to the right indicate full marks.**
- 4. Simple and non-programmable scientific calculators are allowed.**

- Q.1** (a) Discuss the significance of statistical techniques for Environmental Engineering. **03**
(b) Define the following common statistical terms: i. Population, ii. Sample, iii. Variable, iv. Data **04**
(c) Enlist and explain primary and secondary sources for data collection. **07**
- Q.2** (a) Differentiate between mean and median. Give one example with data where median is more suitable than mean. **03**
(b) Explain partition values (quartiles, deciles, percentiles) and their application in environmental data and justify why the inter-quartile range (IQR) is preferred over the range for assessing variability. **04**
(c) The following are the daily Dissolved Oxygen (DO) levels (in mg/L) recorded in a river over 12 days: 6.5, 7.2, 6.8, 7.0, 6.9, 7.1, 7.0, 6.7, 7.3, 6.8, 7.0, 7.2. Calculate the mean, median, mode, range and mean deviation of the DO levels. **07**
- OR
- Q.2** (c) A pilot-scale wastewater treatment unit's efficiency in removing a specific pollutant (in %) was measured over 30 trials: 88, 72, 65, 98, 55, 81, 90, 70, 50, 88, 73, 94, 68, 84, 59, 93, 78, 62, 86, 71, 96, 69, 83, 58, 91, 76, 64, 87, 74, 60. Calculate the Range, Mean Deviation, Standard Deviation, Coefficient of Variation. Also, based on these measures, comment on the consistency of the treatment unit's performance. **07**
- Q.3** (a) Differentiate between histogram and frequency polygon. Discuss their applications in analysing air quality data. **03**
(b) Identify the type of data most suitable for a pie chart and mention the key point to be considered while preparing a pie chart. **04**
(c) A local environmental agency collected daily PM_{2.5} concentration data (in $\mu\text{g}/\text{m}^3$) in Ahmedabad for a month. The data is as follows: PM_{2.5} Concentration Data (n=30, in $\mu\text{g}/\text{m}^3$): 65, 78, 55, 92, 88, 70, 68, 85, 73, 60, 58, 95, 102, 80, 75, 77, 82, 63, 67, 90, 84, 76, 71, 59, 98, 86, 74, 62, 79, 81. Prepare a grouped and cumulative frequency distribution table for PM_{2.5} concentrations, construct a histogram and cumulative frequency curve. Determine and interpret the 90th percentile value in terms of air quality for Ahmedabad. **07**
- OR
- Q.3** (a) Enlist and explain two graphical methods for representing qualitative environmental data with example. **03**
(b) Identify the type of data most suitable for a bar chart and mention the key points to be considered while preparing a bar chart. **04**

- (c) An environmental engineer collected data on the primary type of air pollution control device (APCD) used at 50 local manufacturing facilities. 07

The data categories are: ESP: Electrostatic Precipitator, WSC: Wet Scrubber, BAG: Baghouse Filter, OTH: Other/None
 ESP, WSC, BAG, ESP, BAG, WSC, ESP, BAG, OTH, ESP, WSC, BAG, ESP, BAG, WSC, ESP, BAG, WSC, ESP, BAG, OTH, ESP, WSC, BAG, ESP, WSC, BAG, ESP, WSC, BAG, OTH, ESP, BAG, ESP, WSC, ESP, BAG, WSC, ESP, BAG, OTH, WSC.

i. Prepare a table showing the Frequency and Relative Frequency (%) of each APCD type. ii. Calculate the Central Angle for all four APCD categories and describe how the calculated angles are used to construct the Pie Diagram. iii. Draw the bar diagram representing the frequency of the APCDs.

- Q.4** (a) Explain the concepts of skewness and kurtosis in data distribution. Mention any one measure used for each. 03

- (b) Describe the significance of skewness and kurtosis in analyzing environmental datasets. 04

- (c) An environmental engineer collected daily PM_{2.5} concentration data ($\mu\text{g}/\text{m}^3$) from an urban monitoring station for 15 days: 12, 15, 18, 20, 22, 24, 25, 26, 28, 30, 35, 38, 45, 50, 60. 07

(i) Calculate and interpret the skewness of the data. (ii) Discuss the kurtosis and comment on whether the distribution is more peaked or flat compared to a normal distribution. (iii) Explain what these results indicate about air pollution variability in the area.

OR

- Q.4** (a) Differentiate between correlation and regression with example. 03

- (b) Explain the concept of linear regression analysis in environmental engineering with example. 04

- (c) An environmental engineer studied the relationship between rainfall (mm) and BOD levels (mg/L) in a river over six sampling events. Calculate and interpret the correlation coefficient with regression equation between rainfall and BOD. The data are as follows: 07

Rainfall (mm)	10	20	30	40	50	60
BOD (mg/L)	8	7	6	5	4	3

- Q.5** (a) Differentiate between large and small sample tests with suitable examples. 03

- (b) Explain probability and its use in estimating the risk of exceeding groundwater contamination limits. 04

- (c) An environmental engineer measured lead concentrations (mg/L) in water from two different industrial sites over five days: 07

Day	1	2	3	4	5
Site A	0.12	0.15	0.14	0.13	0.16
Site B	0.1	0.11	0.13	0.12	0.14

Determine if there is a significant difference in the variances of lead concentrations between Site A and Site B along with its interpretation.

OR

- Q.5** (a) Explain the basic laws of probability with an example related to environmental risk assessment. 03

- (b) Explain the concept of Analysis of Variance (ANOVA). Compare one-way and two-way ANOVA used for pollutant levels across different locations and time periods. 04

- (c) An environmental engineer surveyed presence of a particular aquatic species in three different rivers to see if distribution is independent of location:

07

River	Species Present	Species Absent
A	15	5
B	10	10
C	5	15

Check and interpret whether the presence of the species is independent of the river. Explain the ecological implications.
