



**Faculty of Cyber Physical System**  
**Department of IoT and Robotics Engineering (IRE)**  
**OBE Description of “Probability and Statistics”**  
**For the**  
**Undergraduate Programs of BDU**  
**Md. Mehedi Hasan**  
**Lecturer (Mathematics)**  
**Department of General Education, BDU**

**Assignment-1**

**Question-1**

A local gas station collected data from the day’s receipts, recording the gallons of gasoline each customer purchased. The following table lists the frequency distribution of the gallons of gas purchased by all customers on this one day at this gas station.

Gallons of Gas	Number of Customers
0 to less than 4	31
4 to less than 8	78
8 to less than 12	49
12 to less than 16	81
16 to less than 20	117
20 to less than 24	13

- How many customers were served on this day at this gas station?
- Find the class midpoints. Do all of the classes have the same width? If so, what is this width? If not, what are the different class widths?
- Prepare the relative frequency and percentage distribution columns.
- What percentage of the customers purchased 12 gallons or more?
- Explain why you cannot determine exactly how many customers purchased 10 gallons or less.

### Question-2

A data set on money spent on lottery tickets during the past year by 200 households has a lowest value of \$1 and a highest value of \$1167. Suppose we want to group these data into six classes of equal widths.

- Assuming that we take the lower limit of the first class as \$1 and the width of each class equal to \$200, write the class limits for all six classes.
- What are the class boundaries and class midpoints?

### Question-3

Each state collects information on every birth that occurs within its borders. The following data give the 2008 birth rates (number of births per 1000 people) for all of the 56 counties in the state of Montana.

10.1	22.2	15.8	12.2	7.7	3.1	14.5	7.8	13.6	8.8	10.9	8.9	14.7	9.6	14.2	14.9	18.3	22.8	5.4	5.6	
19.6	8.2	9.9	14.7	13.7	10.3	9.7	9.8	8.6	9.4	14.1	12.3	10.5	11.4	2.2	9.8	10.9	4.6	6.6	8.5	10.2
14.4	20.4	18.5	10.8	6.5	11.6	12.1	10.5	9.3	8.1	7.4	10.2	9.7	5.6	14.5						

- Construct a frequency distribution table.
- Calculate the relative frequency and percentage for each class.
- Construct a histogram, bar diagram and a polygon for the birth-rate percentage distribution.
- What percentage of the counties had a birth rate of less than 11 births per 1000 people?

### Question-4

The following data represent the numbers of tornadoes that touched down during 1950 to 1994 in the 12 states that had the most tornadoes during this period. The data for these states are given in the following order: CO, FL, IA, IL, KS, LA, MO, MS, NE, OK, SD, and TX. 1113 2009 1374 1137 2110 1086 1166 1039 1673 2300 1139 5490.

- Construct a frequency distribution table and finding cumulative frequency.
- Calculate the mean and median for these data and find mode.
- Identify the outlier in this data set. Drop the outlier and recalculate the mean and median. Which of these two summary measures changes by a larger amount when you drop the outlier?
- Which is the better summary measure for these data, the mean the median or mode? Explain.
- Draw bar diagram, pie chart, histogram, and polygon ogive line.

### Question-5

The following data give the number of shoplifters apprehended during each of the past 8 weeks at a large department store. 7 10 8 3 15 12 6 11

- a. Find the mean for these data. Calculate the deviations of the data values from the mean. Is the sum of these deviations zero?
- b. Calculate the range, variance, and standard deviation, all quartile, 50<sup>th</sup> percentile.

#### Question-6

Consider the following frequency distribution for a set of data:

Class interval	Frequency
10-20	5
20-30	8
30-40	12
40-50	7
50-60	4

- a) Find the mean of the data.
- b) Determine the median of the data.
- c) Identify the mode, if it exists.
- d) Calculate all quartile, 25<sup>th</sup> and 75<sup>th</sup> percentile
- e) Find mean deviation and coefficient of mean deviation range coefficient, quartiles deviation, variance, standard deviation.
- f) Also draw bar diagram, pie chart, histogram, and polygon ogive line.

#### Question-7

The following are the scores of 30 college students on a statistics test. 75 52 80 96 65 79 71 87 93 95 69 72 81 61 76 86 79 68 50 92 83 84 77 64 71 87 72 92 57 98

- a. Construct a stem-and-leaf display.
- b. Construct a dot plot.
- c. Explain which is more efficient for drawing conclusion about scattering.

#### Question-8

What is kurtosis? What does it measure? How does the measure of kurtosis help in understanding a frequency distribution? Describe the measures kurtosis you are familiar with. The first four moments of a distribution, about 4 were found to be -1.5, 17, -30 and 108.

- a. Calculate the corresponding central moments.
- b. Hence find  $\beta_1$  and  $\beta_2$  and comment on the skewness and kurtosis of the distribution.
- c. Explain which is more efficient for the distribution.

#### Question-9

An auto manufacturing company wanted to investigate how the price of one of its car models depreciates with age. The research department at the company took a sample of eight cars of this model and collected the following information on the ages (in years) and prices (in hundreds of dollars) of these cars.

Age	8	3	6	9	2	5	6	3
Price	45	210	100	33	267	134	109	235

- Construct a scatter diagram for these data. Does the scatter diagram exhibit a linear relationship between ages and prices of cars?
- Find the regression line with price as a dependent variable and age as an independent variable.
- Give a brief interpretation of the values of a and b calculated in part b.
- Plot the regression line on the scatter diagram of part a and show the errors by drawing vertical lines between scatter points and the regression line also find correlation coefficient and regression coefficient and make sure this data is perfectly scatter or not.
- Predict the price of a 7-year-old car of this model.
- Estimate the price of an 18-year-old car of this model. Comment on this finding.

### Question-10

The following table gives data on variables y,  $x_1$  and  $x_2$

y	$x_1$	$x_2$
24	98	52
14	51	69
18	74	63
31	108	35
10	33	88
29	119	54
26	99	51
33	141	31
13	47	67
27	103	41
26	111	46

Using MINITAB, find the regression of y on and using the solution obtained, answer the following questions.

- Write the estimated regression equation.
- Explain the meaning of the estimated regression coefficients of the independent variables.
- What are the values of the standard deviation of errors, the coefficient of multiple determination, and the adjusted coefficient of multiple determination?
- What is the predicted value of y for  $x_1=87$  and  $x_2=54$ ?
- What is the point estimate of the expected (mean) value of y for all elements given that  $x_1=95$  and  $x_2=49$ ?
- Construct a 99% confidence interval for the coefficient of  $x_1$ .

g. Using the 1% significance level, test if the coefficient of  $x_2$  in the population regression model is negative.

**BEST OF LUCK**