

COEP Technological University

Department of Mathematics

(MA- 21001) Probability and Statistics for Engineers

T.Y. B. Tech. Semester V (Computer, Mechanical, Electrical, En TC and Instrumentation Engineering)

Academic Year 2023-24 (Autumn Semester)

Course Coordinator: Dr. Yogita Mahatekar

Tutorial 1 on Unit 1

1. Describe both the population and the observations for the following research questions:

- (a) Evaluation of the satisfaction of employees from an airline.
- (b) Description of the marks of students from an assignment.
- (c) Comparison of two drugs which deal with high blood pressure.

2. Which of the following variables are qualitative, and which are quantitative? Specify which of the quantitative variables are discrete and which are continuous:

Time to travel to work, shoe size, preferred political party, price for a canteen meal, eye colour, gender, wavelength of light, customer satisfaction on a scale from 1 to 10, delivery time for a parcel, blood type, number of goals in a hockey match, height of a child, subject line of an email, Number of books in the classroom, Time it takes for students to finish their quiz, Number of students that have their lunch in the canteen, Car owned by ten friends, Income of 20 employees, Size of clothes as S, M, L, XL, Education of people as High School, Graduate, PG, PhD.

3. Identify the scale of the following variables:

- (a) Political party voted for in an election
- (b) The difficulty of different levels in a computer game
- (c) Production time of a car
- (d) Age of turtles
- (e) Calendar year
- (f) Price of a chocolate bar
- (g) Final ranking at a beauty contest
- (h) Intelligence quotient.

4. Make yourself familiar with the DNase data set from R.

- (a) First, browse through the introduction to R in Appendix A. Then, read in the data.
- (b) View the data both in the R data editor and in the R console.
- (c) Create a new data matrix which consists of the first 5 rows and first 3 variables of the data. Print this data set on the R console. Now, save this data set in your preferred format.
- (d) Add a new variable “product” to the data set which is the product of concentration and density.

5. Identify proper order of various stages in execution of the survey from beginning to end.

- Setting up administrative organization

- Selection, training and supervision of field investigators
- Design of forms
- Processing data
- Control over quality of the field work and field edit
- Follow up of non response
- Preparing Report

(Ans: Setting up administrative organization, Design of forms, Selection, training and supervision of field investigators, Control over quality of the field work and field edit, Follow up of non response, Processing data, Preparing Report)

6. Explain the difference between histogram and bar chart. Give a situation in which one is a better representation than the other.
7. Consider the marks obtained by students in Mathematics, Physics and Chemistry out of 100, 50 and 50 in their board exams in this order for 10 students:

80	45	32	78	43	28	87	42	49	95	45	47	53	32	15
67	23	19	99	50	48	79	45	35	89	39	49	85	36	42

(i) Create a frequency table for ungrouped data for total marks obtained by 10 students.

(ii) Create a frequency distribution for the grouped data with class interval of 10. What is relative frequency?

(iii) Draw a pie chart, histogram and divided bar diagram for the above data to explain some salient features about the data as you feel fit.

(iv) Calculate the mean, median and mode for the above data giving the details. Also compute 20 percent trimmed mean for above data set.

(v) Implement all the above in R.

Refer to the data in Question 7 for the problems 8-14 below:

8. Define at least five different events and find their probabilities.
9. What is the probability that a student scored more than 90 marks in Mathematics and between 70 and 80 (inclusive) in Physics and Chemistry combined together? Hence or otherwise say if the events are independent ?
10. What is the probability that a student scored 45 marks in Chemistry given that he scored at least a total of 150 marks?
11. What is the probability that a student scored between 70 and 80 marks in Mathematics or less than 40 marks in Physics?
12. What is the probability that a student scored at least 80% marks in Physics given that he scored atmost 80% marks in Mathematics and Chemistry combined together?
13. If 5 students are chosen randomly then what is the probability that none of the students scored more than 175 marks out of 200?

14. If it is decided that a student will get a valid score if his total is more than 120 and he will be declared as passed in Mathematics, Physics and Chemistry if he scored more than 70, 30 and 35 marks respectively then
 - (i) draw a Venn diagram to show the number of students passing in the three subjects and
 - (ii) set up a scenario for verifying Baye's theorem and verify it.
15. Consider the following data which gives the percentages of the families that are in the upper income level for some individuals in 15 schools of the city.
72.2, 31.9, 26.5, 29.1, 27.3, 8.6, 22.3, 26.5, 20.4, 12.8, 25.1, 19.2, 24.1, 58.2, 68.1.
Construct a relative frequency histogram of the data.
16. Twelve students compete in a race. In how many ways first three prizes be given?
(Ans: $12 \times 11 \times 10 = 1320$)
17. Suppose $P = \{a \mid a \text{ is an odd prime number } < 7\}$ and $Q = \{b \mid b \in \mathbb{N}, 0 \leq b < 5\}$, where \mathbb{N} is a set of all natural numbers. Find the number of proper subsets of P and Q . (Ans: 3 and 15 respectively.) Hint: Set A with n elements have how many proper subsets??
18. Let 50 patients represent sample units. 20 out of 50 experience stomach ailment after the drug is given. Find sample proportion for which the drug was success and the sample proportion for which drug was not successful. Observe that sample proportion is the sample mean of 1 and 0 where we count 1 if success and 0 stands for failure of drug treatment. ANS: 0.6, 0.4.
19. Suppose $X = \{x \mid x = 3n - 1, n \in \mathbb{N}, n < 3\}$ and $Y = \{y \mid y \text{ is a prime number } < 7\}$. Then find $X \cap Y$. (Ans: $X \cap Y = \{2, 5\}$) Hint: Write X and Y explicitly.
20. Four cards are drawn at random (without replacement) from a well shuffled deck of playing cards. Then find the probability that there is at least one ace among them. (Find answer correct upto four decimal places.) (Ans: 0.2813) Hint: Let us define event A as drawing at least one ace card and use $P(A^c)$.
21. Let E and F be two events with $P(E \cup F) = 0.7$, $P(E) = 0.5$, $P(F) = 0.3$. Find $P(E \cap F^c)$. (Ans: 0.4) Hint: First find $P(E \cap F)$ and use Venn diagram to find desired answer.
22. A fair six-sided dice is rolled twice independently. What is the probability of getting 1 in first roll but not getting 3 or 4 in the second roll? (Ans: $\frac{1}{9}$) Hint: Use concept of independent events
23. Write sample space for the given experiment: Three items are selected at random from a manufacturing process and an item selected is tested for defective (D) or non-defective (N).
24. Two lottery tickets are to be chosen from 20 for first and second prize. Find number of sample points in S . (Ans: $20P_2$)
25. (a) How many ways can five people be lined up to get on a bus? (Ans: $5!$)

(b) If a certain two persons refuse to follow each other, how many ways are possible?

(Ans: $12 \cdot 3! = 72$) $2! \cdot 3! = 12$ || $12 \cdot 3! = 72$

26. A college freshman must take a science course, a social science course and a mathematics course. If he may select any of three sciences, any of four social studies and any of two mathematics courses, how many ways can he arrange his program? (Ans: 24) $4 \cdot 3 \cdot 2$

27. In how many ways can 6 trees can be planted in a circle? (Ans: $5!$) Circle : $(n-1)!$

28. What is the use of harmonic mean in statistics?

29. The average age of 06 persons living in a house is 23.5 years. Three of them are majors and their average age is 42 years. The difference in ages of the three minor children is same. What is the mean of the ages of minor children? (Ans: 5)

30. (Ans: $1 \times 9 \times 9 + 4 \times 1 \times 9 + 4 \times 9 \times 1 = 153$)

31. Two dice are rolled, find the probability that the sum is a) equal to 1; b) equal to 4; c) less than 13. (Ans: a) 0; b) $1/12$; c) 1.)

32. A die is rolled and a coin is tossed, find the probability that the die shows an odd number and the coin shows a head. (Ans: 0.25)

33. A bag I contains 4 white and 6 black balls while another Bag II contains 4 white and 3 black balls. One ball is drawn at random from one of the bags, and it is found to be black. Find the probability that it was drawn from Bag I. (Ans: $\frac{7}{12}$) Hint: Use Bayes rule.

34. If the probability of getting 10 – 20, 21 – 30, 31 – 40, over 40 cars for service at a day is 0.20, 0.35, 0.25, 0.12 respectively. Then find the probability of getting at least 21 cars at that day.

35. A virus has infected 1.8% of a population. A test detects this virus 95% of the time when it is actually present, but it returns a false positive 3% of the time when the virus is not present. If a person at random from this population tests positive for the virus, what is the probability that this person is actually infected? [Round to the nearest percent](Ans: 37 percent) Hint: Use Bayes rule

36. Interpretation of the following data can be done using a bar chart or a Pie chart or both?
1. Ask 20 students their mother tongue. Interpret a bar chart and a pie chart?
2. The pay package given to 50 MBA students are available. Interpret a bar chart and a pie chart?
3. The colour of the shirt worn by 50 students is available. 4. The specializations taken by 40 second year MBA students. 5. The number of students who start their own companies in the last 10 years. 6. Proportion of men and women students in a class 7. Number of different types of defects in manufacturing 8. Number of visits in a website on 5 days in a week 9. Number of journal publications of faculty of a department. 10. Fours and sixes hit by a batsman out of his total career score. 11. Number of customers rating a hotel service as VG, G and poor.

37. Discuss the differences between discrete variables and continuous variables with appropriate examples.
38. How might you convert information captured by continuous variables into categorical variables (grouped variables) for analysis? Explain it with an example.
39. Describe the sub-scales of continuous scale and give an example for each sub-scale.
40. The height (in cm) of 10 people given in ascending order is 164, 165, $\times 3$, 168, $\times 5$, 172, 173, 175, $\times 9$, 176. The unique mode of this data is 175cm, the median is 171cm and the average height is 170.5cm. Calculate the absolute median deviation and the standard deviation for this data.
41. How many arrangements are there of six letters in banana? How many arrangements are there if: (a) The b is followed by an a, (b) The pattern bnn never occurs, (c) The b occurs before any of the a.
42. How many arrangements of 5α s, 5β s and 5γ s are there with at least one β and at least one γ between each successive pair of α s?
43. How many different ways are there to select six hot dogs from three varieties of hot dogs?
44. (a) How many 6 card hands (subsets) can be formed from a standard 52 card deck?
(b) How many 7 card hands that contain exactly two Aces can be formed from a standard 52 card deck? (c) If a 5 card hand is chosen at random, how many ways to choose a 5 card hand is flush (all 5 cards in the hand are in the same suit)?
45. A new board game uses an eight-sided die. Suppose the die is rolled once, what is the probability of getting
(a) an even number ?
(b) a number less than or equal to 5?
(c) an even number greater than 5?
(d) Write event (c) in term of event (a) and (b)
46. The Patil are a family of 5. They are celebrating Diwali and there are 10 gifts, two for each family member. The name tags for each family member have been attached to the gifts. Unfortunately, the name tags on the gifts are damaged by water. Suppose each family member draws two gifts at random. What is the probability that someone
(a) gets his/her two gifts, rather than getting the gifts for another family member?
(b) gets none of his/her gifts, but rather gets gifts for other family members?
28. -> The harmonic mean is typically used whenever different x_i contribute to the mean with a different weight w_i , i.e. when we implicitly assume that the weight of each x_i is not one.