BSCMA1002 : Activity Questions
Week-1

# Contents

# 1 Lecture 1.1

Classification of Statistics, Notion of Sample and Population

A study has been conducted to examine whether the students who had applied for MAEEE entrance exam in 2020 are willing to write the exam in the coming week or not. A group of 10,000 students were randomly selected from the urban areas of India for this study. It is concluded that roughly 80 percent of the students want to write the exam.
Based on the above information, answer the following three questions.

1. Identify the sample and population from the statement given above.

   (a) The sample consists of students from all across India, and population consists of students from the urban areas of India.

   (b) The sample consists of students from the urban areas of India and the population consists of students from all across India.

   (c) The sample consists of students from the urban areas of India and the population consists of selected 10,000 students from the urban areas of India.

   (d) The sample consists of the selected 10,000 students from the urban areas of India and the population consists of students from all across India.

2. Identify the correct statements from the following:

   (a) Aforementioned conclusion made on the basis of the sample is correct.

   (b) Aforementioned conclusion is not correct because the sample is not a good representation of the population.

   (c) The sample is a good representation of the population.

   (d) None of the above.

3. What can you say about the nature of the study?

   (a) The study is descriptive.

   (b) The study is inferential.

4. According to a news report, Facebook is partnering with external researchers from the fields of political science and social media to examine the impact of social networking sites on the society during the 2020 U.S. Presidential elections. Between 200,000 to 400,000 users are expected to opt into the project to see how they behave on Facebook and Instagram.

   Which branch of statistics does the above analysis come under?

   (a) Inferential statistics.

   (b) Descriptive statistics.

5. Choose the correct statements about descriptive statistics from the following:

   (a) Descriptive statistics is concerned with drawing conclusions from the data.

   (b) Descriptive statistics is concerned with describing and summarizing the data.

   (c) Descriptive statistics can be performed either on a sample or a population.

   (d) Descriptive statistics can only be performed on a sample.

   (e) We must always take into account the possibility of chance while studying descriptive statistics.

6. Central Pollution Control Board wants to check the Air Quality Index (AQI) of all the major cities in India after Diwali to decide if air pollution is reaching unhealthy levels. To do so, they have come up with a list of all the major cities and their respective AQI levels and want to determine which cities among the list have an AQI level greater than 100. What kind of statistics does the above analysis come under?

   (a) Descriptive statistics.

   (b) Inferential statistics.

7. Suppose we are interested in studying the height of an average Indian adult above the age of 18. Due to practical constraints, it is not possible to collect the data for the whole country. So we decide to choose a sample of 10,000 people. Which of the following choices for this sample do you think is the most representative of the population?

   (a) A sample consisting of 10,000 people from Karnataka.

   (b) A sample consisting of 10,000 women from South India.

   (c) A sample consisting of 2000 adults from each part of India (i.e., from northern, southern, eastern, western, and middle parts of India).

   (d) A sample consisting of 10,000 people from all over the country who are above 70 kg but below 75 kg.

8. Which one of the following statements best describes the difference between population and sample?

(a) Sample is the entire group, while population is a subset of the sample which is used to describe the parameters of the sample.

(b) Sample and population are synonyms of each other.

(c) Population is the entire group, while sample is a subset of the population which is used to describe the parameters of the population.

(d) Population is all members of an organization while sample is the members of the organization who volunteer to participate in the study.

9. According to the International Cricket Council (ICC), a test cricket ball, when new, should not weigh more than 163 grams. The manufacturers of Duke cricket ball suspected that the average weight of their product is more than 163 grams. To confirm this, 100 Duke cricket balls were selected. Based on this, choose the correct option from below.

   (a) The population is all cricket balls; sample is 100 Duke cricket balls.

   (b) The population is all Duke cricket balls; sample is 100 cricket balls.

   (c) The population is all Duke cricket balls; sample is 100 Duke cricket balls.

   (d) The population is all Duke cricket balls; sample is all Duke cricket balls.

10. A director of an engineering institute wants to evaluate the performance of students of all years and all departments (Computer Science, Mechanical Engineering, Electrical and Electronics Engineering, Electronics and Communication Engineering) in the institute. He randomly selects performance reports of 50 students from the first year Computer Science batch. Based on this, choose the correct option from below.

    (a) The sample chosen is correct as it is collected randomly.

    (b) The sample chosen is wrong as it is not collected randomly.

    (c) The sample chosen is correct as it is a subset of the population.

    (d) The sample chosen is wrong as it is not a representative of the entire population.

11. A construction company wants to start a house construction business in a city. Since there is a large number of houses, the company collects a sample of houses from the city and uses the properties of the sample to study house prices in the city. Which branch of statistics does the above analysis come under?

    (a) Descriptive statistics

    (b) Inferential statistics

12. An online opinion poll was conducted by a film reviewing website to rank Indian films from the 2000s. Among the 1000 films in the list, Bhaag Milkha Bhaag received the maximum number of votes followed by 3 Idiots and Swades respectively. Which branch of statistics does the above analysis come under?

(a) Descriptive statistics

(b) Inferential statistics

# 2 Lecture 1.2

Introduction to data

1. Which of the following statement(s) is(are) true about structured data in a spreadsheet?

   (a) Each observation must have its own row.

   (b) Each variable must have its own column.

   (c) For each variable the same type of value for each case is recorded.

   (d) For each case the same type of value for each variable is recorded.

2. What kind of data is lyrics of a song?

   (a) Unstructured data

   (b) Structured data

3. Data is collected from people of every family living in a locality. The collected information is first name, second name, income, height, weight, blood group, house address, locality. Which among the following collected data is(are) not a variable(s)?

   (a) First name

   (b) Second name

   (c) Income

   (d) Height

   (e) Weight

   (f) Blood group

   (g) House address

   (h) Locality

4. The values of temperature and relative humidity of a room are measured for 24 hours, starting at 9 AM, at time interval of 30 minutes [Table 1.2.1]. What are the case(s) and variable(s) for this data?

   (a) Temperature is the case; relative humidity and time are the variables.

   (b) Temperature is the case; relative humidity is the variable.

   (c) Time is the case; temperature and relative humidity are the variables.

   (d) Relative humidity is the case; temperature is the variable.

   (e) 9:00 AM, 9:30 AM, 10:00 AM are cases; time, temperature, and relative humidity are variables.

| Time | Temperature ($^0C$) | Relative humidity (%) |
|---|---|---|
| 9:00 AM | 27 | 30% |
| 9:30 AM | 29 | 30% |
| 10:00 AM | 31 | 32% |
| 10:30 AM | 31 | 35% |
| 11:00 AM | 35 | 35% |
| 11:30 AM | 36 | 46% |
| 12:00 PM | 38 | 47% |
| 12:30 PM | 39 | 50% |
| 1:00 PM | 38 | 51% |
| 1:30 PM | 38 | 53% |
| 2:00 PM | 35 | 54% |
| 2:30 PM | 31 | 55% |
| 3:00 PM | 31 | 57% |
| 3:30 PM | 29 | 58% |
| 4:00 PM | 28 | 60% |

Table 1.2.1 Temperature and relative humidity variation in a day

5. Which of the following is(are) example(s) of unstructured data?

   (a) Video files

   (b) Image files

   (c) Phone number

   (d) Email address

6. Which of the following statement(s) is(are) true?

   (a) All the structured data have units.

   (b) Comments in a social media platform are structured data.

   (c) Unstructured data cannot be used directly to perform statistical analysis.

   (d) Mathematical operations can be performed on unstructured data.

In a survey, art students are asked to mention their favourite artist and the amount of time they spend in painting outside the classroom. The data collected are given Table 1.2.2.

| S.no. | Name | Gender | Favourite artist | Amount of time(hr) |
|---|---|---|---|---|
| 1. | David | Male | Amrita Shergil | 4 |
| 2. | Lana | Female | M.F. Hussain | 6 |
| 3. | Riya | Female | Frieda Kahlo | 1 |
| 4. | Aakash | Male | ... | ... |
| 5. | Kabir | Male | Leonardo Da Vinci | 2 |
| 6. | Shakshi | Female | Pablo Picasso | 0 |
| 7. | Mohit | Male | Angolie Ela Menon | 3 |
| 8. | Nisha | Female | Judith Leyster | 3 |

Table 1.2.2: Students record

Based on the data in Table 1.2.2, answer the following three questions:

7. Is the data given in Table 1.2.2 structured or unstructured?

   (a) The data is structured.

   (b) The data is not structured.

8. Identify the variable(s) and observation(s) from the dataset given in Table 1.2.2.

   (a) Name is an observation.

   (b) Name, Gender, Amount of time, and Favourite artist are variables.

   (c) Student and Name are variables.

   (d) Each student is an observation.

9. Choose the correct statement(s) from the following:

   Statement 1: Amount of time Aakash spends on painting is 0 hr.
   Statement 2: We cannot make any conclusion about Aakash because data is not available to us.

   (a) Both the statements are correct.

   (b) Statement 1 is correct but statement 2 is not correct.

   (c) Statement 2 is correct but statement 1 is not correct.

   (d) None of the statements are correct.

# 3 Lecture 1.3

Classification of variables

1. In order to forecast tomorrow's temperature in Mumbai, the India Meteorological Department collects the temperature, humidity and rainfall data of Mumbai for all days from the past three months. What kind of data is this?
   (a) Time series data
   (b) Cross-sectional data
2. What kind of variables are comparative forms of adverbs (like more, less; good, better and best etc.)?
   (a) Categorical variables
   (b) Numerical variables
3. What kind of variable is latitude?
   (a) Numerical
   (b) Categorical
   (c) Continuous
   (d) Discrete
4. During the Great Depression, several countries across the world experienced a rise in the rate of unemployment. The unemployment rate between the years 1929 and 1942 in the US was recorded on a monthly basis. Determine whether the data is time series or cross-sectional.
   (a) The data is time series.
   (b) The data is cross-sectional.
5. Number of days a group of students take to finish a school project is a:
   (a) Quantitative and Discrete variable
   (b) Quantitative and Continuous variable
   (c) Qualitative variable
6. What kind of variable is the qualification of a candidate sitting for a job interview?
   (a) Categorical/ Qualitative
   (b) Numerical and discrete
   (c) Numerical and continuous

7. Suppose one is interested in calculating the weight of a bag (in kg) before leaving for the airport. What kind of variable is the weight of the bag?

   (a) Quantitative and discrete

   (b) Quantitative and continuous

   (c) Qualitative/categorical

# 4 Lecture 1.4

Scales of measurement for variables

1. A courier service is interested in knowing customer feedback. The customers are being given the following options to choose from.

   - Not satisfied
   - Somewhat unsatisfied
   - Neutral
   - Satisfied
   - Very Satisfied

   What scale of measurement has been used here?

   (a) Ordinal
   (b) Nominal
   (c) Interval
   (d) Ratio

2. For which of the following scale(s) of measurement is rank or order meaningful?

   (a) Interval scale
   (b) Ratio
   (c) Nominal scale
   (d) Ordinal scale

3. What can you say about the scale(s) of measurement being used for describing hair color such as blonde, brunette, red, and black?

   (a) The scale of measurement is ordinal.
   (b) The scale of measurement is nominal.

4. What is the scale of measurement for phone number?

   (a) Ordinal
   (b) Nominal
   (c) Interval
   (d) Ratio

5. Which of the following is true for a variable with the ratio scale of measurement?

   (a) Addition operation can be performed.

(b) Subtraction operation can be performed.

(c) Multiplication operation can be performed.

(d) Division operation can be performed.

(e) The variable with ratio scale of measurement can have zero as one of its values.

6. If addition, subtraction can be performed on a variable, then the scale(s) of measurement of the variable could be:

   (a) Ratio

   (b) Ordinal

   (c) Interval

   (d) Nominal

# BSCMA1002: Activity Questions
Week-2

**Contents**

## 1 Lecture-2.1

Describing categorical variable using frequency tables

1. The winners of the International Cricket Council(ICC) One Day International(ODI) world cups are as follows:
   West Indies, West Indies, India, Australia, Pakistan, Sri Lanka, Australia, Australia, Australia, India, Australia, England.
   Answer the questions (i) and (ii) based on the given information.

   (i) How many world cups have been played in total?
   - (a) 10
   - (b) 11
   - (c) 12
   - (d) 13

   (ii) Which team won the most number of world cups and how many times?
   - (a) Australia, 4
   - (b) India, 3
   - (c) Australia, 5
   - (d) India, 2

2. Table 2.1.1 shows the causes of deaths in City hospital in 2019. Answer the questions (i), (ii), (iii) and (iv) based on the given data.

| Cause | No. of deaths | Relative frequency |
|---|---|---|
| Cancer | 150 | $t$ |
| Accidents | $x$ | 0.175 |
| Heart disease | $y$ | $z$ |
| Respiratory disease | 250 | 0.25 |
| Others | 300 | 0.3 |

Table 2.1.1: Causes of deaths

(i) How many deaths happened in City hospital in 2019?

(a) 800
(b) 1000
(c) 2000
(d) Data insufficient

(ii) What is the value of $t$?

(a) 0.25
(b) 0.1
(c) 0.2
(d) 0.15

(iii) What is the value of $z$?

(a) 0.125
(b) 0.15
(c) 0.2
(d) Data insufficient

(iv) What is the value of $x + y$?

(a) 175
(b) 200
(c) 250
(d) 300

3. A cab driver got the rating as feedback from 20 customers on a given day which are as follows:

7 9 8 $x$ 6 9 7 5 9 10 7 8 9 8 8 7 10 6 6 9

(i) If the relative frequency of rating 7 is 0.25, then what is the value of $x$?

(a) 7
(b) 8
(c) 9
(d) Data insufficient

(ii) What is the relative frequency of rating 8?

(a) 0.2
(b) 0.25
(c) 0.1
(d) 0.15

4. Table 2.1.2 lists the number of patients visiting a dentist in a week.

| Day | Number of patients checked | Relative frequency |
|---|---|---|
| Sunday | 15 | 0.125 |
| Monday | 20 | |
| Tuesday | 15 | |
| Wednesday | $x$ | $y$ |
| Thursday | 19 | |
| Friday | 20 | |
| Saturday | 11 | |

Table 2.1.2: Number of patients

(i) How many patients visited the dentist in total in that week?

. . . . . .
(Ans:120)

(ii) What is the value of $x$?

. . . . . .
(Ans:20)

(iii) What is the value of $y$?

. . . . . .
(Ans:0.167)

5. In the Summer Olympics of 2016, many countries competed against each other in various sports categories. In each category, gold, silver, and bronze medals were awarded to the top three performing countries respectively. Table 2.1.3 lists the number of medals won by some countries. Based on this dataset, answer the questions (i), (ii), and (iii).

| Country Name | Number of medals |
|---|---|
| Argentina | 4 |
| Australia | 29 |
| China | 70 |
| France | 42 |
| Germany | 42 |
| Great Britain | 67 |
| Italy | 28 |
| Japan | 41 |
| Russia | 56 |
| United States | 121 |
| **Total number of medals** | **500** |

Table 2.1.3: Number of medals won by some countries

(i) Which country won the most olympic medals?

(a) United States
(b) Italy
(c) Great Britain
(d) Argentina

(ii) What is the relative frequency of United States?

---

Answer: 0.242

(iii) What is the relative frequency of European countries (France, Germany, Italy, Russia, and Great Britain)?

---

Answer: 0.47

## 2 Lecture-2.2

Graphical representation of categorical variables

1. Figure 2.1.1 gives data on the number of students playing different sports in Saint Francis School(SFS). Assuming no player plays more than one game, answer the questions (i), (ii) and (iii) based on Figure 2.2.1.

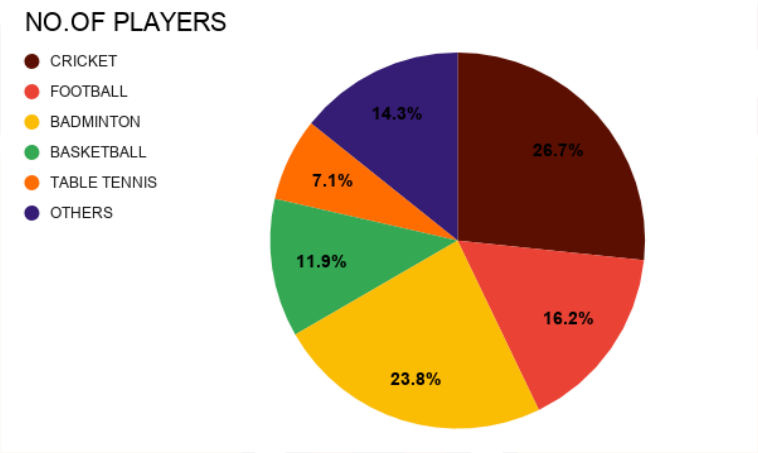



Figure 2.2.1: Number of players playing different sports in SFS

(i) If a total of 210 students play sports in SFS, then how many students play basketball?

(a) 30
(b) 12
(c) 23
(d) 25

(ii) Which sports is played by the most number of students?

(a) Badminton
(b) Football
(c) Cricket
(d) Basketball

(iii) How many students play cricket?

(a) 56
(b) 57
(c) 26
(d) 27

2. Figure 2.2.2 shows the literacy rates according to census of different years from 1951 to 2011.

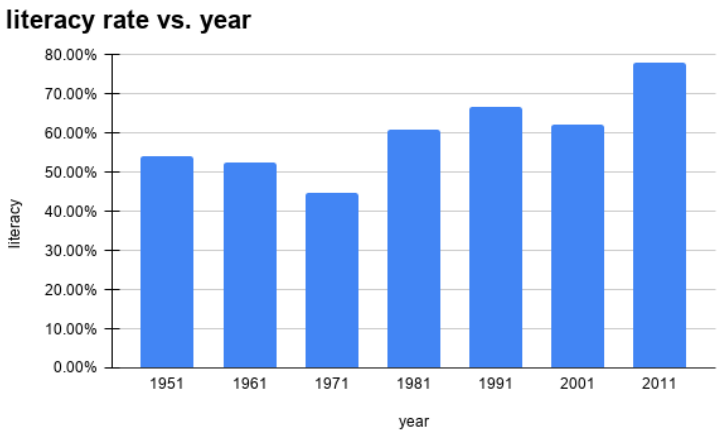



Figure 2.2.2: Literacy rate of different years from 1951 to 2011

(i) Which years have the highest and lowest literacy rates, respectively?

(a) 2011, 1991
(b) 2011, 1951
(c) 1991, 1951
(d) 2011, 1971

(ii) In which decade did the literacy rate decrease the most?

(a) 1971-1981
(b) 1961-1971
(c) 1951-1961
(d) 2001-2011

3. Which is the appropriate Pareto chart for the Figure 2.2.3?

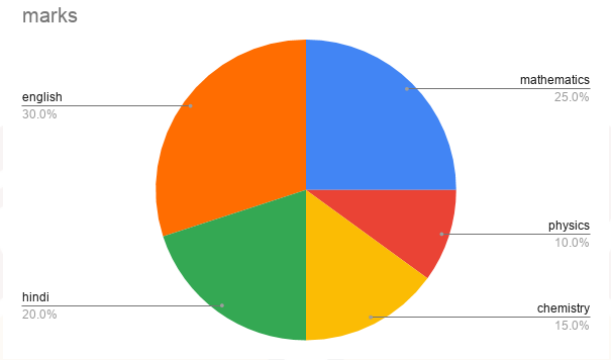
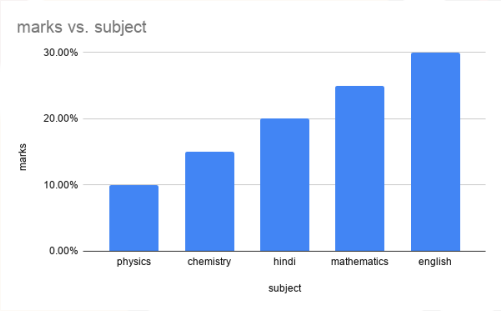
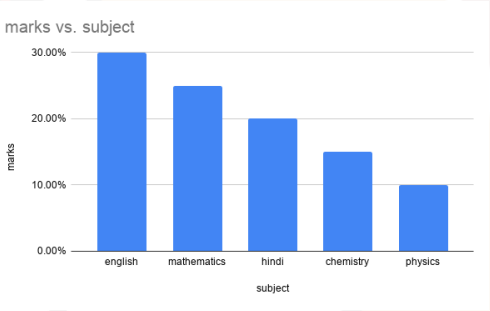



Figure 2.2.3: Marks of students of 12th class




Graph 1




Graph 2

(a) Graph 1

(b) Graph 2

(c) Both

(d) None of the above

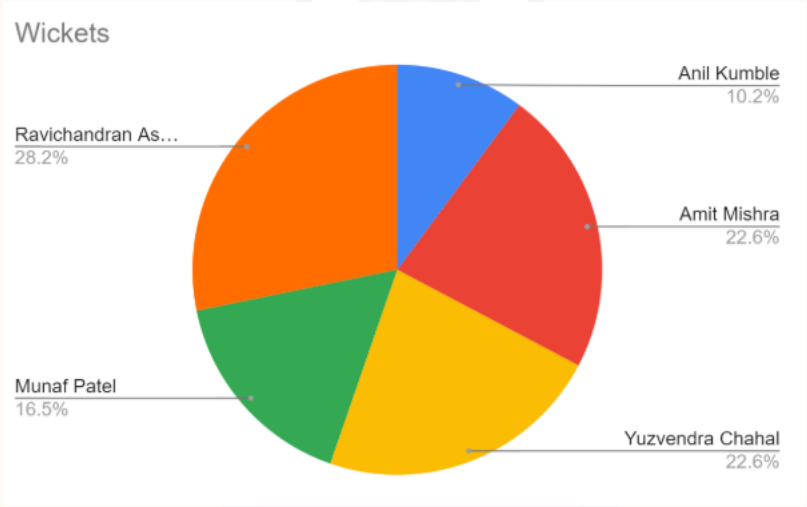
4. Table 2.2.1 shows the wickets taken by Indian bowlers in a certain number of matches in the Indian Premier League(IPL). Based on this data, answer the questions (i), (ii), and (iii).

| Bowler | Wickets |
|---|---|
| Anil Kumble | 45 |
| Amit Mishra | 157 |
| Yuzvendra Chahal | 100 |
| Munaf Patel | 73 |
| Ravichandran Ashwin | 125 |
| **Total** | 500 |

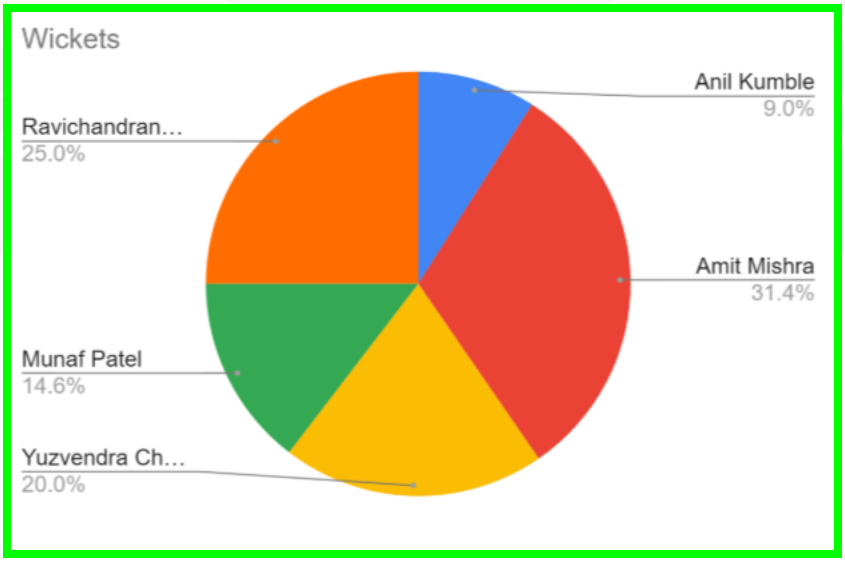
Table 2.2.1: Wickets taken by bowlers

(i) What is the pie chart representation of the data given in Table 2.2.1?

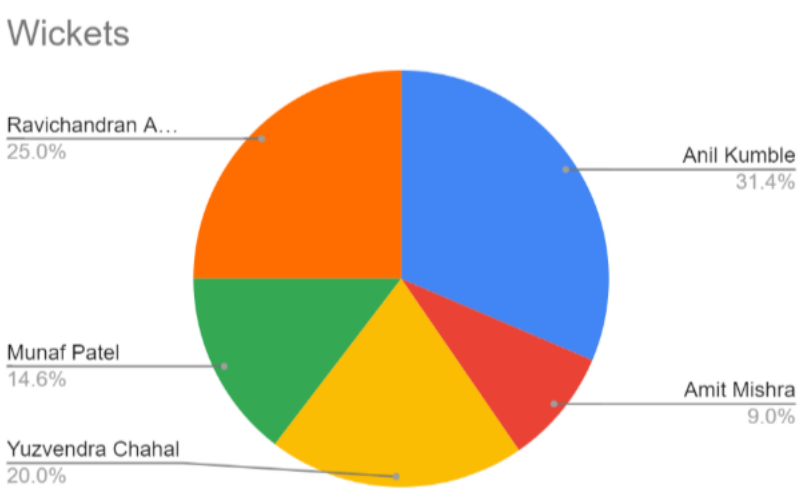
(a)



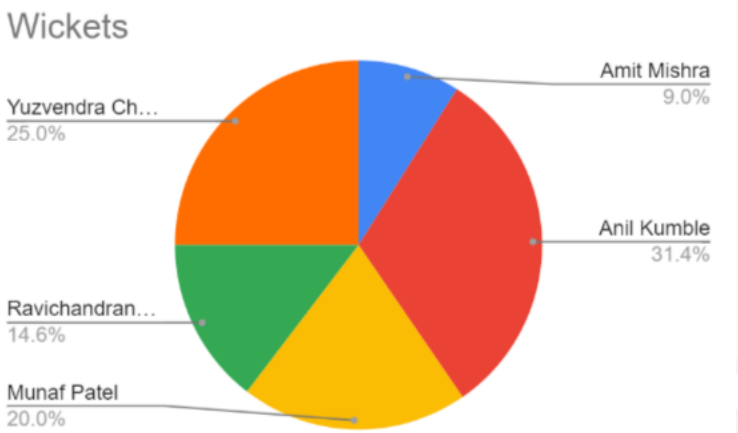

(b)



(c)

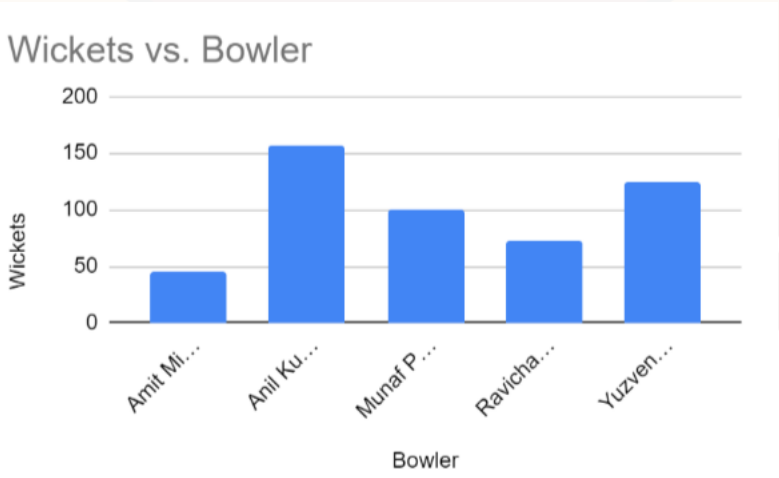



(d)

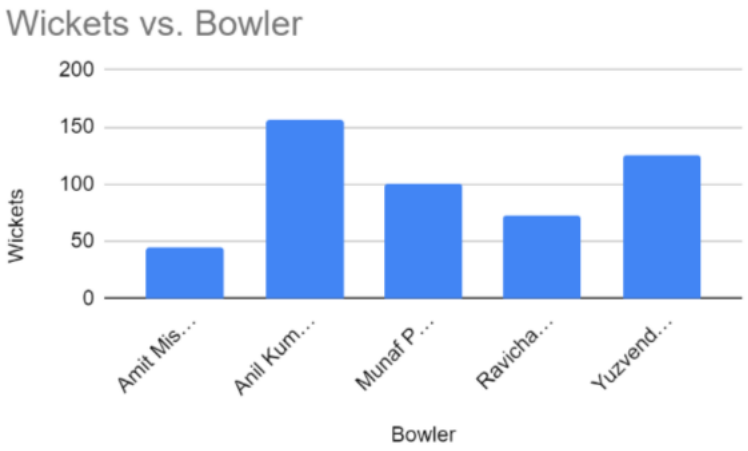



(ii) What is the bar chart representation of the data given in Table 2.2.1?

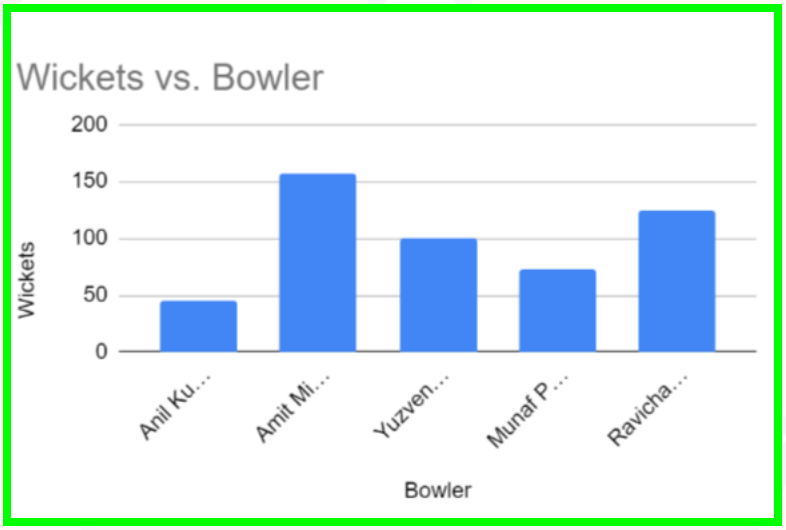
(a)



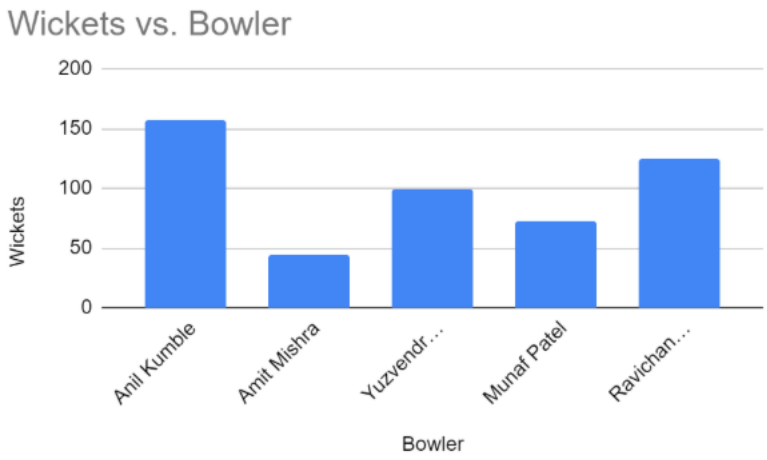
(b)


Wickets vs. Bowler
200
150
100
50
0
Wickets
Amit Mis...
Anil Kum...
Munaf P...
Ravicha...
Yuzvend...
Bowler

(c)

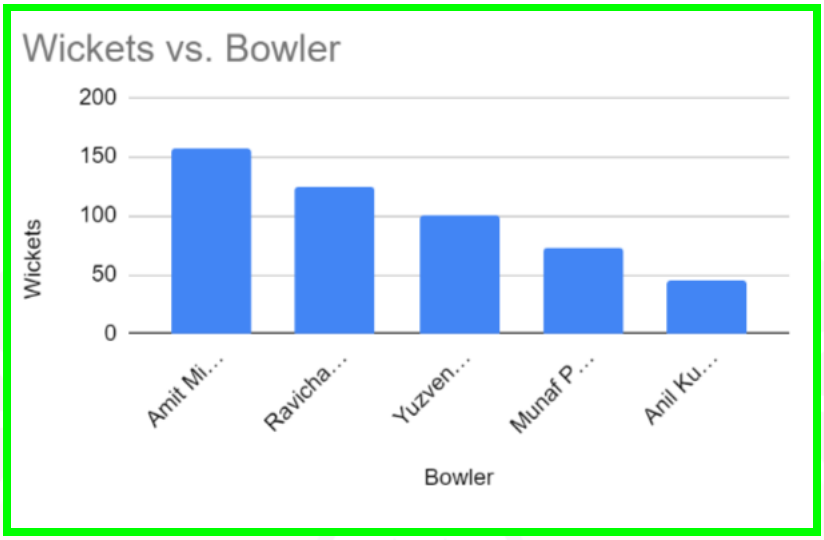

Wickets vs. Bowler
200
150
100
50
0
Wickets
Anil Ku...
Amit Mi...
Yuzven...
Munaf P...
Ravicha...
Bowler

(d)

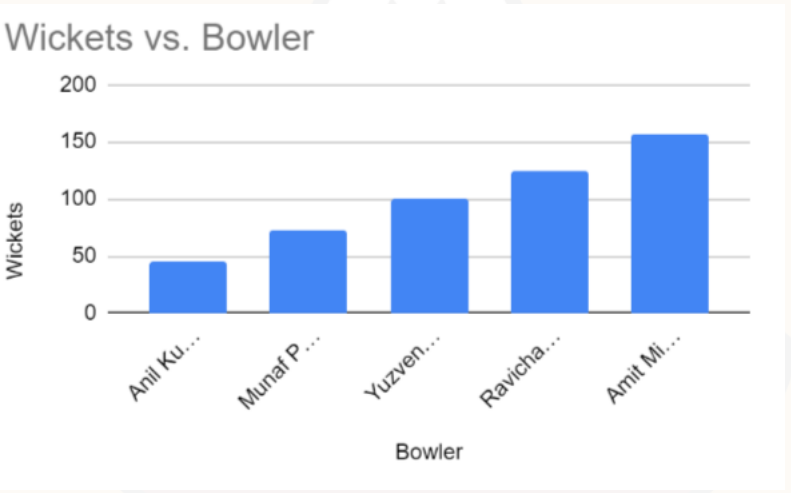

Wickets vs. Bowler
200
150
100
50
0
Wickets
Anil Kumble
Amit Mishra
Yuzvendr...
Munaf Patel
Ravichan...
Bowler

(iii) What is the Pareto chart representation of the data given in Table 2.2.1?

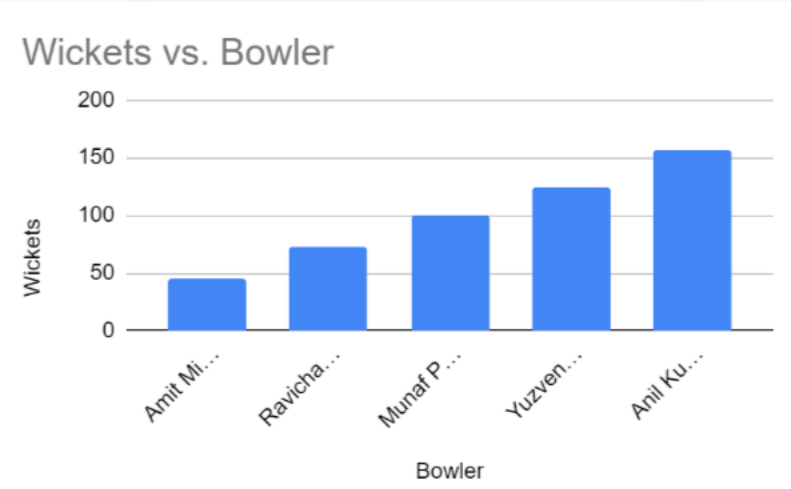
(a)



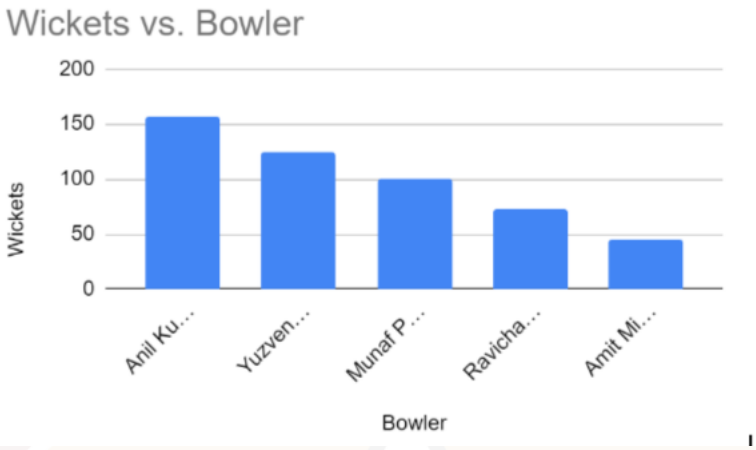

(b)




(c)



(d)




5. Which of the following is/are true about a Pareto chart?

   (a) Pareto chart is same as bar chart if the frequency data is sorted in ascending order.

   (b) Pareto chart can be used to represent a categorical data graphically with an ordinal scale of measurement.

   (c) Pareto chart is same as bar chart if the frequency data is sorted in descending order.

   (d) Pareto chart should not be used to represent a categorical data graphically with an ordinal scale of measurement.

## 3 Lecture-2.3

Best practices while graphing data-1

1. World Health Organisation (WHO) conducted a study to examine how many human lives have been lost during Covid-19 pandemic. It gathered the data across all the countries and represented it graphically. WHO also wants to compare it with another data which shows the percentage of total population of each country affected by it. Which chart will be more suitable for this dataset?

   (a) Bar chart

   (b) Pie chart

   (c) Pareto chart

   (d) None of the above

2. Which of the following statement(s) is(are) true?

   (a) Bar chart is best suited for distribution of construction costs of a house.

   (b) Grouping of categories in frequency table is done when we have categories with same frequencies.

   (c) Pie chart is best suited when we want to emphasize the proportion rather than the frequency of the categories.

   (d) Pie chart and Bar chart are always interchangeable.

3. In which of the following cases is a bar chart better than a pie chart to represent categorical data?

   (a) An e-commerce company wants to know which source generates better revenue.

   (b) An e-commerce company is interested in knowing average transaction by different payment modes.

   (c) A political party wants to assess how the share of votes received by various candidates have changed over the years.

   (d) None of the above

4. In which of the following cases is a Pareto chart a suitable tool to represent a categorical variable graphically?

   (a) To understand the order of the frequency of a categorical variable and compare it with other groups.

   (b) To understand the share or proportion of each group of a categorical variable.

   (c) To compare different groups and not worry about the order of frequency.

   (d) When we have a large number of categories.

5. Which of the following statement(s) is(are) true?

   (a) All the bars in a bar chart can have the same colour.

   (b) Each bar in a bar chart should have a different colour.

   (c) All slices in a pie chart must have the same colour.

   (d) Each slice in a pie chart must have a distinct colour.

## 4 Lecture-2.4

Best practices while graphing data-2

1. The bar chart given in Figure 2.4.1 is misleading because:

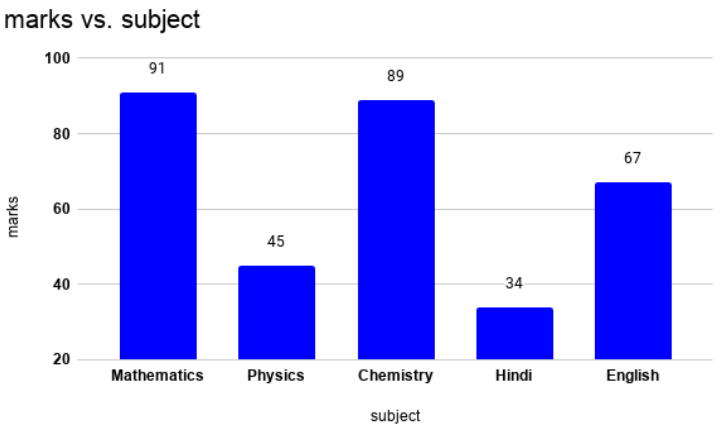



Figure 2.4.1: Marks of students

   (a) Subjects are not arranged in alphabetical order.
   (b) Bars are not arranged in increasing or decreasing order.
   (c) There are no grid lines at the top of the bars.
   (d) The baseline does not start at zero.

2. Road safety organisation in Delhi conducted a study and gathered the data of all fatal car accidents that happened on Yamuna Expressway in last five months. The data shows that 60 times out of 250(24%), car was being driven by women. This data is represented using the pie chart in Figure 2.4.2. Choose the correct statement based on given data.

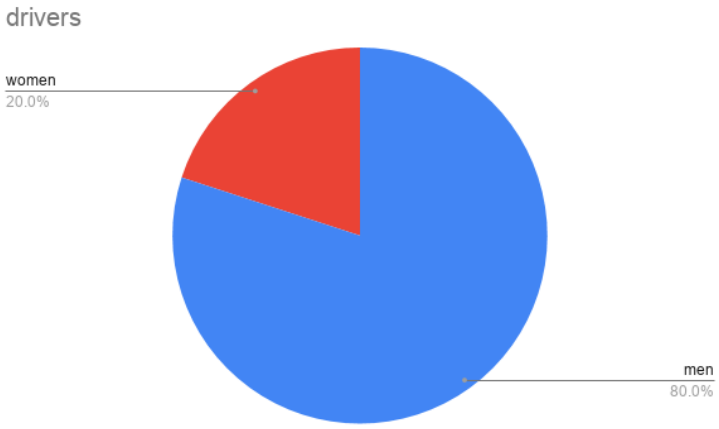



Figure 2.4.2: Drivers during car accidents

(a) Pie chart is misleading because data is rounded off.
(b) Pie chart is misleading as it disobeys area principle.
(c) Pie chart is misleading as it is labelled wrong.
(d) Pie chart is not misleading.

3. Figure 2.4.3 shows the pie chart representation of the results of an online poll conducted by a sports channel to choose the better captain among Virat Kohli and MS Dhoni. Based on this information, choose the correct option(s) from below.

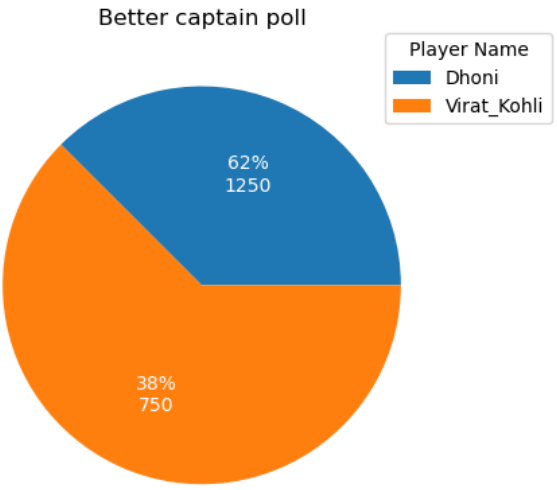



Figure 2.4.3: Captain poll dataset

(a) The pie chart is misleading because the angles of slices in the pie chart are not proportional to the values they represent.

(b) The pie chart is the correct representation because it obeys the area principle.

(c) The pie chart is misleading because the angles of slices in the pie chart are proportional to the values they represent.

(d) The pie chart is misleading because it does not obey the area principle.

4. Figure 2.4.4 shows the bar chart representation of the results of an online poll conducted by IMDb in which a participant votes for their favourite film. Based on this bar chart, choose the correct option(s) from below.

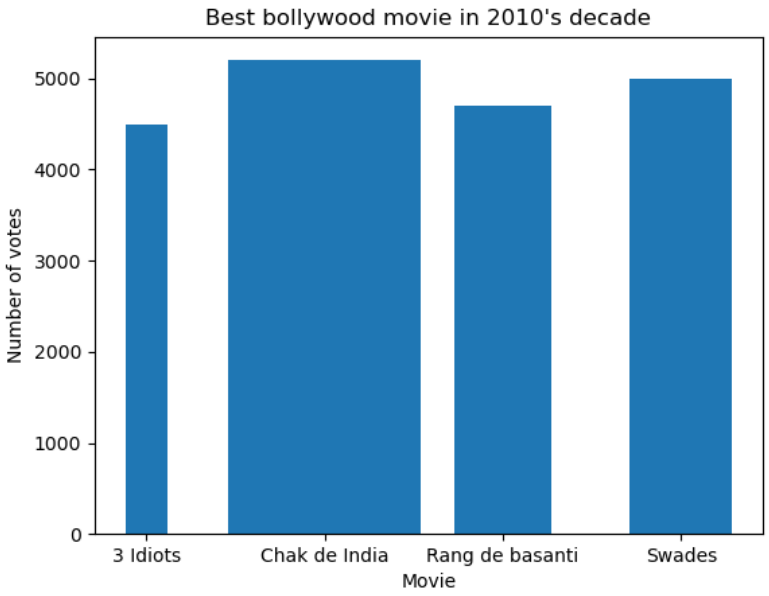



Figure 2.4.4: IMDb movie poll dataset

(a) The bar chart is the correct representation because it obeys the area principle.

(b) The bar chart is not misleading because the width of the bars is the same.

(c) The bar chart is misleading because the width of the bars are not the same.

(d) The bar chart is misleading because it does not obey the area principle.

5. Why is the graph plotted in Figure 2.4.5 representing the share of the votes in an election in the United States of America misleading?

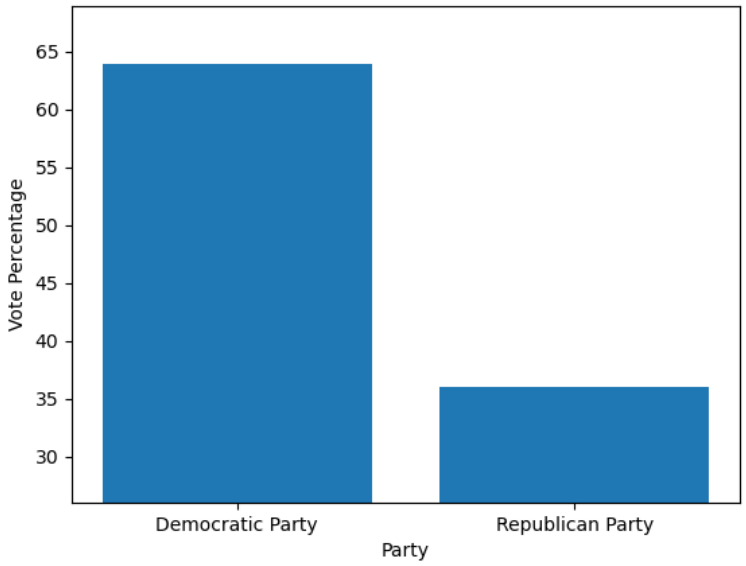



Figure 2.4.5: Share of the votes in an election in the United States of America

(a) Missing baseline.

(b) Manipulated Y-axis.

(c) It is not a misleading graph.

(d) Both missing baseline and manipulated Y-axis.

6. Why is the graph plotted in Figure 2.4.6 misleading?

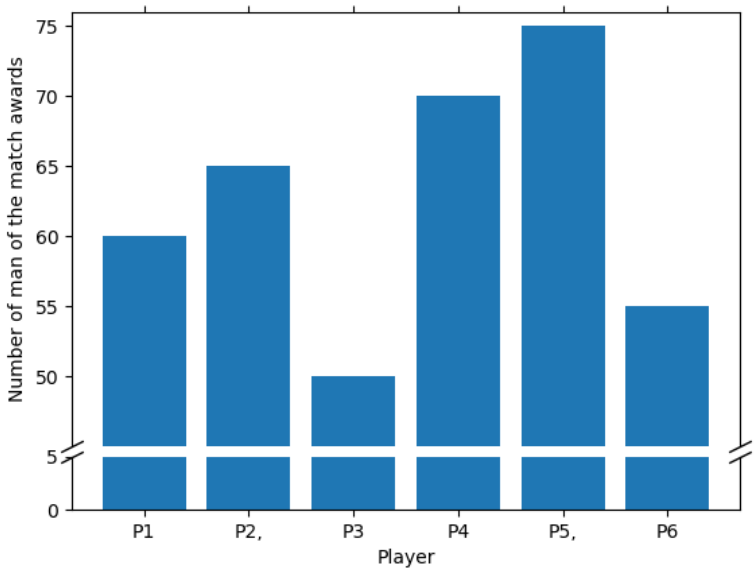



Figure 2.4.6: Numbers of Man of the Match of players

(a) Missing baseline.

(b) Manipulated Y-axis.

(c) It is not a misleading graph.

(d) Both missing baseline and manipulated Y-axis.

7. The share of the Indian telecom market is plotted as the pie chart shown in Figure 2.4.7. Based on this information, choose the correct option(s) from below.

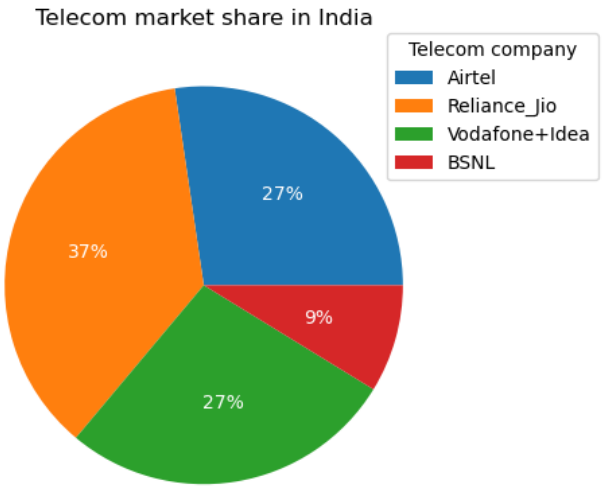



Figure 2.4.7: Telecom market share dataset

(a) The pie chart is not a misleading graph.

(b) The pie chart is misleading.

(c) The pie chart has round-off errors.

(d) The slices of the pie chart do not add up to 100%.

## 5 Lecture-2.5

Descriptive measures of categorical variables

1. Figure 2.5.1 shows the allocation of budget amount under various headings in the 2020 budget presented by the Rajasthan government. Choose the correct options based on the given data:

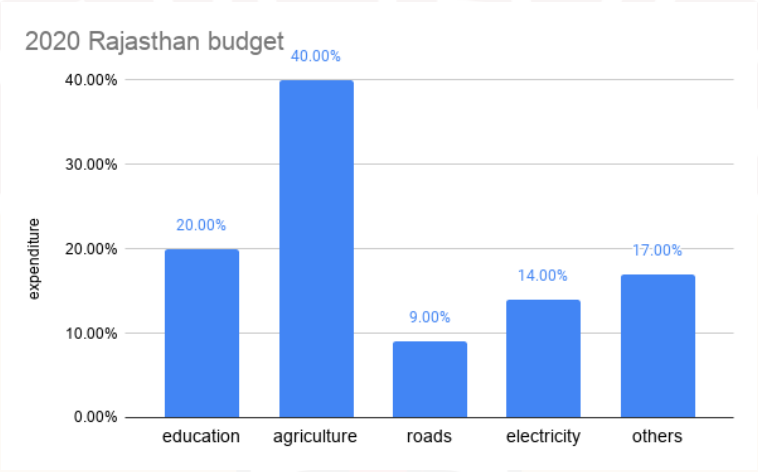



Figure 2.5.1: Rajasthan Budget 2020

(a) Mode of above data is agriculture.

(b) Mode is not defined for categorical variable.

(c) Median of above data is roads.

(d) Median is not defined for this variable.

2. Choose the correct statements:

   (a) For ordinal variables, mode and median can be same.

   (b) For categorical data, mode can not be more than one.

   (c) Mode is always greater than median if defined.

   (d) In Pareto graph, first bar is always mode variable.

3. Consider Figure 2.5.2 and choose the correct options.

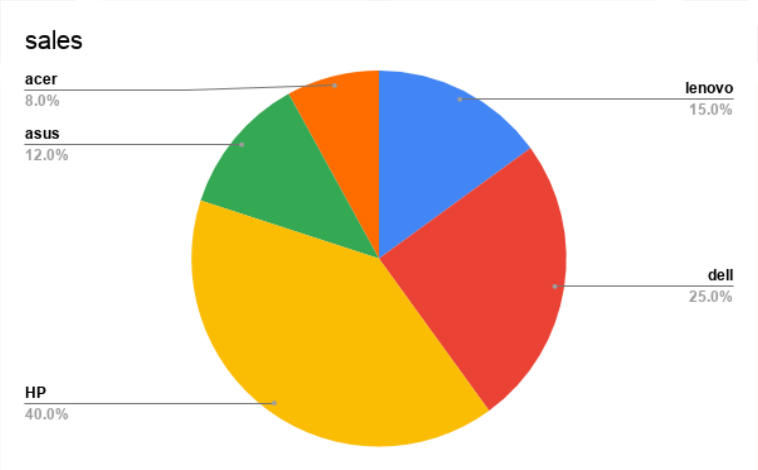



Figure 2.5.2: Sales

   (a) Mode of above data is HP.

   (b) Mode is not defined for categorical data.

   (c) Median of above data is lenovo as its sales percentage is middle value of all five sales.

   (d) Median is not defined for this variable.

4. English teacher has declared the result of class 12 students. Grades of the students are listed in Table 2.5.1 where grades are in the order A> B> C> D> E. Based on this information choose the correct options:

| Grade | No. of students |
|---|---|
| A | 15 |
| B | 20 |
| C | 35 |
| D | 15 |
| E | 6 |

Table 2.5.1: Grades of students

   (a) Mode of above grades is C.

   (b) Mode of above grades is A.

   (c) Median of above grades is C.

   (d) Median is not defined for this data.

5. Which of the following measure(s) of central tendency is(are) applicable for a categorical variable with an ordinal scale of measurement?

   (a) Mean

   (b) Median

   (c) Mode

   (d) None of the above

6. Statement: Mean is a measure of central tendency for a categorical variable.
   Which of the following option(s) is(are) correct about this statement?

   (a) True. A categorical variable can have numerical values and hence the mean of the variable exists.

   (b) False. A categorical variable cannot have numerical values and hence the mean of the variable does not exist.

   (c) True. The order of the data exists in a categorical variable with an ordinal scale of measurement. Thus the mean of the variable exists.

   (d) False. Even though the order of the data might exist in a categorical variable with an ordinal scale of measurement, the mean of the variable will not exist.

7. In a restaurant, a waiter collects feedback from customers asking them to rate the food as good, average, or bad. The data collected by the waiter is given in Table 2.5.2. Based on this data, answer the questions (i) and (ii).

| Customer | Rating of the food |
|---|---|
| Customer 1 | Good |
| Customer 2 | Bad |
| Customer 3 | Bad |
| Customer 4 | Average |
| Customer 5 | Average |
| Customer 6 | Good |
| Customer 7 | Good |
| Customer 8 | Good |
| Customer 9 | Bad |
| Customer 10 | Average |

Table 2.5.2: customer's feedback

(i) Which is the most frequent feedback from the customer?

(a) Good
(b) Bad
(c) Average

(ii) What is the median of the customer's feedback?

(a) Good
(b) Bad
(c) Average
(d) Median does not exist since the feedback variable has nominal scale of measurement.
(e) Median does not exist since the feedback variable has ordinal scale of measurement.

BSCMA1002: Activity Questions
Week-3

# Contents

# 1 Lecture-1

Describing numerical data - Frequency tables for numerical data

1. Table 3.1.1 shows the frequency of outcomes of rolling a dice. Use the information in Table 3.1.1 to answer the question that follows:

| Face Number | Frequency | Relative Frequency |
|---|---|---|
| 1 | 5 | 0.125 |
| 2 | 4 | 0.1 |
| 3 | $x$ | 0.15 |
| 4 | 9 | 0.225 |
| 5 | 8 | 0.2 |
| 6 | 8 | 0.2 |

Table 3.1.1

What is the value of $x$?
Answer: 6

2. In a shopping mall, a customer will get a scratch coupon after shopping for over a certain amount. The customer will get cashback between ₹1 to ₹100. Madhu is a regular customer to the shopping mall. The cashback she received over a period of six months is shown in Table 3.1.2.

| Class | Frequency |
|---|---|
| 1-20 | 5 |
| 20-40 | 10 |
| 40-60 | 6 |
| 60-80 | 8 |
| 80-100 | 6 |

Table 3.1.2

How much percentage of total time, did she receive the cashback worth of ₹40 to ₹60?
Accepted Range: [17 - 18]

3. The number of wickets taken by bowler Sushant in his school cricket matches are given below.
1 2 5 6 4 3 4 2 1 6 4 3 1 3 2
What is the relative frequency corresponding to 3 wickets?
Answer: 0.2

4. The ratings given by the users on the newly launched Samsung mobile phone is given in Figure 3.1.1. How many users gave a rating greater than or equal to 3?

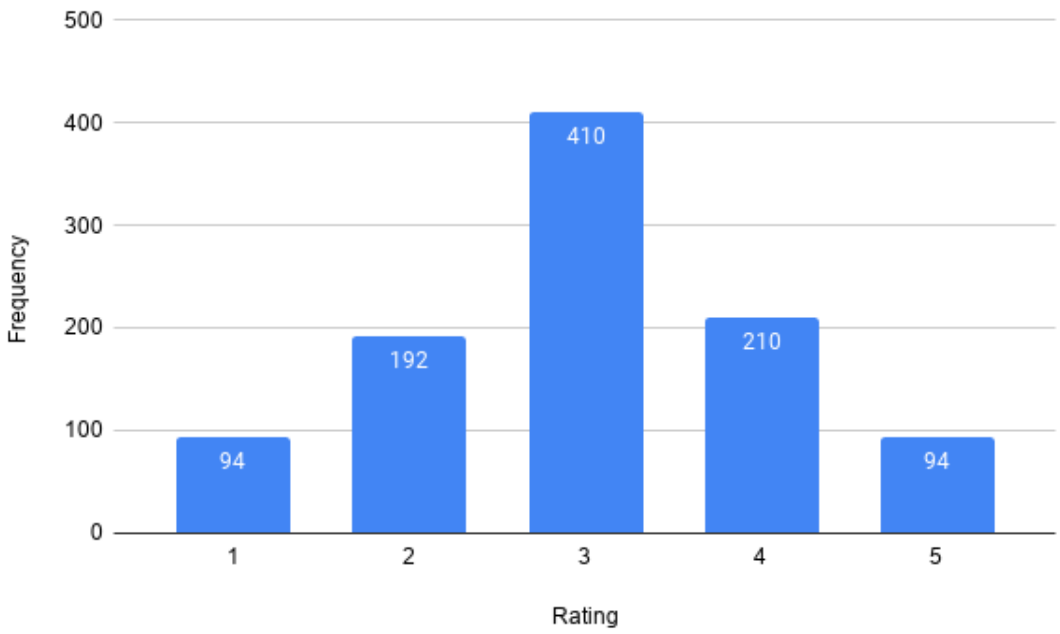



Figure 3.1.1

Answer: 714

5. The marks distribution of students in a coaching center for entrance exams is given in Figure 3.1.2. Based on this information, answer questions (i) and (ii).

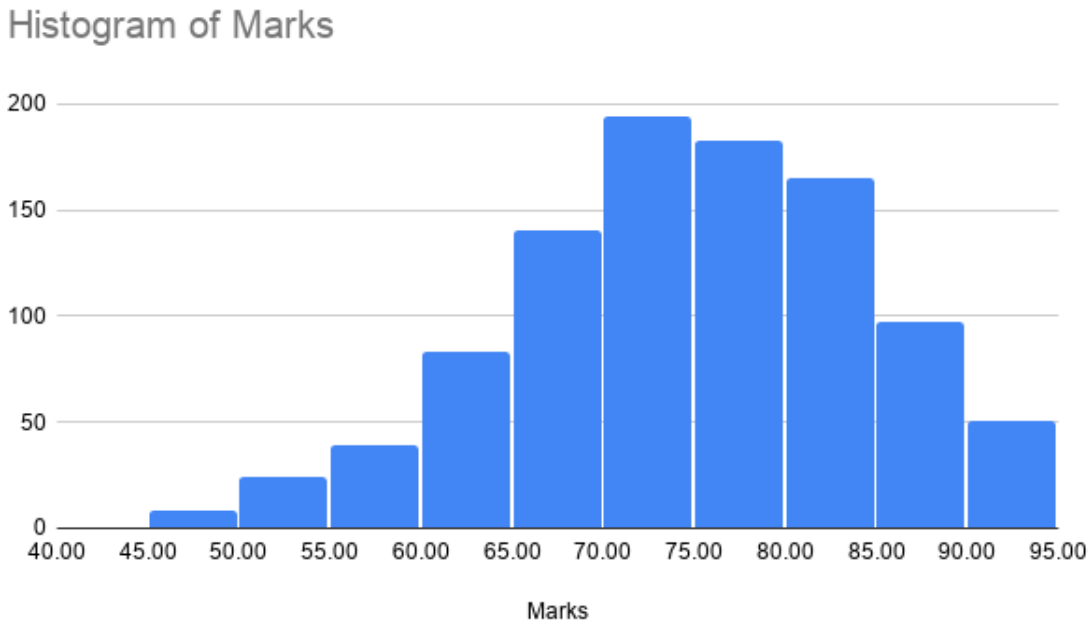



Figure 3.1.2

i What is the class mark of the class interval [55,60)?
Answer: 57.5

ii To which class does a student who got 65 marks belong?

(a) 55-60
(b) 60-65
(c) 65-70
(d) 50-55

# 2 Lecture-2

Describing numerical data - Mean

1. The organiser of a wrestling competition grouped the participating wrestlers into different categories based on their body weights. Table 3.2.1 shows these categories.

| Weight groups (in kg) | Number of wrestlers |
|---|---|
| 60-70 | 7 |
| 70-80 | 5 |
| 80-90 | 4 |
| 90-100 | 4 |

Table 3.2.1

What is the approximate mean of the weights (in kg) of the wrestlers?

(a) 80

(b) 77.5

(c) 87.5

(d) 89

2. The mean of the following data is 26.
19.3 20.1 33.4 23.5 $x$ 29.8

What is the value of $x$?
Answer: 29.9

3. The number of wickets taken by bowler Sushant in his school cricket matches are given below.
1 2 5 6 4 3 4 2 1 6 4 3 1 3 2
What is the average number of wickets taken by Sushant (approximate to nearest integer)?
Answer: 3

4. The ratings given by the audience on a movie review website for a newly released movie is as shown in Figure 3.2.1. What is the average rating given by audience?

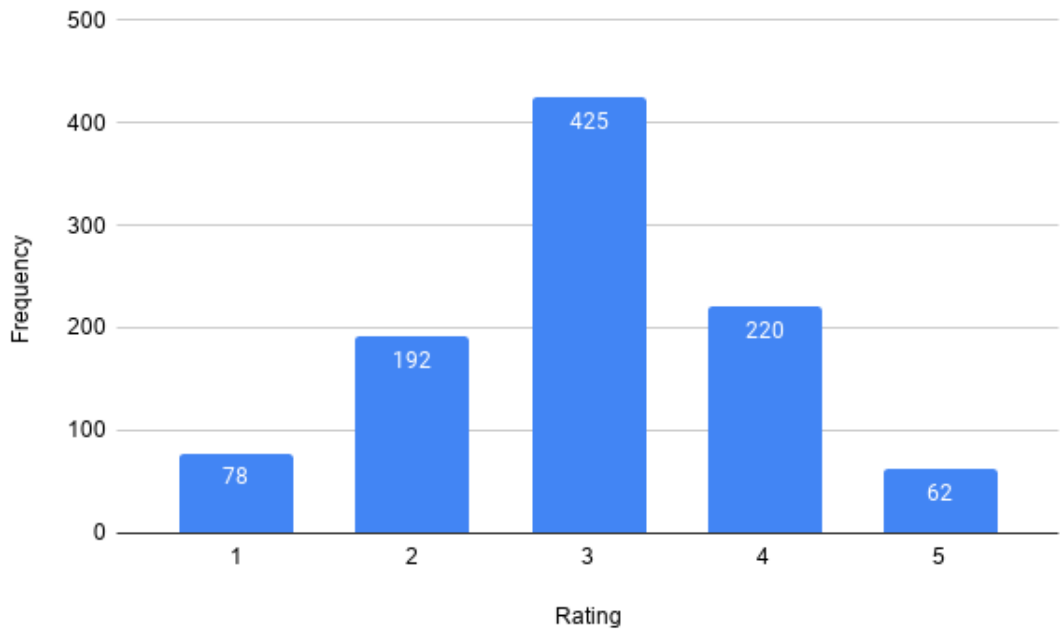



Figure 3.2.1

Answer: [2.9 - 3]

5. Which of the following statement(s) is(are) true?

   (a) Mean is affected by the outliers in the data.
   (b) Adding a constant to the data does not change the mean.
   (c) Mean value calculated for a continuous data from frequency table is equal to the exact mean of the data.
   (d) Mean value calculated for a discrete data from frequency table is equal to the exact mean of the data.

6. What is the mean of the continuous numerical variable shown in Table 3.2.2?

| Class | 10 to 20 | 20 to 30 | 30 to 40 | 40 to 50 | 50 to 60 | 60 to 70 |
|---|---|---|---|---|---|---|
| Frequency | 20 | 25 | 40 | 35 | 30 | 25 |

Table 3.2.2

Answer: 41

# 3 Lecture-3

Describing numerical data - Median and mode

1. During the lock down, the class teacher shared the marks of 10 students of her class with the principal of the school over phone. The principal recorded the marks as follows:
   66 68 70 82 78 67 79 81 75 95
   However, the last student's actual marks was 85 instead of 95.

   i What is the median of the actual marks?
   Answer: 76.5

   ii What is the change in the mode of the marks?

   (a) 82
   (b) 76.5
   (c) 67
   (d) No mode is available for this data.

2. Which of the following statement(s) is(are) true?

   (a) Median is more affected by the outliers in the data than mean.
   (b) Adding a constant to each data point changes the median.
   (c) Median is less affected by the outliers in the data than mean.
   (d) Multiplying each data point by a constant changes the median.

3. The median of a certain dataset is 21. Now, each value in the dataset is incremented by 4 and then multiplied by 4. What is the median of the final dataset?
   Answer: 100

4. The stem and leaf plot shown in Figure 3.3.1 represents ages above fifty years for the persons living in a colony. What is the median of ages?

| Stem | Leaf |
|---|---|
| 5 | 1 2 4 5 7 |
| 6 | 4 7 8 9 |
| 7 | 1 3 5 6 7 8 |
| 8 | 4 5 6 8 9 |

Here 6 | 4 represents 64 years.

Figure 3.3.1

Answer: 72

5. The median of the data $x_1$, $x_2$, $x_3$, ...., $x_n$ is 45. What is the median of the data $3x_1 + 1$, $3x_2 + 1$, $3x_3 + 1$, ...., $3x_n + 1$?
Answer: 136

6. The mode of the marks of students in a class is 66. A teacher awarded 4 additional marks to each student since one question was out of syllabus. What is the new mode of marks after the extra marks were given?
Answer: 70

7. The stem and leaf plot shown in Figure 3.3.2 represents the temperatures at different points inside a furnace in centigrade. What is(are) the mode(s) of the data?

| Stem | Leaf |
|---|---|
| 5 | 1 2 4 5 7 |
| 6 | 4 4 8 9 9 |
| 7 | 1 1 5 6 7 8 |
| 8 | 4 5 6 8 9 |

Here 8 | 2 represents $82^0$C

Figure 3.3.2

(a) 55

(b) 64

(c) 85

(d) 71

(e) 69

8. The mode of the data $x_1$, $x_2$, $x_3$, ...., $x_n$ is 45. What is mode of the data $2x_1 + 5$, $2x_2 + 5$, $2x_3 + 5$, ...., $2x_n + 5$?
Answer: 95

9. The users were asked to give ratings based on user experience by an e-commerce website. The ratings given by users are shown in Figure 3.3.3. What is the mode of the data?

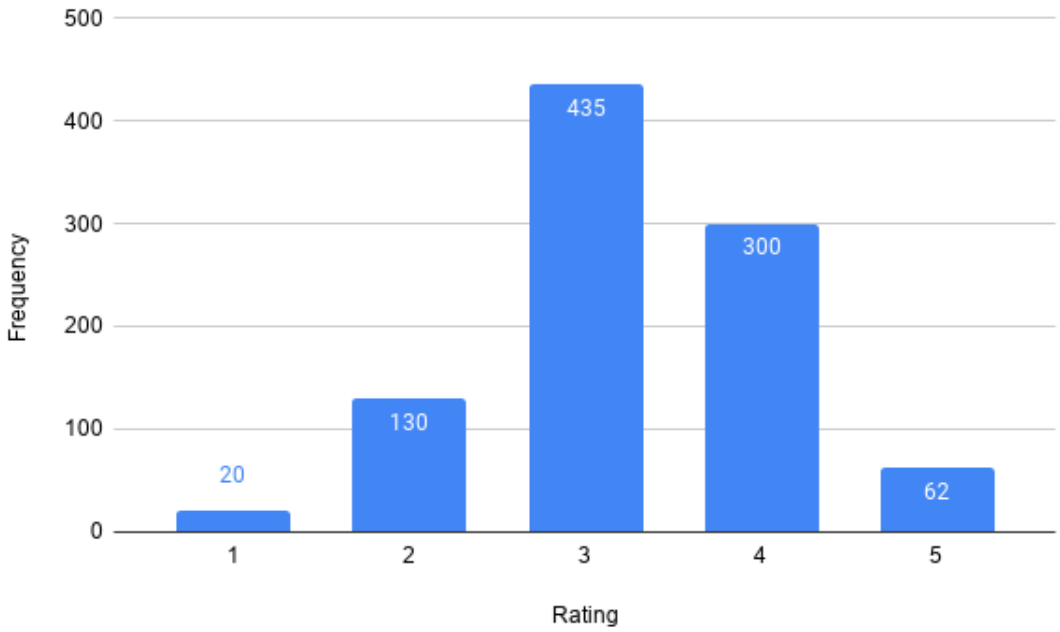



Figure 3.3.3

Answer: 3

# 4 Lecture-4

Describing numerical data - Measures of dispersion

1. A gymnasium owner records the weights of ten of his customers. The weights (in kg) that he recorded are as given below:
   62 85 80 62.5 82 72.4 68 65 73.5 69

   i Calculate the sample variance (in $kg^2$) for the above data.
   Accepted Range: [66 -67]

   ii The gymnasium owner later finds that his weighing machine is defective. The machine shows 3 kg less weight than the actual weight.
   Now, find the new sample variance (in $kg^2$).
   Accepted Range: [66-67]

2. The following data shows the number of orders received by an online shopping website for a cosmetics product for ten consecutive days.
   23 21 28 25 29 32 45 22 28 26

   i What is the sample variance?
   Accepted Range:[47-48]

   ii What is the sample standard deviation?
   Accepted Range: [6.9-7]

3. The heights of ten players in a hockey team are given below (in feet). Based on this information, answer questions (i), (ii), (iii), and (iv).
   5.4 5.5 6 6.1 5.4 5.7 5.9 5.1 5.3 5.7

   i What is the range of the data (in feet)?
   Answer: 1

   ii What is the sample variance of the given data (in $feet^2$)?
   Answer: [0.10 - 0.11]

   iii What is the sample standard deviation of the given data (in feet)?
   Accepted Range: [0.32 - 0.33]

   iv Suppose there was an error in the data collection and the actual value is 0.1 feet more than the collected value for each data point. What is the actual sample standard deviation of heights of the players?
   Accepted Range: [0.32 - 0.33]

4. Which of the following statement(s) is(are) true?

   (a) Standard deviation is affected by outliers.

   (b) Variance is affected by outliers.

   (c) Standard deviation of a numerical data can be negative.

   (d) Standard deviation does not change when we multiply the data by a constant.

# 5 Lecture-5

Describing numerical data - Percentiles and quartiles

1. Which of the following is a measure of dispersion?

   (a) Lower percentile
   (b) IQR (Interquartile Range)
   (c) Second percentile
   (d) Upper percentile

2. The median is

   (a) 25th percentile
   (b) 50th percentile
   (c) 75th percentile
   (d) IQR (Interquartile Range)

3. The stem and leaf plot in Figure 3.5.1 shows the runs scored by 11 players in a school match. What is the 50th percentile of the runs scored?

| Stem | Leaf |
|---|---|
| 1 | 1 3 |
| 2 | 2 5 6 |
| 3 | 3 8 |
| 4 | 2 5 7 9 |

Here 3 | 2 represents 32

Figure 3.5.1

(a) 42

(b) 29.5

(c) 33

(d) 26

4. The number of bicycles sold in a bicycle shop over the first nine days of the month are given below. Based on this information, answer questions (i), (ii), (iii), and (iv).
15 25 35 23 54 37 42 31 29

i What is the value of 9th percentile?
Answer: 15

ii What is the value of 25th percentile?
Answer: 25

iii What is the value of 75th percentile?
Answer: 37

iv What is the Inter Quartile Range (IQR)?
Answer: 12

5. Based on the box plot shown in Figure 3.5.2, answer questions (i), (ii) and (iii).

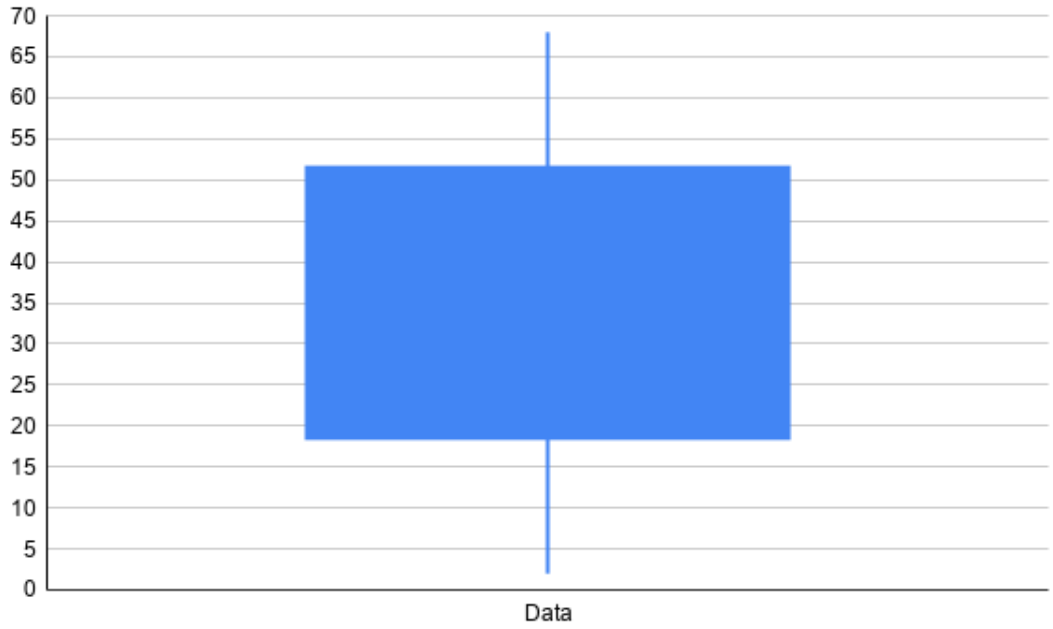



Figure 3.5.2

i What is the approximate first quartile value?

(a) 25
(b) 52
(c) 19
(d) 68

ii What is the value of Inter Quartile Range (IQR)?

(a) 25
(b) 33
(c) 19
(d) 68

iii What is the range of the data?
Accepted Range: 64-66

BSCMA1002: Activity Questions
Week-4

# Contents

# 1 Lecture - 2

Association between two categorical variables: Introduction

1. Data is collected from two websites, A and B, to check whether a customer who visited the site bought an item or not. Total 1500 people visited websites A and B, of which two thirds visited website B. 300 people bought an item from website A. 900 people who visited website B did not buy an item. Based on this data, answer the following questions.

   i. How many people visited website A?
   Answer: 500

   ii. How many people bought an item from website B?
   Answer: 100

   iii. How many people did not buy an item but visited either website A or B?
   Answer: 1100

2. A fruit shop owner wants to understand whether the rating feedback depends on fruits a customer buys. For this the shop owner asked customers to rate the product as good, average or bad and the collected data from 250 customers shown in Table 4.2.1.

| | | Rating | | |
|---|---|---|---|---|
| | | Good | Average | Bad |
| Fruit | Guava | 35 | 15 | 5 |
| | Apple | 50 | 7 | 3 |
| | Orange | 100 | 30 | 5 |

Table 4.2.1: Ratings for fruits purchased

   i. How many customers gave good rating in total for all fruits?
   Answer: 185

   ii. How many customers gave rating less than or equal to average in total for all fruits?
   Answer: 65

   iii. How many customers gave rating greater than or equal to good for guava fruit?
   Answer: 35

# 2 Lecture - 3

Association between two categorical variables: Relative Frequencies

1. In the major e-commerce websites the number of sales on a particular day is given as contingency table [Table 4.3.1].

| | A | B | C | Total |
|---|---|---|---|---|
| Bought item | 1000 | 1500 | 4000 | 6500 |
| Didn't by item | 2000 | 4000 | 20000 | 26000 |
| Total | 3000 | 5500 | 24000 | 32500 |

Table 4.3.1 E-commerce sales data

   i. What is the relative frequency of bought items in website A?
   Answer: 0.333 accepted range 0.333 to 0.34
   ii. What percentage of bought items were from website C on the given day?
   Answer: 61.53 accepted range 61.52 to 61.54
   iii. What is the percentage of customers who bought items after visiting these websites?
   Answer: 20
   iv. Does the customer buying an item if he/she visits a website is associated with the e-commerce company?
      (a) Yes, more customers visiting website A buys an item.
      (b) No, every website has similar row relative frequencies.

2. If there is no association between two categorical variables, then which of the following is true?

   (a) Row relative frequencies of all rows are similar.
   (b) Row relative frequencies of all columns are similar.
   (c) Column relative frequencies of all rows are similar.
   (d) Column relative frequencies of all columns are similar.

# 3 Lecture - 4

Association between two numerical variables: Scatter plot

1. A researcher wants to find out the relation between the age of a person and his Intelligent Quotient (IQ) score. He believes that the IQ score of a person depends on his/her age. The researcher collects the IQ scores of a person at different ages [Table 4.4.1]. Based on this, answer the following questions.

| Age in years | IQ score |
|---|---|
| 10 | 120 |
| 11 | 75 |
| 12 | 135 |
| 13 | 85 |
| 14 | 110 |
| 15 | 98 |
| 16 | 90 |
| 17 | 103 |
| 18 | 101 |
| 19 | 99 |
| 20 | 101 |
| 21 | 105 |
| 22 | 104 |
| 23 | 105 |

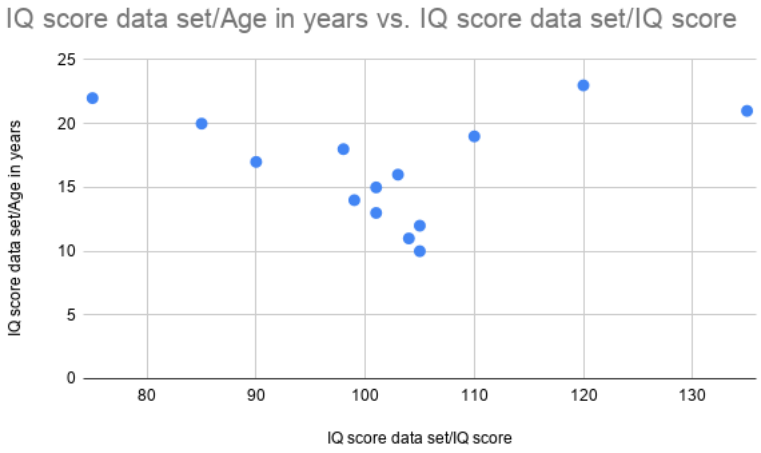
Table 4.4.1: IQ score vs Age of a person in years

i. If a scatter plot is constructed by the researcher, what would be the explanatory variable (X-axis variable)?
Age in years

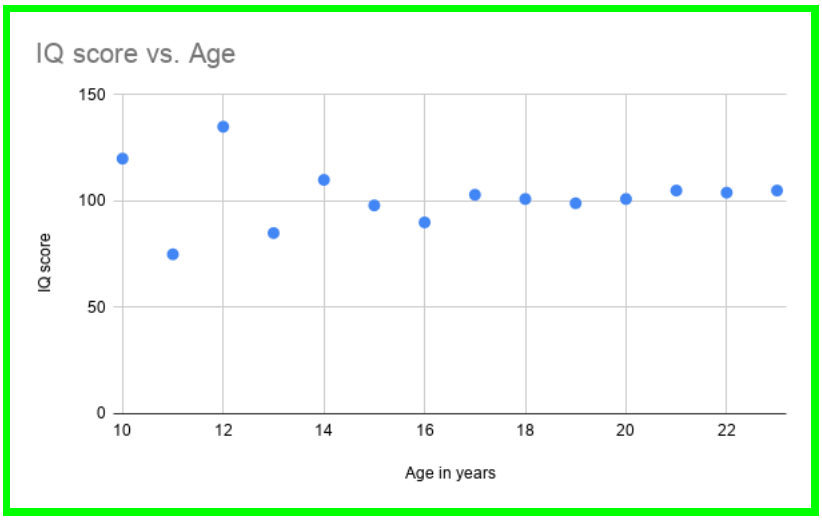
ii. If a scatter plot is constructed by the researcher, what would be the response variable (Y-axis variable)?
IQ score

iii. What is the scatter plot representation of the given data?

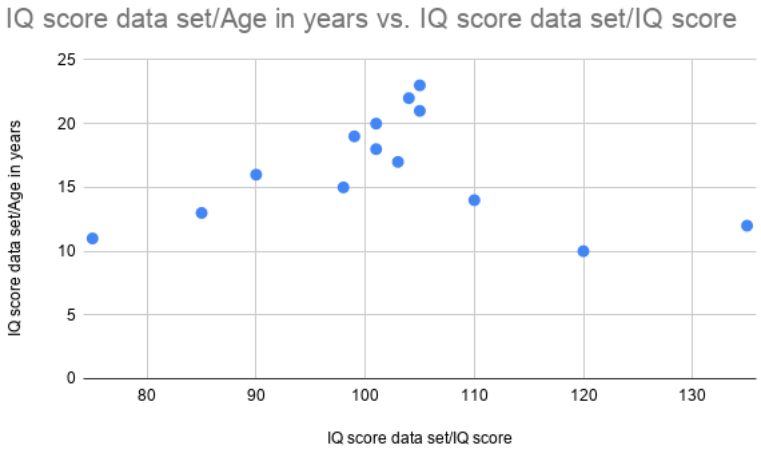
(a)


IQ score data set/Age in years vs. IQ score data set/IQ score
IQ score data set/Age in years
25
20
15
10
5
0
80
90
100
110
120
130
IQ score data set/IQ score

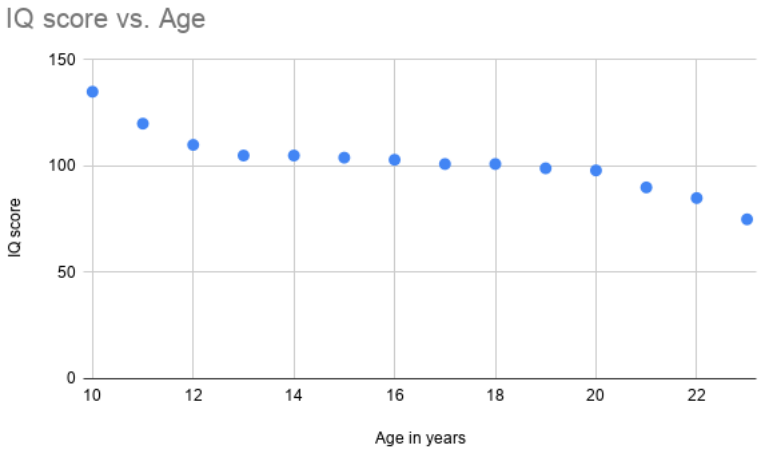
(b)


IQ score vs. Age
IQ score
150
100
50
0
10
12
14
16
18
20
22
Age in years

(c)


IQ score data set/Age in years vs. IQ score data set/IQ score
IQ score data set/Age in years
25
20
15
10
5
0
80
90
100
110
120
130
IQ score data set/IQ score

(d)




iv. Is there a pattern in the scatter plot for the data given in Table 4.4.1?

(a) Yes, the IQ score decreases with an increase in age.

(b) Yes, the IQ score increases with an increase in age.

(c) No, there is no pattern in IQ score and age.

2. Is there a pattern in the scatter plot shown in Figure 4.3.1?

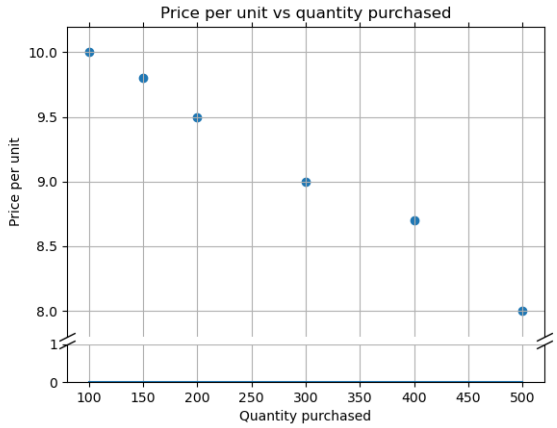



Figure 4.3.1: Scatter plot

(a) Yes, the price decreases with an increase in the quantity purchased.

(b) Yes, the price decreases with a decrease in the quantity purchased.

(c) No, there is no pattern in quantity purchased and the price per unit.

# 4 Lecture - 5

Association between two numerical variables: Describing Association

1. In a mechanical workshop the number of units of products produced per hour versus the average quality rating of the product produced is given in Figure 4.4.1.
   Based on this information, choose the correct option(s).

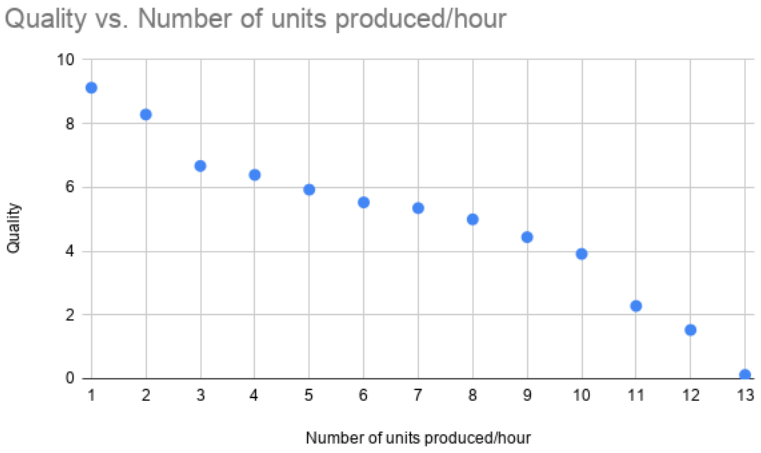



Figure 4.4.1: Quality versus number of products produced

a The curve is trending upwards.

b The curve is trending downwards.

c The scatter plot has outliers.

d The scatter plot does not have outliers.

2. Variation of box office collections with respect to the number of tickets sold is given in Figure 4.4.2. Based on this information, choose the correct option(s).

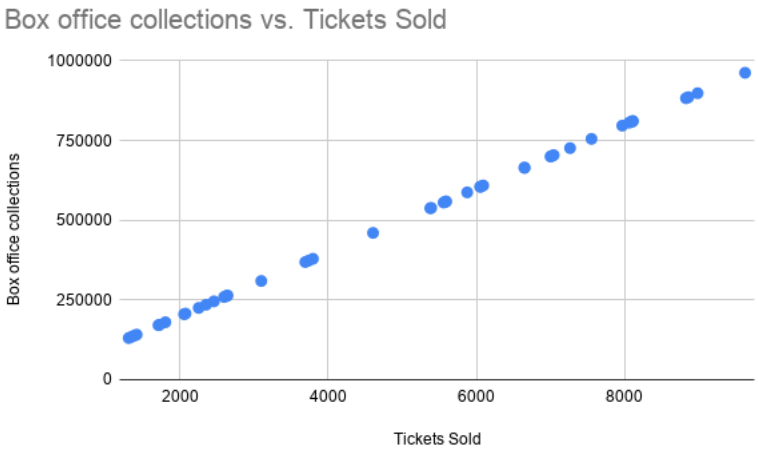



Figure 4.4.2: Box office collections versus tickets sold

(a) The scatter plot is linear.

(b) The scatter plot appears to curve.

(c) The scatter plot has outliers.

(d) The scatter plot does not have outliers.

3. The scatter plot in Figure 4.4.3 shows the surface area of cube for different lengths of cube. Based on the scatter plot, choose the correct option(s).

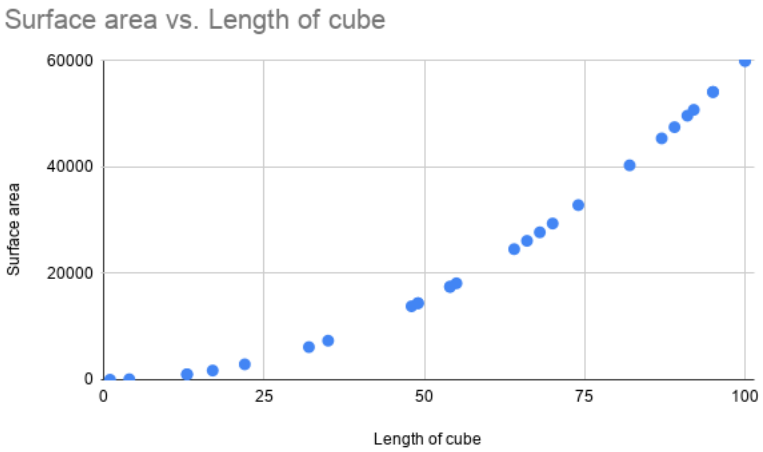



Figure 4.4.3: Surface area versus length of a cube

(a) The curve is trending upwards.

(b) The curve is trending downwards.

(c) The scatter plot is linear.

(d) The scatter plot appears to curve.

(e) The scatter plot has outliers.

(f) The scatter plot does not have outliers.

4. In a small town electricity department is interested in finding the association of power consumed with area of the house. They have collected the data and plotted it as scatter plot in Figure 4.4.4. Based on this information, choose the correct option(s).

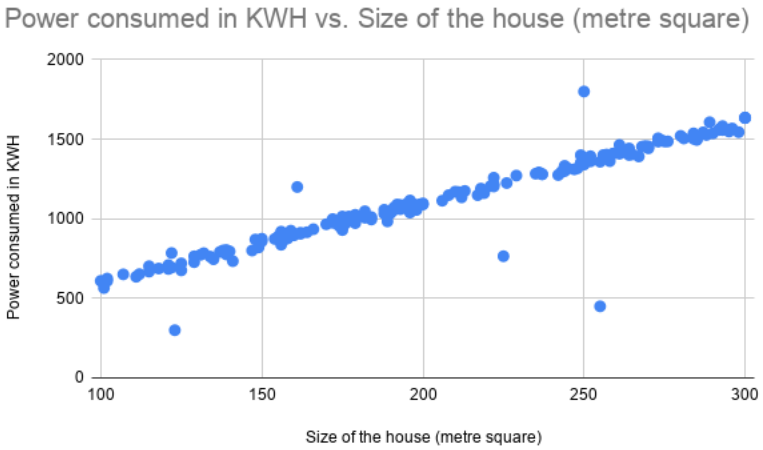



Figure 4.4.4: Power consumed versus Area of the house

(a) The points are variable along the pattern.

(b) The scatter plot is linear.

(c) The scatter plot appears to curve.

(d) The points are tightly clustered.

(e) The scatter plot has outliers.

(f) The scatter plot does not have outliers.

# 5 Lecture - 6

Covariance

1. Covariance between two variables is negative when
   (a) both move in the same direction
   (b) both move in the opposite direction
   (c) values taken by both variables is same
   (d) the respective difference between both variables values is constant
2. Unit of the covariance of the two variables if one of the variable is measured in kilogram and other is measured in metre is
   (a) $\frac{kilogram}{metre}$
   (b) $kilogram^2 \times metre$
   (c) $kilogram \times metre$
   (d) $metre \times kilogram$
   (e) $metre^2 \times kilogram$
   (f) $kilogram^2 \times metre^2$
3. Calculate sample covariance for the following dataset:
   $x$: 2, 4 , 7, 11
   $y$: 30, 49, 60, 55
   (a) 29.25
   (b) 39
   (c) 48.5
   (d) 87.88

# 6 Lecture - 7

Correlation

1. Choose the wrong options among the following.
   (a) Correlation is a measure of quadratic association between two variables.
   (b) Correlation is a measure of linear association between two variables.
   (c) Correlation coefficient can not take negative values.
   (d) If both the variables are taking negative values, does that imply that correlation coefficient between them must be negative.
2. Choose the correct options about values taken by the correlation coefficient $r$ is
   (a) $0 \leq r \leq 1$
   (b) $-1 < r < 1$
   (c) $-1 \leq r \leq 1$
   (d) $-1 \leq r \leq 0$
   (e) $0 \leq r < \infty$
3. The relationship between covariance and correlation of two variables $X$ and $Y$ is
   (a) $Cov(X, Y) = \frac{r}{s_X s_Y}$
   (b) $Cov(X, Y) = r \times s_X s_Y$
   (c) $r = \frac{Cov(X,Y)}{s_X s_Y}$
   (d) $r^2 = \frac{Cov(X,Y)}{s_X s_Y}$
4. Suppose that the value of the correlation coefficient between two variables is -0.85. Then the linear association between them is
   (a) strong and positive
   (b) strong and negative
   (c) weak and negative
   (d) can not conclude anything
5. The covariance between two variables is positive. Then the correlation coefficient
   a) can not take negative values
   b) must be positive

c) can take both negative or positive values depending upon the value of the standard deviation.

d) is always 1

6. If the sample covariance between $x$ and $y$ is 45, sample variance of $x$ is 36 and sample variance of $y$ is 100, then what is the sample correlation coefficient?

(a) 0.55

(b) 0.65

(c) 0.75

(d) 0.85

# 7 Lecture - 8

Fitting a line

1. $R^2$ is
   (a) the measure direction of linear association between two variables.
   (b) the measure of proportion of the variance in the dataset explained by the explanatory variable.
   (c) the measure of goodness of fit of the line to the dataset.
   (d) not the measure of goodness of fit of the line to the dataset.
2. Choose the correct options about values taken by the $R^2$ is
   1. $0 < R^2 < 1$
   2. $0 \leq R^2 \leq 1$
   3. $-1 < R^2 < 1$
   4. $-1 \leq R^2 \leq 0$
   5. $0 \leq R^2 < \infty$
3. The slope and intercept of the line of best fit between price and demand of electronic accessories is 4 and 1 respectively , then the line of the best fit is represented by
   (a) $price = demand + 4$
   (b) $price = 4 \times demand$
   (c) $price = 4 \times demand + 1$
   (d) $price = 4 \times demand - 1$
4. If the slope of the line of best fit is negative, then the correlation between the variables is
   (a) positive
   (b) negative
   (c) zero
   (d) can not conclude anything

# 8 Lecture - 9

Association between numerical and categorical variables

1. Point Bi-serial correlation coefficient is
   (a) the measure of association between two categorical variables.
   (b) the measure of association between two numerical variables.
   (c) the measure of association between numerical and categorical variable.
   (d) the measure of association between explanatory variable and response variable.
2. Compute the value of the Point Bi-serial correlation coefficient from the given set of information:
   $\bar{Y}0 = 937.5, \bar{Y}1 = 455, s_X = 217$
   $\rho_0 = 0.2, \rho_1 = 0.8$
   Answer: 0.88 accepted range 0.86 to 0.90

BSCMA1002: Activity Questions
Week-5

# Contents

# 1 Lecture - 1

Basic principles of counting

1. A student wishes to take either a statistics course or a mathematics course, but not both. If there are two statistics and four mathematics courses for which the student is eligible, then the number of ways the student can choose the course is
   Answer: 6

2. There are thirteen flights from Delhi to Bangalore, and then six flights from Bangalore to Chennai. Riddhi wants to fly from Delhi to Bangalore and then to Chennai. How many choices does she have for her flight plan?

   a) 19
   b) 72
   c) 78
   d) 39

3. Rohan packed four shirts, five pants, and a sweater for his school trip. He will need to choose a shirt and a pant for each outfit and decide whether to wear the sweater. Find the total number of possible outfits.

   a) 20
   b) 10
   c) 40
   d) 9

4. In how many ways can a number be chosen from 1 to 20 such that it is a multiple of 2 or 3?
   Answer: 13

# 2 Lecture - 2

Factorials

1. In how many ways can you arrange the alphabet of the English language, if you were to form all the words that have four alphabets in them?

   a) $\frac{26!}{23!}$

   b) $\frac{26!}{3!}$

   c) $\frac{26!}{4!}$

   d) $\frac{26!}{22!}$

2. The value of the expression $\frac{(n-2)!}{(n-1)!}$ is

   a) $\frac{1}{(n-1)}$

   b) $\frac{(n-2)}{(n-1)}$

   c) $\frac{1}{(n-1)(n-2)}$

   d) $\frac{(n-2)}{(n-1)!}$

3. The value of the expression $\frac{10!}{3!5!}$ is
   Answer: 5040

4. Suppose you have six balls. Each ball has a number marked on it. You also have six slots that you have to fill with the balls. How many different ways can you fill these slots?
   Answer: 720

# 3 Lecture - 3

Permutations: Distinct object

1. In how many ways can a test paper of six questions be attempted, each question being of a TRUE/FALSE type?
   Answer: 729

2. A professor is creating an exam of eleven questions from a test bank of 15 questions. In how many ways can he select and arrange the questions?

   a) $\frac{15!}{11!}$

   b) $\frac{15!}{4!}$

   c) $\frac{15!}{11!4!}$

   d) $\frac{11!}{4!}$

3. In how many ways can four non-identical rings be worn on five fingers?(Assuming people will not wear more than one ring in a finger)
   Answer: 120

4. In how many ways can five distinct books be arranged in two bookshelves?
   Answer: 720

# 4 Lecture - 4

Permutations: Objects not distinct and Circular Permutations

1. Find the number of rearrangements of the letters in the word PROGRAMMING.
   Answer: 4989600
2. In how many ways can the letters of the word SCISSORS be arranged so that all the S's are together?
   Answer: 120
3. Seven people are going to sit at a round table. How many different ways can this be done?
   720
4. Find the number of ways in which eight beads can be arranged to form a necklace.
   Answer: 2520
5. The value of $n$ for which ${}^{n+1}P_2 = 4 \times {}^{n}P_1$ is
   Answer: 3

# 5 Lecture - 5

Combinations

1. Out of a group of six men and eight women, you need to form a committee of three men and five women. In how many ways can the committee be formed?
   Answer: 1120

2. Find the number of subsets of the set {1, 2, 3, 4, 5, 6, 7, 8, 9, 10} having 5 elements.
   Answer: 252

3. You just got three free tickets that allow three people for a movie. If you have six friends who want to come along, how many different groups of three friends could you take with you?
   Answer: 20

4. The value of the expression ${}^{10}C_0 + {}^{10}C_1 + {}^{10}C_2 + {}^{10}C_3 + {}^{10}C_4 + {}^{10}C_5 + {}^{10}C_6 + {}^{10}C_7 + {}^{10}C_8 + {}^{10}C_9 + {}^{10}C_{10}$ is

   (a) $2^{11}$

   (b) $2^{10}$

   (c) $10^2$

   (d) $10^{10}$

5. The value of $n$ for which the ${}^{n}C_2 = {}^{n}C_3$ is
   Answer: 5

6. Fifteen points are chosen in the plane so that no three of them are collinear. How many triangles do they determine?

   (a) ${}^{15}C_3$

   (b) ${}^{15}C_{12}$

   (c) ${}^{15}C_2$

   (d) ${}^{12}C_3$

7. A box contains five red, four white and two blue balls. Three balls are drawn at random. Find out the number of ways of selecting the balls of different colours.
   Answer: 40

# 6 Lecture - 6

Applications

1. Out of six consonants and four vowels, how many words of three consonants and two vowels can be formed?
   Answer: 14400

2. Suppose there are forty men and fifteen women in a party. Each man shakes his hand only with all the men and each woman shakes her hand only with all the women. Find the maximum number of handshakes that take place at the party, if from each group of two persons we have one handshake.
   Answer: 885

3. How many two-digit numbers can be formed by using the digits in 83529, if repetition is not allowed?
   Answer: 20

4. You are going to select first and second place prizes out of six entries. How many possible outcomes are there?
   Answer: 30

BSCMA1002: Activity Questions
Week-6

# Contents

# 1 Lecture - 1

1. The set of all possible outcomes of a random experiment is called
   (a) Event
   (b) Outcome
   (c) Subset
   (d) Sample space
2. Which of the following is a random experiment?
   a) Tossing two coins
   b) Rolling a dice
   c) Guessing a answer in multiple choice questions
   d) Choosing a number at random from 1 to 20
3. The cardinality of the sample space of rolling a pair of dice is
   36
4. Suppose we record the total number of coin tosses in tossing a coin until we see two consecutive tails, then the sample space $S$ is
   a) $S = \{1, 2, 3, 4, ...\}$
   b) $S = \{2, 3, 4, ...\}$
   c) $S = \{3, 4, ...\}$
   d) $S = \{0, 1, 2, 3, 4, ...\}$
5. Choose the incorrect statement among the following.
   1. An outcome is the result of a random experiment.
   2. The set of all possible outcomes of a random experiment is called the sample space.
   3. An event is a subset of the sample space.
   4. None of the above.

# 2 Lecture - 2

1. A single fair die is rolled. Let $A$ be the event "outcome is an even number" and $B$ be the event "outcome is a number divisible by 3". Then the cardinality of the event $A \cup B$ is
   4

2. For the sample space $S = \{1, 2, 3, 4, 5, 6, 7, 8\}, E = \{1, 2, 5\}, F = \{4, 6, 7\}$, and $G = \{1, 3, 8\}$ are the events. Then

   a) $E \cup F = \{1, 2, 5, 6, 7\}$

   b) $E \cap G = \{1\}$

   c) $(F \cup G)^C = \{2, 5\}$

   d) $E \cup (F \cap G) = \{1, 2, 5\}$

3. Which of the following is (are) disjoint events?

   a) A die is rolled. Event $E$ is that it lands on 5, and $F$ is the event that it lands on an even number.

   b) A person is chosen. Event $E$ is that this person was born in India, and $F$ is the event that this person is an Indian citizen.

   c) A woman is chosen. Event $E$ is that she is over 40 years of age, and $F$ is the event that she has been married for over 40 years.

   d) A die is rolled. Event $E$ is that it lands on an odd number, and $F$ is the event that it lands on an even number.

4. A fair die is thrown. If event $A$ represents all the outcomes of getting a natural number less than 4, event $B$ represents all the outcomes of getting an even number and event $C$ represents all the outcomes of getting a number greater than three. Then the set representing the events $A$ and $B$ but not $C$ is:

   (a) {2, 4}

   (b) {2}

   (c) { 1, 2, 5}

   (d) {4}

5. The event of $A$ occurring and $B$ not occurring for two events $A$ and $B$ of a random experiment can be represented as

   (a) $(A \cup B)^C$

(b) $A \cap B^C$

(c) $A \cup B^C$

(d) $(A \cap B)^C$

# 3 Lecture - 3

1. The king, queen and jack of clubs are removed from a deck of 52 playing cards and then the deck is shuffled. Let $E$ be the event of getting a queen and $F$ be the event of getting a club, then the cardinality of event $E \cup F$ is
   13

2. Which of the following is (are) mutually exclusive events?

   a) Selecting a king and selecting a heart from a deck of 52 playing cards.

   b) Selecting a king and selecting a ace from a deck of 52 playing cards.

   c) Turning left or right while walking.

   d) Turning right and scratching your head while walking.

3. Consider a sample space $S$ and three events $E, F$, and $G$ as shown in Figure 6.3.1. What event does the yellow coloured region represent?

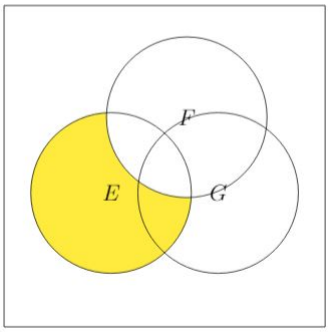



Figure 6.3.1: Venn diagram

   (a) At least one of the events $E, F$, or $G$ occurs.

   (b) $E$ or $G$ occurs, but not $F$.

   (c) $E$ occurs but not $F$.

   (d) Among $E, F$, and $G$, only $E$ occurs.

4. The coloured region is

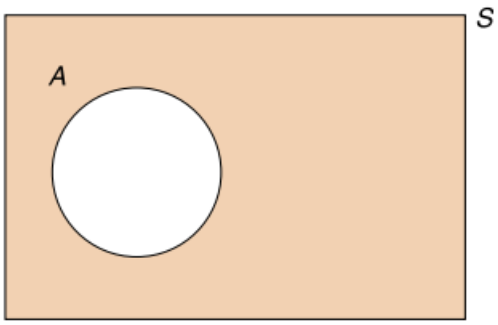



Figure 6.3.2: Venn diagram

(a) $A$

(b) $S \setminus A$

(c) $A^C$

(d) $S \cap A$

# 4 Lecture - 4

1. Which of these numbers cannot be a probability?

   a) 0.5
   b) -0.001
   c) 1.01
   d) 100
   e) 0
   f) 1

2. A single 6 - sided die is rolled. What is the probability of getting a number less than or equal to 6?
   1

3. A customer will purchase a shirt with probability 0.4. The customer will purchase a pant with probability 0.3 and will purchase both a shirt and a pant with probability 0.2. What is the probability that the customer purchases neither a shirt nor a pant?
   0.5

4. Shabnam has 74% chance of receiving an A in either Statistics or Mathematics. Find the probability that she does not receive an A in either Statistics or Mathematics.
   .26

5. Among 40 dieters following a similar routine, 26 lost weight, 8 gained weight, and 6 remained the same weight. If one of these dieters is randomly chosen, find the probability that he or she gained weight.
   0.2

6. 57% students of girls' school wear neither a ring nor a necklace, 30% wear a ring, and 25% wear a necklace. If one of them is randomly chosen, find the probability that she is wearing a ring and a necklace.
   0.12

7. Arjun wants to buy a cricket bat and a cricket ball. The probability of him buying a bat is 0.25. The probability that he will buy a ball is 0.35. The probability that he will buy both a bat and a ball is 0.15.
   What is the probability that he will buy either a ball or a bat?

   a) 0.6
   b) 0.55
   c) 0.95
   d) 0.45

# 5 Lecture - 5

1. In the roll of a die, what is the probability of getting a prime number?
   0.5

2. Find the probability of selecting a red card or a five from a deck of 52 cards.

   a) $\frac{1}{2}$
   b) $\frac{15}{26}$
   c) $\frac{7}{13}$
   d) $\frac{3}{13}$

3. A coin is tossed thrice. What is the probability of getting three consecutive tails?
   0.125

4. What is the probability of drawing a king and an ace consecutively from a deck of 52 cards, with replacement?

   a) $\frac{1}{169}$
   b) $\frac{8}{13}$
   c) $\frac{1}{4}$
   d) $\frac{4}{169}$

5. Suppose that when two dice are rolled, each of the 36 possible outcomes is equally likely. Find the probability that the sum of the outcomes of rolling the two dice is 6.

   a) $\frac{5}{36}$
   b) $\frac{7}{36}$
   c) $\frac{1}{9}$
   d) $\frac{1}{12}$

BSCMA1002: Activity Questions
Week-7

# Contents

# 1 Lecture - 1

Contingency tables

Use the following information to answer the questions from 1 to 6.

A bank wants to analyze how the loan default(Yes or No) and profession of people are related. To study this hypothesis, 2000 people's data were randomly collected from the previous transaction history. The number of people in each category is listed in Table 7.1.1. Suppose that a person is randomly chosen from this group.

| | Profession | | |
|---|---|---|---|
| | Doctor | Teacher | Chef |
| Yes | 150 | 300 | 50 |
| No | 400 | 900 | 200 |

Table 7.1.1: Loan default vs profession

1. What is the probability that the selected person is a teacher?
   0.6

2. What is the probability that the person has defaulted in loan?
   0.25

3. What is the probability that the person has not defaulted in loan?
   0.75

4. What is the conditional probability that the person has defaulted in loan given that the person is a chef?
   0.2

5. What is the conditional probability that the person is a doctor given that the person has defaulted in loan?
   0.3

6. What is the conditional probability that the person has not defaulted in loan given that the person is a teacher?
   0.75

# 2 Lecture - 2

Conditional probability formula

1. A games club has 80 members out of whom 70 play chess, 45 play table tennis, and 35 play both chess and table tennis. If a member of the club is randomly chosen, find the conditional probability that he or she plays table tennis given that he or she plays chess.
   0.5

2. There are $n$ balls in a drawer, of which 7 are red. If 2 balls are chosen randomly, then the probability that both are red is 3/4. Find $n$.
   8

3. A card is randomly selected from a deck of 52 playing cards. What is the conditional probability that the card is an ace given that the card is a face card?
   0

4. Two fair dice are rolled. What is the conditional probability that at least one of the dice lands on 5 given that their sum is 11?
   1

# 3 Lecture - 3

Multiplication rule of probability

1. A box contains 8 white balls and 4 black balls. Two balls are drawn at random one after the other without replacement.Find the probability that both the balls drawn are black.
   (a) $\frac{3}{11}$
   (b) $\frac{1}{11}$
   (c) $\frac{20}{33}$
   (d) $\frac{1}{3}$

2. Suppose you take out two cards from a standard pack of 52 cards one after the other, without replacing the first card. What is the probability that the first card is the ace of club, and the second card is a diamond?
   (a) $\frac{1}{208}$
   (b) $\frac{1}{204}$
   (c) $\frac{7}{26}$
   (d) $\frac{727}{2652}$

3. Consider that you have a box with 5 blue marbles, 3 red marbles, and 4 yellow marbles. You are going to pull out one marble, record its color, put it back in the box and draw another marble. What is the probability of pulling out a blue marble followed by a yellow marble?

(a) $\frac{5}{48}$
(b) $\frac{5}{33}$
(c) $\frac{5}{36}$
(d) $\frac{3}{4}$

4. Suppose that three people are randomly chosen from a group of 6 girls and 8 boys. What is the probability that all three are girls?

(a) $\frac{3}{14}$
(b) $\frac{5}{12}$
(c) $\frac{5}{91}$
(d) $\frac{1}{16}$

# 4 Lecture - 4

Independent events

1. Which of the following options is (are) correct?

(a) If events $A$ and $B$ are disjoint, then $P(A \cup B) = P(A) + P(B)$.
(b) If events $A$ and $B$ are independent, then $P(A \cup B) = P(A) + P(B)$.
(c) If events $A$ and $B$ are disjoint, then $P(A \cap B) = P(A) \times P(B)$.
(d) If events $A$ and $B$ are independent, then $P(A \cap B) = P(A) \times P(B)$.

2. There are two events $A$ and $B$ such that $P(A) = 0.4$ and $P(B) = 0.3$. The events $A$ and $B$ are independent.
i) Find $P(A^C)$
0.6
ii) Find $P(A \cap B)$
0.12
iii)Find $P(A \cup B)$
0.58

3. Let $X$ and $Y$ be two independent events such that $P(X) = 0.4$ and $P(Y) = 0.6$. Find the probability that neither $X$ nor $Y$ will occur.
0.24

# 5 Lecture - 5

Independent events examples

1. Suppose that we roll a pair of fair dice, so each of the 36 possible outcomes is equally likely. Let $A$ denote the event that the first die lands on 5, $B$ be the event that the sum of the outcomes of rolling the pair of dice is 10, and $C$ be the event that the sum of the outcomes of rolling the pair of dice is 7. Then

    1. Event $A$ and event $B$ are independent.
    2. Event $A$ and event $B$ are dependent.
    3. Event $A$ and event $C$ are independent.
    4. Event $A$ and event $C$ are dependent.

2. Suppose that we are going to roll two fair 4 - sided dice. Find the probability that both dice show 2. Enter the answer correct upto 4 decimals accuracy.
   0.0625

3. The probability of a flight being delayed is 0.4. Find the probability of no delays on a round trip.
   0.36

4. A die is thrown twice. What is the probability that both outcomes are prime numbers?
   0.25

# 6 Lecture - 6

Independent events properties

1. A male New Delhi resident is randomly selected. Which of the following pairs of events $A$ and $B$ can reasonably be assumed to be independent?

    a) $A$: He is a doctor.
    $B$: He has brown eyes.

    b) $A$: He had body pain yesterday.
    $B$: He was in an accident yesterday.

    c) $A$: He is wearing blue shoe.
    $B$: He used to drink coffee after reaching the office.

2. Which of the following option(s) is (are) correct?

    (a) If events $A$ and $B$ are independent, then events $A$ and $B^C$ are independent.

(b) If events $A$ and $B$ are independent, then events $A^C$ and $B$ are not independent.

(c) If events $A$ and $B$ are independent, then events $A^C$ and $B^C$ are independent.

(d) If events $A^C$ and $B^C$ are independent, then events $A$ and $B$ are not independent.

3. Event $A$ is attending a math class and event $B$ is attending a statistics class. Suppose $P(A) = 0.8, P(B^C) = 0.4$, and $P(A \cap B) = 0.4$. Are events $A$ and $B$ independent?

   a) Yes

   b) No

4. Let $A, B$ and $C$ be the three events such that $P(A \cap B \cap C) = P(A)P(B)P(C)$. Then, the events $A, B$ and $C$ are independent.

   a) True

   b) False

# 7 Lecture - 7

Bayes' rule

1. Company $A$ produces 30% defective products, Company $B$ produces 40% defective products and $C$ produces 30% defective products. If choosing a company is an equally likely event, then find the probability that the product chosen is not defective.

   a) 1/2

   b) 2/3

   c) 1/3

   d) 1/6

2. In a badminton practice session, the probability that Dishu serves properly is 0.6 and that Rohan serves properly is 0.7. If there are only two players and assume that serving in the practice session is an equally likely event, then find the probability that it is served properly.
   0.65

3. Let there be two newly launched motorcycles $X$ and $Y$. The probability that motorcycle $X$ has good mileage is 0.5 and the probability that motorcycle $Y$ has good mileage is 0.7. If a motorcycle is chosen at random, then find the probability that the motorcycle has good mileage.
   0.60

4. A task is given to 3 students Riddhi, Shivansh, and, Prateek. If the probability of completing the task individually is 1/2, 1/4, and, 1/8 respectively, then find the probability that the task is completed.

a) $\frac{7}{24}$
b) $\frac{7}{8}$
c) $\frac{1}{64}$
d) $\frac{17}{24}$

5. Urn I contains 2 red and 8 black balls while urn II contains 6 red and 4 black balls. One ball is drawn at random from one of the urns, and it is found to be black. Find the probability that it was drawn from urn II.

a) $\frac{1}{3}$
b) $\frac{2}{3}$
c) $\frac{1}{6}$
d) $\frac{3}{4}$

6. A test for a disease gives a correct positive result with a probability of 0.90 when the disease is present, but gives an incorrect positive result with a probability of 0.10 when the disease is not present. If 10% of the population has the disease, and Rohini tests positive, then what is the probability that Rohini has the disease?
0.5

7. Suppose that in IIT Madras 40% of the students are girls. It is known that 70% of the boys play cricket and 60% of the girls play cricket while 40% of the boys play football and 50% of the girls play football. Given that 33% of those who play cricket also play football. A student is randomly chosen and it is known that the student plays cricket, then what is the probability that the student plays football?
0.22

# BSCMA1002: Activity Questions
Week-8

**Contents**

## 1 Lecture-8.1

Random variables - Introduction

1. A random variable is

   (a) A subset of the sample space of a random experiment.

   (b) A relation defined on the sample space.

   (c) A real valued function defined on the sample space.

   (d) None of the above.

2. Range of a random variable $X$ is the set of all possible values that $X$ can take. Let a random variable $Y$ be defined as the number of tosses until first head comes on tossing a coin. Then what will be the range of $Y$?

   (a) {1,2}

   (b) {0,1,2}

   (c) {1,2,3,...}

   (d) {0,1,2,3, ...}

   Use the following information to answer questions (3) and (4).
   Two cards are drawn simultaneously from a well shuffled deck of 52 playing cards. Each of the face cards carries 10 points and other cards carry the same points as its number (assume ace carries one point). Let $X$ be defined as the sum of the points of withdrawn cards.

3. What values does $X$ take?

   (a) 1, 2, ..., 10

   (b) 1, 2, ..., 20

   (c) 2, 3, ..., 10

   (d) 2, 3, ..., 20

4. How many outcomes will result in $X = 19$?
   Answer: 64

5. A random variable $X$ is defined as the time (in hours) from now until the next rain occurs in a certain city. What are the values that X can take?

   (a) $[0, \infty)$

   (b) Natural numbers

(c) $[0, 24]$

(d) $(-\infty, \infty)$

6. Three coins are tossed simultaneously. Let a random variable $X$ be defined as number of heads minus number of tails resulting from the toss of the three coins. What are the values that X can take?

   (a) {0, 1, 2}

   (b) {-2, -1, 0, 1, 2}

   (c) {-3, -1, 1, 3}

   (d) {-3, -2, -1, 0, 1, 2, 3}

7. Two cards are withdrawn from a deck of 52 cards one after the other without replacement. A random variable $X$ is defined as

$$X = \begin{cases} 0 & \text{if both are face cards} \\ 1 & \text{if exactly one card is face card} \\ 2 & \text{if none of them is face card} \end{cases}$$

   What is the probability of $X = 1$?

   (a) 0.18

   (b) 0.36

   (c) 0.053

   (d) 0.60

8. A small village has 120 female voters and 90 male voters. A sample of 10 voters is chosen. Let a random variable $X$ be defined as the number of male voters in the chosen sample. What is the probability that $X = 5$?

   (a) $\frac{^{90}C_5}{^{210}C_{10}}$

   (b) $\frac{^{90}C_5\,^{120}C_5}{^{210}C_{10}}$

   (c) $\frac{^{120}C_5}{^{210}C_{10}}$

   (d) $\frac{^{90}C_5}{^{210}C_5}$

9. An insurance salesman visits up to three clients each day, hoping to sell a new policy. He stops for the day once he makes a sale. Each client independently decides whether to buy a policy. 10% of clients purchase the policy. Let $X$ be defined as the number of sales he made in a day. Find the value of $P(X = 0)$. (Write your answer upto three decimal places)

   Answer: 0.729 Accepted range:[0.7, 0.8]

10. A postman has three posts to deliver in three adjacent houses in his area. But he is not sure which post belongs to which house. He delivers the posts randomly. Let a random variable $X$ be defined as number of houses he delivers the right post. What is the number of relevant outcomes for $X = 2$?

    (a) 0

    (b) 2

    (c) 3

    (d) 6

11. Four students are chosen randomly from a class of 10 boys and 10 girls. A random variable $X$ be defined as the number of girls in chosen students. What will be the cardinality of set of the relevant outcomes for $X = 2$?

    (a) 45

    (b) 120

    (c) 5400

    (d) 2025

## 2 Lecture-8.2

Random variables - Application

1. A machine measures length of the synthetic fabrics and associated imperfections in the fabric. Let $X$ be the error in the measurement of length and $Y$ be the number of imperfections per 10 meters. Choose the correct option(s).

    (a) $X$ is a continuous variable.

    (b) $X$ is a discrete variable.

    (c) $Y$ is a continuous variable.

    (d) $Y$ is a discrete variable.

2. A shopkeeper is interested in knowing the number of customers entering the shop every hour in a day and the total sale of every hour of the day. Let $X$ be the number of customers entering the shop every hour in a day and $Y$ be the time between the arrival of two customer. Choose the correct options.

    (a) $X$ takes value in $[0, \infty)$.

    (b) $X$ takes value in $N \cup \{0\}$ (Where $N$ is the set of natural numbers).

    (c) $Y$ is a discrete variable.

    (d) $Y$ is a continuous variable.

3. There are 10 blue and 5 red balls in a bag. A ball is randomly drawn and its colour is noted and put back in the bag. Again a second ball is drawn and its colour is noted. Let a random variable $X$ be defined as

$$X = \begin{cases} 1 & \text{if both balls are of same colour} \\ 0 & \text{if both balls are of different colour} \end{cases}$$

How many draws will result in $X = 1$?

(a) 125

(b) 25

(c) 100

(d) 150

4. An information source generates symbols at random from a four-letter alphabet {a, b, c, d} with probabilities $P(a) = \frac{1}{2}$, $P(b) = \frac{1}{4}$, and $P(c) = P(d) = \frac{1}{8}$. A coding scheme encodes these symbols into binary codes as follows:

$$a = 0$$

$$b = 10$$

$$c = 110$$

$$d = 111$$

Let $X$ be the random variable denoting the length of the code, that is, the number of binary symbols (bits) (0s and 1s). What is the probability for $X = 3$?

(a) $\frac{1}{2}$

(b) $\frac{1}{4}$

(c) $\frac{1}{8}$

(d) $\frac{1}{16}$

## 3 Lecture-8.3

Random variables - Discrete and continuous random variable

1. Failure rate is the frequency with which an engineered system or component fails, expressed in failures per unit of time. Let $X$ be the time to the failure of an automobile, then $X$ is a:

(a) Discrete random variable

(b) Continuous random variable

2. Let a rod of length 10cm be broken into two parts. Let $X$ denote the length of smaller part, then which of the following statements is (are) true?

   (a) $X$ is a discrete random variable.

   (b) $X$ is a continuous random variable.

   (c) $X$ takes value in [0, 10].

   (d) $X$ takes value in [0, 5].

3. Which of the following are discrete random variables?

   (a) Length of the phone call.

   (b) Number of the customers visiting every hour in the shop.

   (c) Amount of milk produced yearly by a particular cow.

   (d) Number of accidents taking place every hour on NH10.

   (e) Travel time from your home to the office.

   (f) Time interval between two succesive goals in a football match.

   (g) Proportion of people who cast their votes in the favour of National Party.

4. A random variable $X$ is defined as the length of the hypotenuse of the right-angled triangle whose other two sides are determined by roll of two 6-sided dice. Choose the correct option(s).

   (a) Number of values $X$ can take is 36.

   (b) Number of values $X$ can take is 18.

   (c) Number of values $X$ can take is 21.

   (d) $X$ is a discrete variable.

   (e) $X$ is a continuous variable.

## 4 Lecture-8.4

Discrete random variables - Probability mass function- properties

1. A random variable takes the values $1, 2, 3, 4$, and 5. Table 8.5.1 gives the PMF $P(X = x)$ of $X$.

| $X$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $P(X = x)$ | 0 | $t$ | $\frac{t}{2}$ | $4t^2$ | $\frac{3t}{2}$ |

Table 8.5.1: Probability mass function

   Find the probability of $X = 5$ ($P(X = 5)$).

(a) $\frac{1}{4}$

(b) $\frac{3}{4}$

(c) $\frac{3}{8}$

(d) $\frac{1}{8}$

2. Which of the following can never be a probability mass function for a random variable?

(a)

| $X$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $P(X = x)$ | 0 | -0.25 | 0.12 | 0.13 | 1 |

(b)

| $X$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $P(X = x)$ | 0.20 | 0.15 | 0.35 | 0.17 | 0.14 |

(c)

| $X$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $P(X = x)$ | 0.5 | 0.25 | 0.12 | 0.10 | 0.03 |

(d)

| $X$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $P(X = x)$ | 0 | 1 | 0 | 0 | 1 |

Answer: a, b, d

3. Probabilty mass function of a random variable is given by:

$$P(X = x) = \begin{cases} k & \text{if } x = 0 \\ 0.4 + 2k & \text{if } x = 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the value of $k$.

(a) 0.6

(b) 0.3

(c) 0.1

(d) 0.2

4. Keerthana plays a game with Shrikant that involes toss of three coins simultaneously. Let X be a random variable that denotes the number of heads in the toss. Probability distribution of $X$ is given as:

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(X = x)$ | 0.125 | 0.375 | 0.375 | 0.125 |

Find the probability that the number of tails is more than one.

(a) 0.5

(b) 0.125

(c) 0.375

(d) 0.625

5. Let $X$ be the number of candies present in a box. We have the following information:
• There are at most three candies in the box.
• The probabilty of having 2 candies in the box is the same as the probability of having one candy.
• The probability of having no candy in the box is the same as the probability of having 3 candies.
• The probability of having 1 or 2 candies is half of the probability of having 0 or 3 candies.
What will be the PMF of $X$?

(a)

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(X=x)$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{3}$ | $\frac{1}{3}$ |

(b)

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(X=x)$ | $\frac{1}{3}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{3}$ |

(c)

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(X=x)$ | $\frac{1}{3}$ | $\frac{1}{3}$ | $\frac{1}{6}$ | $\frac{1}{6}$ |

(d)

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(X=x)$ | $\frac{1}{6}$ | $\frac{1}{3}$ | $\frac{1}{3}$ | $\frac{1}{6}$ |

Answer: b

6. The probability mass function of a discrete random variable $X$ is given by

$$P(X=x) = k({}^2C_x.{}^3C_{3-x}); \text{for } x = 0, 1, 2$$

Find the value of $k$.

(a) $\frac{1}{10}$

(b) $\frac{1}{8}$

(c) $\frac{1}{5}$

(d) $\frac{1}{3}$

7. A shipment of 20 similar computers to a retail outlet contains 3 defective computers. If a school makes a random purchase of 2 of these computers, find the probability mass function for the number of defectives.

(a)

| $X$ | 0 | 1 | 2 |
|---|---|---|---|
| $P(X=x)$ | $\frac{68}{95}$ | $\frac{51}{190}$ | $\frac{3}{190}$ |

(b)

| $X$ | 0 | 1 | 2 |
|---|---|---|---|
| $P(X=x)$ | $\frac{68}{95}$ | $\frac{41}{190}$ | $\frac{13}{190}$ |

(c)

| $X$ | 0 | 1 | 2 |
|---|---|---|---|
| $P(X=x)$ | $\frac{68}{190}$ | $\frac{51}{190}$ | $\frac{3}{190}$ |

(d)

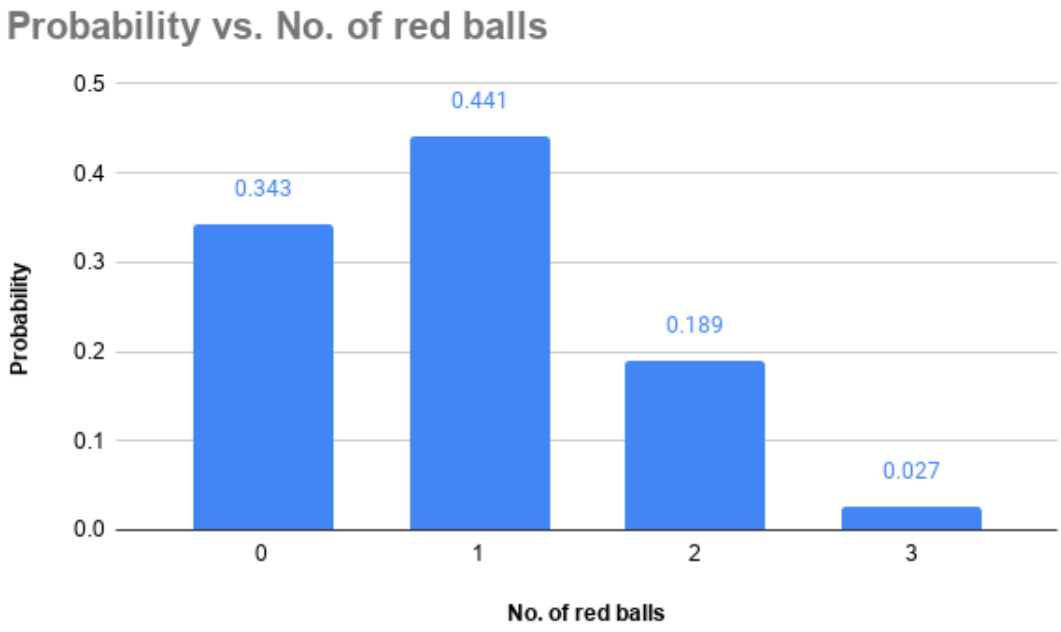
| $X$ | 0 | 1 | 2 |
|---|---|---|---|
| $P(X=x)$ | $\frac{51}{95}$ | $\frac{68}{190}$ | $\frac{3}{190}$ |

Answer: a

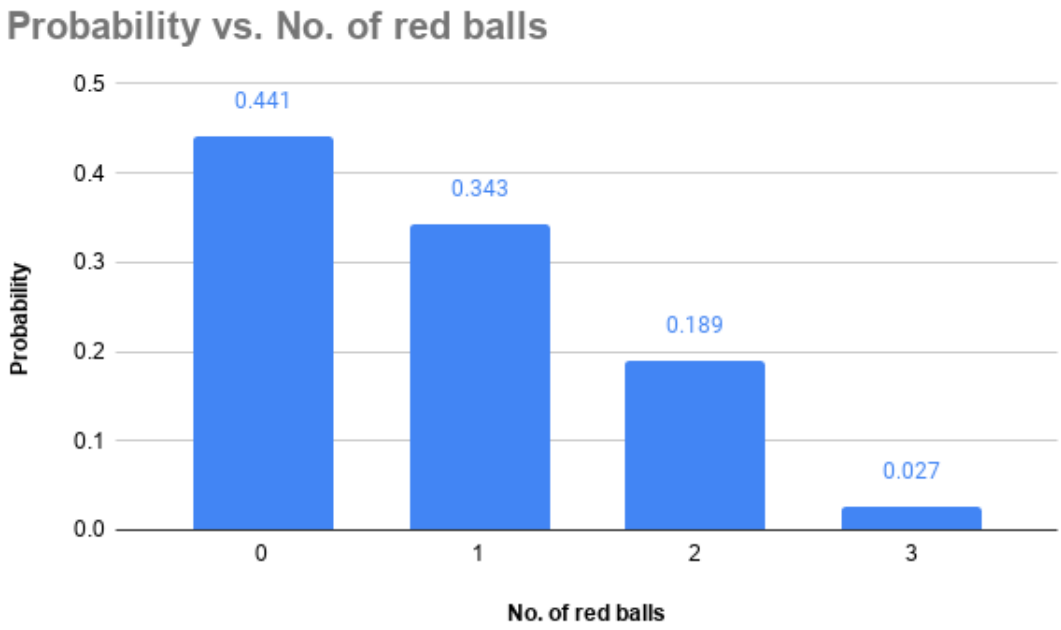
## 5 Lecture-8.5

Discrete random variables - Graph of probability mass function

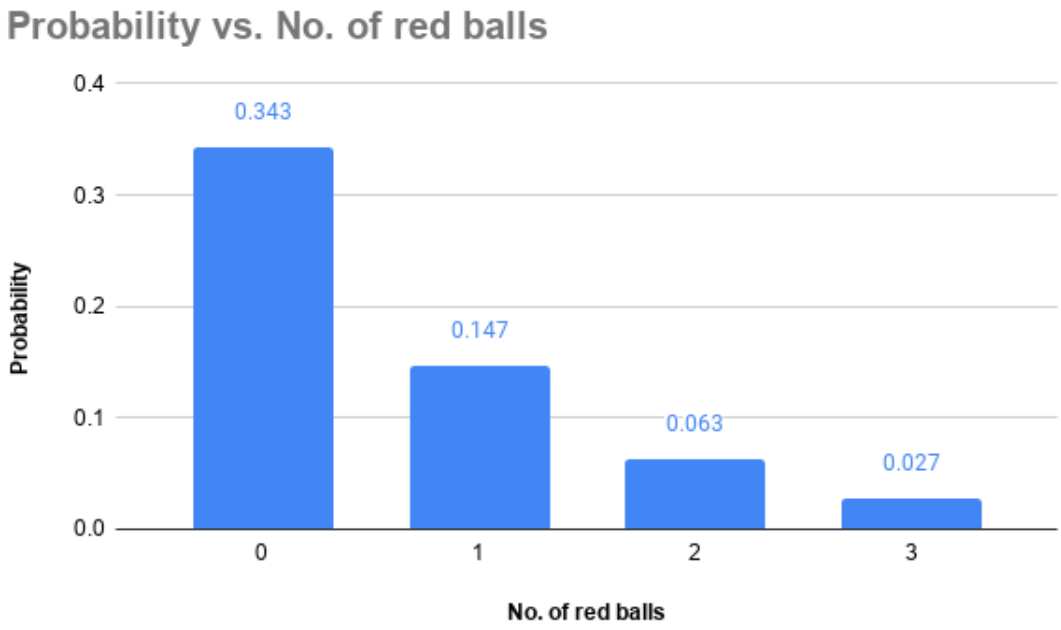
1. There are three red and seven blue balls in a bag. Three balls are randomly drawn one by one with replacement. Let $X$ be the number of red balls in three draws then which of the following is the correct probability mass function of $X$?



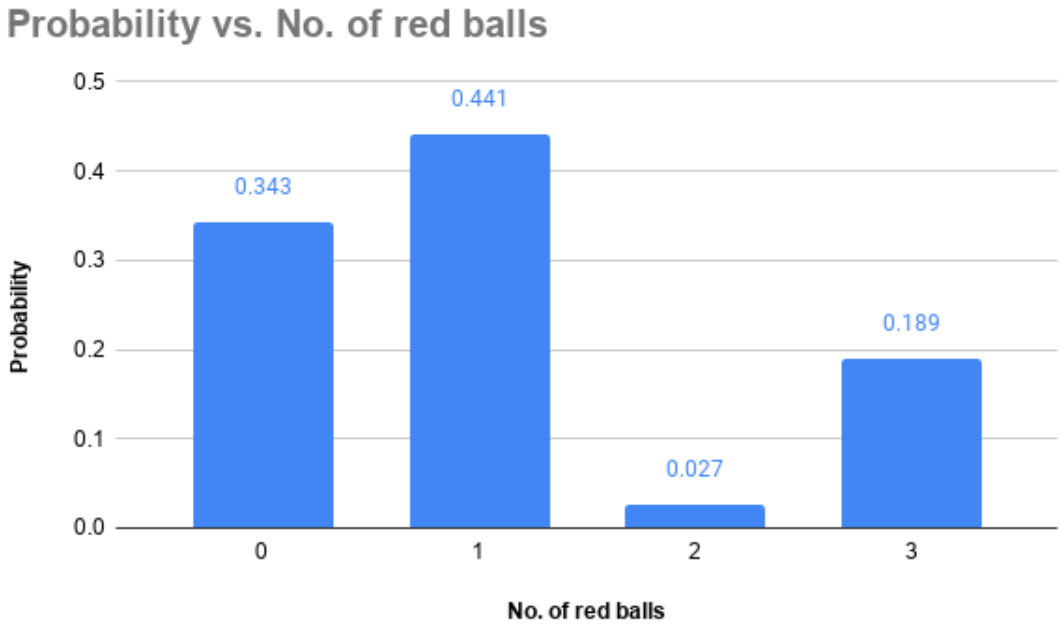

(a)


Probability vs. No. of red balls
0.5
0.4
0.3
0.2
0.1
0.0
Probability
0.441
0.343
0.189
0.027
0
1
2
3
No. of red balls

(b)


Probability vs. No. of red balls
0.4
0.3
0.2
0.1
0.0
Probability
0.343
0.147
0.063
0.027
0
1
2
3
No. of red balls

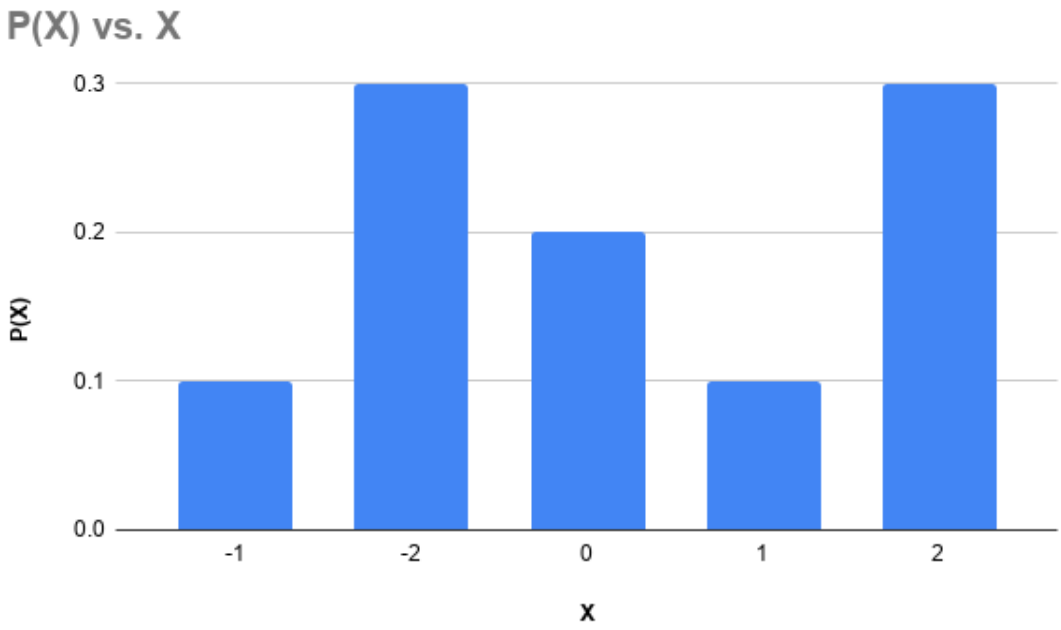
(c)



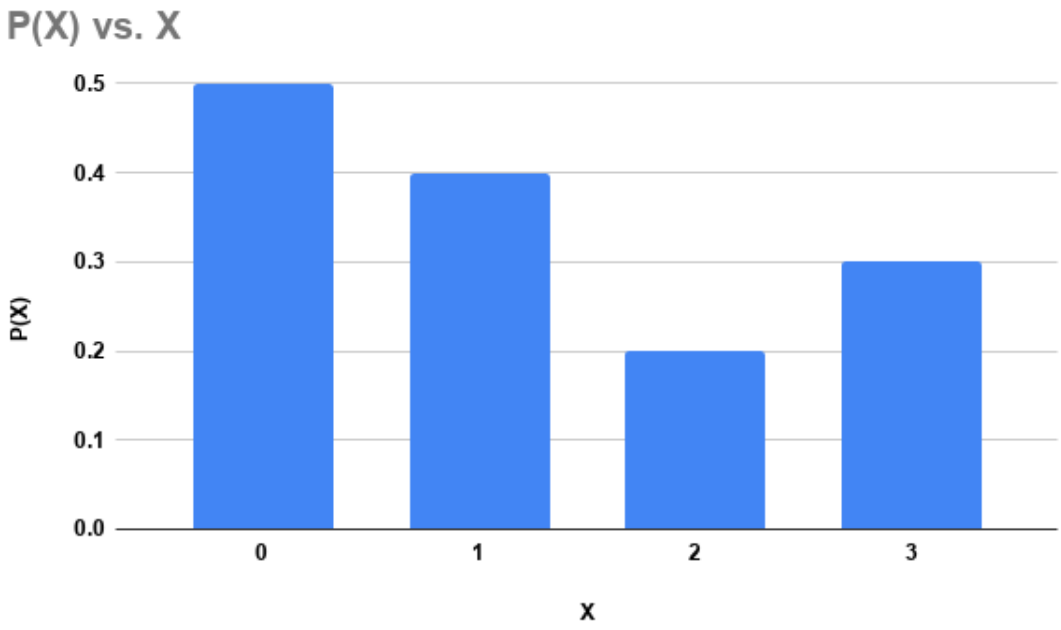

(d)

Answer:(a)

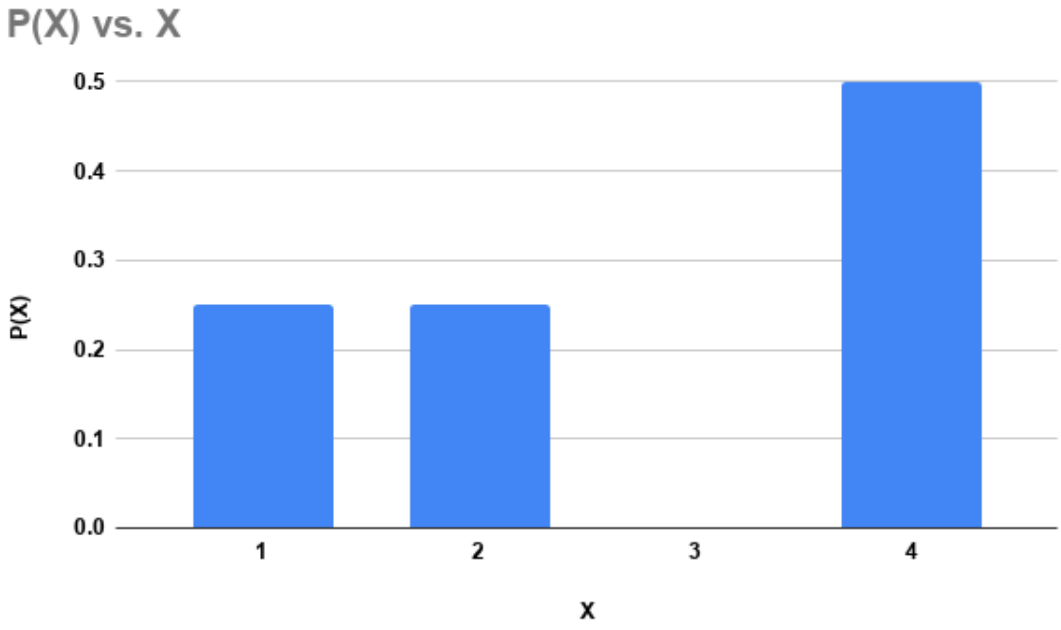
2. Which of the following is/are probability mass function for some random variable?



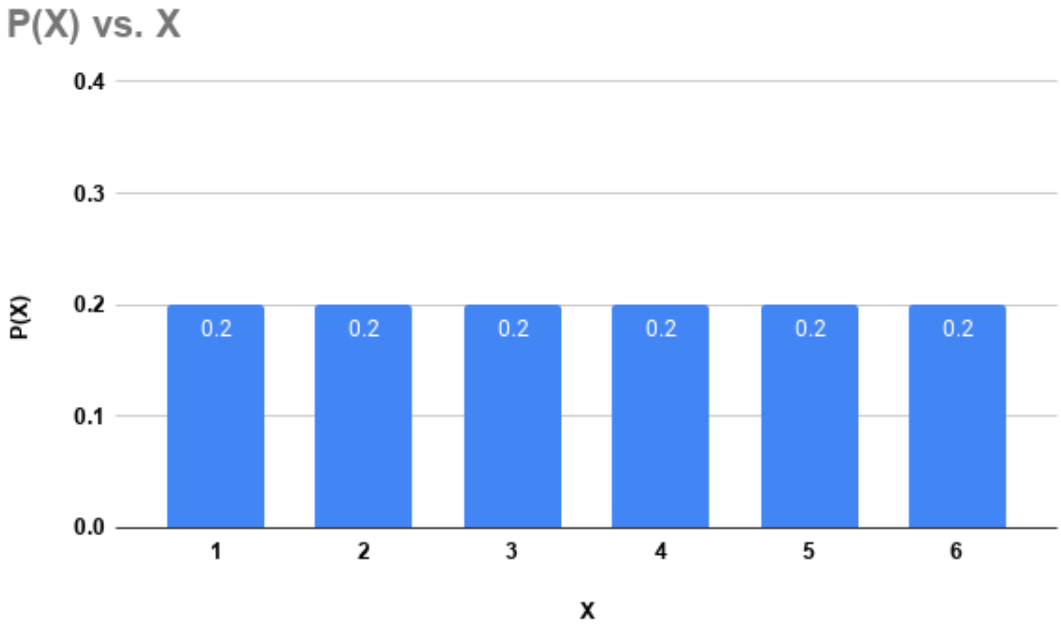
(a)




(b)




(c)




(d)

Answer:(a), (c)

3. An unfair coin is tossed thrice and the distribution of number of tails is given in Figure 8.6.1.

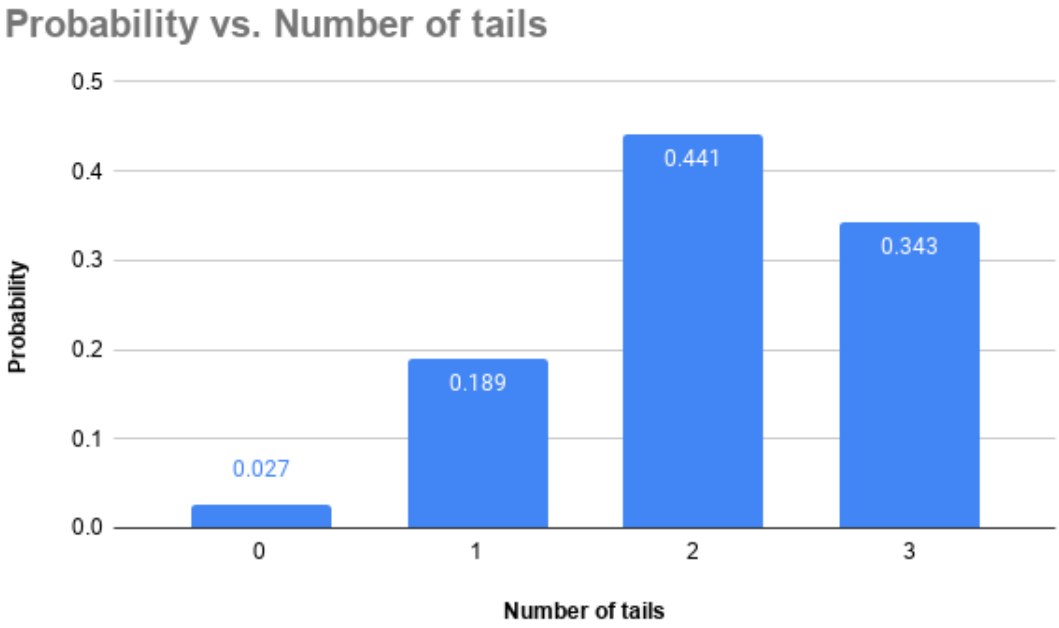



Figure 8.6.1: PMF of number of tails

If the same coin is tossed once then what is the probability that it will show head?
Answer: 0.3

## 6 Lecture-8.6

Discrete random variables - Cumulative distribution function

1. Which of the following statements is/are wrong about the cumulative distributive function $F$ of a discrete random variable $X$?
   (a) If $x_1 \leq x_2$ then $F(x_1) \leq F(x_2)$ where $x_1$, $x_2$ are any two points in $X$.
   (b) Maximum value of $F(x)$ is one.
   (c) It is always a step function for discrete random variable.
   (d) Area under the graph of CDF is always one.
2. Three cards are drawn in succession from a deck with replacement. Find the cumulative distribution function for the number of red cards.

(a)

$$F(x) = \begin{cases} \frac{1}{8} & \text{if } 0 \leq x < 1 \\ \frac{3}{8} & \text{if } 1 \leq x < 2 \\ \frac{3}{8} & \text{if } 2 \leq x < 3 \\ \frac{1}{8} & \text{if } x = 3 \end{cases}$$

(b)

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{1}{8} & \text{if } 0 \leq x < 1 \\ \frac{3}{8} & \text{if } 1 \leq x < 2 \\ \frac{3}{8} & \text{if } 2 \leq x < 3 \\ \frac{1}{8} & \text{if } x \geq 3 \end{cases}$$

(c)

$$F(x) = \begin{cases} \frac{1}{8} & \text{if } 0 \leq x < 1 \\ \frac{1}{2} & \text{if } 1 \leq x < 2 \\ \frac{7}{8} & \text{if } 2 \leq x < 3 \\ 1 & \text{if } x = 3 \end{cases}$$

(d)

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{1}{8} & \text{if } 0 \leq x < 1 \\ \frac{1}{2} & \text{if } 1 \leq x < 2 \\ \frac{7}{8} & \text{if } 2 \leq x < 3 \\ 1 & \text{if } x \geq 3 \end{cases}$$

Answer: (d)

3. Suppose it is known that $X$, the number of cars that arrive at a specific car parking lot during a 1-hour time period, is characterized by the following discrete probability function:

$$P(X = x) = \frac{e^{-5}5^x}{x!} \text{ for } x = 0, 1, 2, ...$$

Find the probability that in a specific 1- hour time period, more than 2 cars arrive at the car parking.

(a) $\frac{e^{-5}5^2}{2!}$

(b) $1 - \frac{e^{-5}5^2}{2!}$

(c) $\sum_{x=0}^{2} \frac{e^{-5}5^x}{x!}$

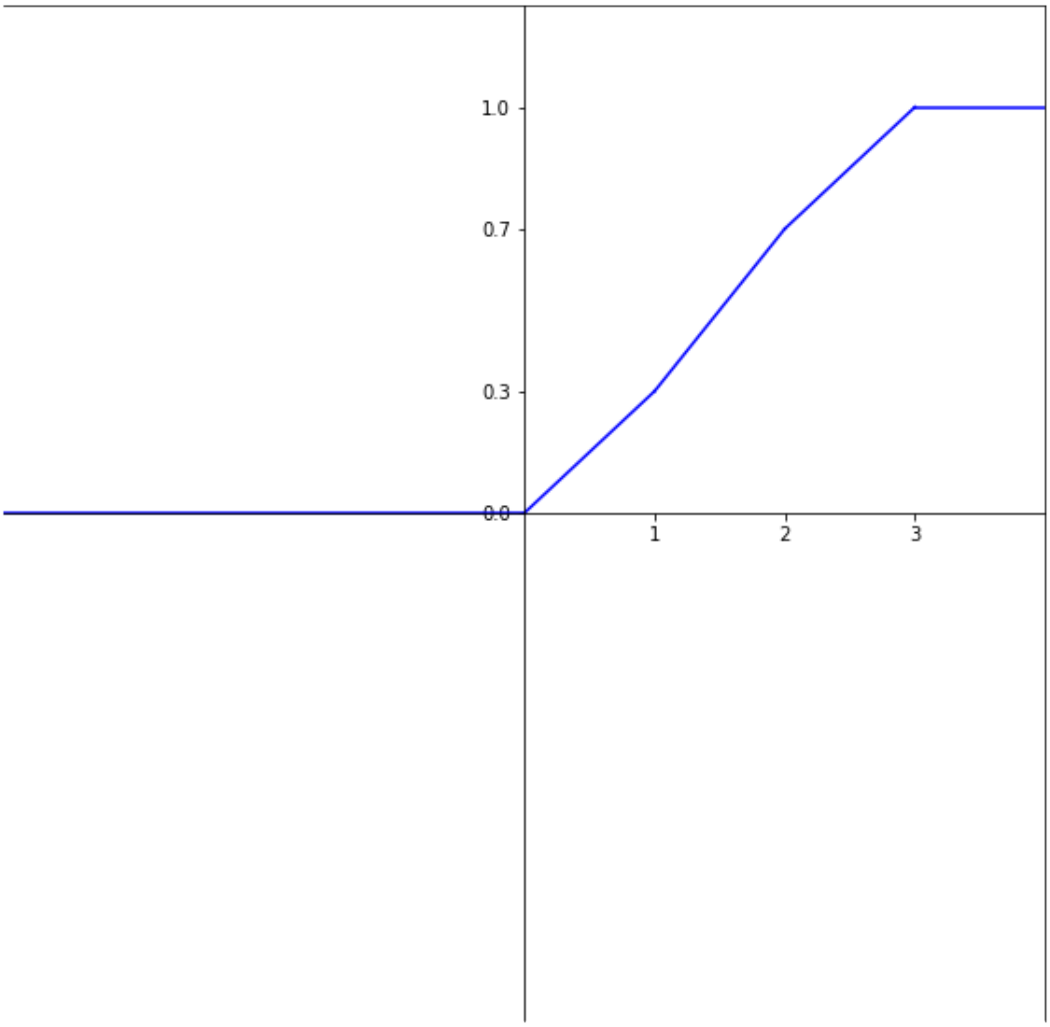
(d) $1 - \sum_{x=0}^{2} \frac{e^{-5}5^x}{x!}$

4. Probability mass function of a discrete random variable $X$ is given in Table 8.7.1.

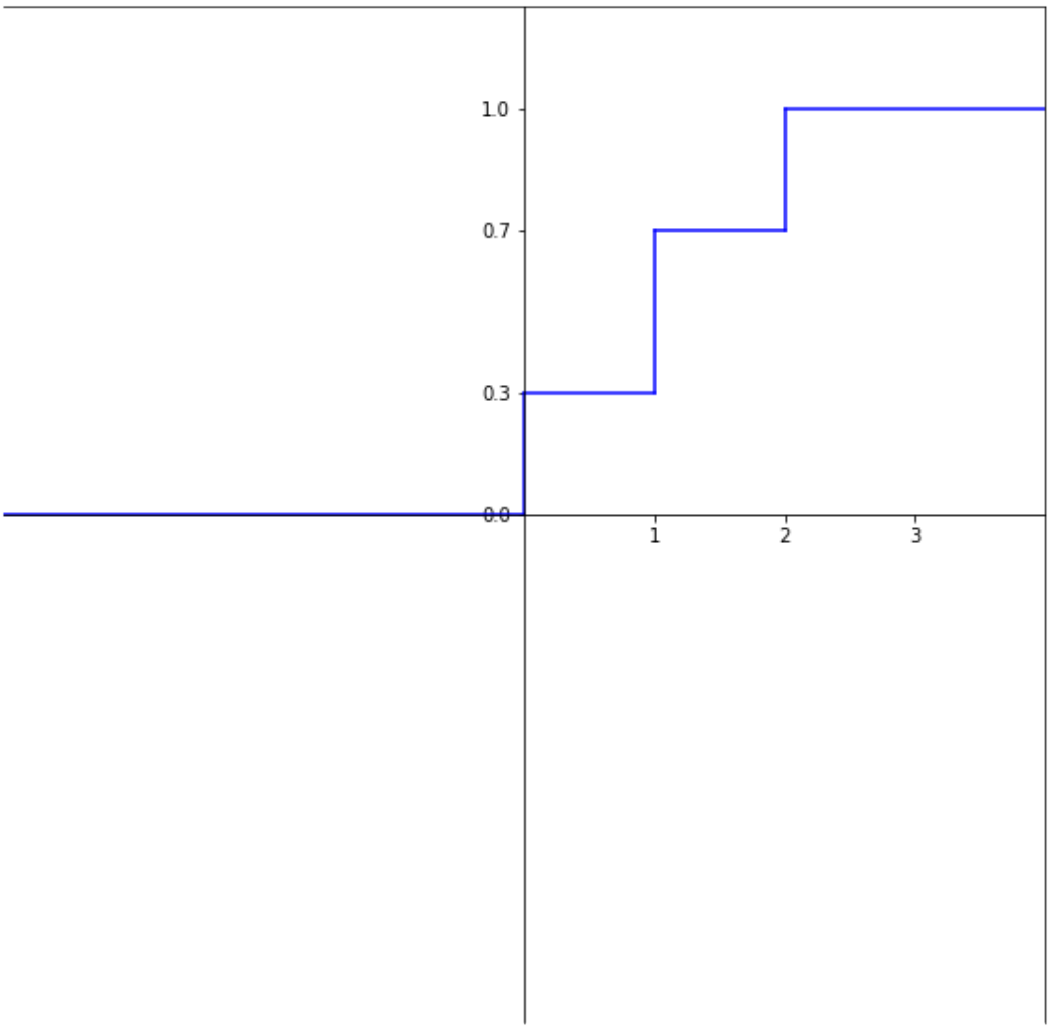
| $X$ | 1 | 2 | 3 |
|---|---|---|---|
| $P(X = x)$ | 0.3 | 0.4 | 0.3 |

Table 8.7.1: PDF of $X$

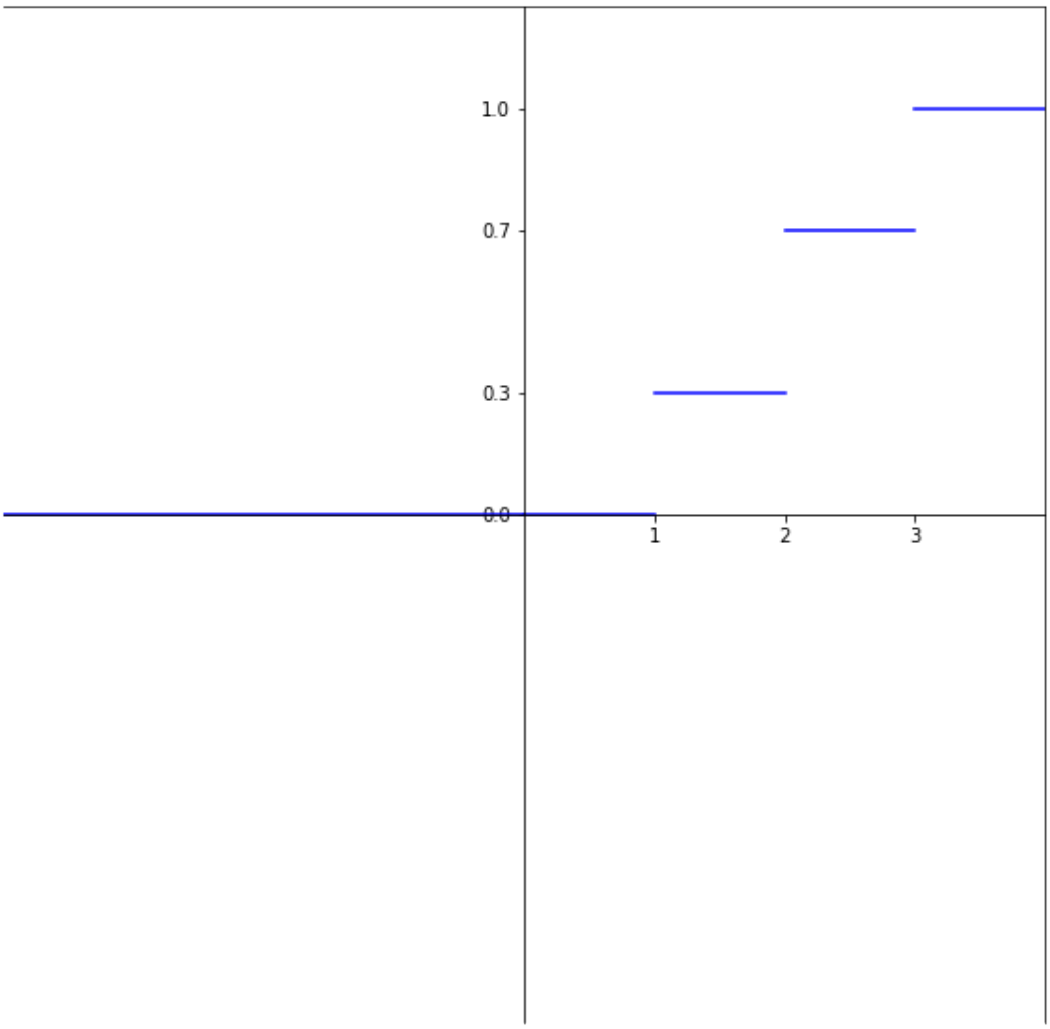
Which of the following graphs is the CDF of $X$?



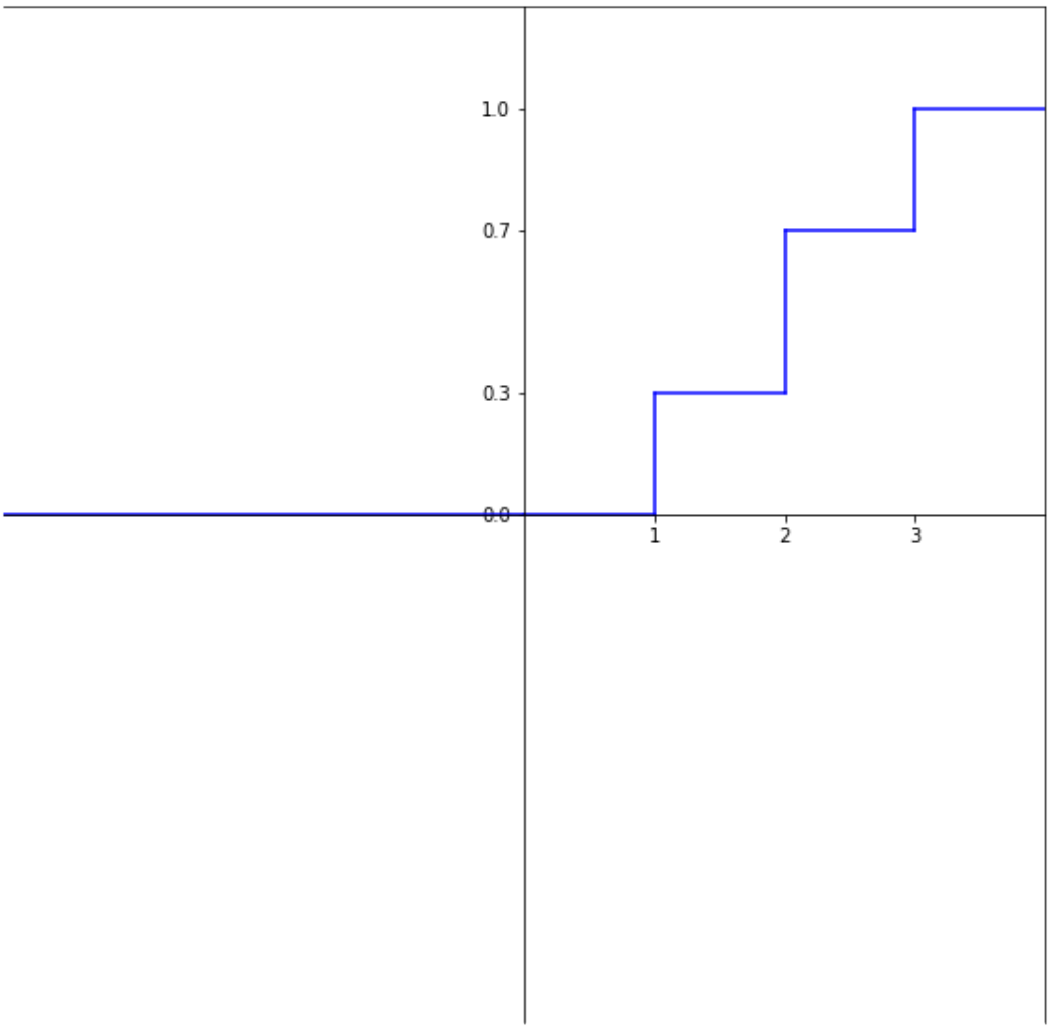

(a)




(b)




(c)




(d)

Answer: (c)

Use the following information to answer the questions (5) and (6).
Cumulative distributive function of the number of defective items in the shipment of every ten items is given as follows:

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ 0.41 & \text{if } 0 \leq x < 1 \\ 0.78 & \text{if } 1 \leq x < 2 \\ 0.91 & \text{if } 2 \leq x < 3 \\ 1 & \text{if } x \geq 3 \end{cases}$$

5. Find the probability that there are 3 defective items in the shipment of ten items.

(write your answer upto two decimal places)
answer: 0.09

6. Find the probability that there are at least 2 defective items in the shipment of ten items. (write your answer upto two decimal places)
answer: 0.22

BSCMA1002: Activity Questions
Week-9

# Contents

# 1 Lecture - 2

Expectation of a random variable

1. Relative frequency of random variable converges to the probability of random variable in experiment of tossing a die:
   (a) If number of observations are infinite.
   (b) If number of observations is finite.
   (c) Irrespective of number of observations.
2. Expectation of random variable is:
   (a) Long-run average value of random variable.
   (b) Short term average value of random variable.
   (c) Always equal to the average value of random variable irrespective of number of observations.
3. Let $X$ be a random variable which takes values $x_1, x_2, x_3, x_4, \cdots, x_{n-1}, x_n$. Expectation of the random variable $X$ is equal to:
   (a)
   $$E(X) = \sum_{i=1}^{i=n} x_i \times P(X = x_i)$$
   (b) If number of observations are infinite and $r(X = x_i)$ is the relative frequency of $X = x_i$,
   $$E(X) = \sum_{i=1}^{i=n} x_i \times r(X = x_i)$$
   (c) If number of observations are finite,
   $$E(X) = \sum_{i=1}^{i=n} x_i \times r(X = x_i)$$
   (d)
   $$E(X) = \sum_{i=1}^{i=n} x_i^2 \times P(X = x_i)$$

Answer: a, b

4. What is the expectation of the following probability mass function of a random variable X represented in Table 9.2.1.

| $X$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $P(X = x_i)$ | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |

Table 9.2.1: Probability mass function of random variable

Answer: 3

5. Expectation of a random variable is always one of the possible values for the random variable.

   (a) True

   (b) False

6. For a bowler in first class cricket, the number of wickets taken by a bowler in a match is the random variable. The outcomes for the random variable for 100 matches he played is given Table 9.2.2. What is the expectation of the random variable? Assume relative frequency of a random variable at $X = x$ is equal to the probability of random variable $P(X = x)$.

| $X$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| Frequency | 15 | 20 | 30 | 15 | 10 | 5 | 3 | 2 |

Table 9.2.2: Wickets taken by a bowler

Answer: 2.22

7. In an online football game, for each goal an online gamer scores he will earn 100 points. The probability mass function of number of goals is given in Table 9.2.3. What is the expected number of points online gamer will earn? Assume relative frequency of a random variable at $X = x$ is equal to the probability of random variable $P(X = x)$.

| X | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| Relative Frequency | 0.15 | 0.20 | 0.15 | 0.3 | 0.10 | 0.05 | 0.05 |

Table 9.2.3: Probability mass function of goals scored

   (a) 215

   (b) 197

(c) 235

(d) 265

8. If $P(X = 0) = 0.2$ in a Bernoulli random variable, then expectation of the Bernoulli random variable is:
Answer:0.8

9. If the expectation of uniform distribution whose probability mass function is given in Table 9.2.4 is 10. Then the value of $n$ is:

| $X$ | 1 | 2 | $\cdots$ | n |
|---|---|---|---|---|
| $P(X = x_i)$ | $\frac{1}{n}$ | $\frac{1}{n}$ | $\cdots$ | $\frac{1}{n}$ |

Table 9.2.4 Probability mass function

Answer: 19

# 2 Lecture 3

Expectation of a random variable: properties of expectation

1. Which of the following statements is/are true.

(a) $E(aX + b) = aE(X) + b$

(b) $E((aX + b)^2) = a^2E(X^2) + 2abE(X) + b^2$

(c) $E((aX - b)^2) = a^2E(X^2) + 2abE(X) + b^2$

(d) $E((aX - b)^2) = a^2E(X^2) - 2abE(X) + b^2$

2. The expected values of outcomes of 10 biased dice is given in Table 9.1.5. What is the expected values of sum of outcomes of the 10 dice?

| Die | Expected value |
|---|---|
| Die 1 | 3.5 |
| Die 2 | 2.5 |
| Die 3 | 1.5 |
| Die 4 | 2.2 |
| Die 5 | 3.7 |
| Die 6 | 4.0 |
| Die 7 | 2.0 |
| Die 8 | 3.0 |
| Die 9 | 5 |
| Die 10 | 2 |

Table 9.1.5: Expected values of outcomes of biased dice

Answer: 29.4

3. In a software office, there are 20 junior developers of which 8 are men and 12 are women. If 5 junior developers are to be promoted to senior level. What is the expected number of women in the promoted developers?
   Answer: 3 .

4. There are 1000 mobile phones in a mobile wholesale shop of which 100 are defective. A retailer buys 20 mobiles from wholesale shop for a total price of 2,00,000 rupees. The retailer gets 2,500 rupees profit if the mobile phone is not defective and 10,000 rupees loss if it is defective for each phone. What is the expected profit (in rupees) earned by the retailer? (Note: Enter the answer with prefix "-" if the net profit is negative. Example "-150".
   Answer: 25000

5. In a biased coin, the probability of getting tails is 0.3. If a coin is tossed 7 times, then what is the total expected number of heads?
   Answer: 4.9

# 3 Lecture - 4

Variance of a random variable

1. Which of the following random variables has the highest variance?

   (a)

   | $X$ | 1 | 2 | 3 | 4 | 5 |
   |---|---|---|---|---|---|
   | $P(X=x)$ | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |

   (b)

   | $X$ | 2 | 4 | 6 | 8 | 10 |
   |---|---|---|---|---|---|
   | $P(X=x)$ | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |

   (c)

   | $X$ | 1 | 2 | 3 | 4 |
   |---|---|---|---|---|
   | $P(X=x)$ | 0.25 | 0.25 | 0.25 | 0.25 |

   (d)

   | $X$ | 2 | 4 | 6 | 8 |
   |---|---|---|---|---|
   | $P(X=x)$ | 0.25 | 0.25 | 0.25 | 0.25 |

2. Variance of a random variable is equal to:

   (a) $E(X^2) + (E(X))^2$

   (b) $E(X^2) - (E(X))^2$

   (c) $E(X^2) - \mu^2$

   (d) $E(X^2) + \mu^2$

3. Which of the following statements is/are true?

   (a) Variance of a random variable can be negative.

   (b) If the variance of a random variable is zero, then there is one/zero possible values for the random variable.

   (c) As the size($n$) of the uniform random distribution increases variance increases.

   (d) Variance of a Bernoulli random variable is maximum when $p = 0.5$.

4. If in a Bernoulli random variable, whose possible values are 0 and 1, $P(X = 1) = 0.5$, then what is the variance of the random variable $X$? Enter the answer up to 2 decimals accuracy.
   Answer: 0.25

5. What is the variance of random variable whose probability distribution is in Table 9.4.1? Enter the answer up to 2 decimals accuracy.

| $X = x$ | 1 | 2 | 3 | 4 | $\cdots$ | 20 |
|---|---|---|---|---|---|---|
| $P(X = x)$ | $\frac{1}{20}$ | $\frac{1}{20}$ | $\frac{1}{20}$ | $\frac{1}{20}$ | $\cdots$ | $\frac{1}{20}$ |

Table 9.4.1 Probability distribution of random variable $X$

Answer: 33.25

# 4 Lecture - 5

Variance of a random variable: properties of variance

1. $V(a^2X + b^2) =$ (where $a$ and $b$ are constants and $X$ is a random variable.)
   (a) $a^2V(X)$
   (b) $a^2V(X) + b^2$
   (c) $a^4V(X)$
   (d) $a^4V(X) + b^4$

2. If the variance of marks obtained by students in Statistics (Var(S)) is 24 and the variance of marks obtained by students in Mathematics (Var(M)) is 26, then what is the value of variance of sum of marks in Statistics and Mathematics Var(S+M)?
   Answer: 50

3. The variance of number of tickets sold in a day in an opera theatre is 25. If for each ticket sold the profit obtained is 25 rupees, then what is the variance of the profit obtained in a day in rupee square?
   Answer: 15625

4. Which of the following statements is/are true?
   (a) If $X$ and $Y$ are independent random variables, then $\text{Var}(X + Y)=\text{Var}(X)+\text{Var}(Y)$.
   (b) If $X$ and $Y$ are dependent random variables, then $\text{Var}(X + Y)=\text{Var}(X)+\text{Var}(Y)$
   (c) If $X_1, X_2, X_3, \cdots, X_n$ are dependent random variable, then $\text{Var}(X_1 + X_2 + X_3 + \cdots + X_n) =\text{Var}(X_1)+\text{Var}(X_2)+\text{Var}(X_3)+\cdots+\text{Var}(X_n)$
   (d) If $X_1, X_2, X_3, \cdots, X_n$ are independent random variables.
   then $\text{Var}(X_1 + X_2 + X_3 + \cdots + X_n) =\text{Var}(X_1)+\text{Var}(X_2)+\text{Var}(X_3)+\cdots+\text{Var}(X_n)$

# 5 Lecture - 6

Standard deviation of a random variable

1. Which of the following statements is/are true?
   (a) SD($X$)=$\sqrt{E(X^2)+\mu^2}$
   (b) SD($X$)=$\sqrt{\text{Var}(X)}$
   (c) Standard deviation of a random variable has the same units as the expectation of a random variable.
   (d) Standard deviation of a random variable has the same units as the units of a random variable.

2. If $X$ and $Y$ are independent random variables and Var($3X$+3)=27, Var($5Y$+10)=150, then what is the value of SD($X+Y$)?
   Answer: 3

3. Among the random variables $X$, $Y$, and $Z$, $Y=2X$ and $X$ and $Z$ are independent. If Var($X$)=2, Var($Z$)=7, then what is the value of SD($X+Y+Z$)?
   Answer: 5

BSCMA1002 : Activity Questions
Week-10

# Contents

# 1 Lecture 10.1

1. Consider the experiment of drawing a card from a well-shuffled deck of 52 cards. If success is defined as getting a red card. Identify the failure.

   (a) Selecting the card of suit heart.

   (b) Selecting the non-red cards.

   (c) Selecting the card of suit club.

   (d) Selecting the face cards.

2. Choose the Bernoulli trial from the following:

   (a) Experiment: Randomly selecting students from the class who have secured more than 90 marks out of 100 in the class test.

   (b) Experiment: Randomly selecting a student from the class and asking their date of birth.

   (c) Experiment: Randomly selecting a product from a manufacturing unit and identifying whether it is defective or not.

3. If $X$ is a Bernoulli random variable with parameter $p = 0.4$, calculate $E(X)$ and $Var(X)$.
   0.4, 0.24

4. Let the pmf of a random variable $X$ be given by:

$$P(X = x) = \begin{cases} 0.1 & x = 1 \\ 0.2 & x = 2 \\ 0.2 & x = 3 \\ 0.3 & x = 4 \\ 0.2 & x = 5 \end{cases}$$

   Now define a Bernoulli random variable $Y$ as

$$Y = \begin{cases} 1 & X \leq a \\ 0 & X > a \end{cases}$$

   where $a$ belongs to the positive integer. For what value of $a$, a Bernoulli trial of $Y$ be considered as the most uncertain Bernoulli trial?
   3

## 2 Lecture 10.2

1. Let $X \sim Binomial(n, p)$, $X$ has the following cdf:

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1024}{3125} & 0 \leq x < 1 \\ \frac{2304}{3125} & 1 \leq x < 2 \\ \frac{2944}{3125} & 2 \leq x < 3 \\ \frac{3104}{3125} & 3 \leq x < 4 \\ \frac{3124}{3125} & 4 \leq x < 5 \\ 1 & x \geq 5 \end{cases}$$

   (a) Find the value of $n$.
   5

   (b) Find the value of $p$ and $q$.
   1/5, 4/5

   (c) Find the value of $P(X \geq 3)$.(Answer upto 3 decimal accuracy)
   0.058 (Accepted range[0.056, 0.060])

2. For 4 independent Bernoulli trials, if $p$ and $q$ represent the probability of success and failure respectively for each trial, find $P(ssfs)$, where $s$ represents the success and $f$ represents the failure.

   (a) $p^2qp$

   (b) $p^3p$

   (c) $p^3q$

   (d) $p^4$

# 3 Lecture 10.3

1. Choose the correct statements from the following for the binomial random variable?

   (a) It is used to model the number of independent Bernoulli trials needed to obtain the first success.

   (b) If $X \sim Binomial(n, p)$, then $X = X_1 + X_2 + \cdots + X_n$, where each $X_i$ is a Bernoulli random variable and are independent, $p$ is the probability of success

   (c) It models the number of successes in $n$ independent Bernoulli trials.

   (d) $E(X) = p, Var(X) = p(1-p)$, where $p$= probability of success, $n \neq 1$.

2. Choose the correct statements from the following:

   (a) If $X \sim Binomial(6, 0.3)$, the graph of the distribution is left skewed.

   (b) If $X \sim Binomial(6, 0.3)$, the graph of the distribution is right skewed.

   (c) If $X \sim Binomial(60, 0.3)$, the distribution will tend to be symmetric.

   (d) If $X \sim Binomial(60, 0.3)$, the distribution will fail to be symmetric.

3. A company is planning on expanding its reach. It is planning to open 4 more branches in the city and is considering 4 different locations for it. The probability of successful opening of a branch at a location is 0.65 and the success of branches are independent of each other.

   (a) Find the probability that exactly 2 branches will be successful.(Enter the anwer upto 3 decimal accuracy.)
   0.310 (Accepted range[0.308, 0.312])

   (b) Identify the shape of the distribution.

      i. The graph is left skewed.
      ii. The graph will be symmetric.
      iii. The graph is right skewed.
      iv. The graph is left skewed and tend to be symmetric.

4. Suppose that 4.5 percent of people living in a particular city own a Creta. A survey has been conducted to know what brand of cars people use in the city. What is the probability that the 12th person being interviewed during the survey is the 2nd one to own a creta?(Enter the answer upto 3 decimal places.)
   0.014 (Accepted range[0.012, 0.016])

# 4 Lecture 10.4

1. Consider a string of length 4 (a sequence of 4 bits) where each bit can be either 0 or 1. Suppose you want to transmit 4 bits wirelessly via a transmission channel. The channel is noisy, i.e., there is a possibility of error. Assume that changing a bit at the $ith$ place is independent from changing the bit at the $jth$ place, $i \neq j$. The probability of 1 bit getting changed is given to be 0.1. Let the random variable $X$ denote the number of bits that changed.

   (a) $X \sim Binomial(n, p)$, find the value of $n$ and $p$.
   4, 0.1

   (b) Choose the correct probability distribution of $X$ from the following.

   i. $P(X = x) = {}^4C_x(0.1)^x(0.9)^{4-x}$
   ii. $P(X = x) = {}^4C_x(0.9)^x(0.1)^{4-x}$
   iii. $P(X = x) = {}^4C_x(0.9)^x(0.1)^{x-4}$
   iv. $P(X = x) = {}^4C_{4-x}(0.9)^{4-x}(0.1)^x$

   (c) What is the probability that the translated message is wrong (Assuming the message is wrong when there is at least 1 bit error)? (Answer upto 4 decimal places)
   0.3439

   (d) An error can be recovered if the number of faulty bits is less than or equal to 1. If we send a string, find the probability that the string is either correct or can be recovered. (Answer upto 4 decimal places).
   0.9477

2. Suppose 5 fair dice are thrown independently. $X$ represents the number of dice where the multiple of 4 appears.

   (a) Find the pmf of $X$.

   i. $P(X = x) = {}^4C_x(\frac{1}{6})^x(\frac{5}{6})^{4-x}$
   ii. $P(X = x) = {}^5C_x(\frac{1}{4})^x(\frac{3}{4})^{x-5}$
   iii. $P(X = x) = {}^5C_x(\frac{1}{2})^x(\frac{1}{2})^{5-x}$
   iv. $P(X = x) = {}^5C_x(\frac{1}{6})^x(\frac{5}{6})^{5-x}$

   (b) Find the probability that at least three of the dice show a multiple of 4. (Enter the answer up to 3 decimal accuracy).
   0.035 (Accepted range:[0.033, 0.037])

3. A person has bought a table from an online furniture store. The seller delivers the disassembled table parts along with some screws to assemble it. The probability of a screw being defective is $\frac{1}{12}$. To compensate for the manufacturing error, the seller sends two extra screws in the package where the table needs exactly 10 screws to assemble.

What is the probability that the buyer will be able to assemble the table? (Enter the answer up to 3 decimal accuracy).
0.928 (Accepted Range [0.926, 0.930])

4. For a mission, helicopter gunships are sent to destroy a critical target. The probability that a gunship will hit the target is 0.65. If at least 5 gunships are needed to destroy the target and a battalion has sent 8 such gunships, then find the probability that the target will be destroyed. (Enter the answer up to 3 decimal accuracy).
0.706 (Accepted Range [0.704, 0.708])

5. Suppose the probability of students going to the school in their locality is 0.04. If 10 kids from a society are picked at random and if going to school in their locality are independent for each kid, what is the probability that 2 among them are going to the school in the locality?(Enter the answer upto 3 decimal accuracy)
0.052 (Accepted range [0.050, 0.054])

# 5 Lecture 10.5

1. Let $X$ be a Binomial random variable with a mean of 5.4 and $n = 9$, calculate the variance.
   2.16

2. Given $X$ is a Binomial random variable such that $E(X) = 6$ and $V(X) = 3.6$. Answer the following questions (a) and (b).

   (a) If $X \sim Binomial(n, p)$, then find the values of $n$ and $p$.

      i. $n = 10, p = 0.2$
      ii. $n = 15, p = 0.4$
      iii. $n = 10, p = 0.4$
      iv. $n = 15, p = 0.2$

   (b) Find $P(X = 3)$. (Enter the answer up to 3 decimal accuracy).
   0.063 (Accepted range[0.061, 0.065])

3. Suppose a fair die is thrown 90 times. The expected number of times a 4 occurs is 15. What is the significance of the expectation?

   (a) If we throw the die for 90 times, for exactly 15 of the outcomes we will get 4.

   (b) If we throw the die for 90 times, we can expect 15 of the outcomes to be 4.

4. Suppose an urn contains marble of 2 different colors, red and yellow. A marble is drawn randomly from the box and its color has been recorded. Then the marble is kept in the box and remixed before the next draw. Suppose the proportion of yellow and red marble in the box is $\pi_1$ and $\pi_2$. Let the urn has 15 marbles out of which 6 are red and a total of 10 marbles is drawn from the urn in the stated manner. Based on the given information, answer the following 2 questions:

   (a) Calculate the value of $\pi_1$ and $\pi_2$.

      i. $\pi_1 = 0.2$ and $\pi_2 = 0.33$
      ii. $\pi_1 = 0.6$ and $\pi_2 = 0.4$
      iii. $\pi_1 = 0.6$ and $\pi_2 = 0.2$
      iv. $\pi_1 = 0.33$ and $\pi_2 = 0.6$

   (b) In the experiment, find the standard deviation of number of yellow balls drawn. (Enter the answer up to 3 decimal accuracy).
   1.549 (Accepted range: [1.547, 1.551])