Statistics week 1 summary

Descriptive statistics - The part of statistics concerned with the description and summarization of data is called descriptive statistics.

Inferential Statistics - The part of statistics concerned with the drawing of conclusions from data is called inferential statistics.

The concept of Population and Sample

- The percentage of all students in India who have passed their Class 12 exams and study engineering.
- The prices of all houses in Tamil Nadu.
- The total sales of all cars in India in the year 2019.
- The age distribution of people who visit a city Mall in a particular month.

Definition –

Population - The total collection of all the elements that we are interested in is called a population.

Sample - A subgroup of the population that will be studied in detail is called a sample.

Data - Data are the facts and figures collected, analyzed, and summarized for presentation and interpretation.

Why do we collect data –

- Interested in the characteristics of some group or groups of • people, places, things, or events.
- Example: To know about temperatures in a particular month
- in Chennai, India.
- Example: To know about the marks obtained by students in
- their Class 12.
- To know how many people like a new song/product/video- collected through comments.

Unstructured data and structured data

- For the information in a database to be useful, we must know
- the context of the numbers and text it holds.
- When they are scattered about with no structure, the information is of very little use.

- Hence, we need to organize data.

- Case (observation): A unit from which data are collected.
- Variable:
- Intuitive: A variable is that \varies".
- Formally: A characteristic or attribute that varies across all • units.
- In our school data set:
- Case: each student
- Variable: Name, marks obtained, Board etc.
- Rows represent cases: for each case, same attribute is recorded • Columns represent variables: For each variables, same type of
- value for each case is recorded.

Classification of data –

- Categorical
- Numerical

- Categorical data
- Also called qualitative variables.
- Identify group membership
- Numerical data
- Also called quantitative variables.
- Describe numerical properties of cases
- Have measurement units
  Measurement units: Scale that de_nes the meaning of • numerical data, such as weights measured in kilograms, prices
- in rupees, heights in centimeters, etc.
- I The data that make up a numerical variable in a data table
- must share a common unit.

Time series - data recorded over time

- Timeplot { graph of a time series showing values in chronological order.

- Cross-sectional - data observed at the same time.

Scales of measurement

Nominal scale - When the data for a variable consist of labels or names used to identify the characteristic of an observation, the scale of measurement is considered a nominal scale. Examples: Name, Board, Gender, Blood group etc.

- There is no ordering in the variable.
- Nominal – Name categories without variables.

Ordinal scale - Data exhibits properties of nominal data and the order or rank of data is meaningful, the scale of measurement is considered a ordinal scale.

- Each customer who visits a restaurant provides a service rating of excellent, good, or poor.

Ordinal – Name categories that can be ordered.

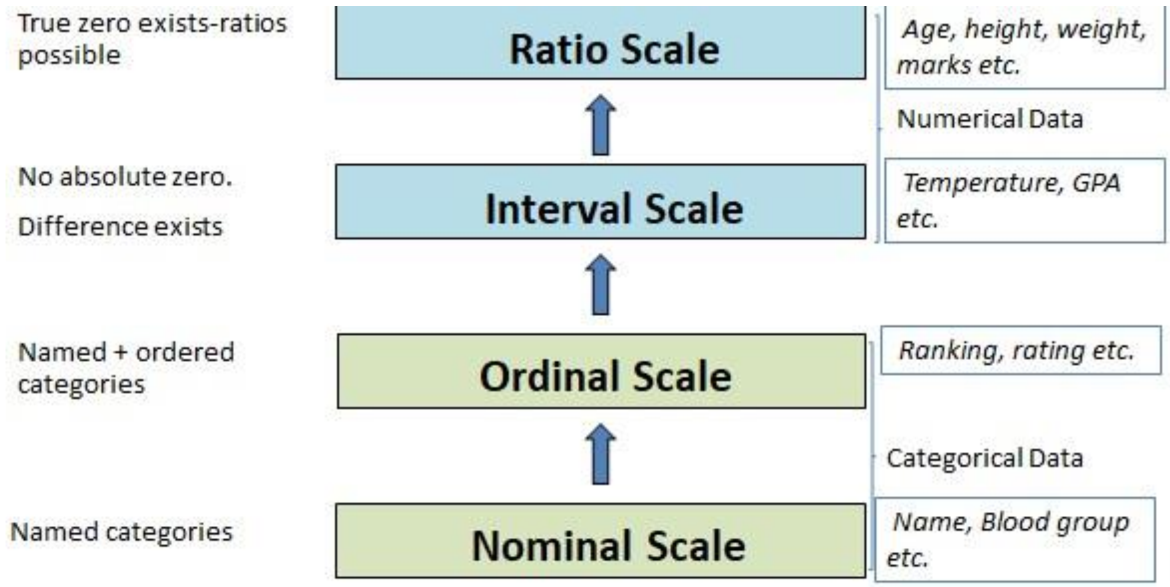
Interval scale - If the data have all the properties of ordinal data and the interval between values is expressed in terms of a fixed unit of measure, then the scale of measurement is interval scale.

- Interval data are always numeric. Can _nd out di_erence between any two values.
- Ratios of values have no meaning here because the value of zero is arbitrary.
- Interval: numerical values that can be added/subtracted (no absolute zero)

Ratio Scale - If the data have all the properties of interval data and the ratio of two values is meaningful, then the scale of measurement is ratio scale.

- Example: height, weight, age, marks, etc.

Ratio: numerical values that can be added, subtracted, multiplied or divided (makes ratio comparisons possible)



Statistics week 2

Frequency Distributions - A frequency distribution of qualitative data is a listing of the distinct values and their frequencies.

Each row of a frequency table lists a category along with the number of cases in this category.

Ex –

Construct a frequency table for the given data

1. A,A,B,C,A,D,A,B,D,C
2. A,A,B,C,A,D,A,B,D,C,A,B,C,D,A
3. A,A,B,C,A,A,B,B,D,C,A,B,C,D,B
4. A, A, B, C,A ,D, A,B,D,C, A,B,C,D,A,C,D,D

The steps to construct a frequency distribution2

- Step 1 List the distinct values of the observations in the data set in the first column of a table.
- Step 2 For each observation, place a tally mark in the second column of the table in the row of the appropriate distinct value.
- Step 3 Count the tallies for each distinct value and record the totals in the third column of the table.

Relative Frequency - The ratio of the frequency to the total number of observations is called relative frequency.

The steps to construct a relative frequency distribution

- Step 1 Obtain a frequency distribution of the data.
- Step 2 Divide each frequency by the total number of observations.

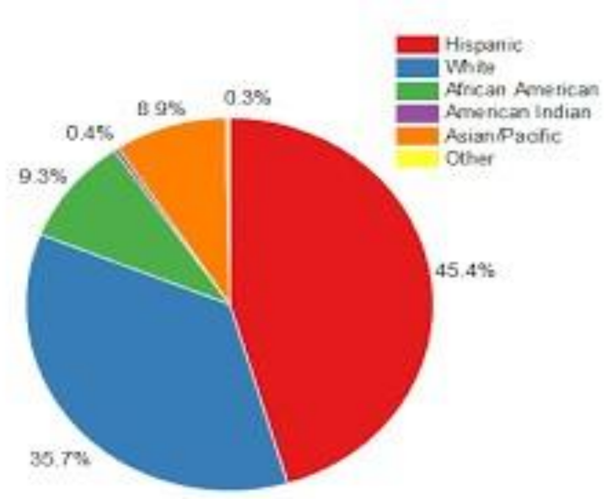
Why relative frequency?

l For comparing two data sets.
l Because relative frequencies always fall between 0 and 1, they provide a standard for comparison.

Pie chart - A pie chart is a circle divided into pieces proportional to the relative frequencies of the qualitative data.

The steps to construct a pie-chart

- Step 1 Obtain a relative-frequency distribution of the data.
- Step 2 Divide a circle into pieces proportional to the relative • frequencies.
- Step 3 Label the slices with the distinct values and their relative frequencies.



1. A pie chart is used to show the proportions of a categorical variable.

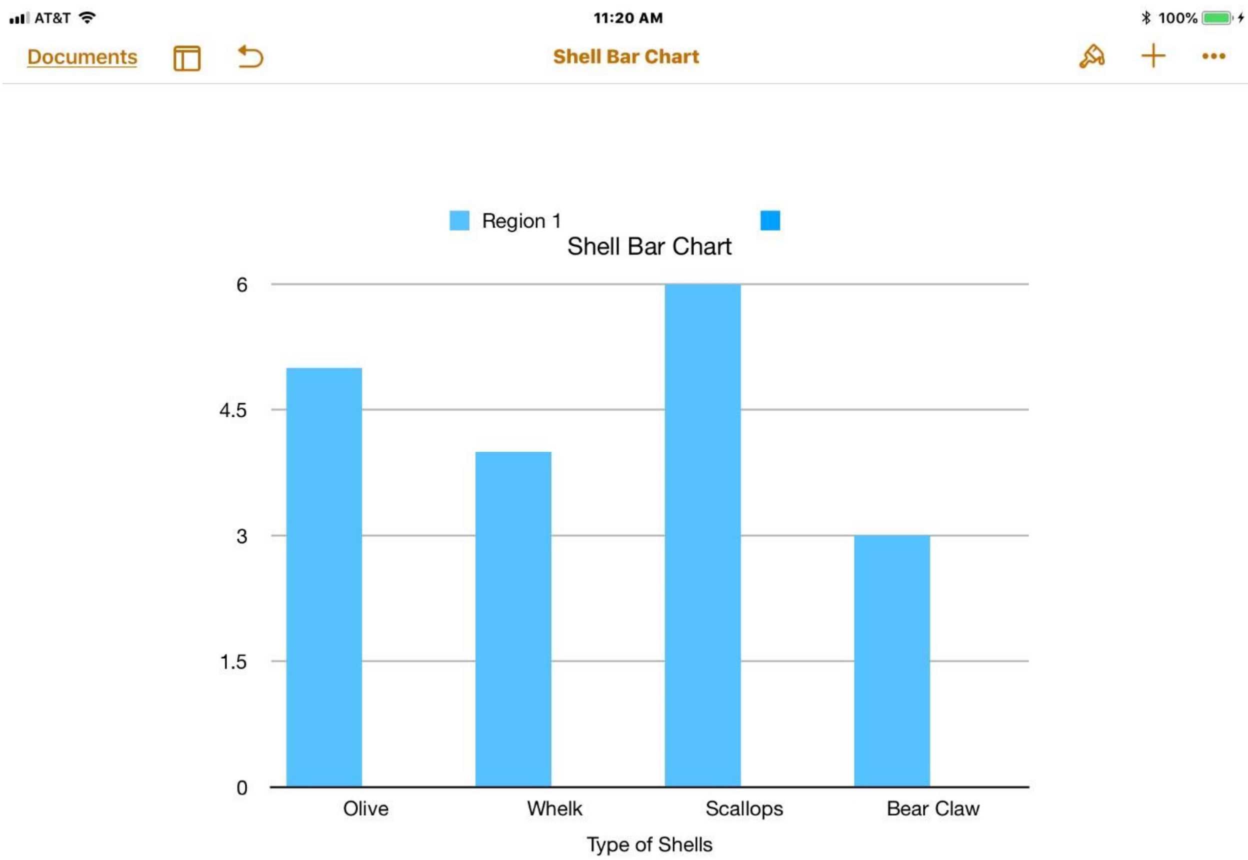
2. A pie chart is a good way to show that one category makes up more than half of the total.

Bar chart - A bar chart displays the distinct values of the qualitative data on a horizontal axis and the relative frequencies (or frequencies or percents) of those values on a vertical axis. The frequency/relative frequency of each distinct value is represented by a vertical bar whose height is equal to the frequency/relative frequency of that value. The bars should be positioned so that they do not touch each other.

Steps to construct a bar chart

To Construct a Bar Chart

- Step 1 Obtain a frequency/relative-frequency distribution of the data.
- Step 2 Draw a horizontal axis on which to place the bars and a vertical axis on which to display the frequencies/relative frequencies.
- Step 3 For each distinct value, construct a vertical bar whose height equals the frequency/relative frequency of that value.
- Step 4 Label the bars with the distinct values, the horizontal axis with the name of the variable, and the vertical axis with \Frequency" /\Relative frequency."



Pareto chart - When the categories in a bar chart are sorted by frequency, the bar chart is sometimes called a Pareto chart. Pareto charts are popular in quality control to identify problems in a business process.

If the categorical variable is ordinal, then the bar chart must preserve the ordering.

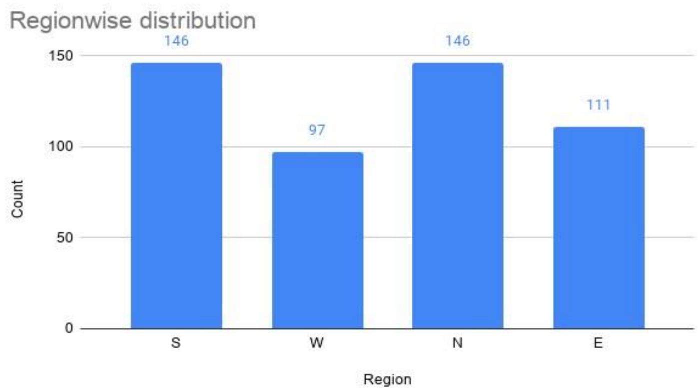
1. A bar chart is used to show the frequencies/relative frequencies of a categorical variable.
2. If ordinal, the order of categories is preserved.
3. The bars can be oriented either horizontally or vertically.
4. A Pareto chart is a bar chart where the categories are sorted by frequency.

The area principle

- Displays of data must obey a fundamental rule called the area principle.
- The area principle says that the area occupied by a part of the graph should correspond to the amount of data it represents.
- Violations of the area principle are a common way to mislead with statistics.

Misleading graphs: violating area principle

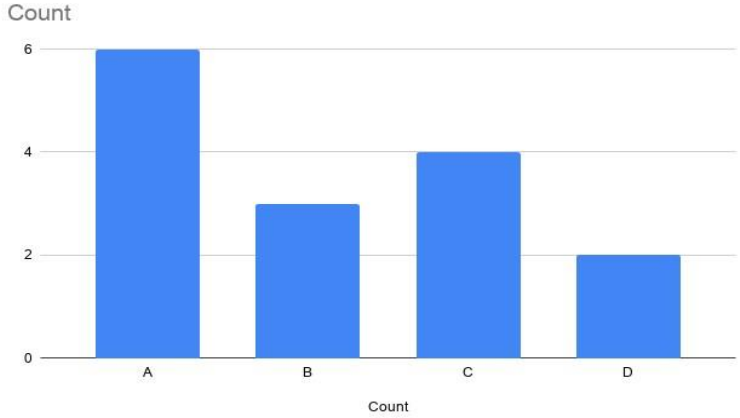
Decorated graphics: Charts decorated to attract attention
often violate the area principle.
No baseline and the chart shows bottles on top of labeled boxes of various sizes and shapes. I Obeys area principle and accurate.



- Another common violation is when the baseline of a bar chart is not at zero.
- Left graph exaggerates the number coming from the South and North. Graph on right shows same data with the baseline at zero.

Mode - The mode of a categorical variable is the most common category, the category with the highest frequency. The mode labels

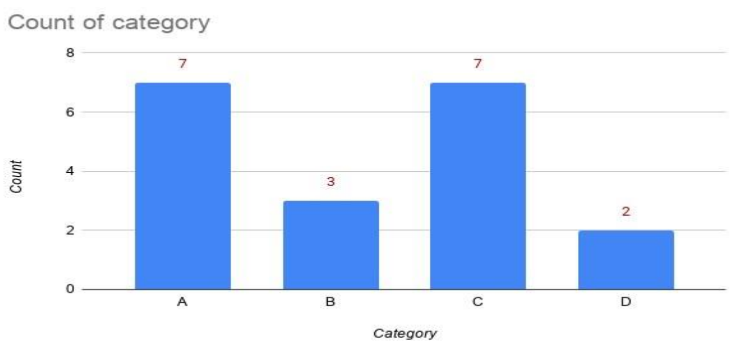
- The longest bar in a bar chart
- The widest slice in a pie chart.
- In a Pareto chart, the mode is the first category shown.



- Let consider the example A,A,B,C,A,D,A,B,C,C, A,B,C,D,A
- The longest bar in a bar chart
- The most common category is "A"

Bimodal and multimodal data

- If two or more categories tie for the highest frequency, the • data are said to be bimodal (in the case of two) or multimodal (more than two).
- Let consider the example A,A,B,C,A,C,A,B,C,C, A,C,C,D,A,A,C,D,B
- Both category "A" and "C" have highest frequency.



- Median - The median of an ordinal variable is the category of the middle
- observation of the sorted values.

- Consider the grades of 15 students which is listed as A,B,B,C,A,D,B,B,A,C, B,B,C,D,A
- The ordered data is A,A,A,A,B,B,B,B,B,B,C,C,C,D,D

- The median grade is the category associated with the 8 observation which is "B".
- Consider the grades of 14 students which is listed as
- A,B,B,C,A,D,B,B,A,C, B,B,C,D
- The ordered data is A,A,A,B,B,B,B,B,B,C,C,C,D,D
- The median grade is the category associated with the 7 or 8 observation which is "B".

Statistics Week 3

Types of variables-

1) Categorical
2) Numerical
   I) Discrete
   II) Continuous

Organizing numerical data

- Recall, a discrete variable usually involves a count of something, whereas a continuous variable usually involves a measurement of something.
- First group the observations into classes (also known as categories or bins) and then treat the classes as the distinct values of qualitative data.
- Once we group the quantitative data into classes, we can construct frequency and relativefrequency distributions of the data in exactly the same way as we did for categorical data.

Organizing discrete data (single value)

- If the data set contains only a relatively small number of distinct, or di_erent, values, it is convenient to represent it in a frequency table.
- Each class represents a distinct value (single value) along with its frequency of occurrence.

Example

- Suppose the dataset reports the number of people in a household. The following data is the response from 15
- individuals.
- 2,1,3,4,5,2,3,3,3,4,4,1,2,3,4

- The distinct values the variable, number of people in each household, takes is 1,2,3,4,5.

Organizing continuous data

Organize the data into a number of classes to make the data understandable. However, there are few guidelines that need to be followed. They are

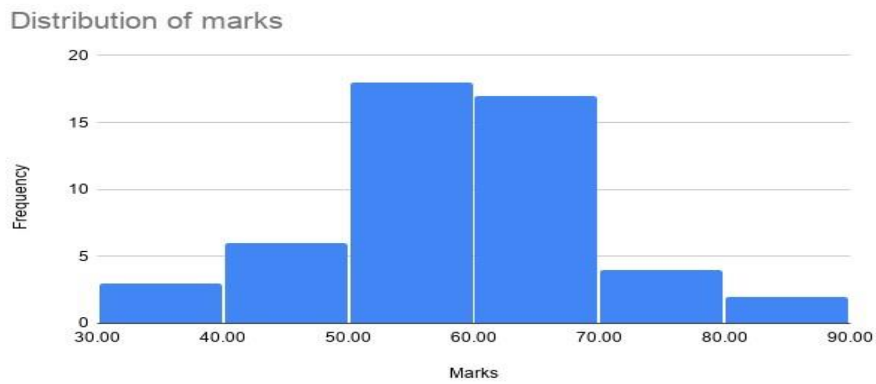
1. Number of classes: The appropriate number is a subjective choice, the rule of thumb is to have between 5 and 20 classes.
2. Each observation should belong to some class and no observation should belong to more than one class.
3. It is common, although not essential, to choose class intervals of equal length.

Some new terms

1. Lower class limit: The smallest value that could go in a class.
2. Upper class limit: The largest value that could go in a class.
3. Class width: The di_erence between the lower limit of a class and the lower limit of the next-higher class.
4. Class mark: The average of the two class limits of a class. 5. A class interval contains its left-end but not its right-end boundary point.

Steps to construct a histogram

- Step 1 Obtain a frequency (relative-frequency) distribution of the data.
- Step 2 Draw a horizontal axis on which to place the classes and a vertical axis on which to display the frequencies (relative frequencies).
- Step 3 For each class, construct a vertical bar whose height equals the frequency (relative frequency) of that class.
- Step 4 Label the bars with the classes, the horizontal axis with the name of the variable, and the vertical axis with \Frequency" (\Relative frequency").



Stem-and-leaf diagram

Definition

- In a stem-and-leaf diagram (or stemplot)1, each observation is • separated into two parts, namely, a stem-consisting of all but the
- rightmost digit-and a leaf, the rightmost digit.
- For example, if the data are all two-digit numbers, then we could let the stem of a data value be the tens digit and the leaf be the ones digit.
- The value 75 is expressed as
  Stem Leaf
  7 | 5
- The two values 75, 78 is expressed as Stem Leaf 7 | 5,8.

Steps to construct stem plot

Step 1 Think of each observation as a stem|consisting of all but the rightmost digit|and a leaf, the rightmost digit.
Step 2 Write the stems from smallest to largest in a vertical column to the left of a vertical rule.
Step 3 Write each leaf to the right of the vertical rule in the row that contains the appropriate stem.
Step 4 Arrange the leaves in each row in ascending order.

Descriptive measures –

- The objective is to develop measures that can be used to summarize a data set.
- These descriptive measures are quantities whose values are determined by the data.

Measures of central tendency: These are measures that indicate the most typical value or center of a data set.

Measures of dispersion: These measures indicate the variability or spread of a dataset.

The mean - The mean of a data set is the sum of the observations divided by the number of observations.

The mean is usually referred to as average.

- Arithmetic average; divide the sum of the values by the number of values (another typical value)
- For discrete observations:
- Sample mean: $x = x_1 + x_2 + : : : + x_{n/n}$
- Population mean: $= x_1 + x_2 + : : : + x_{N/N}$

Mean for grouped data: discrete single value data

$x = f_1x_1 + f_2x_2 + : : : + f_nx_{n/n}$

Mean for grouped data: continuous data

$\_x = f_1m_1 + f_2m_2 + : : : + f_nm_{n/n}$

Here m = midpoint

Adding a constant

Let $y_i = x_i + c$ where c is a constant then $\_y = \_x + c$ (_y is y bar _x is x bar)

Multiplying a constant

Let $y_i = x_i c$ where c is a constant then $\_y = \_xc$

Median - The median of a data set is the middle value in its ordered list.

Steps to obtain median

Arrange the data in increasing order. Let n be the total number of observations in the dataset.

1. If the number of observations is odd, then the median is the observation exactly in the middle of the ordered list, i.e. $\frac{n+1}{2}$ observation
2. If the number of obsevations is even, then the median is the mean of the two middle observations in the ordered list, i.e. mean of $\frac{n}{2}$ and $\frac{n}{2}+1$ observation

Adding a constant

- Let $y_i = x_i + c$ where c is a constant then new median = old median + c.

Multiplying a constant

l Let $y_i = x_i c$ where c is a constant then new median = old median *c

Mode - The mode of a data set is its most frequently occurring value.

Steps to obtain mode

1. If no value occurs more than once, then the data set has no mode.
2. Else, the value that occurs with the greatest frequency is a mode of the data set.

Adding a constant

- Let $y_i = x_i + c$ where c is a constant then new mode = old mode + c

Multiplying a constant

l Let $y_i = x_i c$ where c is a constant then new mode = old mode *c

Measures of dispersion

l To describe that difference quantitatively, we use a descriptive measure that indicates the amount of variation, or spread, in a data set.

○ Such descriptive measures are referred to as

- measures of dispersion, or • measures of variation, or
- measures of spread.

Range - The range of a data set is the difference between its largest and smallest values.

The range of a data set is given by the formula Range = Max - Min where Max and Min denote the maximum and minimum observations, respectively.

Range sensitive to outliers

- Range is sensitive to outliers.
- Though the two datasets di_er only in one datapoint, we can see that this contributes to the value of Range significantly. This happens because the range takes into consideration only the Min and Max of the dataset.

Variance –

- In contrast to the Range, the variance takes into account all the observations.
- One way of measuring the variability of a data set is to consider the deviations of the data values from a central value.

Population variance and sample variance

Recall when we refer to a dataset from a population, we assume the dataset has N observations, whereas, when refer to a dataset from a sample, we assume the dataset has n observations. l The variance is computed using the following formulae l The denominator for computing population variance is N, the total number of observations.
l The denominator for computing sample variance is (n ⍰ 1). The reason for this will be clear in forthcoming courses on statistics.

$$\sigma^2 = \frac{\Sigma (x - \mu)^2}{N}$$ Population Variance

$$s^2 = \frac{\Sigma (x - \bar{x})^2}{n - 1}$$ Sample Variance

$$s = \sqrt{\frac{\Sigma (x - \bar{x})^2}{n - 1}}$$ Sample Standard Deviation

Adding a constant

- Let $y_i = x_i + c$ where c is a constant then new variance = old variance.

Multiplying a constant

- Let $y_i = x_i c$ where c is a constant then new variance = c^2 *old variance.

Standard definition – The quantity which is the square root of sample variance is the sample standard deviation.

Units of standard deviation

- The sample variance is expressed in units of square units if original variable.

Adding a constant

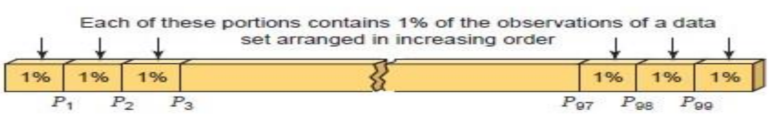
l Let $y_i = x_i + c$ where c is a constant then new variance = old variance.

Multiplying a constant

l Let $y_i = x_i c$ where c is a constant then
new variance = C2 _ old variance, (C2 is c square)

percentiles –

- The sample 100p percentile is that data value having the property that at least 100p percent of the data are less than or equal to it and at least 100(1 - p) percent of the data values are greater than or equal to it.



- If two data values satisfy this condition, then the sample 100p percentile is the arithmetic average of these values.
- Median is the 50th percentile.

Computing Percentile

To find the sample 100p percentile of a data set of size n

1. Arrange the data in increasing order.
2. If np is not an integer, determine the smallest integer greater than np. The data value in that position is the sample 100p percentile.

3. If np is an integer, then the average of the values in positions np and np + 1 is the sample 100p percentile.

Quartiles

Defnition

The sample 25th percentile is called the _rst quartile. The sample50th percentile is called the median or the second quartile. Thesample 75th percentile is called the third quartile.In other words, the quartiles break up a data set into four partswith about 25 percent of the data values being less than the first(lower) quartile, about 25 percent being between the _rst andsecond quartiles, about 25 percent being between the second and third(upper) quartiles, and about 25 percent being larger than the third quartile.

The Five Number Summary

- Minimum
- Q_1: First Quartile or lower quartile
- Q_2: Second Quartile of Median
- Q_3: Third Quartile or upper quartile
- Maximum

The Interquartile Range (IQR) - The interquartile range, IQR, is the difference between the first and third quartiles; that is,
$IQR = Q_3 - Q_1$

- IQR for the example
- First quartile, $Q_1 = 49:75$
- Third quartile, $Q_3 = 68$
- $IQR = Q_3 - Q_1 = 18:25$

Contingency table –

- To understand the association between two categorical variables.
- Learn how to construct two-way contingency table.
- Learn concept of relative row/column frequencies and how to use them to determine whether there is an association between the categorical variables.

Example 1: Gender versus use of smartphone

- A market research _rm is interested in _nding out whether ownership of a smartphone is associated with gender of a student. In other words, they want to _nd out whether more females own a smartphone while compared to males, or whether owning a smartphone is independent of gender. To answer this question, a group of 100 college going children were surveyed about whether they owned a smart phone or not.
- The categorical variables in this example are

Gender: Male, Female (2 categories)- Nominal variable

- ❖ Own a smartphone: Yes,
- ❖ No (2 categories)- Nominal variable

Example 2: Income versus use of smartphone

- A market research _rm is interested in _nding out whether ownership of a smartphone is associated with income of an individual. In other words, they want to find out whether income is associated with ownership of a smartphone.
- To answer this question, a group of 100 randomly picked individuals were surveyed about whether they owned a smart phone or not.
- The categorical variables in this example are
- ❖ Income: Low, Medium, High (3 categories) -Ordinal variable
  - ❖ Own a smartphone: Yes, No (2 categories) - Nominal variable

Row relative frequencies

- What proportion of total participants own a smart phone?
- What proportion of female participants own a smart phone?

Row relative frequency: Divide each cell frequency in a row by its row total.

Column relative frequencies

- What proportion of total participants are female?
- What proportion of smart phone owners are females?

Column relative frequency: Divide each cell frequency in a column by its column total.

Association between two variables

Knowing information about one variable provides information about the other variable.

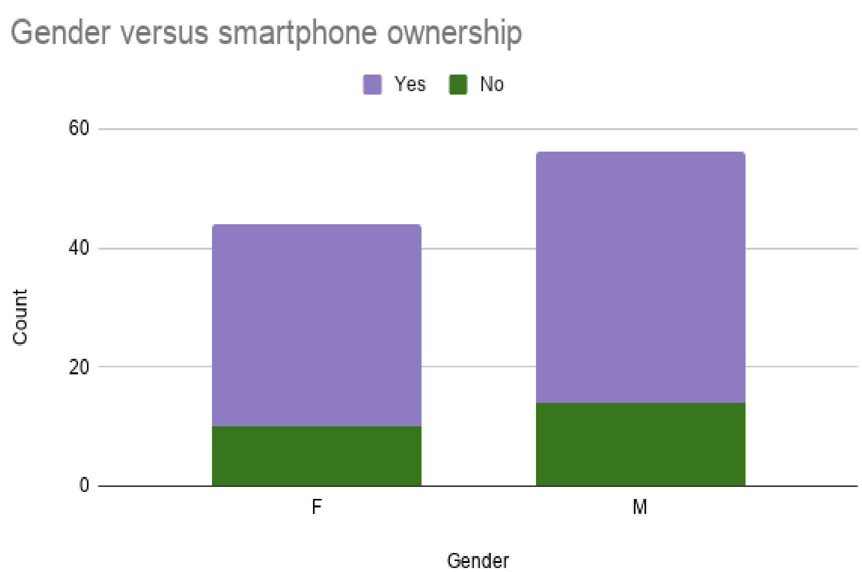
- To determine if two categorical variables are associated, we use the notion of relative row frequencies and relative column frequencies described earlier.

Association between two variables

- If the row relative frequencies (the column relative frequencies) are the same for all rows (columns) then we saythat the two variables are not associated with each other.
- If the row relative frequencies (the column relative frequencies) are different for some rows (some columns) then we say that the two variables are associated with each other.

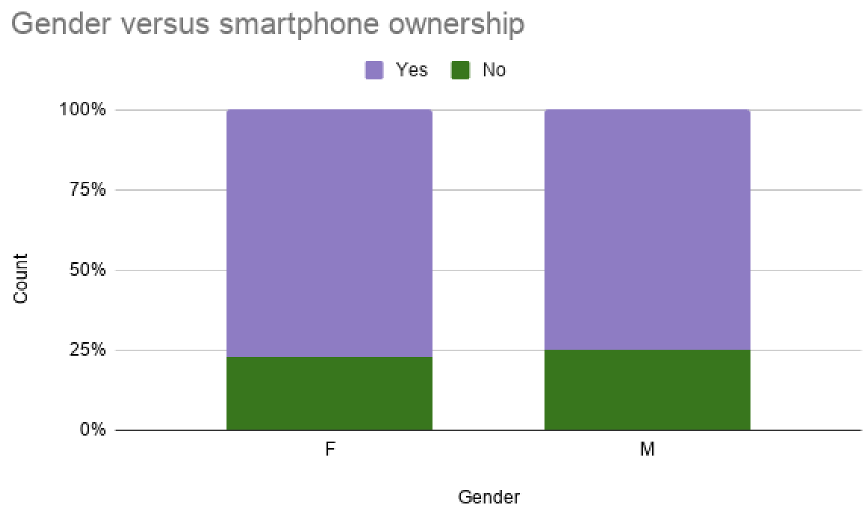
Stacked bar chart

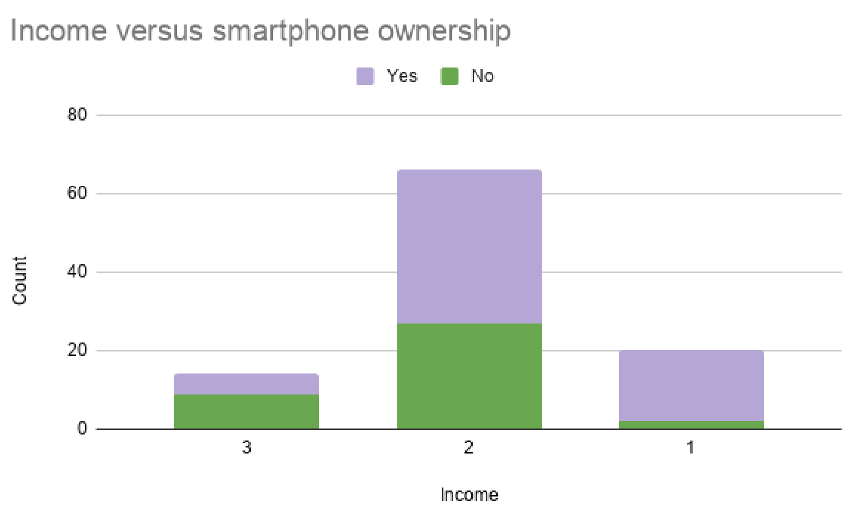
- Recall, a bar chart summarized the data for a categorical variable. It presented a graphical summary of the categorical variable under consideration, with the length of the bars representing the frequency of occurrence of a particular category.
- A stacked bar chart represents the counts for a particular category. In addition, each bar is further broken down into smaller segments, with each segment representing the frequency of that particular category within the segment. A stacked bar chart is also referred to as a segmented bar chart.



Example 1: 100% Stacked bar chart

A 100% stacked bar chart is useful to part-to-whole relationships





Scatter plot - A scatter plot is a graph that displays pairs of values as points on a two-dimensional plane.

We use a scatterplot to look for association between numerical variables.

- To decide which variable to put on the x-axis and which to put on the y-axis, display the variable you would like to explain along the y-axis (referred as response variable) and the variable which explains on x-axis (referred as explanatory variable.)

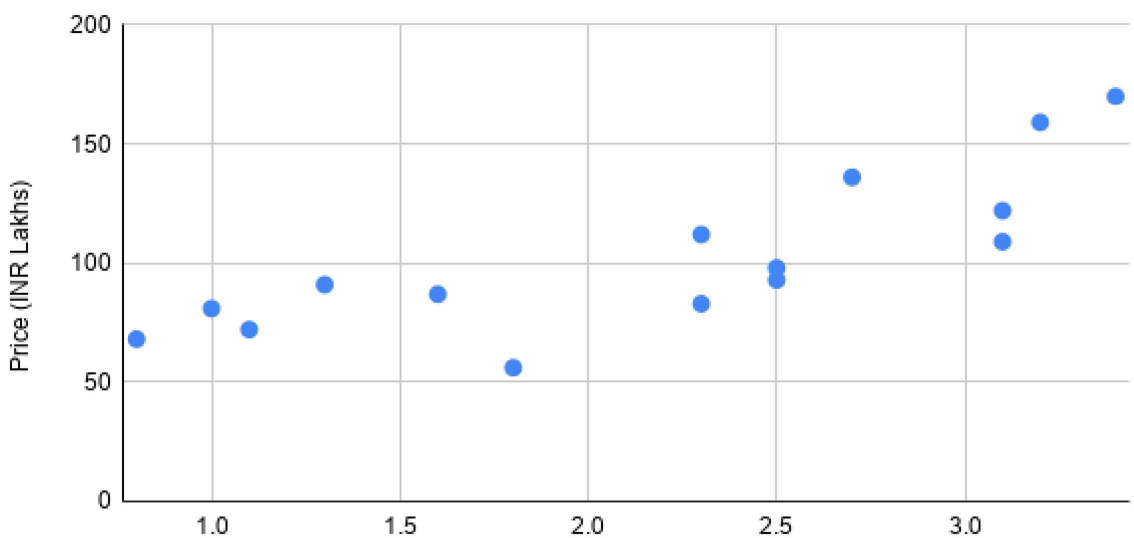
Example 1: Prices of homes

A real estate agent collected the prices of di_erent sizes of homes. He wanted to see what was the relationship between the price of a home and size of a home. In particular, he wanted to know if the prices of homes increased linearly with the size or in any other way? To answer the question, he collected data on 15 homes. The data he recorded was

	Size (100sq feet)	Price INR Lakh
1	0.1	
2	1.1	68
3	1.3	87
4	1.6	45
5	1.8	72
6	2.3	69
7	2.3	36
8	2.5	52
9	2.5	47
10	2.7	85
11	3.1	69
12	3.1	69
13	3.2	52
14	3.4	88
15	3.6	56
16	3.6	66

1. Size of a home measured in 1000 of square feet.
2. Price of a home measured in lakh of rupees.

Scatter plot -



Size (1000 sq ft)

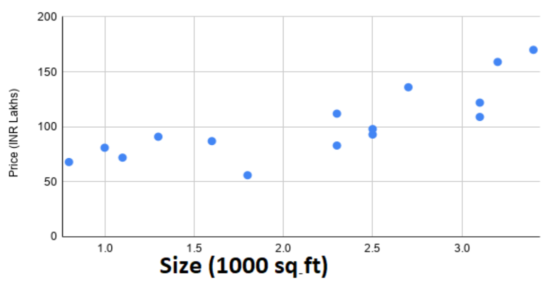
Describing association

When describing association between varaibles in a scatter plot, there are four key questions[1] that need to be answered

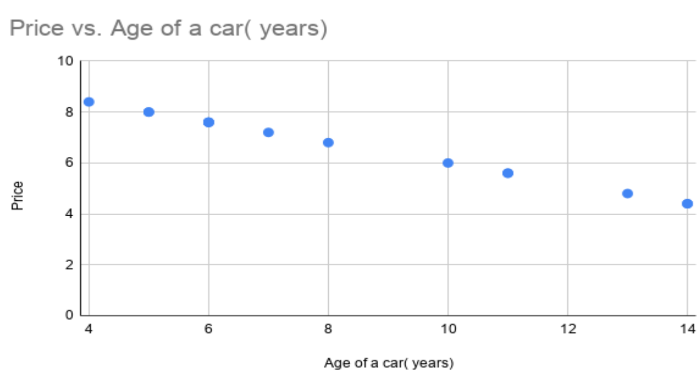
1. Direction: Does the pattern trend up, down, or both?
2. Curvature: Does the pattern appear to be linear or does it curve?
3. Variation: Are the points tightly clustered along the pattern?
4. Outliers: Did you find something unexpected?

Describing association: Direction

Does the pattern trend up, down, or both?



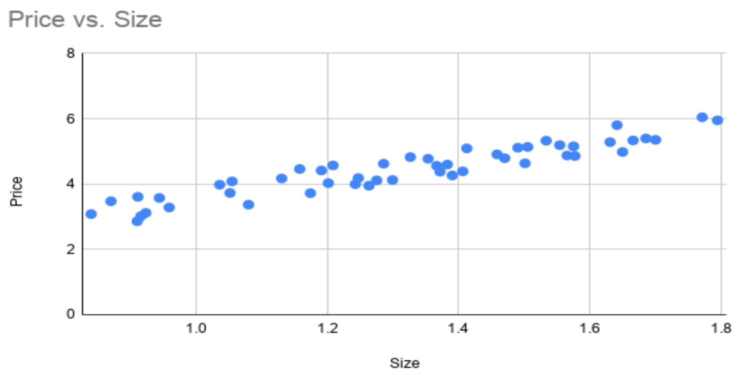
i) UP



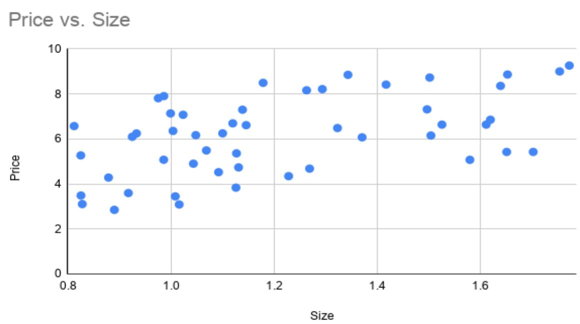
ii) Down

Describing association: Variation

Are the points tightly clustered along the pattern?

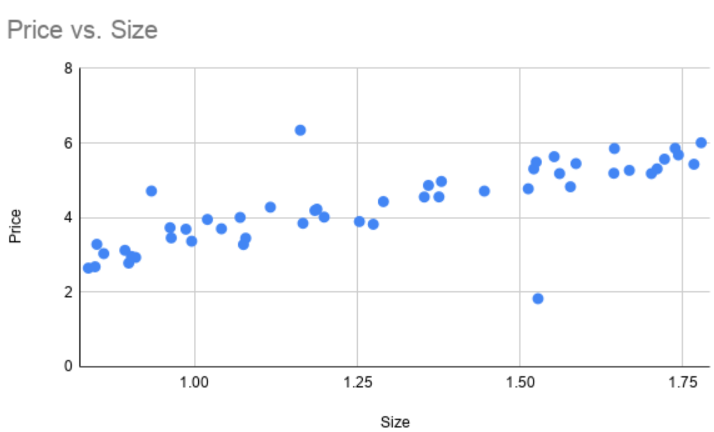


i) Tightly clustered



ii) Variable

Describing association: Outliers



Did you find something unexpected ?

Covariance

Covariance quantifies the strength of the linear association between two numerical variables.

Key observation

l When large (small) values of x tend to be associated with large (small) values of y- the signs of the deviations, (x_i - _x)and (y_i - _y) will also tend to be same.
l When large (small) values of x tend to be associated with small (large) values of y- the signs of the deviations, (x_i - _x)and (y_i - _y) will also tend to be different.

Covariance

Definition

Let x_i denote the i th observation of variable x, and y_i denote thei th observation of variable y. Let (x_i ; y_i) be the i th pairedobservation of a population (sample) dataset having N(n) observations. The Covariance between the variables x and y is given by

- Population covariance: $Cov(x,y) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{N}$
- Sample covariance: $Cov(x,y) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n-1}$

Covariance: Example 1

Age x	Height y	Deviation of x $(x_i - \bar{x})$	Deviation of y $(y_i - \bar{y})$	$(x_i - \bar{x})(y_i - \bar{y})$
1	75	-2	-17.6	35.2
2	85	-1	-7.6	7.6
3	94	0	1.4	0
4	101	1	8.4	8.4
5	108	2	15.4	30.8
				82

- Population covariance: $\frac{82}{5} = 16.4$
- Sample covariance: $\frac{82}{4} = 20.5$

Units of Covariance

- The size of the covariance, however, is di_cult to interpret because the covariance has units.
- The units of the covariance are those of the x-variable times those of the y-variable.

Correlation

- A more easily intepreted measure of linear association between two numerical variables is correlation

- It is derived from covariance.
- To find the correlation between two numerical variables x and y divide the covariance between x and y by the product of the standard deviations of x and y. The Pearson correlation
    - coefficient, r , between x and y is given by

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} = \frac{cov(x,y)}{s_x s_y}$$

Remark

The units of the standard deviations cancel out the units of covariance.

Remark

It can be shown that the correlation measure always lies between -1 and +1.

Correlation: Example 1

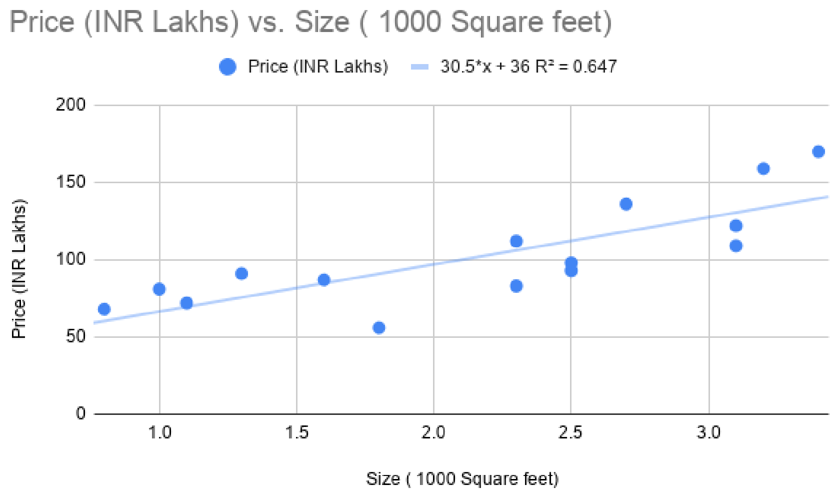
Age x	Height y	sq.Devn of x $(x_i - \bar{x})^2$	sq.Devn of y $(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
1	75	4	309.76	35.2
2	85	1	57.76	7.6
3	94	0	1.96	0
4	101	1	70.56	8.4
5	108	4	237.16	30.8
		10	**677.2**	**82**

- $s_x = 1.58$, $s_y = 13.01$
- $r = \frac{82}{\sqrt{10 \times 677.2}}$ OR $\frac{20.5}{1.58 \times 13.01} = 0.9964$

Summarizing the association with a line

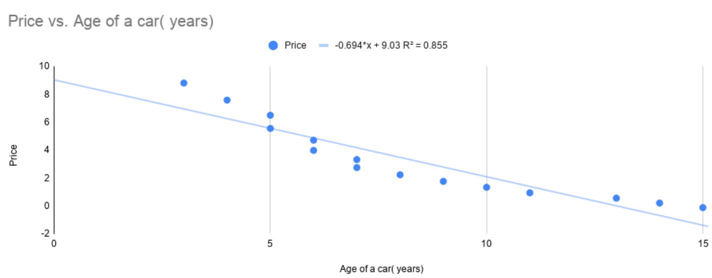
- The strength of linear association between the variables was measured using the measures of Covariance and Correlation.
- The linear association can be described using the equation of a line.

Example 1: Size versus Price of homes: Equation



Equation of the line: Price = 30:5 *Size + 36;
R_2 = 0:647; r = 0:804

Example 2: Age versus Price of cars: Equation



Equation of the line: Price = -0:694 * Age + 9:03;
R_2 = 0:855; r = -0:9247

- Understand the association between a categorical variable and numerical variable.
- Assume the categorical variable has two categories (dichotomous).

Example 1: Gender versus marks DATA

A teacher was interested in knowing if female students performed better than male students in her class. She collected data from twenty students and the marks they obtained on 100 in the subject.

	Gender	Marks
1	F	71
2	F	67
3	F	65
4	M	69
5	M	75
6	M	83
7	F	91
8	F	85
9	F	69
10	F	75
11	M	92
12	F	79
13	M	71
14	M	94
15	F	86
16	F	75
17	F	90
18	M	84
19	F	91
20	M	90

Example 1: Scatter plot

Point Bi-serial Correlation Coe_cient

Let X be a numerical variable and Y be a categorical variable with two categories (a dichotomous variable).

The following steps are used for calculating the Point Bi-serial correlation between these two variables:

Step 1 Group the data into two sets based on the value of the

dichotomous variable Y . That is, assume that the value of Y
is either 0 or 1.

Step 2 Calculate the mean values of two groups: Let Y_0 and Y_1 be the mean values of groups with Y = 0, and Y = 1, respectively.

Step 3 Let p_0 and p_1 be the proportion of observations in a group with Y = 0 and Y = 1, respectively, and s_X be the standard deviation of the random variable X.

The correlation coefficient.

$$r_{pb} = \left(\frac{\bar{Y}_0 - \bar{Y}_1}{s_X} \right) \sqrt{p_0 p_1}$$

Week-5

L.5.1 Permutations & Combinations - Basic principles of counting:

+ variable { Categorical / Numerical }

+ Association btwn variables
  - (i) both categorical
  - (ii) " Numerical (covariances, correlation
  - (iii) One is categorical & Numerical

+ permutation → counting with order
+ Combination → " without order

eg. 1 Buying clothes.

+ gift card → buy shirt/pant. (either of 2)

shirt	pant
Yellow	Black
Blue	Blue
Green	Brown
Red	

→ How many different ways can I use my card?

4 → shirt
3 → pant

(If shirt no pant & vice versa) (Actions are dependent on each)

∴ total choices are 4+3 = 7.

Addition rule of counting:

+ If an action A (shirt) can occur in n_1 (4) different ways, another action B (pant) can occur in n_2 (3) different ways, then the total number of occurrence of the actions A or B is $n_1 + n_2$.

eg: 2 Choose either a pant or a short (last eg!)

Now: Change to one shirt & one pant. (~~No of f~~

Q: No. of choices?

Ans: (Y Bla, Y Blu, Y Brow, = 3

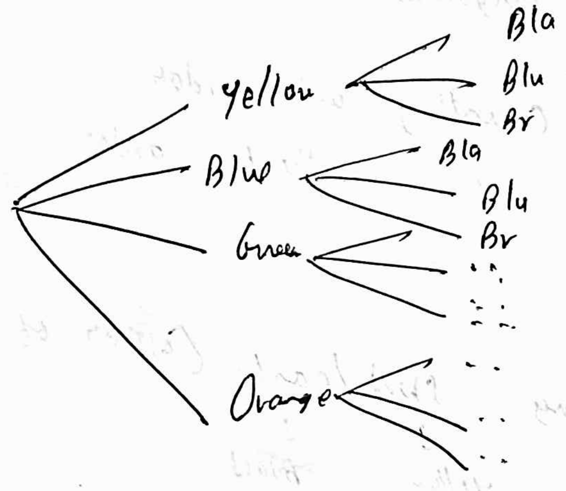
B " B " B " → 3

G " G " G " = 3

O " O " O " = 3

4 4 4 12

Tree



Matching short & pants & shoes:

4 Shirts 3 Pant 2 Shoes

Black

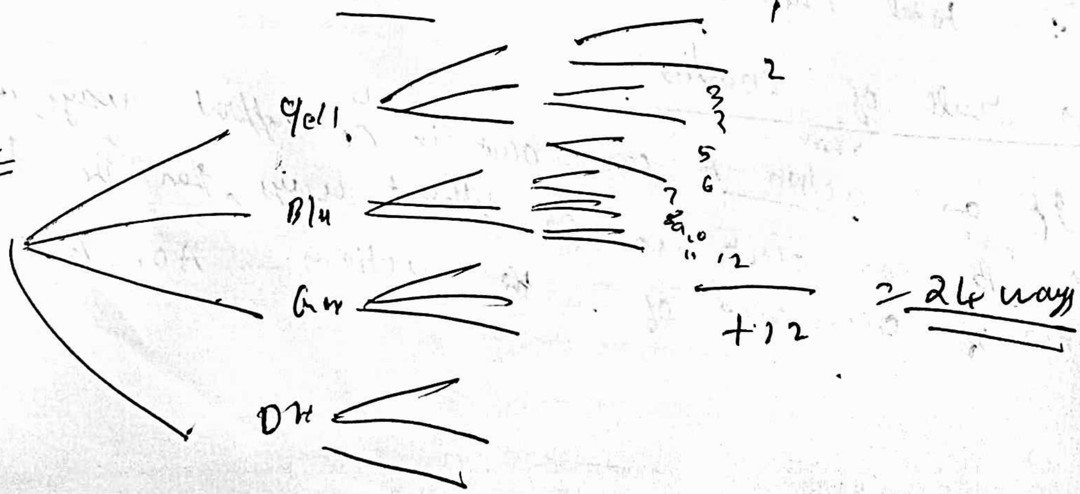
Brow.

↳ 12 pairs + Black 2

12 pairs + Brow

24

Tree



Multiplication rule of counting:

- If an action A can occur in n_1 diff ways, another action B can occur in n_2 diff ways then total no. of occurance of Actions A & B together is $n_1 \times n_2$.

(A or B) => $n_1 + n_2$

- Suppose that r actions are to be performed in a definite order. Further suppose that there are n_1 possibilities for the first actions & that corresponding to each of these possibilities are n_2 possibilities for second action and so on. Then there are $n_1 \times n_2 \times \ldots \times n_r$ possibilities altogether for the r actions

Example - 2 Application: Creating Alpha-numeric code.

Alpha numeric password:

+ Create 6 digit alpha numeric followed by 4 numbers
  + have first two letters
  + Repetition allowed.
    No of ways - $26 \times 26 \times 10 \times 10 \times 10 \times 10 = 6,760,000$
  + Repetition not allowed.
    No. of ways - $26 \times 25 \times 10 \times 9 \times 8 \times 7 = 3,276,000$

Section summary:

- Addition rule of counting ($n_1 + n_2$ => A or B
- Multiplication " " $N_1 \times N_2$ => A & B.

Permutations & combinations - Factorials

Lecture 5.2

Example - 3 Order of finishes in a race.

+ There are 8 athletes, 100 m race, What are possible ways the athletes can finish the race. (assuming no ties)

Position →	1	2	3	4	5	6	7	8
→	P1	P2	P3	P4	P8	P7	P6	P5
(No. of choices)	8	7	6	5	4	3	2	1

∴ possible ways $= 8\times7\times6\times5\times4\times3\times2\times1$

(ie) First place – any of 8, second → any of remaining 7 & so on
last → only one

Factorial

Definition

Product of first n positive integers (counting numbers) is called n factorial & is denoted by $n!$. In symbols

$$n! = n\times(n-1)\times\cdots\times1$$

Remark

By convention $\boxed{0! = 1}$ & $1! = 1$

∴ we represent $8\times7\times6\times5\times4\times3\times2\times1 \Rightarrow 8!$

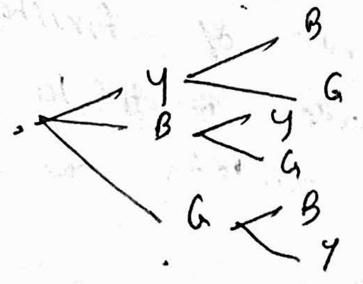
Eg: 4 Choosing shirts (any order of 3 shirts chosen without repetition)

	1	2	3
1	Y	B	G
2	Y	G	B
3	B	Y	G
4	B	G	Y
5	G	Y	B
6	G	B	Y

$\Rightarrow n = 3$

$n! = 3\times2\times1 = 6$

(Yellow, Blue, Green)
↓
possible order of choosing



$= 6.$

Example 5

① $5! = 5\times4\times3\times2\times1 = 120$

② Observe $5! = 5\times4! \Rightarrow 5\times4\times3!$.

$$6! = 6\times\underbrace{5\times4\times3\times2\times1}$$

$$= 6\times5! \Rightarrow 6\times5\times4!$$

$$n! = n\times((n-1)\times\ldots1)$$

$$\boxed{n! = n\times(n-1)!}$$

(ie) in general

③ Observe $5! = 5\times4! = 5\times4\times3!$

• in general, for $i \le n$, we have

$$\boxed{n! = n\times(n-1)\ldots\times(n-i+1)\times(n-i)!}$$

(e) $n = 5$ $i \le n$
$i = 1$
$i = 2$
$i = 3$
$i = 4$
$i = 5$

$5! = n\times4!$

$i=1$ $5\times4!$

$i=2$ $= 5\times4\times3!$

$i=3$ $= 5\times4\times3\times2!$

$i=4$ $= 5\times4\times3\times2\times1$

Example 6: Simplifying expressions

① $\dfrac{6!}{3!} = \dfrac{6\times5\times4\times\overbrace{3\times2\times1}^{3!}}{3!} = 6\times5\times4 = 120$

② $\dfrac{6!\times5!}{3!\times4!} = 6\times5\times4 \cdot \times \dfrac{5\times\cancel{4!}}{\cancel{4!}} = 6\times5\times4\times5 = 600.$

③ Express $25\times24\times23$ in terms of factorials

$$\frac{25\times24\times23\times22\times21\times\ldots 1}{22\times21\times\ldots} = \frac{25!}{22!}$$

Section Summary.

- Introduced factorial notation
- Simplifying expressions

L-5.3 Permutations and Combinations - Permutations: Distinct objects.

Permutation:

A ~ an ordered arrangement of all or some of n objects

(or) ordered arrangement

n-objects (Distinct)

$n=3$ | A | B | C | (here distinct) (no repetition)

(or)

$n=3$ A | A | B $\Rightarrow$ 2 Red one yellow (not distinct)

Example:

Take A - B - C → Possible arrangement - taking all at a time

All (i)

	1	2	3
1	A	B	C
2	A	C	B
3	B	A	C
4	B	C	A
5	C	A	B
6	C	B	A

→ 6 arrangements (all.)

Some

(ii) A, B, C → Possible arrangement - taking two at a time

AB, BA] same

BC, CB] same

AC, CA] same

$n=3$, 2 at a time

↓

$\boxed{\text{6 arrangements}}$

Example - 4

Take A, B, C, D → Possible arrangements taking all at a time

$n=4$ A B C D

$$7DC. \quad \boxed{A} \quad B \quad C \quad D \rightarrow 6 \quad \Rightarrow 4\times 3\times 2\times 1 = \underline{\underline{24}}$$

(B) A C D → 6

(C) A B D → 6

(D). A B C → 6

24

Example

Take A, B, C, D → Possible arrangement, two at a time

AB	BA
AC	CA
AD	DA
BC	CB
BD	DB
CD	DC

$\boxed{12}$ possible arrangement

A < B, C, D

B < A, C, D

C < A, B, D

D < A, B, C

= 12

Permutation formula :

No of possible permutations of r objects, from a collection of n distinct objects given by formula.

$r \le n$,

$n=3$ ABC, $r=3$

$n=3$ ABC, $r=2$

$n=4$ ABCD, $r=4$

$n=4$ ABCD, $r=2$.

1 2 3 ... r

$n \times n-1 \times n-2 \ldots \times \ldots n-(r-1)$ as denoted by ${}^nP_r$

$${}^nP_r = \frac{n\times(n-1)\times(n-2)\ldots(n-r+1)\times(n-r)(n-r-1)\ldots\times 1}{(n-r)(n-r-1)\ldots\times 1}$$

$$\boxed{{}^nP_r = \frac{n!}{(n-r)!}}$$

Repetition is not allowed

• Special cases:

① $^nP_0 = \frac{n!}{(n-0)!} = \frac{n!}{n!} = 1$ → There is only one ordered arrangement of 0 objects

② $^nP_1 = \frac{n!}{(n-1)!} = n$ → There are n ways of choosing one object

③ $^nP_n = \frac{n!}{(n-n)!} = \frac{n!}{0!} = n!$ No repetition → AAB

* Take ABC → Possible arrangement all at a time

$n = 3, \quad r = 3$

$^nP_n = n!$

A B C
A C B
B A C
B C A
C A B
C B A

} 6.

$n = 3$

$3! = 6.$

(i.e) $n = 3, r = 3, \quad ^nP_r = \frac{n!}{(n-r)!} = \frac{3!}{0!} = 6.$

Example

ABC → 2 at a time.

$n = 3, \quad r = 2$

$^nP_r = \frac{n!}{(n-r)!} = \frac{3!}{1!} = 6.$

Take ABCD → all at a time

1 2 3 4

$n = 4, \quad r = 4, \quad ^nP_r = \frac{n!}{(n-r)!} = \frac{4!}{0!} = \frac{4 \times 3 \times 2 \times 1}{1}$

$= 24.$

1 2 3 4

6 + 6 + 6 + 6 = 24

Take ABCD → 2 at a time.

$n = 4, r = 2 \quad ^nP_r = \frac{4!}{2!} = 4 \times 3 = 12$

① ②
A 3
B 3
C 3
D 3 = 12

Application:

- Committee of 8 persons in how many ways can we choose a chairman and vice chairman assuming one person can hold not more than one position

C V.C.
A 7
B 7
C 7
D 7
E
F
G
H

$\boxed{7\times 8}$

$n=8$, $r=2$, No. of ways $= {}^nP_r = \frac{n!}{(n-r)!} = \frac{8!}{(8-2)!} = \cancel{8\times7\times6\times5\times4\times3}$ $\frac{8!}{6!} = 8\times 7 = 56.$

Example

~~no. of~~ 4 digit number formed using 1, 2, 3, 4, 5 no repeat.

1 2 3 4

$n=5$, $r=4$

${}^nP_r = \frac{5!}{(5-4)!} = 5\times4\times3\times2 = 120$

- How many of these will be even?

last digit fixed: 2 or 4

using nos → (i) 1, 3, 4, 5 left ⇒ n; $n=4$, $r=3$

since last digit is fixed, $r=3$.

${}^4P_3 = \frac{4!}{(4-3)!} = 4\times3\times2 = \cancel{24}$

using nos → (ii) 1, 2, 3, 5 → $n=4$, $r=3$

${}^4P_3 = \cancel{8}\ 24$

∴ No. of even $= 48$.

Example

6 ppl. go to cinema. sit in 10 rowed seats (No. of ways they can sit)

(i) $n=10$, $r=6$, ${}^{10}P_6 = 1{,}51{,}200$ (sit anywhere)

(ii) all empty seats next to eachother (together)

$\boxed{\text{Note } r \le n}$

~~8~~ 7 cases

$r=6$

${}^7P_6 = 5040$

Taking all empty chair as a single case (or) chair

| 1234 | 6. |

| 1 | 1234 | ~~5~~ |

| 2 | 1234 | 4. |

} 7 cases

Eg: Take ABC → all at a time – repetition allowed

$3 \times 3 \times 3$

A	A	A
B	B	B
C	C	C

multiplication rule

$= 27.$

A A A
A A B
A A C
A B B
A B C
A C A
A B A
A C A
A C B
A C C
⋮

$9 + 9 + 9 = 27$

$n \to$ no. of distinct object (A B C)

$r \to$ How many times repeated = □ □ □
(or)
No. of combinations continued □ □ □

Eg.

ABC → two at a time

A	A	A	B	B	B	C	C	C
A	B	C	A	B	C	A	B	C

$\underbrace{\qquad\qquad\qquad}_{9}$

Permutation formula:

1	2	3	...	r
□	□	□	...	□
n	n	n		n

$= \boxed{n^r} \to$ repetition

$n \to$ distinct objects

$r \to$ taking r objects from n, $r \leq n$

The no. of possible permutations of r objects from a collection of n distinct object when repetition is allowed is given by the formula

$n \times n \times \ldots \times n.$

and is denoted by n^r.

Eg:

ABC → all at a time $n = 3,\ r = 3,\ n^r = 3^3 = 27$

ABC → two at a time $n = 3,\ r = 2 = 3^2 = 9$

Summary

(i) No. repeats → $^nP_r = \frac{n!}{(n-r)!}$

(ii) allowed → n^r

Lecture-5-4 : Permutations & combinations: Permutations : objects not distinct

- Suppose we want to arrange 'DATA' →

[DAT] → 3 distinct

DA_1TA_2 , $n=4$, $n!$ $4!$

$= 4\times3\times2 = 24$

DA_1TA_2
DA_2TA_1 } → DATA

A_1TA_2D
A_2TA_1D } → ATAD

(ie) for every 2 arrangement I have one arrangement same

(ie) $\frac{24}{2} = 12$ arrangements possible

Example: Rearranging letters.

- Suppose we want to rearrang the letters in the word. "DATA". How many ways?
- There are

$$\frac{4!}{2!} = 12$$

Permutation formula:

No. of permutations of n objects when p of them are of one kind and rest distinct is equal to

$$\frac{n!}{p!}$$

- Suppose we want to rearrange letters in word "STATISTICS". How many ways it can be done

$S_1 T_1 A T_2 I S_2 T_3 I_1 C S_3$

S - 3
T - 3
A - 1
I - 2
C - 1
$\overline{10}$

$\frac{10!}{}$ $n=10$

$p_1 = 3$ of first kind - S
$p_2 = 3$ " 2nd " - T
$p_3 = 1$ " 3rd " - A
$p_4 = 2$ " 4th " - I
$p_5 = 1$ " 5th " - C

permutation formula:

- The no. of permutations of n objects where p_1 of one kind p_2 is of 2nd kind & so on p_k of k^{th} kind is given by

$$\frac{n!}{p_1!\,p_2!\dots p_k!}$$

$$\frac{n=10}{3\,3\,1\,2\,1} = \frac{10!}{3!\,3!\,1!\,2!\,1!} = 50400$$

Section summary:

① No of permutations of n objects when p of them are of one kind & rest distinct is equal to $\frac{n!}{p!}$

② The no. of permutations of n objects where p_1 is of one kind p_2 is of 2nd kind, and so on p_k of k^{th} kind is given by $\frac{n!}{p_1!\,p_2!\dots p_k!}$

Example:

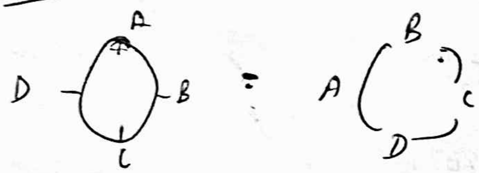
So far we saw in linear arrangement

Now

Arrange in a circle

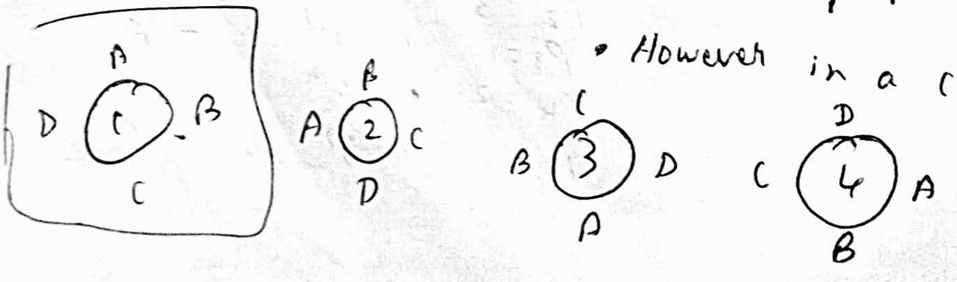
- How many ways can 4 ppl sit in a round table?
- We consider two cases; each selection is called a combination of 3 different objects taken 2 at a time.
  - Clockwise & anticlockwise are diff
  - Clockwise & " are same

ABCD



ABCD ⇒ BCDA

(i) Clock & anticlockwise are different.
- linear permutations of ABCD.
- ABCD, BCDA, CDAB, DABC are diff when people are seated in a row.
- However in a circle



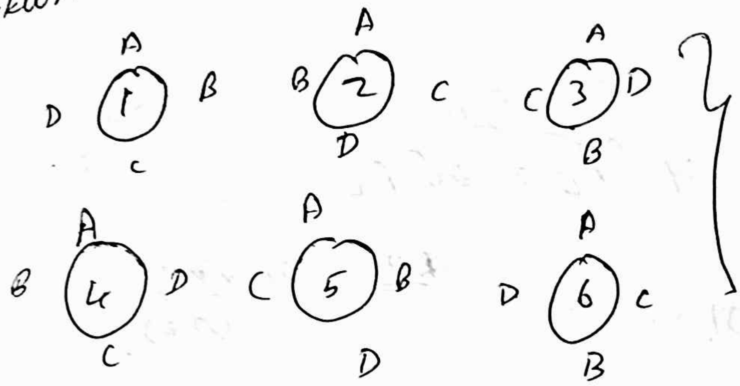
} all are same

6 Possible linear arrangements

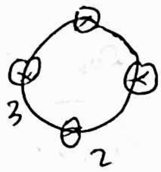
ABCD	BACD
ABDC	BADC
ACBD	BCAD
ACDB	BCDA
ADBC	BDAC
ADCB	BDCA

→ distinct in linear

Clockwise & anticlockwise are different.



A (fixed)

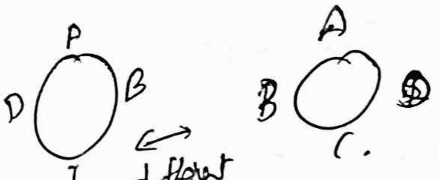


can be filled with three available alphabets (ie) $3! = 6$.

6

- In case of circular & clock & anti different we have one fixed. (ie) & consider others as combo of 3 at a time. without repetition

$A \Rightarrow \begin{pmatrix} BCD \\ CDB \\ DBC \\ DCB \\ BDC \\ CBD \end{pmatrix} = 6$ (ie) $(n-1)!$



different

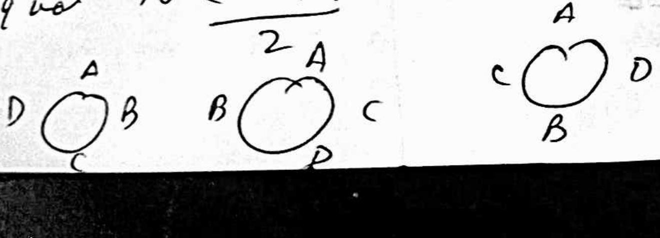
But mirror image

∴ No. of possible circular permutations of n objects if the clockwise & anticlockwise are different $\Rightarrow \underline{(n-1)!}$

& the rational behind this is I fix one of the objects, the other n minus 1 objects can be arranged among themselves in n minus 1 way and this is the same for any of these objects

∴ There are $(n-1)!$ ways of arranging circular permutations of n ~~diff~~ distinct objects (clockwise & anticlockwise are different)

*(when clockwise & anticlockwise are same) – the number of ways n-distinct objects can be arranged in a circle is equal to $\dfrac{(n-1)!}{2}$



$\Rightarrow \dfrac{(n-1)!}{2}$

Eg:

Linear permutation.

arrange 'n' objects in a row, choose r objects from these 'n' & arrange in a linear row.

$^nP_r = \frac{n!}{(n-r)!}$ (without repetition) , with repetition it is n^r

eg: Solving for n.

(i) • Find value of n if $^nP_4 = 20\,^nP_2$

LHS: $\frac{n!}{(n-4)!}$ = RHS $20 \times \frac{n!}{(n-2)!}$

$(n-2)! = 20 \times (n-4)!$

$(n-2)(n-3)\cancel{(n-4)!} = 20 \times \cancel{(n-4)!}$

$(n-2)(n-3) = 20$

$n^2 - 5n - 14 = \cancel{20}\ 0$. $(n+2)(n-7) = 0$

we get, $n = -2$ or $\boxed{n = 7}$

impossible

$\therefore \boxed{n = 7}$

(ii) $\frac{^nP_4}{^{n-1}P_4} = 5/3$

LHS $\frac{n!}{(n-4)!} \Big/ \frac{(n-1)!}{(n-5)!}$

$\frac{n!}{(n-4)!} \times \frac{(n-5)!}{(n-1)!}$

$\frac{n \times \cancel{(n-1)!}\ \cancel{(n-5)!}}{(n-4)\cancel{(n-5)!}\cancel{(n-1)!}} \Rightarrow \frac{n}{n-4} = 5/3$

$3n = 5n - 20$

$2n = 20$

$\boxed{n = 10}$

(iii) Solving r.

$$^5P_r = 2 \times {}^6P_{r-1}$$

LHS $\quad {}^5P_r = \frac{5!}{(5-r)!}$

RHS $\quad {}^6P_{r-1} = \frac{6!}{(6-(r-1))!} = \frac{6!}{(7-r)!}$

$$\frac{5!}{(5-r)!} = \frac{6!}{(7-r)!} \times 2 \longrightarrow (1)$$

$$(7-r)! = (7-r) \times (7-1-r) \times (7-2-r)!$$

$$(7-r)! = (7-r)(6-r)(5-r)!$$

(1) can be reexpressed as

$$\frac{5!}{(5-r)!} = \frac{2 \times 6 \times 5!}{(7-r)(6-r)(5-r)!}$$

$$(7-r)(6-r) = 12$$

$$r^2 - 13r + 42 - 12 = 0$$

$$r^2 - 13r + 30 = 0$$

$$(r-3)(r-10) = 0$$

give $\boxed{r = 10 \text{ (or) } 3.}$

Topic Summary:

1. permutations when n objects are distinct
   - 1.1 → repetition not allowed, $^nP_r = \frac{n!}{(n-r)!}$
   - 1.2 → " allowed (n^r)
2. Permutation when objects are not distinct, P_1, P_2, P_n, $\frac{n!}{P_1! P_2! \ldots P_n!}$
3. Circular Permutation
   - 3.1 → clock & anticlockwise different $(n-1)!$
   - 3.2 → " " same $\frac{(n-1)!}{2}$
4. Solve for r & n using permutation formula

Permutations:

* we saw how to arrange n objects (or) r objects out of n objects.
* Order was important, AB was differnt from BA
⇒ * How can we choose 2 out of 3, (or) 3 out of 10.
In this order is not important AB is same as B

→ Combinations: No. of ways we select r from n.

5.5: Permutations and Combinations: Combinations:

Permutations ⇒ ordered arrangement

Combination:
eg: How many way we select two students from a group of three students? (No orders)

A, B, C

1. AB
2. AC
3. BC

AB BA

No use in order. both are same
No meaning

+ Each selection is called a combination of 3 different objects taken 2 at a time.

3 → 1, 1 3 ↓ n ↓ r

Example:

A, B, C → Possible combinations - taking two at a time,

First Place	Second Place
A	B
A	C
B	C

	FP	SP
1. AB	A	B
	B	A
2. AC	A	C
	C	A
3. BC	B	C
	C	B

6.

No. of combination × 2! = No. of permutations
3

$^nP_r = {}^3P_2 = \frac{3!}{(3-2)!} = 6.$

Total

$n \rightarrow r \rightarrow$ pair

$\rightarrow 1 \rightarrow r!$

$2 \rightarrow r!$

Combination: Notation & formula

- In general, each combinations of r objects from n objects can give rise to $r!$ arrangements.
- The no. of possible combinations of r objects from a collection of n distinct objects is donoted by ^nC_r and is given by. $^nC_r \Rightarrow \underbrace{^nC_r \cdot r!}_{\text{Total no. of permutations}} = {}^nP_r$

n choose r.

- In general, each combination of r objects from n objects can give rise to $r!$ arrangements

$$\boxed{^nC_r = \frac{n!}{r!(n-r)!}} \qquad {}^nC_r = \frac{^nP_r}{r!} = \frac{n!}{r!(n-r)!}$$

$$^3C_2 = \frac{3!}{2!\,1!} = 3$$

- Another common notation is $\binom{n}{r}$ which is also referred to as the binomial coefficient $\binom{n}{r}$, (or) ^nC_r

Some useful results:

① $^nC_r = \frac{n!}{r!(n-r)!} \Leftrightarrow \frac{n!}{(n-r)!\,r!} = {}^nC_{(n-r)} = \frac{n!}{\cancel{(n-(n-r))}!}$

$= \frac{n!}{(n-r)! \cdot (n-(n-r))!}$

$= \frac{n!}{(n-r)! \times r!}$

(ie) selecting r objects from n objects is same as rejecting $n-r$ objects from n objects

② $^nC_n = 1$ & $^nC_0 = 1$ for all values of n

$^nC_n = \frac{n!}{n!\,(n-n)!} = \frac{n!}{n!\,0!} = 1$, $\quad {}^nC_0 = {}^nC_{n-0} = {}^nC_n = 1$

$= \frac{n!}{0!(n-0)!} = \frac{n!}{n!} = 1$

③ $^nC_r = {}^{n-1}C_{r-1} + {}^{n-1}C_r \; : \; 1 \le r \le n$

choose 3 of 5

A B C D E

$n=5$

$r=3$

A is part — A is not a part of 3 (solution) & choose 3 from 4

$$\binom{5}{3} = \text{fix } A \binom{4}{2} + \binom{4}{3}$$

(i.e) Here fix any one of n objects & looked at the situation that "r objects that has n" and "r does not have".

r has n — choose r-1 from remaining n-1 — $\binom{n-1}{r-1}$ + $\binom{n-1}{r}$ — r does not have — remaining (n-1), because eliminating 'n', & choose r

(i.e) $^nC_r = {}^{n-1}C_{r-1} + {}^{n-1}C_r$

Example:

Choosing questions in an exam

① Exam, 12 questions — Part I (7 ques), Part II (5 ques) } = 12 ques

attend total of 8 questions, selecting 3 atleast from each.

∴ In how many ways can a student select the questions

	P I	P II		
8	3	5	$\rightarrow {}^7C_3 \, {}^5C_5 +$	$= \frac{7!}{3!4!} \frac{5!}{5!0!} = 35$
	4	4	$\rightarrow {}^7C_4 \, {}^5C_4 +$	$= \frac{7!}{4!3!} \frac{5!}{4!1!} = 35 \times 5 = 175$
	5	3	$\rightarrow {}^7C_5 \, {}^5C_3$	$= \frac{7!}{5!2!} \frac{5!}{3!2!} = 21 \times 10 = 210$

$\frac{5!}{3!2!} =$

$\boxed{\text{Total} = 420}$

Example: Game of cards:

Club – ♣
spade – ♠
heart – ♥
Diamond – ♦

$\left.\right\}$ 4 suits – $13 \times 4 = 52$ cards

26 Black, 26 Red.

+ Choose 4 from 52 cards.

$$^{52}C_4 = \frac{52!}{4!\,48!} = 270725$$

+ All four cards of same suit

$$^{4}C_1 \times {}^{13}C_4 = \frac{4!}{3!} \times \frac{13!}{4!\,9!} = 2860$$

First choose which suit

+ Cards are of same colour

$$^{2}C_1 \times {}^{26}C_4 = 2 \times \frac{26!}{4!\,22!} = 29900$$

Example: Choosing a cricket team.

+ Select eleven from 17 players only 5 can bowl, requirement is cricket team of 11 must include exactly 4 bowlers? How many ways can the selection be done?

17 → 5 bowl (4) B; 12 no bowl (7 NB)

choose 11 → exactly 4 bowlers

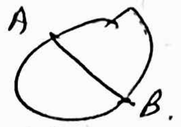
$$^{5}C_4 \; {}^{12}C_7 = 5 \times \frac{12!}{7!\,5!} = 3960 \text{ ways}$$

include exactly 4.

Example : Drawing lines in a circle

- Given n points on a circle, how many lines can be drawn connecting those points?

- $n = 2$ points, one line can be drawn



line segment: AB

$$^2C_2 = \frac{2!}{2!\,0!} =$$

- $n=3$ points three lines can be drawn

$\Rightarrow AB, AC, BC \qquad ^3C_2 = \frac{3!}{2!\,1!} = 3$

- In general, given n points, number of line segments that can be drawn connecting the point is $^nC_2$.

Section Summary :

① Notation & formula for selecting r objects from n objects $\quad ^nC_r = \frac{n!}{r!\,(n-r)!}$

② Some useful Combinatorial identities

Lecture 5.6 Permutations & Combinations - Application

* Important to distinguish b/wn situations involving combinations & situations involving permutations

- Permutation - "order matters", Combination "Order doesn't matter"

$n = 3 \quad A \quad B \quad C$

$r = 2$

AB, BA	AB
AC, CA	AC
BC, CB	BC

Permutation $\qquad$ Combination

Example Finishing a race.

- 8 athletes 100m race

(1) How many diff ways can Gold, Silver, Bronze medals

A B C D E F G H

1 - A
2 - D
3 - F

D	Gold
A	Silver
F	Bronze

They are diffrent
∴ order matter
⇒ Permutation

(2) How many diff ways can you choose the top three athletes to proceed to the next round in the competition?

A B C

B C A

ABC → go to next round
So no order need. → (ie) combination

Soln
(i) Order important, Hence permutation

$n = 8, r = 3 \Rightarrow {}^{8}P_{3} = \frac{8!}{5!} = 8 \times 7 \times 6 = 336$ ways

(ii) No. of order not important ∴ Combination

$n = 8, r = 3, {}^{8}C_{3} = \frac{8!}{5!\,3!} = \frac{336}{6} = 56$

$$\boxed{\begin{aligned} 56 \times r! &= {}^{n}P_{r} \\ 56 \times 3! &= 336 \end{aligned}}$$ → verified

Example: Selecting a team:

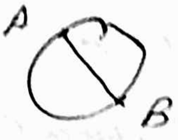
- 40 students
  (1) Choose two leaders?
  (2) Choose Captain & vice captain → having order

Soln (1) ${}^{40}C_{2} = 780$ ways

(2) ${}^{40}P_{2} = 1560$

Example Draw lines in a circle:

- n points on a circle, how many lines can be drawn connecting those points?

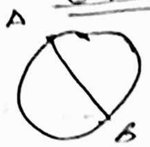


$^nC_2 \rightarrow$ given n points, we can join using $^nC_2$, total no. of lines drawn through n points. The lie on a circle.

eg: If A & B were two location on this point and, going from A to B, to B to A, which was a different case.

$\overline{AB}$, $\overline{BA}$ → direction matters (i.e.) order matters

Combination



Permutation

AB = BA

no order

$n = {}^nC_2$

Order matters

$^nC_2 \times 2! = {}^nP_2$

because every line is giving two directed line

Section Summary:

- Dist. permutation & combination
  - Eg. of situations where permutation is applied, combination is applied.

1. Basic principle of counting: Addition, Multiplication
2. Factorial ($n!$) → Simplified Expression
3. Permutation → Distinct objects 'r' (repeat & non repeats) → Object not distinct. $^nP_r = \frac{n!}{(n-r)!}$
4. Combinations $^nC_r = \frac{n!}{r!(n-r)!}$

$$\boxed{^nC_r \cdot r! = {}^nP_r}$$

5. Distinguish Permutation (or) Combination

drawn

can join
only line
to that

Week-6

Lec-6.1 Probability Basic definitions:

Objectives

1. Uncertainty & random experiment
2. sample spaces, events of random experiments
3. Simple event or compound events.
4. Basic laws of probability
5. Probability of events & use of tree diagram to compute probabilities
6. conditional probability, (i.e., find probability of an event given another event has occurred.
7. Distinguish independent & dependent events

Random Experiment _ Sample Space _ Event

* Venn diagram

Conditional Probability & Bayes Theorem

Intro

eg:
=> 50% chance india will win a toss
=> My guess is 'A' is the right choice
=> 30% chance of rain tomorrow
=> Party ABC will probably win the next election

uncertainty
↓
Not sure of outcome

-> we routinely see or hear claims as the ones mentioned above.

-> To draw valid inference about a population from a sample one needs to know how likely it is that certain event will occur under various circumstances

Descriptive $\xrightarrow{\text{Probability}}$ Inferential

Random experiment

Def: experiment
Any process that produces an observation or outcome

eg: A B

Controlled experiment
Add sugar
&
Same output

* Toss coin (A & B)
Cannot expect same outcome

* Random experiment:
experiment whose outcome is not predictable with certainty (Toss coin)

Remark
Although outcome not known in advance, let us suppose that the set of all possible outcomes is known

Eg: of Random experiment

* Guess answer to a 4 options MCQ
(i) outcome: A B C D

* Order of finish in race with 6 students.
A, B, C, D, E, F
(ii) All possible permutations of ABCDEF.

1	A
2	B
3	C
4	D
5	E
6	F

(iii) Tossing 2 coins & noting the outcomes
Outcome: HH, HT, TH, TT

(iv) Ex: Measuring lifetime (in hours) of a bulb.
Outcome 0 (or) 1 hour 2 (or) 3 hr ... 6.50 or

(v) Throwing a dart on a sq. unit & note the point where it lands.
O/C: Any point in the square (Assuming the dart lands within the square)

(i) ABCD → 4 choices ⇒ discrete
(ii) 1, 2, 3, ... → unique but infinite
(v) point is countable but many.

Sample Space:

Def: (Ω (or, S)) → collection of all basic outcomes
① Mutually exclusive: Only one basic possible outcome that can occur must be outcome can occur
② exhaustive: One basic outcome must occur

Die, $S = \{1, 2, 3, 4, 5, 6\}$ → die
$S = \{H, T\}$ → Coin

eg. of sample space:

set of all outcomes:

(i) $S = \{A, B, C, D\}$
(ii) Order of finish in a race with 6 students
A, B, C, D, E, F
$S = \{A\text{-}F, ABDCEF, \ldots\}$ 6! permutations

(iii) Two Coins:
$S = \{HH, HT, TH, TT\}$

(iv) Measuring lifetime (hours) of a bulb:
$S = \{x : 0 \le x \le \cancel{\infty}\}$

(v) Dart throw on a unit square & not where it lands.

Section Summary
① Random experiment
② Sample space:
$S = \{$Basic outcomes of a Random experiment$\}$

Lecture - 6.2 Probability Events

Events:
Def: → a collection of basic outcomes
→ an event is a subset of sample space
→ an event has occurred if the outcome is in subset

Toss: $S = \{H, T\}$

Roll a die $S = \{1, 2, 3, 4, 5, 6\}$ → collection

Basic outcome: what I have when I roll a die.

$E = \{1, 3, 5\}$, $E = \{2, 4, 6\}$

subset of total sample space

collection of ~~total~~ basic outcomes.

Toss a twice, $S = \{HH, TT, HT, TH\}$.

$B = \{HH\}$.

event also.

Eg:

(i) Guessing answers to a MCQ.

Ans: A; $E = \{A\}$

(ii) Order of finish in a race with 6 students -

A, B, C, D, E, F

Event: "A" finish the race first

$E = \{ABCDEF, ADCDFE, ABDCFE \ldots\}$.

A | B | C | D | E | F

fixed — 5! ways.

(iii) Tossing two coins & noting the outcomes

event: Head on the first toss $E = \{HH, HT\}$

(iv) Measure lifetimes

event: lifetime is less than (or) equal to four hours

$E = \{x : 0 \leq x \leq 4\}$

Event → $\boxed{SET}$

Set operations:
- union
- intersection
- Complement

Union of events:

S = Sample Space

$E, F \subseteq S$

these are event.

→ For any 2 events, E and F, we define the new event $E \cup F$ called the union of events E and F, to consist of all outcomes that are in E or in F or in both E and F.

$\underbrace{E \cup F}_{\text{event}}$ = all outcomes in E (or) F (or) Both

→ The event $E \cup F$ will occur if either E or F occurs

Eg: Union of events.

→ Experiment: Guessing answers to a four option multiple choice question:

Event:

→ answer is A; $E_1 = \{A\}$

" " B; $E_2 = \{B\}$

" " A or B; $E_3 = E_1 = E_1 \cup E_2 = \{A, B\}$.

→ Experiment: Order of finish in a race with 6 students.

A, B, C, D, E, F.

→ A finishes first

$E_1 = \{ABC\ldots, ABCDFE\ldots\ldots AFEDBC\}$ 5!

→ B comes second

$E_2 = \{ABCDEF, CBADEF \ldots\ldots\}$ 5!

→ A comes first or B comes second.

$E_1 \cup E_2 =$

$\{ABCDEF, ABCDFE, ABDCFE\ldots \}$

Here B comes first, Not included.

Eg:

→ Tossing 2 coins and noting the outcomes.

→ head on first $E_1 = \{HH, HT\}$

→ head on second toss $E_2 = \{HH, TH\}$

→ head on 1st or 2nd toss $E_1 \cup E_2 = \{HH, HT, TH\}$

Intersection of events:

→ For given, E & F, $E \cap F$ consist of all outcomes that are in E and F.

→ $E \cap F$ if both E and F occurs.

Eg: Order of finish in race 6 students

→ A finish first

$F_1 = \{ABCDEF, ABCDFE, \ldots\}$

→ B comes second.

$E_2 = \{A, B, C, D, E, F; ABCDFE, \ldots\}$

→ A comes first and B comes second

$E_1 \cap E_2 = \{ABCDEF, ABCDFE, ABDCFE, \ldots ABDCFE\}$ $4!$

Experiment: Tossing 2 coins and noting the outcomes event:

(i) head on the first toss $E_1 = \{HH, HT\}$

(ii) " " " second " $E_2 = \{HH, HT\}$

(iii) head on both toss $E_1 \cap E_2 = \{HH\}$

Null event & disjoint event:

$S = \{HH, HT, TH, TH\}$

E_1 → Head in first toss
E_2 → ~~Head in~~ Tail in first toss. } cannot happen together

$S = \{A, B, C, D\}$ → MCQ, If only one option is correct we can't have 2 at a time

Def: Null event
An event without any outcomes the null event and designate it as ϕ.

Def: disjoint:
If intersection of E and F is the null event. then since E and F cannot simultaneously occur, we say that E and F are disjoint or mutually exclusive

$$E \cap F = \phi.$$

eg: $S = \{H, T\}$
$E = \{H\}$
$F = \{T\}$
$E \cap F = \phi.$ → If I get Head, I cant get tail and vice versa.

Eg: of Null event

Experiment: Guessing answers to a 4 option MCQ.

Event.
→ Answer is A: $E_1 = \{A\}$
" " B : $E_2 = \{B\}$
" " A & B : $E_3 = E_1 \cap E_2 = \phi.$
- we say events E_1 & E_2 are mutually exclusive (or disjoint. Occurence of E_1 disallows occurence of E_2. In other words if my A(B) is my guess then B(A) cannot be my guess.

Complement of an event :-
$S = \{H, T\}$.
$E_1 = \{H\}$ $\qquad S \setminus E = \{T\} @_1 = E^c$.

Def:- Consist of all outcomes in sample spaces that are not in E.

Toss a coin, $S = \{HH, TH, HT, TT\}$

Atleast one is Head. $E = \{HH, HT, TH\}$ → atleast one is a tail.

$E' = \{TT\}$ → Both are Tail

Ex: Toss a coin once

$S = \{H, T\}$

E_1 → head $E_1 = \{H\}$

E_2 → tail $E_2 = \{T\}$

E_1 and E_2 is complement of E_1 (i.e.) $E_2 = E_1^c$

* Tossing 2 coins & noting the outcomes

→ Sample space $S = \{HH, HT, TH, TT\}$

→ Event: Head on first toss $E_1 = \{HH, HT\}$

$E_1^c = \{TT, TH\}$.

Note: → E^c will occur if and only if E does not occur.

→ The complement of sample space is the null set, that is $S^c = \phi$.

Subsets.

Given two events E & F, if all outcome in E are also in F, then we say E is contained in F or E is a subset of F & $E \subset F$.

Ex: Tossing Two coins & noting the outcomes

$S = \{HH, TH, HT, TT\}$

$F = \{HH, HT\}$ → head first toss

$E = \{HH\}$ → head in both

$E \subset F$

Lec-6.3 venn diagram.

deck = 52 cards

Club → 4 clubs → 4 suit × 13 cards = 52.

Suit → 13 cards.

face cards → J, Q, K.

Exper.: Randomly select one card from deck. 80 52 cards

Sample Space S = {collection of all 52 cards}.

Application: playing cards contd.

⇒ Describe event, → card selected is king of hearts

$F = \{K\heartsuit\}$

⇒ E → card selected is a king.

$F = \{1 \times 4\}$

⇒ E → card selected is a heart

$G = \{1 \times 13\} = 13$ outcomes.

⇒ $F \cup G$ → either heart or king.

= 16 possible outcomes.

⇒ A king & A Heart. F → K

G → Heart

$F \cap G$ only one card. = E

Playing cards :

⇒ Let H be event of selecting an Ace. Ace e

E → A heart king.

F → a King

G → a heart

$F \cap G$ ⇒ a king and a heart

= E

H → selecting an Ace. (4)

H = { Ace club

A → heart

A - diamond

A - clover }

H = { A spade A club. A Heart A. diamond}

G → having a heart.

H → {Ace ...

G → {A♡, 2♡ ... }

H & G are not mutually exclusive

* G & H →

I → {relates a Queen}

F → {King ♡, K spade ....} } No common outcome ∴ mutually exclusive

Venn diagram:

A graphical representation that is useful for illustrating logical relations among events in the venn diagram

Sample space:

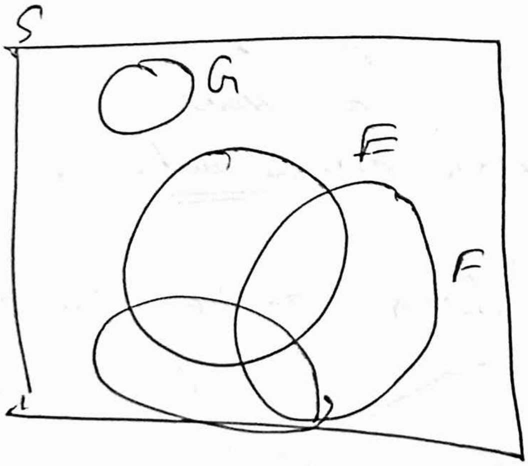
Event E

- union
- Intersection
- Complement

Representation of Sample space:

consists of all possible outcomes and is represented by a large rectangle

* event is a subset of S.



S = {HH, HT, TH, TT}. } same

S = {HH, TH, TT, HT}.

E = {HH, HT}.

F → {TH, HH}.

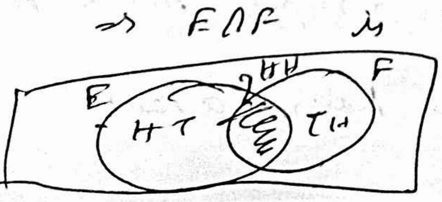
G = {TH, TT}

Representation of event: union & intersection:

⇒ Representation of event:

⇒ $E \cup F$ is entire shaded region

⇒ $E \cap F$ is the shaded in blue region



S: {HH, TH, HT, TT}

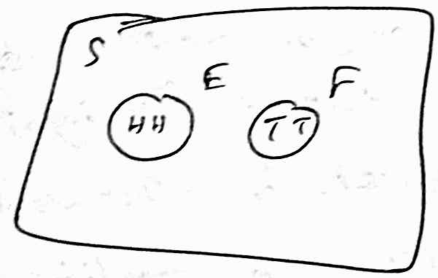
E = {HH, HT} F. {TH, HH}.

disjoint events:

$S = \{HH, T HHT; TT\}$

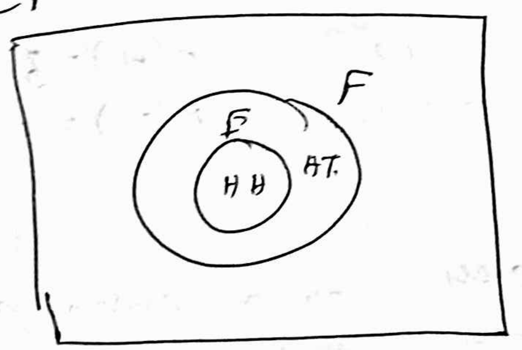
$E = \{HH\}$

$F = \{TT\}$



Subsets:

$E \subset F$



$S = \{HH, HT, TH, TT\}$

$F = \{HH, HT\}$

$E = \{HH\}$.

set of all possible basic outcomes of R.E.

$E \subseteq S.E$, Φ

Topic Summary:

(1) Introduced random experiment, sample space, event

(2) Notion of union, Intersection, complement of events.

(3) representation of sample space, events, using venn diagram.

L6.4 Properties of probability:

3 Main interpretation of probability:

classical approach (A priori (or) theoretical): Let S be the sample space of a random experiment in which there are n equally likely outcomes, and the event E consists of exactly m of these outcomes, then we say the probability of the Event E is m/n and represent it as $P(E) = \frac{m}{n}$.

6 sided die (1-6)

Fair coin

H T

Roll a die.

$S = \{1, 2, 3, 4, 5, 6\}$

Event, E = an even number: $\{2, 4, 6\}$. no. of outcomes in the event are 3.

So, the probability of $S = \{1, 2, 3, 4, 5, 6\}$ $E = \{2, 4, 6\}$.

$P(E) = P(\text{rolling an even number}) = \frac{m}{n} = \frac{3}{6} = \frac{1}{2}$.

This approach assumes all the outcomes are equally likely.

2) Relative frequency (A posteriori (or, empirical)): The probability of an event in an experiment is the proportion (or fraction) of times the event occurs in a very long (theoretically infinite) series of independent repetitions of experiment. In other words, if $n(E)$ is the no. of times E occurs in n repetitions of the experiment, $P(E) = \lim_{n \to \infty} \frac{n(E)}{n}$

eg: toss a coin n.

Trail	1	2	3	4	. . . .	10
output	H	H	T	H	T T H T T	H

$n(H) = 5$
$n(T) = 5$

3) Subjective → assigns a "best guess" by a person, making the statement of the chances, that the event will happen. The probability measures an individual's degree of belief in the event

Probability Axioms:

Consider S. Suppose that for each event E, there is a number, denoted P(E) and called the probability event E, that is in accord with the following three properties (Axioms

Sample space - S

$E \subset S$

|

$P(E)$

(i) For any event E, the probability of E is a number between 0 and 1. (ie) $0 \le P(E) \le 1$.

(ii) The probability of $S = 1$. $P(S) = 1$. (ie) outcome of any random expr will be an element of sample 'S' with probability 1.

(iii) For a sequence of mutually exclusive (disjoint) events,

E_1, E_2

$$\boxed{P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)}$$

Probability of union of disjoint events

The third property can be stated as:

The probability of union of disjoint events is equal to the sum of the probability of these events.

eg: E_1 & E_2 are disjoint.

$$P(E_1 \cup E_2) = P(E_1) + P(E_2)$$

In other words if E_1 & E_2 cannot occur simultaneously, then the probability that the outcome of the experiment is contained in either of E_1 or E_2 is equal to sum of probability that it is in E_1 and the probability that it is in E_2

eg: $S = \{1, 2, 3, 4, 5, 6\}$ → die

$$\left.\begin{aligned} E_1 &\to \text{odd} = \{1, 3, 5\} \\ E_2 &\to \text{even} = \{2, 4, 6\} \end{aligned}\right\} \text{disjoint}$$

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) = P(S) = 1$$

General properties of probability:

Properties 1, 2, 3 can be used to establish some general results concerning probabilities

(1) Probability of complement of an event: $P(E^c) = 1 - P(E)$

⇒ (1) E & E^c → Disjoint / mutually exclusive

⇒ (2) $E \cup E^c = S$

⇒ (3) $P(E \cup E^c) = P(E) + P(E^c)$ → Axiom 3.

RHS - $P(S) = 1$

$$P(E) + P(E^c) = 1$$

$$P(E^c) = 1 - P(E)$$

① Probability of complement of an event: $P(E^c) = 1 - P(E)$

E and E^c are disjoint. Also $E \cup E^c = S$

⇒ Apply property 3 to LHS $P(E \cup E^c) = P(E) + P(E^c)$

⇒ Apply property 2 to RHS $P(S) = 1$.

→ Equating both we get

$$P(E \cup E^c) = P(E) + P(E^c) = P(S) = 1$$

Hence $P(E^c) = 1 - P(E)$.

② ϕ - null event → no outcome

$$P(\phi) = 0.$$

→ $S^c = \phi$.

→ Apply the above property, $P(S^c) = 1 - P(S)$

⇒ Apply property 2: $P(S) = 1$

⇒ Hence $P(\phi) = 0$.

Addition rule of probability:

The following formula relates to probability of the union of events E_1 and E_2 which are not necessarily disjoint to $P(E_1)$, $P(E_2)$ and the probability of intersection of E_1 and E_2 → It is called as addition rule of probability.

For events E_1 & E_2

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) \cancel{+P(E_3)} - P(E_1 \cap E_2).$$

eg: E_1 = King of ♡ - 1

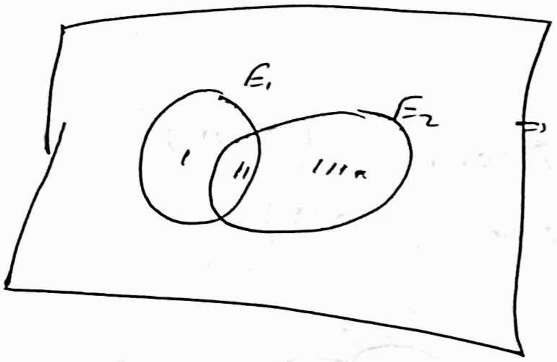
E_2 = hearts → 13

E_3 → King - 4.

what is P, if I have either a heart or a king. (not exclusive)

Proof of Addition rule:

S



$\Rightarrow E_1 \cup E_2 = I \cup II \cup III$

$E_1 = I \cup II$

$E_2 = II \cup III$

$E_1 \cap E_2 = II$

$$P(E_1 \cup E_2 \cup E_3) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$$

$$E_1 \cup E = \underbrace{I \cup II \cup III}$$

3 disjoint sets

$$E_1 = I \cup II$$

$$P(E_1 \cup E_2) = P(I \cup II \cup III) \rightarrow P(I) + P(II) + P(III) \rightarrow ①$$

$$P(E_1) = P(I \cup II) \rightarrow P(I) + P(II) \rightarrow ②$$

$$P(E_2) = P(II \cup III) \rightarrow P(II + P(III) \rightarrow ③$$

$$P(E_1 \cap E_2) = P(II) \longrightarrow ④$$

$$P(III) = P(E_2 - P(II)$$

$$P(E_1 \cup E_2) = P(E_1 + P(E_2) - P(E_1 \cap E_2)$$

Lec. 6.5 – Probability – Applications

eg: shopping shirts & pants:

shirt – 0.3 = $P(S)$

pant = 0.2 = $P(P)$

both = 0.1 = $P(S \cap P)$

∴ neither a shirt nor a pant? $(S \cup P)^c$ $P(S \cup P)' = 1 - 0.4 = 0.6$.

either a pant or shirt = $P(S \cup P)$.

$P(S \cup P) = P(S) + P(P) - P(S \cap P) = 0.3 + 0.2 - 0.1 = 0.4$

$P(S \cup P)^c = 1 - 0.4 = \boxed{0.6}$

eg: Subject grades

$P(S) = 0.4$

$P(M) = 0.6$

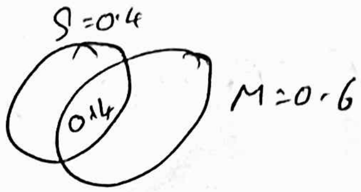
$\cancel{S \cap M} = 0.86$

$P(S \cup M)$

(i) Does not receive A in S (or) M. $= (S \cup M)^c$

(ii) Receive A in both S & M $= (S \cap M)$

Ans :



$S \cup M = 0.4 + 0.6 - S \cap M$

$\cancel{0.86} =$

$S \cap M = 0.4 + 0.6 - 0.86$

(ii) $S \cap M = 0.14 \rightarrow$

(i) 0.14

L-6.6 → Equally likely outcomes:

+ it is natural to assume that each outcome in the sample space S is equally likely to occur.

+ S → consists of N outcomes, say $S = \{1, 2, \dots N\}$.

$\{1\} \{2\} \dots \{N\}$

$S = \{H, T\}$

$\{H\} \{T\}$.

$S = \{1, 2, 3, 4, 5, 6\}$

$\{1\} \{2\} \{3\} \{4\} \{5\} \{6\}$

+ $P(\{i\})$ is the probability of event consisting of the single outcome i.

$S = \{1, 2\}$ $E_1 = \{1\}$ $E_2 = \{2\}$

$E_1 \cup E_2 = S$ $P(S) = 1$. (Axiom-2)

$P(E_1) = P(E_2)$ → Equally likely outcomes

Axiom - (3)

$P(E_1) \cup P(E_2) = P(E_1) + P(E_2)$.

$= P(E) + P(E) = 2P(E)$

$2P(E) = 1$

$P(E) = 1/2$

$P(E_1) = P(E_2) = 1/2$

$E_1 = \{1\}, \quad E_2 = \{2\} \quad E_3 = \{3\}$

$P(E) = 1/3.$

$P(E_1) = P(E_2) = P(E_3) = 1/3.$

$\Rightarrow \cancel{P(\ )} \quad \boxed{P(E) = 1/n}$

$A = \{1, 2, 3\} \qquad P(A)$

$A = \{1\} \cup \{2\} \cup \{3\}$

$P(A) = P[\{1\}] + P\{2\} + P(\{3\}) \quad \cup \{5\}$

$= 1/N + 1/N + 1/N = 3/N + 1/N = \frac{4}{N}.$

$\Rightarrow$ foregoing implies that the probability of any event A is equal to the proportion of the outcomes in the sample space that is in A.

(ie) $P(A) = \dfrac{\text{No. of outcomes in } S \text{ that are in } A}{N}$

eg: Rolling a dice.

$\rightarrow$ Experiment: Roll a fair dice.

$\Rightarrow$ Sample Space; $S = \{1, 2, 3, 4, 5, 6\}$ $\rightarrow$ any one of these outcomes are equally likely to happen.

$P(E_i) = 1/6 \qquad P(\{1\}) = P(\{2\}) = P(\{6\}) = 1/6.$

$\rightarrow$ Define A to be the event the outcome is odd $A = \{1, 3, 5\}$.

$P(A) = 3/6 = 1/2$

$P(A) = P(E_1) + P(E_2) + P(E_5) = 1/6 + 1/6 + 1/6 = 1/2, \quad \frac{n(A)}{N} = \frac{3}{6} = \frac{1}{2}$

$A = E_1 \cup E_2 \cup E_3$

$P(A) = 1/2$

⇒ Let B be the event that the outcome is greater than 4.

$B = \{5, 6\}$.

$P(B) = P(E_5 \cup E_6) = 1/6 + 1/6 = 2/6 = \frac{1}{3}$.

$A = \{1, 3, 5\}$
$B = \{5, 6\}$

$\frac{n(B)}{N} = \frac{2}{6} = 1/3$.

⇒ Let C be event that the outcome is either odd or greater than 4.

$P(C) = P(A \cup B) = P(A) + P(B) - P(A \cap B) = 3/6 + 2/6 - 1/6 = 4/6$

C = Either odd or > 4. (Not mutually exclusive)

$A \cup B = \{1, 3, 5\}$
$\{5, 6\}$

Eg: Playing cards:

Probability card is either red or a queen?

$S = 52$, each is equally likely. $= 1/52$.

$R \to$ Red $= 26 \Rightarrow P(R) = \frac{26}{52} = 1/2$

$Q \to P(Q) = 4/52 = 1/13$.

$P(R \cup Q)$ Not mutually exclusive

$= P(R) + P(Q) - P(R \cap Q)$

$= \frac{26}{52} + \frac{4}{52} - \frac{2}{52}$

$= \frac{28}{52} = \frac{7}{13}$

~~Eg~~ Section Summary:

(1) Interpretations of probability → classical, Frequency, Subjective

(2) Probability axioms → $0 \le P(E) \le 1$; $P(S) = 1$; $P(\cup) = \sum_{i=1}^{\infty} P(E)$

Not necessarily disjoint → (3) Addition rule of probability.

$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$.

When outcomes "equally likely" (E).

1. **Conditional Probability:** We restrict the sample space to a row or column.

Ex: $$P(\text{Female} \mid \text{Doesn't own a phone}) = \frac{10}{24}$$

$$\Updownarrow$$

$$\frac{P(\text{Female} \cap \text{Doesn't own a phone})}{P(\text{Doesn't own a phone})} = \frac{\frac{10}{100}}{\frac{24}{100}}$$

2. **Why conditional probabilities?**

We are often interested in determining probabilities when some partial information concerning the outcome of the experiment is available. In such situations, the probabilities are called conditional probabilities.

3. If each outcome of a finite sample space S is equally likely, then, conditional on the event that the outcome lies in a subset F, all outcomes in F become equally likely. In such cases, it is often convenient to compute conditional probabilities of the form $P(E|F)$ by using F as the sample space.

4. **Conditional Probability Formula:** The probability that event E occurs given that event F occurs (or conditional on event F occurring) is given by

$$P(E|F) = \frac{P(E \cap F)}{P(F)} \; ; \; P(F) > 0$$

↳ non-null subset.

Ex:- Let E denote the event that the sum of the dice is 10 & let F denote the event that the first die lands on 4, then the probability obtained is called the conditional probability of E given that F has occurred. It is denoted by $P(E|F)$.

(5) $P(E \cap F) = P(F) \times P(E|F)$

(6) Generalized multiplication rule :

$$P(E_1 \cap E_2 \cap E_3 \cap \ldots \cap E_n) = P(E_1)\, P(E_2|E_1)\, P(E_3|E_1 \cap E_2) \ldots\ldots P(E_n | E_1 \cap E_2 \ldots E_{n-1})$$

(7) $P(E_1 \cap E_2 \cap E_3 \cap E_4) = P(E_1) \times P(E_2|E_1) \times P(E_3 | E_1 \cap E_2) \times P(E_4 | E_1 \cap E_2 \cap E_3).$

(8) In the case where E is independent of event F if knowing whether F occurs does not affect the probability of E.

$$\therefore P(E|F) = P(E)$$

$$P(E \cap F) = P(F) \times P(E|F)$$

$$= P(F) \times P(E)$$

$$\therefore P(E \cap F) = P(E) \times P(F)$$

(9) In other words, two events are independent if & only if the probability that both occur equals the product of their individual probabilities

(10) Independence of E and F^c :

$$P(E \cap F^c) = P(E) \times P(F^c)$$

If E is independent of F, then the probability of E's occurrence is unchanged by information as to whether or not F has occurred.

(11) E, F, and G are said to be independent if

$P(E \cap F \cap G) = P(E) \times P(F) \times P(G)$

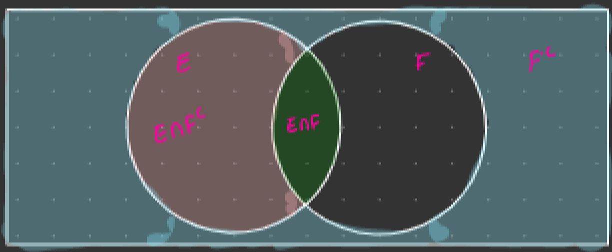
and

$P(E \cap F) = P(E) \times P(F)$

$$P(E \cap G) = P(E) \times P(G)$$

$$P(F \cap G) = P(F) \times P(G)$$

(12) Law of Total probability :



→ E can be expressed as $(E \cap F) \cup (E \cap F^c)$.

→ In other words, for, in order for an outcome to be in E, it must either be in both E & F or be in E but not in F.

$$P(E) = P(E \cap F) + P(E \cap F^c)$$

$$P(E) = P(F)\,P(E|F) + P(F^c)\,P(E|F^c) \quad \text{—— (1)}$$

→ Equ. (1) states that the probability of event E is a weighted average of the conditional probability of E given that F occurs & the conditional probability of E given that F does not occur.

→ Each conditional probability is weighted by the probability of the event on which it is conditioned.

(13) **Rule of total probability:** Suppose that events $F_1, F_2, F_3, \ldots F_k$ are mutually exclusive and exhaustive ; that is, exactly one of the events must occur. Then for any event E,

$$P(E) = P(E|F_1)\,P(F_1) + P(E|F_2)\,P(F_2) + \cdots + P(E|F_k)\,P(F_k)$$

$$P(E_i) = \sum_{i=1}^{k} P(E|F_i)\,P(F_i)$$

14. Baye's Rule:

Suppose we are now interested in the conditional probability of event F conditioned on E. We know,

$$P(F|E) = \frac{P(F \cap E)}{P(E)}$$

From definition,

$$P(F|E) = \frac{P(F \cap E)}{P(E)} = \frac{P(E|F)\,P(F)}{P(F)\,P(E|F) + P(F^c)\,P(E|F^c)}$$

Suppose that events $F_1, F_2, \ldots F_k$ are mutually exclusive & exhaustive; then for any event E,

$$P(F_i|E) = \frac{P(E|F_i)\,P(F_i)}{\sum_{i=1}^{k} P(E|F_i)\,P(F_i)}$$

15. When A & B are independent events,

$$P(A \cup B) = P(A)\,P(B^c) + P(B)$$

A^c & B^c are independent.

WEEK 8

1. **Random Variable:** When a probability experiment is performed, often we are not interested in all the details of the experimental result, but rather are interested in the value of some numerical quantity determined by the result.

   Ex: We may be interested in knowing the sum is 7 & may not be concerned over whether the actual outcome was (1,6), (2,5), (3,4), (4,3), (5,2) or (6,1)

   These quantities of interest, or, more formally, these real-valued functions defined on the sample space, are known as "random variables."

   We may assign probabilities to the possible values of the random variable.

2. **Discrete Random variable:** A random variable that can take on at most a countable number of possible values. Thus, any random variable that can take on only a finite no. or countably infinite no. of different values is discrete.

3. **Continuous random variable:** When outcomes for random events are numerical, but cannot be counted & are infinitely divisible.

④ **Probability Mass Function:** for a discrete random variable X, we define the probability mass function $p(x)$ of X by

$$p(x_i) = P(X = x_i)$$

i.e., $p(x_1)$ for $X = x_1$

$p(x_2)$ for $X = x_2$

$\vdots$

⑤ **Properties of PMF:** The probability mass function $p(x)$ is positive for at most a countable number of values of x. That is, if X must assume one of the values $x_1, x_2, x_3, \ldots$ then

1. $p(x_i) \geq 0$, $i = 1, 2, \ldots$
2. $p(x) = 0$ for all other values of x.

Ex:- Let $X = \{1, 2, 3\}$

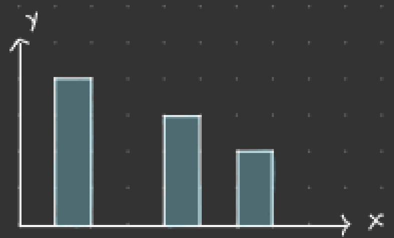
$\therefore p(1) + p(2) + p(3) = 1$

and $p(4) = p(5) = \cdots = 0$.

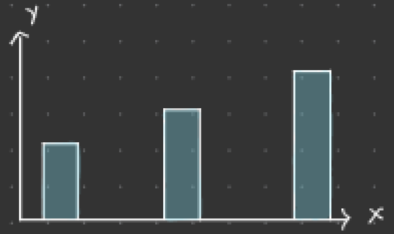
→ Since X must take one of the values x_i, we have

$$\sum_{i=1}^{\infty} p(x_i) = 1$$

⑥ **Positived skewed distribution:**



① Negative Skewed distribution:



② **Cumulative Distribution Function:** The cumulative distribution function (cdf), F, can be expressed:

$$F(a) = P(X \leq a)$$

If X is a discrete random variable whose possible values are $x_1, x_2, x_3, \ldots$ where $x_1 < x_2 < x_3 \ldots$, then the distribution function F of X is a step function.

③ The size of the step at any of the values 1, 2, 3, and 4 is equal to the probability that X assumes that particular value.

WEEK 9

① **Expectation of a random variable:** Let X be a discrete random variable taking values $x_1, x_2, x_3, \dots$ The expected value of X denoted by $E(X)$ and referred to as Expectation of X is given by:

$$E(X) = \sum_{i=1}^{\infty} x_i \times P(X = x_i)$$

② The Expectation of a random variable can be considered the "long-run-average" value of the random variable in repeated independent observations.

③ The expected value tells us what we would expect the avg. of a large number of rolls to be in the long run.

④ The expected value of X is a theoretical average.

⑤ **Bernoulli random variable:** A random variable that takes on either the value 1 or 0 is called a Bernoulli random variable.

→ Let X be a Bernoulli random variable that takes on the value 1 with probability 'p'.

→ The prob. distribution of the random variable is:

X	0	1
$P(X=x_i)$	$1-P$	P

If $P = \frac{1}{2}$, then $E(X) = \frac{1}{2}$.

$\therefore E(X) = 0 \times (1-p) + 1 \times p = P$

⑥ Discrete uniform random variable:

→ Let X be a random variable that is equally likely to take any of the values $1, 2, \ldots n$.

→ p.m.f:

X	1	2	....	n
$P(X=x_i)$	$\frac{1}{n}$	$\frac{1}{n}$	...	$\frac{1}{n}$

$$\therefore E(X) = 1 \times \frac{1}{n} + 2 \times \frac{1}{n} + \cdots + n \times \frac{1}{n} = \frac{1}{n}\left[1+2+\cdots+n\right]$$

$$= \frac{1}{\cancel{n}}\left[\frac{\cancel{n}(n+1)}{2}\right]$$

$$\therefore E(X) = \frac{n+1}{2}$$

⑦ Relative frequency of random variable converges to the probability of random variable in experiment of tossing a die if no. of observations are infinite.

⑧ If no. of obsvs are infinite & $r(X=x_i)$ is the relative frequency of $X=x_i$, $E(X) = \sum_{i=1}^{n} x_i \cdot r(X=x_i)$

⑨ Expectation of a function of a random variable: Let X be a discrete random variable which takes values x_i along with its probability mass function, $P(X=x_i)$. Let 'g' be any real valued function, the expected value of $g(X)$ is:

$$E(g(X)) = \sum_i g(x_i)\, P(X=x_i)$$

Corollary: If 'a' and 'b' are constants, $E(aX+b) = aE(X) + b$

(10) $E(X^2) \neq (E(X))^2$

(11) The expected value of the sum of random variables is equal to the sum of the individual expected values, i.e., let X & Y be two random variables. Then,

$$E(X+Y) = E(X) + E(Y)$$

(12) **Hypergeometric random variable:** Suppose that a sample of size 'n' is to be chosen randomly (w/o replacement) from a box containing N balls, of which 'm' are red and 'N-m' are blue.

Let X denote the no. of red balls selected, then

$$P(X=i) = \frac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}}, \quad i = 0, 1, 2, \ldots, n$$

→ X is said to be a hypergeometric variable for some values of n, m, and N.

(13) $E(X) = \frac{mn}{N}$

(14) Let $X_1, X_2, X_3, \ldots X_k$ be 'k' discrete random variables. Then,

$$E\left(\sum_{i=1}^{k} X_i\right) = \sum_{i=1}^{k} E(X_i)$$

(15) **Variance of a function of a random variable:**

- Let X be a random variable, let 'c' be a constant, then

  → $Var(cX) = c^2 Var(X)$

  → $Var(X+c) = Var(X)$

(16) If a & b are constants, $V(aX+b) = a^2\,Var(X)$

(17) Independent random variables: Random variables X & Y are independent if knowing the value of one of them does not change the probabilities of the other.

(18) Variance of sum of independent random variables is given by:

$$Var(X+Y) = Var(X) + Var(Y)$$

(19) For a hypergeometric random variable,

$$Var(X) = \frac{nm}{N}\left[\frac{(n-1)(m-1)}{(N-1)} + 1 - \frac{nm}{N}\right]$$

(20) Standard deviation of a random variable: The quantity $SD(X) = \sqrt{Var(X)}$ is called the standard deviation of X.

(21) The SD, like the expected value, is measured in the same units as is the random value.

(22) Properties of Standard deviation: Let X be a random variable, let 'c' be a constant, then

$$SD(cX) = c\,SD(X)$$

$$SD(X+c) = SD(X)$$

(23)

	Expected	Variance	S.D.
X	$E(X)$	$V(X)$	$SD(X)$
cX	$c\,E(X)$	$c^2\,V(X)$	$c\,SD(X)$
$X+c$	$E(X)+c$	$V(X)$	$SD(X)$

(21) Hypergeometric random variable:

$$P(X=i) = \frac{{}^{m}C_i \times {}^{N-m}C_{n-i}}{{}^{N}C_n} \quad , \; i = 0, 1, 2, 3, \ldots, n$$