Akash Kumar

Statistics

Data:

Data are fact and figure that are collected analyzed and summerized for presentation and interpretation.

Statistics is art of learning from data

- Descriptive statistics:

It is part of statistics that is concerned with discription and summerization of data. In this we explore data for purpose of analysis.

- Inference statistics:

It is concerned with drawing conclusion from data.

- [ ] Population and Sample:

- Population:

Population is total collection of all elements that we are intrested in is called population.

- Sample:

The subgroup of the population that will be studied in detail is called sample.

- [ ] Structure and unstructured data:

- Structured data:

These are data which are organised

in predefined fashion.

- Unstructured data :
  These are data which are not organised in pre-defined fashion or lack data model.

☐ Variable and cases :

- Cases : It is a unit for which data is collected.

- Variable : It is characteristic that varies across all units.

  Eg : case (each student)
  Variable (name, marks, board, gender...)

Data
- Categorical
  Qualitative variable
- Numerical
  Quantitative Variable
  - Descrete
    in float (measure)
  - Continous
    in intiger (count)

☐ Scale of measurement :-

→ Nominal,
→ Ordinal, } Categorical Data
→ Interval, and
→ Ratio } Numerical Data

○ Nominal scale :

When data is consist of labels or name as characteristic of observation is known as nominal scale.

Eg : Name, Board, Gender, Blood grp, etc.
Weather → comfertable, ok or uncomfertable.

→ Some nominal variable can be numerically coaded. (like M/F = 0/1)
→ It doesn't have any order.

○ Ordinal scale :

When data exhibits property of nominal data but here rank are meaningful, this scale of measurement is considered as ordinal scale.

Eg : Excellent, good or poor.
Cold, warm or hot.

→ In this diff. b/w Excellent to good may not equal to diff. b/w good to poor.

○ Interval scale of measurement :

It has all the property of ordinal data and in this interval b/w values is expressed in term of a fixed unit of measure, than scale of measure is called interval scale.

→ Interval data are always numeric.
can find diff. b/w 2 value.

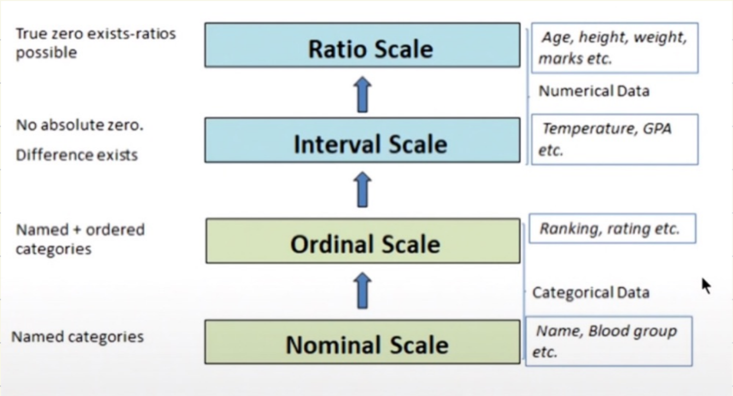
Eg → Temperature 40°c, 20°c, etc.

Here Ratio has no meaning and you, can't tell 40°C is twice hot as 20°C.

○ Ratio scale:

It has all property of interval scale and here ratio of 2 scale are meaningful, then scale of measurement is called ratio scale.

Eg → Height, marks, Run, weight, etc.



○ Categorical data:

☐ Frequency distribution:

Frequency distribution of qualitative data is listing of distinct value and their frequency.

- **Relative frequency**: The ratio of frequency to total no. of observation. It is generaly used to compare 2 data set.

eg:-

AAABBCCCCD

Category	tally	Freq	RF
A	|||	3	.3
B	||	2	.2
C	||||	4	.4
D			1	.1

Total 10

AAAAABCCCCCCDDD

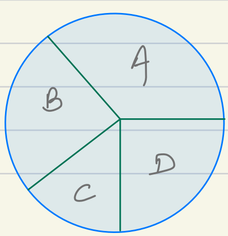
Category	tally	Freq	RF
A	~~||||~~	5	0.3
B			1	0.06
C	~~||||~~		6	0.4
D	|||	3	0.2

Total 15

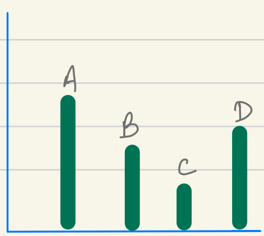
Charts of categorical data:

Barchart and piechart are 2 most common display of categorical variable.

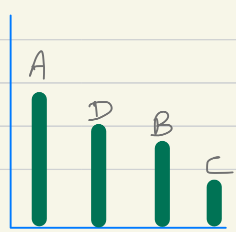
- In pie chart frequency convert in 360° angle.
  We use pie chart to know share of each cat. data.
- Bar display distinct value of qualitative data on horizontal axis with relative frequency.
  We use bar chart to compare each category to other.
- When barchart have shorted frequency it is called pareto chart.
- If catagorical variable is ordinal than chart must be in order (Pareto chart)



Pie chart



Barchart

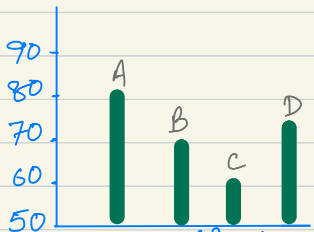


Pareto chart

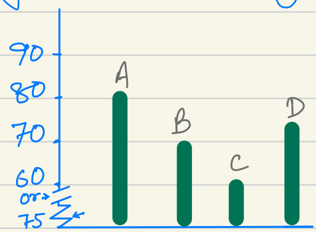
○ Misleading graph: voilate area principle.

→ Area principle say that area occupy by part of graph should correspond to amount of data represented. (Diff. bar should not be of diff. width.)

→ Trunkated graph (baseline of graph not start with zero is not good practice. You may use trunkated graph by using y-axis brake.



(misleading)



y-axis brake

☐ Measure of central tendency:

It is of 2 type in catagorical data:

Median,
Mode.

○ Mode: Mode of of a categorical column is most common category.

Eg → A, B, B, A, C, A, A, C, B → Mode = A

Mode will have largest bar or largest pie.

→ If we have 2 category with highest value than it is bimodel data

→ If there is more than 2 category than it is multimodel.

○ Median : Median is middle observation of sorted value. In ordered form.

Eg → AAABCABAACCCDABC

A A A A A A A B B B C C C C D D

median (even) — B; median (odd) — B; if added — D

It tries to divide data in 2 half.

○ Numerical data:

→ Descrete data : It is considered as count of something.

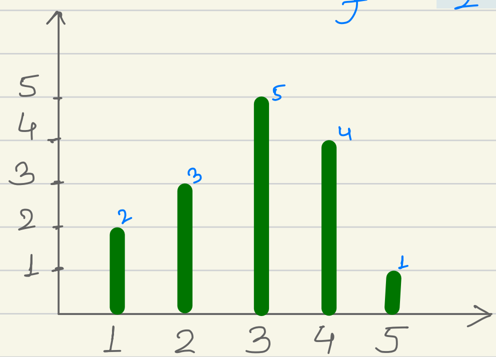
→ If descrete data is single value data than each value treated as categories.

→ By finding frequency you can know about data.

No of people in diff. house → 2,1,3,4,5,2,3,3,3,4,4,1,2,3,4

Converting in catagory :-

	1	2	3	4	5
f	2	3	5	4	1



10 — 20 ⇒ 20-10 = 10 (Class width)

Lower class — 10; Upper class — 20 ⇒ $\frac{20+10}{2}$ = 15 (Class marks)

→ Continous data: It is considered as measure of something.

→ Here organise data in number of classes to make data understandable:

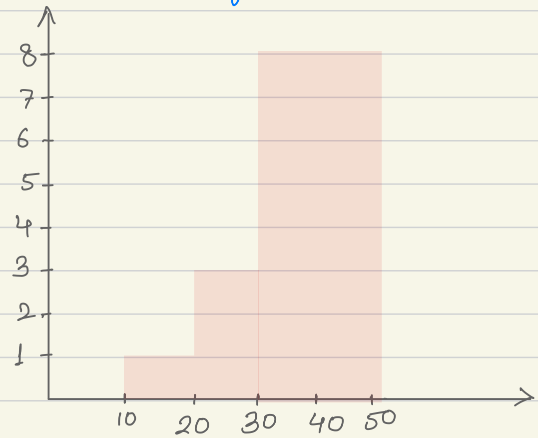
→ Each observation should belong to exactly 1 class.

- Marks is measured not counted so continous data.

Marks of 20 student = 33, 39, 30, 49, 40, 30
30, 40, 30, 11, 27, 34, 45, 48, 41, 43, 21, 47, 36, 33,

Class interval	f	Rf
10 - 20	1	·0
20 - 30	3	·05
30 - 40	8	·4
40 - 50	8	·4
Total	20	1

- Continous data is shown graphical summary through histogram.



○ Stem-and-leaf diagram:
It saperate once & 10th position of dictionary. In smallest to largest.

	Stem		Leaf
Eg → 75 =	7	|	5
75, 78 =	7	|	5, 8

Eg 2 → 15, 22, 29, 36, 31, 23, 45, 10, 25, 28, 48

Steam	leaf
1	0 5
2	23589
3	16
4	58

○ Descreptive measure:
→ Measures of central tendency: It indicate most typical value or centre data set.

→ Measure of dispersion: These measures indicate variablity / spread in data.

□ Measures of central tendency:
It capture centre or typicalness of dataset.
→ Mean
→ Median
→ Mode

○ Mean :- It is most commonly used measure. It is average means sum of all observation divided by no of observation.

$n \to$ Sample Size $\quad$ $N \to$ Population Size

→ mean reffers as average.

For descrete observation:-

▷ Sample mean $= \bar{x} = \dfrac{x_1 + x_2 + x_3 + \dots + x_n}{n}$

▷ Population mean $= \mu = \dfrac{x_1 + x_2 + x_3 + \dots + x_n}{N}$

Example →

	mean
2, 12, 5, 7, 6, 7, 3	$\bar{X} = \dfrac{42}{7} = 7$
2, 105, 5, 7, 6, 7, 3	$\bar{X} = \dfrac{135}{7} = 19.285$
2, 105, 5, 7, 6, 7	$\bar{x} = \dfrac{128}{6} = 21.33$

○

x_i	1	2	3	4	5	Total
f_i	2	3	5	4	1	15
x_if_i	2	6	15	16	5	44

$$\text{mean} = \frac{\sum_{i=n}^{n} f_i x_i}{\sum_{i=1}^{n} f_i} = \bar{X}$$

$$\text{mean} = \frac{44}{15} = 2.93$$

○ Mean for continous data:

Class interval	f	Rf	mp	fimi
10 - 20	1	.0	15	15
20 - 30	3	.05	25	75
30 - 40	8	.4	35	280
40 - 50	8	.4	45	360
Total	20	1		Σfimi 730

$$Avg = \frac{730}{20} = 36.5$$

→ 36.5 is approximation not mean bcz we are not looking to data but only seeing midpoint.

→ If you add constant to every point in dataset than your new $\bar{x}$ = old $\bar{x}$ + constant.

→ If you multiply constant to every point in dataset than your new $\bar{x}$ = old $\bar{x}$ * constant.

→ Highly affected by outlier.

○ Median :-

It is of a data set is middle value of ordared list. It is another frequently used measure of central tendency. It divide dataset in ~~top~~ 50% & botom 50%.

In ordered list: (n = no' of observation)

— If no' of observation is odd than the data in $(n+1)/2$ th value is median.

— If no' of observation is even than the median is avg. of ($\frac{n}{2}$ and $\frac{n}{2}+1$).

Example → 2, 9, 4, 6, 7, 8

In odd.

arrange data → 2, 4, 6, 7, 9

$n = 5$ median = $\frac{5+1}{2} = 3^{rd}$

Median = 3rd element = 6

In even,

2, 4, 6, 7, 8, 9

$n = 6$

median = $\frac{6}{2} = 3$, $\frac{6}{2}+1 = 4$

median = $\frac{6+7}{2} = 6.5$

- Very less affected by outlier.
- If constant is added to each point of data set, length does'n change so new median will be old median + C.
- If constant is multiplied to each data-point. Than due to seme length. median will be old median × C.

Mode :

It is most frequently occuring value of dataset.

- If no value occur more than 1 than there is no mode.

Eg:- 1, 2, 3, 7, 7, 3, 2, 1 mode = 7
2, 9, 3, 4, 2, 7 no mode.

- If constant is added to each point of data set, length does'n change so new mode will be old mode + C.
- If constant is multiplied to each data-point. Than due to seme length. mode will be old mode × C.
- Don't affect by outlier.

Let's compare 2 data:

D1 → 3, 3, 3, 3, 3 D2 = 1, 2, 3, 4, 5

→ finding measure of central tendency :

	D1	D2
Mean	3	3
Median	3	3
Mode	3	—

However mean & median are same for dataset but dataset are not same.

- **Measure of Dispersion:** To describe the above difference quantatively, we use discriptive measure that indicate amount of variation, spread in data. These is called measure of dispersion/variance/Spread.

Measure of dispersion are:
→ Range,
→ Variance,
→ Standard daviation,
→ Interquartile range.

- **Range:-** It is defined as difference b/w largest and lowest value of a dataset.

Range = Largest − Lowest

Let's compare 2 data:

D1 → 3, 3, 3, 3, 3
Range = 3−3 = 0

D2 = 1, 2, 3, 4, 5
Range = 5−1 = 4

D3 = 1, 2, 3, 4, 15
Range = 15−1 = 14

As we can se it do well in D1 & D2 but in D3 not properly.

So, Range is extremly sencitive to outlier.

- **Variance:-** In contrast to range variance takes into account all observations.
  - One way to measure variablity of dataset is consider daviation of data value from centre value. It is affected by outlier.

Eg:-

1 2 3 4 5 centre point $= 3 = \bar{X}$

-2 -1 0 1 2 diffrence from centre

Population variance:

$$\sigma^2 = \frac{(x_1-\mu)^2+(x_2-\mu)^2+\dots+(x_N-\mu)^2}{N}$$

Sample variance:

$$s^2 = \frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+\dots+(x_n-\bar{x})^2}{n-1}$$

Eg:- D1.

	Data	Deviation from mean $(x_i-\bar{x})$	Squared Deviation $(x_i-\bar{x})^2$
1	68	68-59 = 9	81
2	79	79-59 = 20	400
3	38	38-59 = -21	441
4	68	68-59 = 9	81
5	35	35-59 = -24	576
6	70	70-59 = 11	121
7	61	61-59 = 2	4
8	47	47-59 = -12	144
9	58	58-59 = -1	1
10	66	66-59 = -7	49
Total	590	0	1898

Population variance $= \frac{1898}{10} = 189.8$

Sample variance $= \frac{1898}{9} = 210.88$

- If a constant is added to all data point than the variance doesn't change.
- If a constant is multiplied to all data point than new variance = old variance × (constant)^2

- **Standard deviation:** It is square root of variance.
  - Square root of sample variance is sample standard deviation.
  - Square root of population variance is population standard deviation.

Population standard deviation:

$$\sigma = \sqrt{\frac{(x_1-\mu)^2+(x_2-\mu)^2+\dots+(x_N-\mu)^2}{N}}$$

Sample standard deviation:

$$s = \sqrt{\frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+\dots+(x_n-\bar{x})^2}{n-1}}$$

D1. Population std. deviation $= \sqrt{189.8}$

Sample std. deviation $= \sqrt{210.88}$

Unit of standard deviation:

$u^2 \rightarrow$ Variance is recorded as unit^2

$u \rightarrow$ Standard deviation by root convert it back

- If a constant is added to all data point than the standard deviation doesn't change.
- If a constant is multiplied to all data point than new ST deviation = old ST deviation × constant.
- It is affected by outlier.

- **Percentile:-** Percentile indicates the percent of distribution of data.

- To find percentile:-
  - Arrange data in ascending order.
  - Divide data into half if n is even we get 'integer' if it is odd we don't get integer. Than determine smallest integer greater than np. (np = middle of that part)
  - It is value in that position of 100p percentile.
  - But if np is integer in case of even. than take avg of np and np+1.

Example:- Arranged data:

35, 38, 47, 58, 61, 66, 68, 68, 70, 79

P	np		
0.1	1	(35+38)/2	= 36.5
0.25	2.5		= 47
0.5	5	(61+66)/2	= 63.5
0.75	7.5		= 68
1	10		= 79

→ Quartile:

- Sample 25th percentile is called 1st quartile.
- The sample 50th percentile is called median / Second quartile.
- The sample 75th percentile is called 3rd quartile.

The five number summary:
- Minimum
- 1st quartile (lower)
- 2nd quartile (median)
- 3rd quartile (upper)
- Maximum

→ Interquartile range:

The interquartile range (IQR) is diffrence b/w first and third quartile.

$$IQR = Q_3 - Q_1$$

Eg →

1st quartile, $Q1 = 49.75$
3rd quartile, $Q3 = 68$
IQR $(Q_3 - Q_1) = 18.25$

It is also measure of dispersion.

- Contingency table: It is also called two-way frequency table, is a tabular mechanism with atleast two row & two column used in statistics to present categorical in term of frequency count.

Eg:

Ownes phone

Gender	No	Yes	Total
Male	10	34	44
Female	14	42	56
Total	24	76	100

nominal data

By income

Income (coded)	Yes	No	Total
1	2	18	20
2	27	39	66
3	9	5	14
Total	38	62	100

Ordinal data

→ Row relative frequencies: It is dividing each row by its row total.

Eg:

Ownes phone

Gender	No	Yes	Total
Male	10/44	34/44	44
Female	14/56	42/56	56
Total	24/100	76/100	100

nominal data

⇒

Gender	No	Yes	Total
Male	22.7%	77.2%	44
Female	25.0%	75.0%	56
Total	24.0%	76.0%	100

→ Column relative frequencies: It is dividing each column by its row total.

Ownes phone

Gender	No	Yes	Total
Male	10/24	34/76	44/100
Female	14/24	42/76	56/100
Total	24	76	100

nominal data

Gender	No	Yes	Total
Male	41.6%	44.7%	44.0%
Female	58.3%	55.2%	56.0%
Total	24	76	100

→ Association b/w two variable:
It is finding whether information about one variable provide information about another variable.

→ If the row or column relative frequency are same for all rows than two variable are not associated with each other.

Row related frequency

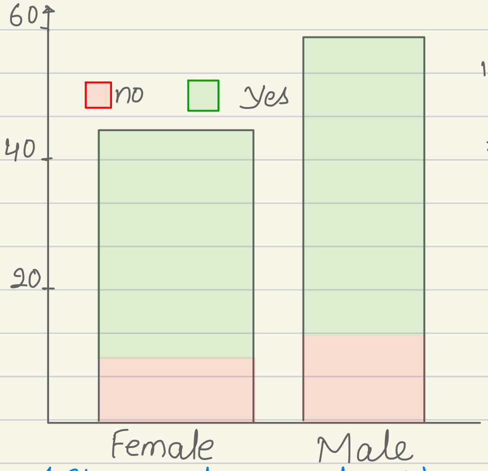
Gender	No	Yes	Total
Male	22.7%	77.2%	44
Female	25.0%	75.0%	56
Total	24.0%	76.0%	100

Both total % are for male % & female %

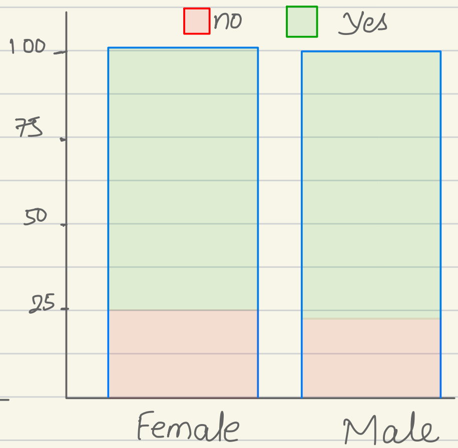
Column related frequency

Gender	No	Yes	Total
Male	41.6%	44.7%	44.0%
Female	58.3%	55.2%	56.0%
Total	24	76	100

Both male & female % are similar to total.



(Standard barchart)



(100% stacked barchart)

→ Stacked barchart :- It summarise data in form of barchart with proportion in respect to category.

→ If the row or column relative frequency are diffrent for some rows than two variable are associated with each other.

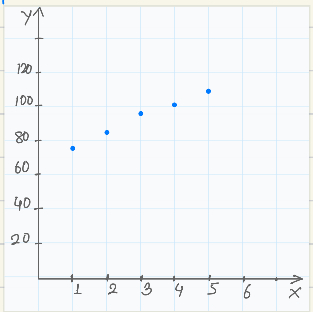
Income level	Yes	No	total
High	10.0%	90.0%	20
Medium	40.9%	59.1%	66
Low	64.1%	35.9%	14
Total	38.0%	62.0%	100

Here, in both row & column freq. are diff. from total so, Income & phone are connected.

Income level	Yes	No	total
High	5.2 %	29.0 %	20.0%
Medium	71.0 %	62.9%	66.0%
Low	23.6 %	8.1 %	14.0%
Total	38	62	100

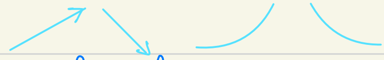
- Scatter plot: It is a graph that displays pair of values as points on 2-D plane.

Age	Height
1	75
2	85
3	94
4	101
5	108

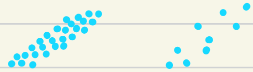


You can describe association b/w variables in scatter plot by answering 4 question:

→ Direction → Does pattern Up, Down or both.



→ Curvature → Is it linear or curve.



→ Variation → tightly clustered → space b/w points.

→ Outliers → Outside pattern (exceptions).

○ Measure of association:
Strength of association b/w 2 variable can be measured with: – Covariance
– Correlation

○ Covariance :- It quantify strength of a linear relation b/w 2 numeric variables. (No units)

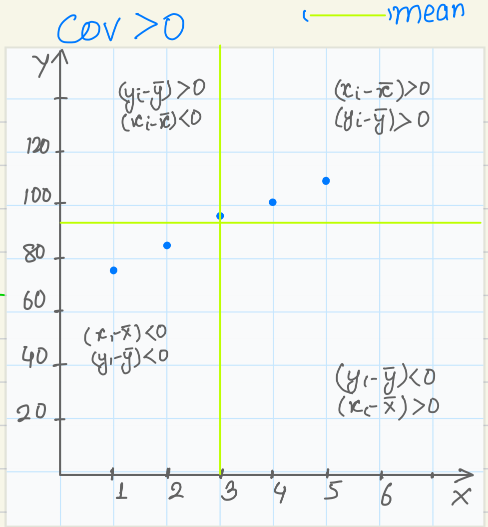
Age	Height	Deviation x $(x_i - \bar{x})$	Deviation y $(y_i - \bar{y})$	$(x_i-\bar{x})(y_i-\bar{y})$
1	75	-2	-17.6	35.2
2	85	-1	-7.6	7.6
3	94	0	1.4	0
4	101	1	8.4	8.4
5	108	2	15.4	30.8
$\bar{x}=3$	$\bar{y}=92.6$			$\Sigma = 82$

①

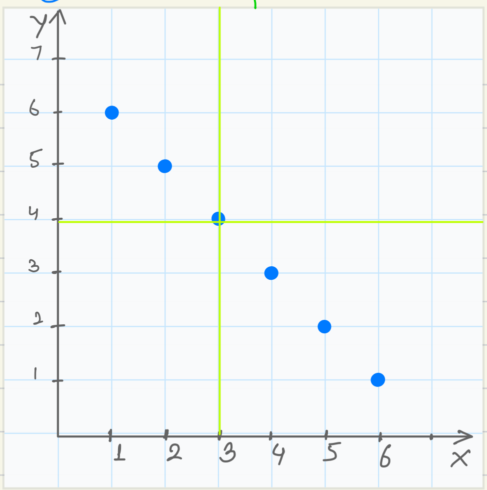
Age	Height	Deviation x	Deviation y	$(x_i-\bar{x})(y_i-\bar{y})$
1	6	-2	2	-4
2	5	-1	1	-1
3	4	0	0	0
4	3	1	-1	-1
5	2	2	-2	-4
$\bar{x}=3$	$\bar{y}=4$			$\Sigma = -10$

②

Cov > 0



Cov < 0



→ When large X is associated with large Y and vice-versa the deviation sign will be same

→ When large X is associated with small Y and vice-versa the deviation sign will diffrent..

The covariance b/w variable x & y is given by :-

- Population covariance :

$$Cov(x,y) = \frac{\sum_{i=1}^{n} (x-\bar{x})(y-\bar{y})}{N}$$

- Sample covariance :

$$Cov(x,y) = \frac{\sum_{i=1}^{n} (x-\bar{x})(y-\bar{y})}{n-1}$$

1. Population covariance $= \frac{82}{5} = 16.4$

   Sample covariance $= \frac{82}{4} = 20.5$

2. Pop. cov. $= \frac{-10}{5} = -2$

   Sam. cov. $= \frac{-10}{4} = -2.5$

- Correlation: It is more easily intepreted measure of linear association b/w two numerical variable. (No units)
  - → It is derived from covariance.
  - → To find correlation b/w two variable x & y divide covariance b/w x & y by product of standard deviation of x & y.

→ $\text{Correlation} = \dfrac{\text{Covariance}}{\text{Standard Deviation}}$

$$\rightarrow \text{Correlation} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}} = \frac{Cov(x,y)}{S_x S_y}$$

Age	Height	Sq. Dev. of X $(X_i-\bar{X})^2$	Sq. Dev. of Y $(y_i-\bar{y})^2$	$(x_i-\bar{x})(y_i-\bar{y})$
1	75	$-2^2 = 4$	$-17.6^2 = 309.76$	35.2
2	85	$-1^2 = 1$	$-7.6^2 = 57.76$	7.6
3	94	$0^2 = 0$	$1.4^2 = 1.96$	0
4	101	$1^2 = 1$	$8.4^2 = 70.56$	8.4
5	108	$2^2 = 4$	$15.4^2 = 237.16$	30.8
$\bar{X} = 3$	$\bar{Y} = 92.6$	$\Sigma 10$	$\Sigma = 677.2$	$\Sigma = 82$

$SD_x = 1.58$, $SD_y = 13.01$, covarriance $= 20.5$

$$\text{Correlation} = \frac{Cov(x,y)}{S_x S_y} = \frac{82}{\sqrt{10 \times 677.2}} \text{ or } \frac{20.5}{1.58 \times 13.01} = 0.9964$$

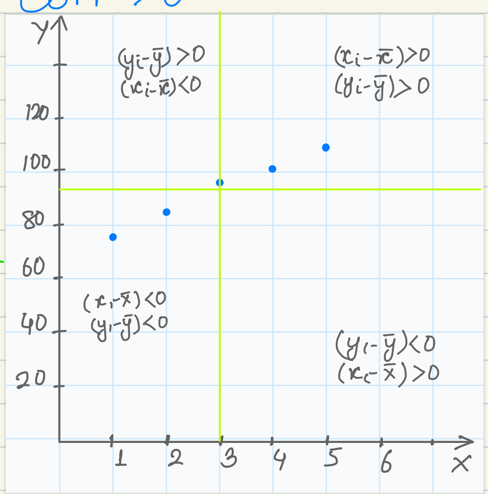
Age	Height	Sq. Dev. of X $(x_i-\bar{x})^2$	Sq. Dev. of Y $(y_i-\bar{y})^2$	$(x_i-\bar{x})(y_i-\bar{y})$
1	6	$-2^2 = 4$	$2^2 = 4$	-4
2	5	$-1^2 = 1$	$1^2 = 1$	-1
3	4	$0^2 = 0$	$0^2 = 0$	0
4	3	$1^2 = 1$	$-1^2 = 1$	-1
5	2	$2^2 = 4$	$-2^2 = 4$	-4
$\bar{x} = 3$	$\bar{y} = 4$	$\Sigma = 10$	$\Sigma\ 10$	$\Sigma = -10$

$SD_x = 1.58$, $SD_y = 1.58$, covarriance =

$$\text{Correlation} = \frac{Cov(x,y)}{S_x S_y} = \frac{-10}{\sqrt{10} \times \sqrt{10}} \text{ or } \frac{-2.5}{1.58 \times 1.58} = -1$$

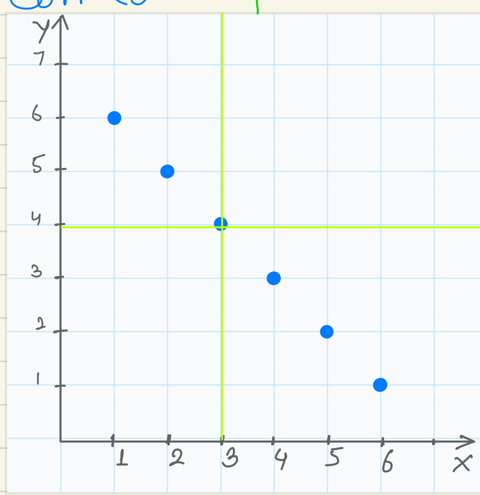
Corr > 0 — mean

1



Corr < 0

2



These linear relation can be summerised through line.

Point Bi-serial Correlation Coefficient :

→ Here we group our data (one numerical and one categorical column and encode our categorical column. eg → male = 0, female = 1)

→ Compute mean value of numerical column in respect to encoded cat. column.
eg → mean of marks of male/Female student.

→ P_0 and P_1 is proportion of group. Eg:-
For 20 student male = 12/20 , female = 8/20

So, Correlation Coffecient :

$$r_{pb} = \left(\frac{\bar{Y}_0 - \bar{Y}_1}{S_x} \right) \sqrt{P_0 P_1}$$

PERMUTATION & COMBINATION

- Adding rule of counting :

If action A can occur in n_1 different ways, another action B can occur n_2 diff. ways, than total number of occurance of action A or B is $N_1 + N_2$.

Eg : There are 4 shirt and 5 paints and we have to select either 1 shirt or 1 paint than no. of possiblity to choose will be $5+4 = 10$ ways.

- Multiplicate rule of counting :

If action A can occur in n_1 different ways, another action B can occur n_2 diff. ways, than total number of occurance of action A and B is $N_1 \times N_2$.

Suppose that 'r' action are there than there are $N_1 + N_2 + \dots + N_r$ possiblity altogether for 'r' actions.

Eg : There are 4 shirt and 5 paints and we have to select either 1 shirt & 1 paint than no. of possiblity to choose will be $5 \times 4 = 20$. ways.
If we add 2 diff. color shoe than it will $5 \times 4 \times 2 = 40$.

Eg: Your task is to calculate how many ways to create 6 digit alpha-numeric password with first 2 digit as alphabet and last 4 digit as numbers.

=> If repitation allowed:

alphabet = 26 ways
numbers = 10 ways

$\underline{26}\ \underline{26}\ \underline{10}\ \underline{10}\ \underline{10}\ \underline{10}$ ways.

$= 26 \times 26 \times 10 \times 10 \times 10 = 6760000.$

If repitation not allowed: $\underline{26}\ \underline{25}\ \underline{10}\ \underline{9}\ \underline{8}\ \underline{7}$

$= 26 \times 25 \times 10 \times 9 \times 8 \times 7 = 3276000$

Eg → There are 8 athelete who take part in 100m race. What are possible ways athelete can finish race.

⇒

Position	1	2	3	4	5	6	7	8
ways	8	7	6	5	4	3	2	4

No. of ways $= 8\times7\times6\times5\times4\times3\times2\times4 = 40,320$

- **Factorial**: The product of first 'n' +ve int (counting number) is called factorial & denoted by $n!$. $n! = n\times(n-1)\times\cdots\times1$

  Note: $0! = 1$

  So, above example can solved as: $8! = 40,320$.

  It also written as: $5! = 5\times4! = 5\times4\times3!$ & so on.

  So, In general, for $i \leq n$ we have,

$$n! = n\times(n-1)\cdots\times(n-i+1)\times(n-i)!$$

Eg: $\frac{6!}{3!} = \frac{6\times5\times4\times\cancel{3\times2\times1}}{\cancel{3\times2\times1}} = 6\times5\times4 = 120$ or $\frac{6\times5\times4\times\cancel{3!}}{\cancel{3!}}$

$\frac{6!\times5!}{3!\times4!} = \frac{6\times5\times4\times3!}{3!}\times\frac{5\times4!}{4!} = 6\times5\times4\times5 = 600$

Eg:

Write $25\times24\times23$ in term of factorials:-

$$= \frac{25\times24\times23\times22\times21\times\cdots\times1}{22\times21\times\cdots\times1} = \frac{25!}{22!}$$

- **Permutation**: It is an ordared arangement of all or some of 'n' object.

  The number of possible permutation of

n = objects
r = positions

'r' objects from a collection of 'n' disticts is given by formula:

$$n \times (n-1) \times \ldots \times (n-r+1)$$

& is denoted by ^nP_r

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

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

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

▷ Special cases:

1.) $^nP_0 = \frac{n!}{(n-0)!} = \frac{n!}{n!} = 1$

2.) $^nP_1 = \frac{n!}{(n-1)!} = n$

3.) $^nP_n = \frac{n!}{(n-n)!} = \frac{n!}{0!} = n!$ where $0! = 1$.

Eg: No. of ways of 8 person committee can be choosen for chairman & vice chairman without repititon.

$\Rightarrow n = 8$, $r = 2$ $= \frac{8!}{(8-2)!} = 56$ ways.

No. of 4 digit no. can be formed using digits 1, 2, 3, 4, 5 if no digit is repeated.

$\Rightarrow n = 5$, $r = 4$ $= \frac{5!}{1!} = 120$

How many even no' can be formed? Same.

$\_\ \_\ \_\ \underline{2}$

$\_\ \_\ \_\ \underline{4}$ fixed

So, $n=4$, $r=3$

$= \frac{4!}{1!} = 24$ for 1 so for both $24 \times 2 = 48$.

6 people 10 chair in linear order. Ways any?

$n=10$, $r=6$

$\Rightarrow {}^{10}P_6 = 151200$

3 characters A, B, C can repeat how many ways to fill 3 box? $\_\ \_\ \_$

3w 3w 3ways $= 3 \times 3 \times 3 = 27$ ways

Same A, B, C with 2 box? $3^3 = 27$

3w 3w $= 3 \times 3 = 9$ ways $3^2 = 9$

The no' of possible permutation of 'r' objects from a collection of 'n' distinct object where repition is allowed is given by. $n \times n \times \cdots \times n$. denoted by n^r.

Example of rearranging objects:

- DATA how many times it can be rearranged?

DATA = $DA_1TA_2 \leftrightarrow DA_2TA_1$ (If we consider A different)

ATAD $A_1TA_2D \leftrightarrow A_2TA_1D$ (Similarly)

4! if total ways to arrange 4 letter in words.
2! is 2 place which are same = 2 'a' in 'data'
'DATA' can be arranged in $4!/2! = 12$ ways.

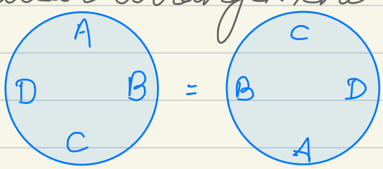
No. of Permutation of n object when p of them are of 1 kind and rest distinct is equal to $\frac{n!}{p!}$

No. of permutation of 'n' object where p_1 is one kind & p_2 is diff. kind and so on so, eq. will be $\frac{n!}{p_1!\,p_2!\ldots p_n!}$

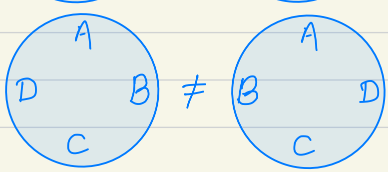
For 'statistics' $n = 10$ $p = 5$ $= \frac{10!}{3! \times 3! \times 1! \times 2! \times 1!} = 50{,}400$

s = 3
t = 3
a = 1
i = 2
c = 1

- Circular arrangement:



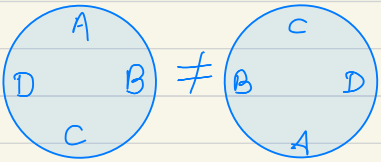
Both are same just rotated.



Mirror images are different.

If clockwise or anticlockwise rotation are same than it will be $\frac{(n-1)!}{2}$

If clockwise or anticlockwise rotation are different means if we fix one point will be $(n-1)!$



Ex. Solve for n.

$^nP_4 = 20\,^nP_2 = \frac{n!}{(n-4)!} = 20 \times \frac{n!}{(n-2)!}$

$(n-2)\times(n-3) = 20$, so, $n = -2$ or $n = 7$

$\frac{^nP_4}{^{n-1}P_4} = \frac{5}{3} = \frac{n!}{(n-4)!} \times \frac{(n-5)!}{(n-1)!} = \frac{5}{3}$

$\Rightarrow \frac{n}{n-4} = \frac{5}{3} \Rightarrow 2n = 20 \Rightarrow n = 10$

Solve for r

LHS RHS

$^5P_r = 2\cdot\,^6P_{r-1} = \;^5P_r = \frac{5!}{(5-r)!}$ $\quad ^6P_{r-1} = \frac{6!}{(6-(r-1))!} = \frac{6!}{(7-r)!}$

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

$\Rightarrow \frac{\cancel{5!}}{\cancel{(5-r)!}} = \frac{2\times 6\times \cancel{5!}}{(7-r)(6-r)\cancel{(5-r)!}}$ $\quad \Rightarrow (7-r)(6-r) = 12$

$\Rightarrow r^2 - 13r + 30 = 0$ $\quad r = 3$ or $r = 10$

- Combination:

In combination $ab = ba$

The number of possible combination of 'r' objects from collection of 'n' distinct object is denoted by ${}^nC_r$ and is given by: ${}^nC_r = \frac{n!}{r!(n-r)!}$

Some useful result:

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

  In other word, selecting 'r' object from 'n' objects is same as rejecting $n-r$ objects from 'n' objects.

- ${}^nC_n = 1$ and ${}^nC_0 = 1$ for all values of n.
- ${}^nC_r = {}^{n-1}C_{r-1} + {}^{n-1}C_r$; $1 \leq r \leq n$

Example → In exam a paper is divided in 2 part with 7 & 5 question. A student need to attend 8 ques & min. of 3 have to done in both.

$\Rightarrow {}^7C_3\,{}^5C_5 + {}^7C_4\,{}^5C_4 + {}^7C_5\,{}^5C_3 = 35 + 175 + 210 = 420$ ways.

Total no. of ways to choose 4 card from 52 cards?

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

All 4 card of same suits?

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

Select 11 players from 17 player where 5 bowler are there and 4 bowler must be their. Ways?

$\Rightarrow {}^5C_4 \times {}^{12}C_7 = \frac{5!}{4!\,1!} \times \frac{12!}{7!\,5!} = 5 \times \frac{9504\cancel{0}}{12\cancel{0}} = 3960$ ways

In general, given 'n' points, number of line segment that can be drawn connecting point is ${}^nC_2$.

- Distinguish situation b/w permutation & combination:

→ Permutation is used when order matters.
→ Combination is used when order don't matter.

Eg:

Perm: No. of ways 8 athelete can come 1st, 2nd & 3rd
Comb: No. of ways 3 athelete from 8 qualify.

Comb: Choose 2 class representative from 40 student.
Perm: Choose 1 V.C & 1 C from 40 students.

Similarly joing lines to dot if:
- Direction matters → Permutation.
- Direction don't matter → Combination.

- Probablity:

- Experiment: It is any process that produce an an observation or outcome.

- Random experiment: It is an experiment whose outcome is not predictable with certainity.

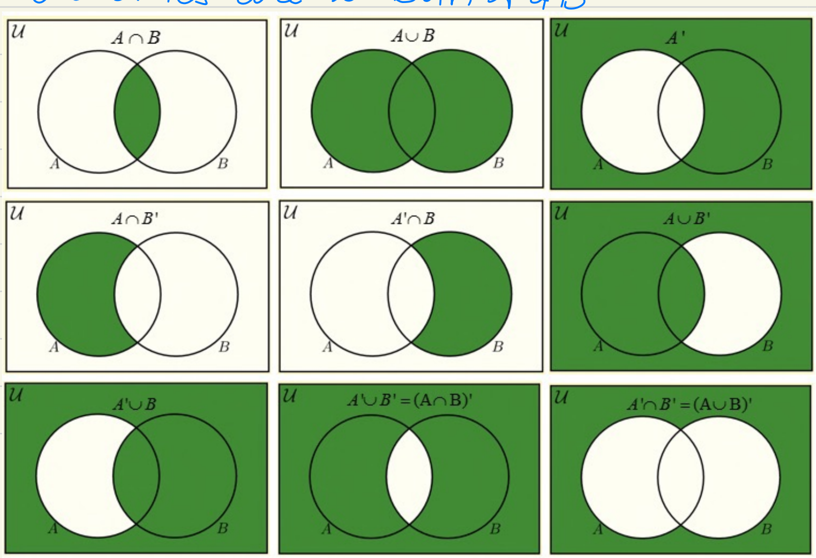
- Sample space (Ω or S): It is collection of all basic outcomes. Ex: Coin toss (H,T), Die roll (1,2,3,4,5,6)

- Event: It is collection of basic outcomes. It is subset of sample space. We say event occured if event is contained in subset. Eg: Die roll (1,3,6), Coin toss (H,H).

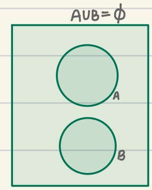
  - Union of event ($\cup$): 2 events A and B, we define new event $A \cup B$ calld union of event A & B, to all outcomes that are in A & B or in both. $A \cup B$ occur if either A or B occur.

  - Intersection of event ($\cap$): For 2 or more event we define $A \cap B$ as intersection of event if all outcomes are in both A & B.

Venn diagram

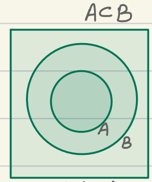


$A \cup B = \phi$



Disjoint event

$A \subset B$



Subset

- Null event and disjoint event :

Event without any outcomes are called null event. Symbol (ϕ)

If intersection of A and B is null event than A & B can't occur together so A and B is disjoint or mutually exclusive event. ($A \cap B = \phi$)

- Compliment of an event :

A compliment of A is denoted by A^c means all outcome of sample space S that are not in A.

Eg: Coin toss. $S = \{HH, HT, TT, TH\}$ $A = \{HT\}$ $A^c = \{HH, TT, TH\}$

The complement of sample set is null set ϕ.

If all outcome of event A are in event B than we can say A is contained by B or A is subset of B. $A \subset B$

Eg. 2 coin toss. $S = \{HH, HT, TT, TH\}$

A = 1st head = $\{HH, HT\}$, Both head = $\{HH\} = B$

we can say $B \subset A$

- Probablity :

Classical def : let S be sample space of random exp. in which 'n' equally likely outcome & E event consist 'm' of these outcome the every event probility is m/n & $P(E) = \frac{m}{n}$

Eg: Even no. in dice roll.

$m = 3 = (2,4,6)$ $n = 6 = (1,2,3,4,5,6) = \frac{m}{n} = \frac{3}{6} = \frac{1}{2}$

Relative frequency : It is calculated by repeating experiments many times. In other words, if $n(E)$ is no. of times E occurs in 'n' repetition of experiment. $P(E) = \lim_{n \to \infty} \frac{n(E)}{n}$

Eg: In 1000 toss you get head 468 times so probablity of head $P(H) = \frac{468}{1000} = .468$ times

Subjective approch : This probablity of event is best guess by person making statement. This probablity is measured by individual's degree of belief in event.

Eg: There is 70% chance you will win tomorow.

- Probablity Axioms :
  - For any event E, the probablity of E is number between 0 and 1. That is, $0 \leq P(E) \leq 1$.
  - The probablity of sample space is 1. (The outcome of exp. will be element of sample space S with $P(E)=1$)
  - The probablity of the union of disjoint event is equal to sum of prob. of these events. $P(E_1 \cup E_2) = P(E_1) + P(E_2)$

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

Eg: Dice roll. E_1 (Getting odd no) $= \{1,3,5\}$ E_2 (Even) $= \{2,4,6\}$

$E_1 = \frac{3}{6}$, $E_2 = \frac{3}{6}$

$$P(E_1 \cup E_2) = \frac{3}{6} + \frac{3}{6} = \frac{6}{6} = 1$$

○ General property of probability :

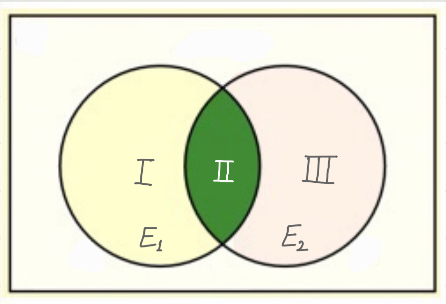
- Probablity of compliment of an event :
  - E and E^c is disjoint , $E \cup E^c = S$
  - $P(S) = 1$, $P(E) + P(E^c) = 1$

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

- Probablity of null event $P(\phi) = 0$
  - $S^c = \phi$

- Addition rule of probablity : It also work for not-disjoint

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



- $E_1 \cup E_2 = I \cup II \cup III$
- $E_1 = I \cup II$
- $E_2 = II \cup III$
- $E_1 \cap E_2 = II$

- $P(E_1 \cup E_2) = P(I \cup II \cup III) \longrightarrow P(I) + P(II) + P(III)$ — (I)
- $P(E_1) = P(I \cup II) \longrightarrow P(I) + P(II)$ — (II)
- $P(E_2) = P(II \cup III) \longrightarrow P(II) + P(III)$ — (III)
- $P(E_1 \cap E_2) = P(II) \longrightarrow P(II)$ — (IV)

From I, II, III & IV we get :

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

Eg→ Customer probablity to buy shirt ·3 & pant ·2 both is ·1. What is probablity of customer to buy neither.

⇒ $S = \cdot 3$, $P = \cdot 2$, $S \cap P = \cdot 1$ then nither will be $P(S \cup P)^c$

$\because P(S \cup P) = P(S) + P(P) - P(P \cap S) = \cdot 3 + \cdot 2 - \cdot 1 = \cdot 4$

Hence, $P(S \cup P)^c = 1 - \cdot 4 = \cdot 6$

	Own smartphone		
Gender	No	Yes	Row t.
Female	10	32	44
Male	14	42	56
Column t.	24	76	100

	Own smartphone		
Gender	No	Yes	Row t.
Female	$\boxed{\cdot 10}$	·32	·44
Male	·14	$\boxed{\cdot 42}$	·56
Column t.	·24	·76	100

Being female not owning smartphone.

Prob. of being male & owning smartphone.

- Joint Probablity: It represent probablity of intersection of 2 or more events.

  Eg: Female & not owning smartphone = ·10

- Marginal probablity: It is probablity of observing an outcome with single attribute, regardless of its other attibutes.

  Eg: P(Female) = $\cdot 10 + \cdot 34 = \cdot 44$ $= P(FO^c) + P(FO)$

  P(Owning Phone) = $\cdot 34 + \cdot 42 = \cdot 76$

- Conditional Probablity:
  - Finding cond. prob. to answer question like:
    - ⇒ 'Among female buyers what is chance of owning phone'
    - ⇒ 'Among people not have phone, how many are male?'

- In this we restrict the sample space of contengency table to a row or column.

Gender	No	Yes	Total
Male	10/44	34/44	44
Female	14/56	42/56	56
Total	24/100	76/100	100

By Row

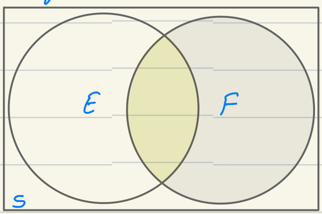
Gender	No	Yes	Total
Male	10/24	34/76	44/100
Female	14/24	42/76	56/100
Total	24	76	100

By Column

$P(E|F)$ ⇒ It is conditional probablity of E given F has occured.

The probablity of event E occurs given that event F occurs is given by:

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



- Multiplication rule: Conditional Probablity formula:
  Multiplying both side with P(F), we get:

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

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

- It states that prob. of both E & F occuring together is equal to probablity that F occurs multiplied by cond. prob. of E given F occurs.
- It is useful to compute prob. of intersection of event

Eg: Of 40 student 23 male 17 female.
Cond 1. (1st st. female & 2nd male)

$$P(F_1 \cap M_2) = P(F_1) \cdot P(M_2|F_1)$$
$$= \frac{17}{40} \cdot \frac{23}{39} = .251$$

- This rule can be generalised intersection of an arbitrary no. of event is reffers as:

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

Eg: A deck of 52 card, divided in 4 pile randomly. What is the probablity of following events.
E_1: Ace of spade is in any of 4 pile.
$P(E_1) = 1$
E_2: Ace of spade & Ace of heart in diff. pile.

On 1st pile Ace of spade of 13 card is 1 card rest 12 card will be choosen from 51 card so, Ans will be its complement. $P(E_2|E_1) = 1 - \frac{12}{51} = \frac{39}{51}$

E_3: AS, AH, AD in diff. pile.

We put AS in P_1 and AH in P_2 so to fill rest 12-12=24 card of P_1 & P_2 we have 50 cards. So, no. of ways AD can be in P_1 & P_2 is complement of $P(E_3|E_1 \cap E_2) = 1 - \frac{24}{50} = \frac{26}{50}$

We got by fixing AS & AH

E_4: $P(E_4|E_1 \cap E_2 \cap E_3) = \frac{13}{49}$ (Fixing 3 card in 3 deck and choosing from 36/49 complement)

E_5: $P(E_1 \cap E_2 \cap E_3 \cap E_4) = \frac{39}{51} \cdot \frac{26}{50} \cdot \frac{13}{49} \approx 0.105$

- Independent event: It can be defined as the upcoming events doesn't depend on previous event.

  Eg: Current cointoss (H, T) don't depend on prev. result.

  Since:

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

  Hence, Two event E and F are independent if

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

  The probablity that both occur equals to the product of their individual probablity.

- Dependent event: Two event that are not independent is dependent.

Eg: Rolling 2 dice.

$$S = \left\{ \begin{array}{l} (1,1),(1,2),(1,3),(1,4),(1,5),(1,6), \\ (2,1),(2,2),(2,3),(2,4),(2,5),(2,6), \\ (3,1),(3,2),(3,3),(3,4),(3,5),(3,6), \\ (4,1),(4,2),(4,3),(4,4),(4,5),(4,6), \\ (5,1),(5,2),(5,3),(5,4),(5,5),(5,6), \\ (6,1),(6,2),(6,3),(6,4),(6,5),(6,6), \end{array} \right\}$$

Define foll. events:

E_1: First outcome is 3

$\Rightarrow 6/36 = 1/6$

E_2: Sum of outcome is 8.

$\Rightarrow 5/36$

E_3: Sum of outcome is 7.

$\Rightarrow 6/36 = 1/6$

Are event E_1 & E_2 independent? No

$\rightarrow E_1 \cap E_2$ is 1st outcome 3 & sum of outcome is 8.

Only event is $(3,5)$ is $1/36$

Since, $1/36 \neq 6/36 \times 5/36$ we see,

$P(E_1 \cap E_2) \neq P(E_1) \times P(E_2)$, So E_1 & E_2 are not independent.

Are event E_1 & E_3 independent? Yes

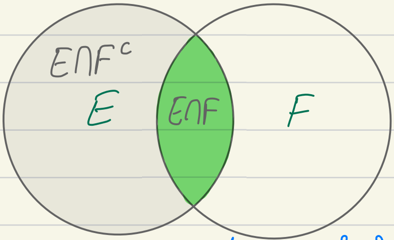
Since, $E_1 \cap E_3 = P(3,5) = 1/36$.

Since $P(E_1 \cap E_3) = P(E_1) \times P(E_3)$

$\Rightarrow 1/36 = 6/36 \times 6/36$ so it is independent.

- Independence of E and F^c

If E and F are independent so does E and F^c.



$E \cap F$ & $E \cup F^c$ is disjoint & mutually exclusive.

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

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

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

$P(E \cap F^c) = P(E)[1 - P(F)]$

$P(E \cap F^c) = P(E) P(F^c)$ So, E & F^c are independent.

- Independent of 3 events:

3 events E, F and G are said to be independent if:

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

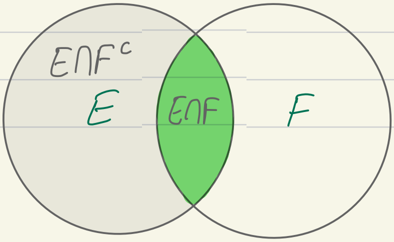
$P(E \cap F) = P(E) \times P(F)$

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

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

If all these 4 happens than it is independent event.

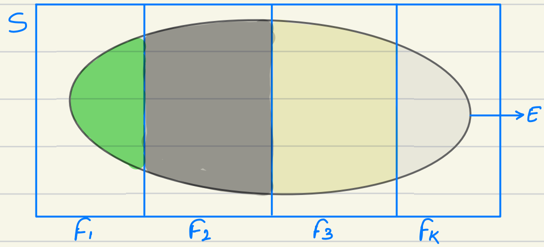
- Law of total probablity: Let E and F be events. In order to outcome to be E it must be in both E & F or be in E and not in F.



$E = (E \cap F) \cup (E \cap F^c)$

– Formula & interpretation:

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



$$E = (E \cap F_1) \cup (E \cap F_2) \cup (E \cap F_3) \cup (E \cap F_K)$$
$$P(E) = P(E \cap F_1) + P(E \cap F_2) + P(E \cap F_3) + P(E \cap F_K)$$
$$P(E) = P(E|F_1) \cdot P(F_1) + P(E|F_2)P(F_2) + P(E|F_3) \cdot P(F_3) + \dots + P(E|F_K)P(F_K)$$

Suppose event F_1, F_2, F_3 & F_K, are mutualy exclusive & exhaustive; that is exactly one of the event must occur. Than for any event E.

$$P(E) = P(E|F_1) \cdot P(F_1) + P(E|F)P(F_2) + P(E|F_3) \cdot P(F_3) + \dots + P(E|F_K)P(F_K)$$

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

- Bayes' rule: Suppose we are intrested in conditional prob. of event F conditioned on E. we know

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

From def:

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

Bayes' rule is suppose that event $F_1, F_2, \dots, F_K$, are mutually exclusive and exhaustive; Than for any event E,

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

- Sample space: A random variable is numerical distribution of outcome of statistical experiment.

Eg: Rolling 2 dice.

$$S = \left\{ \begin{matrix} (1,1),(1,2),(1,3),(1,4),(1,5),(1,6), \\ (2,1),(2,2),(2,3),(2,4),(2,5),(2,6), \\ (3,1),(3,2),(3,3),(3,4),(3,5),(3,6), \\ (4,1),(4,2),(4,3),(4,4),(4,5),(4,6), \\ (5,1),(5,2),(5,3),(5,4),(5,5),(5,6), \\ (6,1),(6,2),(6,3),(6,4),(6,5),(6,6), \end{matrix} \right\}$$

- Of the outcomes, how many outcomes will result in sum of outcomes as 7?
- Of how many outcome will have smaller of the outcome is 3?

(sample space for both question is same.)

Outcome	X	Y	Outcome	X	Y	Outcome	X	Y
(1, 1)	2	1	(1,3)	4	1	(1,5)	6	1
(2, 1)	3	1	(2,3)	5	2	(2,5)	7	2
(3, 1)	4	1	(3,3)	6	3	(3,5)	8	3
(4, 1)	5	1	(4,3)	7	3	(4,5)	9	4
(5, 1)	6	1	(5,3)	8	3	(5,5)	10	5
(6, 1)	7	1	(6,3)	9	3	(6,5)	11	5
(1,2)	3	1	(1,4)	5	1	(1,6)	7	1
(2,2)	4	2	(2,4)	6	2	(2,6)	8	2
(3,2)	5	2	(3,4)	7	3	(3,6)	9	3
(4,2)	6	2	(4,4)	8	4	(4,6)	10	4
(5,2)	7	2	(5,4)	9	4	(5,6)	11	5
(6,2)	8	2	(6,4)	10	4	(6,6)	12	6

Let X denote sum of outcome of 2 rolls.

X takes the value 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 & 12.

X Value	Relevent event	Probablity(x)
2	{(1,1)}	1/36
3	{(1,2), (2,1)}	2/36
4	{(1,3), (2,2), (3,1)}	3/36
...	...	...
10	{(4,6), (5,5), (6,4)}	3/36
11	{(5,6), (6,5)}	2/36
12	{(6,6)}	1/36

Let Y denote lesser one of 2 outcomes.

Y takes value 1, 2, 3, 4, 5 & 6

Y Value	Relevent event	Probablity(x)
1	{(1, 1), (1,2), ..., (5, 1), (6, 1)} = 11	11/36
2	. . . = 9	9/36
3	. . . = 7	7/36
4	. . . = 5	5/36
5	{(5,5), (5,6), (6,5)} = 3	3/36
6	{(6,6)} = 1	1/36
		Answer

Eg: 3 coin are tossed.

S = {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}

Of these 3 tosses How many are head?

" " " " Which result in head first time in pos.?

→ let x = no. of head, X take value 0, 1, 2, 3

Value of X	Relevent event	Probablity(X)
0	{(TTT)}	1/8
1	{(HTT), (THT), (TTH)}	3/8
2	{(HHT), (HTH), (TTH)}	3/8
3	{(HHH)}	1/3

y = Toss with 1st time head, y takes value 1, 2, 3

Value of X	Relevent event	Probablity(X)
1	{(HHH), (HHT), (HTH), (HTT)}	4/8
2	{(THH), (THT)}	2/8
3	{(TTH)}	1/8
Nil	{(TTT)}	1/8

Eg: Life insurence agent has 2 elderly client who are paid 1L upon death.

=> Let A be event for younger among 2 to die and
B be event for older among 2 to die

Conditions may happen	Next year	Money
Both A & B die	$A \cap B$	2L
Only younger die A	$A \cap B^c$	1L
Only elder die B	$A^c \cap B$	1L
Both survive	$A^c \cap B^c$	0

3values

Assume A & B are independent event with probablities $P(A) = \cdot 05$ & $P(B) = \cdot 10$
Let x is total money company pay to their client in case of death (units in lakhs).
So, x can take 3 values. 0, 1 and 2.

Value of x	Relevent event	
0	$A^c \cap B^c$	$P(A^c) \times P(B^c) = \cdot 95 \times \cdot 9 = \cdot 855$
1	$(A \cap B^c) \cup (A^c \cap B)$	$P(A \cap B^c) + P(A^c \cap B) = (\cdot 05 \times \cdot 9) + (\cdot 95 \times \cdot 1) = \cdot 14$
2	$A \cap B$	$P(A) \times P(B) = \cdot 05 \times \cdot 1 = \cdot 005$

- Continous random variable : When outcome for random event are numeric & can't be counted & are infinitely devisible. Eg: Temperature of person, Height of person, etc.
  — It typically involve measuring.

- **Discrete random variable:** It is a random variable that can take on at most a countable number of possible value. Eg: No. of accident, no. of people, etc.
  - Involve counting on no. line.

- **Probablity mass function:**
  In descrete random variable we take countable no. of possible value.
  - So let x be a discrete random variable with 'n' possible value. then it can lebel as $x_1, x_2, \dots, x_n$.
  - For descrete random variable X, we define Probablily mass fn (PMF) $P(x)$ of X by
    $$p(x_i) = P(X = x_i)$$

X	x_1	x_2	$\dots$	x_n
$P(X=x_i)$	$P(x_1)$	$P(x_2)$	$\dots$	$(P x_n)$

  - The PMF $P(x)$ is +ve for atmost a countable no. of value of x. That is x must assume one of the values $x_1, x_2, \dots,$ then
    $$P(x_i) \geq 0, \; i = 1, 2, \dots$$
    $$P(x_i) = 0 \text{ for all other value of } x.$$
    Represented in above tebular form.
  - Since x must take one of the value x_i, we have
    $$\sum_{i=1}^{\infty} P(x_i) = 1$$

Example: Suppose X is a random variable which takes 3 values 0, 1 & 2. with probablity:

X	P(0)	P(1)	P(2)
$P(X=x_i)$	1/4	1/2	1/4

⇒ Condition is all should be more than 0 ✓

⇒ Sum shold be 1. $1/4 + 1/2 + 1/4 = 1$ ✓

- Suppose X is a random variable takes value 0, 1, 2, ... with probablities:

$$p(i) = c\frac{\lambda^i}{i!}, \text{ for some +ve } \lambda$$

Eg: Rolling 2 dice.

$$S = \begin{Bmatrix} (1,1),(1,2),(1,3),(1,4),(1,5),(1,6), \\ (2,1),(2,2),(2,3),(2,4),(2,5),(2,6), \\ (3,1),(3,2),(3,3),(3,4),(3,5),(3,6), \\ (4,1),(4,2),(4,3),(4,4),(4,5),(4,6), \\ (5,1),(5,2),(5,3),(5,4),(5,5),(5,6), \\ (6,1),(6,2),(6,3),(6,4),(6,5),(6,6), \end{Bmatrix}$$

All roll map to its sum: like $(1,1) = 2$ & $(6,6) = 12$

So, X is a random variable maps to sum of outcomes

X	2	3	4	5	6	7	8	9	10	11	12
$P(x=i)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

To verify $\sum_{i=1}^{11} P(X_i) = \frac{36}{36} = 1$

2) lesser value random variable

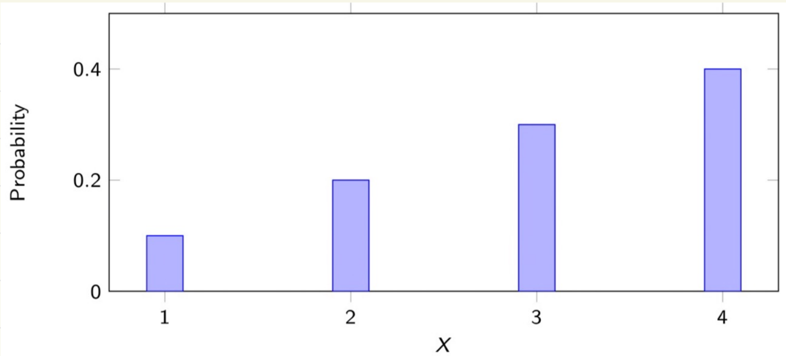
y	1	2	3	4	5	6
$P(y=y_i)$	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{5}{36}$	$\frac{3}{36}$	$\frac{1}{36}$

Sum = 1 So, it is PMF.

- Graph of probablity mass function:
  It is helpful to illustrate mass fn in a graphical format by plotting $P(X=x_i)$ on the y-axis against x_i on the x-axis.
  : Of previous example:

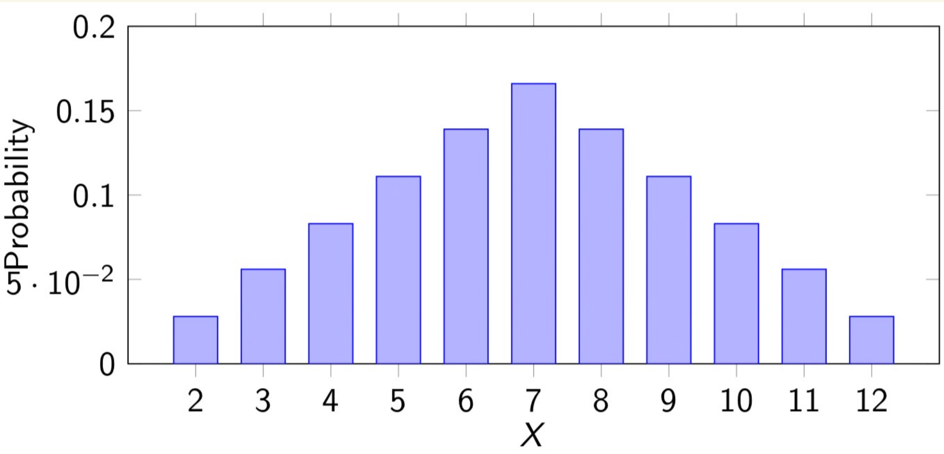
Example: negative skewed distribution

X	1	2	3	4
$P(X = x_i)$	0.1	0.2	0.3	0.4



Rolling a dice twice: X-sum of outcomes

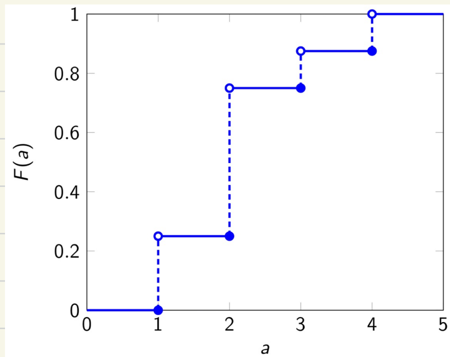
X	2	3	4	5	6	7	8	9	10	11	12
$P(X = x_i)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$



- Cumulative distribution function: It can be expressed by $F(a) = P(X \leq a)$.
  - If X is a descrete random variable whose possible values are $x_1, x_2, x_3, \ldots$, where $x_1 < x_2 < x_3, \ldots$, then the distribution fn F of X is a step function.
  - It will be step fn. So for an example.

X	1	2	3	4
$P(X = x_i)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$

$$F(a) = \begin{cases} 0 & a < 1 \\ \frac{1}{4} & 1 \leq a < 2 \\ \frac{3}{4} & 2 \leq a < 3 \\ \frac{7}{8} & 3 \leq a < 4 \\ 1 & 4 \leq a \end{cases}$$

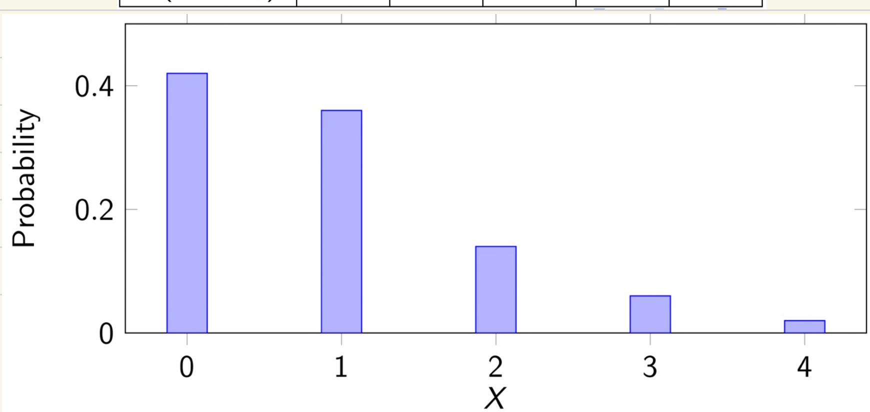


Eg: We want to analyse no. of credit card owned by people.

Random Experiment: Selecting adults at random.

Random variable: No. of CC owned by person.

X	0	1	2	3	4
$P(X = x_i)$	0.42	0.36	0.14	0.06	0.02



Probablity mass-fn

- Describe the distribution
  - The distribution is skewed right with a peak at 0.
  - Number of credit cards owned by people vary between 0 to 4 credit cards.
- Choose an adult at random. Is he/she more likely to have 0 credit cards or 2 or more credit cards?
  - The probability that an adult has no credit cards is 0.42, while the probability of having 2 or more credit cards is about 0.22, so the probability of having no credit cards is higher.
- You take a random sample of 500 people and ask them how many credit cards they own. Would you be surprised at the following:
  - Everyone owns a credit card.
    YES: 42% of adults do not own credit cards. Hence, it is unlikely that every one of the 500 would own credit cards.
  - 72 people respond that they own two credit cards.
    NO: 14% own two credit cards. 14% of 500 $=$ 70, so it is likely that 72 people from sample of 500 own two credit cards.
- Again choose an adult at random. How many credit cards would you "Expect" that person to own?

Eg: Consider a game of rolling dice.
- Even outcome - You loose amt. on dice.
- Odd outcome - You win amt. on dice

Outcome	1	2	3	4	5	6
Winning	+1	-2	+3	-4	+5	-6

Would you play:

1	5	4	6	4
6	4	3	1	5
5	2	1	6	2
6	5	4	5	5
2	4	2	2	5
4	1	3	4	5
6	1	5	2	5
1	6	6	5	2
3	6	5	6	3
5	3	2	5	4
3	3	5	4	4
5	1	2	3	4
3	2	1	1	6
1	3	4	3	4
4	4	5	6	1
3	3	5	1	1
1	6	5	4	3
5	1	4	6	5
4	4	4	3	6
5	5	4	1	3

Rolling 100 times

Outcome	Winning	Frequency	Relative frequency
1	+1	16	0.16
2	-2	10	0.10
3	+3	16	0.16
4	-4	21	0.21
5	+5	23	0.23
6	-6	14	0.14
		100	1

Average winnings: -0.09

Rolling a 1000 times

Outcome	Winning	Frequency	Relative frequency
1	+1	177	0.177
2	-2	177	0.177
3	+3	167	0.167
4	-4	153	0.153
5	+5	163	0.163
6	-6	163	0.163
		1000	1

Average winnings: -0.451

Winning will always in -ve after many hundred rolls so we won't play.
Relative frequency of each outcome is 1/6.

So, gain will be $1\frac{1}{6} - 2\frac{1}{6} + 3\frac{1}{6} - 4\frac{1}{6} + 5\frac{1}{6} - 6\frac{1}{6} = -0.5$

- Expectation of random variable:
Let x be descrete random variable taking values $x_1, x_2, \cdots$ Expected value of x denoted by $E(X)$ reffers as expectation of x given by:

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

It is considered as long run avg. of ind. obs.

Random experiment: Roll a dice once.

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

Random variable X is the outcome of the roll.

The probability distribution is given by

X	1	2	3	4	5	6
$P(X = x_i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$E(X) = 1\frac{1}{6} + 2\frac{1}{6} + 3\frac{1}{6} + 4\frac{1}{6} + 5\frac{1}{6} + 6\frac{1}{6} = 3.5.$

Does this mean that if we roll a dice once, should we expect the outcome to be 3.5?

NO!!-the expected value tells us is what we would expect the average of a large number of rolls to be in the long run.

	100 rolls		1000 rolls		
Outcome	Freq	Rel. Freq	Freq	Rel. Freq	Probability
1	16	0.16	177	0.177	0.166667
2	10	0.1	177	0.177	0.166667
3	16	0.16	167	0.167	0.166667
4	21	0.21	153	0.153	0.166667
5	23	0.23	163	0.163	0.166667
6	14	0.14	163	0.163	0.166667
		3.67		**3.437**	**3.5**

- Notice that average of the rolls need not be exactly 3.5.
- However, we can expect it to be close to 3.5.
- The expected value of X is a theoretical average.

Rolling a dice twice

X is a random variable which is defined as sum of outcomes

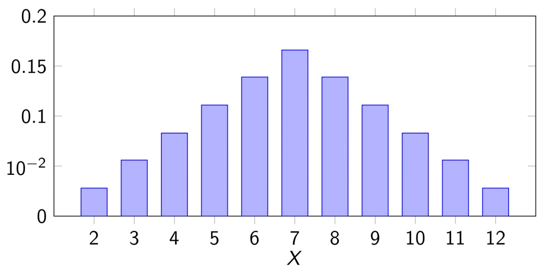
Probability mass function

X	2	3	4	5	6	7	8	9	10	11	12
$P(X = x_i)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

If I rolled two dice a large number of times, what can I expect the average of the sum of the outcomes to be?

$$E(X) = 2 \times \frac{1}{36} + 3 \times \frac{2}{36} + \ldots + 11 \times \frac{2}{36} + 12 \times \frac{1}{36} = 7$$

Interpretation: When two dice are rolled over and over for a long time, the mean sum of the two dice is 7.

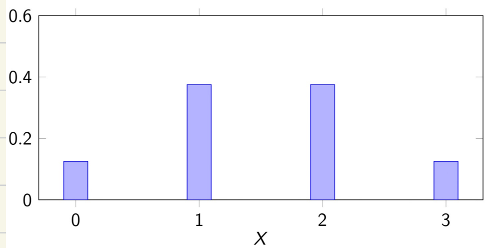


Tossing a coin thrice

- $S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$
- X is the random variable which counts the number of heads in the tosses
  - Probability mass function

X	0	1	2	3
$P(X = x_i)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

  - $E(X) = \sum_{i=0}^{3} x_i p(x_i) =$ $\frac{(0 \times 1) + (1 \times 3) + (2 \times 3) + (3 \times 1)}{8} = \frac{3}{2}$
- Interpretation: When a coin is tossed three times over and over for a long time, the mean number of heads in the three tosses is 1.5.



- Expectation of a fn of a random variable:

It is expected value of probablity by fn g.

Suppose:

x	x_1	x_2	...	x_n
$g(x)$	$g(x_1)$	$g(x_2)$	...	$g(x_n)$
$P(x=x_i)$	$P(x=x_1)$	$P(x=x_2)$		$P(x=x_n)$

Expected value of $g(x)$:

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

If a and b are constant:

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

Eg:

Let X be a discrete random variable with the following distribution

X	-1	0	1
$P(X=x_i)$	0.2	0.5	0.3

Let $Y = g(X) = X^2$. What is $E(Y)$?

$E(Y) = (-1^2) \times 0.2 + 0 \times 0.5 + 1^2 \times 0.3 = 0.5$

Distribution of Y

Y	0	1
$P(Y=y_i)$	0.5	0.5

NOTE: $0.5 = E(X^2) \neq (E(X))^2 = 0.01$

- Expectation of sum of 2 random variable:

→ Expected value of sum of random variable is same as sum of individual random values.

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

Eg:

Two students are randomly chosen from a group of 20 boys and 10 girls. Let X denote the number of boys chosen, and let Y denote the number of girls chosen.

1. Find $E(X)$.
   - X is a Hypergeometric rv. $N=30, m=20, n=2$. Hence $E(X) = \frac{2\times 20}{30} = \frac{4}{3}$
2. Find $E(Y)$.
   - Y is a Hypergeometric rv. $N=30, m=10, n=2$. Hence $E(Y) = \frac{2\times 10}{30} = \frac{2}{3}$
3. Find $E(X+Y)$.
   - $E(X+Y) = E(X)+E(Y) = \frac{4}{3}+\frac{2}{3} = 2$

- Expectation of sum of many rand. var. :-

Expected value of sum of random variable is same as sum of individual random values.
Let, $x_1, x_2, \ldots, x_K$ be K descrete rand. var. Then,

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

Example: Tossing a coin three times

- Toss a coin i times.
- Let Xi be a random variable which equals 1 if the outcome is a head, 0 otherwise.
- E(Xi)=0.5
- X1 +X2 +...+Xn is the total number of heads in n tosses of the coin.
- E(X1 +X2 +...+Xn)=XE(Xi)=0.5×n
- For n=3,X1+X2+X3 is equal to the number of heads in three tosses of a coin.

E(X1 +X2 +X3) = 3×0.5 = 1.5

This is the same expectation of number of heads in three tosses of a coin.

The expected value of a random variable gives the weighted average of possible values of random variable. It don't tell about variation & spread.

- Variance of a random variable :

Let expected value of a random variable X be μ.
— Let X be rand. var. with expected value μ, than the variance of X, denoted by Var(X) or V(X), defined by

$$Var(X) = E(X-\mu)^2$$

X measures the square of diff. of rand. var. from its μ, on the average.

$Var(X) = E(X-\mu^2)$

$(X-\mu)^2 = X^2 - 2X\mu + \mu^2$

Using expectation property:

$E(X^2 - 2X\mu + \mu^2) = E(X^2) - 2\mu E(X) + \mu^2$ same as.

$Var(X) = E(X^2) - \mu^2$

X is a random variable which is defined as sum of outcomes

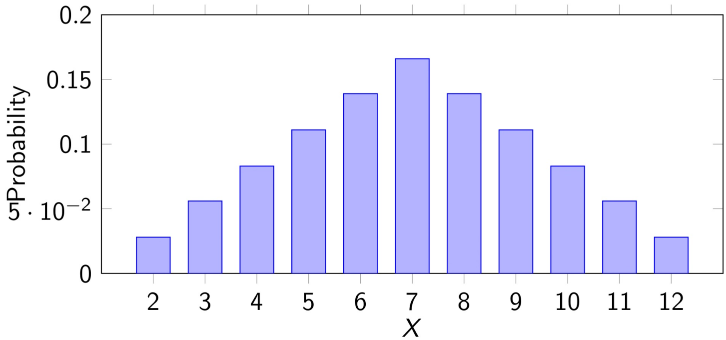
Probability mass function

X	2	3	4	5	6	7	8	9	10	11	12
X^2	4	9	16	25	36	49	64	81	100	121	144
$P(X = x_i)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

$E(X) = 2 \times \frac{1}{36} + 3 \times \frac{2}{36} + \ldots + 11 \times \frac{2}{36} + 12 \times \frac{1}{36} = 7$

$E(X^2) = 4 \times \frac{1}{36} + 9 \times \frac{2}{36} + \ldots + 121 \times \frac{2}{36} + 144 \times \frac{1}{36} = 54.833$

$Var(X) = 54.833 - 49 = 5.833$



- Barnauli Random Variable:

The Rand. Var. that takes either 0 or 1 is called Barnauli Random Variable.

X	0	1
X^2	0	1
$P(X = x_i)$	$1-p$	p

Expected value of a Bernoulli random varaible:

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

Variance of a Bernoulli random variable:

$Var(X) = p - p^2 = p(1-p)$

Descrete uniform Random variable:

let X be a rand. var. likely to take any value 1, 2, …, n.

PMF:

X	1	2	…	n
X^2	1	4	…	n^2
$P(X = x_i)$	$\frac{1}{n}$	$\frac{1}{n}$	…	$\frac{1}{n}$

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

$E(X^2) = \frac{(n+1)(2n+1)}{6}$

$Var(X) = \frac{n^2-1}{12}$

Variance of a fn of random variable:

let X be a random variable, let c be a constant, than,

$$Var(cX) = c^2 Var(X)$$

$$Var(X+c) = Var(X)$$

So, If a & b are constant,

$$V(aX+b) = a^2 Var(X)$$

Proof:

$$E(aX+b) = a\mu + b,$$

$$Var(aX-b) = E(aX+b-a\mu-b)^2 = E(a^2(X-\mu)^2)$$

$$= a^2 E(X-\mu)^2 = a^2 Var(X)$$

Variance of sum of 2 random variable:

We know, $E(X+Y) = E(E) + E(Y)$

let X & Y are 2 random variable.

Let $X = X$ & $Y = X$, so (V= variance)

$$V(X+Y) = V(X+X) = V(2X) = 4V(X)$$

So, $Var(X+X) = Var(2X) = 4 Var(X) \neq Var(X) + Var(X)$

Independendent random variable:

X & Y are independent R.V if knowing value of one of them don't change prob. of other.

Ex. → Roll a dice twice. $S = \{(1,1), \dots, (6,6)\}$

→ X & Y outcome of 1st & 2nd dice.

→ Knowing $X = i$ don't change Y taking value $1, 2, \dots, 6$.

→ So X & Y are ind. rand. var.

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

Example: Rolling a dice twice

- Let X be the outcome of a fair dice. Let Y be the outcome of another fair dice.
- We know $E(X) = E(Y) = 3.5$
- $X + Y$ is the sum of outcomes of both the dice rolled together. Then, we know $E(X+Y) = E(X) + E(Y) = 3.5 + 3.5 = 7$.
- We also know $Var(X) = Var(Y) = 2.917$
- X and Y are independent, hence, $Var(X+Y) = 2.917 + 2.917 \approx 5.83$, which is the same as what we obtained earlier applying the computational formula.

Variance of sum of many independent random variable:

The result that ind. var. variance is equal to sum of variance holds for not only 2 but any rand. variable.

→ Let $X_1, X_2, \dots, X_K$ be k discrete rand. var. then

$$Var\left(\sum_{i=1}^{K} X_i\right) = \sum_{i=1}^{K} Var(X_i)$$

Example: Tossing a coin three times

- Toss a fair coin i times.
- Let X_i be a random variable which equals 1 if the outcome is a head, 0 otherwise.
- $E(X_i) = 0.5$, $Var(X_i) = 0.25$
- $X_1 + X_2 + \ldots + X_n$ is the total number of heads in n tosses of the coin.
- $Var(X_1 + X_2 + \ldots + X_n) = \sum_{i=1}^{n} Var(X_i) = 0.5 \times n$
- For $n = 3$, $X_1 + X_2 + X_3$ is equal to the number of heads in three tosses of a coin.

$$Var(X_1 + X_2 + X_3) = 3 \times 0.25 = 0.75$$

This is the same as variance of number of heads in three tosses of a coin.

○ Standard deviation of random variable:

To calculate standard deviation of random var. we use $SD(X) = \sqrt{Var(X)}$ (+ve sq. root of variance)

– Properties of std. dev.:-

$$SD(cX) = cSD(X)$$

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

Eg, $Var(X)=4$, $SD(3X) = 3\sqrt{4} = 6$

$Var(2X+3)=16$, $\Rightarrow 4\,Var(X) = 16 \Rightarrow Var(X) = 4$ so, $SD(X) = 2$

Condition	Expectation	Variance	Stand. Dev.
X	$E(X)$	$V(X)$	$SD(X)$
cX	$cE(X)$	$c^2V(X)$	$cSD(X)$
$X+c$	$E(X)+c$	$V(X)$	$SD(X)$

Examples:

Application: family bonus

Sanjay and Anitha are a married couple who work for the same company. Anitha's Diwali bonus is a random variable whose expected value is ₹15,000 and standard deviation is ₹3,000. Sanjay's bonus is is a random variable whose expected value is ₹20,000 and standard deviation is ₹4,000. Assume the earnings of Sanjay and Anitha are independent of each other. What is the expected value and standard deviation of the total family bonus.

- Let X denote Anita's bonus. Given $E(X) = 15,000, SD(X) = 3,000$.
- Let Sanjay's bonus be Y. Given $E(Y) = 20,000, SD(Y) = 4,000$
- $E(X+Y) = E(X) + E(Y) =$ ₹35,000
- $SD(X+Y) = \sqrt{Var(X) + Var(Y)} =$ ₹5,000

Application: Lawyer's fees

- A lawyer must decide whether to charge a fixed fee of ₹25,000 or to take a contingency fee of ₹50,000 if she wins the case (and ₹0 if she loses).
- She estimates that her probability of winning is 0.5 (equal chance of winning or losing).
- Determine the expectation and standard deviation of her fee if
  - a She charges a fixed fee. $E(X) = 25,000, SD(X) = 0$
  - b She charges a contingency fee. $E(X) = 25,000, SD(X) = 25,000$

□ Bernoulli trial: A trial or an experiment whose outcomes can be classified as either success or failure is Bernoulli trial.

$S = \{Success, failure\}$

Let X be a rand. var. which takes 1 as success and 0 as failure. so X is bernoulli rand. var..

Examples of Bernoulli trials

- Experiment: Tossing a coin: $S = \{Head, Tail\}$
  - Success: Head
  - Failure: Tail
- Experiment: Rolling a dice: $S = \{1, 2, 3, 4, 5, 6\}$
  - Success: Getting a six.
  - Failure: Getting any other number.
- Experiment: Opinion polls: $S = \{Yes, No\}$
  - Success: Yes
  - Failure: No
- Experiment: Salesperson selling an object: $S = \{Sale, No\ sale\}$
  - Success: Sale
  - Failure: No sale
- Experiment: Testing effectiveness of a drug: $S = \{Effective, Not\ effective\}$
  - Success: Effective
  - Failure: Not effective

Non Bernoulli trial

- Experiment: Randomly choosing a person and asking their age.
- Not Bernoulli- Outcomes are not 2.

○ Bernouli random variable: It is rand. var. which takes either 0 or 1. X is a bernoulli rand. var. that take value 1 with probablity P means 0 = 1-P.

→ Probab. dist. is:

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

→ Expected value of bernoulli rand variable:

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

→ Variance of bernoulli rand. var.:

$$V(X) = P - P^2 = P(1-P)$$

Variance of bernouli distribution:

The largest variance occur when $P = 1/2$, when success & failure are equally likely.

In other word most uncertain bernoulli trials are those with largest variance, resembles toss of fair coin.

□ Independent and identically distributed Bernoulli trials:

A collection of Bernoulli trials is defined as IID Bernoulli trials.

A coll. of rand. var. is IID if they are ind. & share a common prob. dist..

- Ind. trial → Tossing 2 coin each time.
- Non ind. trial → Choosing 1 of 3 ball without replacement.

○ Binomial random variable:

- Suppose N ind. trials are performed with result 'success' as 'P' or 'failure' with probablity '1-P'.
- Let X is total no. of success that occurs in 'n' trials, that X is binomial rand. var. with parameter n & P.

n = no. of ind. trial (Here n is fixed.)

P = Prob. of success.

n=3 independent trials, X = number of successes

Let $n = 3$ independent Bernoulli trials.
Let p is probability of success.
X= number of successes in 3 independent trials.
The probabilities of outcomes of the independent trials are

S.No	Outcome	Number of successes	Probabilities
1	(s,s,s)	3	$p \times p \times p$
2	(s,s,f)	2	$p \times p \times (1-p)$
3	(s,f,s)	2	$p \times (1-p) \times p$
4	(s,f,f)	1	$p \times (1-p) \times (1-p)$
5	(f,s,s)	2	$(1-p) \times p \times p$
6	(f,s,f)	1	$(1-p) \times p \times (1-p)$
7	(f,f,s)	1	$(1-p) \times (1-p) \times p$
8	(f,f,f)	0	$(1-p) \times (1-p) \times (1-p)$

The probability distribution of X

X	0	1	2	3
$P(X=i)$	$(1-p)^3$	$3 \times p \times (1-p)^2$	$3 \times p^2 \times (1-p)$	p^3

n independent trials, X = number of successes

- Let there be n independent Bernoulli trials.
- Let p is probability of success.
- X= number of succeses in n independent trials.
- The probabilities of outcomes of the independent trials are

S.No	Outcome	Number of successes	Probabilities
1	(s,s,...,s)	n	$p \times p \times \ldots \times p$
2	(s,s,...,f)	$n-1$	$p \times p \times \ldots \times (1-p)$
3	(s,...,f,s)	$n-1$	$p \ldots \times p \times (1-p) \times p$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
2^{n-2}	(f,...,s,f)	1	$(1-p) \times (1-p) \ldots \times p \times (1-p)$
2^{n-1}	(f,f,...,s)	1	$(1-p) \times (1-p) \ldots \times p$
2^n	(f,f,...,f)	0	$(1-p) \times (1-p) \ldots \times (1-p)$

n independent trials, X = number of successes

- Consider any outcome that results in a total of i successes.
  - This outcome will have a total of i successes and $(n-i)$ failures.
  - Probability of i success and $(n-i)$ failures $= p^i \times (1-p)^{(n-i)}$

.

- There number of different outcomes that result in i successes and $(n-i)$ failures$=\binom{n}{i}$
- The probability of i successes in n trials is given by

$$P(X=i) = \binom{n}{i} \times p^i \times (1-p)^{(n-i)}$$

☐ Binomial rand. var. :-

X is binomial rand. var. with parameter n & p that represent no. of success in ind. bernouli trials when each trial is a success with prob. p. X ~~takes~~ value $0, 1, 2, \ldots, n$ with the probablity:

$$P(X=i) = \binom{n}{i} \times p^i \times (1-p)^{(n-i)}$$

Example: Tossing a coin thrice

- $S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$
- Success= head, Failure = tails
- X is the random variable which counts the number of heads in the tosses. $n = 3$ $p = 0.5$
  - Probability mass function

X	0	1	2	3
$P(X = x_i)$	$\binom{3}{0}\frac{1}{2}^0\frac{1}{2}^3$	$\binom{3}{1}\frac{1}{2}^1\frac{1}{2}^2$	$\binom{3}{2}\frac{1}{2}^2\frac{1}{2}^1$	$\binom{3}{3}\frac{1}{2}^3\frac{1}{2}^0$
$P(X = x_i)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

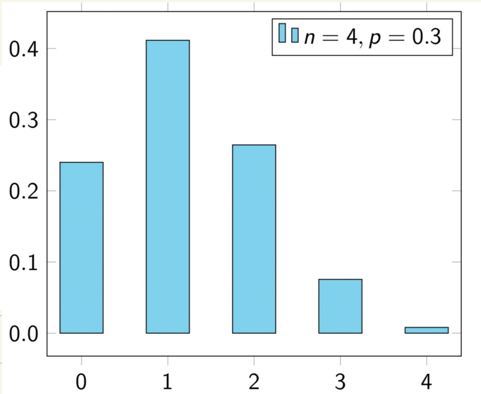
Let $n = 4$ independent Bernoulli trials.

Let $p = 0.3$ is probability of success.

Let X = number of successes in 4 independent trials.

The probability distribution of X

X	0	1	2	3	4
$P(X = i)$	0.2401	0.4116	0.2646	0.0756	0.0081



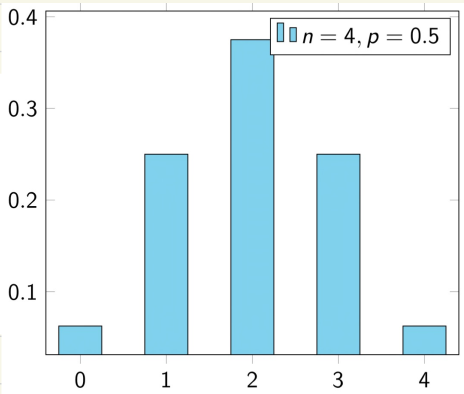
Let $n = 4$ independent Bernoulli trials.

Let $p = 0.5$ is probability of success.

Let X = number of successes in 4 independent trials.

The probability distribution of X

X	0	1	2	3	4
$P(X = i)$	0.0625	0.25	0.375	0.25	0.0625



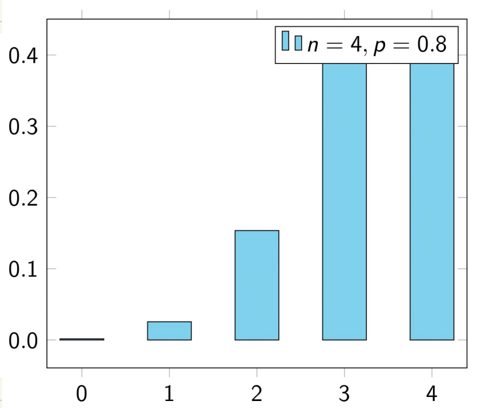
Let $n = 4$ independent Bernoulli trials.

Let $p = 0.8$ is probability of success.

Let X = number of successes in 4 independent trials.

The probability distribution of X

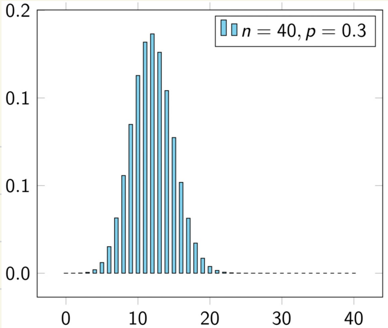
X	0	1	2	3	4
$P(X = i)$	0.0016	0.0256	0.1536	0.4096	0.4096

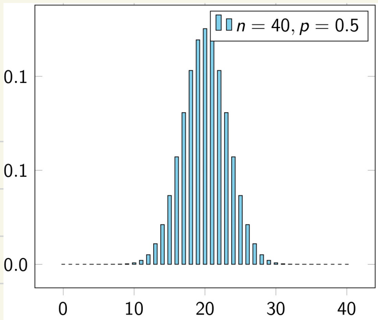


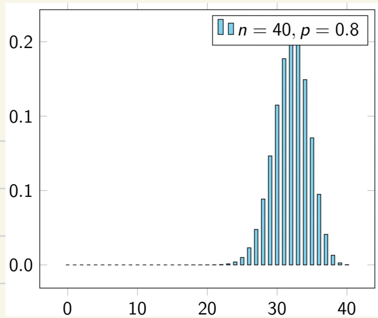
A binomial distribution is

- right skewed if $p < 0.5$
- is symmetric if $p = 0.5$
- is left skewed if $p > 0.5$

. We demonstrate the same for $n = 4$ and different p







Example of binomial distribution:

Consider a company that sells goods in packs of three.

The production process of the goods is not very good and results in 10% of goods being defective.

The company believes that customers will not complain if one out of three in a pack is of bad quality, however, will complain if more than two out of three are of bad quality.

The company wants to keep number of complaints low, say at 3%.

How do we help the company analyse the situation?

Random experiment: Choosing an item and noting its quality.
$S = \{Good, Bad\}$

- Success: good
- Failure: Bad

Given probability of a defective item is 0.1. Hence, Probability of good=p= 0.9.

We want to know number of good items in a pack of three. Hence $n = 3$

Let X = number of good in pack of three. X is a Binomial random variable with $n = 3, p = 0.9$.

The distribution of X is given by

X	0	1	2	3
$P(X = x_i)$	$\binom{3}{0}\frac{9}{10}^0\frac{1}{10}^3$	$\binom{3}{1}\frac{9}{10}^1\frac{1}{10}^2$	$\binom{3}{2}\frac{9}{10}^2\frac{1}{10}^1$	$\binom{3}{3}\frac{9}{10}^3\frac{1}{10}^0$
$P(X = x_i)$	0.001	0.027	0.243	0.729

Customers will complain if they find more than one defective in the pack of three.

- $P(X \leq 1)$

The distribution of X is given by

X	0	1	2	3
$P(X = x_i)$	0.001	0.027	0.243	0.729

$P(X \leq 1) = 0.001 + 0.027 = 0.028$

2.8% is less than 3% which was the goal set by the company- goal achieved.

However, if the company set 2.5% as their threshold then 2.8% would have been more than 2.5% and company would not have achieved its goal.

n							
3	X	0	1	2	3		
	$P(X = i)$	0.001	0.027	0.243	0.729		
4	X	0	1	2	3	4	
	$P(X = i)$	1E-04	0.0036	0.0486	0.2916	0.6561	
5	X	0	1	2	3	4	5
	$P(X = i)$	1E-05	0.00045	0.0081	0.0729	0.32805	0.59049

Examples :

Rolling a dice

Roll four fair dice. Define success as getting a six. Find the probability that

a 6 appears at least once.
b 6 appears exactly once.
b 6 appears at least twice.

- Let $X =$ the number of sixes in four rolls of the dice. Then $X \sim B(4, 1/6)$.
- The pmf is given by

X	0	1	2	3	4
$P(X = i)$	0.4823	0.3858	0.1157	0.0154	0.0008

a 6 appears at least once=0.5177.
b 6 appears exactly once $=$ 0.3858.
c 6 appears at least twice $=$ 0.1319.

Example: Defective ball bearings

Each ball bearing produced is independently defective with probability 0.05. If a sample of 5 is inspected, find the probability that

a None are defective.
b Two or more are defective.

- Let $X =$ the number of defectives in sample of five. Then $X \sim B(5, 0.05)$.
- The pmf is given by

X	0	1	2	3	4	5
$P(X = i)$	0.7738	0.2036	0.0214	0.0011	0.0000	0.0000

a None are defective$= P(X = 0) = 0.7738$
b Two or more are defective$=P(X \geq 2) = 0.0225$

Example: Satellite functioning

A satellite system consists of 4 components and can function if at least 2 of them are working. If each component independently works with probability 0.8, what is the probability the system will function?

- Let $X =$ the number of components among four that are functioning. Then $X \sim B(4, 0.8)$.
- The pmf is given by

X	0	1	2	3	4
$P(X = i)$	0.0016	0.0256	0.1536	0.4096	0.4096

a System will function if$X \geq 2$, $P(X \geq 2) = 0.9728$

A multiple-choice examination has 4 possible answers for each of 5 questions. What is the probability that a student will get 4 or more correct answers just by guessing?

a Let X be number of correct responses. $X \sim B(5, 1/4)$
The pmf is given by

X	0	1	2	3	4	5
$P(X = i)$	0.2373	0.3955	0.2637	0.0879	0.0146	0.0010

b Student will get 4 or more correct answers just by guessing
$X \geq 4, P(X \geq 4) = 0.0156$

- Expectation and variance of binomial random variable

Binomial rand. var. is no. of success in 'n' indep. trial when each trial is a success with prob. p.

Written as $X = X_1 + X_2 + \ldots + X_n$

- $X = X_1 + X_2 + \ldots + X_n$
- $E(X) = E(X_1 + X_2 + \ldots + X_n) = E(X_1) + E(X_2) + \ldots + E(X_n)$
  - $E(X) = p + p + \ldots + p = np$
- $V(X) = V(X_1 + X_2 + \ldots + X_n) = V(X_1) + V(X_2) + \ldots + V(X_n)$
  - $V(X) = p(1-p) + p(1-p) + \ldots + p(1-p) = np(1-p)$

From fact, exp. of sum of rand. var. is sum of its exp.:

$E(X) = np$

&, Var. of sum of ind. var. is sum of their variance:

$Var(X) = np(1-p)$

Example: Tossing a coin 500 times

If a fair coin is tossed 500 times, what is the standard deviation of the number of times that a head appears?

Let X = the number of heads in 500 tosses of a fair coin. Then $X \sim B(500, 1/2)$. $V(X) = 125$, $SD(X) = \sqrt{125} = 11.1803$

Finding probability given expectation and n

The expected number of heads in a series of 10 tosses of a coin is 6. What is the probability there are 8 heads?

Let X be number of heads. $X \sim B(10, p)$.

1. Since $E(X) = np$; $10p = 6$, hence $p = 0.6$
2. Prob there are 8 heads; $P(X = 8) = 0.121$

Find pmf given Expectation and Variance

If X is a binomial random variable with expected value 4.5 and variance 0.45, find

a $P(X = 3)$

b $P(X \geq 4)$

- $X \sim B(n, p)$
- $np = 4.5$, $np(1-p) = 0.45$.
- Solving gives $n = 5$ and $p = 0.9$

a $P(X = 3) = 0.0729$

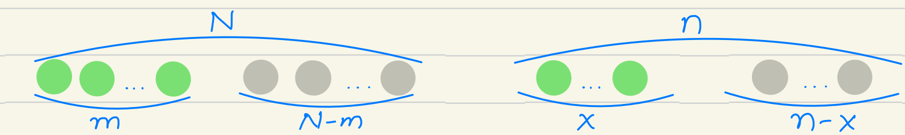
b $P(X \geq 4) = 0.9185$

Hypergeometric distribution:

For this to work,

→ Population must be divided into only 2 ind. subset.
→ Exp must have changing subset with each experiment.
→ Sampling without replacement.

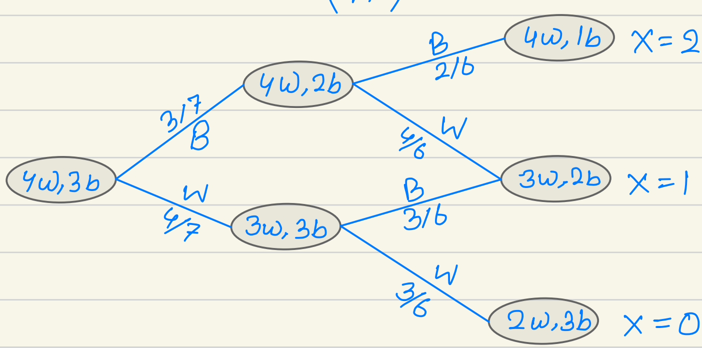
If we randomly select 'n' item without replacement from a set of N item in which 'm' items are 1 type & 'N-m' are another type



let x be no. of items of type 1, then PMF of descrete rand. var. X is called hypergeometric dist.

$$P(X=x) = \frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}} \quad ; \; x = 0,1,\dots,n$$

$$; x \leq n, x \leq m, n-x \leq N-m$$



Examples: Choosing balls without replacement

A bag consists of 50 balls of which 30 are white and 20 are blue. A student randomly samples five balls without replacement. Let X be the number of blue balls selected.

- Here, $N = 50, n = 5, m = 20$
- X takes values: $0, 1, 2, 3, 4, 5$
- $P(X = i) = \dfrac{\binom{20}{i}\binom{30}{5-i}}{\binom{50}{5}}; i = 0, 1, 2, 3, 4, 5$
- The pmf

i	0	1	2	3	4	5
P(X=i)	0.07	0.26	0.36	0.23	0.07	0.01

Example: Voters

- Assume there are 150 female voters and 250 male voters in a particular locality. If a group of twenty five voters is selected at random, then the probability that ten of the selected voters would be female can be calculated with the help of hypergeometric probability distribution.
- In this case: $N = 400, n = 25, m = 150$
- $$P(X = 10) = \frac{\binom{150}{10}\binom{250}{15}}{\binom{400}{25}}$$

Example: Defectives-1

In a batch of 10 computer parts it is known that there are three defective parts. Four of the parts are selected at random to be tested. Define the random variable X to be the number of working (non defective) computer parts selected

$$\overbrace{\underbrace{\boxed{N}\boxed{N}\boxed{N}\boxed{N}\boxed{N}\boxed{N}\boxed{N}}_{m=7}\ \underbrace{\boxed{D}\boxed{D}\boxed{D}}_{N-m=3}}^{N=10} \curvearrowright \overbrace{\underbrace{\boxed{N}\boxed{N}}_{x=2}\ \underbrace{\boxed{D}\boxed{D}}_{n-x=2}}^{n=4}$$

This is a Hypergeometric distribution with
$N = 10, n = 4, m = 7$

The pmf $P(X = x) = \dfrac{\binom{7}{x}\binom{3}{4-x}}{\binom{10}{4}}; x = 0, 1, 2, 3, 4$

Example: Defectives-2

In a batch of 10 computer parts it is known that there are three defective parts. Four of the parts are selected at random to be tested. Define the random variable X to be the number of defective computer parts selected

$$\overbrace{\underbrace{\boxed{N}\boxed{N}\boxed{N}\boxed{N}\boxed{N}\boxed{N}\boxed{N}}_{N-m=7}\ \underbrace{\boxed{D}\boxed{D}\boxed{D}}_{m=3}}^{N=10} \curvearrowright \overbrace{\underbrace{\boxed{N}\boxed{N}}_{x=2}\ \underbrace{\boxed{D}\boxed{D}}_{n-x=2}}^{n=4}$$

This is a Hypergeometric distribution with
$N = 10, n = 4, m = 3$

The pmf $P(X = x) = \dfrac{\binom{3}{x}\binom{7}{4-x}}{\binom{10}{4}}; x = 0, 1, 2, 3$

Example: Sampling from a deck of cards

Take a deck of 52 cards. Draw five cards from the deck. Let the random variable X denote the number of aces in the random sample of five cards. What is the probability distribution of X?

This is a Hypergeometric distribution with
$N = 52, n = 5, m = 4$

The pmf is given by

$$P(X = x) = \frac{\binom{4}{x}\binom{48}{5-x}}{\binom{52}{5}}; x = 0, 1, 2, 3, 4$$

Expectation

Let X follow a hypergeometric distribution in which n objects are selected from N objects with m of the objects being one type, and $N-m$ of the objects being a second type. What is the expected value of X?

$$E(X) = \frac{nm}{N}$$

Expectation- proof

$$E(X) = \sum_x x \frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}$$

Now, $\binom{m}{x} = \frac{m!}{x!(m-x)!}$, and, $\binom{N}{n} = \frac{N!}{n!(N-n)!} =$
$\frac{N(N-1)!}{n\cdot(n-1)!(N-n)!} = \frac{N}{n}\cdot\frac{(N-1)!}{(n-1)!(N-1-(n-1))!} = \frac{N}{n}\cdot\Big(\begin{matrix}N -\\ n -\end{matrix}$

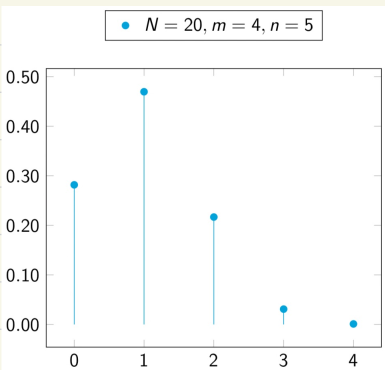
Hence, $E(X) = \sum_x x \frac{\frac{m\cdot(m-1)!}{x!(m-x)!}\binom{N-m}{n-x}}{\frac{N}{n}\cdot\binom{N-1}{n-1}} =$

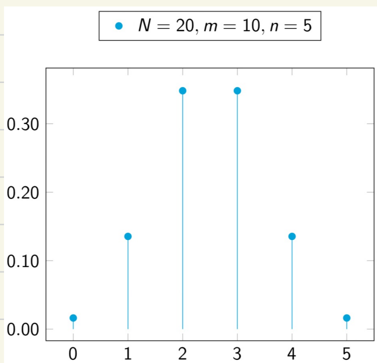
$$\sum_x \frac{nm}{N}\frac{\frac{(m-1)!}{(x-1)!(m-1-(x-1))!}\binom{(N-1)-(m-1)}{(n-1)-(x-1)}}{\binom{N-1}{n-1}} = \frac{nm}{N}$$

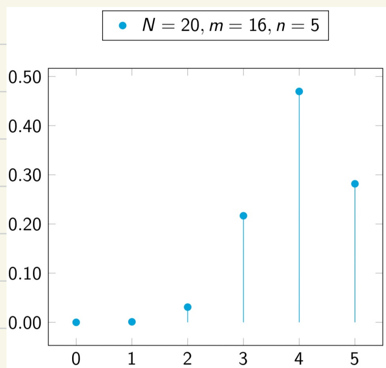
○ Variance:

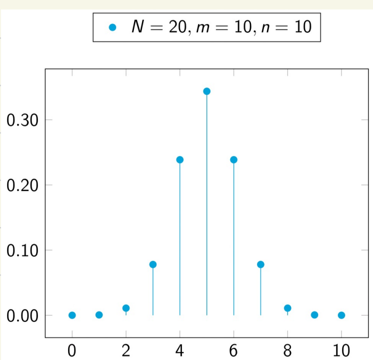
let X follow a hypergeometric distribution in which 'n' objects are selected from N objects with 'm' of object in one type & N-m of objects being second type.

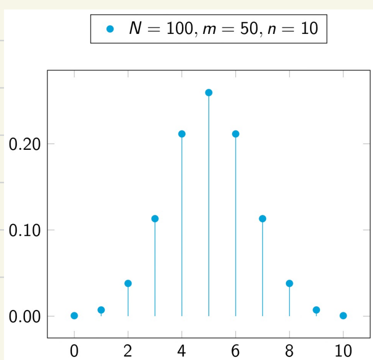
$$Var(X) = n\left(\frac{m}{N}\right)\left(\frac{N-m}{N}\right)\left(\frac{N-n}{N-1}\right)$$

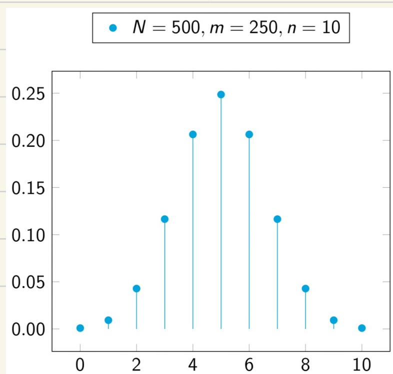






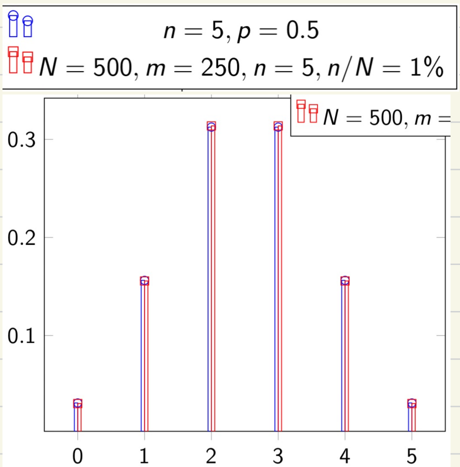


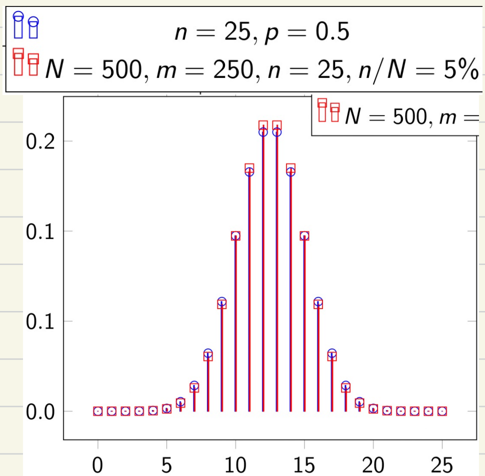


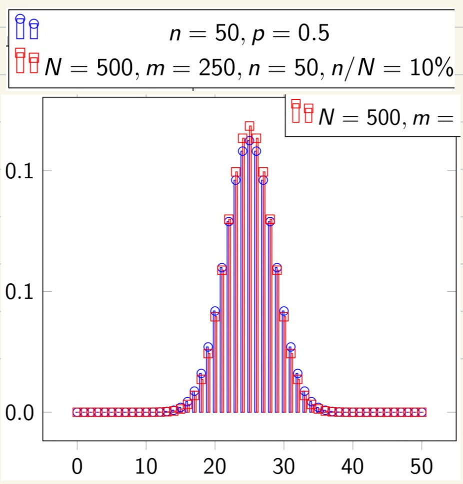


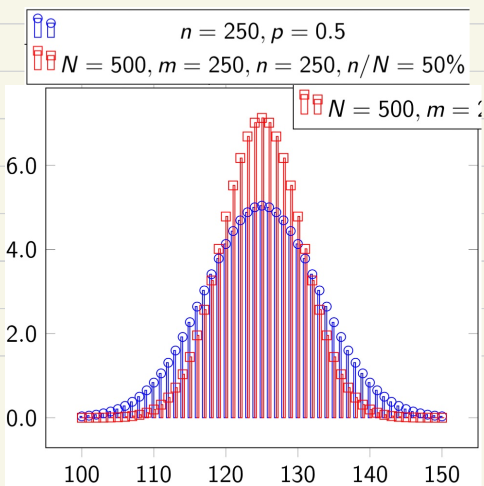
Expectation and variance

- $X \sim Hypergeometric(N, m, n)$
  - $E(X) = \dfrac{nm}{N}$
  - $Var(X) = n\frac{m}{N}\frac{N-m}{N}\frac{N-n}{N-1}$
- $Y \sim Bin\left(n, \dfrac{m}{N}\right)$
  - $E(X) = \dfrac{nm}{N}$
  - $Var(X) = n\frac{m}{N}\frac{N-m}{N}$
- $\frac{N-n}{N-1}$ is known as finite population correction
  - For $n = 1$, replacement has no effect both are Bernoulli trial
  - For $n = N$, the whole population is sampled- hence variance is zero.
- If the population N is very large compared to the sample size n (i.e. $N >> n$) then Hypergeometric(N, m, n) is about Binomial$(n, \dfrac{m}{N})$.







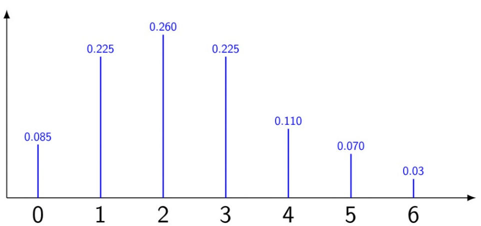


- Poisson distribution: It give probab. of no. of events occuring in fixed interval of time or space.
  - Assuming these events happens with known avg. rate, λ, & ind. of time since last event.
  - X denote no. of times an event occurs.
  - We say $X \sim \text{Poisson}(\lambda)$ & used as approx. of binomial dist. if prob. of success is 'small' & no. of trial is 'large'.

Motivation example

Consider a researcher who is observing the number of vehicles that pass a busy traffic intersection in a day. She collects data comprising of 1000 one minute intervals and tabulates the same in form of a frequency table given below.

Number of vehicles	0	1	2	3	4	5	> 6
Count	80	225	260	225	110	70	30



x	Freq f	Rel Freq f_r	$f_r x$	$f_r x^2$
0	80	0.08	0	0
1	225	0.225	0.225	0.225
2	260	0.26	0.52	1.04
3	225	0.225	0.675	2.025
4	110	0.11	0.44	1.76
5	70	0.07	0.35	1.75
6	30	0.03	0.18	1.08
	1000	1	2.39	7.88

Mean$= 2.39$

Variance $= 7.88 - 2.39^2 = 2.16$

Observations

Number of vehicles passing a traffic intersection are at random and independently of each other

The average number of vehicles per minute is about 2.39 which is equivalent to 143 per hour.

Question: What is the appropriate probability distribution to model the number of vehicles passing a traffic intersection?

- Poisson

Poisson as Binomial approximation

Define "success" as exactly one event happening in a short interval of length δt

The n events happening in interval of length t can be viewed as n successes happening in n intervals of length δt, with each one of them being an independent and identical trial.

Hence the problem can be viewed as a $Bin\left(n, p = \dfrac{\lambda}{n}\right)$ experiment.

Derivation- contd

$$= \binom{n}{x}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{n(n-1)\dots(n-x+1)}{x!}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-x}$$

$$= \frac{\lambda^x}{x!}\left(\frac{n(n-1)\dots(n-x+1)}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-x}$$

$$= \frac{\lambda^x}{x!}\left(\frac{n^x(1-\frac{1}{n})\dots(1-\frac{(x-1)}{n})}{n^x}\right)\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-x}$$

Now let's make the intervals very small, i.e, $\delta t \longrightarrow 0$ or $n \longrightarrow \infty$

$$= \lim_{n\longrightarrow\infty}\binom{n}{x}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$= \lim_{n\longrightarrow\infty}\frac{\lambda^x}{x!}\underbrace{\left(\frac{n^x(1-\frac{1}{n})\dots(1-\frac{(x-1)}{n})}{n^x}\right)}_{\to 1 \text{ as } n\to\infty}\underbrace{\left(1-\frac{\lambda}{n}\right)^n}_{\to e^{-\lambda} \text{ as } n\to\infty}\underbrace{\left(1-\frac{\lambda}{n}\right)^{-x}}_{\to 1 \text{ as } n\to\infty}$$

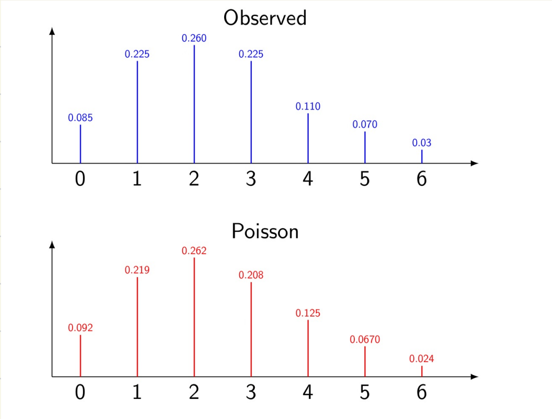
$$= \frac{\lambda^x}{x!}e^{-\lambda}$$

- Probablity mass fn. of Poisson: This distribution is avg. no of λ events per interval, is poison descrete rand. var. $X \sim Poisson(\lambda)$

$$P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \dots$$

→ Rand. var. X represent no. of interval per time interval.
→ e is mathematical constant 2.718.

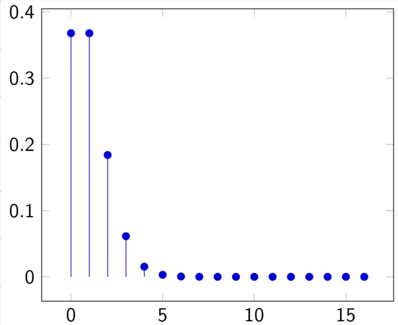
x	freq	Prob
0	80	0.092
1	225	0.219
2	260	0.262
3	225	0.208
4	110	0.125
5	70	0.060
6	30	0.024



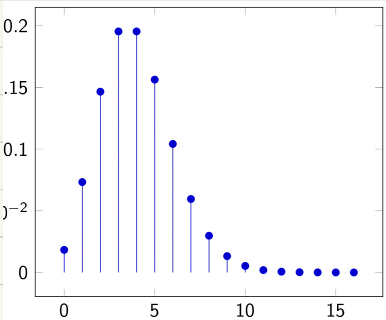
The shape of the Poisson distribution depends on the value of the parameter λ.

If λ is small the distribution has positive skew, but as λ increases the distribution becomes progressively more symmetrical.

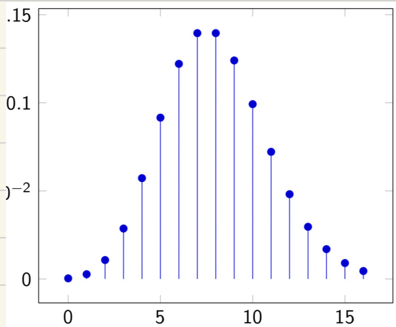
Graph of pmf for $\lambda = 1$



Graph of pmf for $\lambda = 4$



Graph of pmf for $\lambda = 8$



○ Expectation of poison distribution: 'λ'

$$
\begin{aligned}
E(X) &= \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} \\
&= \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda}{x} \frac{\lambda^{(x-1)}}{(x-1)!} \\
&= e^{-\lambda}\lambda \sum_{x=1}^{\infty} \frac{\lambda^{(x-1)}}{(x-1)!} \\
&= e^{-\lambda}\lambda \left(\frac{\lambda^0}{0!} + \frac{\lambda^1}{1!} + \dots \right) \\
&= e^{-\lambda}\lambda e^{\lambda} \\
&= \lambda
\end{aligned}
$$

○ Variance of poison distribution: 'λ^2'

$$
\begin{aligned}
E(X(X-1)) &= \sum_{x=0}^{\infty} x(x-1) \frac{e^{-\lambda}\lambda^x}{x!} \\
&= \sum_{x=0}^{\infty} x(x-1) \frac{e^{-\lambda}\lambda^2}{x(x-1)} \frac{\lambda^{(x-2)}}{(x-2)!} \\
&= e^{-\lambda}\lambda^2 \sum_{x=2}^{\infty} \frac{\lambda^{(x-2)}}{(x-2)!} \\
&= e^{-\lambda}\lambda^2 \left(\frac{\lambda^0}{0!} + \frac{\lambda^1}{1!} + \dots \right) \\
&= e^{-\lambda}\lambda^2 e^{\lambda} \\
&= \lambda^2
\end{aligned}
$$

Variance of Poisson distribution

Now, $E(X^2) = E(X(X-1)) + E(X) = \lambda^2 + \lambda$.

Hence

$$Var(X) = E(X^2) - (E(X))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

For a Poisson random variable $X \sim Poisson(\lambda)$, both the expected value and the variance of X are equal to λ.

Examples of Poisson distribution

Events occurring in fixed interval of time

1. Number of vehicles passing through a traffic intersection in a fixed time interval of one minute.
2. Number of people withdrawing money from a bank in a fixed time interval of fifteen minutes.
3. Number of telephone calls received per minute at a call center

Events occurring in fixed interval of space

1. Number of typos (incorrect spelling) in a book.
2. Number of defects in a wire cable of finite length.
3. Number of defects per meter in a roll of cloth

Modeling number of accidents

Suppose the number of accidents per week in a factory can be modeled by the Poisson distribution with a mean of 0.5.

1. Find the probability that in a particular week there will be less than two accidents?
   Let X be the number of accidents per week in the factory.
   We have $X \sim Poisson(\lambda = 0.5)$
   Need to find $P(X \leq 2) = \sum_{i=0}^{2} \frac{e^{-0.5} \times 0.5^i}{i!} =$
   $=0.6065 + 0.3033 + 0.0758 = 0.9856$

Modeling the number of defects

Suppose the number of defects in a wire cable can be modeled by the Poisson distribution with a of 0.5 defects per meter.

1. Find the probability that a single meter of wire will have exactly 3 defects
   Let X be the number of defects per meter
   We have $X \sim Poisson(\lambda = 0, 5)$
   Need to find $P(X = 3) = \frac{e^{-3} \times 0.5^3}{3!} = 0.0126$

Modeling number of killings

The number of dogs that are killed on a particular stretch of road in Chennai in any one day can be modeled by a Poisson(0.42) random variable.

1. Calculate the probability that exactly two dogs are killed on a given day on this stretch of road.
   Let $X=$ number of dogs killed in one day
   $X \sim Poisson(\lambda = 0.42)$
   $P(X = 2) = \frac{e^{-0.42} \times 0.42^2}{2!} = 0.058$
2. Find the probability that exactly two dogs are killed over a 5-day period on this stretch of road.
   Let $X=$ number of dogs killed in five day
   $X \sim Poisson(\lambda = 2.1)$
   $P(X = 2) = \frac{e^{-2.1} \times 2.1^2}{2!} = 0.27$

Modeling typos

A typist makes 1500 mistakes in a book of 500 pages. Let X be number of mistakes per page. Then, $X \sim Poisson(3)$
On how many pages would you expect to find

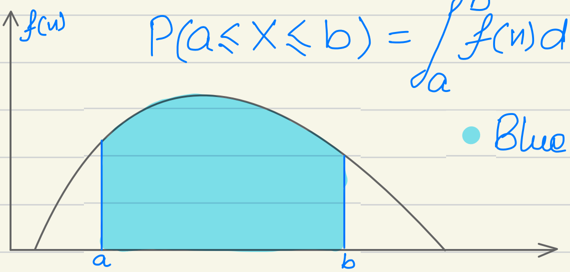
1. no mistake
   $P(X = 0) = \frac{e^{-3} \times 3^0}{0!} = 0.0498$
   Number of pages with no mistakes $= 0.0498 \times 500 \approx 25$ pages
2. one mistake
   $P(X = 1) = \frac{e^{-3} \times 3^1}{1!} = 0.1494$
   Number of pages with no mistakes $= 0.1494 \times 500 \approx 75$ pages
3. three or more mistakes
   $P(X >= 31) = 1 - \frac{e^{-3} \times 3^0}{0!} - \frac{e^{-3} \times 3^1}{1!} - \frac{e^{-3} \times 3^2}{2!} = 0.5768$
   Number of pages with no mistakes $= 0.5768 \times 500 \approx 288$ pages

- Continous random variable : It is the one that has possible values that form an interval along real line. It can assume any value in an interval.

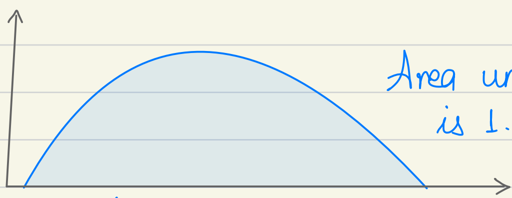
- Probablity density fn :
  Every Continous random variable X has a curve associated with it.
  Probablity dist. curve of a continous rand. var. is also called probab. density fn. denoted by $f(x)$.

Area under PDF :
- Consider any 2 pt. 'a' & 'b', where 'a' is less than 'b'.
- The prob. that X assumes a value that lies b/w a & b is equal to area under curve b/w a & b.
  $P(X \in [a,b]) = P(a \leq X \leq b)$ is area under curve b/w a & b.

$$P(a \leq X \leq b) = \int_a^b f(x)\,dx$$



- Blue area b/w 0 & 1.

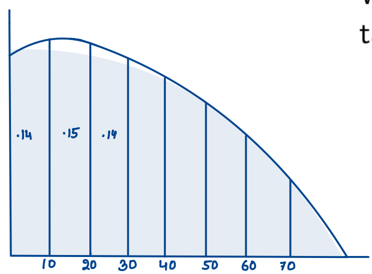


Area under entire curve is 1.

- Here open & close interval don't matter.

Example :

Figure below is a probability density function for the random variable that represents the time (in minutes) it takes a repairer to service a television. The numbers in the regions represent the areas of those regions.



What is the probability that the repairer takes

1. Less than 20 =0.29
2. Less than 40 =0.56
3. More than 50 =0.33
4. Between 40 and 70 minutes to complete a repair? =0.27

- Cumulative dist. fn :

For continous rand. var X

$$F(a) = P(X \leqslant a) = \int_{-\infty}^{a} f(x)dx$$

Since, prob. of cont. rand. var. X assume single value is always zero so it will be same as above.

- Expectation & variance:

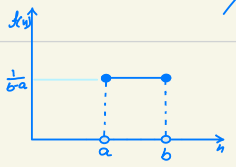
Expected value : $E(X) = \int x f(x) dx$

Variance : $\int (x-E(X))^2 f(x) dx$

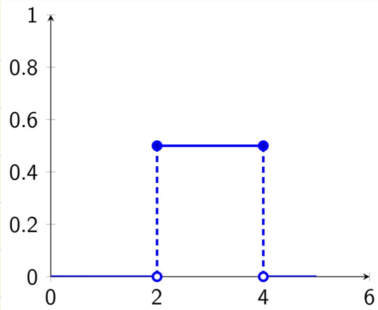
- Uniform distribution : A continous rand. var. has a uniform distribution, denoted $X \sim U(a,b)$.

Probab. density fn is:

$$f(x) \begin{cases} 1/(b-a) & a<x<b \\ 0 & \text{otherwise} \end{cases}$$



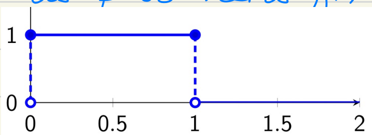
Graph of pdf of a Uniform distribution $U(2,4)$



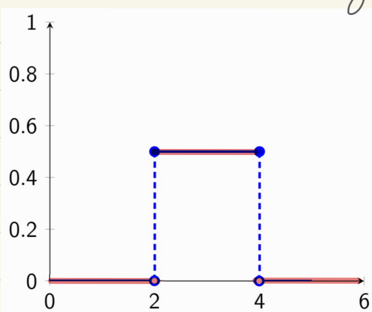
$a=2$
$b=4$
$b-a=2$
hence, $\frac{1}{b-a} = \frac{1}{2} = 0.5$

- Standard uniform distribution:

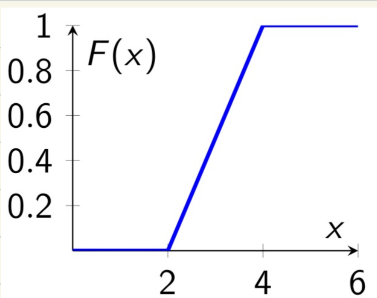
A rand. var. has std. uniform dist. with min 0 & max 1 if its prob. dens. fn. given by: $f(x) \begin{cases} 1 & 0<x<1 \\ 0 & 0 \text{ otherwise} \end{cases}$



- Cumulative dist. of Uniform dist:

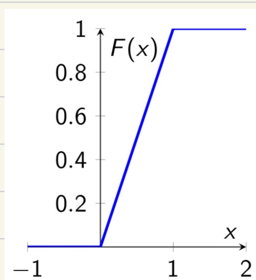


$$f(x) \begin{cases} 0 & \text{for } (x<a) \\ \frac{x-a}{b-a} & \text{for } (x \in [a,b)) \\ 1 & \text{for } (x \geq b) \end{cases}$$



When $x<a$ $F(x)=0$
When $x>b$ F(x) =1
$a<x<b$ The slope of the line between and is $\frac{1}{(b-a)}$.

For $X \sim U(0,1)$



$$F(x) = \begin{cases} 0 & \text{for } x<0 \\ x & \text{for } x \in [0,1) \\ 1 & \text{for } x \geq 1 \end{cases}$$

○ Expectation & variance of uniform rand. var.:

□ Expectation of $(X \sim U(a,b))$;

$$E(X) = \frac{a+b}{2}$$

Derive: $E(X) = \int_a^b x f(x)\, dx = \int_a^b x \frac{1}{b-a}\, dx = \frac{1}{b-a} \frac{x^2}{2}\Big|_a^b = \frac{b+a}{2}$

□ Variance of $X \sim U(a,b)$;

$$Var(X) = \frac{(b-a)^2}{12}$$

Derive ⇒ $Var(X) = E(X^2) - [E(X)]^2 = \frac{b^2+a^2+ab}{3} - \left(\frac{b+a}{2}\right)^2 = \frac{(b-a)^2}{12}$

Example: Computing probabilities given distribution

Suppose that X is a uniform random variable over the interval $(0, 1)$. Find

1. $P(X > 1/3) = 2/3$
2. $P(X \leq 0.7) = 0.7$
3. $P(0.3 < X \leq 0.9) = 0.6$
4. $P(0.2 \leq X < 0.8) = 0.6$

You are to meet a friend at 2 p.m. However, while you are always exactly on time, your friend is always late and indeed will arrive at the meeting place at a time uniformly distributed between 2 and 3 p.m. Find the probability that you will have to wait

1. At least 30 minutes
2. Less than 15 minutes
3. Between 10 and 35 minutes
4. Less than 45 minutes

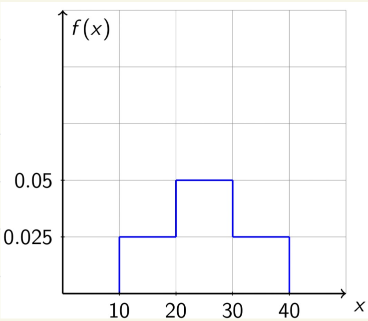
Example: Question

Answer

Let X denote the amount of time you will have to wait. $X \sim U(0, 60)$

1. At least 30 minutes $= P(X \geq 30) = 30/60 = 1/2$
2. Less than 15 minutes $= P(X < 15) = 15/60 = 1/4$
3. Between 10 and 35 minutes $= P(10 \leq X \leq 35) = 25/60 = 5/12$
4. Less than 45 minutes $= P(X < 45) = 45/60 = 3/4$

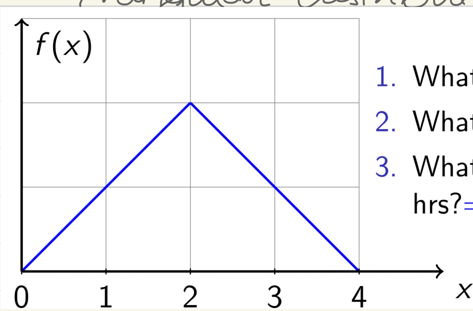
- **Non-uniform distribution**: These are dist. which are not same at all point of time in a interval.



Let X be amount of playing time in minutes. Find the probability that the player plays

1. Over 20 minutes $= P(X > 20) = 0.5 + 0.25 = 0.75$
2. Less than 25 minutes$= P(X < 25) = 0.25 + 0.25 = 0.5$
3. Between 15 and 35 minutes$= P(15 \leq X \leq 35) = 0.125 + 0.5 + 0.125 = 0.75$
4. More than 35 minutes$= P(X > 35) = 0.125$

- **Trangular distribution**: To understand it do a ex. of Amt. of time someone studied for 4 hr.

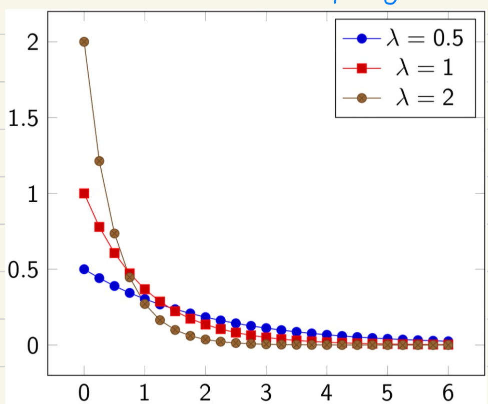


1. What is the height of the curve at the value 2?$= 1/2\,unit$
2. What is the probability she will study more than 3 hrs?$= 1/8$
3. What is the probability she will study between 1 and 3 hrs?$= 3/4$

- **Exponential distribution**: A rand. var. whose pdf is given as

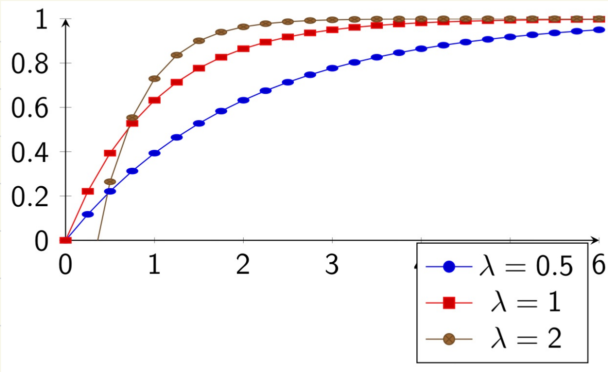
$$f(x) \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

is said to be exp. distribution with perameter λ.



– cdf of exp. fn:

$$F(a) = P(X \leq a)$$
$$= \int_0^a \lambda e^{-\lambda x} dx$$
$$= -e^{-\lambda x} \Big|_0^a$$
$$= 1 - e^{-\lambda a}$$



○ Exponent and variance of exp. distribution:

$$X \sim exp(\lambda)$$
$$E(X) = \frac{1}{\lambda}$$
$$E(X^2) = \frac{1}{\lambda^2}$$
$$Var(E) = \frac{1}{\lambda^2}$$

In practice, the exponential distribution often arises as the distribution of the amount of time until some specific event occurs.

Example:

- Suppose that the length of a phone call in minutes is an exponential random variable with parameter $\lambda = 0.1$. If someone arrives immediately ahead of you at a public telephone booth, find the probability that you will have to wait
  - a more than 10 minutes
  - b between 10 and 20 minutes.

Solution Let X denote the length of the call made by the person in the booth. $X \sim exp(0.5)$

a more than 10 minutes $= P(X > 10) = e^{-1} \approx 0.368$

b between 10 and 20 minutes$=$
$P(10 < X < 20) = F(20) - F(10) = e^{-1} - e^{-2} \approx 0.233$