Probabily

and

Statistics

24-1-23

Tuesday

Lecture # 1

population: all observations

sample: representative data

- where there is data, there is statistics
- statistics can be defined as:

⇒ numerical collection of data

⇒ descipline

⇒ quantities calculated from sample

- parameters ⇒ quantities calculated from populated

- variables

⇒ qualitative: cannot be expressed numerically e.g hair colour, gender

⇒ quantitative: can be expressed numerically e.g eye-sight

1. discrete: includes jumps/gaps. e.g no of siblings (whole numbers)
2. continuous : continuous variation of data e.g time, weight

statistics

descriptive

- measurement of central tendency:

⇒ mean

⇒ median

⇒ mode

- measurement of dispertion:

⇒ standard deviation

⇒ variation

* if variation is less, average is good

infuential

- andaza lagana on the basis of sample.

⇒ estimation

⇒ prediction

Mean: average. (flaw: out liers/ extreme values)

1. unqrouped data:

$$\bar{x} = \frac{\Sigma x}{n}$$

2. grouped data:

$$\bar{x} = \frac{\Sigma fx}{\Sigma f} \qquad (\Sigma f = n)$$

Median: divides sorted data into two halves

1. if $\frac{n}{2}$ is not an integer (odd):

$\left(\frac{n+1}{2}\right)^{th}$ value is answer

2. if $\frac{n}{2}$ is an integer (even):

$$\left(\frac{\frac{n}{2} + \left(\frac{n}{2}+1\right)}{2}\right)^{th}$$ value is answer

1.1 The following measurements were recorded for drying time in hours of a certain brand of lytix paint of prime time

3.4, 2.5, 4.8, 2.9, 3.6, 2.8, 3.3, 5.6, 3.7, 2.8, 4.4, 4.0, 5.2, 3.0, 4.8

sorted data: 2.5, 2.8, 2.8, 2.9, 3.0, 3.3, 3.4, 3.6, 3.7, 4.0, 4.4, 4.8, 4.8, 5.2, 5.6

mean $= \frac{56.8}{15} = 3.79$

median $= 3.6$

mode:

2, 3, 5, 7, 5, 8, 10, 5, 3

mode = 5, 7

⇒ bimodal (has 2 modes)

* unimodal (has 1 mode)
* multimodal

H.W

1.2 for practise

Assignment Q's:

1.3 part c, d

1.4 part a

1.21 (length of power failure), 1.30 (lifetime of frosted lamps)

Lecture # 2

Mean grouped data:

Q. calculate the mean weight of apples from given data organized in a frequency (grouped data) distribution

weight (grams)	f	X (midpoint)	fX
65-84	9	74.5	670.5
85-104	10	94.5	945
105-124	17	114.5	1946.5
125-144	10	134.5	1345
145-164	5	154.5	772.5
165-184	4	174.5	698
185-204	5	194.5	972.5
	$\Sigma f = 60$		$\Sigma fx = 7350$

mean for grouped data:

$$\bar{X} = \frac{\Sigma fx}{\Sigma f} = \frac{7350}{60} = 122.5 \text{ grams}$$

Median grouped data:

Q. The weight in mg of 2538 seeds of a pine leaf are as follows

	weight (grams)	f	class boundaries	commulative frequency
	10-24.9	16	9.95-24.95	16
	25-39.9	68	24.95-39.95	84
	40-54.9	204	39.95-54.95	288
	55-69.9	233	54.95-69.95	521
q_1	70-84.9	240	69.95-84.95	761
q_2	85-99.9	655	84.95-99.95	1416
q_3	100-114.9	803	99.95-114.95	2219
	115-129.9	294	114.95-129.95	2513
	130-144.9	21	129.95-144.95	2534
	145-159.9	4	144.95-159.95	2538
		$\Sigma f = 2538$		

$$l + \frac{h}{f}\left(\underset{\Sigma f/2}{\frac{n}{2}} - c\right)$$

l = lower class boundary of median class

h = interval difference

f = frequency of median class

c = commulative frequency of preceding class

$$\Rightarrow 84.95 + \frac{15}{655}(1269 - 761)$$

$$= 96.58 \text{ mg}$$

Scanned with CamScanner

Date:

Mode of grouped data: (previous example)

(no need of cf)

$$l + \frac{(f_m - f_1)}{(f_m - f_1) + (f_m - f_2)} \times h$$

l = lower class boundary of model class

h = interval difference

f_m = frequency of model class

f_1 = frequency of preceding class

f_2 = frequency of next class

$\Rightarrow$ putting the values

$$99.95 + \frac{(803-655)}{(803-655)+(803-294)} \times 15$$

$= 103.32$

Trim mean:

- to treat out liers — trimmed by x% → iska mean
- data sort kina paray ga

1.1. Trim by 20%

~~2.5~~, ~~2.8~~, ~~2.8~~, 2.9, 3.0, 3.3, 3.4, 3.6, 3.7, 4.0, 4.4, 4.8, ~~4.8~~, ~~5.2~~, ~~5.6~~

20% of 15 = 3

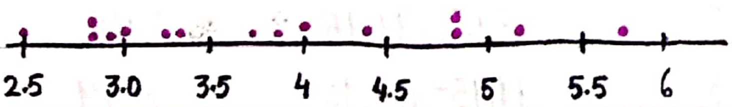
Trim 3 from both sides then find mean

$$\Rightarrow \frac{33.1}{9}$$

$= 3.67$

Dot plot:

1.1.

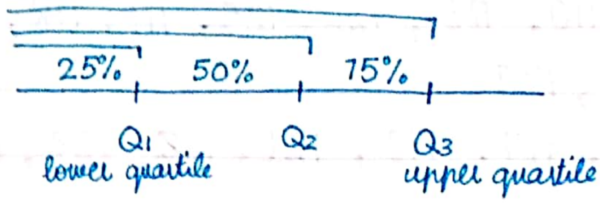


30-1-23 Wednesday

Lecture # 3

Quartile :

3 values which divide data in four equal parts.



ungrouped :

odd :

$$Q_1 = \left(\frac{n+1}{4}\right)^{th} \text{value}, \quad Q_2 = \left(\frac{2n+1}{4}\right)^{th} \text{value}, \quad Q_3 = \left(\frac{3n+1}{4}\right)^{th} \text{value}$$

even :

$$Q_1 = \frac{1}{2}\left[\left(\frac{n}{4}\right)^{th} \text{value} + \left(\frac{n}{4}+1\right)^{th} \text{value}\right] \quad Q_3 = \frac{1}{2}\left[\left(\frac{3n}{4}\right)^{th} \text{value} + \left(\frac{3n}{4}+1\right)^{th} \text{value}\right]$$

Deciles :

9 values which divide data in ten equal parts.

ungrouped :

odd :

$$D_j = \left(\frac{jn+1}{10}\right)^{th} \text{value} \qquad j = 1, 2, 3, \dots, 9$$

even :

$$D_j = \frac{1}{2}\left[\left(\frac{jn}{10}\right)^{th} \text{value} + \left(\frac{jn}{10}+1\right)^{th} \text{value}\right]$$

Percentiles :

99 values which divide data in 100 equal parts.

ungrouped :

odd :

$$P_i = \left(\frac{in+1}{100}\right)^{th} \text{value} \qquad i = 1, 2, 3, \dots, 99$$

even :

$$P_i = \left[\left(\frac{in}{100}\right)^{th} \text{value} + \left(\frac{in}{100}+1\right)^{th} \text{value}\right]$$

Date:

Q. A laptop manufacturing company is interested in determining the life time of a certain type of battery. A sample in hrs of life is as follows:

123, 116, 122, 110, 175, 126, 125, 111, 118, 117, 115, 121, 129, 131, 127

sorted: 110, 111, 115, 116, 117, 118, 121, 122, 123, 125, 126, 127 129 131 175

⇒ Find lower and upper quartile

⇒ comment on 7^{th} decile

⇒ obtain 45^{th} percentile

as $\frac{n}{4}$ is not an integer, data is odd

lower quartile:

$\left(\frac{15+1}{4}\right)^{th}$ value $= 4^{th}$ value $= 116$

upper quartile:

$\left(\frac{3(15)+1}{4}\right)^{th}$ value $= 12^{th}$ value $= 127$

7^{th} decile:

$D_7 = \left(\frac{(7\times15)+1}{10}\right)^{th}$ value $= 11^{th}$ value $= 126$

45^{th} percentile:

$P_{45} = \left(\frac{45(15)+1}{100}\right)^{th}$ value $= 7^{th}$ value $= 121$

Quartiles :

Grouped data:

Q_1 : $\frac{n}{4}$ (which lies in one of the groups)

$$Q_1 = l + \frac{h}{f}\left(\frac{n}{4} - c\right)$$

$$Q_3 = l + \frac{h}{f}\left(\frac{3n}{4} - c\right)$$

$$Q_2 = l + \frac{h}{f}\left(\frac{n}{2} - c\right)$$

Deciles :

D_j : $\frac{jn}{10}$ (lies in which group)

$$D_j = l + \frac{h}{f}\left(\frac{jn}{10} - c\right)$$

Percentiles :

P_i : $\frac{in}{100}$ (lies in which group)

$$P_i = l + \frac{h}{f}\left(\frac{in}{100} - c\right)$$

Scanned with CamScanner

Date:

weight (milligrams)	No. of seeds f	class boundaries	cF
10 - 24.9	16	9.95 - 24.95	16
25 - 39.9	68	24.95 - 39.95	84
40 - 54.9	204	39.95 - 54.95	288
55 - 69.9	233	54.95 - 69.95	521
70 - 84.9	240	69.95 - 84.95	761
85 - 99.9	655	84.95 - 99.95	1416
100 - 114.9	803	99.95 - 114.95	2219
115 - 129.9	294	114.95 - 129.95	2513
130 - 144.9	21	129.95 - 144.95	2534
145 - 159.9	4	144.95 - 159.95	2538

$$Q_1 = \frac{2538}{4} = 634.5$$

$$Q_1 = 69.95 + \frac{15}{240}\left(634.5 - 521\right) = 77.04$$

Scanned with CamScanner

1-2-23 Monday

Lecture # 4

Five point summary:

minimum, Q1, Q2, Q3, maximum ⇒ five points which define the whole data

Box and whisker plot:

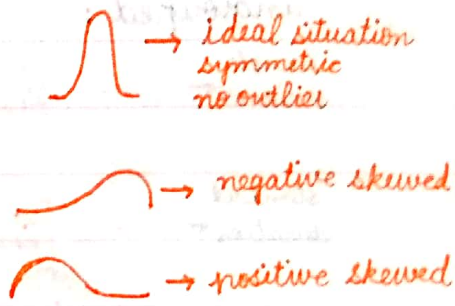
⇒ to determine general direction of data ⇒ to detect outliers

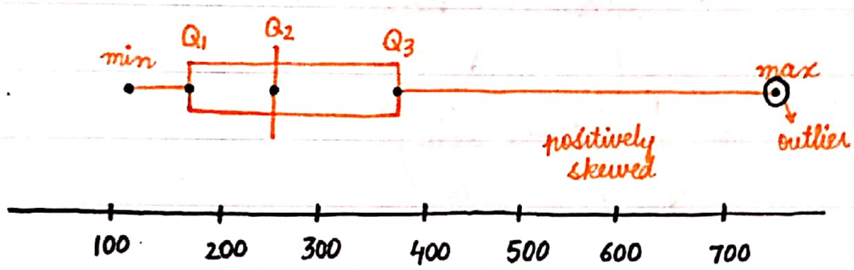
Q. The time in s that it took for each of 12 vehicles to exit a parking lot in a specific location is

145, 105, 260, 330, 250, 195, 375, 480, 150, 180, 420, 750. Draw a box plot

sorted: 105, 145, 150, 180, 195, 250, 260, 330, 375, 420, 480, 750

min = 105, max = 750, $Q_1 = 165$, $Q_2 = 255$, $Q_3 = 397.5$





upper fence / upper limit: $Q_3 + 1.5(Q_3 - Q_1) = 397.5 + 1.5(397.5 - 165) = 746.25$

lower fence / lower limit: $Q_1 - 1.5(Q_3 - Q_1) = 165 - 1.5(397.5 - 165) = -183.75$

interquartile range

IQ Range: dispersion of middle 50% of data

$Q_3 - Q_1$

Semi Interquartile Range: dispersion of middle 25% of data

$$\frac{Q_3 - Q_1}{2}$$

Measures of dispertion:

⇒ two different data can have same mean, median and mode so central tendency approach fails there.

Date:

variance:

⇒ affected by outliers

⇒ square of standard deviation

⇒ whenever we have to find measure of dispertion and method is not mentioned, find variance

⇒ mean of squares of deviations of all the observations from their mean

⇒ variance = $(\text{unit})^2$

standard deviation = units

unqrouped:

sample variance ← $S^2 = \frac{\Sigma(x-\bar{x})^2}{n} = \frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2$

standard deviation ← $S = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}} = \sqrt{\frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2}$

Example:

45, 32, 37, 46, 39, 36, 41, 48, 36

$\bar{x} = \frac{\Sigma x}{n} = \frac{360}{9} = 40$

x	$x-\bar{x}$	$(x-\bar{x})^2$	x^2
45	+5	25	2025
32	−8	64	1024
37	−3	9	1369
46	+6	36	2116
39	−1	1	1521
36	−4	16	1296
41	+1	1	1681
48	+8	64	2304
36	−4	16	1296
	iska sum 0 ayega	232	14632

variance = $\frac{232}{9}$

= 25.78

standard deviation = 5.08

grouped:

$$s^2 = \frac{\sum f(x-\bar{x})^2}{\sum f} = \frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2$$

$$s = \sqrt{\frac{\sum f(x-\bar{x})^2}{\sum f}} = \sqrt{\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2}$$

weights	f	x	fx	fx^2 ($fx \times x$)
65-84	9	74.5	670.5	49952.25
85-104	10	94.5	94.5	89302.5
105-124	17	114.5	1946.5	222874.25
125-144	10	134.5	1345	180902.5
145-164	5	154.5	772.5	119351.25
165-184	4	174.5	698	121801
185-204	5	194.5	972.5	189151.25
			7350	973335

$$s^2 = \frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2$$

$= 1216$

$s = 34.9$

population variance:

$$\sigma^2 \text{ (sigma)} = \frac{\sum (X-\mu)^2}{N}$$ $\mu \rightarrow$ population mean

$$\sigma = \sqrt{\frac{\sum (X-\mu)^2}{N}}$$

⇒ variance and standard deviation are independent of origin (constant addition or subtraction from data) and scale (multiplication by a value)

⇒ $\frac{\sum (x-\bar{x})}{n-1}$ → unbiased this formula is precise and also used

$\frac{\sum (x-\bar{x})}{n}$ → biased we will mostly use this formula

Date:

A home theatre in a box is the easiest and cheapest way to provide surround sound for a home entertainment centre. A sample of prices is shown here. The prices are for the models with a DVD player and for models without a DVD player.

Models with DVD players	Prices in dollars (x)	$\bar{x}=410$ $(x-\bar{x})^2$	Models without DVD players	Prices in dollars (x)	$\bar{x}=310$ $(x-\bar{x})^2$
Sony HT-1800DP	450	1600	Sony HI DDW 750	300	100
Sony HT-C800DP	300	12100	Kenwood HTB-306	300	100
Panasonic SC-HT1900	400	100	RCA RT-2600	360	2500
Panasonic SC-MTI	500	8100	Kenwood HTB-206	290	400
LG HTD-330DV	400	100	Panasonic SC-HT110	300	100

$S = 66.3$ $S = 25.3$

6-2-23 Lecture #5 Monday

Graphical Representation

Graphs | Diagrams

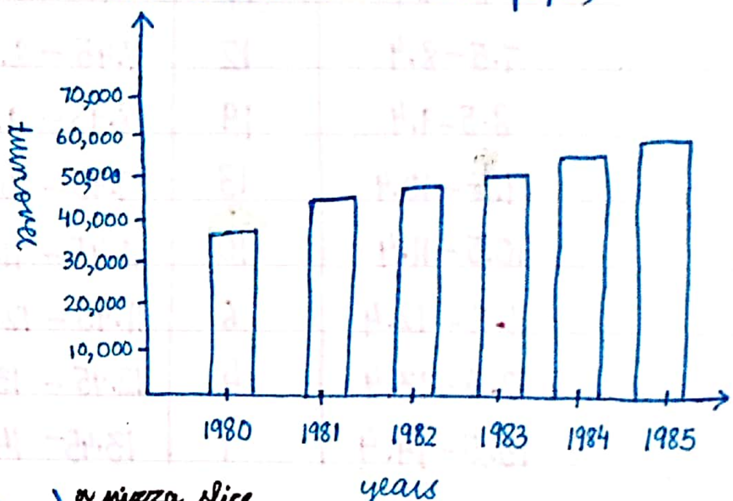
* for frequency distribution data, take category on x-axis and frequency on y-axis (for Bar chart)

- **Bar Chart**:

Draw a simple bar diagram to represent the turnover of a company for 6 years

Years	Turnover (rupees)
1980	38000
1981	45000
1982	48000
1983	52500
1984	55000
1985	58000

Bar chart: (contains gaps)



- **Pie diagram**: (or sector diagram) or pizza slice diagram

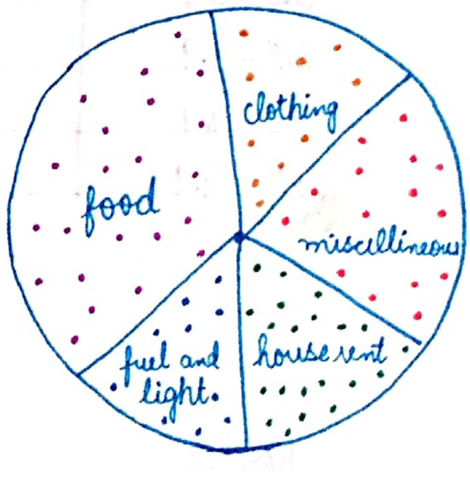
$$\text{Angle} = \frac{\text{component part}}{\text{total geometry}} \times 360^\circ$$

(whole quantity)

Represent the total expenditure and expenditures on various items of a family by a pie diagram

items	expenditure (in 1000	angles of the sectors
food	50	120°
clothing	30	72°
house rent	20	48°
fuel and light	15	36°
miscillineous	35	84°
	total = 150	

* differenciate by colours or patterns

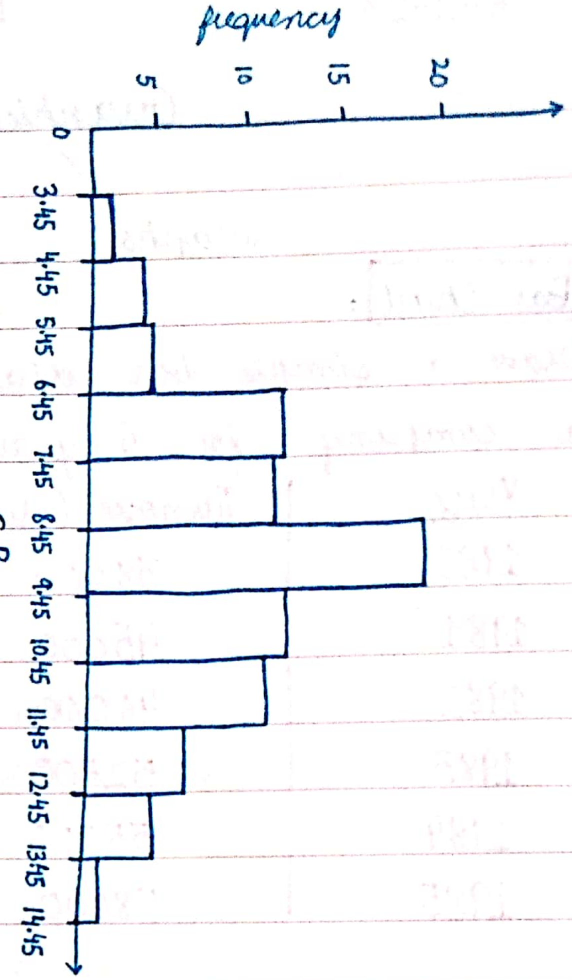


Date:

- **Histogram** : ⇒ for quantitative data

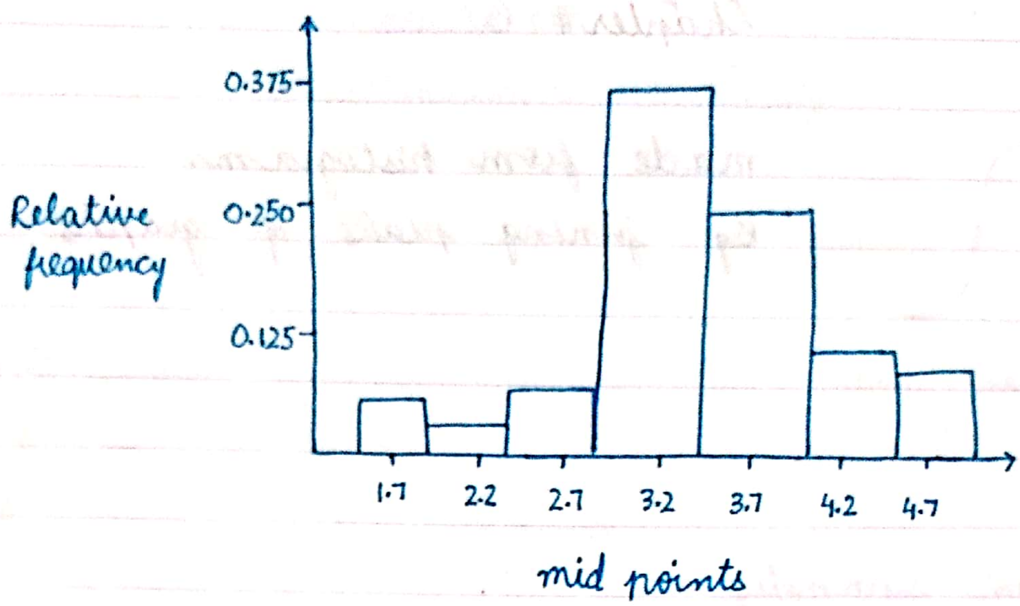
⇒ no spaces

Death rates	f	C.B
3.5 – 4.4	1	3.45 – 4.45
4.5 – 5.4	4	4.45 – 5.45
5.5 – 6.4	5	5.45 – 6.45
6.5 – 7.4	13	6.45 – 7.45
7.5 – 8.4	12	7.45 – 8.45
8.5 – 9.4	19	8.45 – 9.45
9.5 – 10.4	13	9.45 – 10.45
10.5 – 11.4	10	10.45 – 11.45
11.5 – 12.4	6	11.45 – 12.45
12.5 – 13.4	4	12.45 – 13.45
13.5 – 14.4	1	13.45 – 14.45
	88	



- **Relative frequency distribution** : or Relative frequency histogram

Class intervals	Class midpoints	f	Relative frequency $f/\Sigma f$
1.5 – 1.9	1.7	2	0.550
2.0 – 2.4	2.2	1	0.025
2.5 – 2.9	2.7	4	0.100
3.0 – 3.4	3.2	15	0.375
3.5 – 3.9	3.7	10	0.250
4.0 – 4.4	4.2	5	0.125
4.5 – 4.9	4.7	3	0.075
		40	

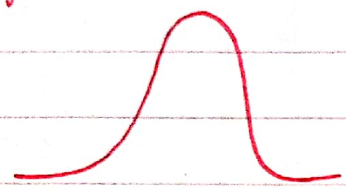

0.375
0.250
0.125
Relative frequency
1.7
2.2
2.7
3.2
3.7
4.2
4.7
mid points

Wednesday

Date: 8-2-23

Lecture # 6

Symmetric :

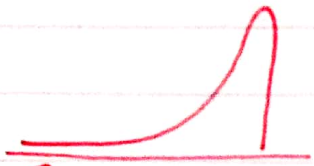


made from histogram.
By joining peaks of graphs

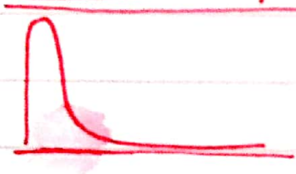
mean = median = mode

skewness :
departure from symmetry

negatively skewed :
mean < median < mode



positively skewed :
mean > median > mode



skewness using quartiles (Box and whisker) :

If $Q_3 - Q_2 = Q_2 - Q_1$ symmetrical
If $Q_3 - Q_2 > Q_2 - Q_1$ positively skewed
If $Q_3 - Q_2 < Q_2 - Q_1$ negatively skewed

13-2-23 Monday

Lecture # 7
Introduction to Probability

Set:

Collection of distinct objects e.g group of students etc.

Sample space:

A list of all possible outcomes / elements / sample spaces.

sample space ← S.S = {H, T} ↗ sample point

"Which produces different results on similar condition."

Trial:

Single performance of an experiment.

event outcome of interest S.S = {H, T}

A : Head occurs

Simple Event:

It shows only one point / event

B = {(1, 1)}

Compound Event:

It shows multiple sets

B = {(5,5), (4, 6), (6,4)}

Date:

Disjoint / Mutually exclusive event:

When 2 events do not occur at the same time

e.g: you get only 1 answer when dice is rolled

Not mutually exclusive:

When 2 events can occur at the same time.

e.g: A card can have 9 and red

Equally likely event:

50/50 chances

e.g: head / tail

Exhaustive event:

Mutually exclusive events ka union = sample space

S.S = {H, T}

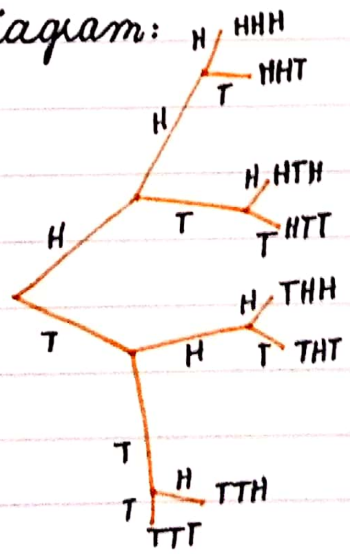
A = {H}

B = {T}

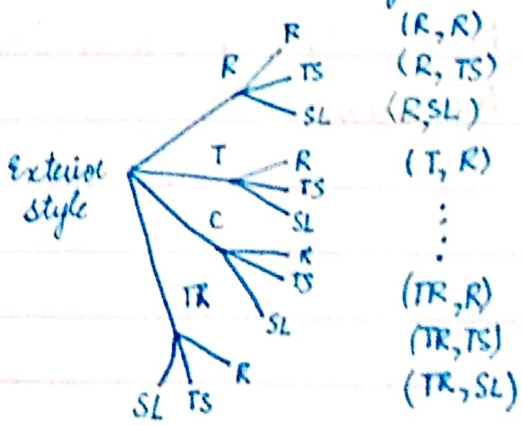
AUB = {H, T}

AUB = S.S

Tree diagram:



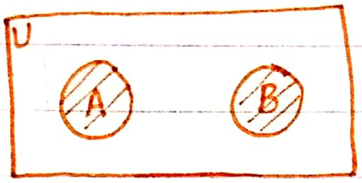
Tree diagram:



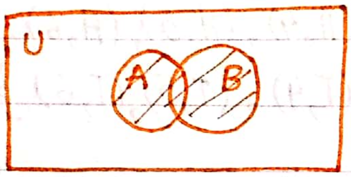
Quantitative measure of uncertainty:

$$P(A) = \frac{\text{Favourable outcomes}}{\text{total outcomes}}$$

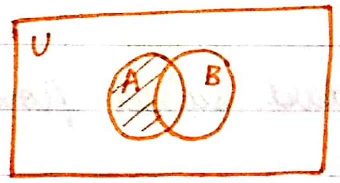
Venn diagram:



AUB



AUB



A-B

Counting sample points:

Sample techniques: 1. Multiplication 2. Permutation 3. Combination

Rule of multiplication:

- Total outcomes for 3 TA's among 35 students of a class
  $= 35 \times 34 \times 33 = 39270$ ways/combinations

- Sam is going to assemble a computer by himself. He has choice of chips from 2 brands, a harddrive from 4, memory from 3 and accessory bundle from 5 local stores. How many diff ways can Sam order the parts?
  $5 \times 4 \times 3 \times 2 = 120$ ways

- A developer offers prospective home buyers a choice of tudor, rustic, colonial and traditional exterior styling in ranch, two storey and split level floor planks. In how many way can a buyer order 1 of these homes? $4 \times 3 = 12$

Date: 15-2-23 Wednesday

Lecture # 8

Counting sample points

1) Rules of multiplication:

n_1 , n_2 $n_1 \cdot n_2$

$SS = \{H, T\}$

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

$2 \times 6 = \{(H,1), (H,2), (H,3), (H,4), (H,5), (H,6),$
$(T,1), (T,2), (T,3), (T,4), (T,5), (T,6)\}$

Permutation:

A permutation is any ordered subset from a set of n distinct objects

- A B C

3 choices for first letter
2 choices for second letter
1 choice for third letter

$3!$
for n objects
⇒ they can be arranged in $n!$ ways

- A B C D E F G

four letter word
7 choices for first letter
6 choices for second letter
5 choices for third letter
4 choices for fourth letter

$7 \times 6 \times 5 \times 4$

$^nP_r = {}^7P_4 = \frac{7!}{(7-4)!}$

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

r = jitni arrangemen

Q. A club consists of 4 members. How many sample points are in the sample space when 3 officers President, secretary and tresurer are to be chosen.

$$^{n}P_{r} = \frac{4!}{(4-3)!} = 24$$

Q. 8 people A, B C D E F G and H are arranged randomly in a line. What is the probability that A and B are a) next to each other. b) not together

a) $\underline{A\ B}$ C D E F G H $7!$

$\underline{B\ A}$ C D E F G H $7!$

$2 \times 7!$

probability $= \frac{2 \times 7!}{8} = \frac{1}{4}$

P(A and B are together) $= \frac{1}{4}$

b) P(A and B are not together) $= 1 - \frac{1}{4} = \frac{3}{4}$

Permutation for identical objects:

The number of distinct permutations of n objects of which n_1 are of one kind, n_2 are of second kind, ... n_k are of k kind

then

$$\frac{n!}{n_1!\, n_2! \ldots n_k!}$$

Date:

Q. How many different letter arrangements can be made from the letter words STATISTICS?

$$\frac{10!}{3!\,3!\,2!\,1!\,1!}$$

Combination:

A combination is any subset of r objects, selected without regard to their order, from a set of n distinct objects.

A B C D

four sets of 3 letters:

ABC	ABD	ABD	BCD	$^nC_r = \frac{^nP_r}{r!}$
ACB	⋮	⋮	⋮	
BAC				$\binom{n}{r} = \frac{n!}{(n-r)!\,r!}$
BCA				
⋮				
3!	3!	3!	3!	$r!$

Q A three persons committee is to be formed from a list of 4 persons. How many sample points are associated with experiment

$$\frac{4!}{(4-3)!\,3!} = 4 \text{ ways}$$

Q A team of 5 people which must contain 3 men and 2 women is chosen from 8 men and 7 women. How many different teams can be selected.

$$\binom{8}{3}\binom{7}{2} = 1176$$

How can you love,
if you don't love yourself
if you don't love yourself
you can love no one else

Counting techniques:

* $n!$ for n distinct objects.
* $\dfrac{n!}{n_1!\,n_2!\ldots n_k}$ for objects that are similar.
* $(n-1)!$ circular arrangement.
* $^nP_r = \dfrac{n!}{(n-r)!}$ $n=$ total, $r=$ selected for ordered selection ⇒ Permutation (one by one selection)
* $^nC_r = \dfrac{n!}{r!\,(n-r)!}$ for un-ordered selection ⇒ Combination (once at start selection)

Q. In a college football training session, the defensive coordinator needs to have 10 players standing in a row. Among these 10 players there is 1 freshman, 2 sophomores, 4 juniors and 3 seniors. How many different ways can they be arranged in row so that their class level can be distinguished.

$$\frac{n!}{n_1!\,n_2!\ldots n_k!}$$

Q. In how many ways can a carivan of 8 covered wagons from Arizona be arranged in a circle.

$(8-1)! = 7! = 5040$

Date:

Q
a) How many distint permutations can be made from the letters of the word COLUMNS.
b) How many of these permutations start with the letter M.

a) $7! = 5040$

b) $(n-1)! = 6! = 720$

Q. In one year 3 awards (research, teaching and service) will be given to a class of 25 graduates in management department. If each student can receive atmost 1 reward, how many selections are possible.

$$^{n}P_{r} = {}^{25}P_{3} = \frac{25!}{(25-3)!} = 13800$$

H.W Q 2.32 , Q2.36 , Q2.37 , Q2.46 , Q2.39

20-2-23 Monday

Probability of an event

Lecture # 9

Addition rule of probability: * jab OR ayega

1. $P(A \text{ or } B) = P(A) + P(B)$ ⇒ for mutually exclusive events
2. $P(A \text{ or } B) = P(A) + P(B) - P(A \cup B)$ ⇒ for not mutually exclusive events

Q. 2 fair dice are thrown. A prize is won if the total score on the two rolls is 4 OR if each individual score is over 4.

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

Prize is won if

A = total score on the two rolls is 4

B = each individual score is over 4

$A = \{(1,3), (2,2), (3,1)\}$ $P(A) = 3/36$

$\{(5,5), (5,6), (6,5), (6,6)\}$ $P(B) = 4/36$

$P(A \cup B) = P(A) + P(B)$

$= 3/36 + 4/36 = 7/36$

Q. A prize is won if the total score on 2 rolls is 10 OR if each indivial score is over 4. Probability that prize is won?

$A = \{(5,5), (5,6), (6,5), (6,6)\}$ $P(A) = 4/36$

$B = \{(4,6), (5,5), (6,4)\}$ $P(B) = 3/36$

$A \cap B = \{(5,5)\}$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$= 4/36 + 3/36 - 1/36 = 6/36 = 1/6$

Date:

Q. In a group of 20 adults, 4 out of 7 women, and 2 out of 13 men wear glasses. What's the probability that a person selected at random from the group is a woman or someone who wear glasses.

G = person who wears glasses
G' = person who does not wear glasses
M = person who wears glasses is a man
W = person who wears glasses is a woman

	M	W	
G	2	4	6
G'	11	3	14
	13	7	20

$P(W \cup G) = ?$

$P(W) = 7/20 \qquad P(G) = 6/20$

$P(W \cap G) = 4/20$

$P(W \cup G) = P(W) + P(G) - P(W \cap G)$

$= 7/20 + 6/20 - 4/20$

$= 9/20$

Q. A box contains 50 bolts and 150 nuts. Half of the bolts and half of the nuts are rusted. If one item is selected at random, what is the probability that it is rusted or a bolt?

R = item is rusted

$P(R) = 100/200 = 1/2$

B = item is bolt

$P(B) = 50/200 = 1/4$

A = rusted and bolt

$P(A) = 25/200 = 1/8$

$P(R \cup B) = 1/2 + 1/4 - 1/8$

$= 5/8$

Conditional Probability :

* some condition is given and sample space is reduced.

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$ B ⇒ event which has already occured

Q. Consider a class of 30 students of whom 17 are girls and 13 are boys. Suppose further that 5 of the girls and 6 of the boys are L-handed. If a student selected at random is a girl, what's the probability that person is left handed.

L = left handed $\qquad P(L) = 11/30 \qquad P(L \cap G) = 5/30$

G = girl $\qquad P(G) = 17/30$

$$P(L/G) = \frac{P(L \cap G)}{P(G)}$$

$$= \frac{5/30}{17/30} = 5/17$$

Probability of an Event

* Axiomatic approach

Q. 2 coins are tossed. what is the probability that at least 1 head occurs?

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

$$S = \{w, w, w, w\} \qquad 4w = 1, \quad w = \frac{1}{4}$$

$$S = \left\{\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right\}$$

Let A represents the occurence of atleast one head

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

$$P(A) = \left\{\frac{1}{4} + \frac{1}{4} + \frac{1}{4}\right\}$$

$$= \frac{3}{4}$$

Date:

Q A dice is rolled in such a way that an even number is twice as likely to occur as an odd number. If E is an event that a number <4 occurs on a single roll of dice, then find probability of E.

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

$$= \{w, w, w, w, w, w\}$$

$$= \{w, 2w, w, 2w, w, 2w\} \qquad 9w = 1, \quad w = \frac{1}{9}$$

$$= \left\{\frac{1}{9}, \frac{2}{9}, \frac{1}{9}, \frac{2}{9}, \frac{1}{9}, \frac{2}{9}\right\}$$

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

$$= \frac{1}{9} + \frac{2}{9} + \frac{1}{9}$$

$$P(A) = \frac{4}{9}$$

* classical approach

$$P(A) = \frac{n}{N}$$

$n \rightarrow$ favourable outcome

$N \rightarrow$ total outcome that are equally likely

Q. A statistics class for enigineers consist of 25 industrial, 10 mechanical, 10 electrical and 8 civil enigneering students. If the person is randomly selected to ans a question, find probability that selected student is:

a) industrial engg. major $\quad 25/53$

b) Electrical or civil engg major $\quad 18/53$

EASE

Q. In a poker hand consisting of 5 cards, find the probability of holding 2 aces and 3 jacks.

Total possible ways $= {}^{52}C_5 = 2598960$

Total favourable ways $= {}^4C_2 \times {}^4C_3 = 24$

Let A represents the the event of occurence of 2 Aces and 3 jacks

$$P(A) = \frac{24}{2598960} = 0.9 \times 10^{-5}$$

Q. An employer wish to hire 3 ppl from a group of 15 applicants, 8 men and 7 women all of whom are equally qualified to fill the position. If he selects the 3 participants at random, what's the probability of getting

a) all 3 men

b) at least one will women

a) $P(A) = \frac{n(A)}{n(S)}$

$= \frac{{}^8C_3}{{}^{15}C_3}$

$= \frac{8}{65}$

b) $P(B) = \frac{n(B)}{n(S)}$

$= \frac{{}^7C_1\,{}^8C_2 + {}^7C_2\,{}^8C_1}{{}^{15}C_3}$

$= 0.87$

22-2-23

Lecture # 10 Wednesday

Q. A man tosses 2 fair dice what's the conditional probability that the sum on 2 dice will be 7, given the condition that

i) Sum is odd

ii) Sum is greater than 6

iii) Two dice shows the same outcome

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

A : sum is 7

$A = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$ $P(A) = 6/36$

B : sum is odd

$B = \{(1,2),(1,4),(1,6),(2,1),(2,3),\ldots,(6,5)\}$ $P(A) = 18/36$

C : sum is greater than 6

$C = \{(1,6),(2,5),(2,6),(3,4),\ldots,(6,6)\}$ $P(C) = 21/36$

D : Two dice shows the same number

$D = \{(1,1),(2,2),(3,3),(4,4),(5,5),(6,6)\}$ $P(D) = 6/36$

i) $A \cap B = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$

$P(A \cap B) = 6/36$

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{6/36}{18/36} = \frac{1}{3}$$

ii) $$P(A/C) = \frac{P(A \cap C)}{P(C)} = \frac{6/36}{21/36} = \frac{2}{7}$$

iii) $$P(A/D) = \frac{P(A \cap D)}{P(D)} = \frac{0}{6/36} = 0$$

Date:

Permutation Practise Qs.

Q. Find the no. of ways of arranging 6 women and 3 men in a row, so that no 2 men are standing next to one another.

$\downarrow W_1 \downarrow W_2 \downarrow W_3 \downarrow W_4 \downarrow W_5 \downarrow W_6 \downarrow$

$6! \times 7 \times 6 \times 5$

Additive law of Probability

Q. What is the probability of getting a total of 7 or 11 when a pair of dice is rolled?

$A = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$

$P(A) = \frac{6}{36}$

$B = \{(5,6), 6,5\}$

$P(B) = \frac{2}{36}$

$P(A \cup B) = P(A) + P(B)$

$= \frac{6}{36} + \frac{2}{36} = \frac{8}{36}$

Q. A card is drawn at random from a deck of cards. What's the probability that it is diamond, face or king.

$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$

A : card is diamond

B : card is face

C : card is king

$P(A \cup B \cup C) = \frac{13}{52} + \frac{12}{52} + \frac{4}{52} - \frac{3}{52} - \frac{1}{52} - \frac{4}{52} + \frac{1}{52}$

$= \frac{22}{52}$

Date:

Law of complementation

$P(A) + P(A)' = 1$

Q. If the probality is that a mechanic will service 3, 4, 5, 6, 7 or 8 or more cars on any working day are respectively 0.12, 0.19, 0.28, 0.24, 0.10 & 0.07. What's the probability that he will service at least 5 cars on his next working day?

Probability $= 1 - (0.12 + 0.19)$
$= 0.69$

Conditional Probability

Probability of A given that B $\quad P(A/B) = \dfrac{P(A \cap B)}{P(B)}$

if $P(A/B) = P(A)$ then events are independent

Q. The probability that a regularly scheduled flight departs on time is $P(D) = 0.83$. The probability that it arrives on time is $P(A) = 0.82$ and the probability that it departs and arrives on time is $P(A \cap B) = 0.78$. Find the probability that a plane

i) arrives on time given that it departed on time
ii) departed on time given that it has arrived on time

i) $P(A/D) = \dfrac{P(A \cap D)}{P(D)} = 0.94$ $\qquad P(A/D') = \dfrac{P(A \cap D')}{P(D')}$

ii) $P(D/A) = \dfrac{P(D \cap A)}{P(A)} = 0.95$ $\qquad P(D') = 1 - P(D)$

$P(A \cap D') = P(A) - P(A \cap D)$

EASE

1-3-23 Wednesday

Lecture # 11

Multiplication rule for dependent events

If eik event se baki events affect ho rahay hon.

$$P(B/A) = \frac{P(A \cap B)}{P(A)}$$

$$\boxed{P(A \cap B) = P(A).P(B/A)}$$

A and B

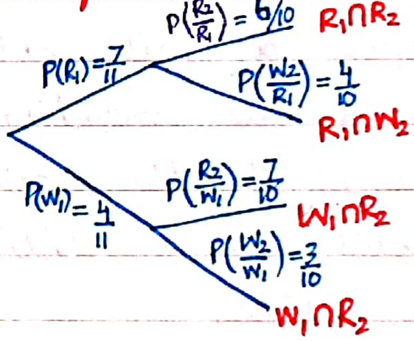
Q. Suppose a jar contains 7 red balls and 4 white balls. 2 balls are sellcted without replacement, what is the probability that i) both are red ii) first white and second red iii) both are of same colour?

R_1 = First red ball

R_2 = Second red ball

W_1 = First white ball

W_2 = Second white ball



i) $P(R_1 \cap R_2) = P(R_1).P(R_2/R_1)$

$$= \frac{7}{11} \cdot \frac{6}{10}$$

$$= \frac{42}{110}$$

ii) $P(W_1 \cap R_2) = P(W_1).P(R_2/W_1)$

$$= \frac{4}{11} \cdot \frac{7}{10} = \frac{28}{110}$$

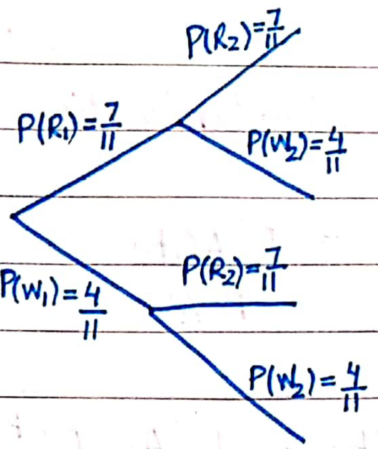
iii) $P(R_1 \cap R_2)$ or $P(W_1 \cap W_2) = P(R_1).P\left(\frac{R_2}{R_1}\right) + P(W_1).P\left(\frac{W_2}{W_1}\right)$

$$= \frac{7}{11} \cdot \frac{6}{10} + \frac{4}{11} \cdot \frac{3}{10}$$

$$= \frac{54}{110}$$

Date:

Multiplication rule for independent events:



$$\boxed{P(A \cap B) = P(A).P(B)}$$

Probability that second ball is red:

$$P(R_2) = P(W_1 \cap R_2) \text{ OR } P(R_1 \cap R_2)$$
$$= P(W_1)P(R_2) + P(R_1).P(R_2)$$
$$= \frac{4}{11}.\frac{7}{11} + \frac{7}{11}.\frac{7}{11} = \frac{7}{11}$$

Q. In a carnival game a contestent has to first spin a fair coin and then roll a dice whose faces are mentioned 1 to 6. The contestent wins a prize if the coin shows head and the dice shows score below 3. Find probability that contestent wins a prize.

$$P(A \cap B) = P(A).P(B)$$

A: coin shows head

B: dice shows below 3

$$P(A \cap B) = \frac{1}{2}.\frac{2}{6} = \frac{1}{6}$$

Bayes theorem OR Bayes rule

⇒ conditional probability ko fuse ke kay further conditional probability bnayeen gy

$$\boxed{P(A/B) = \frac{P(A \cap B)}{P(B)}}$$

sum of probabilities

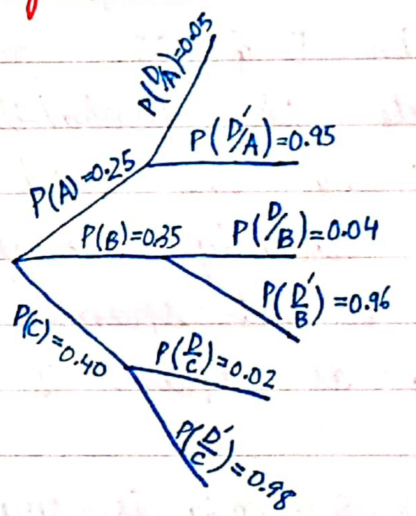
Q. In a bolt factory machines A, B and C manufactures 25%, 35% and 40% of total outcomes respectively. Of their outputs 5%, 4% and 2% respectively are defective. A bolt is selected at random and found to be defective. What's the probability that it came from C.

D: bolt is defective

$P(D/A) = 0.05$

$P(D/B) = 0.04$

$P(D/C) = 0.02$



$$P(C/D) = \frac{P(C \cap D)}{P(D)}$$

$$P(D) = P(D \cap A) + P(D \cap B) + P(D \cap C)$$

$$= P(A).P\left(\frac{D}{A}\right) + P(B).P\left(\frac{D}{B}\right) + P(C).P\left(\frac{D}{C}\right)$$

$$= 0.25 \times 0.05 + 0.35 \times 0.04 + 0.40 \times 0.02$$

$$P\left(\frac{C}{D}\right) = \frac{P(C \cap D)}{P(D)}$$

$$= \frac{P(C)\ P(D/C)}{P(D \cap A) + P(D \cap B) + P(D \cap C)}$$

$$= \frac{0.40 \times 0.02}{0.03}$$

$$= 0.26$$

Date:

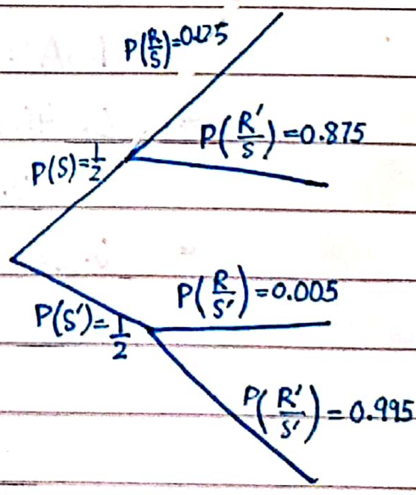
Application of Bays theorem:
Baysian span filter:

Q. Suppose that we have found that the word 'Rolex' occurs in 250 of 2000 msgs known to be spam and in 5 of 1000 msgs known not to be spam. Estimate the probability that an incoming msg containing the word 'Rolex' is spam assuming that it is equally likely that an incoming msg is spam or not spam. If our threshold for rejecting a msg as spam is 0.9, will we reject such msgs

S = msg is spam
S' = msg is not spam
R = rolex word occurs
R' = rolex word does not occur

$$P(S/R) = \frac{P(S \cap R)}{P(R)}$$

$$P(R) = P(R \cap S) + P(R \cap S')$$
$$= P(S).P\left(\frac{R}{S}\right) + P(S').P\left(\frac{R}{S'}\right)$$



$$= \frac{1}{2} . 0.125 + \frac{1}{2} \times 0.005$$

$$= 0.065$$

$$P(S/R) = \frac{P(S).P(R/S)}{P(S \cap R) + P(S' \cap R)}$$

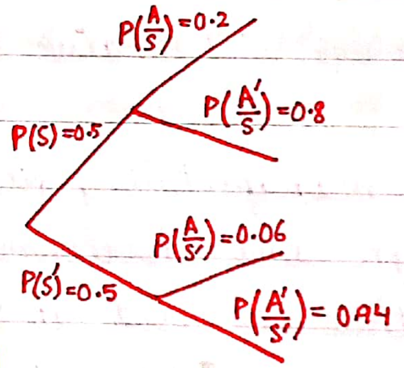
$$= \frac{1/2 \times 0.125}{0.065} = 0.962$$ we reject such messages as spam

6-3-23 Monday

Lecture # 12

Example:

Suppose that we train a Bayesian spam filter on a set of 2000 spam messages and 1000 messages that are not spam. The word "stock" appears in 400 spam messages and 60 messages that are not spam. And the word "undervalued" appears in 200 spam messages and 25 messages that are not spam. Estimate the probability that an incoming message containing both the words "stock" and "undervalued" is spam, assuming that we have no prior knowledge about whether it is spam. Will we reject such messages as spam when we set the threshold at 0.9?



A = word is stock

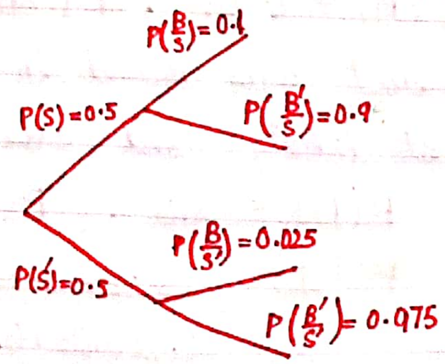
B = word is undervalued

we have to find

$$P\left(\frac{S}{A\cap B}\right) = \frac{P(S).P\left(\frac{A}{S}\right).P\left(\frac{B}{S}\right)}{P(S).P\left(\frac{A}{S}\right).P\left(\frac{B}{S}\right)+P(S')P\left(\frac{B}{S'}\right)P\left(\frac{B}{S'}\right)}$$

$$= \frac{0.5\times0.2\times0.1}{(0.5\times0.2\times0.1)+(0.5\times0.06\times0.025)}$$

$$= 0.93 > 0.9$$



we reject such messages as spam.

Date:

Q. Two cards are drawn in succession without replacement from an ordinary deck of cards. Find probability that an event $A_1 \cap A_2 \cap A_3$ occurs where A_1 = first card is ace red, A_2 = second card is 10 or Jack, A_3 = third card is >3 but <7.

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)\, P(A_2/A_1)\, P(A_3/A_1 \cap A_2)$$

$$= \frac{2}{52} \times \frac{8}{51} \times \frac{12}{50} = \frac{8}{5525}$$

Q. The prob that a man will be alive in 25 year is $3/5$ and the probability that his wife will be alive in 25 years is $2/3$. Find the probability that:

i) Both will be alive $P(A \cap B) = P(A)\,P(B) = \frac{3}{5} \times \frac{2}{3} = \frac{2}{5}$

ii) Only man will be alive $P(A \cap \bar{B}) = \frac{3}{5} \times (1 - \frac{2}{3}) = \frac{1}{5}$

iii) Only wife will be alive $P(\bar{A} \cap B) = (1 - \frac{3}{5})(\frac{2}{3}) = \frac{4}{15}$

iv) Atleast 1 will be alive $P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{3}{5} + \frac{2}{3} - \frac{2}{5} = \frac{13}{15}$

v) No one will be alive $P(\bar{A} \cup \bar{B}) = 1 - P(A \cup B) = 1 - \frac{13}{15} = \frac{2}{15}$

Q. A manufacturing firm employs three analytical plans for the design and development of a particular product. For cost reasons, all three are used at varrying times. In fact, plan 1, 2 and 3 are used for 30%, 20% and 50% of the products respectively. The defect rate is different for the 3 plans and are as follows:

$P(D/P_1) = 0.01$, $P(D/P_2) = 0.03$, $P(D/P_3) = 0.02$. If a random product was found to be defected, which one of the plans is most likely used and responsible?

$P(P_1) = 3/10$, $P(P_2) = 2/10$, $P(P_3) = 5/10$

$$P(P_1/D) = \frac{0.3 \times 0.01}{0.2 \times 0.03 + 0.3 \times 0.01 + 0.5 \times 0.02} = 0.158$$

$$P(P_2/D) = \frac{0.3 \times 0.01}{0.3 \times 0.01 + 0.3 \times 0.01 + 0.5 \times 0.02} = 0.316$$

$$P(P_3/D) = \frac{0.5 \times 0.02}{0.3 \times 0.01 + 0.3 \times 0.01 + 0.5 \times 0.02} = 0.526$$

Practise: 2.81, 2.83, 2.89, 2.75, 2.76, 2.94

EASE

Date: 8-3-23

Lecture # 13
Chapter # 3

Wednesday

Random variable

$SS = \{HH, HT, TH, TT\}$
$\quad\quad\;\; 2 \quad\; 1 \quad\; 1 \quad\; 0$

X: no of heads
$X: 0, 1, 2$

X	$f(x)$
0	$1/4$
1	$2/4$
2	$1/4$
	1

$f(x) \geq 0$
$\Sigma f(x) = 1$
$P(X=x) = f(x)$

Eg 3.8 A shipment of 20 similar lappies to a retail outlet contains 3 that are defective. If a school makes a random purchase of 2 of these lappies, find the probability distribution for the number of defectives.

Total	Good	Defective
20	17	3

X: No of defective
$X: 0, 1, 2$

$$P(X=0) = \frac{\binom{3}{0}\binom{17}{2}}{\binom{20}{2}}$$

Probability function ← $P(X=x) = \frac{\binom{3}{x}\binom{17}{2-x}}{\binom{20}{2}} = f(x)$

X	$f(x)$ or $P(x)$	
0	$68/95$	$P(X=0)$
1	$51/190$	$P(X=1)$
2	$3/190$	$P(X=2)$
	1	

$f(x) \geq 0$
$\Sigma f(x) = 1$
$P(X=x) = f(x)$

Date:

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

X : No of tails

X : 0, 1, 2, 3

X	$f(x)$
0	$1/8$
1	$3/8$
2	$3/8$
3	$1/8$
	1

$$\frac{\binom{3}{x}}{8}$$

Discrete probability function

OR

Probability mass function

$f(x) = P(X = x)$
(exact wohi probability)

Commulative distribution function

OR

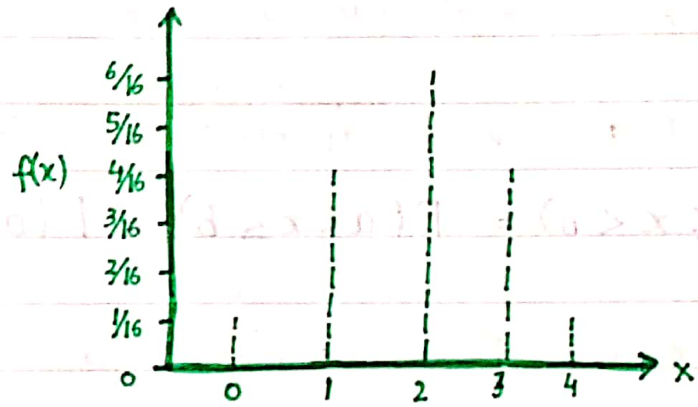
Distribution function $\Rightarrow$ $F(x) = P(X \leq x)$
(whan tk ki probability ka sum)

$$f(x) = \frac{1}{16}\binom{4}{x} \quad \text{for } x = 0, 1, 2, 3, 4$$

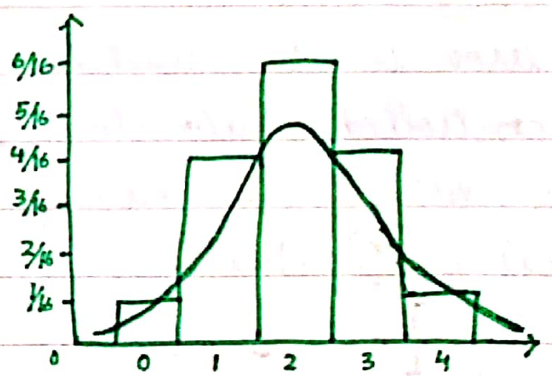
x	$f(x) = 1/16 \binom{4}{x}$	$F(x)$
0	$1/16$	$1/16$
1	$4/16$	$1/16 + 4/16 = 5/16$
2	$6/16$	$5/16 + 6/16 = 11/16$
3	$4/16$	$11/16 + 4/16 = 15/16$
4	$1/16$	$15/16 + 1/16 = 1$

$$f(2) = F(2) - F(1)$$
$$= \frac{11}{16} - \frac{5}{16}$$
$$= \frac{6}{16}$$

$$F(x) = \begin{cases} 0 & \text{for} \quad x<0 \\ 1/16 & \text{for} \quad 0 \le x < 1 \\ 5/16 & \text{for} \quad 1 \le x < 2 \\ 11/16 & \text{for} \quad 2 \le x < 3 \\ 15/16 & \text{for} \quad 3 \le x < 4 \\ 1 & \text{for} \quad x \ge 4 \end{cases}$$



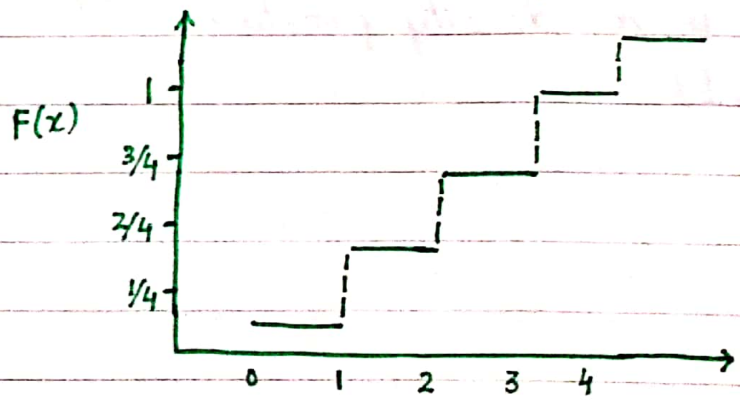
Probability mass function



Probability histogram

Practise Qs:

3.1, 3.2, 3.22, 3.10



Discrete commulative distribution function

Date: 13-3-23 Monday

Lecture # 14

Density probability function:

The function $f(x)$ is a probability density function for continuous random variable x if

1. $f(x) \geq 0$
2. $\int_{-\infty}^{\infty} f(x)\, dx = 1$
3. • $P(a \leq x \leq b) = \int_a^b f(x)dx$ * jab pdf (probability density function) kay through probability nikalein
   • $P(a \leq x \leq b) = F(b) - F(a)$ * jab cdf (commulative distributive function) kay through probability nikalein

* $P(a < x < b) = P(a \leq x < b) = P(a < x \leq b) = P(a \leq x \leq b)$

Pg 89 Exp 3.11

Q. Suppose that the error in the reaction temperature in Centigrade for a controlled laboratory experiment a continuous random variable X having the probability density function $f(x) = \begin{cases} \frac{x^2}{3} & -1<x<2 \\ 0 & \text{else where} \end{cases}$

a) verify that $f(x)$ is a density function * integral should be = 1

b) find $p(0 \leq x \leq 1)$

a) $\int_{-\infty}^{\infty} f(x)\, dx = 1$

$= \int_{-1}^{2} \frac{x^2}{3}\, dx = 1$

$= \frac{1}{3} \left| \frac{x^3}{3} \right|_{-1}^{2}$

$$= \frac{1}{3}\left[\frac{2^3}{3} + \frac{1}{3}\right]$$

$$= \frac{9}{9} = 1$$

b) $\int_0^1 \frac{x^2}{3} dx$

$$= \frac{1}{9}\left| x^3 \right|_0^1$$

$$= \frac{1}{9}[1-0]$$

$$= \frac{1}{9}$$

Commulative distributive function :

$$F(x) = P(x \leq x) = \int_{-\infty}^{x} f(x)\, dx$$ lower limit to x

* For previous question

$$= \int_{-1}^{x} \frac{x^2}{3} dx = \frac{1}{3}\left[\frac{x^3}{3}\right]_{-1}^{x}$$

$$= \left|\frac{x^3}{9}\right|_{-1}^{x}$$

$$= \frac{1}{9}[x^3+1]$$

$$F(x) = \frac{x^3+1}{9}$$

$P(0 \leq x \leq 1)$ using $F(x)$:

$$P(a \leq x \leq b) = F(b) - F(a)$$

$$P(0 \leq x \leq 1) = \frac{1+1}{9} - \frac{1}{9}$$

$$= \frac{1}{9}$$

Date:

Question 3.30 :

$$f(x) = \begin{cases} k(3-x^2) & -1<x<1 \\ 0 & \text{elsewhere} \end{cases}$$

a) determine k

b) find F(x)

c) find $p(x<\frac{1}{2})$ using F(x)

a) $\int_{-1}^{1} k(3-x^2)\, dx = 1$

$k\int_{-1}^{1} 3\, dx - k\int_{-1}^{1} x^2\, dx = 1$

$3k\,|x|_{-1}^{1} - k\left|\frac{x^3}{3}\right|_{-1}^{1} = 1$

$6k - \frac{2}{3}k = 1$

$\frac{16}{3}k = 1$

$k = \frac{3}{16}$

c) using p.d.f

$p(x<\frac{1}{2})$

$\int_{-1}^{1/2} \frac{3}{16}(3-x^2)dx$

b) $F(x) = \int_{-1}^{x} \frac{3}{16}(3-x^2)\, dx$

$= \frac{3}{16}\int_{-1}^{x}(3-x^2)\, dx$

$$= \frac{3}{16}\left[3x - \frac{x^3}{3}\right]_{-1}^{x}$$

$$= \frac{3}{16}\left[\frac{9x-x^3}{3} - \left(-3 - \left(\frac{-1}{3}\right)\right)\right]$$

$$= \frac{1}{16}(-x^3+9x+8)$$

$$= \frac{-x^3}{16} + \frac{9x}{16} + \frac{1}{2}$$

Q 3.36 On a laboratory assignment if the equipment is working, the density function of the observed outcome X is $f(x) = 2(1-x) \quad 0<x<1$

$0 \quad$ otherwise

a) $P(x \leq 1/3)$

b) $P(x > 0.5)$

c) Find distribution function and use it to evaluate part 3

a) $P(x \leq 0.33) = \int_{-\infty}^{+\infty} fx\,dx$

$$= \int_{0}^{0.33} fx\,dx$$

$$= \int_{0}^{0.33} 2(1-x)\,dx$$

$$= 2\left[x - \frac{x^2}{2}\right]_{0}^{0.33}$$

$$= 2\left[0.33 - \frac{(0.33)^2}{2}\right]$$

$$= 0.55 = \frac{5}{9}$$

b) $P(x>0.5) = \int_{-\infty}^{+\infty} fx\,dx$

$$= \int_{0.5}^{1} 2(1-x)\,dx$$

$$= \frac{1}{4} = 0.25$$

Date:

Probability using cdf

- $P(a<x<b) = F(b)-F(a)$
- $P(x>b) = 1-F(b)$
- $P(x\leq b) = F(b)$

c) $F(x) = \int_{-\infty}^{\infty} f(x)\,dx$

$= \int_{0}^{x} 2(1-x)\,dx$

$= 2\left[x - \frac{x^2}{2}\right]_0^x$

$= \cancel{2}\left(\frac{2x-x^2}{\cancel{2}}\right)$

$= 2x - x^2$

$F(0.5) = 2(0.5)-(0.5)^2 = 0.75$

$1-F(0.5) = 1-0.75 = 0.25$

Q. Consider the density function $f(x) = k\sqrt{x}$ $0<x<1$

0 otherwise

a) Evaluate 'k'

b) Find $F(x)$ & use it to evaluate $P(0.3<x<0.6)$

a)

$f(x) = k\sqrt{x}$

$k\int_0^1 \sqrt{x}\,dx = 1 = \frac{2k}{3}\left[x^{3/2}\right]_0^1$

$= \frac{2k}{3} = 1$

$\Rightarrow k = \frac{3}{2}$

b)

For $0\leq x<1$, $F(x) = \frac{3}{2}\int_0^x \sqrt{t}\,dt$

$= \left[t^{3/2}\right]_0^x$

$= x^{3/2}$

Now,

$$F(x) = \begin{cases} 0 & x<0 \\ x^{3/2} & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$$

$$P(0.3 < X < 0.6) = F(0.6) - F(0.3)$$
$$= (0.6)^{3/2} - (0.3)^{3/2}$$
$$= 0.3004$$

Q. An investment firm offers its customers muncipal bonds that mature after varying no. of offers given that the cdf of T, the no. of years of maturity for a randomly selected bond $F(t) = \begin{cases} 0 & t<1 \\ 1/4 & 1 \le t < 3 \\ 1/2 & 3 \le t < 5 \\ 3/4 & 5 \le t < 7 \\ 1 & t \ge 7 \end{cases}$

Find

a) $P(T=5)$ * hint $F(5) - F(4)$

b) $P(T>3)$ * hint $1 - F(3)$

c) $P(1.4 < T < 6)$

d) $P(T \le 5 / T \ge 2)$

a) $P(T=5) = F(5) - F(4)$
$$= \frac{3}{4} - \frac{1}{2}$$
$$= \frac{1}{4}$$

b) $P(T>3) = 1 - F(3)$
$$= 1 - \frac{1}{2} = \frac{1}{2}$$

Date:

c) $P(1.4 < T < 6) = F(6) - F(1.4)$

$$= \frac{3}{4} - \frac{1}{4}$$

$$= \frac{1}{2}$$

d) $\frac{P(2 \leq T \leq 5)}{P(T \geq 2)} = \frac{F(5) - F(2)}{1 - F(2)}$

$$= \frac{3/4 - 1/4}{1 - 1/4}$$

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

Q 3.27 The time of failure in hrs of an iron piece of equipment used in a manufacturer DVD player has the density function

$$f(x) = \begin{cases} \frac{1}{2000} \exp\left(\frac{-x}{2000}\right) & x \geq 0 \\ 0 & x < 0 \end{cases}$$

a) Find $F(x)$ (cdf)

b) determine the probability that the component last more than 1000 hrs before the component needs to be replaced.

c) determine the probability that the component fails before 2000 hrs.

20-3-23 Monday

P.d.f

Q. $f(x) = \begin{cases} 2x & 0<x<1 \\ 0 & \text{otherwise} \end{cases}$

$$P\left[x \le \tfrac{1}{2} \,/\, \tfrac{1}{3} \le x \le \tfrac{2}{3}\right]$$

$$P(A/B) = \frac{P(A\cap B)}{P(B)}$$

$$= \frac{P[x \le \frac{1}{2} \cap \frac{1}{3} \le x \le \frac{2}{3}]}{P(\frac{1}{3} \le x \le \frac{2}{3})}$$

$$= \frac{P(\frac{1}{3} \le x \le \frac{1}{2})}{P(\frac{1}{3} \le x \le \frac{2}{3})}$$

$$= \frac{\int_{1/3}^{1/2} 2x \, dx}{\int_{1/3}^{2/3} 2x\,dx}$$

$$= \frac{5/36}{1/3}$$

$$= \frac{5}{12}$$

Q. $F(x) = \begin{cases} 0 & \text{for } x<0 \\ \frac{2x^2}{5} & \text{for } 0<x\le 1 \\ -\frac{3}{5}+\frac{2}{5}\left(3x-\frac{x^2}{2}\right) & 1<x\le 2 \\ & x>2 \end{cases}$

Find pdf $f(x)$ and $P(|x|<1.5)$

$$f(x) = \frac{d}{dx}F(x)$$

Q.

$$f(x) = \begin{cases} 0 & \text{for } x<0 \\ \frac{4x}{5} & \text{for } 0<x\le 1 \\ \frac{2}{5}(3-x) & \text{for } 1<x\le 2 \\ 0 & \text{for } x>2 \end{cases}$$

Solution:

$$P(-1.5 \le x \le 1.5) = \int_{-1.5}^{0} 0\,dx + \int_{0}^{1} \frac{4x}{5}\,dx + \int_{1}^{1.5} \frac{2}{5}(3-x)\,dx$$

$$= \frac{4}{5}\left|\frac{x^2}{2}\right|_0^1 + \left|\frac{2}{5}\left(3x-\frac{x^2}{2}\right)\right|_1^{1.5}$$

$$= 0.75$$

Joint Probability Distribution

- $f(x, y) = P(x = x \text{ and } y = y)$ ← Joint probability function
- $F(x, y) = P(x \le x \text{ and } y \le y)$ ← Joint commulative distribution function

- $f(x_i, y_j) \ge 0$

$i = 1, 2, \ldots, m$
$j = 1, 2, \ldots, n$

- $\sum_i \sum_j f(x_i, y_j) = 1$

Joint probability distribution of x and y

$x \backslash y$	y_1	y_2	...	y_j	...	y_n	$P(x=x_i)$
x_1	$f(x_1,y_1)$	$f(x_1,y_2)$	...	$f(x_1,y_j)$	...	$f(x_1,y_n)$	$g(x_1)$
x_2	$f(x_2,y_1)$	$f(x_2,y_2)$	...	$f(x_2,y_j)$	...	$f(x_2,y_n)$	$g(x_2)$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
x_i	$f(x_i,y_1)$	$f(x_i,y_2)$	...	$f(x_i,y_j)$	...	$f(x_i,y_n)$	$g(x_i)$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
x_m	$f(x_m,y_1)$	$f(x_m,y_2)$	...	$f(x_m,y_j)$	...	$f(x_m,y_n)$	$g(x_m)$
$P(y=y_j)$	$h(y_1)$	$h(y_2)$	...	$h(y_j)$	...	$h(y_n)$	1

Marginal function of x:

$$g(x_i) = \sum_{j=1}^{n} f(x_i, y_j)$$

$$= f(x_i,y_1) + f(x_i,y_2) + \dots + f(x_i,y_n)$$

$$g(x_i) = f(x = x_i)$$

Marginal function of y:

$$h(y_j) = \sum_{i=1}^{m} f(x_i, y_j)$$

$$= f(x_1,y_j) + f(x_2,y_j) + \dots + f(x_m,y_j)$$

$$h(y_j) = f(y = y_j)$$

Date:

Exp 3.14: 2 ball pens are selected at random from a box that contains 3 blue pens, 2 red pens and 3 green pens. If X is a number of blue pens and Y is number of red pens selected, find a) joint probability function $f(x,y)$, b) $P(x+y \leq 1)$, c) marginal probability distribution of x and y, d) The conditional probability distribution of $f(x/1)$, e) verify whether x and y are independent or not.

a)

B	R	G	
3	2	3	= 8

$SS = \binom{8}{2}$

x = no of blue pens

y = no of red pens

x : 0, 1, 2

y : 0, 1, 2

$$f(x,y) = \frac{\binom{3}{x}\binom{2}{y}\binom{3}{2-x-y}}{\binom{8}{2}}$$

x \ y	0	1	2	g(x)
0	3/28	6/28	1/28	10/28
1	9/28	6/28	0	15/28
2	3/28	0	0	3/28
h(y)	15/28	12/28	1/28	

b) $P(x+y \leq 1) = f(0,0) + f(0,1) + f(1,0)$

$$= \frac{3}{28} + \frac{6}{28} + \frac{9}{28} = \frac{18}{28}$$

c)

x	g(x)
0	10/28
1	15/28
2	3/28
	1

y	h(y)
0	15/28
1	12/28
2	1/28

d)

conditional probability distribution:

$$f(x_i/y_j) = P\left(\frac{x = x_i}{y = y_j}\right)$$

$$= \frac{P(x = x_i \text{ and } y = y_j)}{P(y = y_j)}$$

$$= \frac{f(x_i, y_j)}{h(y_j)}$$

$$f(y_j/x_i) = P\left(\frac{y = y_j}{x = x_i}\right)$$

$$= \frac{P(y = y_j \text{ and } x = x_i)}{P(x = x_i)}$$

$$= \frac{f(x_i, y_j)}{g(x_i)}$$

we have to find:

$$f(x/1) = P\left(\frac{x = x}{y = 1}\right)$$

$$= \frac{P(x = x \text{ and } y = 1)}{P(y = 1)}$$

$$= \frac{f(x,1)}{h(1)}$$

$$h(1) = f(0,1) + f(1,1) + f(2,1)$$

$$= 6/28 + 6/28 + 0 = 12/28$$

$$f(0/1) = \frac{f(0,1)}{h(1)} = \frac{6/28}{12/28} = \frac{1}{2}$$

$$f(1/1) = \frac{f(1,1)}{h(1)} = \frac{6/28}{12/28} = \frac{1}{2}$$

$$f(2/1) = \frac{f(2,1)}{h(1)} = 0$$

x	$f(x/1)$
0	½
1	½
2	0
	1

e) as $f(0,1) \neq g(0)\,h(1) \Rightarrow \frac{3}{14} \neq \frac{5}{14} \cdot \frac{3}{7}$ so events are dependent

Date: 22-3-23 Wednesday

Q. $P(x > x) = e^{-\lambda x}$ $x > 0$

i) Find cdf of x

ii) Find $P(T < x < 2T)$ where $T = \frac{1}{\lambda}$

Sol:

i) $F(x) = P(x \le x) = 1 - P(x > x)$

$= 1 - e^{-\lambda x}$

$P(a \le x \le b) = F(b) - F(a)$

$P(T \le x \le 2T) = 1 - e^{-\lambda(2T)} - [1 - e^{-\lambda(T)}]$

$= 1 - e^{-\lambda(2\frac{1}{\lambda})} - (1 - e^{-\lambda(1/\lambda)})$

$= 0.233$

Continuous bivariate probability density function

$f(x,y) \ge 0$

$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\, dy\, dx = 1$

$P(a \le x \le b, c \le y \le d) = \int_a^b \int_c^d f(x,y)\, dy\, dx$

$F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x,y)\, dy\, dx$

Margina function of x

$g(x) = \int_{-\infty}^{\infty} f(x,y)\, dy$

$h(y) = \int_{-\infty}^{\infty} f(x,y)\, dx$

Conditional pdf of x given y

$$f\left(\frac{x}{y}\right) = \frac{f(x,y)}{h(y)}$$

$$f\left(\frac{y}{x}\right) = \frac{f(x,y)}{g(x)}$$

- if $f(x,y) = g(x) \cdot h(y)$ then x and y are independent
- if $f(x,y) \neq g(x) \cdot h(y)$ then x and y are dependent

Q. $f(x,y) = \frac{1}{8}(6-x-y)$ $\quad 0 \le x \le 2$, $2 \le y \le y$

$\quad\quad 0$ $\quad$ otherwise

a) verify that $f(x,y)$ is a joint density function

b) $P(x \le 3/2,\ y \le 5/2)$

c) Find marginal pdf of x and y i.e., $g(x)$ and $h(y)$

d) Find the conditional p.d.f $f(x/y)$ and $f(y/x)$

a) $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dy\,dx = 1$$

$$\int_0^2 \int_2^4 \frac{1}{8}(6-x-y)\,dy\,dx = 1$$

$$= \frac{1}{8}\int_0^2 \left[6y - xy - \frac{y^2}{2}\right]_2^4 dx$$

$$= \frac{1}{8}\int_0^2 [(24-4x-8)-(12-2x-2)]\,dx$$

$$= \frac{1}{8}\int_0^2 (6-2x)\,dx$$

Date:

$$= \frac{1}{8}\left[6x - x^2\right]_0^2$$

$$= \frac{1}{8}\left[(12-4)-0\right]$$

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

$$= 1$$

b) $P(x \leq 3/2 \, , \, y \leq 5/2)$

$$\int_0^{3/2}\int_2^{5/2} \frac{1}{8}(6-x-y)\,dy\,dx$$

$$= 9/32$$

c) *Marginal pdf of x and y*

$$g(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$$

$$= \int_2^4 \frac{1}{8}(6-x-y)\,dy$$

$$= \frac{1}{8}\int_2^4 (6-x-y)\,dy$$

$$= \frac{1}{8}\left[6y - xy - \frac{y^2}{2}\right]_2^4$$

$$= \frac{1}{8}\left[(24-4x-8)-(12-2x-2)\right]$$

$$= \frac{1}{8}\left[6-2x\right]$$

$$= \frac{1}{4}(3-x) \quad 0 \leq x \leq 2$$

$$h(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$$

$$= \int_{0}^{2} \frac{1}{8}(6-x-y)\,dx$$

$$= \frac{1}{8}\left[6x - \frac{x^2}{2} - xy\right]_0^2$$

$$= \frac{1}{8}\left[(12-2-2y) - 0\right]$$

$$= \frac{1}{4}(5-y) \qquad 2 \le y \le 4$$

d)

$$f(x/y) = \frac{f(x,y)}{h(y)}$$

$$= \frac{1/8\,(6-x-y)}{1/4\,(5-y)}$$

$$= \frac{6-x-y}{2(5-y)}$$

$$f(y/x) = \frac{f(x,y)}{g(x)}$$

$$= \frac{1/8\,(6-x-y)}{1/4\,(3-x)}$$

$$= \frac{(6-x-y)}{2(3-x)}$$

Date: 27-3-23 Monday

$\boxed{\text{Mathematical expectation}}$ ⇒ mean

$$\mu_x = E(x) = \sum_{i=1}^{n} x\, f(x) \quad \text{(Discrete variable)}$$

$$\mu_x = E(x) = \int_{-\infty}^{\infty} x\, f(x) \quad \text{(continuous variable)}$$

variance :

$$\text{variance} = E(x^2) - [E(x)]^2$$

$$E(x^2) = \sum x^2 f(x) \quad \text{(Discrete)}$$

$$E(x^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx \quad \text{(continuous)}$$

Alternative :

$$\sigma^2 = \text{var}(x) = E(x-\mu)^2 = \sum (x-\mu)^2 f(x) \quad \text{(Discrete)}$$

$$\sigma^2 = \text{var}(x) = E(x-\mu)^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx \quad \text{(continuous)}$$

Q. what is the mathematical expectation for the no. of heads when 3 fair coins are tossed.

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

x : no of heads

X : 0, 1, 2

x	$f(x)$	$x\,f(x)$	$x^2 f(x)$
0	1/8	0	0
1	3/8	3/8	3/8
2	3/8	6/8	12/8
3	1/8	3/8	9/8
	1	12/8	24/8

$E(x) = \sum x\, f(x)$

$E(x) = \frac{12}{8} = 1.5$

1.5 heads per toss

for variance :

$$var(x) = E(x^2) - [E(x)]^2$$

$$E(x^2) = \Sigma x^2 f(x)$$

$$= 24/8$$

$$var(x) = \frac{24}{8} - (1.5)^2$$

$$= 0.75$$

Q. If the continuous random variable r.v x has p.d.f

$$f(x) = \begin{cases} \frac{3}{4}(3-x)(x-5) & 3 \le x \le 5 \\ 0 & \text{otherwise} \end{cases}$$

Find mean, variance and SD of x

$$E(x) = \int_{-\infty}^{\infty} x \, f(x) \, dx$$

$$= \int_3^5 x \left[3/4 (3-x)(x-5) \right] dx$$

$$= \frac{3}{4} \int_3^5 (-x^3 + 8x^2 - 15x) dx$$

$$= \frac{3}{4} \left[\frac{-x^4}{4} + \frac{8x^3}{3} - \frac{15x^2}{2} \right]_3^5$$

$$= \frac{3}{4} \left[\left(\frac{-125}{12} + \frac{64}{4} \right) \right]$$

$$= 4$$

Date:

$$E(x^2) = \int_{-\infty}^{\infty} x^2 f(x)\, dx$$

$$= \int_3^5 x^2 \left[\frac{3}{4}(3-x)(x-5)\right] dx$$

$$= \frac{3}{4}\int_3^5 (-x^4 + 8x^3 - 15x^2)\, dx$$

$$= \frac{3}{4}\left[-\frac{x^5}{5} + \frac{8x^4}{4} - \frac{15x^3}{3}\right]_3^5$$

$$= \frac{81}{5}$$

$$var(x) = E(x^2) - [E(x)]^2$$

$$= \frac{81}{5} - (4)^2$$

$$var(x) = 0.2$$

standard deviation $= \sqrt{var(x)}$

$$= \sqrt{0.2}$$

$$= 0.447$$

Q.

$$f(x) = \begin{cases} x^2/3 & -1 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

Find $E(x)$ and $var(x)$

$$E(x) = \int_{-1}^{2} x . f(x)\, dx$$

$$= \int_{-1}^{2} x . \frac{x^2}{3}\, dx$$

$$= \frac{5}{4}$$

$$E(x^2) = \int_{-1}^{2} x^2 \cdot \frac{x^2}{3} \, dx$$

$$= \frac{11}{5}$$

$$var(x) = E(x^2) - [E(x)]^2$$

$$var(x) = \frac{11}{5} - \left(\frac{5}{4}\right)^2$$

$$var(x) = 0.64$$

Chapter 4: Def 4.1, Theorem 4.1, Def 4.3, Theorem 4.2

day / date: 29-3-23 Wednesday

Expectation of a function of random variable

$g(x)$ or $h(x)$

$E[g(x)] = \sum g(x) f(x)$ (discrete)

$E[g(x)] = \int_{-\infty}^{\infty} g(x) f(x)\, dx$ (continuous)

where $g(x)$ is any function of x

Q. suppose that the no. of cars X that pass through a car wash b/w 4-5pm on any sunny friday has a following probability distribution

X	$P(X=x) = f(x)$	$2x-1$	$(2x-1)(f(x))$
4	1/12	$2(4)-1=7$	7/12
5	1/12	9	9/12
6	1/4	11	11/4
7	1/4	13	13/4
8	1/6	15	15/6
9	1/6	17	17/6
			12.67

let

let $g(x) = 2x-1$ represent the amount of money in dollars paid to the attendant by the manager. Find the attendants expected earning for the particular time period.

$E[(g(x)] = \sum (2x-1)(f(x))$

$E(2x-1) = 12.67$

Q. Let x be a random variable with pdf

$$f(x) = \begin{cases} 2(x-1) & 1<x<2 \\ 0 & \text{otherwise} \end{cases}$$

find the expected values of $g(x) = 2x-1$ and $g(x) = x^2$

$$E(2x-1) = \int_{-\infty}^{\infty} (2x-1)\, f(x)\, dx$$

$$= \int_1^2 (2x-1)\, 2(x-1)\, dx$$

$$= 2\int_1^2 (2x^2-3x+1)\, dx$$

$$= 2\left| \frac{2x^3}{3} - \frac{3x^2}{2} + x \right|_1^2$$

$$= 2\left[\left(\frac{16}{3} - 6 + 2\right) - \left(\frac{2}{3} - \frac{3}{2} + 1\right)\right]$$

$$= \frac{7}{3}$$

$$E[g(x)] = E(x^2)$$

$$= \int_1^2 x^2 f(x)\, dx$$

$$= \int_1^2 x^2 \cdot 2(x-1)\, dx$$

$$= 2\int_1^2 (x^3-x^2)\, dx$$

$$= 2\left| \frac{x^4}{4} - \frac{x^3}{3} \right|_1^2$$

$$= 2\left[\left(4-\frac{8}{3}\right) - \left(\frac{1}{4} - \frac{1}{3}\right)\right]$$

$$= \frac{17}{6}$$

variance:

day / date:

this formula is easy

$$\sigma_x^2 = var(x) = E(x-\mu)^2 \quad \text{OR} \quad E(x^2) - [E(x)]^2$$

$$\sigma_{g(x)}^2 = var[g(x)] = E[g(x) - \mu_{g(x)}]^2$$

$$var[g(x)] = E[g(x)^2] - [E g(x)]^2$$

Q. Find variance of $g(x) = 2x+3$

x	$f(x)$	$2x+3$	$(2x+3)f(x)$	$4x^2-12x+9$	$(4x^2-12x+9)f(x)$
0	1/4	3	3/4	9	9/4
1	1/8	5	5/8	1	1/8
2	1/2	7	7/2	1	1/2
3	1/8	9	9/8	9	9/8
			6		4

$$var[g(x)] = E[g(x) - \mu_{g(x)}]^2$$

$$\mu_{g(x)} = E(2x+3) = 6$$

$$= E[2x+3-6]^2$$

$$\sigma^2 = E(2x-3)^2$$

$$= E(4x^2 - 12x + 9)$$

Q. Let x be a random variable with pdf

$$f(x) = \begin{cases} x^2/3 & -1<x<2 \\ 0 & \text{otherwise} \end{cases}$$

Find variance of $g(x) = 4x+3$

$$\sigma^2_{g(x)} = E[g(x) - \mu_{g(x)}]^2$$

$$\mu_{g(x)} = \int_{-1}^{2} (4x+3)\, f(x)\, dx$$

$$E(4x+3) = \int_{-1}^{2} (4x+3) \cdot \frac{x^2}{3}\, dx$$

$$= 8$$

$$\sigma^2_{g(x)} = E[4x+3-8]^2$$

$$= E(4x-5)^2$$

$$= \int_{-1}^{2} (4x-5)^2 \frac{x^2}{3}\, dx$$

$$= \frac{1}{3}\int_{-1}^{2} (16x^4 - 40x^3 + 25x^2)\, dx$$

$$= \frac{1}{3}\left| \frac{x^5}{5} - \frac{40x^4}{4} + \frac{25x^3}{3} \right|_{-1}^{2}$$

$$= \frac{51}{5}$$

Mathematical expectation Q's

day / date:

Q. A lot containing 9 components is sampled by a quality inspector. The lot contains 4 good components and 3 defective components. A sample of 3 is taken by inspector. Find the expected value of the no of good components in this sample.

$$E(x) = \sum x P(x) \qquad x = 0, 1, 2, 3$$

$$P(x = x) = \frac{\binom{4}{x}\binom{3}{3-x}}{\binom{7}{3}}$$

$$E(x) = 0\left(\frac{1}{35}\right) + 1\left(\frac{12}{35}\right) + 2\left(\frac{18}{35}\right) + 3\left(\frac{4}{35}\right)$$

$$= \frac{12}{7} = 1.7$$

Q4.3 Let x be random variable that denotes the life in hrs of a certain electronic device. the prob density function is

$$f(x) = \begin{cases} \dfrac{20,000}{x^3} & x > 100 \\ 0 & \text{otherwise} \end{cases}$$

find expected life of device?

$$E(x) = \int_{-\infty}^{\infty} x f(x)\, dx$$

$$= \int_{100}^{\infty} x \frac{20,000}{x^3}\, dx$$

$$= 20,000 \left| \frac{x^{-1}}{x^{-1}} \right|_{100}^{\infty}$$

$$= -20,000\left[-\frac{1}{100}\right] = 200$$

on avg component's life is 200

day / date:

Q4.8. Let the random variable x represents the no of vehicles that are used for official business purposes on any working day. The prob distribution for company A is

X	f(x)
1	0.3
2	0.4
3	0.3

and that for company B is

X	f(x)
0	0.2
1	0.1
2	0.3
3	0.3
4	0.1

show that the variance of the prob distribution for comp B is greater thatn that company A.

Company A:

$var(x) = E(x-\mu)^2 p(x)$

$= \Sigma (x-\mu)^2 p(x)$

$\mu = E(x) = \Sigma x P(x)$

$= 1\times0.3 + 2\times0.4 + 3\times0.3 = 2$

$var(A) = 0.3 + 0 + 0.3 = 0.6$

Company B:

$var(x) = E(x-\mu)^2 P(x)$

$= \Sigma (x-\mu)^2 P(x)$

$\mu = E(x) = \Sigma x P(x)$

$= 0 + 1\times0.1 + 2\times0.3 + 3\times0.3$

$+ 4\times0.1 = 2$

$var(x) = 4(0.2) + 1(0.1) + 0$

$+ 0.3 + 4(0.1)$

$= 1.6$

~~4.9~~

Q. Let the random variable x represents the no of defective parts of a machine when 3 parts are sampled from a production time. The following is probability distribution of x

x	$P(x)$	$x\,P(x)$	$x^2 P(x)$
0	0.51	0	0
1	0.38	0.38	0.38
2	0.10	0.2	0.4
3	0.01	0.03	0.09
		0.61	0.87

$$\text{var}(x) = E(x^2) - (E(x))^2$$
$$= 0.87 - (0.61)^2$$
$$= 0.5$$

4.50 For a laboratory assignment if an equipment is working the density function of the observed outcome x is

$$f(x) = \begin{cases} 2(1-x) & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

find variance and standard deviation.

a) $$E(x^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx$$
$$= \int_0^1 x^2\, 2(1-x)\,dx$$
$$= \int_0^1 2x^2 - 2x^3\,dx$$
$$= 2\left[\frac{x^3}{3} - \frac{x^4}{4}\right]_0^1$$
$$= 2\left(\frac{1}{3} - \frac{1}{4}\right)$$
$$= 0.167$$

b) $$\sqrt{0.167}$$
$$= 0.4$$

Chapter # 5 day / date: 3-4-23 Monday

Binomial Distribution

1. The outcome of each trial may be classified into one of the two categories (success or failure).
2. The repeated trials are all independent.
3. The probability of success remains constant throughout the trial.
4. The experiment is repeated a fixed number of times say n.

Probability mass function: $\boxed{P(x=x) = \binom{n}{x} p^x q^{n-x}}$ $0 \leq x \leq n$

used to define binomial distribution

parameter : n, p

mean $= np$

variance $= npq$

x = no. of success

p = probability of success

q = probability of failure

n = no of trials

Q. A and B play a gam in which A has the probability of winning is 2/3 in a series of 8 games. a) what's the probability that a will win exactly 4 games? b) what's the probability that A will win at least 4 games?

a) $P(x=x) = \binom{n}{x} p^x q^{n-x}$

$P(x=4) = \binom{8}{4} \left(\frac{2}{3}\right)^4 \left(\frac{1}{3}\right)^{8-4}$

$P(x=4) = 0.1707$

p(winning) = 2/3

p(not winning) = q

$q = 1-p = 1-2/3 = 1/3$

b) $P(x \geq 4) = P(x=4) + P(x=5) + P(x=6) + P(x=7) + P(x=8)$

$= \binom{8}{4}\left(\frac{2}{3}\right)^4\left(\frac{1}{3}\right)^4 + \binom{8}{5}\left(\frac{2}{3}\right)^5\left(\frac{1}{3}\right)^3 + \binom{8}{6}\left(\frac{2}{3}\right)^6\left(\frac{1}{3}\right)^2 + \binom{8}{7}\left(\frac{2}{3}\right)^7\left(\frac{1}{3}\right)^1 + \binom{8}{8}\left(\frac{2}{3}\right)^8\left(\frac{1}{3}\right)^0$

day / date:

e.g 5.3

Q. A large chain retailer purchases a certain kind of electronic device from a manufacturer. The manufacturer indicates that defective rate of device is 3%

a) The inspector randomly picks 20 items from a shipment. what's the prob that there will be atleast 1 defective item among these 20?

b) Suppose that the retailer receives 10 shipments in a month and the inspector randomly tests 20 devices per shipment. what's the prob that there will be exactly 3 shipments each containing at least 1 defective device among the 20 that are defective and tested from the shipment?

a) $P(x \geq 1) = 1 - P(x < 1)$

$= 1 - P(x = 0)$

$= 1 - \left[\binom{20}{0} (0.03)^0 (0.97)^{20} \right]$

$= 0.4562$

$n = 20$
$p = 0.03$
$q = 0.97$

b) $P(x = 3) = \binom{10}{3} (0.4562)^3 (1 - 0.4562)^7$

$= 0.1602$

$n = 10$
$p = 0.4562$
$q = 1 - 0.4562$

Binomial Distribution

day / date:

5.1

Q. The prob that a certain kind of component will survive a shock test is 3/4. Find the prob that exactly 2 of the next 4 components tested survived.

$$P(x=2) = \binom{4}{2}\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^2$$

$$= \frac{27}{128}$$

5.2

Q. The prob that a patient recovers from a rare blood disease is 0.4. If 15 patients are known to have this disease, what's the prob that

a) at least 10 survive

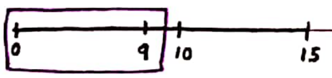
b) from 3 to 8 survive

c) exactly 5 survive

a) $P(x \geq 10) = P(x=10) + P(x=11) + P(x=12) + P(x=13) + P(x=14) + P(x=15)$

$1 - \sum_{x=0}^{9} b(x,n,p)$

$$= 0.0338$$



1 - this gives answer

b) $P(3 \leq x \leq 8) = \sum_{x=0}^{8} b(x,n,p) - \sum_{x=0}^{2} b(x,n,p)$

$$= 0.905 - 0.0271$$

$$= 0.8779$$

c) $P(x=5) = \binom{15}{5}(0.4)^5(0.6)^{10}$

$$= 0.186$$

variance = 3.6

mean = 6

day / date:

Q. A nationwide servey of clg seniors in a uni revealed that almost 70 % disapproved smoking. If 12 seniors are selected at random and asked their opinions, find the prob that the number who disapprove smoking is

a) anywhere from 7 to 9

b) atmost 5

a) $P(7 \leq x \leq 9) = \binom{12}{7}(0.7)^7(0.3)^5 + \binom{12}{8}(0.7)^8(0.3)^4 + \binom{12}{9}(0.7)^9(0.3)^3$

$= 0.6294$

b) $P(x \leq 5) = \binom{12}{0}(0.7)^0(0.3)^{12} + \dots + \binom{12}{5}(0.7)^5(0.3)^7$

$= 0.0386$

Scanned with CamScanner

Multinomial Distribution

1. The outcome of each trial may be classified into one of the k mutually exclusive categories $C_1, C_2, \ldots, C_k$.
2. The repeated trials are all independent.
3. The probability of i^{th} outcome is P_i, which remains constant and $\Sigma P_i = 1$.
4. The experiment is repeated a fixed number of times.

day / date:

$$P(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k) = \binom{n}{x_1, x_2, \dots, x_k} P_1^{x_1} P_2^{x_2} \dots P_k^{x_k}$$

$$\binom{n}{x_1, x_2, \dots, x_k} = \frac{n!}{x_1!\, x_2! \dots x_k!}$$

$$\boxed{\begin{array}{l} \sum_{i=1}^{k} P_i = 1 \\ \sum_{i=1}^{k} x_i = n \end{array}}$$

Parameters : $n, P_1, P_2, \dots, P_k$

Mean = $nP_i \quad (1 \leq i \leq k)$

Variance = $nP_i q_i$

Exp 5.7

Q. The complexity of arrivals and departures of planes at an airport is such that computed simulation is often used to model the ideal conditions. For a certain airport with 3 runways, it is known that in the ideal setting the following are the probabilities that the individual runways are accessed by a randomly arriving commercial jet.

Run way 1 : $P_1 = 2/9$
Run way 2 : $P_2 = 1/6$
Run way 3 : $P = 11/18$

what's the prob that 6 randomly arriving airplanes are distributed in the following fashion.

Run way 1 : 2 Airplanes
Run way 2 : 1 Airplanes
Run way 3 : 3 Airplanes

$$P(X_1 = 2, X_2 = 1, X_3 = 3) = \binom{6}{2 \quad 1 \quad 3} \left(\frac{2}{9}\right)^2 \left(\frac{1}{6}\right)^1 \left(\frac{11}{18}\right)^3$$

$$= \frac{6!}{2!\ 1!\ 3!} \left(\frac{2}{9}\right)^2 \left(\frac{1}{6}\right)^1 \left(\frac{11}{18}\right)^3$$

$$= 0.1127$$

day / date:
Practise Q's: 5.22, 5.20

Multinomial Distribution Q's

Q5.23

Q. The probabilities are 0.4, 0.2, 0.3 and 0.1 respectively that a delegate to a certain convention arrived by air, bus, automobile or train. What's the prob that among 9 delegates randomly selected at this convenvention 3 arrived by air, 3 arrived by bus, 1 arrived by train and 2 arrived by train

$$P(x_1=3, x_2=3, x_3=1, x_4=2) = \binom{9}{3\ 3\ 1\ 2}(0.4)^3(0.2)^3(0.3)^1(0.1)^2$$

$$= \frac{9!}{3!\ 3!\ 1!\ 2!}(0.4)^3(0.2)^3(0.3)^1(0.1)^2$$

$$= 0.0077$$

Practise Q's : 5.30, 5.31

day / date: 5-4-23
Wednesday

Hyperbgeometric Distribution

⇒ outcomes classified into two categories. success or failure
⇒ trials are dependent
⇒ Probability of success changes on each trial
⇒ Experiment is repeated a fixed no. of times

$$P(x=x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

$k = 0, 1, 2, \ldots, n$

$x = 0, 1, 2, \ldots, k$ (when $k < n$)

or

$x = 0, 1, 2, \ldots, n$

N : No of objects/units in the population.
n : No of objects/units in the sample.
k : No of success in population
N-k : No of failures in the population
x : No of success in a sample
n-k : No of failures in sample

Parameters: N, n, k

$\text{Mean} = n \cdot \frac{k}{N}$

$P = \frac{k}{N}$ probability of success

$$\text{Variance} = n\frac{k}{N}\overbrace{\left(1-\frac{k}{N}\right)}^{q}\left(\frac{N-n}{N-1}\right)$$

Q. A box contains 4 red and 6 black balls. A sample of 4 balls is selected without replacement. Let x be no of red balls contained in the sample. Then find the probability of exactly 4 red balls.

Black	Red	
6	4	= 10

$N = 10$
$n = 4$
$k = 4$

$$P(x=4) = \frac{\binom{4}{4}\binom{6}{0}}{\binom{10}{4}}$$

$$= \frac{1}{210}$$

day / date:

Q. Names of 5 men and 5 women are written on slips of paper and placed in a hat. 4 names are drawn. What's the prob that 2 are men.

M	W	
5	5	= 10

$N = 10$
$n = 4$
$k = 5$
$x = 2$

$$P(x=2) = \frac{\binom{5}{2}\binom{5}{2}}{\binom{10}{4}}$$

$$= \frac{10}{21}$$

Hyper Mathematical Distribution

Q. lots of 40 components each are deemed unacceptable if they contain 3 or more defectives. The procedure for sampling a lot is to select 5 components at random and 2 reject the lot if a defective is found. What is the probability that exactly 1 defective is found in the sample if there are 3 defectives in the entire lot.

$$P(x=1) = \frac{\binom{3}{1}\binom{37}{4}}{\binom{40}{5}}$$
$$= 0.3011$$

$N = 40$
$n = 5$
$k = 3$
$x = 1$
$N-k = 37$

- If $\frac{n}{N} \leq 0.05$ then binomial approximation can be used

but we require p

$$\Rightarrow p = \frac{k}{N}$$

Q. A manufacturer of automobile tyres reports that among a shipment of 5000 sent a local distributor, 1000 are slightly blemished. If one purchases 10 of these tyres random from the distributor, what's the prob that exactly 3 are blemished?

$N = 5000$
$n = 10$
$\frac{n}{N} = 0.002$
$k = 1000$
$p = \frac{k}{N} = 0.2$

By binomial:

$$P(x=3) = \binom{10}{3} 0.2^3\, 0.8^7$$
$$= 0.2013$$

By hyper:

$$P(x=3) = \frac{\binom{1000}{3}\binom{4000}{7}}{\binom{5000}{10}} = 0.2015$$

Table 82 → Poisson values

Poisson Distribution

day / date: 17-4-23 Monday

Avg no. of occurrences in a given time or space or length.

$$P(x, \lambda) = \frac{e^{-\lambda t}(\lambda t)^x}{x!} \qquad \mu = \lambda t$$

$$= \frac{e^{-\mu}\mu^x}{x!} \qquad x = 0, 1, 2, \ldots$$

λ : avg no of occurences per time, distance, area

t : specific time, distance, area or volume

x : no of occurences in time t

e : 2.71828

parameter = $\mu = \lambda t$

mean = μ

variance = μ

Ex. 5.16

Q. 10 is the avg no of oil tankers arriving each day at a certain port. The facilities at the port can handle at most 15 tankers per day. Find prob that on a given day tankers have to be turned away?

$$P(x > 15) = 1 - P(x \le 15) \qquad \mu = \lambda t = 10$$

$$= 1 - [P(x=0) + P(x=1) + \ldots + P(x=15)]$$

$$= 1 - \left[\frac{e^{-10}10^0}{0!} + \frac{e^{-10}10^1}{1!} + \frac{e^{-10}10^2}{2!} + \ldots + \frac{e^{-10}10^3}{3!}\right]$$

$$= 0.0487$$

Ex. 5.17

Q. During a laboratory experiment, the avg no radioactive particles passing through a counter in 1 ms is 4. What's the prob that 6 particles enter the counter in a given ms?

$$\mu = \lambda t = 4 \times 1 = 4$$

$$P(x=6) = \frac{e^{-\mu}\mu^x}{x!}$$

$$= \frac{e^{-4}4^6}{6!}$$

$$= 0.1042$$

day / date:

Approximation of binomial distribution by poisson:

if n is large and p is small

$n \geq 20$ and $p \leq 0.05$

μ (mean of binomial) $= np$ $b(x,n,p) \rightarrow P(x,\lambda)$

Q. In a certain industrial facility accidents occur infrequently. It is known that the prob of an accident on any given day is 0.005 and accidents are independent of e/o.

a) what's the prob that in any period of 400 days, there will be an accident on 1 day?

b) what's the prob that there are atmost 3 days with an accident?

$p <<$
$n >>$

$$\mu = np = 400 \times 0.005 = 2$$

$$P(x=1) = \frac{e^{-2}\,2^1}{1!}$$

$$= 0.27$$

$$P(x \leq 3) = P(x=0) + P(x=1) + P(x=2) + P(x=3)$$

$$= \frac{e^{-2}\,2^0}{0!} + \frac{e^{-2}\,2^1}{1!} + \frac{e^{-2}\,2^2}{2!} + \frac{e^{-2}\,2^3}{3!}$$

$$= 0.857$$

wo experiments jin ka graph symmetric ho

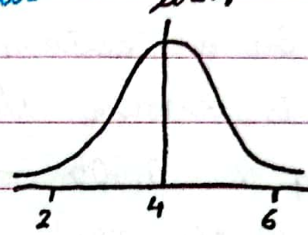
Normal Distribution

day / date: 19-4-23 Wednesday

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad N(x, \mu, \sigma)$$

$-\infty < x < +\infty$

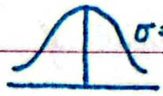
μ : location parameter
σ : shape parameter



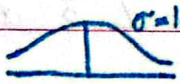
Parameters: μ (mean) and σ (standard deviation)

$$Z = \frac{x-\mu}{\sigma}$$

$$f(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

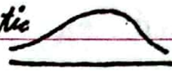


$-\infty < Z < \infty$



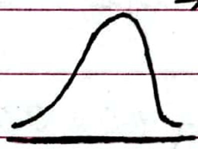
$\boxed{\mu = E(z) = 0}$ $\boxed{\sigma = v(z) = 1}$

- platykurtic curve



- normal curve ⇒ mesokurtic curve

- leptokurtic curve

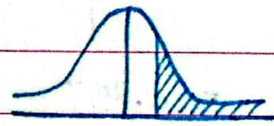


6.2 Q. Given a standard normal distribution, find the area under the curve that lies

a) to the right of $z = 1.84$ $\qquad P(Z > 1.84)$

b) b/w $z = -1.97$ and $z = 0.86$

a) $P(z > 1.84) = 1 - P(z \leq 1.84)$
$= 1 - 0.9671$
$= 0.0329$



b) $P(-1.97 \leq z \leq 0.86) = P(z \leq 0.86) - P(z \leq -1.97)$
$= 0.8051 - 0.0244$
$= 0.7807$

Q 6.4. $\mu = 50$, $\sigma = 10$
$P(45 \leq x \leq 62) = ?$

The values of Z corresponding to $x_1 = 45$ and $x_2 = 62$ are

$$z_1 = \frac{45-50}{10} = -0.5 \quad \text{and} \quad z_2 = \frac{62-50}{10} = 1.2$$

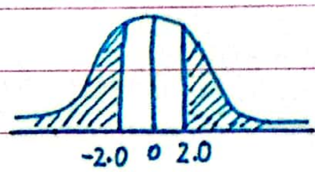
e.g 6.5

day / date:

$$P(45 \le x \le 62) = P(-0.5 \le z \le 1.2)$$
$$= P(z \le 1.2) - P(z \le 0.5)$$
$$= 0.8849 - 0.3085 = 0.5764$$

Q6.9. In an industrial process, the diameter of a ball bearing is an important measurement. The buyer sets specifications for the diameter to be 3.0 ± 0.01cm. The implication is that no part falling outside these specifications will be accepted. It is known that in the process the diameter of a ball bearing has a normal distribution with mean $\mu = 3.0$ and standard deviation $\sigma = 0.005$. On avg, how many manufactured ball bearings will be scrapped?

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{2.99 - 3.0}{0.005} = -2.0$$



$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{3.01 - 3.0}{0.005} = 2.0$$

$1 - P(-2.0 \le z \le 2.0)$

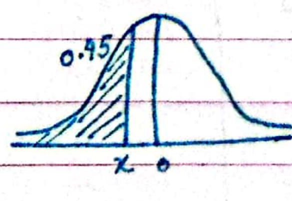
$$P(-2.0 \le z \le 2.0) = P(z \le 2.0) - P(z \le -2.0)$$
$$= 0.9772 - 0.0228$$
$$= 0.9544$$
$$1 - P(-2.0 \le z \le 2.0) = 1 - 0.9544$$
$$= 0.0456 \times 100$$
$= 4.56\%$ will be scrapped

Using normal curve in reverse

Q 6.6. Given a normal distribution with $\mu = 40$ and $\sigma = 6$, find x that has
a) 45% of area to the left b) 14% of area to the right

a) $P(z \le -0.13) = 0.45$

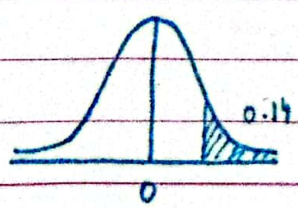
$$z = \frac{x - \mu}{\sigma}$$



$$x = \mu + \sigma z = 40 + 6(-0.13) = 39.22$$

b) $P(z \le 1.08) = 0.86$

$$x = \mu + \sigma z$$
$$= 40 + 6(1.08)$$
$$= 46.48$$



Q's of Normal distribution

e.g 6.8, 6.11
day / date:
ex 6.13, 6.14, 6.11 till C, 6.8 (first 3 part)

6.7
Q. A certain type of storage battery last on avg 3 years with SD of 0.5 year. Assuming the battery life is normally distributed find prob that a given battery will last less than 2.3 years.

6.4

$$P(x < 2.3) = P\left(\frac{x-\mu}{\sigma} < \frac{2.3-3}{0.5}\right)$$

$$= P(z < -1.4)$$

$$= 0.0808$$

6.12
Q. The loaves of bread distributed to local stores by a certain bakery have an avg length of 30 cm and SD of 2cm Assuming that the lengths are normally distributive, what's the prob that

a) longer than 31.7 cm
b) b/w 29.3 and 33.5 cm
c) shorter than 25.5 cm

a) $P(x > 31.7) = P\left(\frac{x-\mu}{\sigma} > \frac{31.7-30}{2}\right)$

$$= P(z > 0.35)$$

$$= 1 - P(z \leq 0.35)$$

$$= 1 - 0.8023 = 0.1977$$

c) $P(x < 25.5)$

$$= P\left(z < \frac{25.5-30}{2}\right)$$

$$= P(z < -2.25)$$

$$= 0.0122$$

b) $P(29.3 \leq z \leq 33.5) = P\left(\frac{29.3-30}{2} < x < \frac{33.5-30}{2}\right)$

$$= P(-0.35 \leq z \leq 1.75)$$

$$= P(z \leq 1.75) - P(z \leq -0.35)$$

$$= 0.9599 - 0.3632$$

$$= 0.5967$$

day / date: 26-4-23
Wednesday

Statistical Inference

- Estimation
  - point estimation — eik point ki form mai estimation dena
  - interval estimation — eik interval ki form mai estimation dena
- Testing of hypothesis

Hypothesis or Statistical hypothesis:

- Null hypothesis H_0 — equality ati hai e.g $\leq, \geq, =$
- Alternative hypothesis H_1 or H_A — equality nahi ati e.g $<, >, \neq$

* we will take either of the two hypothesis

$\left.\begin{matrix} \mu = 60 \\ \sigma = 4 \end{matrix}\right\}$ simple hypothesis

$\left.\begin{matrix} \mu > 60 \\ \sigma = 40 \\ \text{or} \\ \mu = 60 \\ \sigma > 40 \end{matrix}\right\}$ composite hypothesis

Test statistic:

Errors:

Type I error: rejection of true hypothesis

Type II error: acception of false hypothesis

True situation	Decision: Donot reject or Accept H_0	Decision: Reject H_0
H_0 is true	Correct Decision (no error)	Type I error
H_0 is false	Type II error	Correct Decision (no error)

level of significance

$\alpha = P(\text{Type I error})$
$= P(\text{Reject } H_0 / H_0 \text{ is true})$

day / date:

General procedure for testing of hypothesis:

i) State the null and alternate hypothesis

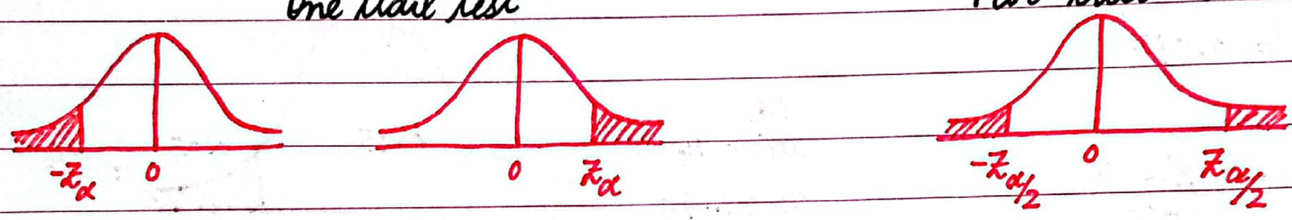
ii) choose α. If it's not given, put it equal to 5%

iii) choose an appropriate test statistic and establish the critical region based on α

iv) Reject H_0 if computed value of test statistic is in the critical region, otherwise do not reject.

v) Draw conclusions

One trail test — Two trail test



$H_1 : \mu < \mu_0$ — $H_1 : \mu > \mu_0$ — $H_0 : \mu = \mu_0$, $H_1 : \mu \neq \mu_0$

shaded region ⇒ critical region / rejection region

plain region ⇒ acceptance region

Testing hypothesis about mean: (single sample)

i) Testing on a single mean when population variance is known (Z - test)

ii) Testing on a single mean when population variance is unknown (t-test)

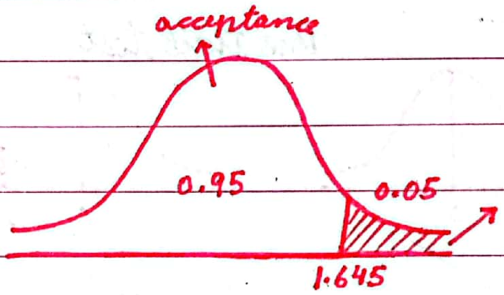
Q. A random sample of 100 recorded deaths in the US during the past year showed an average life span of 71.8 years. Assuming a population's standard deviation of 8.9 years, does this seem to indicate that the mean life span today is greater than 70 years. Use a 5% level of significance.

1) $H_0 : \mu \leq 70$ or $H_0 : \mu = 70$
$H_1 : \mu > 70$

2) $\alpha = 0.05$

3) $z = \dfrac{\bar{X} - \mu}{\sigma / \sqrt{n}}$

CR
reject H_0 if
$z > 1.645$



4) $z = \dfrac{71.8 - 70}{8.9 / \sqrt{100}}$

$= 2.02$

5) Decision: Reject H_0
Conclusion: Mean life span is greater than 70

Ex Q's :
10.23, 10.26, 10.29

day / date:

t- distribution

hypothesis testing procedure : (one sample mean)

i) $H_0 : \mu = \mu_0$ | $\mu \leq \mu_0$ | $\mu \geq \mu_0$

$H_1 : \mu \neq \mu_0$ | $\mu > \mu_0$ | $\mu < \mu_0$

ii) level of significance: $\alpha = 5\%$

iii) test- statistics :

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

iv) critical region :

v: degrees of freedom

$v : n-1$

if $H_1 : \mu > \mu_0$ $\quad t \geq t_{\alpha(v)}$

if $H_1 : \mu < \mu_0$ $\quad t \leq -t_{\alpha(v)}$

if $H_1 : \mu \neq \mu_0$ $\quad t \leq -t_{\alpha/2(v)}$ & $t \geq t_{\alpha/2(v)}$

v) conclusion

confidence interval estimation

$$\bar{x} \pm Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$\bar{x} \pm t_{\alpha/2(v)}\left(\frac{s}{\sqrt{n}}\right)$$

day / date: 3-5-23 Wednesday

Statistical Inferences

29 10.4

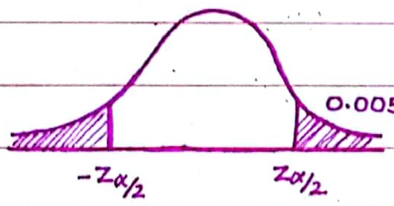
Q. A manufacturer of sports equipments has developed a new synthetic fishing line that the company claims has a mean breaking strength of 8 kg and SD of 0.5 kg. Test the hypothesis that $\mu = 8$ against the alternative that $\mu \neq 8$ if a random sample of 50 lines is tested and found to have a mean breaking strength of 7.8 kg. Use a 0.01 level of significance.

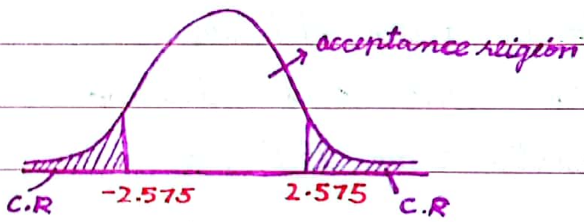
1) $H_0 = \mu_0 = 8$

   $H_1 = \mu \neq 8$

2) $\alpha = 0.01$ $\quad \frac{\alpha}{2} = 0.005$

3) $z = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}}$





Reject H_0 if $z \leq -2.575$ and $z > 2.575$

4) $z = \dfrac{7.8 - 8}{0.5/\sqrt{50}} = -2.83$

5) Decision : Reject H_0

Average breaking strength is not equal to 8.

day / date:

Testing of mean when population S.D or population variance is unknown

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

eg 10.5

Q. The Edison Electric Institute has published figures on the number of kilowatt hours used annually by variance home appliances. It is claimed that a vacuum cleaner uses an average of 46 kilowatt hours per year. If a random sample of 12 homes included in a planned study indicates that vacuum cleaners use an average of 42 kilowatt hours per year with a standard deviation of 11.9 kilowatt hours. Does this suggest at the 0.05 level of significance that vacuum cleaners use, on average, less than 46 kilowatt hours annually? Assume the population of kilowatt hours to be normal.

1) $H_0 : \mu \geq 46$ or $\mu = 46$

$H_1 : \mu < 46$

2) $\alpha = 0.05$

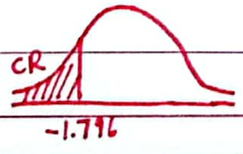
3) $t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$ with degrees of freedom $= n-1 = 12-1 = 11$

critical region

Reject H_0 if $t < t_{(\alpha, n-1)}$

$t < t_{(0.05, 11)}$

$t < -1.796$



4) $t = \frac{42 - 46}{11.9/\sqrt{12}}$

$= -1.16$

5) Do not reject H_0

conclusion: The avg number of kilowatt hours used annually by home vacuum cleaners is not less than 46

day / date: 15-5-23
Monday

$\hat{Y} = b_0 + b_1 X$

Regression

Qa. In an experiment to measure the stiffness of a spring, The length of the spring under different loads was measured as follows

Loads (X)	Lengths (Y)	XY	X^2	Y^2
3	10	30	9	100
5	12	60	25	144
6	15	90	36	225
9	18	162	81	324
10	20	200	100	400
12	22	264	144	484
15	27	405	225	729
20	30	600	400	900
22	32	704	484	1024
28	34	952	784	1156
130	220	3467	3288	5486

Find the regression eq. appropriate for predicting the length, given the weight on the spring

$$\hat{Y} = b_0 + b_1 X$$

$$b_1 = \frac{n\Sigma XY - \Sigma X \Sigma Y}{n\Sigma X^2 - (\Sigma X)^2} = \frac{10(3467) - (130)(220)}{10(2288) - (130)^2} = 1.02$$

$$b_0 = \bar{Y} - b_1\bar{X} = \frac{\Sigma Y}{n} - b_1 \frac{\Sigma X}{n}$$

$$= \frac{220}{10} - (1.02)\left(\frac{130}{10}\right)$$

$$= 8.74$$

$$\Rightarrow \hat{Y} = 8.74 + 1.02X$$

day / date:

if load was $X = 40$ what would be value of Y?

$\Rightarrow$ $Y = 8.74 + 1.02(40)$

$\hat{Y} = 49.54$

* $\hat{Y}$ hat represents that this is estimated value of Y.

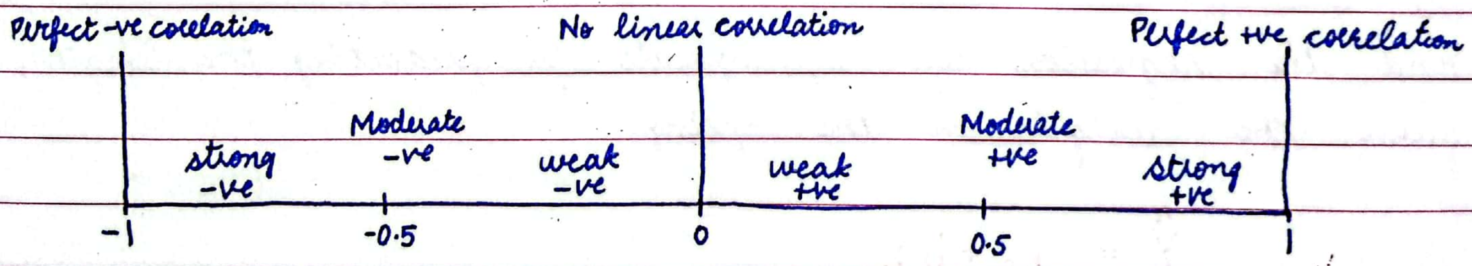
b. Find corelation cofficient : Ranges from -1 to 1

$$r = \frac{n\sum XY - \sum X \sum Y}{\sqrt{\left(n\sum X^2 - (\sum X)^2\right)\left(n\sum Y^2 - (\sum Y)^2\right)}}$$

$$r = \frac{10(3467) - (130)(220)}{\sqrt{\left(10(2288) - (130)^2\right)\left(10(5484) - (220)^2\right)}}$$

$$r = 0.97$$

There is strong +ve corelation b/w length and loads since r is very close to 1.



how much variablility of dependent can be explained by independent variable. day / date:

Coefficient of determination:

$$R^2 = 1 - \frac{\Sigma(Y-\hat{Y})^2}{\Sigma(Y-\bar{Y})^2} = 1 - \frac{\Sigma Y^2 - b_0\Sigma Y - b_1\Sigma XY}{\Sigma Y^2 - \frac{(\Sigma Y)^2}{n}}$$

$$= 1 - \frac{5486 - 8.74(220) - (1.02)(3467)}{5486 - \frac{(220)^2}{10}}$$

$$R^2 = 0.958$$

as $R^2 = 0.958$ indicates that 95.8% of the variability in Y, the length of the spring is explained by its linear relationship with X the weight.

* R^2 is square of r.

KAGHAZ