PROBABILITY & STATISTICS

Observation: Any recording of information whether it is numerical or categorical is called an observation

Experiment: Experiment any process which generate any set of data

Sample space: The set of all possible outcomes of an experiment is called an sample space.

$S = \{H, T\}$

Sample Space

element or sample point (finite or infinite)

Ex:1 set of cities with population over 10 million

$S = \{x / x$ is having population one crore $\}$

$S = \{t / t \geq 0\}$ $\quad t =$ time

Ex:2 Throwing a die

$S_1 = \{1, 2, 3, 4, 5, 6\}$ more information

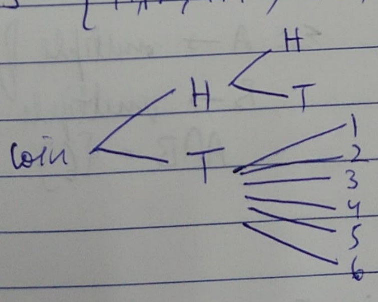
$S_2 = \{odd, even\}$

[Both are correct but S_1 gives more info about the outcomes so S_1 preferred]

Example 1

An experiment consists of flipping a coin and then flipping it a second time, if head occurs; if tail occurs in 1st flip, the die is tossed.

$SS = \{HH, HT, T1, T2, T3, T4, T5, T6\}$



Q2 3 items are selected at random from a manf process. Each item inspected and classified as defective and non-defective.

A2 SS = {DDD, DDN, DND, DNN, NDD, NDN, NND, NNN}

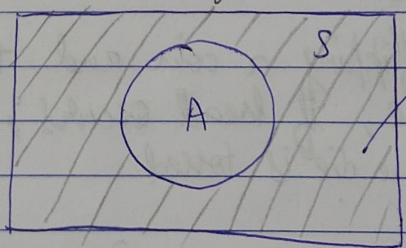
Events
↳ subset of sample space

Ex1 $S = \{1, 2, 3, 4, 5, 6\}$
event A → multiple of 3
$A = \{3, 6\} \subseteq S$

Ex2 Bulb lives atleast 5 years
$S = \{t \mid t \geq 0\}$
$A = \{t \mid t \geq 5\} \subset S$

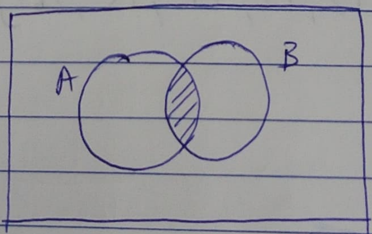
Complement of A

(1)



→ shaded $\bar{A}$

(2) $A \cap B$



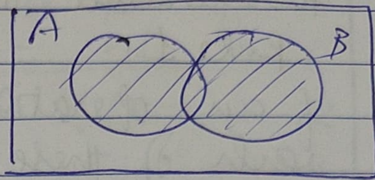
Ex: A → multiple of 2
B → multiple of 3
$A \cap B = \{6\}$

The intersection of A and B in the event containing all events common to A and B

③ A∪B

~~Ex: A:~~

An event containing all the elements of A and B & both



Mutually exclusive / Disjoint

~~if A∩~~

$\boxed{\text{if } A \cap B = \phi}$

Fundamental Theorem of multiplication

Rule 1:

If an operation can be performed in m_1 ways and if for each of these ways a second operation can be performed in m_2 ways then the two operations can be performed together in $m_1 \times m_2$ ways

Ex: Throwing a pair of dice

Q: Neha is going to assemble a computer by herself. She has choice of chips from two brands, a harddrive from 4 and memory from 3 and accessory bundle from 5 brands. How many diff ways can Neha find

chip	2	$2 \times 4 \times 3 \times 5 = 120$ ways
harddrive	4	
memory	3	
accessory bundle	5	

Permutation

no of arrangements of n objects $= n!$

no of permutation of n distinct object's taken one at a time

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

Ex: Taking two letters at a time out of 4

Ans $4P_2 = \frac{4!}{(4-2)!} = \frac{24}{2} = 12$

Q: In 1 year, 3 distinct awards are to be given to 25 students, if each student received atleast 1 award, How many possible selection are there

$$25P_3 = 25 \times 24 \times 23$$

Combination: selection of items from set of n items

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

Q. A boy asked his mother for 3 arcade and 2 sports. It has 10 arcade and 5 sports games

$\Rightarrow \; ^{10}C_3 \times {^5}C_2$

Probability of an event

$S = \{S_1, S_2, S_3 \dots S_n\}$

$A = \{S_1, S_2, S_3 \dots S_n\}$

$P(S) = 1$ (probability sum of all elements of S)

$0 \le P(A) \le 1$

$P(\phi) = 0$

Q1 A coin is tossed twice, what is the probability that atleast one head occurs?

$S = \{\underbrace{HH, HT, TH, TT}_{\text{equally likely}}\}$

$4w = 1$

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

or simply $P(S_1) = \frac{1}{4} = \frac{\text{no of event}}{\text{Total event}}$

$P(\text{atleast one head}) \longrightarrow = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$

Q2 A die is loaded in such a way that an even no is twice as likely to occur than an odd no (a) If A is an event that the no less than 4 occurs in a single roll of die, Find prob. of A

$S = \{1,2,3,4,5,6\}$ $P(\text{even}) = 2\omega$

$A = \{1,2,3\}$ $P(\text{odd}) = \omega$

$P(S) = \omega + 2\omega + \omega + 2\omega + \omega + 2\omega$

$= 9\omega = 1$

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

$P(A) = \frac{1}{9} + \frac{2}{9} + \frac{1}{9} = \frac{4}{9}$

(b) B is an event having even no. and C is an event having number multiple of 3

$B = \{2,4,6\}$

$C = \{3,6\}$

$B \cup C = \{2,3,4,6\}$

$P(B \cup C) = \frac{2}{9} + \frac{1}{9} + \frac{2}{9} + \frac{2}{9} = \frac{7}{9}$

$B \cap C = \{6\}$

$P(B \cap C) = \frac{2}{9}$

CONCEPT

$S = \{S_1, S_2, S_3 \ldots\ldots, S_N\}$ $n \leq N$

$A = \{S_1, S_2, S_3 \ldots\ldots, S_N\}$

$P(S) = 1$

$P(S_i) = \frac{1}{N}$

$P(A) = P(S_1) + P(S_2) + P(S_3) + \ldots P(S_n)$

$P(A) = \frac{N}{N} = 1$

Additive rule

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

   If $A \cap B = \phi$

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

2. $P(A \cup B \cup C) = P(A) + P(B) + P(C) - [P(A \cap B) + P(B \cap C) + P(C \cap A)] + P(A \cap B \cap C)$

Ex: After two companies, Neha assesses that probability of getting offer from company A is 0.8 and that of company B is 0.6. If she believes that she will receive offers from both the companies is 0.5, find $P(A \cup B) = ?$

$P(A) = 0.8$

$P(B) = 0.6$

$P(A \cap B) = 0.5$

$P(A \cup B) = 0.8 + 0.6 - 0.5 = 0.9$

* $\boxed{P(A) + P(A') = 1}$

~~A~~ A' $A \cup A' = S$

$P(A \cup A') = P(S) = 1$

$P(A) + P(A') + \underbrace{P(A \cap A')}_{0} = 1$

so $\underline{P(A) + P(A') = 1}$

Ex: If the probability that a automobile mechanic will serve 3, 4, 5, 6, 7, 8 or more cases on any particular day is 0.12, 0.19, 0.28, 0.24, 0.1 and 0.07. What is the probability that he service at least 5 cases on next day

$X = $ (no of cases serviced)

$P(X \geq 5) = 1 - P(x < 5)$

$= 1 - [P(3) + P(4)]$

$= 1 - [0.12 + 0.19]$

$= 0.69$

Conditional Probability

A has occurred, then find probability of A

$$P(B/A) = \frac{P(A \cap B)}{P(A)} \quad ; \quad P(A) > 0$$

probability of B given
A already occurred

Ex: Event B is getting a perfect square when a die is tossed, the die is constructed such as probability of even no. is twice as likely to occur as odd no. A is getting a no greater than 3

$B = \{1, 4\}$ perfect square

$P(\text{even}) = \frac{2}{9}$ $\quad P(\text{odd}) = \frac{1}{9}$ $\qquad P(A \cap B) = \frac{2}{9}$

$P(B) = \frac{2}{9} + \frac{1}{9} = \frac{1}{3}$

$A = \{4, 5, 6\}$

$P(A) = \frac{2}{9} + \frac{2}{9} + \frac{1}{9} = \frac{5}{9}$

$P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{2}{9} \times \frac{9}{5} = \frac{2}{5}$

Ex:

Gender	Employed	Unemployed	Total
Female	140	260	400
Male	460	40	500
	600	300	900

S is the population of adults in a town who have completed degree

S = completed degree
M → a man is chosen
E → employed is chosen

$P(M/E) \rightarrow \dfrac{P(M \cap E)}{P(E)} = \dfrac{460}{600} = \dfrac{23}{30}$

Ex A is an event ace is chosen and B being an event spade is chosen

$P(A) = \dfrac{1}{13}$, $P(B) = \dfrac{1}{4}$

$P(B/A) = \dfrac{P(A \cap B)}{P(A)} = \dfrac{1/52}{4/52} = \dfrac{1}{4}$

$P(B/A) = P(B)$ & $P(A/B) = P(A)$
∴ B & A are independent events

Multiplicative Rule

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

~~Ex: suppose a fuse box has 20 fuses of which 5~~

Q: Suppose a fuse box has 20 fuses of which 5 are defectives. If two fuses are selected at random and removed from the box without replacing the first one. What is the probability that both fuses are defectives

A: A : first fuse is defective
B : second fuse is defective

$$P(A \cap B) = P(B/A)\,P(A)$$
$$= P(A/B)\,P(B)$$

$$P(A) = \frac{5}{20} = \frac{1}{4}$$

$$P(B/A) = \frac{4}{19}$$

so $P(A \cap B) = \frac{1}{4} \times \frac{4}{19} = \frac{1}{19}$

* If A & B are independent then
$P(B/A) = P(B)$ & $P(A/B) = P(A)$

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

Q: A small town has a fire engine (FE) and ambulance (AM). The probability of having FE available is 0.98 and that of AM = 0.92. In case of injury, find out probability that both are available, consider AM & FE as independent

$P(FE) = 0.98$, $P(AM) = 0.92$

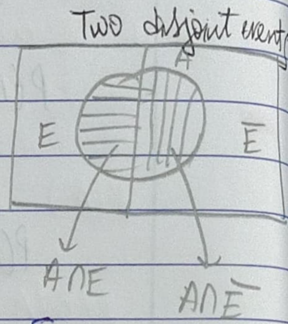
$$P(FE \cap AM) = P(FE) \cdot P(AM)$$
$$= 0.98 \times 0.92$$
$$= 0.902$$

Q. 36 of employed and 12 of unemployed are members of rotary club. Find possibility that an individual picked at random is member of rotary club

A : individual is a member of rotary club

$P(E) = 2/3$ $\qquad P(A/E) = \dfrac{n(A \cap E)}{n(E)} = \dfrac{36}{600} = \dfrac{3}{50}$

Two disjoint events

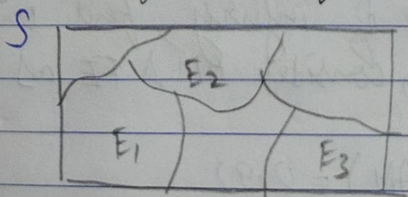


$P(\bar{E}) = \dfrac{1}{3}$ $\qquad P(A/\bar{E}) = \dfrac{n(A \cap \bar{E})}{n(\bar{E})} = \dfrac{12}{300} = \dfrac{1}{25}$

$\left[A = (A \cap E) \cup (A \cap \bar{E})\right]$

$$
\begin{aligned}
P(A) &= P\left[(A \cap E) \cup (A \cap \bar{E})\right] \\
&= P(A \cap E) + P(A \cap \bar{E}) \\
&= P(A/E)\,P(E) + P(A/\bar{E})\,P(\bar{E}) \\
&= \frac{3}{50} \times \frac{2}{3} + \frac{1}{25} \times \frac{1}{3} \\
&= \frac{1}{25} + \frac{1}{75} \\
&= \frac{4}{75}
\end{aligned}
$$

* In case of n no. of disjoint events



$E_1, E_2, E_3 \ldots E_n$ are disjoint events

$E_1 \cup E_2 \cup E_3 \cdots \cup E_n = S$

$$P(A) = P(A/E_1)\,P(E_1) + P(A/E_2)\,P(E_2) + \cdots + P(A/E_n)\,P(E_n)$$

$$\boxed{P(A) = \sum_{i=1}^{n} P(A/E_i) \cdot P(E_i)}$$

Q. In a certain machinary plant, B_1, B_2, B_3 make upto 30%, 45%, 25% of the products. Now the probability that the product turns up defective is 2%, 3%, 2%. Suppose a product is selected at randomn, find probability of it being defective.

A : The product is defective
B_1 : product by B_1
B_2 : " " B_2
B_3 : " " B_3

$P(B_1) = 0.3$, $P(B_2) = 0.45$, $P(B_3) = 0.25$
$P(A/B_1) = 0.02$, $P(A/B_2) = 0.03$, $P(A/B_3) = 0.02$

$$P(A) = P(A/B_1) \times P(B_1) + P(A/B_2) \times P(B_2) + P(A/B_3) \times P(B_3)$$
$$= 0.3 \times 0.02 + 0.45 \times 0.03 + 0.25 \times 0.02$$
$$= 0.0245$$

Baye's Theorem

$E_1, E_2, \ldots E_n$ constitute parts of S such that $P(E_i) \neq 0 \ \forall i$, then for any event A in S such that $P(A) \neq 0$ then probability of E_i given A

$$\boxed{P(E_i/A) = \frac{P(E_i \cap A)}{\sum_{i=1}^{n} P(A/E_i)\, P(E_i)}}$$

Q: If the product is defective, find probability of it being manufactured by B_3

$$P(B_3/A) = \frac{P(B_3 \cap A)}{P(A/B_1)P(B_1) + P(A/B_2)P(B_2) + P(A/B_3)P(B_3)}$$

$P(B_3 \cap A) \rightarrow P(A/B_3) \cdot P(B_3)$

$$= \frac{0.02 \times 0.25}{0.02 \times 0.3 + 0.03 \times 0.45 + 0.02 \times 0.25}$$

$$= \frac{0.0050}{0.0245} = \frac{10}{49}$$

① # To prove

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

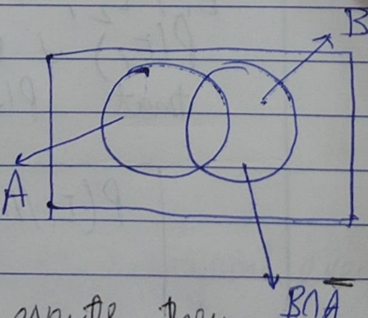
$$P(A \cup B) = P(A \cup (B \cap \bar{A}))$$

∵ Now since $P(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{i=1}^{n} P(A_i)$

$A_1, A_2, \dots A_n$ are disjoint

$$P(A \cup B) = P(A) + P(B \cap \bar{A})$$

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



② To prove $P(\phi) = 0$

$$S \cup \phi = S$$

$$P(S \cup \phi) = P(S) = 1$$

$$P(S) + P(\phi) = 1$$

Thus, $P(\phi) = 0$

③ (a) # * If A and B are independent events then

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

Then A and $\bar{B}$ are also independent events

$$P(A \cap \bar{B}) = P(A) - P(A \cap B) = P(A) - P(A) \cdot P(B)$$

$$= P(A)\,[1 - P(B)]$$

$$= P(A)\,P(\bar{B})$$

Hence Proved!

(3)(b) $\bar{A}$ and $\bar{B}$ are independent events

$$P(\bar{A} \cap \bar{B}) = P(\overline{A \cup B})$$
$$= 1 - P(A \cup B)$$
$$= 1 - [P(A) + P(B) - P(A \cap B)]$$
$$= 1 - [P(A) + P(B) - P(A) \cdot P(B)]$$
$$= 1 - P(A) - P(B)(1 - P(A))$$
$$= (1 - P(A))(1 - P(B))$$
$$= P(\bar{A})\, P(\bar{B})$$

Hence proved!

(4) $P(A \cap B) \leq P(A)$

$A \cap B \leq A$ ($A \cap B$ is part of A)

so $P(A \cap B) \leq P(A)$

Similarly $P(A \cap B) \leq P(B)$

(5) $P(A \cap B) \leq P(A) \leq P(A \cup B)$

$A \leq A \cup B$

so $P(A) \leq P(A \cup B)$

(6) $P(A \cap B) \leq P(A) \leq P(A \cup B) \leq P(A) + P(B)$

using additive rule

$$P(A \cup B) = P(A) + P(B) \underbrace{- P(A \cap B)}_{\text{subtracting a non-negative quantity}}$$

Hence $P(A \cup B)$ is less than $P(A) + P(B)$

Q. If $P(A) = 3/4$ and $P(B) = 5/8$

Then prove

(i) $P(A \cup B) \geq \frac{3}{4}$ (ii) $\frac{3}{8} \leq P(A \cap B) \leq \frac{5}{8}$

$P(A \cup B) \geq P(A)$
$\geq 3/4$

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

$P(A \cap B) \geq \frac{3}{4} + \frac{5}{8} - 1$

$P(A \cap B) \geq \frac{11}{8} - 1$ (as $P(A \cup B) \leq 1$)

$P(A \cap B) \geq \frac{3}{8}$

STATISTICS

Data $\longrightarrow$ (a) discrete (b) continuous

(i) qualitative (ii) quantitative

DESCRIPTIVE STATISTICS

① Mean $= \bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i}$$

Q: find the arithmetic mean of following frequency distribution

x	f
1	5
2	9
3	12
4	17
5	14
6	10
7	6

arithmetic mean

$$\bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i}$$

$\Sigma f_i x_i = 5 + 18 + 36 + 68 + 70 + 60 + 42$

$= \cancel{289}\ 299$

$\Sigma f_i = 73$ so $\bar{x} = \frac{299}{73} = 4.09$

Data as class intervals

Q Ex: Marks

Marks	x	f
0 - 10	5	12
10 - 20	15	18
20 - 30	25	27
30 - 40	35	20
40 - 50	45	17
50 - 60	55	6

$$\Sigma f_i x_i = \frac{2800}{100} = 28$$

* If x_i and f_i are very large values then

* $$\boxed{\bar{x} = A + \frac{\sum f_i d_i}{\sum f_i}}$$

Odd no of obs : middle one A
even no of obs : any of middle 2 can be A

A = assumed mean

$d_i = x_i - A$

Ex-

	x	f	d_i	$f_i d_i$
	20	8	−20	−160
	30	12	−10	−120
any of these A here let A = 40	40	20	0	0
	50	10	10	100
	60	6	20	120
	70	4	30	120

$$\frac{\sum f_i d_i}{\sum f_i} = \frac{60}{60} = 1$$

$$\bar{x} = A + \frac{\sum f_i d_i}{\sum f_i} = 40 + 1 = 41$$

Ex:

$\boxed{A = 25}$

Marks	no of student	x_i	d_i
0−10	5	5	−20
10−20	10	15	−10
20−30	25	25	0
30−40	30	35	10
40−50	~~45~~ 20	45	20
50−60	~~55~~ 10	55	30

$f_i d_i$
-100
-100
0
300
400
300

$\Sigma f_i d_i = 800$

$\bar{x} = 25 + \frac{800}{100} = 33$

★ $\boxed{\bar{x} = A + \frac{\Sigma f_i d_i}{N} \times h}$ $\left(\text{here } d_i = \frac{x_i - A}{h} \quad h = \text{class size} \right)$

Ex:

∗ C = common factor

$\boxed{C = 5}$ here

Marks	x	f	$x_i - A$	$d_i = \frac{x_i - A}{5}$
0-10	5	5	-40	-8
10-30	20	12	-25	-5
30-60	45	25	0	0
60-100	80	8	35	7

$[A = 45]$

$f_i d_i$
-40
-60
0
56

$\Sigma f_i d_i = -44$

$\bar{x} = 45 + \left(\frac{-44}{50} \times 5 \right)$

$= 45 - 4.4$

$= 40.6$

② Median:

Step 1: Arrange the data in either ascending / descending order

Step 2: ➢ If no of obs is <u>odd</u> then $x_{(n+1)/2}^{th}$ element gives median

➢ if no of obs is <u>even</u> then $\frac{1}{2}\left[x_{n/2} + x_{n/2+1}\right]^{th}$ term gives median

Ex: 141, 143, 132, 161, 122, 155, 137

Ans Ascending : ~~132~~ 122, 132, 137, 141, 143, 155, 161

7 observations so median $= \left(\frac{7+1}{2}\right)^{th}$ term

$= 4^{th}$ term $= 141$

Discrete data

① Arrange in ascending / descending order
② find out cumulative frequency
③ find the c.f just ~~above~~ greater than N/2, the corresponding of x is median

Income Ex:	15000	15500	16800	18000	18500	17800
people =	24	26	16	20	6	30

Income	People	cf
15000	24	24
15500	26	50
16800	16	66
17800	30	~~86~~ 96
18000	20	116
18500	6	122

$\Sigma f = 122 = N$

$\frac{N}{2} = \frac{122}{2} = 61$, cf just greater is 66

so median is 16800

Continuous data

$$\boxed{\text{Median} = l + \frac{h}{f}\left[\frac{N}{2} - c.f\right]}$$

choose median class $cf \geq \frac{N}{2}$

l = lower limit

h = class-size

f = corresponding frequency

N = Total frequency

$c.f$ = c.f of prev. to med class

marks	students	c.f
5-10	7	7
10-15	15	22
15-20	24	46
20-25	31	77
25-30	42	~~109~~ 119
30-35	30	149
35-40	26	175
40-45	15	190
45-50	10	200
	200	

$N = 200$

$N/2 = 100$

c.f greater = 119

so median class = 25-30

$= 25 + \frac{5}{42}\left[100 - 77\right]$

$= \cancel{25}\ 27.738$

Mode

Ex:

Marks	Students (I) f	f (II)	f (III)	f (IV)
10	8	20		
15	12		48	56
20	36	71		
25	35		63	81
30	28	46		
35	18		27	
40	9			

f (V)	f (VI)
83	99
55	

choose the highest no of out all f(i)

36	20
71	20, 25
63	25, 30
81	25, 30, 35
83	15, 20, 25
99	20, 25, 30

20	25	30	35	~~ ~~
4	5	3	1	~~ ~~

Since 25 occurs 5 times here
so mode = 25

Mode - Continuous series

$$\text{Mode} = L + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times h$$

h = class size
f_1 = freq of modal class
f_0 = freq prev modal
f_2 = freq next modal
L = Lower limit of modal class

Modal class : The class with highest frequency is modal class

Ex:

Marks	f
0-10	3
10-20	5
20-30	7
30-40	10
40-50	12 (f_0)
50-60	15 (f_1)
60-70	12 (f_2)
70-80	6
80-90	2
90-100	8

$$\text{Mode} = 50 + \frac{15-12}{30-12-12} \times 10$$

$$= 50 + \frac{1}{2} \times 10$$

$$= 55$$

Another way, with median & mean

$$\boxed{\text{Mode} = 3\,\text{Median} - 2\,\text{Mean}}$$

Q. The algebraic sum of all deviations of a set of values from their arithmetic mean is (zero) 0

$$\sum_{i=1}^{n} f_i(x_i - \bar{x}) = 0 \text{ (To prove)}$$

$(d_i = x_i - \bar{x})$

$$\sum_{i=1}^{n} f_i(x_i - \bar{x}) = \sum_{i=1}^{n} f_i x_i - \sum_{i=1}^{n} f_i \bar{x}$$

constant so comes out summation

$$= \underbrace{\sum_{i=1}^{n} f_i x_i} - \bar{x} \sum_{i=1}^{n} f_i$$

$$\left(\bar{x} = \frac{\sum f_i x_i}{\sum f_i} \quad \sum f_i x_i = \bar{x} \sum f_i \right)$$

$$\Rightarrow \bar{x} \sum f_i - \bar{x} \sum f_i$$

$$\Rightarrow 0$$

Q. The sum of squares of deviations of a set of values is minimum when taken about mean

To prove $\sum d_i = \sum f_i (x_i - A)^2$ when $A = \bar{x}$

so now $d_i = f_i (x_i - A)^2$

$$X \Rightarrow \sum_{i=1}^{n} d_i = \sum_{i=1}^{n} f_i (x_i - A)^2$$

differentiate $\dfrac{dX}{dA} = -2 \sum_{i=1}^{n} f_i (x_i - A) = 0$

$$A = \frac{\sum f_i x_i}{\sum f_i} = \bar{x}$$

$$\frac{d^2 X}{dA^2} = \cancel{2A} = 2 \sum f_i$$

(positive value so minima at $A = \bar{x}$ proved)

Measures of Dispersion

① Range: Maximum value − Minimum value

② Quartile Deviation : Q_1, Q_2, Q_3 are quartiles which divide any 100 points in $\boxed{4}$ equal parts

0 — Q_1 — Q_2 — Q_3 — 100

$\boxed{Q_2 \text{ Median}}$

$Q_1 \rightarrow$ lower quartile
$Q_3 \rightarrow$ upper quartile

Inter quartile range $= Q_3 - Q_1$

Quartile deviation $= \dfrac{Q_3 - Q_1}{2}$

How to find Q_1 & Q_3?

$$\text{Median} = L + \frac{h}{f}\left[\frac{N}{2} - c.f\right] \quad \text{check } c.f \text{ greater than } \frac{N}{2}$$

$$Q_1 = L + \frac{h}{f}\left[\frac{N}{4} - c.f\right] \quad \text{〃 〃 〃 〃 } \frac{N}{4}$$

$$Q_3 = L + \frac{h}{f}\left[\frac{3N}{4} - c.f\right] \quad \text{〃 〃 〃 〃 } \frac{3N}{4}$$

Q. Find the median and quartile for the data

x:	0	1	2	3	4	5	6	7	8
f:	1	9	26	59	72	52	29	7	1
$c.f$	1	10	36	95	167	219	248	255	256

$\dfrac{N}{2} = \dfrac{256}{2} = 128$, $cf >$ than 128 is 167

median for discreate is $cf > \dfrac{N}{2}$ class

so median $= 4$

$Q_1 = \frac{N}{4} = \frac{256}{4} = 64$

$Q_1 = 3$

$Q_2 = \frac{3N}{4} = \frac{3 \times 256}{4} = 192$

$Q_3 = 5$

Quartile deviation $= \frac{5-3}{2} = \underline{1}$

Q: To find median and quartiles

Marks	No. of Student	Cf
0-10	15	15
10-20	20	35
20-30	25	60
30-40	~~60~~ 24	84
40-50	10	94
50-60	33	127
60-70	71	198
70-80	51	249

$\frac{N}{2} = \frac{249}{2} = \cancel{124.5}\ 124.5 = \cancel{125}$

Median class 50-60 $\Rightarrow$ $\cancel{L}\ 50 + \frac{10}{33}\left[124.5 - 94\right]$

$= 50 + \frac{10}{33} \times 30.5$

$= \cancel{59.3939}$

$= 59.24$

$Q_1 \rightarrow \frac{N}{4} = \frac{249}{4} = 62.25$

$C.f = 84$

median $\cancel{Q}$ $= 30 + \frac{10}{24}\left[62.25 - 60\right]$

$= 30.9375$

$Q_3 \rightarrow \frac{3N}{4} = 186.75$

$C.f = 198$

$\Rightarrow 60 + \frac{10}{71}\left[186.75 - 127\right]$

$\Rightarrow 68.41$

(3) Mean Deviation

$$\boxed{M.D = \frac{1}{N} \sum f_i |x_i - A|}$$

↳ A can be mean, median, mode

By default, if nothing is specified then about median

Ex: Calculate mean deviation

(1) → mean deviation about median

x	f	C_f	x_i - median)
10	3	3	-2
11	12	15	-1
12	18	33	0
13	12	45	1
14	3	48	2
	48		

$\frac{N}{2} = \frac{48}{2} = 24$

$C.f \geq$ is 33

so median $= 12$

① Sum of all deviations is zero

$\lvert x_i - \text{median} \rvert$	$f_i \lvert x_i - 12 \rvert$
2	6
1	12
0	0
1	12
2	6
	36

$\Rightarrow$

Mean deviation $= \dfrac{\sum f_i \lvert x_i - \text{median} \rvert}{\sum f_i} = \dfrac{36}{48}$

② $\rightarrow$ Mean deviation about mean

Q. Calculate mean deviation from mean

x	f	$f_i x_i$	$\lvert x_i - \bar{x} \rvert$
2	2	4	6
4	2	8	4
6	4	24	2
8	5	40	0
10	3	30	2
12	2	24	4
14	1	14	6
16	1	16	8
	20	160	

$$\frac{\sum f_i x_i}{\sum f_i} = \frac{160}{20} = 8$$

$$\bar{x} = 8$$

$f_i \lvert x_i - \bar{x} \rvert$
12
8
8
0
6
8
6
8
56

$$\frac{\sum f_i \lvert x_i - \bar{x} \rvert}{\sum f_i} = \frac{56}{20} = 2.8$$

standard Deviation (Root mean square deviation)

☆ $$\boxed{S = \sqrt{\frac{1}{N}\sum_{i=1}^{n} f_i (x_i - A)^2} = \sigma}$$

$d_i = x_i - A$

If $d_i = x_i - \bar{x}$ then just

① $$S = \sqrt{\frac{1}{N}\sum_{i=1}^{n} f_i (x_i - \bar{x})^2} = \sigma$$

☆ $$\boxed{\sigma^2 = \text{variance} = \frac{1}{N}\sum_{i=1}^{n} f_i (x_i - \bar{x})^2}$$

$$= \frac{1}{N}\sum_{i=1}^{n} f_i (x_i^2 + \bar{x}^2 - 2x_i\bar{x})$$

$$= \frac{1}{N}\sum f_i x_i^2 + \frac{\bar{x}^2}{N}\underbrace{\sum f_i}_{N} - \underbrace{\frac{2\bar{x}}{N}\sum_{i=1}^{n} x_i f_i}_{\bar{x}}$$

$$\sigma^2 = \frac{1}{N}\sum f_i x_i^2 - \bar{x}^2$$

② $$\boxed{\sigma = \sqrt{\frac{\sum f_i x_i^2}{N} - \bar{x}^2}}$$

☆ Change of origin

$\Rightarrow$ $d_i = x_i - A$ (multiply f_i on both sides)

$\Rightarrow$ $f_i d_i = f_i x_i - A f_i$ (Summation & divide N)

$\Rightarrow$ $\frac{\sum f_i d_i}{N} = \frac{\sum f_i x_i}{N} - \frac{A\sum f_i}{N}$

$= \bar{d} = \bar{x} - A$

$\Rightarrow \quad \bar{d} = \bar{x} - A$

$\Rightarrow \quad d_i - \bar{d} = d_i - \bar{x} + A \qquad (d_i = x_i - A)$

$\Rightarrow \quad d_i - \bar{d} = x_i - \bar{x}$

so $\sigma_x = \sqrt{\frac{1}{N} \sum f_i (x_i - \bar{x})^2}$

(3) $\boxed{\sigma_d = \sqrt{\frac{1}{N} \sum f_i (d_i - \bar{d})^2}}$ (replace each $x_i \to d_i$, $\bar{x} \to \bar{d}$)

$\sigma_x = \sigma_d = \sqrt{\frac{1}{N} \sum f_i d_i^2 - \left(\frac{\sum f_i d_i}{N}\right)^2}$ (4)

★ change of origin and scale

$\Rightarrow \quad d_i = \frac{x_i - A}{h}$

$\Rightarrow \quad h d_i = x_i - A$

$\Rightarrow \quad h \bar{d} = \bar{x} - A$

$\Rightarrow \quad h(d_i - \bar{d}) = x_i - \bar{x}$

(5) $\sigma_x = \sqrt{\frac{1}{N} \sum_{i=1}^{n} f_i (h(d_i - \bar{d}))^2}$

$= h \sqrt{\frac{1}{N} \sum_{i=1}^{n} f_i (d_i - \bar{d})^2}$

(6) $\boxed{\sigma_x = h \sigma_d} = h \sqrt{\frac{1}{N} \sum f_i d_i^2 - \left(\frac{\sum f_i d_i}{N}\right)^2}$

Relation b/w S and σ

$$S = \sqrt{\frac{1}{N}\sum f_i (x_i - A)^2}$$

$$\sigma = \sqrt{\frac{1}{N}\sum f_i (x_i - \bar{x})^2}$$

$$S^2 = \frac{1}{N}\sum f_i (x_i - \bar{x} + \underbrace{\bar{x} - A}_{\bar{d}})^2$$

$$S^2 = \frac{1}{N}\sum f_i (x_i - \bar{x} + \bar{d})^2$$

$$S^2 = \frac{1}{N}\sum f_i \left((x_i - \bar{x})^2 + \bar{d}^2 + 2(x_i - \bar{x})\bar{d}\right)$$

$$S^2 = \sigma^2 + \frac{\bar{d}^2}{N}\sum f_i + \underbrace{\frac{2\bar{d}}{N}\sum f_i (x_i - \bar{x})}_{\text{sum of all mean dev} = 0}$$

$$\boxed{S^2 = \sigma^2 + \bar{d}^2}$$

* Hence standard deviation is least possible root mean square deviation

Ex: find standard dev

$$\Rightarrow \bar{x} = \frac{70}{10} = 7$$

x_i	$d_i = x_i - \bar{x}$	$(d_i)^2$
5	-2	4
8	1	1
7	0	0
11	4	16
9	2	4
10	3	9
8	1	1
2	-5	25
4	-3	9
6	[illegible]	[illegible]

$\Sigma(d_i)^2 = \frac{70}{10} = 7$

$\sigma = \sqrt{7} = 2.65$

Q:

x_i	f_i	$d_i = x_i - A$ (where $A = 55$)	$d_i = \frac{x_i - A}{h}$ ($h=10$)	$f_i(d_i)^2$	$f_i d_i$
25	3	-30	-3	27	-9
35	61	-20	-2	244	-122
45	132	-10	-1	132	-132
55	153	0	0	0	0
65	140	10	1	140	140
75	51	20	2	204	102
85	2	30	3	18	6
	542			765	-15

$\bar{x}_1 = \quad \sigma = h\sqrt{\frac{1}{N}\Sigma f_i(d_i - \bar{d})^2}$

$\sigma = 10 \times \sqrt{\frac{765}{542} - \frac{225}{(542)^2}}$

$\sigma = 11.8$

Ex: find SD and coeff of variation

Marks	no of students	x_i	$d_i = \frac{x_i - 45}{10}$	$(d_i)^2$	$f_i d_i$
0-10	5	5	-4	16	-20
10-20	10	15	-3	9	~~-150~~ -30
20-30	20	25	-2	4	-40
30-40	40	35	-1	1	-40
40-50	30	45	0	0	0
50-60	20	55	1	1	20
60-70	10	65	2	4	20
70-80	~~40~~ 4	75	3	9	12
	139				-78

$f_i(d_i)^2$

80

90

80

40

0

20

40

36

386

$$\sigma = 10 \times \sqrt{\frac{386}{139} - \left(\frac{-78}{139}\right)^2}$$

$$\sigma = 15.69 \approx 15.7$$

coffiicient of variation $= \frac{\sigma}{\bar{x}} \times 100$

$=$

$$\bar{x} = A + \frac{\sum f_i d_i}{N} \times h$$

$$\bar{x} = 45 - \frac{78}{139} \times 10$$

$$\bar{x} = 39.4$$

$$\Rightarrow \frac{15.7}{39.4} \times 100 = 39.84$$

Moments

(about A) $\mu_r' = \frac{1}{N} \sum_i f_i (x_i - A)^r$

(about mean) $\mu_r = \frac{1}{N} \sum_i f_i (x_i - \bar{x})^r$

$r=1$ mean deviation

$r=2$ variance

$r=3$ skewness

$r=4$ kurtosis

NASA

☆ for $\boxed{r=0}$ $\mu_0' = \frac{1}{N}\Sigma fi = \frac{N}{N} = 1$

☆ for $\boxed{r=1}$ $\mu_1' = \frac{1}{N}\Sigma fi(x_i - \bar{x}) = 0$ (sum of all mean devi is 0)

<u>Relation b/w μ_r & μ_r'</u>

$\mu_1' = \frac{1}{N}\Sigma fi di$ $\qquad \bar{\mu} = A + \frac{\Sigma fidi}{N}$ μ_1'

$$\Rightarrow \left[\mu_1' = \bar{\mu} - A\right]$$

Now, $\mu_r = \frac{1}{N}\Sigma fi(x_i - \bar{x})^r$

$$= \frac{1}{N}\Sigma fi(\underbrace{x_i - A}_{di} + \underbrace{A - \bar{x}}_{-\mu_1'})^r$$

$$\mu_r = \frac{1}{N}\Sigma fi(di - \mu_1')^r$$

$$\mu_r = \frac{1}{N}\Sigma fi(di^r - {}^rC_1 di^{r-1}\mu_1' + {}^rC_2 di^{r-2}(\mu_1')^2 + \cdots (-1)^r(\mu_1')^r)$$

☆ $$\left[\mu_r = \frac{1}{N}\Sigma fi di^r - \frac{1}{N} r \overbrace{\Sigma fi di^{r-1}}^{\mu_{r-1}'} \mu_1' + \frac{{}^rC_2}{N}\Sigma fi di^{r-2}(\mu_1')^2 + \frac{(-1)^r}{N}\Sigma fi (\mu_1')^r\right]$$

☆ for $\boxed{r=2}$ $\mu_2 = \overbrace{\frac{1}{N}\Sigma fi di^2}^{\mu_2'} - \overbrace{\frac{2}{N}\Sigma fi di}^{\mu_1'}\mu_1' + \frac{1}{N}\overbrace{\Sigma fi}^{N}(\mu_1')^2$

$$\mu_2 = \mu_2' - 2\mu_1'\mu_1' + \mu_1'^2$$

$$\boxed{\mu_2 = \mu_2' - (\mu_1')^2}$$

* for $\boxed{r=3}$: $\cancel{\Sigma f_i}$ $\frac{\Sigma f_i d_i^3}{N} - \frac{3}{N}\Sigma f_i d_i^2 \mu_1' + \frac{3}{N}\Sigma f_i d_i (\mu_1')^2$

$- \frac{\Sigma f_i (\mu_1')^3}{N}$

$$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 3(\mu_1')^3 - (\mu_1')^3$$
$$= \mu_3' - 3\mu_2'\mu_1' + 2(\mu_1')^3$$

* for $\boxed{r=4}$: $\mu_4 = \frac{\Sigma f_i d_i^4}{N} - \frac{4}{N}\Sigma f_i d_i^3 \mu_1' + \frac{\cancel{4}6}{N}\Sigma f_i d_i^2 (\mu_1')^2$

$- \frac{4}{N}\Sigma f_i d_i (\mu_1')^3 + \frac{1}{N}\Sigma f_i (\mu_1')^4$

$$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'(\mu_1')^2 - 4(\mu_1')^4 + (\mu_1')^4$$
$$= \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'(\mu_1')^2 - 3(\mu_1')^3$$

Q. Calculate the first 4 moments about mean

(A=4)

x	f_i	$d_i = x_i - A$	$(d_i)^2$	$(d_i)^3$	$(d_i)^4$
0	1	-4	16	-64	256
1	8	-3	9	-27	81
2	28	-2	4	-8	16
3	56	-1	1	-1	1
4	70	0	0	0	0
5	56	1	1	1	1
6	28	2	4	8	16
7	8	3	9	27	81
8	1	4	16	64	256
	256				

$f_i d_i$	$f_i(d_i)^2$	$f_i(d_i)^3$	$f_i(d_i)^4$
−4	16	−64	256
−24	72	−216	648
−56	112	−224	448
−56	56	−56	56
0	0	0	0
56	56	56	56
56	112	224	448
24	72	216	648
4	16	64	256
0	512	0	2816

$$\mu_1' = \frac{\Sigma f_i d_i}{N} = 0$$

$$\mu_2' = \frac{\Sigma f_i d_i^2}{N} = \frac{512}{256} = 2$$

$$\mu_3' = \frac{\Sigma f_i d_i^3}{N} = 0$$

$$\mu_4' = \frac{\Sigma f_i d_i^4}{N} = \frac{2816}{256} = 11$$

$$\bar{x} = A + \frac{\Sigma f_i d_i}{N}$$

$\bar{x} = 4$ so moments about mean would be same as about A

so $\mu_r' = \mu_r$

Q: Calculate first 4 moments about mean

$A = 45$
$h = 10$

Wage	No. of persons	x_i	$d_i = \frac{x_i - 45}{10}$
0-10	15	5	-4
10-20	23	15	-3
20-30	35	25	-2
30-40	49	35	-1
40-50	32	45	0
50-60	28	55	1
60-70	12	65	2
70-80	6	75	3
	200		

$f_i d_i$	$f_i (d_i)^2$	$f_i (d_i)^3$	$f_i (d_i)^4$
-60	240	-960	3840
-63	207	-621	1863
-70	140	-280	560
-49	49	-49	49
0	0	0	0
28	28	28	28
24	48	96	192
18	54	162	486
-178	766	-1624	7018

$$\mu_1' = \frac{\sum f_i d_i}{N} = \frac{-178}{200} = -0.89$$

$$\mu_2' = \frac{\sum f_i d_i^2}{N} = \frac{766}{200} = +3.85$$

$$\mu_3' = \frac{\sum f_i d_i^3}{N} = \frac{-1624}{200} = -8.12$$

$$\mu_4' = \frac{\sum f_i d_i^4}{N} = \frac{7018}{200} = 35.09$$

$\mu_1 = 0$

$\mu_2 = \mu_2' - (\mu_1')^2 = \cancel{8}.$

$= 3.83 - (0.89)^2$

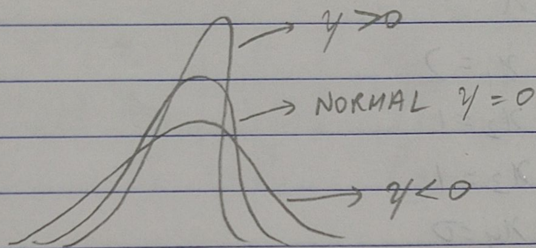
$= 2.04$

$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3$

$= 0.7$

$\mu_4 = \cancel{\mu}\ 22.5$

Kurtosis



$\beta_2 = \dfrac{\mu_4}{\mu_2^2}$

$\gamma_2 = \beta_2 - 3$

for curve which is normal curve, $\gamma = 0$
curve less than NC, $\gamma < 0$
curve greater than NC, $\gamma > 0$

Random variable

3 items, ~~ND~~ N → non def, D → def

$S = \{NNN, DDD, NND, NDN, NDD, DND, DDN, DNN\}$

X = random variable

SS	X
NNN	0
DDD	3
NDD	2

→ It's the function ~~whose domain~~ that associates a real number with each element in sample space

Q. Two balls are drawn in succession w/o replacement from a basket containing 4 red and 3 Black Balls. The possible outcomes & value of random variable X where X = no. of red balls are?

X = no. of red balls

SS	X
RR	$x_1 = 2$
BR	$x_2 = 1$
RB	$x_3 = 1$
BB	$x_4 = 0$

Q. 3 elements are returned to Neha, Smita, Anjali, who had prev. made them. Then possible outcomes of random variable where X = correct helmet to correct person

SS	X
correct ⟵ NSA	3
NAS	1
SNA	1 ⟶ only one match is correct
SAN	0 ⟶ no match is correct
ANS	0
ASN	1

★ Discrete random variable (DRV)
represents count data such as no. of defective items in n items, fatalities on road

* continuous random variable EC (CRV)
represents measured data like height, weight and temperature distance or life periods

* Discrete probability distribution

for Neha, Smita, Anjali case

X	0	1	3
P(X)	1/3	1/2	1/6
	P_1	P_2	P_3

$P(X=x) = f(x)$ → probability mass function
→ probability f^n
→ prob. distribution

* The set of ordered pair $(x, f(x))$ is a prob. mass f^n or in short (PMF) or prob f^n of a discrete random variable (RV) if

(i) ▷ for each possible outcome x, $f(x) \geq 0$

(ii) ▷ $\sum_x f(x) = 1$ $(1/3 + 1/2 + 1/6 = 1)$

(iii) ▷ $P(X=x) = f(x)$

Q. A shipment of 20 similar laptops to a retail outlet contains 3 defective laptops. If a school makes a random purchase of 2 laptops. find the prob. distribution for no of defectives

X = no of defective laptops

(P.T.O)

X	0	1	2
P(X)	$\frac{^3C_0\ ^{17}C_2}{^{20}C_2} = \frac{68}{95}$	$\frac{^3C_1\ ^{17}C_1}{^{20}C_2} = \frac{51}{190}$	$\frac{^3C_2\ ^{17}C_0}{^{20}C_2} = \frac{3}{190}$

Sum of $P_1 + P_2 + P_3 = \frac{136 + 51 + 3}{190} = 1$

Cumulative distribution f^n

$$F(x) = P(X \le x)$$
$$= \sum_{t \le x} f(t)$$

NOTE
$F(x) = CDF$
$f(x) = PMF$

CDF of a discrete random variable with PMF $f(x)$ is sum of all PMF less than equal to any constant given to it

$$F(2) = f(0) + f(1) + f(2)$$
$$= \frac{68}{95} + \frac{51}{190} + \frac{3}{190}$$
$$= 1$$

$$F(1) = f(0) + f(1)$$
$$= \frac{68}{95} + \frac{51}{190}$$
$$= \frac{187}{190}$$

$$F(0) = f(0)$$
$$= \frac{68}{95}$$

~~PROPE~~

* PROPERTIES OF CDF

i) $F(x)$ is a non-decreasing f^n of x i.e. if $x_1 < x_2$ then $F(x_1) \leq F(x_2)$

ii) $F(-\infty) = 0$ and $F(\infty) = 1$

iii) $P(X = x_i)$ or $f(x_i) = F(x_i) - F(x_{i-1})$

Again take Neha, Anjali, Smita case

x	0	1	3
$P(x)$	$1/3$	$1/2$	$1/6$

$$CDF = \begin{cases} 0 & x < 0 \\ 1/3 & 0 \leq x < 1 \\ 5/6 & 1 \leq x < 2 \\ 5/6 & 2 \leq x < 3 \\ 1 & 3 \leq x < \infty \end{cases}$$

Q: If a random variable X takes values 1 to 4 such that ~~2 times P(X)~~

$$2P(X=1) = 3P(X=2) = P(X=3) = 5(P(X=4))$$

find prob distribution (PMF) and CDF

let $2P(X=1) = 3P(X=2) = P(X=3) = 5P(X=4)) = k$

so since PMF sum is 1 then

$$\frac{k}{2} + \frac{k}{3} + k + \frac{k}{5} = 1$$

$$\Rightarrow \frac{15k + 10k + 30k + 6k}{30} = 1$$

$$\Rightarrow k = 30/61$$

$P(X=1) = \frac{15}{61}$

$P(x=2) = \frac{10}{61}$

$P(X=3) = \frac{30}{61}$

$P(X=4) = \cancel{6} \frac{6}{61}$

X	1	2	3
$f(x)$	15/61	10/61	30/61
$F(x)$	15/61	25/61	55/61

x vs PMF

x vs CDF

Q: A random variable X has a following probability function

X	0	1	2	3	4	5	6	7	8
$P(X=x)$	a	3a	5a	7a	9a	11a	13a	15a	17a

i) find the value of a

ii) find $P(X<3)$, $P(0<X<3)$, $P(X\geq 3)$

iii) Find the cdf ~~cont.~~ (cumulative distri f^n) of x

i) $\sum_{x \in x} f(x) = 1 \Rightarrow 81a = 1$

$a = \frac{1}{81}$

ii) $P(X<3) = P(0) + P(1) + P(2)$

$= 9a = \frac{9}{81} = \frac{1}{9}$

$P(0<X<3) = P(1) + P(2)$

$= \frac{8}{81}$

$P(X\geq 3) = 1 - P(X<3)$

$= \frac{8}{9}$

(iii)

X	0	1	2	3	4	5	6	7	8
cdf	a	$4a$	$9a$	$16a$	$25a$	$36a$	$49a$	$64a$	$81a$
	$\frac{1}{81}$	$\frac{4}{81}$	$\frac{1}{9}$	$\frac{16}{81}$	$\frac{25}{81}$	$\frac{36}{81}$	$\frac{49}{81}$	$\frac{64}{81}$	1

Continuous distribution f^n

If X is a random variable defines on continuous scale then $P(X=x)=0$

$$P(a \le x < b) = P(a<x<b) + P(x=a)$$
$$= P(a<x<b)$$

Probability density $f^n \longrightarrow f(x)$
(Pdf)

The function $f(x)$ is a pdf for a continuous random variable X defined on a set of real no.

If

a) $f(x) \ge 0$, $x \in R$

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

c) $P(a<X<b) = \int_a^b f(x)dx$

Q. Suppose that an error in the reaction temp. for a controlled lab experiment is a continuous random variable X having pdf

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

(i) verify that $f(x)$ is a density f^n

ii) determine $P(0<X\le 1)$

(i) $f(x) = \frac{x^2}{3} \ge 0 \; (x \in (-1,2))$

(ii) $\int_{-\infty}^{\infty} f(x)dx = \int_{-1}^{2} \frac{x^2}{3}dx = \left[\frac{x^3}{9}\right]_{-1}^{2} = \frac{8}{9} + \frac{1}{9} = 1$

ii) $P(0<X\le 1) = \int_0^1 \frac{x^2}{3}dx = \left[\frac{x^3}{9}\right]_0^1$
$= \frac{1}{9}$

Cumulative distribution function

$F(x)$ of a continuous random variable X with density function $f(x)$

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)\,dx \quad \text{or} \quad \int_{-\infty}^{x} f(t)\,dt$$

$$P(a < X < b) = F(b) - F(a)$$

__Example__ Find $F(x)$ and use it to evaluate $P(0 < X < 1)$

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

$$F(x) = \begin{cases} 0 & x < -1 \\ \dfrac{x^3+1}{9} & -1 < x < 2 \\ 1 & x > 2 \end{cases}$$

$$F(x) = P(-1 < X < 2) = \int_{-\infty}^{x} f(x)\,dx = \int_{-\infty}^{x} \frac{x^2}{3}\,dx$$

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

$$= 0 + \int_{-1}^{x} \frac{x^2}{3}\,dx$$

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

so $P(0<X<1) = F(1) - F(0)$

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

$= \frac{1}{9}$

Q. $f(x) = 6x(1-x) \quad 0 \le x \le 1$

check whether it is pdf or not

$f(x) \ge 0$ in the given interval (non-negative)

$$\int_{-\infty}^{\infty} f(x)dx = \int_0^1 6(x(1-x))\,dx$$

$$= \int_0^1 (6x - 6x^2)\,dx$$

$$= \left[3x^2 - 2x^3\right]_0^1$$

$$= 1$$

Q. Find value of K for PDF : $f(x) \begin{cases} Kx^2, & 0 \le x \le 3 \\ 0, & \text{otherwise} \end{cases}$

Also compute $P(1 \le X \le 2)$

and distribution f^n

(1) $\int_0^3 Kx^2 dx = 1$

$\left[\frac{Kx^3}{3}\right]_0^3 = 1$

$\Rightarrow 9K = 1$

$\Rightarrow K = 1/9$

(2) $P(1 \le X \le 2) = \frac{1}{9}\int_1^2 x^2 dx = \frac{7}{27}$

(3) $F(x) \begin{cases} 0 & x<0 \\ x^3/27 & 0 \le x \le 3 \\ 1 & x > 3 \end{cases}$

[P.T.O]

$$\Rightarrow \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{3x} \cancel{x^3}\, kx^2$$

$$\Rightarrow 0 + \int_{0}^{x} \frac{x^2}{9}$$

$$\Rightarrow \left[\frac{x^3}{27}\right]_0^x \Rightarrow \frac{x^3}{27}$$

Q. If the density f^n of a continuous random variable is given by

$$f(x) = \begin{cases} ax & 0 \le x \le 1 \\ a & 1 \le x \le 2 \\ 3a - ax & 2 \le x \le 3 \\ 0 & \text{elsewhere} \end{cases}$$

find value of a

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

$$\int_{0}^{1} ax\,dx + a\int_{1}^{2} dx + a\int_{2}^{3} (3-x)\,dx = 1$$

$$\left[\frac{ax^2}{2}\right]_0^1 + \left[ax\right]_1^2 + a\left(3x - \frac{x^2}{2}\right)_2^3 = 1$$

$$\Rightarrow \frac{a}{2} + a + \frac{9a}{2} - 4a = 1$$

$$\Rightarrow 2a = 1$$

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

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

$$P(0 \le x \le 1) = \int_{-\infty}^{x} f(x)\,dx = \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{x} \frac{x}{2}\,dx$$

$$= 0 + \frac{x^2}{4}$$

$$P(1 \le x \le 2) = \int_{-\infty}^{x} f(x)\,dx = \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{1} f(x)\,dx + \int_{1}^{x} f(x)\,dx$$

$$= 0 + \frac{1}{4} + \int_{1}^{x} \frac{1}{2}\,dx$$

$$= \frac{2x-1}{4}$$

$$P(2 \le x \le 3) = \int_{-\infty}^{x} f(x)\,dx = \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{1} f(x)\,dx + \int_{1}^{2} f(x)\,dx + \int_{2}^{x} f(x)\,dx$$

$$= \cancel{0 + \frac{1}{4} + \frac{3}{4}}$$

$$= \frac{3}{4} + \int_{2}^{x} \left(\frac{3}{2} - \frac{x}{2}\right) dx$$

$$= \frac{3}{4} + \left(\frac{3}{2}x - \frac{x^2}{4}\right)_{2}^{x}$$

$$= \frac{3}{4} + \left(\frac{3x}{2} - \frac{x^2}{4} - \frac{6}{2} + \frac{4}{4}\right)$$

$$= \frac{3}{4} + \left(\frac{3x}{2} - \frac{x^2}{4} - 2\right)$$

$$\Rightarrow -\frac{5}{4} + \frac{3x}{2} - \frac{x^2}{4} - \frac{5}{4}$$

$f(x,y) \rightarrow$ unless mentioned otherwise means x & y are ~~dependent on each other~~

Joint Probability Distribution

$X, Y \quad , \quad X = x, \; Y = y$

for discrete $P(X=x) = f(x)$
for joint (Both simultaneously) $= P(X=x, Y=y) = f(x,y) \; \forall \; x,y$

→ The value $f(x,y)$ gives the probability that outcome x and y occur at the same time

Q. If an 18-wheeler vehicle has to have it's tyres serviced and X represent the no. of kilometres these tyres have driven and Y is for no. of tyres that need to be replaced then $f(30000, 5)$

X: no. of kms
Y: no. of tyres to be replaced.

Definition: $f(x,y)$ is a joint probability distribution or PMF of discrete random variable X & Y such that

i) $f(x,y) \geq 0 \quad \forall \; x,y$

ii) $\sum_x \sum_y f(x,y) = 1$

iii) $P(X=x, Y=y) = f(x,y)$

iv) $P[(X,Y) \mid (X,Y) \in A]$

Q. 2 ball point ~~pens~~ pens are selected at random from a box that contains 3 blue pens, 2 red pens and 3 green pens.

X: no. of blue pens selected
Y: no. of red pens selected

i) Then find the joint probability distributions $f(x,y)$
ii) $P[(X,Y) \in A]$ where $A: \{(x,y) \mid x+y \leq 1\}$

$(x,y) = \{(2,0), (0,2), (1,0), (0,1), (1,1), (0,0)\}$

(i) $P(X=2, Y=0) = \dfrac{3C_2}{8C_2} = \dfrac{3}{28}$

$y \backslash x$	0	1	2	ROW SUM
0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$15/28$
1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{\cancel{8}\,6}{14}$
2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
COLSUM	$5/14$	$15/28$	$3/28$	

$$P(X=x, Y=y) = \frac{3c_x \, 2c_y \, 3C_{2-(x+y)}}{8C_2}$$

(ii) $P[(X,Y) \mid X+Y \leq 1] = f(0,0) + f(1,0) + f(0,1)$

$= \dfrac{3}{28} + \dfrac{9}{28} + \dfrac{3}{14}$

$= \dfrac{18}{28}$

$= \dfrac{9}{14}$

Continuous random variable Joint function

If X & Y are continuous, then $f(x,y)$ has some properties

(P.T.O)

DEFINITION

$f(x,y)$ is a joint density f^n of continuous random va[riables]
X & Y if

① $f(x,y) \geq 0 \ \forall \ (x,y) \in A$

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

③ $$P[(x,y) \in A] = \iint_A f(x,y)\,dx\,dy$$

Ex: A business operates both driving (drive-in) and walking (walk-in) facility

X: proportion of time for which drive-in used

Y: proportion of time for which walk-in used

suppose that joint ~~distri~~ density f^n of X & Y is

$$f(x,y) \begin{cases} \frac{2}{5}(2x+3y) & 0 \leq x \leq 1,\ 0 \leq y \leq 1 \\ 0 & \text{O.W} \end{cases}$$

(i) verify condition ②

(ii) $P[(X,Y) \in A]$ where $A = \left\{(x,y) \mid 0 \leq x \leq \ldots,\ \frac{1}{4} < y \leq \ldots \right.$

$$\Rightarrow \frac{2}{5}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(2x+3y)\,dx\,dy$$

$$\Rightarrow \frac{2}{5}\int_0^1\int_0^1(2x+3y)\,dx\,dy$$

$$\Rightarrow \frac{2}{5}\int_0^1\left[x^2+3xy\right]_0^1 dy$$

$$\Rightarrow \frac{2}{5}\int_0^1 1+3y\ dy$$

$$\Rightarrow \frac{2}{5}\left[y+\frac{3y^2}{2}\right]_0^1$$

$$\Rightarrow \frac{2}{5} \times \frac{5}{2} = 1$$

(ii) $\frac{2}{5}\int_{1/4}^{1/2}\int_{0}^{1/2}(2x+3y)\,dx\,dy$

$\frac{2}{5}\int_{1/4}^{1/2}\left(x^2+3xy\right)_0^{1/2}dy$

$\Rightarrow \frac{2}{5}\int_{1/4}^{1/2}\frac{1}{4}+\frac{3y}{2}\,dy$

$\Rightarrow \frac{2}{5}\left[\frac{y}{4}+\frac{3y^2}{4}\right]_{1/4}^{1/2}$

8 + 8 − 3

$\Rightarrow \frac{2}{5}\left[\frac{1}{8}+\frac{3}{16}-\frac{1}{16}-\frac{3}{64}\right]$

$\Rightarrow \frac{2}{5}\times\left[\frac{13}{64}\right]$

$\Rightarrow \cancel{\frac{26}{64}}\quad \frac{26}{320}=\frac{13}{160}$

Q. The joint density function of X and Y is

$$f(x,y)=\begin{cases}2e^{-x}e^{-2y}, & 0<x<\infty,\ 0<y<\infty\\ 0 & \text{else where}\end{cases}$$

compute

i) $P[X>1,\ Y<1] = \int_0^1\int_1^{\infty} f(x,y)\,dx\,dy$

$= \int_0^1\int_1^{\infty} 2e^{-x}e^{-2y}\,dx\,dy$

$$\int_0^1 \left[-2e^{-x}\right]_1^{\infty} e^{-2y}\,dy$$

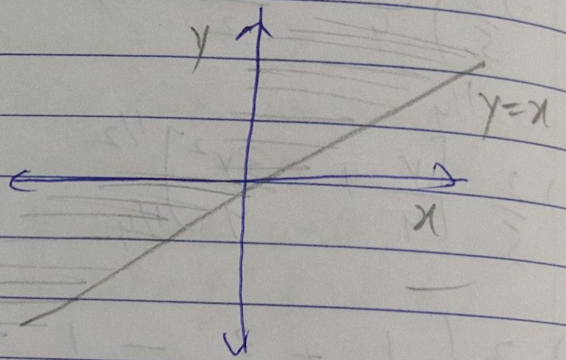
$$\left[-2+\frac{2}{e}\right]\left[\frac{-e^{-2y}}{2}\right]_0^1$$

$$\Rightarrow \left[-1+\frac{1}{e}\right]\left[-e^{-2}+1\right]$$

$$\Rightarrow e^{-2}-e^{-3}-1+e^{-1}$$

? Do again

iii) $P(X<Y)=\iint f(x,y)\,dx\,dy$



$$\int_0^{\infty}\int_0^{y} f(x,y)\,dx\,dy$$

$$\Rightarrow 2\int_0^{\infty}\left[-e^{-x}\right]_0^{y} e^{-2y}\,dy$$

$$2\int_0^{\infty}\left[-e^{-3y}+e^{-2y}\right]dy$$

$$\Rightarrow 2\left[\frac{+e^{-3y}}{3}-\frac{e^{-2y}}{2}\right]_0^{\infty}$$

$$\Rightarrow \cancel{2\left[\frac{1}{3}-\frac{1}{2}\right]}$$

$$\Rightarrow 2\left[-\frac{1}{3}+\frac{1}{2}\right]$$

$$\Rightarrow \frac{1}{6}\times 2=\frac{1}{3}$$

? Do

(iii) $P(X < a)$

$$2\int_0^{\infty}\int_0^{a} e^{-x} e^{-2y}\, dx\, dy$$

=) .

Relation between CDF ($F(x)$) and PDF ($f(x)$)

$$F(x) = \int_{-\infty}^{x} f(x)\, dx$$ use leibnitz rule to prove

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

marginal distribution

given $f(x,y) \longrightarrow$ joint pmf or joint pdf

X $\qquad$ Y

$g(x)$ $\qquad$ $h(y)$

marginal distribution for X & Y

$x = 0, 1$
$y = 0, 1, 2$

Example
$g(0) = f(0,0) + f(0,1) + f(0,2)$
$g(1) = f(1,0) + f(1,1) + f(1,2)$

[A] FOR DISCRETE CASE : ① $g(x) = \sum_{y} f(x,y)$

② $h(y) = \sum_{x} f(x,y)$

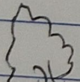
y \ x	0	1	2	Row sum	
0	3/28	9/28	3/28	15/28	$h(0)$
1	3/14	3/14	0	6/14	$h(1)$
2	1/28	0	0	1/28	$h(2)$
Colsum	5/14	15/28	3/28	Sum = 1	
	$g(0)$	$g(1)$	$g(2)$		

~~$g(0)$~~

For continuous random variable X & Y

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

$g(x)$ & $h(y)$ are marginal totals of the respective columns and rows



In this question find $g(x)$ and $h(y)$

$$g(x) = \frac{2}{5}\int_0^1 (2x+3y)\,dy = \frac{2}{5}\left[2xy + \frac{3y^2}{2}\right]_0^1$$

$$\Rightarrow \frac{2}{5}\left[2x + \frac{3}{2}\right]$$

$$\Rightarrow \frac{8x+6}{10} = \frac{4x+3}{5}$$

$$h(y) = \frac{2}{5}\int_0^1 (2x+3y)\,dx = \frac{2}{5}\left[x^2 + 3xy\right]_0^1$$

$$\Rightarrow \frac{2}{5} + \frac{6}{5}y$$

$$\Rightarrow \frac{2+6y}{5}$$

$$g(x) \begin{cases} \frac{4x+3}{5}, & 0<x<1 \\ 0 & \text{otherwise} \end{cases} ; \quad h(y) \begin{cases} \frac{2+6y}{5} & 0<y<1 \\ 0 & \text{o.w} \end{cases}$$

conditional distribution

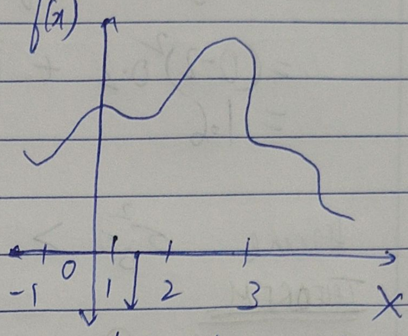
$$P(X=x / Y=y) = \frac{f(X=x, Y=y)}{h(y)} = \frac{f(x,y)}{f(y)}$$

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

variance of random variable

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

where E is expectation



$g(x) = (X-\mu)^2$

μ (mean) $= E(x)$

&

$\text{variance}(X) = E(g(x))$

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

$= \sum_x (x-\mu)^2 f(x)$ (for discrete X) *

* X is CTS (continuous)

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

0.1
0.6
0.9
0.4
2.0

Q: let the random variable X represent no of automobiles

Company A

x	1	2	3
$f(x)$	0.3	0.4	0.3

Company B

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

show that variance of company B > variance of company A

① $E(x) = 1(0.3) + 2(0.4) + 3(0.3) = 2$

$\sigma_A^2 = \sum_1^3 (x-2)^2 f(x)$

$= (1-2)^2(0.3) + 0 + (3-2)^2 0.3$

$= 0.6$

② $E(x) = 0 + 1(0.1) + 2(0.3) + 3(0.3) + 4(0.1)$

$= 2$

$$\sigma_B^2 = \sum_0^4 (x-2)^2 f(x)$$

$$= (0-2)^2 0.2 + (1-2)^2 0.1 + 0 + (3-2)^2 0.3 + (4-2)^2$$

$$= 1.6$$

Hence $\sigma_B^2 > \sigma_A^2$ proved

THEOREM

Q. The variance of r.v X is

$$\boxed{\sigma^2 = E(x^2) - (E(x))^2}$$

PROOF

$$\sigma^2 = \sum_x (x-\mu)^2 f(x)$$

$$= \sum_x x^2 f(x) + \mu^2 \overbrace{\sum_x f(x)}^{\sum PMF = 1} - 2\mu \sum_x x f(x)$$

$$= E(x^2) + \mu^2 (1) - 2\mu(\mu)$$

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

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

Q.

x	0	1	2	3
$f(x)$	0.51	0.38	0.1	0.01

~~XLYAD~~

$$E(x) = 0 + 1(0.38) + 2(0.1) + 3(0.01)$$

$$= 0.61$$

Ex $g(x) = 2X + 3$

x	0	1	2	3
$f(x)$	1/4	1/8	1/2	1/8

find variance $\sigma^2_{g(x)}$

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

$= \sum_{0}^{3} (2x+3) f(x)$

$= 3\left(\frac{1}{4}\right) + 5 \times \left(\frac{1}{8}\right) + 7 \times \frac{1}{2} + 9 \times \frac{1}{8}$

$\Rightarrow \frac{6+5+28+9}{8} = \frac{48}{8} = 6$

$\Rightarrow 6$

$\sigma^2_{g(x)} = \sum_3 (2x+3-6)^2 f(x)$

$= \sum_{0} (2x-3)^2 f(x)$

$= 4$

Q. $f(x) = \begin{cases} x^2/3 & -1 < x < 2 \\ 0 & \text{otherwise} \end{cases}$ find $\sigma^2_{g(x)}$

$g(x) = 4X + 3$

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

$= \int_{-1}^{2} (4x+3)\left(\frac{x^2}{3}\right)$

$= \int_{-1}^{2} \frac{x^4}{3} + \frac{3x^3}{3} = \frac{16}{3} + \frac{8}{3} - \frac{1}{3} + \frac{1}{3} = \frac{24}{3}$

$= 8$

$$\text{variance}(X) = \int_{-1}^{2} (g(x) - 8)^2 f(x)\, dx$$

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

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

$$= \int_{-1}^{2} \frac{16x^{\cancel{3}4}}{3} + \frac{25x^2}{3} - \frac{40x^3}{3}$$

$$\Rightarrow \cancel{256} \quad \cancel{\frac{64}{3}} + \frac{16x^5}{15} + \frac{25x^3}{9} - \frac{40x^4}{12}$$

$$\Rightarrow \frac{16 \times 32}{15} + \frac{25 \times 8}{9} - \frac{40 \times 16}{12} - \frac{16}{3} - \frac{25}{3} - \frac{40}{3}$$

$$\Rightarrow \frac{16 \times 27}{15} + \frac{25 \times 5}{9} - \frac{40 \times 20}{12}$$

$\Rightarrow$

* **NOTE** : Properties of Expectation

① * If a & b are constants then

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

LHS $\int_{-\infty}^{\infty}(aX+b)f(x)\,dx$

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

$$= aE(X)+b$$

→ If $b=0$, $E(aX) = aE(X)$

→ If $a=0$, $E(b) = b$

Q. $f(x) = 2X-1$

$$E(2X-1) = 2E(X)-1$$
$$= 2\times 12.67 - 1$$
$$= 24.54$$

② * $g(X)$, $h(X)$ 2 functions of X

$$E[g(x) \pm h(x)] = E(g(x)) \pm E(h(x))$$

Q.

x	0	1	2	3
$p(x)$	1/3	1/2	0	1/6

find $E(X-1)^2$

$$E(X-1)^2 = E(X^2+1-2X)$$
$$= E(X^2) + E(1) - 2E(X)$$
$$= 2 \cancel{+} -1$$
$$= \underline{1}$$

NASA

(3) * $E[g(x,y) \pm h(x,y)] = E[g(x,y)] \pm E[h(x,y)]$

(4) * If $g(x,y) = X$ & $h(x,y) = Y$

then $E[X \pm Y] = E(X) \pm E(Y)$

(5) * If X & Y are independent random variable then $E(XY) = E(X)E(Y)$

x & y

$f(x,y) = g(x)h(y)$

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

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

$$= E(x)\,E(y)$$

Properties of Variance

$$Var(x) = E(x^2) - (E(x))^2$$

1. $Var(b) = E(b^2) - (E(b))$ (b = constant)
$= b^2 - b^2$
$= 0$

2. $var(ax) = E(a^2x^2) - (E(ax))^2$
$= a^2E(x^2) - a^2(E(x))^2$
$= a^2[E(x^2) - E(x)^2]$
$= a^2 Var(x)$

Ex: $Var(4X - 2Y + 6)$ $\sigma_x^2 = 9$, $\sigma_Y^2 = 3$

$= Var(4X) - Var(2Y) + Var(6)$
$= 16 \times 9 - 4 \times 3 + 0$
$= 132$

Co-Variance of X & Y

$$\sigma_{xy} = E[(X - \mu_x)(Y - \mu_y)]$$

$$= \sum_y \sum_x (x - \mu_x)(y - \mu_y) f(x,y)$$

when ~~if~~ x & y are discreate
when x & y are continous

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

* The sign of covariance indiates relation b/w 2 random variables X & Y
  - +ve → increasing relation
  - −ve → decreasing relation

$$\sigma_{xy} = E(XY) - E(X)E(Y)$$

* **PROOF**

$$\sum_y \sum_x xy f(x,y) - \mu_y \sum_y \sum_x x f(x,y) - \mu_x \sum_y \sum_x y f(x,y)$$
$$+ \mu_y \mu_x \sum_y \sum_x f(x,y)$$

$= E(XY) - \mu_Y\mu_X - \mu_X\mu_Y + \mu_X\mu_Y$

$\Rightarrow E(XY) - \mu_X\mu_Y$

$\Rightarrow \boxed{E(XY) - E(X)E(Y)} = \sigma_{XY}$

~~Ball point example from Past~~

Given $f(x,y)$

$E(X) = \sum_x x\, g(x)$ or $\sum_x \sum_y x f(x,y)$

$= \cancel{0 \times 5}\ 0 + 1\left(\frac{15}{28}\right) + 2\left(\frac{3}{28}\right)$

$\Rightarrow \frac{21}{28} = \frac{3}{4}$

$E(Y) = \sum_y y\, h(y)$

$= 0 + 1\left(\frac{3}{7}\right) + 2 \times \frac{1}{28}$

$= \frac{12}{28} + \frac{2}{28} = \frac{1}{2}$

$E(XY) = \sum_y \sum_x xy\, f(x,y) = \frac{3}{14}$

$\sigma_{XY} = \frac{3}{14} - \frac{3}{4} \times \frac{1}{2}$

$= \frac{3}{14} - \frac{3}{8}$

$= \frac{-9}{56}$; X & Y are -vely related

Ex: $f(x,y) = \begin{cases} 5xy & 0 \le x \le 1 \\ & 0 \le y \le x \\ 0 & \text{otherwise} \end{cases}$

$$g(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = 8x\int_0^x y\,dy = 8x\left(\frac{y^2}{2}\right)_0^x$$
$$= 4x^3$$

Ind $g(x) = \begin{cases} 4x^3 & 0 \le x \le 1 \\ 0 & \text{o.w} \end{cases}$

$$h(y) = \int_{-\infty}^{\infty} f(x,y)\,dx = y\int_0^1 8x\,dx = y\left(4x^2\right)_0^1$$
$$= 4y$$

$h(y) = \begin{cases} 4y & 0 \le y \le \cancel{1} \\ 0 & \text{ow} \end{cases}$

$$E(x) = \int_0^1 x g(x)\,dx = \int_0^1 4x^4\,dx = \frac{4x^5}{5} = \frac{4}{5}$$

$$E(y) = \int_0^x y\, \cancel{g}\cdot h(y)\,dy = \int_0^x 4y^2\,dy = \frac{4y^3}{3} = \frac{4}{3}$$

$$E(xy) = \int_0^1\int_0^1 4x^4 \, 4y^2\,dx\,dy$$

$$16\int_0^1\int_0^1 x^4 y^2\,dx\,dy$$

$$= 16\int_0^1 \left(\frac{x^5}{5}\right)_0^1 y^2\,dy$$

$$= 16\int_0^1 \frac{y^2}{5}\,dy = \frac{16\,y^3}{15} = \frac{16}{15}$$

$\Rightarrow$

☆ If X, Y are independent then $E(XY) = E(X)E(Y)$
hence $\sigma_{xy} = 0$