Probability and statistics

Statistics:

According to R.A. Fisher, statistics is defined as collection of data, analysis of data and interpretetion of data.

Types of data

Grouped / frequency distribution
ungrouped / Raw data
open
closed

Grouped data.

If the data is in the form of class intervals and frequencies together then the data is called as grouped data

(OR)

Distributing the frequenceies to their corresponding class intervals is knows as frequency distribution

Closed data:

If the class intervals are in continuous form without any discontinuity then the data is called as closed data otherwise it is open data

ungrouped data:

If the data contains only observations without any class interv then the data is known as ungrouped (or) raw data

Mean (Average): $\overline{X}$

In ungrouped data, $\overline{X} = \dfrac{\sum_{i=1}^{n} x_i}{n}$

$n \rightarrow$ No. of observations

In grouped data, $\overline{X} = \dfrac{\sum f_i x_i}{N}$

x_i = Midpoint of each class = $\dfrac{UL+LL}{2}$
f_i = frequency of each class
N = Sum of frequencies

Note: To apply any statistical Model, the data must be closed data.

Median:

In ungrouped data,

If 'n' is odd then the middle observation itself is a median

If 'n' is even then the average between the middle observations is median,

provided

1) The data should be rearranged either in increasing order (or) decreasing order
2) the no. of observations above the middle = no. of observations below the middle

) In grouped data

$$\text{Median } (Md) = L + \frac{\left(\frac{N}{2} - m\right)}{f} * C$$

L = Lower limit of the ideal class

f = Frequency for the ideal class

m = Cumulative frequency for above the ideal class

C = Size of class (ie. class Interval size)

Eg: find the median for the following grouped data

Class Interval (CI)	frequency
0-10	3
10-20	5
20-30	7
30-40	2
40-50	1

Sol:

CI	frequency	c.f
0-10	3	3
10-20	5	8
20-30	7	15
30-40	2	17
40-50	1	18

← Ideal class

cf = cumulative frequency

$N = 18$

$\frac{N}{2} = 9 \Rightarrow$ It is between 8 & 15 (cf)

9 is moving towards 15

$\therefore$ Ideal class $\Rightarrow$ 20-30

$\therefore$ $L = 20$

$m = 8$, $f = 7$

$$Md = 20 + \left(\frac{9-8}{7}\right) * 10$$

$$= 20 + 10/7$$

$$= 21.4$$

Note:

If the first class is ideal then the cumulative frequency and frequency are equal ($m = f$)

Mode :

13

I) In ungrouped data,

The most frequently repeated observation is known as Mode.

Eg :

p) 1, 2, 3, 4, 5, 2, 3, 6, 7, 11, 12, 2, 3, 14, 2

Mode = 2 ← Unimodel data

p) 1, 2, 3, 4, 5, 2, 3, 6, 7, 11, 12, 2, 3, 14, 2, 26, 3, 49

Mode = 2, 3 ← bimodel data

p) 0, 1, 6, 49, 57, 21, 102, 191

Mode = ? ← Mode does not exist as no number is repeated

II) In grouped data

~~Median~~ Mode is defined as an emperical relation, as follows

$$\text{Mode } M_0 = 3\,\text{Median} - 2\,\text{Mean}$$

(OR)

$$M_0 = L + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right) * C$$

L = Lowerlimit of Ideal class

$\Delta_1 = f - f_{-1}$ (above Ideal class)

$\Delta_2 = f - f_{+1}$ (below Ideal class)

Eg: find mode for the following grouped data

CI	frequency
0-2	11
2-4	14
4-6	17
6-8	08
8-10	04

sol: The highest frequency class is Ideal class

	CI	frequency	
	0-2	11	
	2-4	14	f_{-1}
L	4-6	17	f ← Ideal class
	6-8	08	
	8-10	04	f_{+1}

$L = 4$

$\Delta_1 = f - f_{-1} = 3$

$\Delta_2 = f - f_{+1} = 9$

$C = 2$

$M_0 = 4 + \left(\frac{3}{3+9}\right) * 2$

$= 4 + \frac{6}{12}$

$= 4.5$

note:

1. If the maximum frequencies are repeated first, last and in between, then select inbetween as ideal class.
2. If the maximum frequencies are repeated in between, select randomly (bimodel grouped data)
3. If all the frequencies are equal, mode is undefined ($\because \Delta_1 = \Delta_2 = 0 \Rightarrow 0/0$ form)
4. If the maximum frequencies are repeated first and last select randomly (first $\Rightarrow f_{-1} = 0$. last $\Rightarrow f_{+1} = 0$)

Measures of central tendency

- Mean
  Median
  Mode
- out of these, Mean is comparatively best Measurement because it considers all data.
- with these measures, we cant identify the uniformity, regularity, consistency of data. For this we require measure of dispersion

Measures of dispersion/variability

Range
Quartile Deviation (Q.D.)
Mean Deviation (M.D.)
standard Deviation (S.D.)
Coefficient of variation (C.V.)

Range:

$$\text{Range} = \left.\begin{array}{l}\text{Max} - \text{Min (or)} \\ G.V - L.V\end{array}\right\} \text{grouped (or) ungrouped data}$$

G.V → Greatest value
L.V → Least value

standard deviation:

standard deviation, $SD = \sqrt{\text{variance}}$

$$\text{Variance} = (SD)^2$$

- variance is used to measure the differences, deviations within the group.

In ungrouped data

$$\text{Variance} = \sigma_x^2 = \frac{\Sigma(x_i - \bar{x})^2}{n} \rightarrow \text{sum of the squares of deviation from mean}$$

$$= \frac{\Sigma x_i^2}{n} - \cancel{\Sigma}(\bar{x})^2$$

Note:

- Lesser variance is more consistent (or) more uniform
- variance can never be negative
- variance of constant is zero
- Sum of the squares of deviations from mean should be uniform
- If variances are equal for groups, then greater mean is more consistent

I) In grouped data

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

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

$\longrightarrow$ for both grouped, ungrouped data

$6QD = 5MD = 4SD$

$\Rightarrow 6QD = 4SD$	$5MD = 4SD$
$QD = \frac{4}{6} SD$	$MD = \frac{4}{5} SD$
$\boxed{QD = \frac{2}{3}\sigma}$	$\boxed{MD = \frac{4}{5}\sigma}$

$\rightarrow$ $C.V = \frac{S.D}{Mean} \times 100$

$= \frac{\sigma}{\bar{x}} \times 100$

Note:

Lesser σ $\Rightarrow$ lesser C.V

$\Rightarrow$ Data is more consistent (or) uniform

3/10/11

Biased coin $\Rightarrow$ Both sides Head (or) Both sides tail

Random experiment:

Unpredictable outcomes of an experiment is known as a random experiment

Eg: 1) Tossing a unbiased coin (Both faces are different)

2) Rolling a die

Drawing a card from a pack of 52 cards

Sample space:

The collection of all possible outcomes of an experiment is known as sample space and is denoted by 'S'

Event:

The outcomes of an experiment is known as event and is denoted by E

Mathematically event is a subset of sample space

$$\boxed{E \subseteq S}$$

Probability:

The probability of an event is defined as the ratio betwe the favourable number of cases to the event and the no. of outcomes of the event (The outcomes are mutually exclusive and exhaustive events)

15

$$P(E) = \frac{m}{n} \quad (m \leq n)$$

Axiomatic approach to probability (or) Rules of probability.

1. Sum of probabilities of all events in sample space is one

$$P(S) = 1$$

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

$P(H) = \frac{1}{2}$

$P(T) = \frac{1}{2}$

$P(S) = 1$

2. $0 \leq P(E) \leq 1$

ie. The probability of an event ranges from 0 to 1

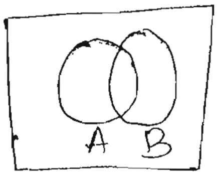
If $P(E) = 0$ then it is impossible event and is denoted by 'ϕ'

ie. $P(\phi) = 0$

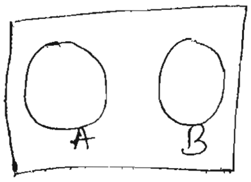
If $P(E) = 1$ then it is certain event (or) sure event.

3. $P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i)$ iff E_i's are disjoint / mutually exclusive events

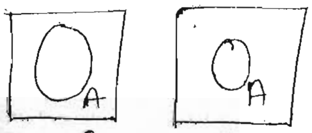
ie probability of sample space is equal to the sum of probabilities of individual events in sample space provided the events are mutually exclusive events



· Dependent events



Mutually exclusive events



S_1 S_2

Independent events

Note:

· occurance of one event doesnot depends upon occurance of other events in same sample space then those events are called as mutually exclusive events

→ Let A & B are mutually exclusive events then

$$A \cap B = \phi$$

$$\Rightarrow P(A \cap B) = 0$$

: occurance of one event doesnot depends upon the occurance of same event in a different sample space. then those events are known as independent points

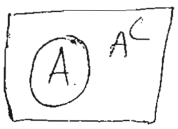
: Mutually exclusive events never be independent and Independent events never be mutually exclusive events

Results:

1. $P(S) = 1$
2. $0 \leq P(E) \leq 1$
3. $P(A^c) = 1 - P(A)$

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

Complementary theorem of probability



$A \cup A^c = S$

$P(A \cup A^c) = P(S)$

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

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

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

4. If A & B are events   Event → default dependent events

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

If A&B are mutually exclusive events

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

→ $P(A+B) = P(A) + P(B)$

$P(A+B+C) = P(A) + P(B) + P(C)$

'+' indicates that events are mutually exclusive

This is Addition theorem of probability

5. Multiplication theorem of probability (Dependent events)

→ If A & B are events

$P(A\cap B) = P(A)\cdot P(B/A)$ (B depends on A)

$= P(B)\, P(A/B)$ (A depends on B)

Known event   Known event

$P(B/A)$ (or) $P(A/B)$ is conditional probability

Unknown event   Unknown event

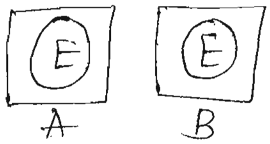
→ If A, B, C are ~~the~~ three events

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

6. Multiplication theorem (Independent)

A & B are independent iff

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



i.e. $E_1, E_2, E_3, \ldots E_n$ are independent iff

$$P\left(\bigcap_{i=1}^{n} E_i\right) = \prod_{i=1}^{n} P(E_i)$$

i.e. there 'n' sample spaces

But in all sample spaces, event is same.

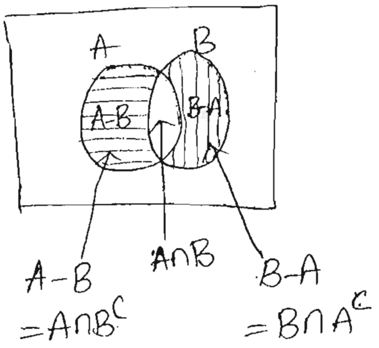
- If A & B are two events then

$P(A\cap B^c) = P(A) - P(A\cap B)$

$P(A^c\cap B) = P(B) - P(A\cap B)$

$P(A^c\cap B^c) = P(\overline{A\cup B})$

$= 1 - P(A\cup B)$



** $P(A\Delta B) = P(A\cap B^c) + P(A^c\cap B)$

$A\cap B^c \Rightarrow$ occuring only event A

$A^c\cap B \Rightarrow$ occuring only event B

$A^c\cap B^c \Rightarrow$ neither A nor B (none of A, B)

$A\Delta B \Rightarrow$ only once
ie. If A occurs then B doesnot occur
If B occurs then A doesnot occur

→

$$P(A^c/B) = \frac{P(A^c\cap B)}{P(B)} \quad \text{(from multiplication theorem)}$$

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

$$= 1 - \frac{P(A\cap B)}{P(B)}$$

$$\boxed{P(A^c/B) = 1 - P(A/B)}$$

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

$$= \frac{P(A) - P(A\cap B)}{1-P(B)} \quad (P(B)\neq 1)$$

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

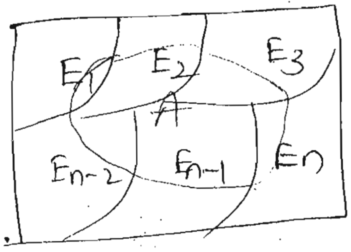
$$= \frac{1-P(A\cup B)}{1-P(B)} \quad (P(B)\neq 1)$$

NOTE:

If A and B are independent events $A^c \cap B$, $A \cap B^c$, $A^c \cap B^c$ are also independent.

8) ** **Baye's theorem:**

If $E_1, E_2, E_3, \dots E_n$ are the mutually exclusive events $(P(E_i) \neq 0)$ such that A is an arbitrary event which is a subset of $\bigcup_{i=1}^{n} E_i$ then



$$P(A) = P(A \cap E_1) + P(A \cap E_2) + \dots + P(A \cap E_n)$$

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

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

17

This is total probability.

Reverse probability

A - unknown event

E_i's - known event

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

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

steps in Baye's theorem:

1. Identify the known events in data (mutually exclusive)
2. Select the unknown event (It should be a part of known events)
3. write the probability of unknown in terms of known events ($P(A/E_i)$)
4. Find the total probability of unknown event
5. Evaluate reverse probability for known events

note:

→ 1-coin → 2
2-coin → 2^2
⋮
n-coin → 2^n

1-die → 6
2-dice → 6^2
⋮
n-dice → 6^3

52 cards

suit:	K	Q	J	
(13) Hearts	1	1	1	← Red (26)
(13) Diamonds	1	1	1	
(13) club	1	1	1	Black (26)
(13) Spade	1	1	1	
	4	4	4	= 12

12 → No. of face cards

atleast	min	$\geq$
atmost	max	$\leq$
and	product	$\cap$
or	addition	$\cup$

→ Addition theorem

either or

atleast once

atmost once

or

Multiplication theorem:

simultaneously

successively

as well as

one after the other

one by one

alternatively

and

) If 3 coins are tossed at a time, find the probability of getting atmost one head

sol: $n(S) = 2^3 = 8$ S = {HHH, HHT, HTH, THH, HTT, TTH, THT, TTT}

$X \leftarrow$ No. of heads

$$P(\text{atmost one Head}) = P(X \leq 1)$$
$$= P(X=1) + P(X<1)$$
$$= P(X=1) + P(X=0)$$
$$= \frac{3}{8} + \frac{1}{8}$$
$$= \frac{1}{2}$$

) Find the probability that atleast one tail got when 3 coins tossed

$X \leftarrow$ No. of tails

$$P(X \geq 1) = 1 - P(X<1)$$
$$= 1 - P(X=0)$$

(or)

$P(X \geq 1) = P(X=1) + P(X>1)$

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

$= \frac{3}{8} + \frac{3}{8} + \frac{1}{8}$

$= \frac{7}{8}$

P) when 3 coins are tossed, find the probability of getting atleast one Head and atmost one tail.

Sol:

$\left.\begin{matrix} T\ H\ H \\ H\ T\ H \\ H\ H\ T \end{matrix}\right\}$ 1 Tail

$H\ H\ H \}$ 0 Tail

$P(\text{Required}) = \frac{4}{8} = \frac{1}{2}$

P) 4 coins are tossed at a time. find the probability that atleast two heads and atleast two tails

Sol: $n(S) = 2^4 = 16$

HHTT $\Rightarrow \frac{4!}{2!\,2!} = 6$ i.e. $4C_2 \leftarrow$ favourable cases

No. of repetitions

i.e.
HHTT
HTHT
HTTH
THTH
TTHH
THHT

$P(\text{Required}) = \frac{6}{16} = \frac{3}{8}$

Note:

	Heads		Tails	
getting 0 heads	$4C_0$	= 1 =	$4C_4$	← getting 4 tails
" 1 "	$4C_1$	= 4 =	$4C_3$	← " 3 "
" 2 "	$4C_2$	= 6 =	$4C_2$	← " 2 "
" 3 "	$4C_3$	= 4 =	$4C_1$	← " 1 "
" 4 "	$4C_4$	= 1 =	$4C_0$	← " 0 "
		16		

when 4 coins are tossed

) A coin is repeated 5 times. Find the probability that Head appears odd number of times

Sol: $n(S) = 2^5 = 32$

$P(\text{required}) = \frac{5C_1 + 5C_3 + 5C_5}{2^5}$

$= \frac{16}{32} = \frac{1}{2}$

) If coin is repeated 'n' times. Find the probability of getting tail ~~odd~~ even no. of times.

Sol: $P(\text{required}) = \frac{nC_0 + nC_2 + nC_4 + \cdots + nC_n}{2^n}$

$= \frac{2^{n-1}}{2^n} = \frac{1}{2}$

Note: $nC_0 + nC_2 + nC_4 + \cdots + nC_n = nC_1 + nC_3 + nC_5 + \cdots + nC_{n-1} = 2^{n-1}$

) A coin is repeated 5 times. Find the probability that the no. of Heads are more than the no. of tails

Sol:

No. of heads

Heads — $5C_0 \quad 5C_1 \quad 5C_2 \quad 5C_3 \quad 5C_4 \quad 5C_5$

tails — $5C_5 \quad 5C_4 \quad 5C_3 \quad 5C_2 \quad 5C_1 \quad 5C_0$

No. of tails

$P(\text{Required}) = \frac{5C_3 + 5C_4 + 5C_5}{2^5} = \frac{1}{2}$

) Two dice are rolled. Find the probability that the first die should contain a prime number or a total of 8.

Sol: A: getting prime number on first die

B: getting a total (Sum) of 8.

$n(S) = 6^2 = 36$

Prime no. on first die

(2,1) (2,2) (2,3) (2,4) (2,5) (2,6) — 6

(3,1) — — — 6

(5,1) — — — 6

$P(A) = \frac{18}{36}$

$= \frac{1}{2}$

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

8 8

Sum = 8

$P(B) = \frac{5}{36}$

$P(A \cap B) = \frac{3}{36}$

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

$= \frac{18}{36}+\frac{5}{36}-\frac{3}{36} = \frac{20}{36} = \frac{5}{9}$

P) Two dice are rolled. Find the probabability that neither sum 7 nor sum 11

Sol: A: getting sum 7. $\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$

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

B: getting sum 11 $\{(5,6),(6,5)\}$

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

$P(A^c\cap B^c) = P(\overline{A\cup B})$

$= 1-P(A\cup B)$

$= 1-\{P(A)+P(B)-P(A\cap B)\}$

$= 1-\{\frac{6}{36}+\frac{2}{36}-0\}$

$= 1-\frac{2}{9} = \frac{7}{9}$

P) Two dice are rolled two times. Find the probability that for getting a sum 7

i) atleast once ii) only once iii) Twice

Sol: Two dice are rolled two times
$\Downarrow$
Independent

Same common event: getting 7.

A: Getting sum 7 first time

B: Getting sum 7 second time

$P(A) = \frac{1}{6} \Rightarrow P(A^c) = \frac{5}{6}$

$P(B) = \frac{1}{6} \Rightarrow P(B^c) = \frac{5}{6}$

i) $P(\text{atleast once}) = P(A\cup B)$
$\Downarrow$
addition-theorem

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

$= P(A)+P(B) - P(A)\cdot P(B)$

$= \frac{2}{6}-\frac{1}{36} = \frac{11}{36}$

ii) $P(\text{only once}) = P(A \Delta B)$

$= P(A \cap B^c) + P(A^c \cap B)$

$= P(A) \cdot P(B^c) + P(A^c) \cdot P(B)$ (As Independent)

$= \frac{1}{6} \cdot \frac{5}{6} + \frac{5}{6} \cdot \frac{1}{6}$

$= 10/36 = 5/18$

iii) $P(\text{Twice}) = P(A \cap B) = P(A) \cdot P(B) = \frac{1}{6} \cdot \frac{1}{6} = 1/36$

ie. first time aswell as second time ie. first time and second time

∴ product theorem

) 4 cards are drawn at random from a ~~card~~ pack of 52 cards. Find the probability that

i) All 4 are drawn from the same suit

Here drawing 4 cards at a time from a single suit

$$P(\text{required}) = \frac{\overset{H}{^{13}C_4} + \overset{D}{^{13}C_4} + \overset{C}{^{13}C_4} + \overset{S}{^{13}C_4}}{^{52}C_4} = \frac{4 \cdot {^{13}C_4}}{^{52}C_4}$$

ii) No two cards are from same suit

Not considering two at a time from a single suit

$$P(\text{required}) = \frac{^{13}C_1 \cdot {^{13}C_1} \cdot {^{13}C_1} \cdot {^{13}C_1}}{^{52}C_4}$$

⇒ Taking a single card from each suit at a time.

ie Drawing one by one

Product

) A card is drawn from the pack of 52 cards. Find the probability that it is

i) neither a diamond nor a face card

$P(D) = \frac{13}{52}$ $P(D \cap F) = \frac{3}{52}$

$P(F) = 12/52$

$P(D^c \cap F^c) = 1 - P(\overline{D \cup F})$

$= 1 - \{P(D) + P(F) - P(D \cap F)\}$

$= 1 - 22/52 = 30/52$

ii) either As or '10'

$P(A) = \frac{4}{52}$

$P(B) = \frac{\cancel{10}\,4}{52}$

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

$= \frac{4}{52} + \frac{4}{52} - 0$

$= \frac{8}{52} = \frac{2}{13}$

Q) A determinant is chosen from the set of all determinant of order 2 Matrix with the elements 0 and (or) 1. Find the probability that the chosen determinant is non zero

Sol:

$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ $n(S) = 2 \cdot 2 \cdot 2 \cdot 2 = 16$

$\Delta = ad - bc \neq 0$

case (i): $\Delta = +1$

$\Rightarrow a = d = 1$ and atleast one of b, c is '0'

$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$ $\begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix}$ $\begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix}$ $\Rightarrow$ 3 possibilities

case (ii): $\Delta = -1$

$\Rightarrow b = c = 1$ and atleast one of a, d is '0'

$\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}$ $\begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix}$ $\begin{vmatrix} 0 & 1 \\ 1 & 1 \end{vmatrix}$ $\Rightarrow$ 3 possibilities.

$P(\text{non zero } \Delta) = \frac{3+3}{16} = \frac{6}{16}$

$P(\text{non negative } \Delta) = 1 - \frac{3}{16} = \frac{13}{16}$ (−ve Δ)

$P(\text{zero } \Delta) = 1 - \frac{6}{16} = \frac{10}{16}$

Q) A die is rolled. If the no. is odd number, find the probability for getting prime number.

sol: A: getting odd ← known event

$P(A) = \frac{3}{6}$

P: getting prime ← odd event

$P(P/A) = \frac{P(P \cap A)}{P(A)} = \frac{2/6}{3/6} = 2/3$

) A number is chosen from 100 numbers. $\{00, 01, 02, \ldots 99\}$. Let 'x' denotes the sum of digits on the number. 'y' denotes the product on the number. Find the probability that $P(X=9/Y=0)$

sol: $P(X=9/Y=0) = \frac{P(X=9 \cap Y=0)}{P(Y=0)}$

$= \frac{2/100}{19/100} = \frac{2}{19}$

) A, B, C are three players tossing the same coin on the condition that one who gets the head first wins the game. If A starts the game, what are the winning chances of player C in 3rd trial.

sol: $P(H) = P(T) = \frac{1}{2}$

$P(H) = p = \frac{1}{2}$

$P(T) = q = \frac{1}{2}$

Probability winning 'C' in 1st trial $\Rightarrow P(\text{win } C) = q \cdot q \cdot p = q^2 p$

" 2nd trial $\Rightarrow q \cdot q \cdot q \cdot q \cdot q \cdot p = q^3 \cdot q^2 p$

" 3rd trial $\Rightarrow q^6 \cdot q^2 p$

$P(\text{winning } C \text{ in 3rd trial}) = q^3 \cdot q^3 \cdot q^2 p$

$= \frac{1}{2^8} \cdot \frac{1}{2} = \frac{1}{2^9}$

Random variable | Expectation

connecting the outcome of an experiment with the real values is known as random variable (1 dimensional random variable). The corresponding data is known as univariate data.

2D random variable: connecting the two outcomes at a time with one real number and those two outcomes must be from same sample space. The corresponding data is known as bivariate data.

Random variable

- Discrete random variable ⇕ Probability mass function $-p(x)$ ⇕ discrete distribution ⇕ Eg: binomial, poisson
- continuous random variable ⇕ probability density function $-f(x)$ ⇕ continuous distribution ⇕ Eg: Normal

Expectation (Mean | average).

$E(X) = \sum_{x=0}^{n} x \cdot p(x)$ where x is discrete random variable

$= \int_{-\infty}^{\infty} x f(x)\, dx$ x is continuous random variable

variance:

$V(X) = E(X^2) - (E(X))^2$

$= \Sigma x^2 p(x) - (\Sigma x\, p(x))^2$ x is discrete r.v.

$= \int x^2 f(x)\, dx - \left[\int x f(x) dx\right]^2$ x is continuous r.v

properties of Expectation:

- If x is a random variable and 'a' is a constant

$E(ax) = aE(x)$

- If x and y are random variables

$$E(x+y) = E(x) + E(y)$$
$$E(x-y) = E(x) - E(y)$$

- If x and y are random variables

$$E(x \cdot y) = E(x) \cdot E(y/x)$$ — conditional expectation
$$= E(y) \cdot E(x/y)$$

- x and y are independent random variables iff

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

- If $y = ax+b$ (a, b are constants) then

$$E(y) = aE(x) + b \qquad (\because E(\text{constant}) = \text{constant})$$

roperties of variance:

If 'x' is a random variable and 'a' is a constant

$$V(ax) = a^2 V(x)$$
$$V(-y) = (-1)^2 V(y) = V(y)$$

If x and y independent random variables

$$V(x+y) = V(x) + V(y)$$
$$V(x-y) = V(x) + V(-y)$$
$$= V(x) + V(y)$$

$$V(x \pm y) = V(x) + V(y)$$

- If 'a' and 'b' are constants, x & y are independent random variables

$$V(ax-by) = a^2 V(x) + b^2 V(y)$$
$$V\left(\frac{x}{a} - \frac{y}{b}\right) = \frac{1}{a^2} V(x) + \frac{1}{b^2} V(y)$$

- $Y = ax + b$; a, b are constants

$V(y) = V(ax+b)$

$= V(ax) + V(b)$

$= a^2 V(x) + 0$ $\quad (\because V(\text{const}) = 0)$

$= a^2 V(x)$

- If x and y are random variables

$V(x+y) = V(x) + V(y) + 2\,\text{cov}(x, y)$

$V(x-y) = V(x) + V(y) - 2\,\text{cov}(x, y)$

$\boxed{\text{cov}(x, y) = E(x \cdot y) - E(x) \cdot E(y)}$

- $\text{cov}(x, x) = E(x^2) - E(x) \cdot E(x) = E(x^2) - [E(x)]^2$

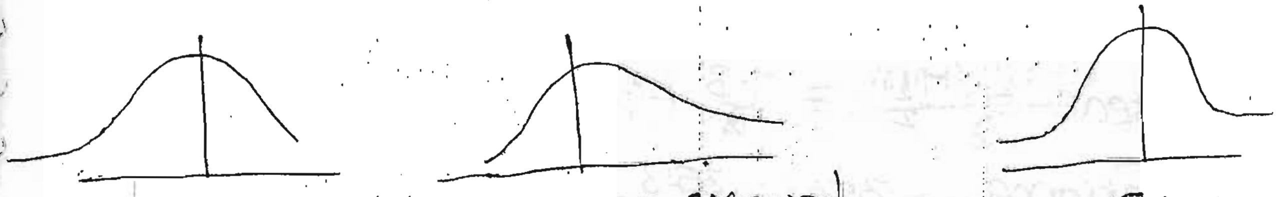
$= V(x)$

- $\text{cov}(a, b) = E(ab) - E(a) \cdot E(b)$

$= ab - ab = 0$

Note:
If x and y are independent random variables, $\text{cov}(x, y) = 0$ but converse need not be true

Skewness:
Lack of symmetry



-ve skewed   +ve skewed   symmetrical

$\beta_1 = \dfrac{\mu_3^2}{\mu_2^3}$ $\qquad \mu_3 =$ 3rd central moment

$\mu_2 =$ variance

Note: If $\mu_3 = 0$ then the curve is symmetric otherwise +vely skewed

Find the expectation of the no. on die when it is thrown

x	1	2	3	4	5	6
$P(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Probability

Mean $= \sum_{1}^{6} x \cdot P(x)$

$= 1\cdot\frac{1}{6} + 2\cdot\frac{1}{6} + 3\cdot\frac{1}{6} + 4\cdot\frac{1}{6} + 5\cdot\frac{1}{6} + 6\cdot\frac{1}{6} = \frac{21}{6} = \frac{7}{2}$

Find the variance for the single die

Sol: variance $= E(x^2) - E(x)$

$E(x^2) = 1^2\cdot\frac{1}{6} + 2^2\cdot\frac{1}{6} + 3^2\cdot\frac{1}{6} + 4^2\cdot\frac{1}{6} + 5^2\cdot\frac{1}{6} + 6^2\cdot\frac{1}{6} = \frac{91}{6}$

$[E(x)]^2 = (\frac{7}{2})^2 = \frac{49}{4}$

variance $= \frac{91}{6} - \frac{49}{4} = \frac{35}{12}$

Note: The mean and variance for the sum of the number on dice is

Mean $= \frac{7n}{2}$ n is no. of dice

variance $= \frac{35n}{12}$

) 3 unbiased dice are thrown. find the mean and and variance for the sum of the numbers on them

Sol: Mean $= \frac{7n}{2} = \frac{7\cdot 3}{2} = 21$

variance $= \frac{35n}{12} = \frac{35\cdot 3}{12} = \frac{35}{4}$

) A man has given 'n' keys of which 1 fits the lock. He tries them successively without replacement to open the lock. what is the probability that the lock will be open at the rth trial. Also determine mean and variance

Note: with replacement implies that independent trials.
without replacement implies that dependent trials.

Sol: probability of opening lock in

1^{st} trial $= \frac{1}{n}$

2^{nd} trial $= \frac{1}{n-1}$

3^{rd} trial $= \frac{1}{n-2}$

Probability of opening lock successfully 1st time in 2^{nd} trial $= (1-\frac{1}{n})(\frac{1}{n-1}) = \frac{1}{n}$

Probability of opening lock 1st success in 3^{rd} trial
$= (1-\frac{1}{n})(1-\frac{1}{n-1})\frac{1}{n-2}$
$= \frac{1}{n}$

Probability of opening lock 1st success in r^{th} trial $= \frac{1}{n}$

$\therefore P(x) = \frac{1}{n}$

Mean $= E(x) = \sum_{1}^{n} x \cdot P(x)$
$= P(x)\,(1+2+ \cdots +n)$
$= \frac{1}{n} \cdot \frac{n(n+1)}{2} = \frac{n+1}{2}$

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

$E(x^2) = \sum_{1}^{n} x^2 P(x)$
$= (1^2+2^2+ \cdots +n^2) \cdot \frac{1}{n} = \frac{n(n+1)(2n+1)}{6} \cdot \frac{1}{n} = \frac{(n+1)(2n+1)}{6}$

$V(x) = \frac{(2n+1)(n+1)}{6} - \frac{(n+1)^2}{4} = \frac{n+1}{2}\left[\frac{2n+1}{3} - \frac{n+1}{2}\right] = \frac{n^2-1}{12}$

Note: The probability of r^{th} trial with replacement $= q^{r-1} \cdot p$

P) If x is continuous random variable,

$f(x) = K \cdot x^2 e^{-x} \quad 0 \le x \le \infty$

find K, Mean, variance

$\int_0^\infty k \cdot x^2 \cdot e^{-x}\,dx = 1 \Rightarrow k\int_0^\infty x^2 e^{-x}\,dx = 1$

we have $\int_0^\infty e^{-x}\cdot x^{n-1}\,dx, (n>0) = \Gamma n$
$= (n-1)!$

$\Gamma 1 = 1, \quad \Gamma\frac{1}{2} = \sqrt{\pi}$

$\Rightarrow k\int_0^\infty x^2 e^{-x}\,dx = 1$

$k(2!) = 1 \Rightarrow k = \frac{1}{2}$

~~variance $\leq 12-9 \leq 3$~~

Mean $= E(x) = \int_0^\infty x\cdot f(x)\,dx$

$= \frac{1}{2}\int_0^\infty x^2 \cdot x \cdot e^{-x}\,dx$

$= \frac{1}{2}\int_0^\infty x^3 e^{-x}\,dx$

$= \frac{1}{2}\cdot 3! = 3$

variance $= \int_0^\infty x^2 f(x)\,dx - [E(x)]^2$

$\int_0^\infty x^2 f(x)\,dx = \frac{1}{2}\int_0^\infty x^4 e^{-x}\,dx = \frac{4!}{2} = 12$

$\therefore$ variance $= 12 - 9 = 3$

) If x and y are random variables and expectation $E(x) = 10$ $V(x) = 25$ find +ve values a, b such that $y = ax - b$ has expectation $= 0$ and variance $= 1$

Sol: $E(y) = 0$
$aE(x) - b = 0$
$a(10) - b = 0$

$V(y) = 1$
$a^2 V(x) - 0 = 1$
$a^2 \cdot 25 = 1 \Rightarrow a = \frac{1}{5}$

$\therefore b = 2$

Binomial distribution:

x is said to be a binomial random variable if it allows values from 0 to n with the parameters n, p and its probability mass function is

$B(x; n, p) = P(x) = nC_x\, p^x q^{n-x} \quad 0 \leq x \leq n, \quad p+q = 1$

$= 0$ otherwise

Conditions:

1. observations are independent ('n' is small)
2. probability of success is constant ('p' is Large)
3. Mean > variance ($E(X) > V(X)$)

Properties:

1) $E(x) =$ mean $= np$, variance $= V(x) = \mu_2 = npq$

2) $\mu_3 = npq(q-p)$, $\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{n^2p^2q^2(q-p)^2}{(npq)^3} = \frac{(q-p)^2}{(npq)} = \frac{(1-2p)^2}{npq}$

3) Moment generating function (MGF)

$M_x(t) = E(e^{tx}) = (q+pe^t)^n$

characteristic function (cf) : $\phi_x(t) = E(e^{itx}) = (q+pe^{it})^n$

- Note: if $p = \frac{1}{2}$ then binomial distribution is symmetric i.e $\beta_1 = 0$ otherwise +vely skewed
- Sum of independent binomial variables is also binomial variabl
- The moment generating function is used for finding the addition and difference b/n random variables with probabili function
- The characteristic function used for finding the convolution and ratios b/n the random variables with probability fund

P) Find the probability of getting '9' exactly 2 in 3 times with a pair of dice

Sol: $n = 3$, $P(\text{sum}=9) = \frac{4}{36} = \frac{1}{9} \in \{(5,4)\ (4,5)\ (6,3)\ (3,6)\}$

$\therefore q = \frac{8}{9}$

$P(X=2) = {}^nC_x\, p^x q^{n-x} = 3C_2 \cdot (\frac{1}{9})^2 \cdot \frac{8}{9} = \frac{8}{243}$

P) The probability of a man hitting target is $\frac{1}{3}$. ① If he fires 5 times, what is the probability of his hitting target atleast twice ② how many times must he fire so that the probability of hitting target is atleast once is more than 90%.

Sol: $p = \frac{1}{3} \Rightarrow q = \frac{2}{3}$, $n = 5$

1) $P(X \geq 2) = 5C_2\, p^x q^{n-x} + 5C_3\, p^3q^2 + 5C_4\, p^4 q + 5C_5\, p^5$ (or) $1 - (5C_0\, q^5 + 5C_1\, p^1 q^4) = \frac{131}{243}$

2) $P(x \geq 1) > 0.9 \Rightarrow 1 - P(X=0) > 0.9 \Rightarrow 1 - nC_0\, q^n > 0.9$

$q^n < 0.1 \rightarrow (\frac{2}{3})^n < 0.1 \quad \therefore n = 5.62$

2) Two dice are rolled 120 times. Find the average number of times in which the number on first die is exceeds the number on the second die.

Sol: $n = 120$.

(2,1) (3,1), (4,1), (5,1), (6,1)

(3,2) (4,2) (5,2) (6,2)

(4,3) (5,3) (6,3)

(5,4) (6,4)

(6,5)

$$p = \frac{15}{36}$$

$$\text{Average} = E(X) = np$$

$$= 120 \cdot \frac{15}{36} = 50$$

) X and Y are the two binomial random variables $X \sim B(2,p)$, $Y \sim B(4,p)$. If $P(X \geq 1) = \frac{5}{9}$ then find $P(Y \geq 1)$?

Sol:

$$P(X \geq 1) = \frac{5}{9}$$

$$1 - P(X=0) = \frac{5}{9}$$

$$1 - q^n = \frac{5}{9}$$

$$q^n = \frac{4}{9}$$

$n=2$ for X $\quad q^2 = \frac{4}{9} \Rightarrow q = \frac{2}{3}$

$$\therefore p = \frac{1}{3}$$

$$P(Y \geq 1) = 1 - P(Y=0)$$

$$= 1 - q^n$$

$$= 1 - \left(\frac{2}{3}\right)^4 = 1 - \frac{16}{81} = \frac{65}{81}$$

Poisson distribution: Eg: Measuring accidents in a city

- It is a discrete distribution and is an extension of binomial distribution
- It is meant for rate of arrival
- It is used for measuring defective probability
- It best suits for rare occurances
- It is purely dependent
- If 'x' is said to be a poisson variable, it allows the values from 0 to ∞ with a parameter $\lambda (>0)$ and its probability mass function is

$$P(x; \lambda>0) = P(X) = \frac{e^{-\lambda} \lambda^x}{x!} \qquad \lambda>0, \quad 0 \le x < \infty$$

$$= 0 \qquad \text{otherwise}$$

conditions:

1. observations are infinitely large $(n \to \infty)$
2. success is very small $(p \to 0)$
3. $np = \lambda \Rightarrow p = \frac{\lambda}{n}$

$$P(x; np) = \frac{e^{-np} (np)^x}{x!}$$

(Approximation of binomial distribution)

Poisson process:

$$p(x; \lambda, t) = \frac{e^{-\lambda t} (\lambda t)^x}{x!}$$

properties:

$E(X) = \text{mean} = \lambda$

$\text{variance} = \mu_2 = \lambda$

$\mu_3 = \lambda$

$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{\lambda^2}{\lambda^3} = \frac{1}{\lambda}$

$\Rightarrow \beta_1 > 0$ ~~it~~ is positively skewed. Never be symmetric.

MGF $\Rightarrow M_X(t) = e^{\lambda(e^t - 1)}$

characteristic function (c.f) $\Rightarrow \phi_X(t) = e^{\lambda(e^{it}-1)}$

Note:

In poisson distribution mean = variance = parameter = λ

~~Some~~ sum of the independent poisson random variables is also a poisson random variable

$\therefore$ It is always +vely skewed

1) A telephone switchboard receives 20 calls on an average during an hour. Find the probability that for a period of 5 minutes

arrived

$\therefore$ poisson

1) no call is received
2) exactly three calls are received
3) atleast two calls are received

Sol:

for 60 min $\lambda = 20$

1 min $\Rightarrow \lambda = \frac{20}{60}$

$\therefore$ for 5 min $\lambda = 5 \times \frac{20}{60} = 1.65$

1) $P(X=0) = \frac{e^{-1.65}(1.65)^0}{0!} = e^{-1.65}$

2) $P(X=3) = \frac{e^{-1.65}(1.65)^3}{3!}$

3) $P(X \geq 2) = 1 - P(X<2)$

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

$= 1 - e^{-1.65} - e^{1.65}(1.65)$

) X_1 and X_2 are two independent poisson random variables with variances 1, 2. Find ~~$P(X_1) + P(X_2) = 4$~~ $P(X_1 + X_2) = 4$

sum of random variables can be done using MGF

sum of independent random variables (poisson) is also a poisson random variable

$$P(X_1+X_2=K) = \frac{e^{-(\lambda_1+\lambda_2)} \cdot (\lambda_1+\lambda_2)^K}{K!}$$

$$P(X_1+X_2=4) = \frac{e^{-(1+2)} (1+2)^4}{4!} = \frac{81 e^{-3}}{24}$$

P) X and Y are independent poisson random variables such that $P(X=1) = P(X=2)$ and $P(Y=2) = P(Y=3)$ find $V(3X-4Y)$

Sol: $P(X=1) = P(X=2)$ and λ be parameter

$$\Rightarrow \frac{e^{-\lambda} \cdot \lambda^1}{1!} = \frac{e^{-\lambda} \lambda^2}{2!} \Rightarrow \lambda = 2$$

$\therefore$ variance = mean = 2

$V(X) = E(X) = 2$

$P(Y=2) = P(Y=3)$ and θ be parameter

$$\frac{e^{-\theta} \theta^2}{2!} = \frac{e^{-\theta} \cdot \theta^3}{3!} \Rightarrow \theta = 3$$

$\therefore V(Y) = E(Y) = 3$

$$V(3X-4Y) = 9V(X) + 16V(Y)$$
$$= 9 \cdot 2 + 16 \cdot 3 = 66$$

P) If 'X' is a poisson random variable and $E(X^2) = 6$ find $V(X)$.

Sol: we have $E(X) = V(X) = \lambda$

But $V(X) = E(X^2) - (E(X))^2$

$$\lambda = 6 - \lambda^2 \Rightarrow \lambda^2 + \lambda - 6 = 0$$
$$\lambda = +2, \cancel{-3}$$

$$\therefore V(X) = \lambda = 2$$

Normal distribution (Gaussian distribution)

- It is a continuous distribution

If 'x' is said to be a normal random variable, it allows the values from $(-\infty, +\infty)$ with mean $= E(x) = \mu$ and variance $= V(x) = \sigma^2$ then the random variable is known as normal random variable and its density function is

$$N(x; \mu, \sigma^2) = f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad \begin{array}{l} -\infty < x < +\infty \\ -\infty < \mu < +\infty \\ 0 < \sigma < \infty \end{array}$$

$$= 0 \quad \text{otherwise}$$

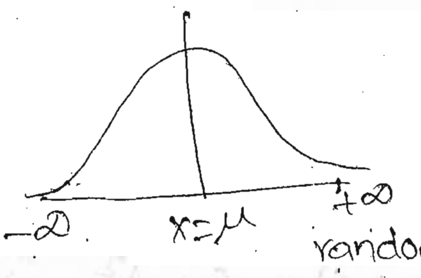
standard normal random variable :

If 'x' is a normal random variable with mean=0 and variance = 1 then the random variable is known as standard normal variable and its density function is

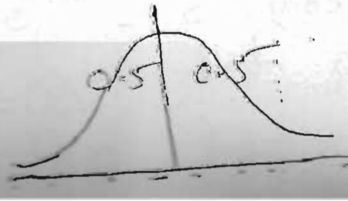
$$N(x; 0, 1) = f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

mathematically, standard normal variate is denoted by z and is $z = \frac{X - E(X)}{\sqrt{V(X)}}$; $-3 \leq Z \leq +3$

normal curve



standard normal curve



Areas under standard normal curve :

1) $P(Z \le z_0) = 0.5 + A$ $\quad z_0$ is +ve

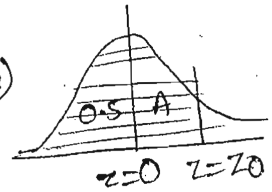
2) $P(Z \le z_0) = 0.5 - A$ $\quad z_0$ is −ve

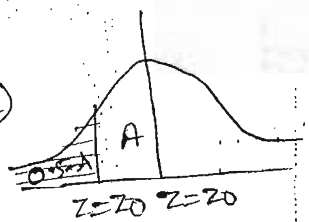
3) $P(z_1 \le Z \le z_2) = A_1 + A_2$ $\quad$ if z_1 is −ve and z_2 is +ve

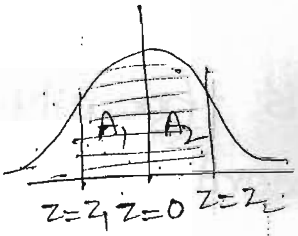
4) $P(z_1 \le Z \le z_2) = A_2 - A_1$ $\quad$ if z_1 and z_2 are +ve (or) z_1 and z_2 are −ve

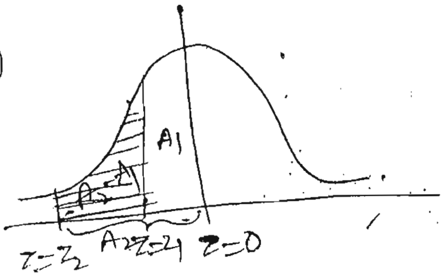
5) $P(Z \ge z_0) = 0.5 + A$ $\quad z_0$ is −ve

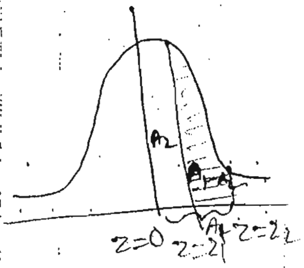
6) $P(Z \ge z_0) = 0.5 - A$ $\quad z_0$ is +ve

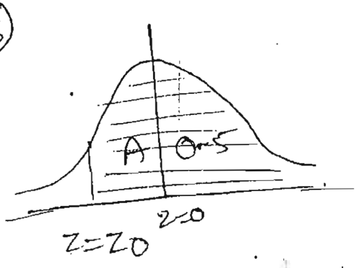
1) 

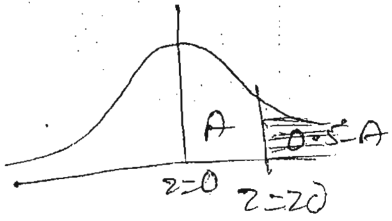
2) 

3) 

4) 

(OR) 

5) 

6) 

P) X is normally distributed with mean = 20 and standard deviation 3.33 find the probability between 21.11 and 26.66. The area under the normal curve $z=0$ to $z=0.33$ is 0.1293 and $z=0$ to $z=2$ is 0.4772

Sol:

Sol: $\mu = E(x) = 20$, $\sigma = 3.33$

$$P(\underset{x_1}{21.11} \le x \le \underset{x_2}{26.66}) = ?$$

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{21.11 - 20}{3.33} \doteq 0.33$$

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{26.66 - 20}{3.33} = 2$$

$$\therefore P(21.11 \le x \le 26.66) = P(0.33 \le z \le 2)$$

$$= 0.4772 - 0.1293$$

$$= 0.3479$$

) x is normally distributed with mean $= 30$ and standard deviation $= 5$. find $P(|x-30| \le 5)$. The area under normal curve $z=0$ to $z=1$ is 0.3413

Sol. $E(x) = \mu = 30$, $\sigma = 5$

$$P(|x-30| \le 5) = P(-5 \le x - 30 \le +5)$$

$$= P(\underset{x_1}{25} \le x \le \underset{x_2}{35})$$

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{25-30}{5} = -1$$

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{35-30}{5} = +1$$

$$\therefore P(|x-30| \le 5) = P(-1 \le z \le 1)$$

$$= 0.3413 + 0.3413$$

$$= 0.6826$$

properties:

$E(x) = \text{mean} = \mu$

$V(x) = \text{variance} = \mu_2 = \sigma^2$

$\mu_3 = 0$

$\beta_1 = 0$

MGF: $M_X(t) = e^{t\mu + \frac{t^2\sigma^2}{2}}$

CF: $\phi_X(t) = e^{it\mu - \frac{t^2\sigma^2}{2}}$

Note:

1. The sum of the independent normal random variables is also a normal random variable
2. The difference between the independent normal random variables is also a normal random variable. (Linear combination)

2/1/12

Uniform distribution (rectangular distribution):

- It is continuous distribution
- It is the only distribution that has constant probability function.
- If 'x' is said to be a uniform random variable over the interval (a, b) $a<b$ and its probability density function is

$$U(a,b) = f(x) = \frac{1}{b-a} \qquad a<x<b, \; a<b$$

$$= 0 \qquad \text{otherwise}$$

Note:

1. If 'x' is said to be a uniform random variable over the interval $(-a, a)$, its density function is

$$f(x) = \frac{1}{2a} \qquad -a<x<a$$

$$= 0 \qquad \text{otherwise}$$

Properties:

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

$V(X) = \mu_2 = \frac{(b-a)^2}{12}$

$\mu_3 = 0$

$\beta_1 = 0$

- MGF : $M_X(t) = \dfrac{e^{bt} - e^{at}}{t(b-a)}$

CF : $\phi_X(t) = \dfrac{e^{ibt} - e^{iat}}{it(b-a)}$

Note:

· The sum of the independent uniform random variables is also a uniform random variable

) If 'x' is uniform random variable, with mean 1 and variance is 4/3 find $P(X<0)$

sol: Mean = $E(X) = 1$

$\dfrac{a+b}{2} = 1 \Rightarrow a+b = 2$

variance = $V(X) = \dfrac{4}{3}$

$\dfrac{(b-a)^2}{12} = \dfrac{4}{3} \Rightarrow b-a = \pm 4$

but $b > a$

$\therefore b - a = 4$

$b + a = 2$

$2b = 6 \Rightarrow b = 3$

$a = -1$.

Interval is $(-1, 3)$

$U(-1,3) = f(x) = \dfrac{1}{3-(-1)} = \dfrac{1}{4}$

$P(X<0) = \int_{-1}^{0} f(x)\,dx$

$= \int_{-1}^{0} \frac{1}{4}\,dx = \frac{1}{4}$

) If 'x' is uniform random variable over the interval (i.e. $(-\alpha, \alpha)$ $(\alpha>0)$ find α such that $P(|X|<1) = P(|X| \geq 1)$

sol: $P(|X|<1) = P(|X| \geq 1)$

$= 1 - P(|X|<1)$

$2P(|x|<1) = 1$

$P(-1 < x < +1) = \frac{1}{2}$

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

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

$\frac{1}{\alpha}(1+1) = 1 \Rightarrow \alpha = 2$

Exponential distribution:

- continuous distribution
- used for measuring service rate / rate of work done
- If 'x' is said to be exponential random variable defined in the interval $0 \leq x < \infty$ with a parameter $\theta\,(>0)$ and its probability density function is

$$E(x;\theta>0) = f(x) = \theta e^{-\theta x} \quad \theta>0,\ 0 \leq x < \infty$$

$$= 0 \quad \text{otherwise}$$

Properties:

$\text{Mean} = E(x) = \frac{1}{\theta}$

$\text{variance} = V(x) = \mu_2 = \frac{1}{\theta^2}$

$V = \frac{\text{Mean}(M)}{\theta}$

$V = M$ if $\theta = 1$

$V > M$ if $0 < \theta < 1$

$V < M$ if $\theta > 1$

$\mu_3 = \frac{2}{\theta^3}$

$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \left(\frac{2}{\theta^3}\right)^2 \times \left(\frac{\theta^2}{1}\right)^3 = 4$

+vely skewed

MGF: $M_X(t) = \left(1-\frac{t}{\theta}\right)^{-1}$

CF: $\phi_X(t) = \left(1-\frac{it}{\theta}\right)^{-1}$

Note:

1. The sum of independent exponential random variables is a gamma random variable with two parameters and its mean and variance are

$$E(X) = \frac{n}{a}$$

$$V(X) = \frac{n}{a^2}$$

2. This distribution can also be known as memory less distribution (or) forgetfullness distribution (It is measured by conditional probability)

1) If 'x' is a exponential random variable with mean 2 find ① $P(X>10)$ ② $P(X>13/P.X\geq 11)$

Sol: Mean = 2

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

$f(x; \theta=\frac{1}{2}) = \frac{1}{2}e^{-\frac{x}{2}}$

1) $P(X>10) = \int_{10}^{\infty} f(x)\,dx$

$= \int_{10}^{\infty} \frac{1}{2}e^{-x/2}\,dx = \frac{1}{2}\cdot\left[\frac{e^{-x/2}}{-\frac{1}{2}}\right]_{10}^{\infty}$

$= -\left[0 - e^{-5}\right] = e^{-5}$

2) $P(X>13/PX\geq 11)$

$= P(X>2) = \int_{2}^{\infty} f(x)\,dx$

$= \int_{2}^{\infty} \frac{1}{2}e^{-x/2}\,dx = e^{-1} = \frac{1}{e}$

P) When a database requests arrives at a server, it passes in sequence through 5 processes. Each processor follows an exponential random variable with the parameter 2.5. Let 'T' be the total service time occurred by the request in server, find mean and variance of T.

Sol: Here T represents sum of the 5 independent exponential random variables

$\therefore$ T is a gamma random variable with the parameters $\theta = 2.5$, $n = 5$

$$E(T) = \frac{n}{a} = \frac{5}{2.5} = 2$$

$$V(T) = \frac{n}{a^2} = \frac{5}{(2.5)^2} = 0.8$$

$\rightarrow$

Baye's theorem problems:

P) A die is rolled. If the number is odd, find the probability of prime number.

Sol: odd number $\rightarrow$ known event

Prime number $\rightarrow$ unknown event

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

$$P(O) = 3/6$$

$$P(P/O) = \frac{P(P \cap O)}{P(O)} = \frac{2/6}{3/6} = 2/3$$

P) A number is chosen from 100 numbers $\{00, 01, 02, \ldots 99\}$ Let 'x' denotes the sum of digits on the number and 'y' denotes the product of digits on the number.

Find the probability that $x = 9$ given $Y = 0$

Sol: $$P(X=9/Y=0) = \frac{P(X=9 \cap Y=0)}{P(Y=0)} = \frac{2/100}{19/100} = \frac{2}{19}$$

(00, 01, 02, 03, 04, 05, 06 - 09, 10, 20, 30, 40, - 90)

1) There are 3 coins. of these two are unbiased. one is a biased coin with two heads. A coin is drawn at random and tossed two times. It appears head on both times. Find the probability, that the selected coin is a biased coin

Sol:

$$\underline{UB} \qquad \underline{B}$$

$$2 \qquad 1 = 3$$

$$P(UB) = \frac{2}{3} \qquad P(B) = \frac{1}{3}$$

E: Getting a head 2 times

$$P(E/UB) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

$$P(E/B) = 1 \cdot 1 = 1$$

$$P(E) = P(UB \cap E) + P(B \cap E)$$

$$= P(UB) \cdot P(E/UB) + P(B)\, P(E/B)$$

$$= \frac{2}{3} \cdot \frac{1}{4} + \frac{1}{3} \cdot 1 = \frac{1}{2}$$

$$\text{Required} = P(B/E) = \frac{P(B \cap E)}{P(E)} = \frac{\frac{1}{3} \cdot 1}{\frac{1}{2}} = \frac{2}{3}$$

2) ~~playno~~ player A is speaking ~~through~~ truth 4 out of 5 times. A die is rolled ~~aedo~~. He reports that it is a 2. what is the probability that actually there was a 2

Sol:

$$P(T) = \frac{4}{5} \qquad P(F) = \frac{1}{5}$$

E: Reporting a 2.

$$P(E/T) = \frac{1}{6}$$

$$P(E/F) = \frac{5}{6}$$

$$P(E) = P(E \cap T) + P(E \cap F)$$

$$= P(T)\, P(E/T) + P(F) \cdot P(E/F)$$

$$P(T/E) = \frac{P(T\cap E)}{P(E)} = \frac{\frac{4}{5}\cdot\frac{1}{6}}{\frac{9}{30}} = 4/9$$

P) A letter is known to have come from either ~~Tatanagar~~ TATANAGER (or) CALCUTTA. on the envelope just two consecutive letters TA are visible. Find the probability that the letter has come from JATANAGER

Sol: $P(T) = \frac{1}{2}$ $\quad P(C) = \frac{1}{2}$.

E: Getting TA

$P(E/T) = \frac{2}{8}$

$P(E/C) = \frac{1}{7}$

$\overset{8}{T}\overset{7}{A}\overset{6}{T}\overset{5}{A}\overset{4}{N}\overset{3}{A}\overset{2}{G}\overset{1}{E}R$

$\overset{1}{C}\overset{2}{A}\overset{3}{L}\overset{4}{C}\overset{5}{U}\overset{6}{T}\overset{7}{T}A$

$$P(E) = P(E\cap T) + P(E\cap C)$$
$$= P(T)\,P(E/T) + P(C)\,P(E/C)$$
$$= \frac{1}{2}\cdot\frac{2}{8} + \frac{1}{2}\cdot\frac{1}{7} = \frac{11}{56}$$

$$P(T/E) = \frac{P(T\cap E)}{P(E)} = \frac{\frac{1}{2}\cdot\frac{2}{8}}{\frac{11}{56}} = 7/11$$

P) In answering a question on multiple choice test, a student either knows the answer (or) guess the answer. Let 'p' be the probability that student knows the answer to a question and '1-p' be the probability that student guess the answer. Assume that if the student guess the answer to a question, will be correct with the probability $\frac{1}{5}$. what is the conditional probability that if student knew the answer to question given that he answered it correctly.

Sol: $P(K) = P \quad P(G) = 1-P$

E: Answering correctly

$P(E/K) = 1$

$P(E/G) = \frac{1}{5}$

$P(E) = P \cdot 1 + (1-P)\frac{1}{5} = \frac{4P+1}{5}$

$$P(K/E) = \frac{P(K \cap E)}{P(E)}$$

$$= \frac{P \cdot 1}{\frac{4P+1}{5}} = \frac{5P}{4P+1} \quad 0 \leq P \leq 1$$