Probability

- Probability is the likelihood or the chance that an uncertain event would occur. Probability is a number that ranges from 0 to 1. 0 for an event which cannot occur. 1 for an event certain to occur.

- Need: To indicate the chances that something will occur, we use its probability and assign a certain value it. Thus, probability is both the language and measure of uncertainty. The probability concepts help to transform a problem into a logical model and the decision maker can use probabilities as an aid to its solution.

$$P(A) = \frac{\text{No of favourable Outcomes}}{\text{Total No. of Outcomes}}$$

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

- Events → Equally likely $(P(A) = P(B))$

  → Mutually Exclusive / Disjoint (can't occur together, $A \cap B = \phi$)

  → Exhaustive $(A \cup B \cup C = \text{sample space})$

  → Independent (A has no effect on occurrence of B)

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

  if A and B are mutually exclusive:

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

  if A and B are exhaustive:

  $$P(A \cup B) = 1 \text{ as } n(A \cup B) = n(S)$$

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

- Conditional Probability

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

  given that B has occured what are the chances of A occuring

- Multiplication theorem: $P(A \cap B) = P(B) \times P(A/B) = P(A) \times P(B/A)$

If A and B are independent

$$P(A/B) = P(A)$$

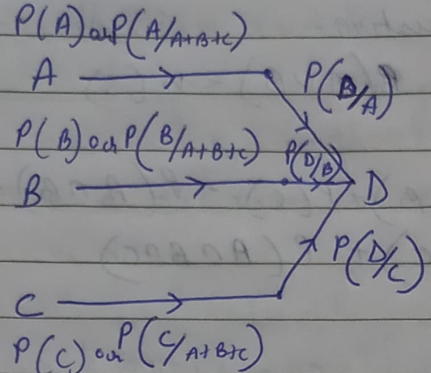
$$\Rightarrow P(A \cap B) = P(A) \times P(B)$$

If A, B and C are independent

$$P(A \cap B \cap C) = P(A) \times P(B) \times P(C)$$

~~Bayes Total Probability~~

- Bayes theorem:



Total Probability = $P(A) \times P(D/A) + P(B) \times P(D/B) + P(C) \times P(D/C)$

Probability that route followed for D is A =

$$P(A/D) = \frac{P(A) \times P(D/A)}{\text{Total Probability}}$$

- Random Variable : A variable is random if it takes different values as a result of the outcomes of a random experiment. eg: ~~afo~~ before tossing a coin or dice we can't predict the result.

- Continuous random variable is a random variable that ~~has only cont~~ can assume any value in given range. eg: Time taken to complete a 60 minutes test lies in the range 0 to 60, $0 \leq x \leq 2$

- Discrete random variable can assume only discrete values. $X = 1, 2, 3 \ldots$

Discrete	Continuous
Countable	Uncountable
Discrete Points	Continuous Intervals
Measured at Exact points	Measured over intervals
Points have probability	Points have no probability

- Probability Distribution: Distribution which are not obtained by actual observation or experiments but are deduced mathematically under certain assumptions. It is also called theoretical frequency distribution or modal or expected frequency distribution.

Probability Distribution

- Discrete
  - Binomial
  - Poisson
- Continuous
  - Exponential
  - Normal

- Binomial: number of trials $= n$, number of success $= r$
  success $= p$ and failure $= q$
  $p+q=1$, both are mutually exclusive and all trials are independent
  probability of p and q remain constant
  $n-r =$ number of failure

$$P(r) = {}^{n}C_{r}\, p^{r} q^{n-r} = \frac{n!}{(n-r)!\,r!} \cdot p^{r} \cdot q^{n-r}$$

mean = np   S.D = $\sqrt{npq}$ variance = npq

- Both mean and variance are finite, regardless of the values of n and p

- As the number trials increase, distribution becomes more bell-shaped and approaches a normal distribution.

- As the probability of success p approaches 0 or 1, the distribution becomes more peaked and less spread out with a smaller variance.

- Poisson: If value of n is very large and the value of p is too small and np is finite number then Binomial is not suitable. So poisson is suitable when successful events are few

$$P(r) = \frac{e^{-m} \cdot m^r}{r!}$$

where m = mean = np and e = 2.71828 (exponent)

- value of both mean and variance = m i.e. np

- It is always positively skewed but degree of skewness decreases with increase in mean (expected value)

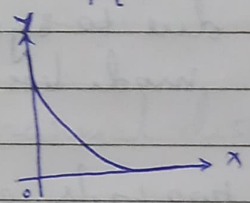
- Exponential: It is useful in describing the completion time of a task of the interval between two tasks.

$$f(x) = \frac{1}{m} e^{-(x/m)} \quad , \; m = \text{mean} \; , \; x = \text{time interval}$$

- Poisson and exponential are interrelated as poisson ~~do~~ tells about number of occurrences during an interval while exponential tells about the interval between two occurrences.

- In exponential, S.D = mean , mean $= \frac{1}{m}$

- $f(x \leq x_0) = 1 - e^{-(x/m)}$
- $f(x \geq x_0) = me^{-(x/m)}$



- The average time between events is the reciprocal of the rate parameter

- Normal (Gaussian Distribution): In this the variable can take on any value within a given range and in which the probability distribution is continuous.

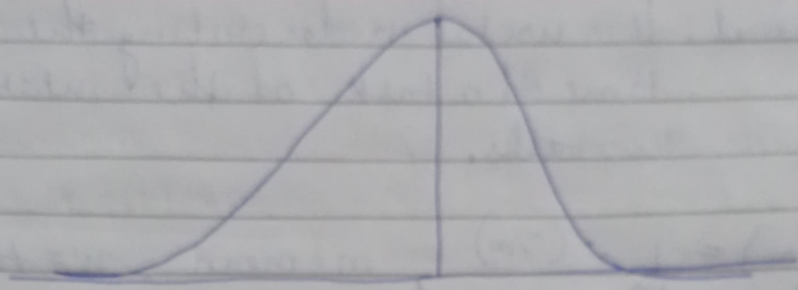
- it is used as approximation to other probability distribution.

$$y(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-m)^2/2\sigma^2}$$

$e = 2.71828$ $\quad$ σ = S.D. $\quad$ x = particular value of random variable

$\pi = 3.14$

m = expected value / mean $\quad$ $y(x)$ = height of curve at x.



- curve has a single peak due to which it is bell shaped and unimodal.

- mean of normally distributed curve lies at centre.

- due to symmetry of curve mean, median and mode lies at the centre.

- two tails extend indefinitely and never touch horizontal axis.

- area under curve is 1

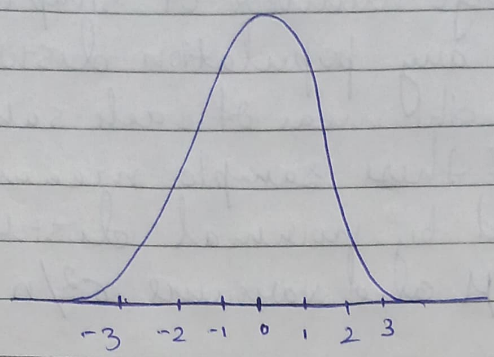
- mean affects position while S.D affects height

- Normal distribution with m = 0 and S.D = 1 is Standard Normal Distribution. It is possible to convert normal into standardized variable z. This is called z-transformation

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

- In Standard Normal Distribution intervals increase by 1

- Z-score tells us how many S.D an observation is from the mean.



Z-score of -2 tells us we are 2 S.D. left of the mean

- it also allows us to calculate how much area that specific ~~data~~ z-score is associated with.

- Empirical Rule:
  - 68% of all ~~data~~ the values in a normal distribution lie within ±1 S.D from mean
  - 95.5% of all the values lies with ±2 S.D from mean.
  - 99.7% of all values lie within ±3 S.D from mean.

- Chebyshev's Theorem:
S.D. measures the variation among observations. If S.D value is small, then values are closer to the mean. This theorem allows us to determine the proportion of observations that fall within a specified number of S.D from the mean value.

$$\% \text{ of observations} = 1 - \frac{1}{z^2}, \quad z = \text{z-score or distance of S.D from mean.}$$

- Central Limit Theorem:

If we select a large number of simple random samples from any population distribution and determine the mean of each sample, the distribution of these sample means will tend to be described by normal distribution with a mean μ and variance σ^2/n

Utility is that it requires no conditions on distribution pattern of the individual random variable being summed.