## Bernouli Distribution:

Descrete probability distribution of a random variable which takes the value 1 with probability p and the value 0 with probability $q = 1-p$

In a easy way it can be said that,

It is a model for the set of possible outcomes of any single experiment that asks a yes-no question.

Key things to remember in Bernouli Distribution:-

1. Discrete Random Variable {PMF}
2. Outcomes are binary [(0,1), (Head, tail), (yes, no)]
3. Example- Tossing a coin.

   $Pr(Head) = 0.5$ (suppose p)

   $Pr(Tail) = 1-0.5 \; (1-p)$

4. Example→ Pass or fail in exam

   $Pr(Pass) = p = 0.7$

   $Pr(Fail) = q = 1-p = 1-0.7 = 0.3$

## Parameters:

$0 \leq p \leq 1$

$q = 1-p$

$K = \{0,1\}$ [yes or no]

① **PMF:** (Probability Mass Function)

$$PMF = p^K * (1-p)^{1-K}$$

Here, $K \in \{0,1\}$

if $K=1$, $P_r(K=1) = p^1 * (1-p)^{1-1}$

$= p * 1$

$= p$

$p$ = probability of success

$n$ = no of trials

$K$ = no of success.

if, $K=0$ then $P_r(K=0) = p^0 * (1-p)^{1-0}$

$= 1-p$

**PMF simplified:**

$$PMF \begin{cases} q = 1-p & \text{if, } K=0 \\ p & \text{if, } K=1 \end{cases}$$

**Mean of Bernouli Distribution:**

$$E(K) = \sum_{i=0}^{K} K \cdot P(K)$$

if, suppose

$P_r(k=1) = 0.6$

$P_r(k=0) = 0.4$

$= [0 \times P(0) + 1 \times P(1)]$

$= [0 \times 0.4 + 1 \cdot (0.6)]$

$= 0.6$

$= p$

So, $p$ is the mean of Bernouli Distribution

CamScanner

## Median of Bernoulli Distribution:

$$\text{Median} \begin{cases} 0 & \text{if } p < 0.5 \\ [0,1] & \text{if } p = 0.5 \\ 1 & \text{if } p > 0.5 \end{cases}$$

Based on problem Statement

## Variance:

$$Var = p*(1-p)$$
$$= p*q \quad [q = 1-p]$$

## Std:

$$Std = \sqrt{pq}$$

## Binomial Distribution: $B(n,p)$

The binomial distribution with parameters n and p is the discrete probability distribution of the number of successes in a sequence of n independent experiments, each asking yes-no question, and each with it's own boolean valued outcome: Success (p) or failure $(q = 1-p)$

## Parameters:

$n$ = number of trials

Suppose I am tossing a coin 7 time

So, trial $n = 7$

$p$ = probabity $[0,1]$ → Success probability for each trial

$q = p - 1$

Key things to note:

⓪ For discrete random variable (PMF Function)

① Every outcome is binary

② This experiment is performed for $n$ trials. where each trail is a bernouli distribution

③ Every single trial from the $n$ trial is called Bernouli distribution.

**Support:** $K = \{1, 2, 3, \ldots n\}$ → Number of success for $n$ trials

**PMF:** $Pr(K, n, p) = {}^{n}C_{K}\, p^{K} (1-p)^{n-K}$ $\left\{ K \in 0 \xrightarrow{to} n \right.$

**Mean:** $np$, **Variance:** $npq$ **std:** $\sqrt{npq}$

**Poisson Distribution:**

→ For Discrete Random variable (PMF Function)

→ Describes the number of events in a fixed time interval.

Example: Number of people ($n$) visiting hospital every hour
↳ fixed time

## Sampling Distribution:

Sample distribution is a probability distribution that describes the statistical properties of a sample statistic (such as the sample mean or sample proportion) computed from multiple independent samples of the same size from a population.

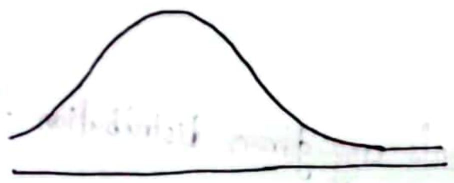
Suppose you have the population dataset of Bangladesh which is about salaries of people.

Sample 1 [x x x . . . . . . .] → 50 people → $\overline{x_1}$

Sample 2 [x x x x . . . . . .] → 50 people → $\overline{x_2}$

⋮

Sample n [x x x . . . . . .] → 50 people → $\overline{x_3}$



Now you are taking samples from the population where each sample has 50 people in it. Like this you take 100 samples. Suppose $\overline{x_1}$ is the mean of the first sample. Similarly you will get 100 mean values from 100 sample list. You can also get the variance, std etc. Their distribution (mean, median, variance, std, etc) will be called Sampling Distribution.

## Central Limit Theorem:

The mean sets we have taken from the sampling distribution, that would be Normally distributed. No matter how the data is distributed. After we get the mean set from the n number of samples, those mean values together will be normally distributed. That's what the central limit theorem says.

But some conditions has to be followed:

1) Every sample size should be large enough ($\geq 30$)

2) The sample is drawn from a finite population or infinite population with a finite variance

3) The random variables in the sample are independent and identically distributed.

if population mean and variance was $(\mu, \sigma^2)$
then the mean and variance of the Normal Distribution will be $\left(\mu, \frac{\sigma^2}{n}\right)$

($n$ = sample size)