# Binomial Probability Distribution

---

---

In binomial probability distribution, the number of `Success` in a sequence of n experiments, where each time a question is asked for yes-no, then the boolean-valued outcome is represented either with success/yes/true/one (probability p) or failure/no/false/zero (probability q = 1 − p).

---

## Binomial Distribution Formula

**P(x:n,p) = {n!/(x!(n-x)!)} .px.(q)n-x**

Where,

n = the number of experiments

x = 0, 1, 2, 3, 4, ...

p = Probability of Success in a single experiment

q = Probability of Failure in a single experiment = 1 – p

---

++ ## Binomial Distribution Vs Normal Distribution

The main difference between the binomial distribution and the normal distribution is that binomial distribution is discrete, whereas the normal distribution is continuous. It means that the binomial distribution has a finite amount of events, whereas the normal distribution has an infinite number of events. In case, if the sample size for the binomial distribution is very large, then the distribution curve for the binomial distribution is similar to the normal distribution curve.

---

REF: Binomial Distribution - Definition, Formula & Examples | Probability (byjus.com)

# Summary

---

---

## Statistics

The statistical analysis is meant to collect and study the information available in large quantities.

### Descriptive Statistics

In this type of statistics, the data is summarised through the given observations. The summarisation is one from a sample of population using parameters such as the mean or [standard deviation].

Descriptive statistics are also categorised into four different categories:

- Measure of frequency
- Measure of dispersion
- Measure of central tendency
- Measure of position

The frequency measurement displays the number of times a particular data occurs. Range, Variance, Standard Deviation are measures of dispersion. It identifies the spread of data. Central tendencies are the mean, median and mode of the data. And the measure of position describes the percentile and quartile ranks.

REF : Types of Statistics (Descriptive & Inferential) (byjus.com)

# The distribution of The data

**1. Right-skewed** : Mean greater than Median

**2. Left-skewed** : Mean less than Median

**3. Symmetric** (frequently normally distributed) : Mean equals Median

The `mode` of a distribution is essentially the tallest bar in a histogram.

## Outliers

`outliers` are points that fall very far from the rest of our data points. This influences measures like the mean and standard deviation much more than measures associated with the five number summary.

## Identifying Outliers

1. Sorting your values from low to high and checking minimum and maximum values.
2. Visualizing your data with a box plot and looking for outliers.
3. Using the interquartile range to create fences for your data.

---

# Simpson's Paradox

`Trust No one - certainty is a LIE`

*Simpson's paradox* occurs when groups of data show one particular trend, but this trend is reversed when the groups are combined together.

---

---