Inferential Statistics

Probability

Permutations

A permutation is a way of arranging a selected group of objects in such a way that the order is of significance. The list below shows some examples where permutation is used to count the number of ways in which a particular sequence of events can occur:

- All possible four-letter words that can be formed using the alphabets R, E, A and D is $P(4,3)$.
- All possible ways in which the final league standings of the eight teams can be in an IPL tournament 8!
- All possible ways that a group of 10 people can be seated in a row in a cinema hall is $10!$
- And so on.

If there are n objects that are to be arranged among r available 'spaces', then the number of ways in which this task can be completed is $P(n,r) = \frac{\text{n!}}{(n-r)!}$. If there are n 'spaces' as well, then the number of ways would be just $n!$. Here $n!$ (Pronounced as n factorial) is simply the product of all the numbers from n till 1 and is given by the formula $n! = n \times (n-1) \times (n-2) \times \ldots \times 3 \times 2 \times 1$.

Combinations

Combinations is the number of ways of selecting r elements out of n distinct objects where the order is not important.

Some other examples of combinations are as follows:

- The number of ways in which you can pick three letters from the word UPGRAD is $C(5,3) = \frac{5!}{3!(5-3)!}$
- The number of ways a team can win three matches in a league of five matches is $C(5,3) = \frac{5!}{3!(5-3)!}$.
- The number of ways in which you can choose 13 cards from a deck of 52 cards is $C(52,13)$.
- And so on.

The formula for counting the number of ways to choose r objects out of a set of n objects is $C(n,r) = \frac{n!}{r!(n-r)!}$.

Probability

The formula for calculating the probability is Probability

$$= \frac{Number\ of\ desired\ outcomes}{Total\ number\ of\ possible\ outcomes}$$

Example: Suppose a coin is tossed twice. What is the probability of getting head twice?

Answer: If a coin a tossed twice, then the number of all possible of outcomes is $\{(Tail, Tail), (Tail, Head), (Head, Tail), (Head, Head)\}$. As per the question, occurring head twice happens only once. So, the probability of getting head is $= \frac{1}{4}$.

Probability values have the following two major properties:

- Probability values always lie in the range $[0,1]$. The value is 0 in the case of an impossible event (like the probability of you being in Delhi and Mumbai at the same time) and 1 in the case of a sure event (like the probability of the sun rising in the east tomorrow).
- The probabilities of all outcomes for an experiment always sum up to 1. For example, in a coin toss, there can be two outcomes, heads or tails. The probability of both of the outcomes is 0.5 each. Hence, the sum of the probabilities turns out to be $0.5 + 0.5 = 1$.

Some important concepts:

- **Experiment:** An experiment is essentially a repeatable scenario wherein a particular event may or may not occur. It is of two types:
  - Deterministic experiment: Outcome is the same every time.

    - Random experiment: Outcome can take many possible values. Throughout the majority of our business analytics course, we'll focus on random experiment only.
- **Sample Space:** A sample space is nothing but the list of all possible outcomes of a random experiment. It is denoted by S={x:x is a possible outcome}. For example, in the coin toss example, the sample space S={H,T}, where H=heads and T=tails.
- **Event**: It is a subset, i.e., a part of the sample space that you want to be true for your probability experiment. For example, if in a coin toss you want heads to be the desired outcome, then the event becomes {H}. As you can see clearly, {H} is a part of {H,T}

Types of Events

An event is an outcome or a set of outcomes of a random experiment.

- **Independent Events:** If you have two or more events and the occurrence of one event has no bearing whatsoever on the occurrence/s of the other event/s, then all the events are said to be independent of each other. For example, the chances of rain in Bengaluru on a particular day has no effect on the chances of rain in Mumbai 10 days later. Hence, these two events are independent of each other.
- **Disjoint or Mutually Exclusive Events:** Now two or more events are mutually exclusive when they do not occur at the same time, i.e., when one event occurs, the rest of the events do not occur. For example, if a student has been assigned grade C for a particular subject in an exam, he or she cannot be awarded grade B for the same subject in the same exam. So, the events that a student gets a grade of B or C in the same subject are mutually exclusive or disjoint.

Two or more events cannot be independent and disjoint simultaneously. Two events can be either exclusively independent or exclusively disjoint. Here are some examples.

- The events ‘Customer A buys the product’ and ‘Customer B buys the product’ are independent whereas the events ‘Customer A buys the product' and ‘Customer A doesn’t buy the product’ are disjoint.
- The events ‘You’ll win Lottery A’ and ‘You’ll win Lottery B’ are independent whereas ‘You’ll win Lottery A’ and ‘You won’t win Lottery A’ are disjoint events.

Complement Rule of Probability:

Disjoint events have one special property that is intuitive and easy to understand. For example, if A and B are 2 disjoint events, let’s say A=Event that it rains today and B=Event that it doesn′t rain today and you know the $P(A)=0.3$. It is nothing but $1-P(A)=1-0.3=0.7$. This is something known as the complement rule for probability. It states that if A and A′ are two events which are mutually exclusive/disjoint and are complementary/negation of each other (you can read A′ as not A), then $P(A)+P(A')=1$.

In the above example, B is the complement of A and hence $P(B)=1-P(A)=1-0.3=0.7$. Here are some examples, where you can use this rule to find the probability of the complement of an event.

- If the probability that a customer buys a product is 0.4, then the probability that he/she doesn’t buy the product is 0.6;
- If the probability that you win the lottery is 33% then the probability that you don’t win the lottery is 67%;
- And so on.

This rule is basically an extension from the basic rule of probability.

Addition Rule of Probability

When there are two individual probabilities of two events A and B, denoted by $P(A)$ and $P(B)$, the addition rule states that the probability of the event that either A or B will occur is given by $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, where $P(A \cup B)$ denotes the probability that either the event A or B occurs, $P(A)$ denotes the probability that only the event A occurs, $P(B)$ denotes the probability that only the event B occurs, and $P(A \cap B)$ denotes the probability that both the events A and B occur simultaneously.

Note: The symbols ∪ and ∩ are obtained from the world of 'set theory' and are used to denote union and intersection, respectively. You don't need to learn about them in detail right now. All you need to learn are the meanings of the probability terms mentioned above. Also, we would be skipping the proof of the formula right now. Another important thing to note here is that the formula given above works for all types of events A and B, irrespective of the fact that they're independent or disjoint, etc.

You can also read $P(A \cup B)$ as $P(either\ event\ A\ or\ B\ occurs)$ and $P(A \cap B)$ as $P(both\ events\ A\ and\ B\ occur)$.

So, for disjoint events A and B, $P(A \cap B) = 0$ since both cannot occur simultaneously. Hence, the formula can be rewritten as $P(A \cup B) = P(A) + P(B)$.

Example: What is the probability of drawing a card which is either a heart or a face card out of 52 cards in a deck?

Let, A = drawing a heart

B = drawing a face card

So, P(A) $= \frac{13}{52} = \frac{1}{4}$ and P(B) $= \frac{12}{52} = \frac{3}{13}$.

Now, $P(A \cap B) = P(A\ card\ which\ is\ heart\ as\ well\ as\ face\ card) = \frac{3}{52}$

So, $P(Drawing\ a\ heart\ or\ a\ face\ card) = P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{13}{52} + \frac{12}{52} - \frac{3}{52} = \frac{22}{52} = \frac{11}{26}$.

Multiplication Rule of Probability

When an event A is not dependent on event B and vice versa, they are known as independent events. And the multiplication rule allows us to compute the probabilities of both of them occurring simultaneously, which is given as $P(A \cap B) = P(A) \times P(B)$.

Now, this rule can be extended to multiple independent events where all you need to do is multiply the respective probabilities of all the events to get the final probability of all them occurring simultaneously. For example, if you have four independent events A, B, C and D, then $P(A \cap B \cap C \cap D) = P(A) \times P(B) \times P(C) \times P(D)$.

Example: What is the probability of drawing a card which is a heart and a face card out of 52 cards in a deck?

Let, A = drawing a heart

B = drawing a face card

So, $\mathrm{P(A)} = \frac{13}{52} = \frac{1}{4}$ and $\mathrm{P(B)} = \frac{12}{52} = \frac{3}{13}$.

Now, $P(A \cap B) = P(A\ card\ which\ is\ heart\ and\ face\ card) = P(A) \times P(B) = \frac{13}{52} \times \frac{12}{52} = \frac{3}{52}$.

Comparison between Addition Rule and Multiplication Rule

Both addition rule and multiplication rule allow you to compute the probabilities of the occurrence of multiple events. However, there is a key difference between the two, which should help to decide when to use which rule.

- The addition rule is generally used to find the probability of multiple events when either of the events can occur at that particular instance. For example, when you want to compute the probability of picking a face card or a heart card from a deck of 52 cards, a successful outcome occurs when either of the two events is true. This includes either getting a face card, a heart card, or even both a face and a heart card. This rule works for all types of events.
- The multiplication rule is used to find the probability of multiple events when all the events need to occur simultaneously. For example, in a coin toss experiment where you toss the coin three times and you need to find the probability of getting three heads at the end of the experiment, a successful outcome occurs when you get a head in the first toss, a head in the second toss and a head in the third toss as well. This rule is used for independent events only.

Three important approaches of probability:

- **Equal likelihood approach**
- **Relative frequency approach**
- **Judgemental approach**

- **The equal likelihood approaches**
    - If an experiment has n simple outcomes, then this method will assign a probability of 1n to each outcome.
    - For example, 16 is the probability that the upper surface of an ordinary die will display six pips when it is rolled, considering it is equally likely that each of the six surfaces of the die will be displayed.
- **The relative frequency approaches**
    - In this method, probabilities are assigned based on previous experiments or historical trends.
    - For example, life insurance companies will consider providing insurance to a young person over an old person. This is because intuition and numerous studies suggest that the likelihood of the survival of a young person is higher than that of an old person.

- **The judgemental approach**
    - Sometimes, the approaches might not necessarily be effective or too many factors might be involved in understanding the likelihood of an event. In such cases, you will have to rely on your own judgement in drawing a conclusion.
    - For example, before launching the world's first electric vehicle, Tesla had no way of predicting whether it will be a success or a failure. The company had to rely on their intelligent judgements about the market and the product.

Three Types of Probability

The three major types of probability are as follows:

- Marginal probability
- Joint probability
- Conditional probability

To understand the different types of probability, Ankit has considered the following sample of people and the fraction of them belonging to 'Male', 'Female','Emplòyed' or Unemployed' as follows:

	Men	Woman	Total
Employed	0.1	0.45	0.55
Unemployed	0.25	0.2	0.45
Total	0.35	0.65	1.00

To understand the three types of probability, let's consider a scenario in which there are two types of events: Event A and Event B.

Event A: The person is a 'Man'.

Event B: The person is 'Employed'.

Joint probability

- This is the probability of two or more events happening simultaneously.
- If A and B are two events, then the joint probability of the two events is written as $P(A \cap B)$. Essentially, it is the probability of Event A and Event B occurring simultaneously.
- In the below example, the joint probability of these two events is the probability of a person being both a man and employed; hence, it is 0.1.

Joint Probability of person being 'Male' and 'Employed'

	Men	Woman	Total
Employed	0.1	0.45	0.55
Unemployed	0.25	0.2	0.45
Total	0.35	0.65	1.00

Marginal probability

- This is the probability of an event happening irrespective of the outcomes of another event also happening simultaneously.
- Considering the same example as above, suppose we want to find the probability of a person being employed, irrespective of whether the person is a man or woman.
- Since this probability is independent of the gender, this probability is a marginal probability
- As shown in the table given below, this marginal probability is 0.55.

Marginal Probability
Probability of person being Employed

	Men	Woman	Total
Employed	0.1	0.45	0.55
Unemployed	0.25	0.2	0.45
Total	0.35	0.65	1.00

Conditional probability

- This is the probability of an event happening given that another event has already occurred.
- If A and B are two events, then the conditional probability of Event A occurring, given that Event B has occurred, is written as $P(A|B)$.
- The general multiplication rule links the three types of probability in the following manner:
  - $$\text{Conditional Probability of A given B} = P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{\text{Joint Probability of A and B}}{\text{Marginal Probability of B}}$$
  - $$\text{Conditional Probability of B given A} = P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{\text{Joint Probability of A and B}}{\text{Marginal Probability of A}}$$
- If A = Person being male and B = Person being Unemployed, then the probability of a person being unemployed given that he is a male $= \frac{0.25}{0.35} = 0.714$.

Joint Probability of person being 'Male' and 'Unemployed'

	Men	Woman	Total
Employed	0.1	0.45	0.55
Unemployed	0.25	0.2	0.45
Total	0.35	0.65	1.00

Marginal Probability of person being Male

Probability Distributions

Introduction to Probability

Example of statistics in everyday life:

- In cricket, a coin is flipped to decide which team will choose to bat first. There are two teams, and each team has a 50% chance of winning the toss. Hence, the probability of both the teams winning the toss is 0.5.
- The weather report for the day conveys that there is a 40% chance that your city will experience rain. This indicates that the probability of rain is 0.4.
- Now, probability also affects your day-to-day decision-making, some of which are as follows:
    - If the probability of rain is **high**, then you will choose to carry an umbrella with you.
    - If there is a **high probability** that the stock prices will increase, then you will decide to invest in that particular stock.

- You may be tempted to purchase a lottery ticket in the hope that you will win the lottery even after knowing that the chances or probability of you winning the lottery is **close to nil**.

What is a random variable?

A random variable is a variable whose values are the outcomes derived from a random experiment. Let's understand this using the example of a weather forecast. Considering the probability of rain is 0.4, the outcome of this experiment is that 'it will rain'.

Suppose X is the random variable.

then Probability of $(X = \text{'}raining\text{'}) = 0.4$

Here, the probability of the random variable X for the outcome 'raining' is 0.4. However, there is a 60% probability of no rain. Hence, the probability of the random variable $X = \text{'}it\ will\ not\ rain\text{'}\ is$ 0.6. Here, the random experiment is whether it'll rain or not.

What is a sample space?

The sample space represents all the possible values of a random variable of a particular experiment.

Tossing of a coin

What is the probability that the outcome of a coin toss will be Heads? The possible outcomes of this experiment are {Heads, Tails}. This is the sample space. Let's consider a random variable X such that,

X=0 if the coin surface is '**Heads**', and

X=1 if the coin surface is '**Tails**'.

Here, the coin toss is the random experiment.

Number of favourable outcomes =1 i.e., {Heads}

Total number of outcomes = sample space =2 i.e., {Heads,Tails}

Hence, $P(X = 0) = \frac{1}{2} = 0.5$

Let's take another example.

Rolling of a die

Let X represent the number of pips (or dots) that the upper surface of the die displays after it is rolled.

What is the probability of getting the outcome X=1?

If you roll the die, there is only one favourable outcome, i.e., the upper surface shows one pip. However, the total possible outcomes are {1,2,3,4,5,6}. This is also your sample space. Hence, there are total of 6 possible outcomes.

Probability Distribution

Upgrad Ball Game

In a bag, there are two balls out of which one is red, and another is blue. A ball is drawn from the bag and after noting down its color, it is again placed in the same bag. This process is repeated 4 times. If you get four red balls, then you will win Rs. 200 else Rs. 10 will be lost. If this game runs in a casino, then there are five different possibilities out of which only one outcome will help you to win. Rest of the possibilities will help the house to win the game.

we need to find out if the game will be profitable for the players or for us (i.e., the house) in the long run. The three-step process for this is:

1. Find all the possible combinations.
2. Find the probability of each combination.
3. Use the probabilities to estimate the profit/loss per player.

Let us define the random variable X where, X= **Number of red balls**

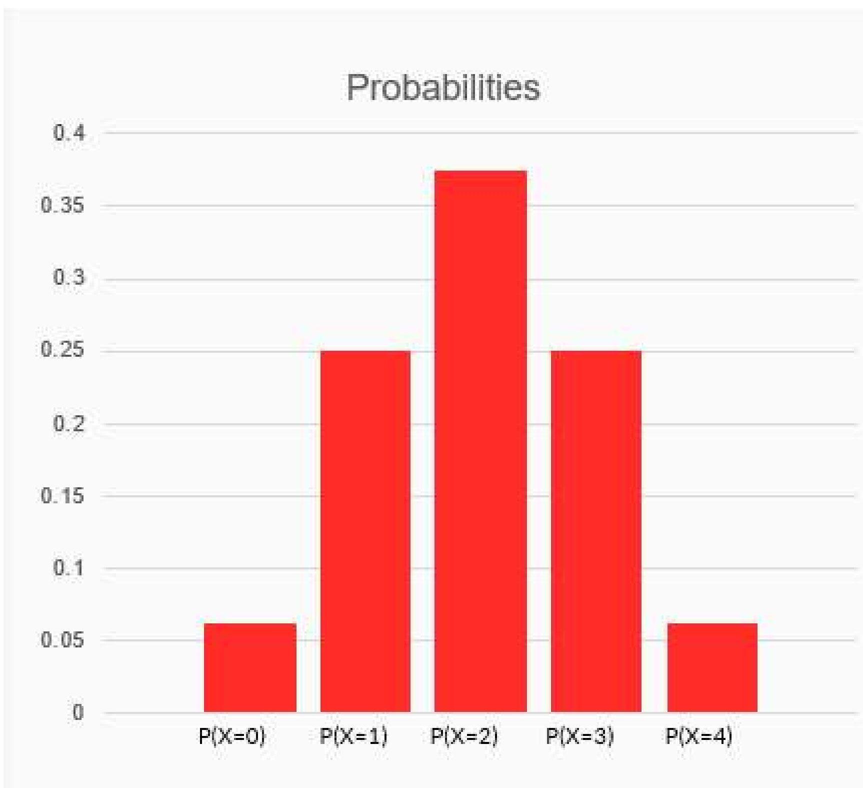
the random variable X can take a total of 5 values, i.e., 0, 1, 2, 3, and 4, and the probabilities of X taking each of these values came out to be:

- $P(X = 0) = 1 \times (\frac{1}{2}) \times (\frac{1}{2}) \times (\frac{1}{2}) \times (\frac{1}{2}) = \frac{1}{16}$ (0 red balls and 4 blue balls)
- $P(X = 1) = 4 \times (\frac{1}{2}) \times (\frac{1}{2}) \times (\frac{1}{2}) \times (\frac{1}{2}) = \frac{4}{16}$ (1 red ball and 3 blue balls)
- $P(X = 2) = 6 \times (\frac{1}{2}) \times (\frac{1}{2}) \times (\frac{1}{2}) \times (\frac{1}{2}) = \frac{6}{16}$ (2 red balls and 2 blue balls)

- $P(X = 3) = 4 \times (\frac{1}{2}) \times (\frac{1}{2}) \times (\frac{1}{2}) \times (\frac{1}{2}) = \frac{4}{16}$ (3 red balls and 1 blue ball)
- $P(X = 4) = 1 \times (\frac{1}{2}) \times (\frac{1}{2}) \times (\frac{1}{2}) \times (\frac{1}{2}) = \frac{1}{16}$ (4 red balls and 0 blue balls)

The number multiplied at the beginning of each equation indicates the total number of permutations possible for that particular outcome.

Furthermore, the plot of these probabilities to get a visual idea of what the probability of each outcome looks like:



Expected Value

Again, the three-step process we followed to find whether the upGrad red ball game was profitable for the players or for the house:

1. Find all the possible combinations.
2. Find the probability of each combination.
3. Use the probabilities to estimate the profit/loss per player.

The expected value of a variable X is the value of X you would 'expect' to get after performing the experiment once. It is also called the expectation, average, and mean value. Mathematically speaking, for a random variable X that can take values x_1, x_2, x_3,..., x_n, the expected value (**EV**) can be given as:

$$EV = x_1 \times P(X = x_1) + x_2 \times P(X = x_2) + x_3 \times P(X = x_3) + \ldots\ldots + x_n \times P(X = x_n)$$

The expected value should be interpreted as **the average value** you get after the experiment has been conducted **an infinite number of times**. For example, the expected value for the number of red balls is 2. This means that if we conduct the experiment (play the game) infinite times, the average number of red balls per game would end up being 2.

In this case, $EV = (200 \times 0.0625) + (-10 \times 0.9375) = 3.125$

So, if one person play the ball game 1000 times, then the casino will loose $3.125 \times 1000 = Rs. 3125$.

This is clearly a disaster for the house. To overcome this, any of the following three options can be tried out:

- Decrease the prize money.
- Increase the penalty.
- Decrease the player's chances of winning.

Method to calculate the expected value:

calculating the expected value is a three-step process:

1. Define the random variable (X).
2. Calculate the probability distribution $P(X)$. You'll need to calculate it on your own.
3. Plug the above two terms in the following formula:

$$E[X] = \sum(X \times P(X))$$

Discrete Probability Distribution

In the upGrad ball game described before, we define the random variable X as the number of red balls obtained, and the values that X could take were 0, 1, 2, and 3. Now, you might have noticed that these values are discrete, i.e., **in this case, the random variable** X **can only take discrete values and never take continuous values in between, such as, say 0.99 or 1.01**. Such events where the random variable can only take discrete values are represented using discrete probability distributions.

Uniform distribution: If all the outcomes of an event have equal probability, then it follows a uniform distribution. For example, rolling a die is uniformly distributed as the probability of all outcomes is $\frac{1}{6}$.

Cumulative Probability Distribution: Recalling the upGard ball game. As per the rule of the game, by drawing four red balls $P(X = 4)$ one person can win else the house will win. You might be interested in the probability of the win by the house only. In that case, if X is a random variable, then you need $P(X \leq 3)$. This is called cumulative probability.

So, cumulative probability $= F(X) = P(X \leq x)$

E.g. $F(2) = P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$

Binomial Distribution

Let's understand the binomial experiment with the help of an example. **Find the probability of getting Heads twice when tossing a coin five times**. Considering this example, the properties of a binomial experiment are as follows:

Note: One must understand the **difference between an event and an experiment**. Each coin toss is defined as a single event or a trial. Tossing this coin five times is the experiment. An experiment consists of various events. So, the probability of each event (success of failure) will be known to us.

- The experiment consists of a sequence of n **identical and independent trials**. Therefore, none of the trials affects the probability of the subsequent trial. In the coin-tossing experiment, each trial is a single coin toss, and each coin toss is independent of the other coin tosses.

- Essentially, each trial has two possible outcomes: **'Success' and 'Failure'**. In the coin-tossing experiment, there are two possible outcomes: 'Heads' and 'Tails'. Getting a 'Heads' will be counted as a success.

- **The probability of a success that is equal to p** is the same for each trial. Consequently, the **probability of a failure that is equal to $(1 - p)$** is constant for each trial. In the coin-tossing example, the probability of getting 'Heads' is $p = 0.5$, and probability of 'Tails' is $(1 - p) = 0.5$.

Calculating the binomial probability

The objective of this experiment is to find the **probability of getting exactly x successes in n trials**.

Let x be the random variable for this distribution, and it is defined as the number of successes in n trials. So, the values are as follows:

$P(x) =$ **Probability of x successes in** n **trials**

Note that x being a random variable can take **discrete values between** 0 **and $\boldsymbol{n}$**.

Using the example explained above, let's find the probability of getting two successes in five trials through the following steps:

- The first step is to find the different combinations of getting two successes and three failures. The 10 combinations that can be found are as follows:

$$\{SSFFF, SFSFF, SFFSF, SFFFS, FSSFF, FSFSF, FSFFS, FFSSF, FFSFS, FFFSS\}$$

- We can directly calculate the number of combinations using the combinations formula as follows:

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

Where n = number of trails, k = number of successes.

On substituting $n = 5$ and $k = 2$ from our example, we get the value 10.

- Now, let S = Success and F = Failure.

  As getting each combination is equally likely, we can directly calculate the probability using 10 multiplied by the probability of a single combination. So, the **probability of two successes in five trials will be equal to:**

$$= \mathbf{10 \times P(SSFFF)}$$

- As you learnt previously, for independent trials, the probability can be obtained by multiplying the probability at each trial. So, the equation will be as follows:

$$10 \times P(SSFFF) = 10 \times P(S) \times P(S) \times P(F) \times P(F) \times P(F)$$
$$= 10 \times P(S)^2 \times (1 - P(S))^3$$

Knowing that the value of $P(S) = 0.4$, we can substitute this value and calculate the **probability of getting two success in five trials is** $= \mathbf{10 \times 0.4^2 \times 0.6^3 = 0.3456}$

Hence the probability of getting two heads in five coin tosses = 0.3456

- You can match your values with those given in the table below.

Probability of Success

p

n	r	.01	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	.55	.60	.65	.70	.75	.80	.85	.90	.95
2	0	.980	.902	.810	.723	.640	.563	.490	.423	[illegible]	.303	.250	.203	.160	.123	.090	.063	.040	.023	.010	.002
	1	.020	.095	.180	.255	.320	.375	.420	.455	[illegible]	.495	.500	.495	.480	.455	.420	.375	.320	.255	.180	.095
	2	.000	.002	.010	.023	.040	.063	.090	.123	[illegible]	.203	.250	.303	.360	.423	.490	.563	.640	.723	.810	.902
3	0	.970	.857	.729	.614	.512	.422	.343	.275	[illegible]	.166	.125	.091	.064	.043	.027	.016	.008	.003	.001	.000
	1	.029	.135	.243	.325	.384	.422	.441	.444	[illegible]	.408	.375	.334	.288	.239	.189	.141	.096	.057	.027	.007
	2	.000	.007	.027	.057	.096	.141	.189	.239	[illegible]	.334	.375	.408	.432	.444	.441	.422	.384	.325	.243	.135
	3	.000	.000	.001	.003	.008	.016	.027	.043	[illegible]	.091	.125	.166	.216	.275	.343	.422	.512	.614	.729	.857
4	0	.961	.815	.656	.522	.410	.316	.240	.179	[illegible]	.092	.062	.041	.026	.015	.008	.004	.002	.001	.000	.000
	1	.039	.171	.292	.368	.410	.422	.412	.384	[illegible]	.300	.250	.200	.154	.112	.076	.047	.026	.011	.004	.000
	2	.001	.014	.049	.098	.154	.211	.265	.311	[illegible]	.368	.375	.368	.346	.311	.265	.211	.154	.098	.049	.014
	3	.000	.000	.004	.011	.026	.047	.076	.112	[illegible]	.200	.250	.300	.346	.384	.412	.422	.410	.368	.292	.171
	4	.000	.000	.000	.001	.002	.004	.008	.015	[illegible]	.041	.062	.092	.130	.179	.240	.316	.410	.522	.656	.815
5	0	.951	.774	.590	.444	.328	.237	.168	.116	[illegible]	.050	.031	.019	.010	.005	.002	.001	.000	.000	.000	.000
	1	.048	.204	.328	.392	.410	.396	.360	.312	.259	.206	.156	.113	.077	.049	.028	.015	.006	.002	.000	.000
	2	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.346	.337	.312	.276	.230	.181	.132	.088	.051	.024	.008	.001
	3	.000	.001	.008	.024	.051	.088	.132	.181	[illegible]	.276	.312	.337	.346	.336	.309	.264	.205	.138	.073	.021
	4	.000	.000	.000	.002	.006	.015	.028	.049	.077	.113	.156	.206	.259	.312	.360	.396	.410	.392	.328	.204
	5	.000	.000	.000	.000	.000	.001	.002	.005	.010	.019	.031	.050	.078	.116	.168	.237	.328	.444	.590	.774

Number of trials

Number of Successes

Probability of getting 2 successes in 5 trials

So, the formula for finding **binomial probability** is given by:

$$P(X = r) = C(n, r)(p)^r (1 - p)^{n-r}$$

Where n is **the number of trials**, p is the **probability of success**, and r is the **number of successes after n trials**.

Continuous Probability Distribution

First, let's understand this using a continuous variable, say, the weight of a person. You know that the weight of a person is a continuous variable, i.e., it need not take specific discrete values. For example, the weight of a person can be 60 kg, or 60.1 kg, or 60.2 kg and so on.

But if we have such variables where the values can be continuous, does our old way of measuring probability work? For instance, if, say I want to find out the probability of a person's weight being exactly 60 kg, the probability is very less because the person might have a weight of 60.01 kg or 59.99 kg. The probability of a person's weight thus, being exactly 60 kgs is very less; we might even say that this probability is practically zero. So if the exact probabilities are zero, then how do we represent these numbers. The answer is simple - using cumulative probabilities.

But when continuous probability distributions is used, cumulative probability becomes quite significant.

Consider the table below.

Random Variable (X)	$F(X) = P(X \leq x)$
50	0.167
55	0.250
60	0.367
65	0.467
70	0.550
75	0.650
80	0.733
85	0.833
90	0.933
95	0.967
100	1

From the table above, the probability of a person whose weight is 50kg or less is $P(x \leq 50) = 0.167$. The probability of a person whose weight is less than 60kg is $P(x \leq 60) = 0.367$. So, the person whose weight is between 50kg and 60 kg is $P(x \leq 60) - P(x \leq 50) = 0.367 - 0.167 = 0.200$.

Notice that the data provided in the probability table is given in terms of cumulative probability because **in this case, the random variable X (which is the weight in this case) is a continuous variable.** X can take any value, like 60.01 or 59.99 and so on. Therefore, in this case, we cannot assign probabilities to each number, as the number of possibilities is infinite.

For instance, if somebody asks what the probability of an employee's weight is being exactly 60 kg, i.e., P(X=60), **it will be zero**. This is because the random variable X could be 59.99 or 60.01 or any other number which is close to 60 but not exactly 60, and **the chances of an employee's weight being exactly 60 kg out of infinite possibilities are zero**. However, if someone asks you what the probability of an employee's weight is being between 60 kg and 65 kg, you can still give a relevant probability number, because a lot of people will fall in that range.

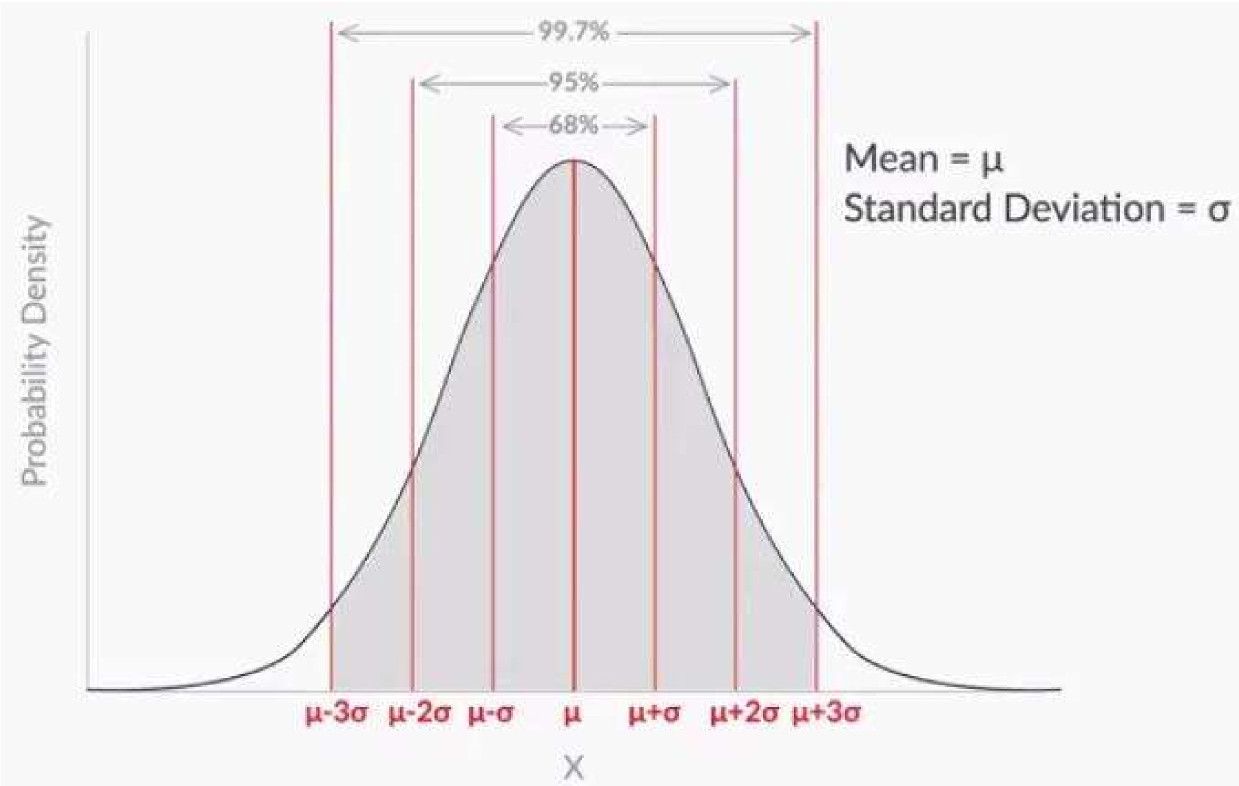
Normal Distribution

Normal distribution is the most commonly occurring probability distribution in day-to-day life.

In the industry, normal distribution and central limit theorem (which will also be covered) have significant applications, especially in quality control. For instance, consider the example, 'Weights of Coca-Cola bottles'. The fact that the weights are normally distributed can help the company determine whether the weights are in an acceptable range, as they wouldn't want a large number of bottles to be overweight. On the other hand, regulatory departments can use this fact to determine if Coca-Cola is, indeed, providing the promised weights to the customers. This is again just one of the many applications of normal distribution.

Normal distribution is an extremely symmetric distribution with its mean, median, and mode all lying exactly at the centre. Moreover, all data that is normally distributed follows the 1-2-3 rule. This rule states that there is a:

1. 68% probability of the variable lying within 1 standard deviation of the mean.
2. 95% probability of the variable lying within 2 standard deviations of the mean.
3. 99.7% probability of the variable lying within 3 standard deviations of the mean.

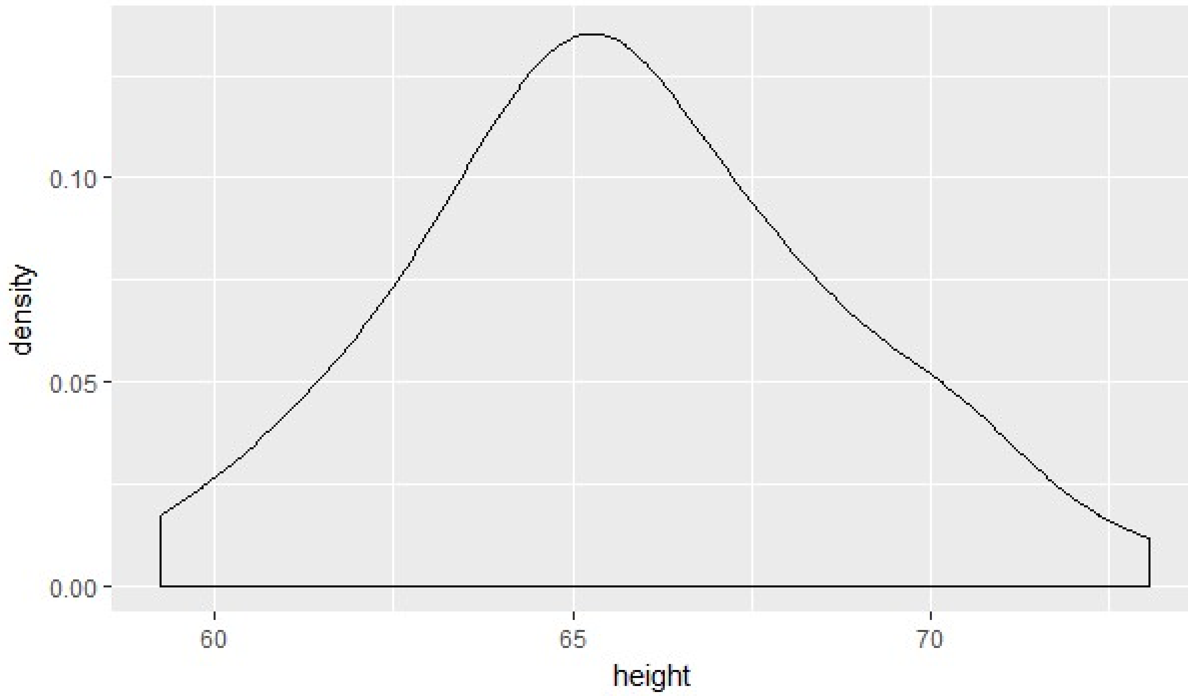


1-2-3 rule for a normal distribution

This is actually like saying that if you buy a loaf of bread everyday and measure it, then (assuming mean weight = 100 g, standard deviation = 1 g):

- For 5 days every week (57~68%), the weight of the loaf you buy that day will be within (100−1)=99 g and (100+1)=101 g.
- For 5 days every week (2021~95%), the weight of the loaf you buy that day will be within (100−2)=98 g and (100+2)=102 g.
- For 5 days every week (364365~99.7%), the weight of the loaf you buy that day will be within (100−3)=97 g and (100+3)=103 g.

A lot of naturally occurring variables are normally distributed. For example, the height of a group of adult men would be normally distributed. To try this out, we asked 50 male employees at the upGrad office for their height and then plotted the probability density function using that data.



Height of 50 male upGrad employees - Normally distributed

The above data is roughly normal.

Standard Normal Distribution:

In normal distribution, it does not matter what the value of μ and σ is. If we want to find the probability, all we need to know is how far the value of X is from μ, specifically, what multiple of σ is the difference between X and μ.

the standardised random variable is an important parameter. It is calculated as:

$$Z = \frac{X - \mu}{\sigma}$$