# **Basic Concepts of Probability**

* A probability is a number that reflects the chance or likelihood that a particular event will occur.
* Probabilities can be expressed as proportions that range from 0 to 1, and they can also be expressed as percentages ranging from 0% to 100%.
* A probability of 0 indicates that there is no chance that a particular event will occur, whereas a probability of 1 indicates that an event is certain to occur.
* A probability of 0.45 (45%) indicates that there are 45 chances out of 100 of the events occurring.

The concept of probability can be illustrated in the context of a study of obesity in children 5-10 years of age who are seeking medical care at a particular paediatric practice. The population (sampling frame) includes all children who were seen in the practice in the past 12 months and is summarized below.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Age (years)** | | | | | |  |
|  | **5** | **6** | **7** | **8** | **9** | **10** | **Total** |
| **Boys** | 432 | 379 | 501 | 410 | 420 | 418 | 2,560 |
| **Girls** | 408 | 513 | 412 | 436 | 461 | 500 | 2,730 |
| **Totals** | 840 | 892 | 913 | 846 | 881 | 918 | 5,290 |

## Unconditional Probability

* If we select a child at random (by simple random sampling), then each child has the same probability (equal chance) of being selected, and the probability is 1/N, where N=the population size.
* Thus, the probability that any child is selected is 1/5,290 = 0.0002.
* In most sampling situations we are generally not concerned with sampling a specific individual but instead we concern ourselves with the probability of sampling certain types of individuals.
* For example, what is the probability of selecting a boy or a child 7 years of age?
* The following formula can be used to compute probabilities of selecting individuals with specific attributes or characteristics.

**P(characteristic) = # persons with characteristic / N**

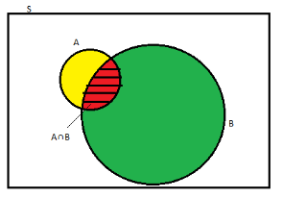
Try to figure these out before looking at the answers:

1. What is the probability of selecting a boy?
2. What is the probability of selecting a 7 year-old?
3. What is the probability of selecting a boy who is 10 years of age?
4. What is the probability of selecting a child (boy or girl) who is at least 8 years of age?
5. If we select a child at random, the probability that we select a boy is computed as follows P(boy) = 2,560/5,290 = 0.484 or 48.4%.
6. The probability of selecting a child who is 7 years of age is P(7 years of age) = 913/5,290 = 0.173.
7. P(boy who is 10 years of age) = 418/5,290 = 0.079.
8. P(at least 8 years of age) = (846 + 881+ 918)/5,290 = 2,645/5,290 = 0.500.

# **Conditional Probability**

**Conditional probability** is known as the possibility of an event or outcome happening, based on the existence of a previous event or outcome. It is calculated by multiplying the probability of the preceding event by the renewed probability of the succeeding, or conditional, event.

The probability of occurrence of any event A when another event B in relation to A has already occurred is known as conditional probability. It is depicted by P(A|B).



As depicted by the above diagram, sample space is given by S, and there are two events A and B. In a situation where event B has already occurred, then our sample space S naturally gets reduced to B because now the chances of occurrence of an event will lie inside B.

As we have to figure out the chances of occurrence of event A, only a portion common to both A and B is enough to represent the probability of occurrence of A, when B has already occurred. The common portion of the events is depicted by the intersection of both the events A and B, i.e. A ∩ B.

## Formula

When the intersection of two events happen, then the [formula for conditional probability](https://byjus.com/conditional-probability-formula/) for the occurrence of two events is given by;

**P(A|B) =**

**Or**

**P(B|A) = N(A∩B)/N(A)**

Where P(A|B) represents the probability of occurrence of A given B has occurred.

N(A ∩ B) is the number of elements common to both A and B.

N(B) is the number of elements in B, and it cannot be equal to zero.

**Example 1: Two dies are thrown simultaneously, and the sum of the numbers obtained is found to be 7. What is the probability that the number 3 has appeared at least once?**

Solution: The sample space S would consist of all the numbers possible by the combination of two dies. Therefore S consists of 6 × 6, i.e. 36 events.

Event A indicates the combination in which 3 has appeared at least once.

Event B indicates the combination of the numbers which sum up to 7.

A = {(3, 1), (3, 2), (3, 3)(3, 4)(3, 5)(3, 6)(1, 3)(2, 3)(4, 3)(5, 3)(6, 3)}

B = {(1, 6)(2, 5)(3, 4)(4, 3)(5, 2)(6, 1)}

P(A) = 11/36

P(B) = 6/36

A ∩ B = 2

P(A ∩ B) = 2/36

Applying the conditional probability formula we get,

P(A|B) = P(A∩B)/P(B) = (2/36)/(6/36) = ⅓

**Example 2: The table below shows the occurrence of diabetes in 100 people. Let D and N be the events where a randomly selected person "has diabetes" and "not overweight". Then find P(D | N).**

|  |  |  |
| --- | --- | --- |
|  | Diabetes (D) | No Diabetes (D') |
| Not overweight (N) | 5 | 45 |
| Overweight (N') | 17 | 33 |

**Solution:**

From the given table, P(N) = (5+45) / 100 = 50/100.

P(D ∩ N) = 5/100.

By the conditional probability formula,

P(D | N) = P(D ∩ N) / P(N)

= (5/100) / (50/100)

= 5/50

= 1/10

**Answer: P(D | N) = 1/10.**

**Example 3: The probability that it will be sunny on Friday is 4/5. The probability that an ice cream shop will sell ice creams on a sunny Friday is 2/3. Then find the probability that it will be sunny and the ice cream shop sells the ice creams on Friday.**

**Solution:**

Let us assume that the probabilities for a Friday to be sunny and for the ice cream shop to sell ice creams be S and I respectively. Then,

P(S) = 4/5.

P(I | S) = 2/3.

We have to find P(S ∩ I).

We can see that S and I are dependent events. By using the dependent events' formula of conditional probability,

P(S ∩ I) = P(I | S) · P(S) = (2/3) · (4/5) = 8/15.

**Answer: The required probability = 8/15.**

**Bayes' Theorem**

Bayes' Theorem is a way of finding a [probability](https://www.mathsisfun.com/data/probability.html) when we know certain other probabilities.

The formula is:

P(A|B) =

|  |  |  |
| --- | --- | --- |
| Which tells us: |  | how often A happens *given that B happens*, written **P(A|B)**, |
| When we know: |  | how often B happens *given that A happens*, written **P(B|A)** |
|  |  | and how likely A is on its own, written **P(A)** |
|  |  | and how likely B is on its own, written **P(B)** |

**Example:**

Let us say P(Fire) means how often there is fire, and P(Smoke) means how often we see smoke, then:

P(Fire|Smoke) means how often there is fire when we can see smoke  
P(Smoke|Fire) means how often we can see smoke when there is fire

So the formula kind of tells us "forwards" P(Fire|Smoke) when we know "backwards" P(Smoke|Fire)

### Example:

### dangerous fires are rare (1%)

### but smoke is fairly common (10%) due to barbecues,

### and 90% of dangerous fires make smoke

We can then discover the **probability of dangerous Fire when there is Smoke**:

P(Fire|Smoke) =P(Fire) P(Smoke|Fire)/**P(Smoke)**

=1% x 90% / **10% = 0.01\*0.9 /0.1 = 0.09 = 9%**

=9%

So it is still worth checking out any smoke to be sure.

### Example: Picnic Day

You are planning a picnic today, but the morning is cloudy

* Oh no! 50% of all rainy days start off cloudy!
* But cloudy mornings are common (about 40% of days start cloudy)
* And this is usually a dry month (only 3 of 30 days tend to be rainy, or 10%)

**What is the chance of rain during the day?**

We will use Rain to mean rain during the day, and Cloud to mean cloudy morning.

The chance of Rain given Cloud is written P(Rain|Cloud)

So let's put that in the formula:

P(Rain|Cloud) = P(Rain) P(Cloud|Rain)/ **P(Cloud)**

* P(Rain) is Probability of Rain = 10%
* P(Cloud|Rain) is Probability of Cloud, given that Rain happens = 50%
* P(Cloud) is Probability of Cloud = 40%

P(Rain|Cloud) = 0.1 x 0.5 / **0.4**  = .125

Or a 12.5% chance of rain. Not too bad, let's have a picnic!

# **Random Variable**

Sashi plans to begin selling ice-creams in the month of March. She manufactures a batch of 1200 waffle-cones for the same. She then examines random waffle-cones and judges each cone as either “defective” or “non-defective”. She decides to examine 3 waffle-cones.

* The process of observation of activity is termed as an **experiment**.
* The results of the observation are termed as **outcomes** of the experiment.
* **Random experiments** are those experiments whose outcomes can't be predicted.

Examining a random waffle-cone and judging each cone as either defective or non-defective, is the experiment in the above scenario.

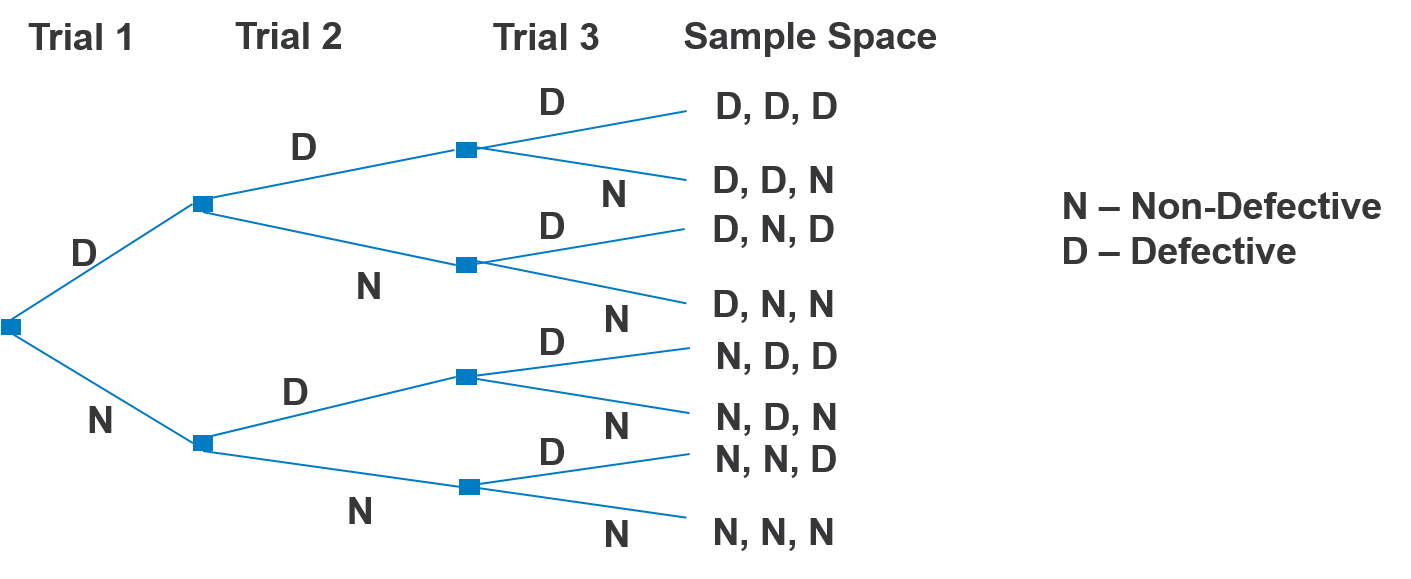
The outcome of the above experiment can be either defective(**D**) or non-defective(**N**).

Since the exact outcome of the experiment (**D** or **N**) cannot be predicted, the above experiment is a random experiment.

* Individual repetitions of the same random experiment are termed as the**trial**.
* The set of all possible outcomes in a random experiment is called **Sample space**.

Examining a waffle-cone is a trial within the experiment composed of examining 3 waffle cones.

The sample space can be defined for the experiment using a tree diagram as shown below:



One particular outcome or a set of some outcomes from the entire sample space is termed as an **event**.

In the experiment, the event of interest can be composed of all the outcomes in which the total no. of defective cones is two. Then the event **E**is

**E = {DDN, DND, NDD}**

Hence, there are 3 ways in which the above event can occur.

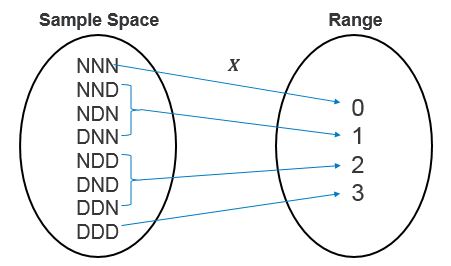
In the experiment, the event of interest can be composed of all the outcomes in which the total no. of defective cones is one or more. Then the event**E**is

**E = {DDD, DNN, NDN, NND, DDN, DND, NDD}**

Hence, there are 7 ways in which the above event can occur.

A function **X** can be defined or explained on the Sample Space as a relation where each independent sample space outcome is mapped to a numerical value, based on the event of interest.

For example, if you go through the total number of defective cones in 3 trials then, **X** can be depicted as follows:



**X** is called as **Random Variable**.

* A real-valued function, defined over a sample space is called a random variable.
* Only one real value is assigned by function to each individual outcome.
* **X** or any other uppercase letter denotes a random variable.
* A lowercase letter is used, for example, **x**  denotes the real value that can be mapped with a random variable map and its each outcome inside the sample space.
* **X** is not a variable like in Algebra. e.g. x+2 = 7, where x value is unknown.
* **X** is a function and can be depicted as follows:

**X = {x1, x2, x3, ....} or X = xk, where k = 1,2,3,...**

You can explain random variable **X** in the ice-cream scenario as follows:

X = Total no. of defective cones in 3 trials = {0,1,2,3}

You may know/be aware of the possible values of X i.e. 0,1,2,3 before to begin the experiment, you cannot be sure what values will be taken at the experiment end.

Also, **X**can assume different values each time the experiment is performed.

Due to its trial to trial variability in value and non-predictable nature, **X** is called a random variable.

# **Types of Random Variables**

A random variable may either be **Discrete** or **Continuous**.

* A random variable having a finite range is called a **Discrete Random Variable**.

For example, the total number of scratches found on a glass surface or proportion or percentage of defective boxes among 1000 tested

In real time experiments, you could record the total no of transmitted bits in error transmissions as integers or in fractions like 0.0042 proportion of the total 10,000 transmitted bits. But the fractional count can be put as numbers on a number line. So, whenever, the count is limited to a finite point on the real number line, those random variables are called discrete random variables.

* Similarly, a random variable having values within real numbers interval is called a **continuous random variable**.

For example, Height of a child measured between age 2 and 6, Weight of a person logged over a span of 2 years.

Sometimes you can observe that calculations (like heights, weights, current in a wire) assume the value in a range of the real numbers (always true theoretically). There is the possibility for an arbitrary precision in the calculation. However, in real life application, you may round off the value to a nearest 10th or 100th of a unit. Continuous random variables are such random variables that represent this type of range of values.

Sometimes, when the range of possible values is very high, you can consider a discrete random variable **X** is continuous. For example, a multi-meter output that displays the current at the nearest 100th of a milliampere. Although the measurements are limited, and it could be perceived to be discrete random variable it can also be considered as continuous random variables if needed for simplification in computation.

Possible outcomes for **X**= Number of defective cones in 3 trials

i.e., X = {0,1,2,3}

**X** is a discrete random variable because it results in values that are finite in nature and are from the above set.

In the ice-cream scenario, some of the other examples that can be considered as discrete random variables are:

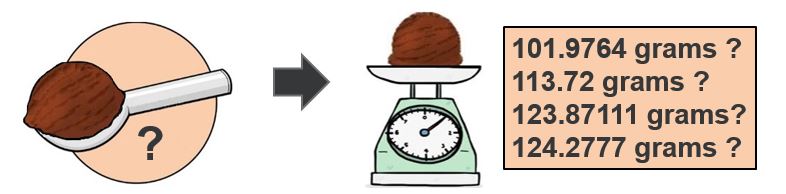
* Number of customers expected to arrive between 2 pm and 3 pm
* Number of ice-creams bought by children below 12 years of age
* Number of ice-cream cones that break/leak while being filled with ice-cream
* Number of ice-creams sold in a month
* Number of ice-creams sold in a year

Sashi uses a 4oz dish spoon to serve ice-cream equivalent to 113 grams, per scoop.

Let us consider a random variable, **X**= Weight of one scoop of ice-cream.

Let us consider, the permissible error margin is 10% for the weight of a single scoop of ice-cream, so our **X**can assume any values in the interval **[101.7 grams, 124.3 grams]**and outcomes could be

X = {101.7, 101.88, ..................., 110.346, .................. 124.29}



If you consider any two values in the interval [101.7,124.3] say, 102 and 103, here you can get any value in the interval [102, 103]. You can assume any value between 102.5 and 102.9 for the weight of a single scoop.

In such a scenario, wherein an interval, **X**can assume any value, is defined to be a continuous random variable.

Few other examples for continuous random variables :

* The exact time taken by Sashi to serve a customer in the [10 sec, 15 sec] interval with an accuracy of 100th of a second.
* The precise diameter of a single ice-cream scoop that Sashi serves, assuming the average diameter of a single scoop that Sashi serves lies in the interval [2 inches, 2.25 inches]
* The exact thickness of the waffle-sheet with which the waffle-cones are made from, assuming the average thickness of a waffle-sheet lies in the interval [1.4 mm, 1.8 mm]

# **Probability Distribution**

By studying Distributions you can understand how the random variable behaves. When the possibility of random variable values is associated with each of its probabilities, you get its Probability Distribution.

The probability distribution is usually represented through either a table or a Graph(usually a histogram).

Recall, for a Finite Sample Space **S,**then the probability **P(A),**is a real number assigned to the event **A** such that **0<=P(A)<= 1** and **P(S)=1**

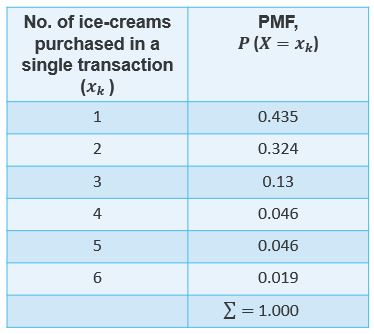
# **Types of Probability Distributions**

Probability Distributions can be Discrete or Continuous.

* The associated probability distribution for a random variable with discrete values is called a **Discrete Probability Distribution**
  + Discrete Probability Distributions are described by using the**Probability Mass Function (PMF)**.
* The associated probability distribution for a random variable with continuous (or approx. continuous) values is called a **Continuous Probability Distribution**
  + Continuous Probability Distributions are described by using the **Probability Density Function (PDF)**.

# **Probability Mass Function PMF**

In the table given below:



The probability of a customer buying exactly one ice-cream is 0.435 i.e.,**P(X=1)=0.435**i.e.,**43.5%.**

Find the probability of next customer buying exactly 4 ice-creams?

# **Types of Discrete Distributions**

* The Discrete Uniform Distribution
* The Binomial Distribution
* The Negative Binomial Distribution
* The Poisson Distribution