# Introduction to Probability

February 7, 2021

## 0.1 About Probability

**Event:** In probability theory, an *event* is one or more of the possible outcomes of doing some thing. Example: If we toss a coin then getting a head would be an event and getting a head would be an event and getting a tail would be another event.

**Experiment:** The activity that produces such an event is referred to in probability theory as an experiment.

**Sample Space:** The set of all possible outcomes of an experiment is called the sample space for the experiment. In coin toss experiment, the sample space is:

S = {head, tail}

```
[14]:  #import packages
       import pandas as pd
       import numpy as np
       import seaborn as sns
       import matplotlib.pyplot as plt

       %reload_ext autoreload
       %autoreload 2
       %matplotlib inline
       %config InlineBackend.figure_format = 'retina'

       sns.set_style("white")
       sns.set_context("talk")
```

## 0.2 Events

Events can be:

- Independent(each event is not affected by other events)

- Dependent(also called "conditional", where an event is affected by other events)

- Mutually Exclusive(events can't happen at the same time)

- Collectively exhaustive list(When a list of the possible events that can result from an experiment includes every possible outcome)

**Note:** Collectively exhaustive events: In probability theory and logic, a set of events is jointly or collectively exhaustive if at least one of the events must occur. For example: when rolling a

six sided die, the outcomes 1,2,3,4,5,and 6 are collectively exhaustive, because they encompass the entire range of possible outcomes. Another way to describe collectively exhaustive events, is that their union must cover all the events within the entire sample space.

**Exercise:** Southern Bell is considering the distribution of funds for a campaign to increase long distance calls within North Carolina. The following table lists the markets that the company considers worthy of focused promotions:

Market Segment

Cost of Special Campaign Aimed at Group

Minorities

350,000

Business People

550,000

Women

250,000

Professionals and white collar workers

200,000

Blue Collar workers

250,000

There is up to $ 800,000 available for these special campaigns.

  (a) Are the market segments listed in the table collectively exhaustive? Are they mutually exclusive?
  (b) Make a collectively exhaustive and mutually exclusive list of the possible events of the spending decision.
  (c) Suppose the company has decided to spend th eentire $ 800,000 0n special campaigns. Does this change your answer to part(b)? If so what is your new answer?

There are three types of Probability, like three types of conceptual approaches to solve probabilistic problem.

  1. Classical Approach

  2. Relative frequency Approach

  3. Subjective approach

### 0.2.1  Classical Probability:

This is same as the common probability theory we use.

$$Probability \ of \ an \ event = \frac{number \ of \ outcomes \ where \ the \ event \ occurs}{total \ number \ of \ possible \ outcomes}$$

```
[1]: #implementing the classical probability with python
     def event_probability(event_outcomes, sample_space):
         """
         count the probability of an event.

         Paramrter Description:
         ----------------------

         event_outcomes: number of outcomes where the event occurs
         sample_space: total number of possible outcomes

         Returns:
         ----------
         probability of the event in percentage format

         Example:
         ----------
         >>> print(event_probability(13,52))
         >>> 25.0

         """
         probability = (event_outcomes/sample_space)*100
         return round(probability,1)
```

```
[2]: # defining sample space
     cards = 52

     #event outcomes
     hearts = 13

     hearts_probability = event_probability(hearts,cards)
     print(str(hearts_probability) + '%')
```

Classical probability is often called a **priori probability** because if we keep using orderly examples such as fair coins, unbiased dice, and standard decks of cards, we can state the answer in advance (a priori) without tossing a coin, rolling a die, or drawing a card. We don't have to perform experiment to make probability of those events rather we can make statements based on logical reasoning.

**Problems with classical approach:** This approach to probability is useful when we deal with card games, dice games, coin tosses, and the like, but has serious problems when we try to apply it to the less orderly decision problems we encounter in management. The classical approach to probability assumes a world that does not exist. It assumes away situations that are very unlikely but that could conceivably happen. Real life sutuations, disorderly and unlikely as they often are, make it useful to define probabilities in other ways.

### 0.2.2 Relative frequency appraoch

Suppose if we want to know complex questions like "What is the probability that I will live to be 85?" or "What are the chances that I will blow on eof my stereo speakers if I turn my 200-watt amplifier up to wide open?". We quickly see that we may not be able to state in advance, without experimentation, what these probabilities are.

In 1800s, British Statisticians, interested in theoretical foundation for calculating risk of losses in life insurance, began defining probabilities from statistical data collected on births and deaths. Today this approach is called the ***Relative Frequency of Occurance***. It defines probabilitiy of either:

1. The observed relative frequecny of an event in a very large number of trials, or
2. The proportion of times that an event occurs in the long run when conditions are stable.

The method uses the relative frequencies of past occurrences as probabilities. we determine how often something has happened in the past and use that figure to predict the probability that it will happen again in the future.

For showing the example we can a see 100 coin toss example that how is it effecting the 0.5 probability of getting either head or tail from a coin toss.

```
[4]: toss = pd.read_csv("data/CoinToss.csv")
```

```
[5]: toss
```

```
[5]:     Toss Coin1 Coin2  Coin1_h  Coin2_h
     0      1     H     T        1        0
     1      2     T     H        0        1
     2      3     T     T        0        0
     3      4     H     T        1        0
     4      5     T     H        0        1
     ..   ...   ...   ...      ...      ...
     95    96     H     T        1        0
     96    97     T     T        0        0
     97    98     T     T        0        0
     98    99     T     H        0        1
     99   100     H     H        1        1

     [100 rows x 5 columns]
```

```
[6]: toss = toss[['Toss','Coin1']]
```

```
[7]: toss.rename(columns = {'Toss':'Count', 'Coin1':'Coin_result'}, inplace = True)
```

```
/usr/local/lib/python3.9/site-packages/pandas/core/frame.py:4300:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-
```

```
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  return super().rename(
```

[8]: `toss`

[8]:
```
     Count Coin_result
0        1           H
1        2           T
2        3           T
3        4           H
4        5           T
..     ...         ...
95      96           H
96      97           T
97      98           T
98      99           T
99     100           H

[100 rows x 2 columns]
```
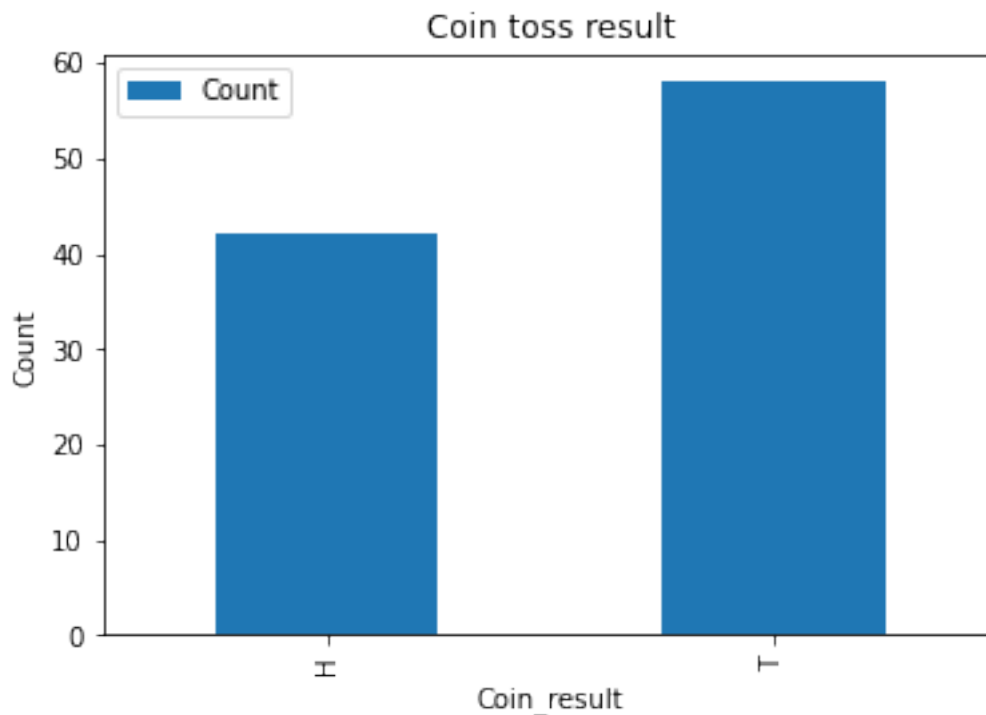
[9]: `toss_sum = toss.groupby(['Coin_result']).count()`

[10]: `toss_sum`

[10]:
```
             Count
Coin_result
H               42
T               58
```

[11]:
```
toss_sum.plot(kind = "bar", title = "Coin toss result")
plt.ylabel('Count')
plt.show()
```

So here we can see if we toss a coin 100 times the result is not quite 0.5 in case. The amount of tail is prone to high. So that's an example of using relative frequency of occurance.

One difficulty with the relative frequency approach is that people often use it without evaluating a sufficient number of outcomes. If you heard someone say, "My aunt and uncle got the flu this year, and they are both over 65, so everyone in that age bracket will probably get the flu," So, it is clear that his observation were having insufficient data for establishing a relative frequency of occurance probability.

### 0.2.3   Subjective Probabilities

Subjective probability refers to the probability of something happening based on an individual's own experience or personal judgment. A subjective probability is not based on market data or historical information and differs from person to person. In other words, it is created from the opinion of an individual and is not based on fact.

In most types of probability, quantitative information is interpreted to determine the likelihood of something happening. However, subjective probability does not base its probability on quantitative information, is affected by personal beliefs, and contains no formal calculations.

Example:

An analyst is asked the probability of the S&P 500 will hit all-time highs in the coming months. The analyst looks at past trends and current market conditions and estimates that the probability of the S&P 500 will hit all-time highs at 20%.

### 0.2.4 Important concept on classical probability problems

During solving any problem in classical approach, be sure to check whether the situation is "with replacement" after each draw or "without replacement". Let's see why is it important:

Let's say you had a population of 7 people, and you wanted to sample 2. Their names are:

tanvir, tawhid, farabi, apon, mamun, sourav, anik You could put their names in a hat. If you sample with replacement, you would choose one person's name, put that person's name back in the hat, and then choose another name. When you sample with replacement, your two items are independent. In other words, one does not affect the outcome of the other. You have a 1 out of 7 (1/7) chance of choosing the first name and a 1/7 chance of choosing the second name. p(tanvir,tawhid) = (1/7) * (1/7) = 0.02 p(farabi,apon) = (1/7) * (1/7) = 0.02 and so on,

but when we are doing **without replacement:** But now, your two items are dependent, or linked to each other. When you choose the first item, you have a 1/7 probability of picking a name. But then, assuming you don't replace the name, you only have six names to pick from. That gives you a 1/6 chance of choosing a second name. The odds become: p(tanvir,tawhid)= (1/7) * (1/6) = 0.024 p(farabi,apon) = (1/7) * (1/6) = 0.024

Let's test our concept: Classify the following probability estimates as to their type (classical, relative frequecy or subjective): (a) The probability of scoring on a penalty shot in ice hockey is 0.47 (b) The probability that the current mayor will resign is 0.85 (c) The probability of rolling two sixes with two dice is 1/36 (d) The probability that a president elected in a year ending in zero will die in office is 7/10 (e) The probability that you will go to Europe this year is 0.14

## 0.3 Probability Rules

Most managers who use probabilities are concerned with two conditions:

1. The case where one event or another will occur.
2. The situation where two or more events will both occur.

**Marginal Probability or unconditional probability:** A single probability means that only one event can take place. It is called conditional probability.

**Addition Rule for mutually exclusive events:** If two events are mutually exclusive, we can express this probability using the addition rule for mutually exclusive events.

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

**Addition Rule for events that are not mutually exclusive:** if two events are not mutually exclusive, it is possible for both events to occur. In this case we will have to omit the probability of happening both event.

$$P(A \ or \ B) = P(A) + P(B) - P(AB)$$

There is a nice diagrammatic way to illustrates these probabilities rules. We use a pictorial representation called a **_Venn Diagram_**, after the nineteenth century English mathematician John

Venn. In these diagram, the entire sample space is represented nu a rectangle, and events are represented by parts of the rectangle. If two events are mutually exclusive their parts of the rectangle will not overlap each other otherwise if two events are not mutually exclusive then they will overlap.

Sample space

A    B

Two mutually exclusive events

Sample space

A   B

Two  nonexclusive events
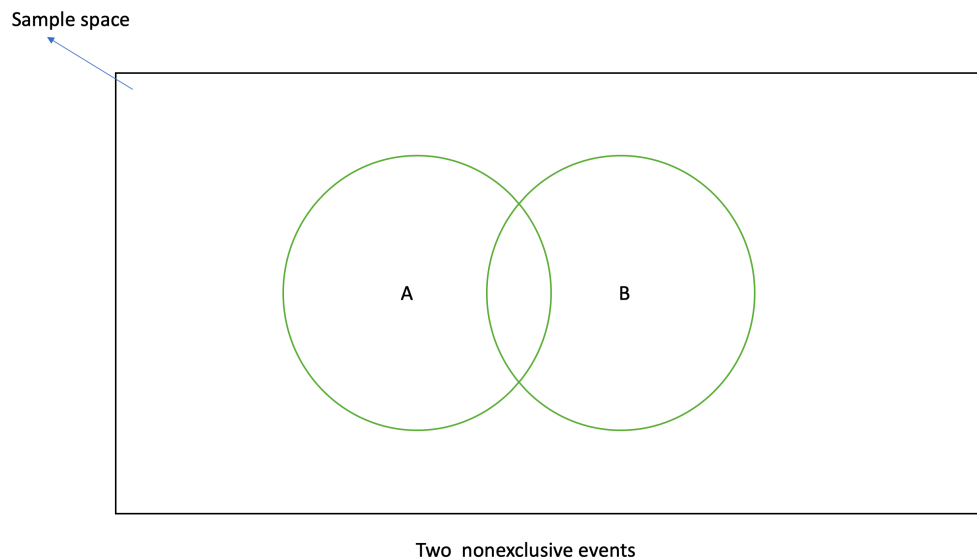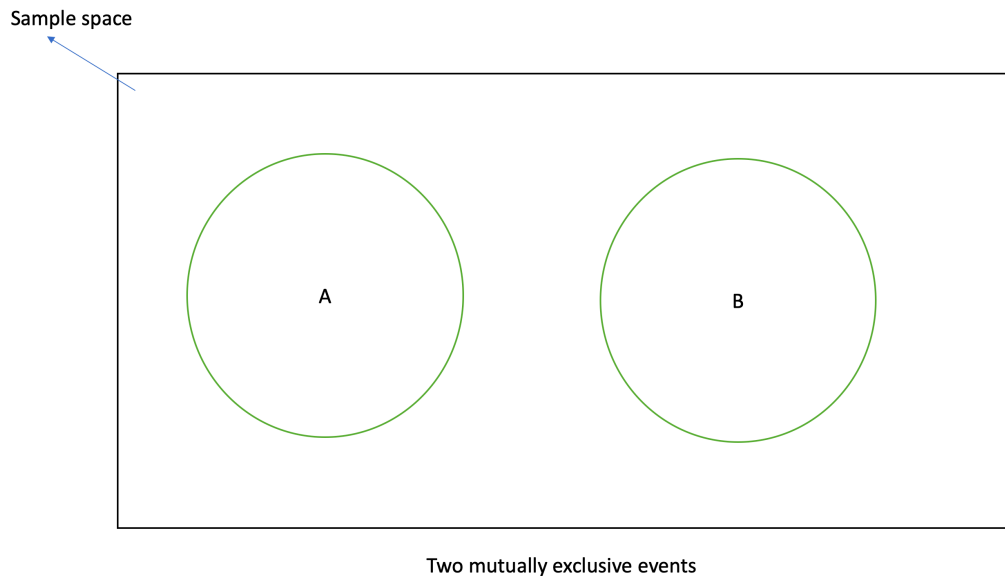
## 0.4   Probabilities under conditions of statistical independence

When two events happen, the outcome of the first event may or may not have an effect on the outcome of the second event. That is, the events may be either dependent or independent. Here we will examine the events that are statistically independent: The occurance of one event has no effect on the probability of the occurence of any other event. There are three types of probabilities under statistical independence:

1. Marginal
2. Joint
3. Conditional

### 0.4.1 Marginal Probabilities under statistical Independence

Marginal or unconditional probability is the simple probability of occurance of an event. This is true for every unbiased coin toss. Now imagine we have a biased or unfair coin that has been altered in such a way that head occurs 0.90 of the time and tails occurs 0.10 of the time. On each individual toss, $P(H) = 0.90$ and $P(T) = 0.10$. The outcome of any particular toss is completely unrealted to the outcomes of the tosses that may preceede or follow it. So for biased coin as well this is true.

### 0.4.2 Joint Probabilities under statistical Independence

The probability of two or more independent events occuring together or in succession is the product of their marginal probabilities. Mathematically, this is stated (for two events):

$$
P(AB) = P(A) * P(B)
$$

where, $P(AB)$ = probability of events A and B occuring together or succession - this is known as join probability $P(A)$ = marginal probability of event A occuring $P(B)$ = marginal probability of event B occuring

**Probability Tree:** We can make the probabilities of events event more explicit using a probability tree. It is showing the possible outcomes and their respective probabilities for one toss of a fair coin . Let's contruct a probability tree for tossing coin thrice.
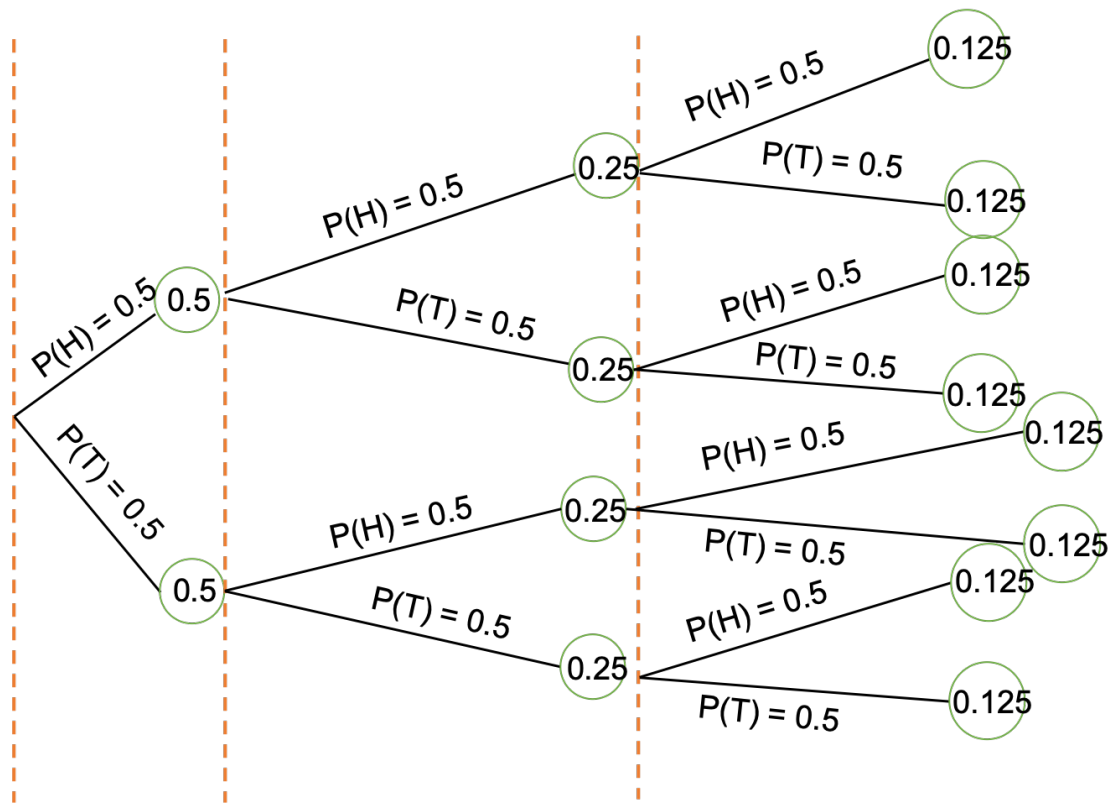
So from the tree we are getting that, if we toss a fair coin and want to know the probability that all three tosses will result in heads. so,

$$P(H_1 H_2 H_3) = P(H_1) * P(H_2) * P(H_3) = 0.5 * 0.5 * 0.5 = 0.125$$

### 0.4.3 Conditional Probabilities under statistical Independence

Conditional probability is the probability that a second event (B) will occur of a first event (A) has already happened.



P(B|A)

The **probability** of **event B** given that **event A** has occurred
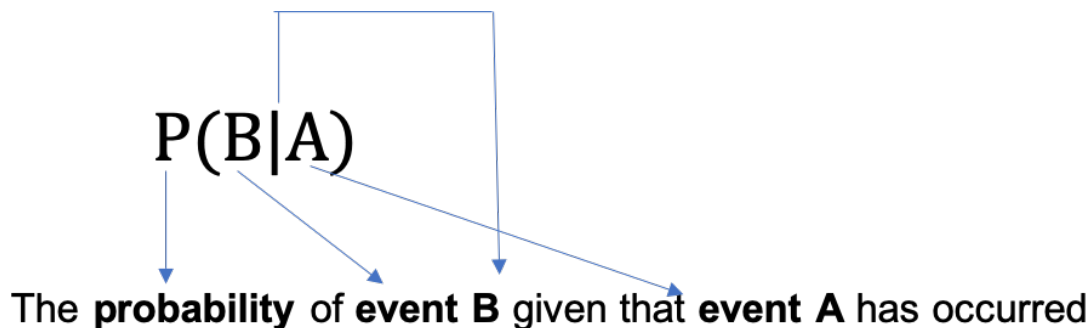
For the statistically independent events, the conditional probability of event B given that event A has occured is simply the probability of event B:

$$P(B|A) = P(B)$$

**Example:**

The health department routinely conducts two independent inspections of each restaurant, with the restaurant passing only if both inspectors pass it. Inspector A is very experienced, and, hence, passes only 2 percent of restaurants that actually do have health code violations. Inspector B is less experienced and passes 7 percent of restaurants with violations. What is the probability that:

    (a) Inspector A passes a restuarant, given that inspector B has found a violation?
    (b) A restuarant with a violation is passed by the health department?

## 0.5 Probabilities under condition of statistical Dependence

Statistical dependence exists when the probability of some event is dependent on or affected by the occurance of some other event. 3 types of probabilities is possible:

1. Conditional
2. Joint
3. Marginal

### 0.5.1 Conditional probabilities under statistical dependence

Suppose you have two events, event A and event B. The conditional probability means event A has occured, taking into account that event B has definitely occured during trial. Explaining it with example would be better:

Assume that we one box containing 10 balls distributed as follows:

- Three are colored and dotted
- One is colored and striped
- Two are gray and dotted
- Four are gray and striped

The probability of drawing any one ball from this is 0.1. Since there are 10 balls, each with equal probability of being drawn.

Now, Suppose someone draws a colored ball from the box. What is the probability that it is dotted? What is the probability that it is striped?

Answer: This question can be expressed symbolically as P(D|C) or what is the conditional probability that this ball is dotted given that it is colored?

So, we already have the condition that the previous ball that is drawn was colored. So we will omit the possibility of having gray balls. We will just look for colored and dotted and colored and stiped balls. Total numbers are 4 in this case. That means we have 4 colored balls. Now among them, 3 of them are dotted. so the probability becomes *3/4 = 0.75* and 1 of them is striped so probability is: *1/4 = 0.25* We have our answer now.

But how can we generalize it and make a equation out of it?

To calculate the probability of dotted given colored, P(D|C), we divided the probability of colored and dotted balls(3 out of 10 or 0.3) by the probability of colored balls (4 out of 10 or 0.4):

$$P(D|C) = \frac{P(D \cap C)}{P(C)}$$

### 0.5.2 Implementation the problem in python

```python
[2]: def event_probability(given_condition_value, required_condition_value):
         """
         count the conditional probability of an event.

         Paramrter Description:
         ----------------------

         given_conditon_value: The probability of given condition will occur among␣
     ↪total value
         required_condition_value: The probaility of possible final outcome that we␣
     ↪want.

         Returns:
         ----------
         probability of the event in percentage format

         Example:
         ----------
         >>> print(event_probability(0.3,0.4))
         >>> 75.0%

         """
         conditional_probability = (required_condition_value/
     ↪given_condition_value)*100
         return round(conditional_probability,1)
```

```python
[10]: probability_colored_and_dotted_balls = 3/10
      probability_of_colored_balls = 4/10

      probability =␣
       ↪event_probability(probability_of_colored_balls,probability_colored_and_dotted_balls)
      print((str(probability) + '%'))
```

```
75.0%
```

### 0.5.3 Joint probabilities under statistical dependence

Joint dependent probability means the probability of occuring two statistically dependent probability at once. Simply the formula will be changed to:

$$P(D \cap C) = P(D|C) \times P(C)$$

12

Here, $P(D \cap C)$ = Joint probaibility of events D and C happening together or in succesion $P(D|C)$ = Probability of event D given that event C has happened $P(C)$ = Probability that event A will happen

### 0.5.4   Marginal probabilities under statistical dependence

Marginal Probabilities under statistical independence are computed by summing up the probabilities of all the joint events in which the simple event occurs.

## 0.6   Probability formula table under statistical independence and dependence

| Types of Probability | Symbol | Formula under statistical Independence | Formula under statistical dependence |
|---|---|---|---|
| Marginal | P(A) | P(A) | sum of the probabilities of the joint events in which A occurs |
| joint | P(AB) or P(BA) | P(A) * P(B) or P(B) * P(A) | P(A\|B) * P(B) or P(B\|A) * P(A) |
| Conditional | P(B\|A) or P(A\|B) | P(B) or P(A) | P(BA) / P(A) or P(AB) / P(B) |