**Probability and Statistics-1**

**Outline**

* **Probability definition**
  + Event
  + Experiment
  + Sample space
* **Probability formula**
* **Laws of porbability**
* **Probability frequentist approach**
* **Marginal probability**
* **Independent events**
* **Mutually exclusive events**
* **Joint probability**
* **Conditional probability**
* **Law of total probability**

**Probability definition**

**What is probability ?**

* Probability is **how likely an event is to occur**.
* Used whenever there is a **uncertainity in the event**
* The probability of an event is a number **between 0 and 1**.
  + 0 indicates impossibility of the event
  + 1 indicates certainty of the event
* E.g., choosing cup with money is not certain.

**What is an event ?**

* It is desired outcome out of all possible outcomes of an experiment.
* It is a subset of the respective sample space.

Well, there are 2 new terms in this definition

1. Experiment
2. Sample space

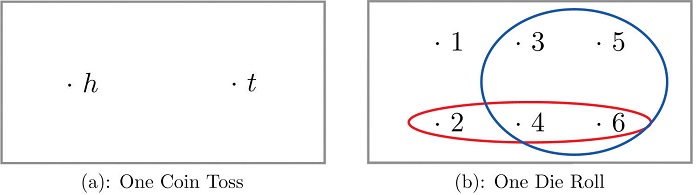
Lets understand what they are

**What is an experiment ?**

* It is a **well-defined set of possible outcomes**
* It can be **infinitely repeated**.
* It can be one of:
  1. **Random**
     + More than 1 possible outcome
     + Eg: Tossing coin can lead to heads/tails
  2. **Deterministic**
     + Only 1 possible outcome
     + So probability of that outcome becomes 1
     + Eg: Adding two given numbers

**And what is a sample space ?**

* Sample Space - **Set of all possible outcomes** of an experiment.
* Denoted using set notation (S)



**Probability Formula**

**Now how can we calculate Probability of an event?**

* We know that there can be multiple outcomes of an experiment
* The chances of each outcome might be same or different
* When each outcome is equally likely to occur:
  + **Probability = (number of desired outcomes) / (total number of possible outcomes)**
  + For eg: If we toss a coin, both heads and tails are equally likely to occur
  + So P(Heads) = P(Tails) = 1/2 = 0.5
* Sometimes outcomes are not equally likely to occur
  + Lets say, you found out that a coin is biased
  + It was observed through multiple trials that P(Head) = 2\*P(Tails)
  + **How can you find P(Heads) and P(Tails) through this ?**
    - We know P(Heads) + P(Tails) = 1
      * Because these 2 are the only possible outcomes
    - So, 2\*P(Tails)+P(Tails)=1 (By substitution)
    - => P(Tails)=1/3 = 0.33
    - => P(Heads)=1-P(Tails)=2/3=0.667

**Laws of Probability**

The laws of probabilty are:

1. 0 <= P(E) <= 1 (Probability lies between 0 and 1)

* P(E) = 0 means event cannot occur
  + For eg: Rolling a die and getting a 7 on it
* P(E) = 1 means the event will definitely occur
  + Eg: Rolling a die and getting a number between 1 and 6

1. P(S) = 1. (Sample space has a probability of 1)

* Eg: If we roll a dice, then probability of getting any number between [1,6],  
  P(1 to 6) = 1

**Casino Vegas Game 2 (Probability Frequentist Approach)**

* You and your friend want to play another game at casino.
* This game involves rolling a dice
* If the dice shows the number of your choice you win $10.

Lets say you choose the **number 6**

**What do you think is the probability of winning ?**

* Experiment: Rolling a dice
* Sample space: 6 outcomes; {1, 2, 3, 4, 5 ,6}
* Event: Getting 6 on the dice

**So what should be the probability of winning ?**

* P(Getting 6) = 1/6 = 0.166

Now lets suppose you went around asking everyone who played the game, whether they won or lost.

After asking 100 people, you came to know that only 20 won that game.

**So what can you say about the probability of winning from this ?**

* Its 20/100 = **0.2**
* But ideally around 16-17 people should have won to get a probability of 0.166.

**Why are we not getting a probab of 0.166 ?**

* We have asked only 100 people.
* The **sample size is very small**
* So there is no certainity that we will only get 0.166 probability

**Do you think this probab will remain the same if we ask 1000 or maybe 10000 people ?**

* No
* The probab will start getting closer to 0.166.

**So whats the solution ?**

* Perform the experiment larger number of times.
* Ideally, if you perform the experiment for **infinite times probability becomes ~ 0.166**.

Lets simulate this game to see this for ourselves

Lets suppose we roll a die n number of times

If we get 6 x times, then probability of getting a 6 is going to be x/n

We want x/n ~ 0.166

Lets see at what sample size this happens

import random

for i in [5, 10, 100, 10000, 1000000]:

win = 0

GAMES = int(i)

for i in range(GAMES):

win += int(random.randint(1,6)==6)

print("No. of trials: {}, P(Win) : {} ".format(GAMES, win/GAMES))

No. of trials: 5, P(Win) : 0.2

No. of trials: 10, P(Win) : 0.3

No. of trials: 100, P(Win) : 0.16

No. of trials: 10000, P(Win) : 0.1677

No. of trials: 1000000, P(Win) : 0.16702

**What can we infer from this simulation ?**

* This simulation shows that you can never be certain that chances of winning are exactly 16.6%
* If the number of experiments tends to the infinity then the probability becomes 0.166
* Experimental probability is unstable and converges to theoretical probability as more experiments are conducted

**Casino Vegas Game 3 (Marginal, Independent events, mutually exclusive, Joint, Conditional)**

* At the casino, you want to play another game, which involves cards.
* The dealer picks one card out of the deck
* You have to guess which category the card belongs to
* The categories to choose from are:
  1. Face Cards
  2. Color card (Red/Black)
  3. Number card
  4. Odd number
  5. Even number cards.
  6. Suite (Spade, Diamond, Clubs, Hearts)

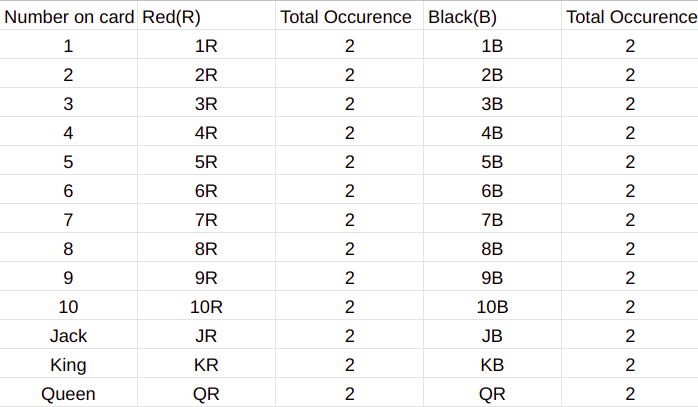
Lets say you chose a red colour card and your friend chose even card

You want to know the chances of both of you winning the game

**What is the experiment, event and sample space in this case ?**

* Experiment : Picking a card from the deck
* Sample space : All cards in deck (Total 52)
* Event :
  + Card being even numbered and red in colour
  + Lets break this into 2 events
    1. The card is even numbered - Denoted by A
    2. It is red - Denoted by B

Lets see the possible outcomes of the experiement



**Marginal Probability**

**Now, What about probability of getting a card with even number ?**

* P(Card = Even) = P(A)
* P(A) = Total occurences of even card/Total Sample size
* Total occurences of even card = Sum of even rows = 20
* So, P(A) = 20/52 = 5/13
* It doesn’t depend upon the colour of cards

**Whats the probability of getting a red card ?**

* P(Red card) = P(B)
* P(B) = (Total occurences of red card)/Total sample size
* Total occurences of red card = Sum of the 3rd col = 26
* So P(B) = 26/52 = 1/2
* It doesn’t depend on what number the card has

The probability of such events which does not get affected by other variables is called **Marginal Probability**

**Joint probability**

Now we want to find P(A and B)P(A and B)

The probability of multiple events occuring together is called **joint probability**

* It is the probability of the intersection of two or more events. P(A∩B)P(A∩B)
* Joint probability is symmetric, meaning P(A∩B)=P(B∩A)P(A∩B)=P(B∩A)

**Can you calculate P(A and B) through this ?**

* P(Even ∩ Red) = P(A ∩ B)
* P(A∩B)=Total occurences of Red and Even number togetherTotal sample spaceP(A∩B)=Total occurences of Red and Even number togetherTotal sample space
* P(A∩B)=Total occurences of A and B togetherTotal sample spaceP(A∩B)=Total occurences of A and B togetherTotal sample space

So $P(A ∩ B) = \frac{10}{52} $

**Union**

**But now you also want to find out the probability of either one of you two winning**

Find: **P(Even card OR Red card) = P(A OR B) = P(A ∪ B)**

Going back to our outcomes,

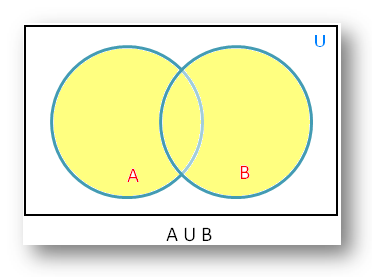
* Desired outcome:
  + The card is **either red or even or both**
  + i.e. Either A or B or both are true

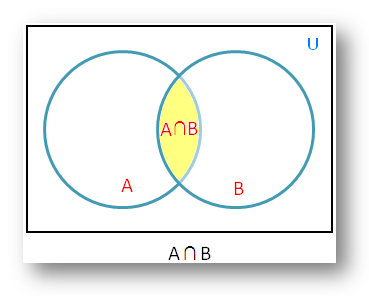
1. Both A and B are true:  
   - This is the case with A∩BA∩B  
   - Total: 10 outcomes
2. A is true (B may or may not be true)  
   - All even cards (Red/Black)  
   - Total: 20 outcomes
3. B is true (A may or may not be true)  
   - Includes all red cards (Number/face)  
   - Total: 26 outcomes

**So what should be**A∪BA∪B**?**

* **Is it 20+26+10 = 46 ?**

**Whats the problem with this ?**

* Lets understand this through venn diagram
* **Union**:  
  
* **Intersection**:



* When we do A+B+A∩BA+B+A∩B we are adding the intersection area thrice:
  1. Through term A
  2. Through term B
  3. Through term A ∩ B itself

**So now what do you think**A∪BA∪B**is going to be ?**

* $(A \cup B) = A + B - (A \cap B) $

Hence

P(Even OR red card)=P(A∪B)P(Even OR red card)=P(A∪B)

=P(A)+P(B)−P(A∩B)=P(A)+P(B)−P(A∩B)

=(20/52)+(26/52)−(10/52)=(20/52)+(26/52)−(10/52)

=36/52=36/52

**Disjoint events**

**Now lets suppose, you choose Jack and your friend chooses even numbered card**

**But can a card be Jack and even-numbered at the same time ?**

* Obviously not

Two events **that cannot occur at the same time or simultaneously** are said to be **mutually exclusive**.

Mutually exclusive events are also called **disjoint events**.

**Probability of Mutually Exclusive Events is:**

* **P (A and B) = 0**
* P(A∩B)=ϕP(A∩B)=ϕ

**Independent events**

Now lets say you chose a red card and won the game

Because of this you play another game and again chose a red card

Remember, the dealer puts the chosen card back into the deck

So, The experiment and sample space remains the same for both the events

**What is the probability that you will win again ?**

**What will be the event ?**

* Getting a red card in second game given you got a red card in the first one
* Breaking this event into 2:
  + Event A: Getting a red card in the first game
  + Event B: Getting a red card in the second game

**But will B actually depend on A ?**

* No matter what you get in first game, the second game will obviously not be affected by it
* Because the sample space remains the same

Such events whose occurrence is not dependent on any other event are called **Independent events**.

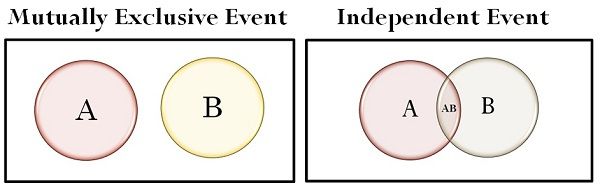
So, P(Winning second game) = P(Getting a red card) = 26/52=1/2

**Now you also want to find the chances of winning both the games i.e. P(A ∩ B)**

**But then, are independent events mutually exculsive as well ?**

* No
* If A occurs then B cannot occur and vice-versa
* This means that occurence of A and B affect each other
* So they are not independent events
* Also independent events need not be mutually exclusive
* This means they can occur simultaneously or one after the other or not at all.

Lets look at a venn diagram to understand this



**So, what do you think**P(A∩B)P(A∩B)**this is going to be ?**

* We have already established that A and B are independent
* P(A) = 1/2
* P(B) = 1/2
* So, \**P(A ∩ B) = P(A)P(B) = 1/4*

**Conditional Probability**

Now suppose you bet that the card is Diamond

Lets suppose you somehow got to know that the card is going to be Red

**What are the chances that you will win ?**

**What is the experiment, event and sample space in this case ?**

* Experiment : Picking a card from the deck
* Sample space : All cards in deck (Total 52)
* Event :
  + Here there are 2 events
    1. The card is Diamond - Denoted by A
    2. The card is red coloured - Denoted by B

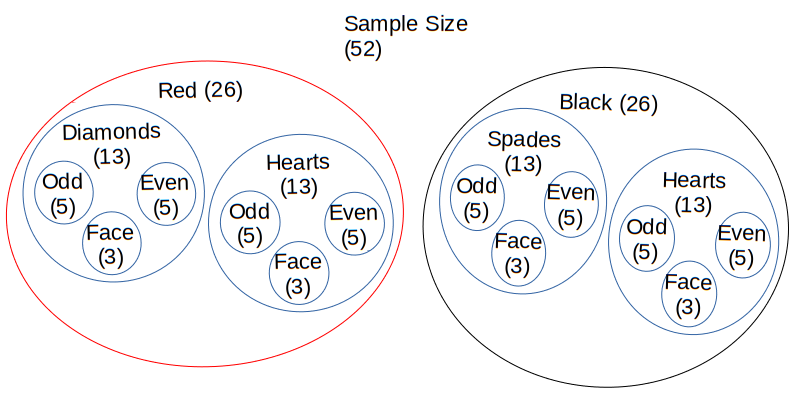
We want to find P(A given B occurs)P(A given B occurs)

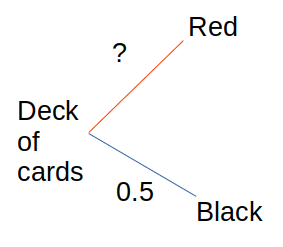
The probability of an event occurring, based on the occurrence of a previous event is called **Conditional probability**

* Notation : P(A|B)P(A|B) : A and B are the events.
* Meaning : Probability that event A happens, given that event B has already happend.

Lets calculate this using venn diagrams

**What is the probability of a card being red ?**





P(Red)=P(B)P(Red)=P(B)

P(B)=26/52P(B)=26/52

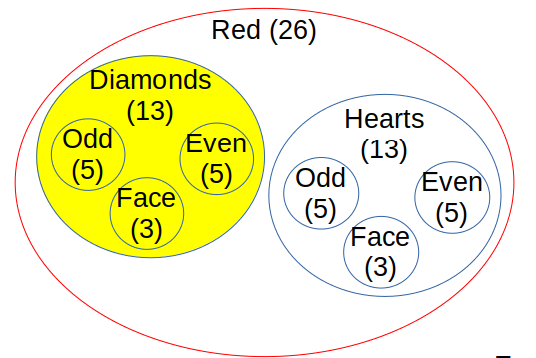
P(B)=0.5P(B)=0.5

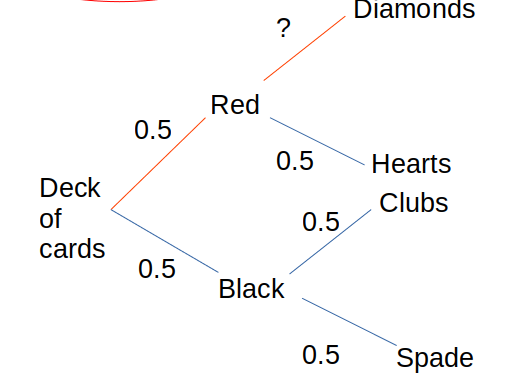
The sample space now reduces to 26 cards

**What is P(Card=Diamond) in this sample space?**

Out of 26 red cards 13 are Diamond

**So what is**P(Card=Diamond|Card=Red)P(Card=Diamond|Card=Red)**?**





* P(Card=Diamond|Card=Red)=P(A|B)P(Card=Diamond|Card=Red)=P(A|B)
* Reducing sample space from **52 to 26 red cards**

P(A|B)=(13/26)=1/2=0.5P(A|B)=(13/26)=1/2=0.5

Notice how P(A) changed

In original sample space P(A) = 13/52 = 1/4

Now in the reduced sample space it becomes 1/2

**But can we devise a mathematical formula for this ?**

* Notice that in the first step we are reducing the sample space
* After that the sample space **reduces to the sample space of Event B**
* In last step we take **intersection of A and B**

**So, How can we calculate**P(A|B)P(A|B)**?**

P(A|B)=P(A∩B)P(B)P(A|B)=P(A∩B)P(B)

**Lets play this game again now**

This time lets choose an even card

Suppose You came to know that the card dealer picked is red

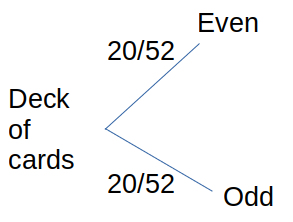
**What are your chances of winning now ?**

* The experiment and sample space remains the same
* Events:
  1. Card picked is even numbered : A
  2. Card picked is red : B

Find: P(Card=Even|Card=Red)P(Card=Even|Card=Red)

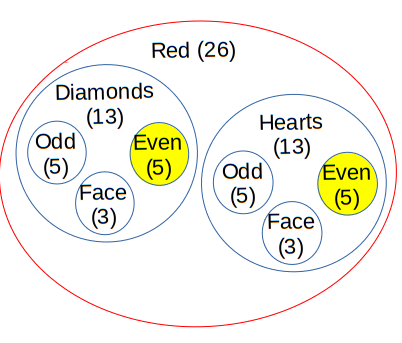
Lets draw tree diagrams again to understand this

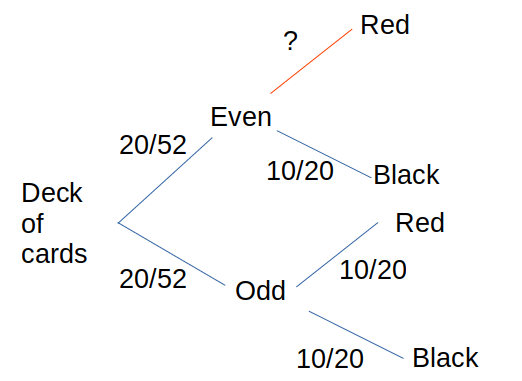
We already have P(Card = Red) i.e. P(B)



The sample space reduces

**Whats P(Card=Even) in this reduced sample space?**





Out of 26 red cards:

* 10 are even
* 10 are odd
* 6 are face

**So what is the probability that we win ?**

P(Card=Even|Card=Red)=P(Card=Even and Card=Red)P(Card=Red)P(Card=Even|Card=Red)=P(Card=Even and Card=Red)P(Card=Red)

P(A∩B)=Card=Even and Card=RedTotal cardsP(A∩B)=Card=Even and Card=RedTotal cards

=>P(A∩B)=10/52=>P(A∩B)=10/52

P(B)=26/52P(B)=26/52

=>P(A|B)=(10/52)/(26/52)=>P(A|B)=(10/52)/(26/52)

P(A|B)=10/26P(A|B)=10/26

P(A|B)=5/13=0.38P(A|B)=5/13=0.38

**But wait, whats**P(A)P(A)**here ?**

* P(A) = P(Card=Even)
* P(A) = 20/52
* P(A) = 10/26

This means,P(A|B)=P(A)P(A|B)=P(A)

=>P(A∩B)P(B)=P(A)=>P(A∩B)P(B)=P(A)  
=>P(A∩B)=P(A)∗P(B)=>P(A∩B)=P(A)∗P(B)

But we know this is possible only when events are independent

Thus, **Card=Even and Card=Red are independent events**

**Law of total probability**

Now lets take another example

**Lets say the machines at the casino are sourced from two different manufactures**

* 60% of the machines are delivered from factory A
* Rest 40% are delivered from factory B.
* Machines made in X, works perfectly in 99% of cases.
* Machines made in Y, works perfectly in 95% of cases.

**What is the probability that a machine works perfectly?**

**What are the events in this case ?**

* Event A - Machine works perfectly
* Event Bx - Machine is made in Factory x
* Event By - Machine is made in Factory y.

**So what do you have to find?**

Find : P(A)

**And what is the sample space ?**

* S=Bx∪ByS=Bx∪By

**Can events Bx and By occur together ?**

* No, they are mutually exhaustive
* Bx∩By=ϕBx∩By=ϕ

**So how can we calculate**P(A)P(A)**?**

* P(A)=P(A∩Bx)+P(A∩By)P(A)=P(A∩Bx)+P(A∩By)
* P(A)=P(A|Bx).P(Bx)+P(A|By).P(By)P(A)=P(A|Bx).P(Bx)+P(A|By).P(By)
* This is known as the **law of total probability**.
* P(A)=0.99∗0.6+0.95∗0.4P(A)=0.99∗0.6+0.95∗0.4
* P(A)=0.974P(A)=0.974

This is all about probability theory for now

As we move towards the lectures involving DS/ML you will realise how this forms the basis for many of the algorithms

Last changed by

**Scaler Topics**Scaler Topics

 338

Published on [**HackMD**](https://hackmd.io/)