# Probability and Statistics
# EC184406

Computer Engineering, Institut Teknologi Sepuluh Nopember

# Basic Probability

# Basics of probability

- By definition, $\boldsymbol{P(A)}$ corresponds to the probability that among all the possible events, event $\boldsymbol{A}$ occurs.
- This probability has to be a number between $0$ and $1$
- being 0 as it is impossible that this event occurs, and 1 as it is certain that this event occurs.

# Basics of probability

- The simplest example occurs when flipping a coin: the chances of the event being **heads** are the same as being **tails : 50%**. This probability can be expressed as a quotient:

$$P(heads) = \frac{\text{desired event}}{\text{possible events}}$$

- where there's 1 possibility that the event is the one we desire (heads), and there are 2 possible events when flipping a coin. Then,


$$P(heads) = P(tails) = \frac{1}{2}$$


- The sum of all the possible events is called **sample space** ($\Omega$). The probability of the sample space will always be 1 ($P(\Omega) = 1$).
- In the coin example, sample space would be $\Omega = \{heads, tails\}$
- When we flip the coin, we are certain that the resulting event is going to be **heads** or **tails**, so it is always going to be an event from $\Omega$.

- The set with any events is called **empty set** ($\emptyset$). The probability of the empty set will always be 0 ($P(\emptyset) = 0$).
- In the coin example, when we flip it, we are certain that there's going to be a resulting event **heads** or **tails**, so it is never going to be empty.

- For example, when throwing a dice the sample space would have six possible events: $\Omega = \{1,2,3,4,5,6\}$.
- Then the probability of having a result of 1 would be 1 possibility divided by 6 possible events:

$$P(1) = \frac{1}{6}$$

- which is the same for the six numbers of the dice. The sum of all the probabilities of the possible events would be:

$$P(1) + P(2) + P(3) + P(4) + P(5) + P(6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{6}{6} = 1$$

- If, for example, we would want to know the probability of having a result of 2 or 3 with the dice, that would be:

$$P(2 \cup 3) = \frac{\text{desired events}}{\text{possible events}} = \frac{2}{6}$$

- **NOTE: we use the sign ∪ to denote OR, and the sign ∩ to denote AND**

- See that the probability P(2∪3), in this case, is the same as doing: $P(2) + P(3) = \frac{2}{6}$

- Generally, the probability of one event or another would be:

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

- In this case, event 2 and event 3 happening at the same time is impossible, so $P(A \cap B) = 0$.

- Finally, let's resume the axioms seen in this section:

  - $0 \leq P(A) \leq 1$
  - $P(\Omega) = 1$
  - $P(\emptyset) = 0$
  - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

# Joint probability

- A joint describes the probability of two events happening, that can be expressed as

$$P(A \cap B) = P(A, B) = P(A\ AND\ B)$$

- If **A** and **B** are independent, then

$$P(A, B) = P(A) \cdot P(B)$$

# Joint probability

- For example, if we had two dice, now $P(2 \cap 3)$ would represent the probability that we obtain a result of 2 on one dice and a result of 3 on the other dice.

- The dice do not depend on each other, so the results are independent. That's why the probability of obtaining a 2 and a 3 will be

$$P(2 \cap 3) = P(2) \cdot P(3) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

# Conditional probability

- Conditional probability stands for the probability of an event **A** happening if, previously, another event **B** happened. This is expressed as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

# Independent event

- if A and B are independent, then $P(A \cap B) = P(A) \cdot P(B)$ and

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \cdot P(B)}{P(B)} = P(A)$$

- which means that we are not interested in what happened with B, to calculate what happened with A.
- In the example of two dice, if one dice got a 3, the probability of the other dice getting a 2 would be

$$P(2|3) = \frac{P(2 \cap 3)}{P(3)} = P(2) = \frac{1}{6}$$

# Independent event

- If $A$ and $B$ are not independent,
  then $P(A \cap B)$ needs to be calculated.
- For example, if we have a bag with 5 balls, **3 are red** and **2 are blue**. If we first take one ball without looking at them, the probability that it is a red one would be

$$P(red) = \frac{3}{5}$$

# Dependent Event

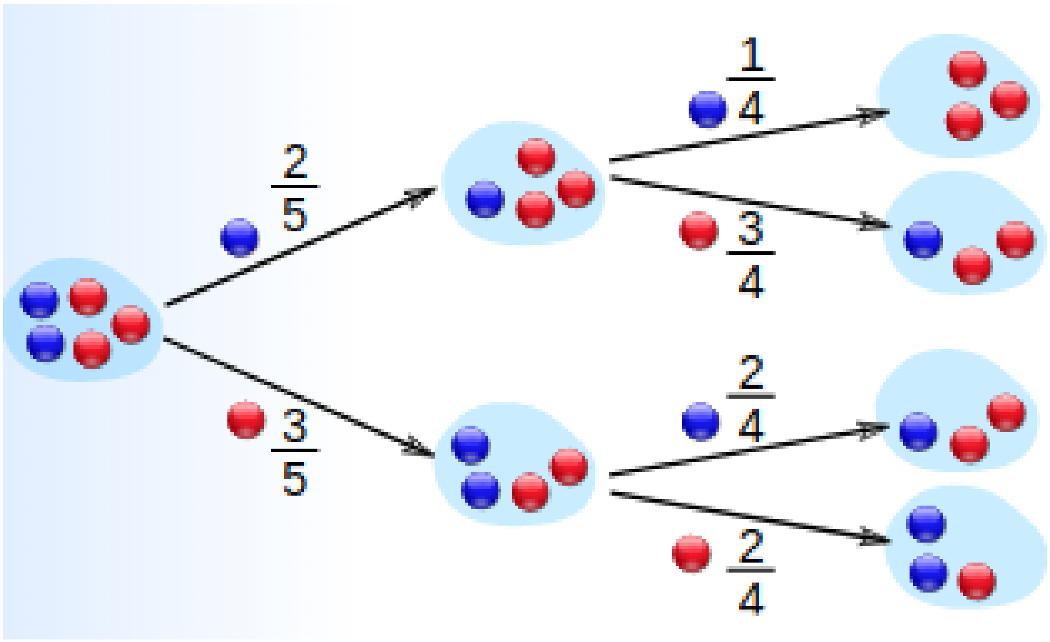
- The probability of taking a blue one would be

$$P(blue) = \frac{2}{5}$$

- However, if first we took one blue ball, 3 red balls and 1 blue ball would remain in the bag. What's the probability of taking a red ball in those conditions?

$$P(red|blue) = \frac{3}{4}$$

# Dependent Event




What is the probability of drawing two blue marbles without replacement?

Event A: Drawing 1st Blue Marble
Event B: Drawing 2nd Blue Marble

$$P(A \cap B) = P(A) \,.\, P\,(B\ after\ A)$$

$$P(A \cap B) = \frac{2}{5} \,.\, \frac{1}{4} = \frac{1}{10}$$

This is what we called chainrule!

# Chain rule

- From the definition of conditional probability, there's the following definition of a chain rule:

$$P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

- In the red and blue balls example, the probability of taking a blue ball and then a red one would be

$$P(blue \cap red) = P(red|blue) \cdot P(blue) = \frac{3}{4} \cdot \frac{2}{5} = \frac{6}{20}$$

# Bayes' rule

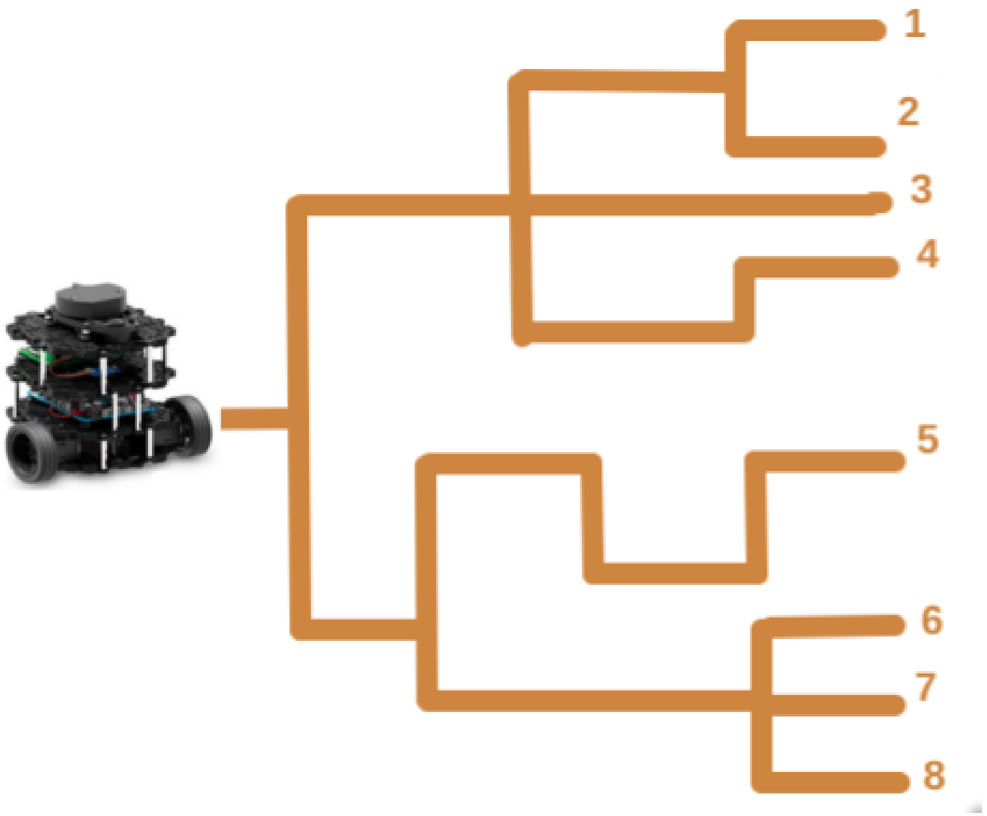
- Finally, all these concepts can be put together in one fundamental rule of probability: **Bayes' rule**:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

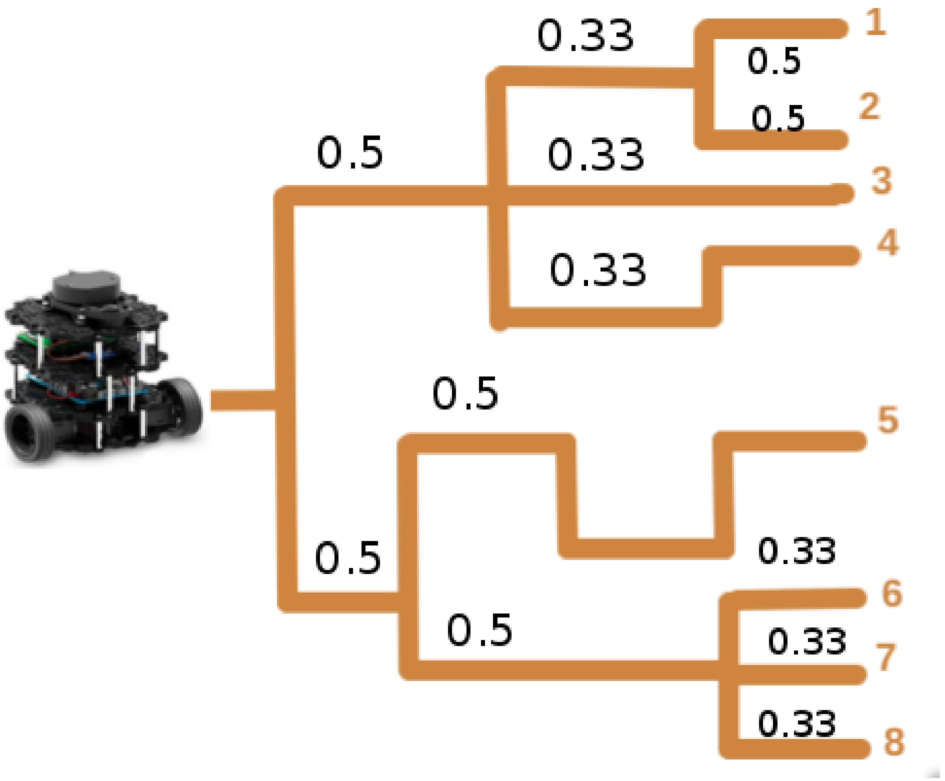
- where $P(A)$ is called the prior, $P(A|B)$ the posterior, $P(B|A)$ the likelihood, and $P(B)$ the evidence.
- This rule is helpful when $P(B|A)$ is easier to obtain than $P(A|B)$.

# Exercise

- Let's imagine we have a robot that is trying to escape from a maze. The robot goes forward by default, and when it reaches a crossroad, it has to decide between going left, straight, or right.




calculate the probability of the eight possible outcomes by applying conditional probability each time the robot passes a crossroad




$$P(\text{Door } 1) = P(left|left) \cdot P(left|left) \cdot P(left)$$

$$P(\text{Door } 2) = P(right|left) \cdot P(left|left) \cdot P(left)$$

$$P(\text{Door } 3) = P(front|left) \cdot P(left)$$

$$P(\text{Door } 4) = P(right|left) \cdot P(left)$$

$$P(\text{Door } 5) = P(left|right) \cdot P(right)$$

$$P(\text{Door } 6) = P(left|right) \cdot P(right|right) \cdot P(right)$$

$$P(\text{Door } 7) = P(front|right) \cdot P(right|right) \cdot P(right)$$

$$P(\text{Door } 8) = P(right|right) \cdot P(right|right) \cdot P(right)$$

# Python snippet

```
P_door1 = 0.5*0.33*0.5
print("P(Door 1) = ",P_door1)
P_door2 = 0.5*0.33*0.5
print("P(Door 2) = ",P_door2)
P_door3 = 0.5*0.33
print("P(Door 3) = ",P_door3)
P_door4 = 0.5*0.33
print("P(Door 4) = ",P_door4)
P_door5 = 0.5*0.5
print("P(Door 5) = ",P_door5)
P_door6 = 0.33*0.5*0.5
print("P(Door 6) = ",P_door6)
P_door7 = 0.33*0.5*0.5
print("P(Door 7) = ",P_door7)
P_door8 = 0.33*0.5*0.5
print("P(Door 8) = ",P_door8)
```

```
('P(Door 1) = ', 0.0825)
('P(Door 2) = ', 0.0825)
('P(Door 3) = ', 0.165)
('P(Door 4) = ', 0.165)
('P(Door 5) = ', 0.25)
('P(Door 6) = ', 0.0825)
('P(Door 7) = ', 0.0825)
('P(Door 8) = ', 0.0825)
```

With these results, the robot is more likely to be at door 5.