Quantifying Uncertainty

**Introduction**

We are interested in intelligent agents that take rational decision based on specific goals and the likelihood to achieve them.

Probability theory is a tool for representing and summarising the uncertainty of the agent’s world.

A rational agent selects the next action based on the average of all the possible outcomes, weighted by probabilities

**Variables and Events**

A random variable can represent any property or feature of a process

E.g. the possible outcome of a die roll



One such outcome could be the event

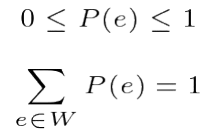


In this case the probability of the event is



**Probability**

Basic axioms



The probability is always between 0 and 1, and the sum of the probabilities of all the possible events e in a given world “W” is 1.

In the die-roll case indeed



**Multiple Variables**

We can also calculate joint probability

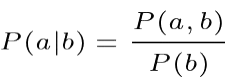


More generally we can write the probability of multiple events as



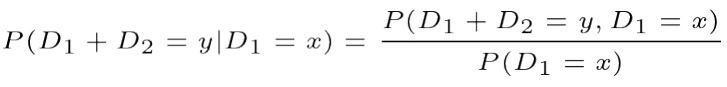
**Conditional Probability**

The conditional probability of a “given” b

Probability of a given b = joint probability of a and b divided by the probability of b

Is the probability of event ***a*** to be true given event ***b*** is true.

E.g. the probability of the sum of two dice given the outcome of the first dice

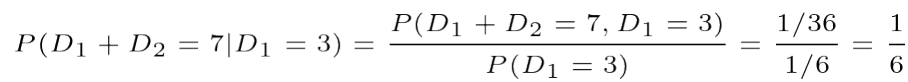


Example



There are 6 x 6 = 36 possible outcomes, but only 1 case for (D2 = 4) in which 

Therefore



(1/6+1/6) + (2/6+2/6) + (3/6+3/6) + (4/6+4/6) + (5/6+5/6) + (6/6+6/6)

Independence

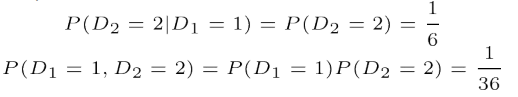
Two events a and b are independent if



Or equivalently



For example



**Conditional Independence**

Two events a and b are conditionally independent given a third event c if



Or equivalently



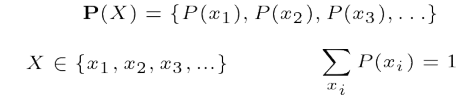
For example





**Probability Distributions**

The probability distribution P(X) is the set of probabilities assigned to each possible value of X



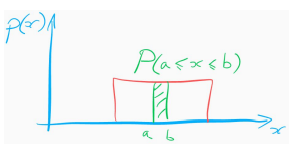
For example the die-roll has a (uniform) probability distribution



**Probability Density Function**

For a continuous variable, the distribution is represented by a probability density function (pdf)



This distinction is due to the infinite number of possible values of a continuous variable 

Bayes Rule

From the definition of conditional probability, we can write the **product rule**



But because P(a, b) = P(b, a), we can also write the **Bayes rule**



Another way of writing the same



Law of Total Probability (marginalisation)