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

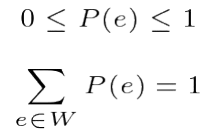


In this case the probability of the event is



**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

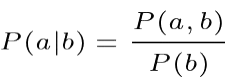


More generally we can write the probability of multiple events as



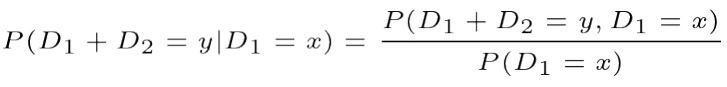
**Conditional Probability**

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

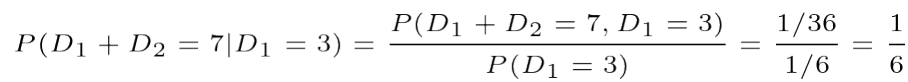


Example



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

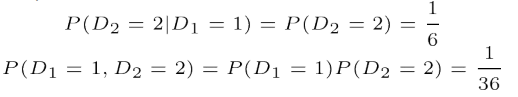
Two events a and b are independent if



Or equivalently



For example



**Conditional Independence**

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



Or equivalently



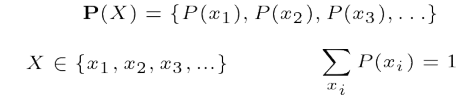
For example





**Probability Distributions**

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



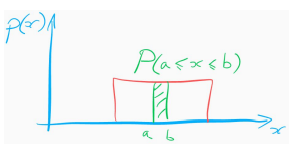
For example the die-roll has a (uniform) probability distribution



**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**

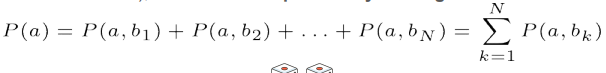


Another way of writing the same



**Law of Total Probability (marginalisation)**

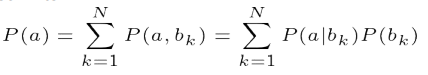




For example, in a two-dice roll 



We can also write



**Example (from Pearl & Mackenzie’s “The Book of Why”)**

Probability of 40 years old woman of getting breast cancer within a year



Probability of positive mammogram test given that one has cancer (sensitivity)

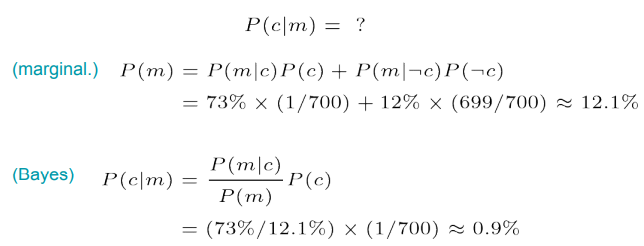


Probability of positive mammogram test given that one doesn’t have cancer (false positive rate)



What is the probability of breast cancer given a positive test?





**Expected Value**

Expected value or mean of a random variable X



Expected value of a function g(X)



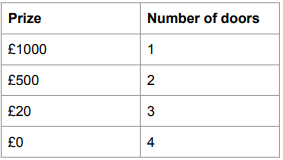
Expected value of Y conditional on X = x



Example

Imagine a game show with some money prizes hidden behind 100 doors

What is the expected value of the prize for opening one of them?

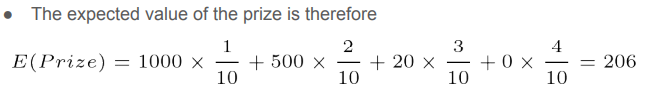


1/10 = £1000

2/10 = 500

3/10 = £20

4/10 = £0



On average people would earn £206 because of the probability of earning money from opening the doors.

**Variance, standard deviation and Covariance**

The variance of a random variable indicates how much “spread out” it is from the mean 



The standard deviation is simply the square root of the variance

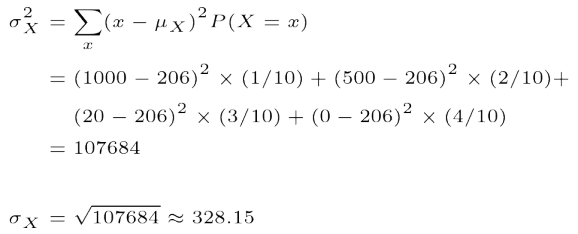


The covariance of two random variable indicates the degree to which they vary together, or are “associated”



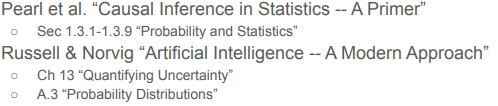
**Example**

**For the 10-doors game show, where  = 206**

****

There is a variation of 328.15 across each people getting £206 so it is 206 +- 328.15

Reading



<https://attendance.lincoln.ac.uk/>

Workshop: 967926