Mathematical Concepts (G6012)

Lecture 18

Thomas Nowotny
Chichester I, Room CI-105
Office hours: Tuesdays 15:00 - 16:45
T.Nowotny@sussex.ac.uk

PROBABILITY THEORY

Disclaimer

- We are only going to do so-called discrete probability spaces here. Some of what follows does not directly generalize to continuous probability spaces (e.g. the power set as the domain of the probability measure …)

Power set

For a set S the power set $\mathcal{P}(S)$ is the set of all subsets, e.g.

$$S = \{1, 2, 3\}$$

$$\mathcal{P}(S) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \underbrace{\{1,2,3\}}_{S}\}$$

Probability measure

Definitions

- Ω : Set of "elementary events"
- A **probability measure** is a function

$$P : \mathcal{P}(\Omega) \to [0,1]$$
$$\omega \mapsto P(\omega)$$

Typical example

- Outcome of throwing a die (singular of dice …):

$$\Omega = \{1,2,3,4,5,6\}$$

$$P(\{i\}) = \frac{1}{6}$$
$$P(\{i,j\}) = \frac{2}{6} = \frac{1}{3}$$
$$P(\Omega) = 1$$

Probability

We call the subsets of Ω **events** and for an event $A \subset \Omega$ we call $P(A)$ the **probability** of the event A .

Note: This is all there is, if people write $P(x < 5)$ this really means

$$P(\{x \in \Omega : x < 5\})$$

where (Ω, P) is an underlying **probability space**.

Probability space

For (Ω, P) to be a proper probability space, the following conditions must hold:

$$P(\Omega) = 1$$

$$P(\emptyset) = 0$$

$$P(A \cup B) = P(A) + P(B) \quad \text{if} \quad A \cap B = \emptyset$$

(Additivity)

Typical example: dice

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

$P(\{i\}) = \frac{1}{6}$ for $i = 1, \ldots, 6$

Event $A = \{\text{number even}\}$

$$\begin{aligned} P(A) &= P(\{2, 4, 6\}) \qquad (\{2,4,6\} = \{2\} \cup \{4\} \cup \{6\}) \\ &= P(\{2\}) + P(\{4\}) + P(\{6\}) \\ &= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2} \end{aligned}$$

Probability of the complement

Let $A \subset \Omega$ be an event.

$A^C = \Omega \backslash A$ or, equivalently, $A \cup A^C = \Omega$

and $A \cap A^C = \emptyset$

Therefore,

$$P(A \cup A^C) = P(A) + P(A^C) = P(\Omega) = 1$$

$$P(A^C) = 1 - P(A)$$

Examples for dice: **BB**

BB Example complements

Example 1:

$$P(\{\text{even}\} = 1 - P(\{\text{odd}\})$$

Example 2:

$$P(\{1, 2, 3, 4, 5\}) = 1 - P(\{6\})$$

Example 3: Throwing 10 dice

$$P(\{\text{at least one 1}\}) = 1 - P(\{\text{no 1}\})$$

$$= 1 - \prod_{i=1}^{10} P(\{\text{die } i \text{ not } 1\}) = 1 - \left(1 - P(\{1\})\right)^{10}$$

Independence

Definition: Two events A and B are **independent** if

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

Warning: Often confused with A and B being disjoint or exclusive, i.e. $A \cap B = \emptyset$

Typical example: two dice

$$\Omega = \{(i,j) : i = 1, \ldots 6, j = 1 \ldots 6\}$$

Result die 2

Result die 1

$$P(\{(i,j)\}) = \frac{1}{36} \quad \text{for all} \quad (i,j) \in \Omega$$

Note: Defining P on "elementary events" is enough because of additivity!

Example continued

$$A = \{i = 1\} = \{(i,j) : i = 1\}$$ "First die shows a 1"

$$B = \{j = 3\} = \{(i,j) : j = 3\}$$ "Second die shows a 3"

$$P(A) = P(\{(1,1)\}) + \ldots + P(\{(1,6)\})$$
$$= \frac{1}{36} + \ldots + \frac{1}{36} = \frac{6}{36} = \frac{1}{6}$$

Similarly, $P(B) = \frac{1}{6}$

Example continued

$$P(A \cap B) = P(\{(1,3)\}) = \frac{1}{36}$$

$$P(A) \cdot P(B) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

Therefore, A and B are independent:

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

Other example (one die)

$$A = \{i \text{ is even}\} \quad P(A) = \frac{1}{2}$$

$$B = \{i \geq 2\} \quad P(B) = \frac{5}{6}$$

$$P(A \cap B) = P(\{2, 4, 6\}) = \frac{1}{2}$$

$$P(A) \cdot P(B) = \frac{1}{2} \cdot \frac{5}{6} = \frac{5}{12}$$

A and B are not independent!

But … (still one die)

$$A = \{i \text{ is even}\} \quad P(A) = \frac{1}{2}$$

$$B = \{i \geq 3\} \quad P(B) = \frac{4}{6} = \frac{2}{3}$$

$$P(A \cap B) = P(\{4, 6\}) = \frac{2}{6} = \frac{1}{3}$$

$$P(A) \cdot P(B) = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$$

So, here, A and B are independent!

Useful tool: tree graphs

Throwing two dice again:

$$\Omega = \{(i,j) : i = 1,\ldots 6, j = 1\ldots 6\}$$

Result die 1 (for i)

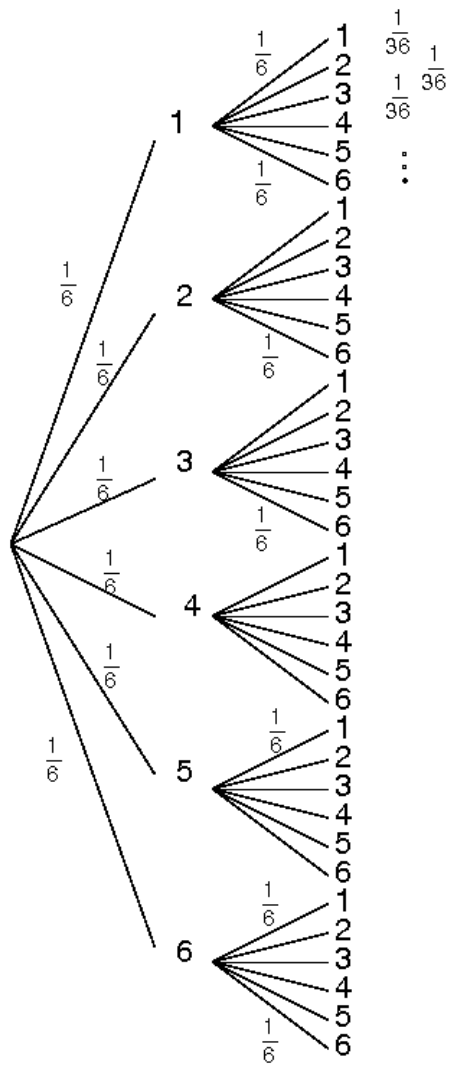
Result die 2 (for j)

We can understand this as the combination of two independent experiments (see above):

$$P(\{i\}) = \frac{1}{6} \qquad P(\{j\}) = \frac{1}{6}$$

& construct a tree graph to calculate probabilities: **BB**

BB Tree graph for 2-dice exp.



Each branch is one outcome.

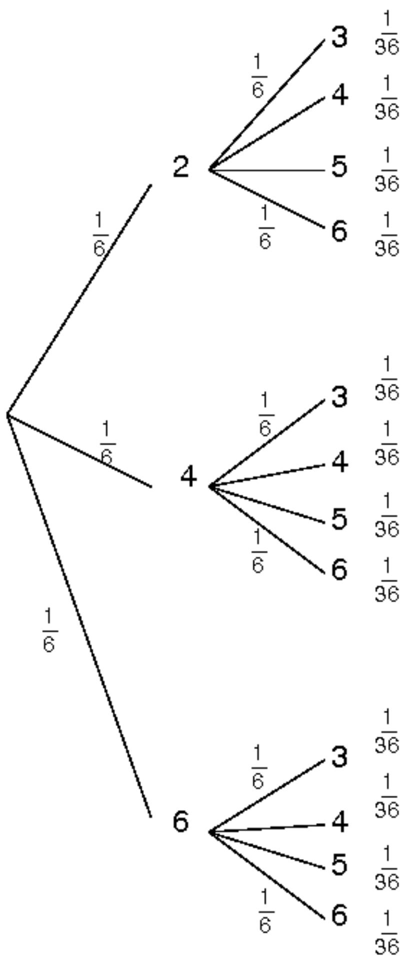
At each branch we note the probability of that outcome ($\frac{1}{6}$).

The total probability of the outcomes is the product of the probabilities along branches ($\frac{1}{36}$) .

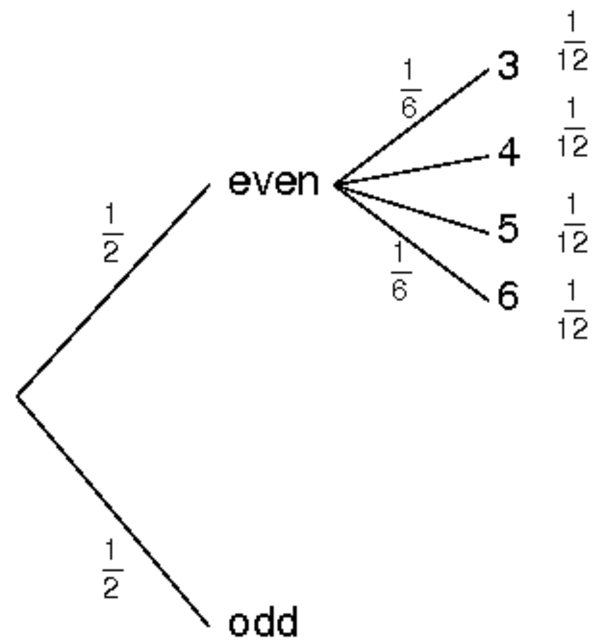
I practice we would draw only branches we are interested in …

BB Tree graphs, 2 dice

$$P(\{(i,j) : i \text{ even}, j \geq 3\})$$



$$12 \cdot \frac{1}{36} = \frac{1}{3}$$



$$4 \cdot \frac{1}{12}$$

In both cases the result is $\frac{1}{3}$
(as it should)

Tree graph summary

- Very useful to find all cases & calculate probabilities correctly
- Only works if there are independent stages (like throwing one die, then the other)
- It is important to do it tidily to not miss anything (!)
- One can often save time and effort by drawing reduced graphs.

Conditional probabilities

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

Read **"Probability of A given B ."**

It is called a **conditional probability**.

For independent events:

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

Example: One die

$$A = \{i \text{ even}\} \qquad P(A) = \frac{1}{2}$$

$$B = \{i \geq 2\} \qquad P(B) = \frac{5}{6}$$

$$P(A \cap B) = P(\{2, 4, 6\}) = \frac{1}{2}$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{2}}{\frac{5}{6}} = \frac{1}{2} \cdot \frac{6}{5} = \frac{3}{5}$$

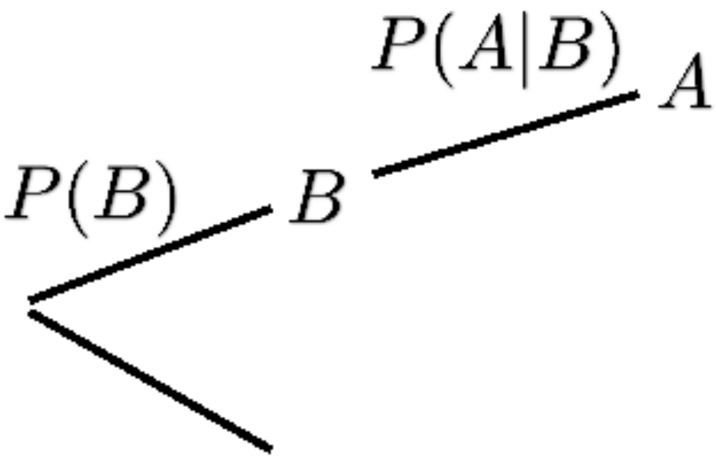
Conditional probabilities

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

$$\Leftrightarrow \; P(A \cap B) = P(A|B)\, P(B)$$

In a sense, this is what we have been doing with the tree graphs all along:

Tree graph with conditional probabilities



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

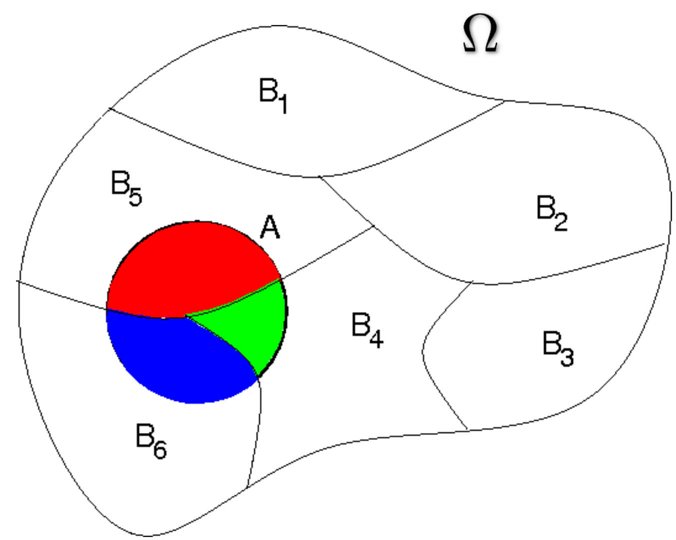
Conditional probabilities

Furthermore, if B_i is a set of disjoint events that covers Ω , i.e.

$$B_i \cap B_j = \emptyset, \quad i \neq j \qquad \text{and} \qquad \bigcup_{i=1}^{n} B_i = \Omega$$

$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$$

$$= \sum_{i=1}^{n} P(A \cap B_i)$$



Bayes rule

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

Proof:

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

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

In other words Bayes theorem = definition of conditional probability.