

MAT1830

Lecture 21: Probability

Probability - why should you care?

Probability is really important in computer science for

- ▶ understanding the performance of random algorithms
- ▶ modelling user behaviour and system requirements
- ▶ understanding errors and malfunctions
- ▶ determining the security of cryptosystems
- ▶ machine learning, reliability modeling, simulation algorithms, data mining, speech recognition etc etc.

Probability gives us a way to model random processes mathematically. These processes could be anything from the rolling of dice, to radioactive decay of atoms, to the performance of a stock market index. The mathematical environment we work in when dealing with probabilities is called a probability space.

21.1 Probability spaces

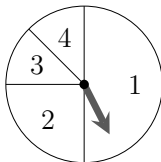
We'll start with a formal definition and then look at some examples of how the definition is used.

A *probability space* consists of

- a set S called a *sample space* which contains all the possible *outcomes* of the random process; and
- a *probability function* $\Pr : S \rightarrow [0, 1]$ such that the sum of the probabilities of the outcomes in S is 1.

Each time the process occurs it should produce exactly one outcome (never zero or more than one). The probability of an outcome is a measure of the likeliness that it will occur. It is given as a real number between 0 and 1 inclusive, where 0 indicates that the outcome cannot occur and 1 indicates that the outcome must occur.

Example.



The spinner above might be modeled by a probability space with sample space $S = \{1, 2, 3, 4\}$ and probability function given as follows.

$$\Pr(s) = \begin{cases} \frac{1}{2} & \text{for } s = 1 \\ \frac{1}{4} & \text{for } s = 2 \\ \frac{1}{8} & \text{for } s = 3 \\ \frac{1}{8} & \text{for } s = 4. \end{cases}$$

It can be convenient to give this as a table:

| s | 1 | 2 | 3 | 4 |
|----------|---------------|---------------|---------------|---------------|
| $\Pr(s)$ | $\frac{1}{2}$ | $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{1}{8}$ |

Example. Rolling a fair six-sided die could be modeled by a probability space with sample space $S = \{1, 2, 3, 4, 5, 6\}$ and probability function \Pr given as follows.

| | | | | | | |
|----------|---------------|---------------|---------------|---------------|---------------|---------------|
| s | 1 | 2 | 3 | 4 | 5 | 6 |
| $\Pr(s)$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ |

A sample space like this one where every outcome has an equal probability is sometimes called a *uniform sample space*. Outcomes from a uniform sample space are said to have been taken *uniformly at random*.

21.2 Events

An *event* is a subset of the sample space.

An event is just a collection of outcomes we are interested in for some reason.

Example. In the die rolling example with $S = \{1, 2, 3, 4, 5, 6\}$, we could define the event of rolling at least a 3. Formally, this would be the set $\{3, 4, 5, 6\}$. We could also define the event of rolling an odd number as the set $\{1, 3, 5\}$.

The probability of an event A is the sum of the probabilities of the outcomes in A .

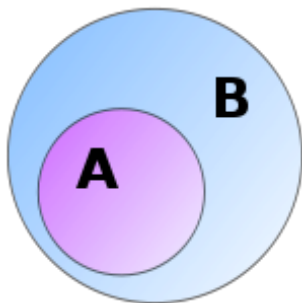
21.2 Events

An *event* is a subset of the sample space.

An event is just a collection of outcomes we are interested in for some reason.

Example. In the die rolling example with $S = \{1, 2, 3, 4, 5, 6\}$, we could define the event of rolling at least a 3. Formally, this would be the set $\{3, 4, 5, 6\}$. We could also define the event of rolling an odd number as the set $\{1, 3, 5\}$.

The probability of an event A is the sum of the probabilities of the outcomes in A .



Example. In the spinner example, for the event $A = \{1, 2, 4\}$, we have

$$\begin{aligned}\Pr(A) &= \Pr(1) + \Pr(2) + \Pr(4) \\ &= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} \\ &= \frac{7}{8}.\end{aligned}$$

In a uniform sample space (where all outcomes are equally likely) the probability of an event A can be calculated as:

$$\Pr(A) = \frac{\text{number of outcomes in } A}{\text{number of outcomes}} = \frac{|A|}{|S|}.$$

Consider a probability space with sample space $\{1, 2, 3, 4, 5\}$ and probability function given by

| | | | | | |
|----------|----------------|---------------|---------------|---------------|----------------|
| s | 1 | 2 | 3 | 4 | 5 |
| $\Pr(s)$ | $\frac{1}{10}$ | $\frac{1}{5}$ | $\frac{2}{5}$ | $\frac{1}{5}$ | $\frac{1}{10}$ |

Let E be the event that the outcome is even and P be the event that the outcome is prime.

Question What sets do E and P correspond to?

Answer $E = \{2, 4\}$ $P = \{2, 3, 5\}$.

Question What are $\Pr(E)$ and $\Pr(P)$?

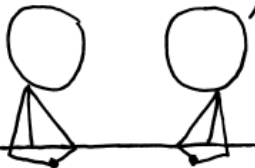
Answer

$$\Pr(E) = \Pr(2) + \Pr(4) = \frac{1}{5} + \frac{1}{5} = \frac{2}{5}$$

$$\Pr(P) = \Pr(2) + \Pr(3) + \Pr(5) = \frac{1}{5} + \frac{2}{5} + \frac{1}{10} = \frac{7}{10}.$$

A WEIGHTED RANDOM NUMBER
GENERATOR JUST PRODUCED
A NEW BATCH OF NUMBERS.

LET'S USE THEM TO
BUILD NARRATIVES!



ALL SPORTS COMMENTARY

21.3 Operations on events

Because events are defined as sets we can perform set operations on them. If A and B are events for a sample space S , then

- $A \cup B$ is the event “ A or B ,”
- $A \cap B$ is the event “ A and B ,”
- \overline{A} is the event “not A .”

We always take the sample space as our universal set, so \overline{A} means $S - A$.

Remember the probability space with sample space $\{1, 2, 3, 4, 5\}$ and probability function given by

| | | | | | |
|----------|----------------|---------------|---------------|---------------|----------------|
| s | 1 | 2 | 3 | 4 | 5 |
| $\Pr(s)$ | $\frac{1}{10}$ | $\frac{1}{5}$ | $\frac{2}{5}$ | $\frac{1}{5}$ | $\frac{1}{10}$ |

Let E be the event that the outcome is even and P be the event that the outcome is prime. So $E = \{2, 4\}$ and $P = \{2, 3, 5\}$.

Question What are the events $E \cap P$, \overline{P} and $\overline{E} \cup P$?

Answer $E \cap P = \{2\}$ $\overline{P} = \{1, 4\}$ $\overline{E} \cup P = \{1, 2, 3, 5\}$

Flux Exercise

If A and B are any two events then $\Pr(A \cap B)$

A. $= \Pr(A) + \Pr(B)$

B. $= \Pr(A) \times \Pr(B)$

C. $= \Pr(A) - \Pr(B)$

D. $= \Pr(A) \div \Pr(B)$

E. $\leq \min(\Pr(A), \Pr(B))$

Answer E. In general $\Pr(A \cap B)$ depends on the events A and B , but the only thing you can be sure of is that it is no more than the probability of either of the individual events.

21.4 Probabilities of unions

We saw in the section on the inclusion-exclusion principle that $|A \cup B| = |A| + |B| - |A \cap B|$ for finite sets A and B . We have a similar law in probability.

For any two events A and B ,

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B).$$

Example. In our die rolling example, let $A = \{1, 2\}$ and $B = \{2, 3, 4\}$ be events. Then

$$\Pr(A \cup B) = \frac{|A|}{|S|} + \frac{|B|}{|S|} - \frac{|A \cap B|}{|S|} = \frac{2}{6} + \frac{3}{6} - \frac{1}{6} = \frac{2}{3}.$$

Two events A and B are *mutually exclusive* if $\Pr(A \cap B) = 0$, that is, if A and B cannot occur together. For mutually exclusive events, we have

$$\Pr(A \cup B) = \Pr(A) + \Pr(B).$$

Mutually exclusive events are ones that **can't occur together**.

For example, if we roll a standard die then:

- ▶ “the number rolled is at least 5” and “the number rolled is 2” are mutually exclusive.
- ▶ “the number rolled is even” and “the number rolled is 6” are **not** mutually exclusive.

For **any** two events A and B we have

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B).$$

If A and B are mutually exclusive events then $\Pr(A \cap B) = 0$.

So $\Pr(A \cup B) = \Pr(A) + \Pr(B)$.

21.5 Independent events

We say that two events are *independent* when the occurrence or non-occurrence of one event does not affect the likelihood of the other occurring.

Two events A and B are *independent* if

$$\Pr(A \cap B) = \Pr(A)\Pr(B).$$

Independent events are ones that **don't affect each other's likelihood**.

For example, if we roll a standard die then:

- ▶ “the number rolled is at least 5” and “the number rolled is even” are independent

Intuitively: An even number has base probability $\frac{1}{2}$ and still has probability $\frac{1}{2}$ if we know the number rolled was at least 5.

Similarly ≥ 5 has base probability $\frac{1}{3}$ and still has probability $\frac{1}{3}$ if we know the number rolled was even.

Formally: $\Pr(\geq 5) \Pr(\text{even}) = \frac{1}{3} \times \frac{1}{2} = \frac{1}{6}$. Also $\Pr(\geq 5 \cap \text{even}) = \Pr(6) = \frac{1}{6}$.

- ▶ “the number rolled is at least 4” and “the number rolled is even” are **not** independent.

Intuitively: An even number has base probability $\frac{1}{2}$ and but has probability $\frac{2}{3}$ if we know the number rolled was at least 4.

Similarly ≥ 4 has base probability $\frac{1}{2}$ but has probability $\frac{2}{3}$ if we know the number rolled was even.

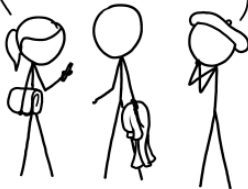
Formally: $\Pr(\geq 4) \Pr(\text{even}) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$. But $\Pr(\geq 4 \cap \text{even}) = \Pr(4, 6) = \frac{1}{3}$.

Don't mix up independent with mutually exclusive!

WE SHOULD GO TO THE NORTH BEACH.
SOMEONE SAID THE SOUTH BEACH HAS
A 20% HIGHER RISK OF SHARK ATTACKS.

YEAH, BUT STATISTICALLY, TAKING
THREE BEACH TRIPS INSTEAD OF TWO
INCREASES OUR ODDS OF GETTING
SHOT BY A SWIMMING DOG CARRYING
A HANDGUN IN ITS MOUTH BY **50%!**

OH NO! THIS IS
OUR THIRD TRIP!



REMINDER: A 50% INCREASE
IN A TINY RISK IS **STILL TINY.**

Example. A binary string of length 3 is generated uniformly at random. The event A that the first bit is a 1 is independent of the event B that the second bit is a 1. But A is *not* independent of the event C that the string contains exactly two 1s.

Formally, the sample space is $S = \{000, 001, 010, 011, 100, 101, 110, 111\}$ and $\Pr(s) = \frac{1}{8}$ for any $s \in S$. So,

$$A = \{100, 101, 110, 111\} \quad \Pr(A) = \frac{1}{2}$$

$$B = \{010, 011, 110, 111\} \quad \Pr(B) = \frac{1}{2}$$

$$C = \{011, 101, 110\} \quad \Pr(C) = \frac{3}{8}$$

$$A \cap B = \{110, 111\} \quad \Pr(A \cap B) = \frac{1}{4}$$

$$A \cap C = \{101, 110\} \quad \Pr(A \cap C) = \frac{1}{4}$$

So $\Pr(A \cap B) = \Pr(A)\Pr(B)$ but $\Pr(A \cap C) \neq \Pr(A)\Pr(C)$.

Flux Exercise

An Australian adult is chosen uniformly at random. Consider the events that the chosen person is (a) A professional basketballer, (b) A jockey, (c) A lawyer, (d) Over 180cm tall, (e) 75 years old. Which pair of these events is closest to being independent?

A: (a) and (d)

B: (a) and (e)

C: (b) and (d)

D: (c) and (d)

E: (d) and (e)

Answer D. If someone is over 75 they are less likely to be employed, and they are also a bit less likely to be tall (why?). If they are tall it makes them more likely to be a basketballer and less likely to be a jockey. On the other hand being tall doesn't affect your likelihood to be a lawyer (much)...

21.6 Warning

Random processes can occur in both discrete and continuous settings, and probability theory can be applied in either setting. In this lecture, and in the next four lectures, we are discussing only the discrete case. Many of the definitions and results we state apply only in this case. Our definition of a probability space, for example, is actually the definition of a discrete probability space, and so on.

The discrete setting provides a good environment to learn most of the vital concepts and intuitions of probability theory. What you learn here is very useful in itself, and will act as a good base if you go on to study continuous probability.