MATH2750 Introduction to Markov Processes

Matthew Aldridge

University of Leeds, 2020–2021

Contents

Schedule

Week 1 (25–29 January)

- About the module
- **Section 1**: Stochastic processes and the Markov property
- **Section 2**: Random walks
- **Problem Sheet 1**
- **Lecture**: Tuesday at 1400 (link in Minerva)
- **Drop-in sessions**: Tuesday or Wednesday (Teams)

About MATH2750

This module is **MATH2750 Introduction to Markov Processes**. The module manager and lecturer is Dr Matthew Aldridge, and my email address is m.aldridge@leeds.ac.uk.

Organisation of MATH2750

This module lasts for 11 weeks. The first nine weeks run from 25 January to 26 March, then we break for Easter, and then the final two weeks run from 26 April to 7 May.

Notes and videos

The main way I expect you to learn the material for this course is by reading these notes and by watching the accompanying videos. I will set two sections of notes each week, for a total of 22 sections.

Reading mathematics is a slow process. Each section roughly corresponds to one lecture last year, which would have been 50 minutes. If you find yourself regularly getting through sections in much less than an hour, you're probably not reading carefully enough through each sentence of explanation and each line of mathematics, including understanding the motivation as well as checking the accuracy.

It is possible (but not recommended) to learn the material by only reading the notes and not watching the videos. It is not possible to learn the material by only watching the videos and not reading the notes.

You are probably reading the web version of the notes. If you want a PDF copy (to read offline or to print out), then click the PDF button in the top ribbon of the page. (Warning: I have not made as much effort to make the PDF neat and tidy as I have the web version.)

Since we will all be relying heavily on these notes, I'm even more keen than usual to hear about errors mathematical, typographical or otherwise. Please, please email me if think you may have found any.

Problem sheets

There will be 10 problem sheets; Problem Sheet n covers the material from the two sections from week n (Sections $2n-1$ and $2n$), and will be discussed in your workshop in week $n+1$.

Lectures

There will be one online synchronous "lecture" session each week, on Tuesdays at 1400, with me, run through Zoom.

This will not be a "lecture" in the traditional sense of the term, but will be an opportunity to re-emphasise material you have already learned from notes and videos, to give extra examples, and to answer common student questions, with some degree of interactivity.

I will assume you have completed all the work for the previous week by the time of the lecture, but I will not assume you've started the work for that week itself.

I am very keen to hear about things you'd like to go through in the lectures; please email me with your suggestions.

Workshops

There will be 10 workshops, starting in the second week. The main goal of the workshops will be to go over your answers to the problems sheets in smaller classes. You will have been assigned to one of three workshop groups, meeting on Mondays or Tuesdays, led by Jason Klebes, Dr Jochen Voss, or me. Your workshop will be run through Zoom or Microsoft Teams; your workshop leader will contact you before the end of this week with arrangements.

My recommended approach to problem sheets and workshops is the following:

- Work through the problem sheet before the workshop, spending plenty of time on it, and making multiple efforts at questions you get stuck on. I recommend spending *at least three hours* on each problem sheet, in more than one block. Collaboration is encouraged when working through the problems, but I recommend writing up your work on your own.
- Take advantage of the smaller group setting of the workshop to ask for help or clarification on questions you weren't able to complete.
- After the workshop, attempt again the questions you were previously stuck on.
- If you're still unable to complete a question after this second round of attempts, *then* consult the solutions.

Assessments

There will be four pieces of assessed coursework, making up a total of 15% of your mark for the module. Assessments 1, 2 and 4 will involve writing up answers to a few problems, in a similar style to the problem sheets, and are worth 4% each. (In response to previous student feedback, there are fewer questions per assessment.) Assessment 3 will be a report on some computational work (see below) and is worth 3%.

While you may want to discuss an assessment with others in advance of completing it by yourself, copying is not allowed and will be dealt with in accordance with University rules.

The assessments deadlines are:

- Assessment 1: Thursday 11 February 1400 (week 3)
- Assessment 2: Thursday 18 March 1400 (week 8)
- Assessment 3 (Computing Worksheet 2): Thursday 25 March 1400 (week 9)
- Assessment 4: Thursday 6 May 1400 (week 11)

Work will be submitted via Gradescope.

Computing worksheets

There will be two computing worksheets, which will look at the material in the course through simulations in R. This material is examinable. You should be able to work through the worksheets in your own time, but if you need help, there will be optional online drop-in sessions in the weeks 4 and 7 with Muyang Zhang through Microsoft Teams. (Your computing drop-in session may be listed as "Practical" on your timetable.)

The first computing worksheet will be a practice run, while a report on the second computing worksheet will be the third assessed piece of work.

Drop-in sessions

If you there is something in the course you wish to discuss in detail, the place for the is the optional weekly drop-in session. The drop-in sessions are an optional opportunity for you to ask questions you have to a member of staff – nothing will happen unless you being your questions.

You will have been assigned to one of three groups on Tuesdays or Wednesdays with Nikita Merkulov or me. The drop-in sessions will be run the Microsoft Teams. Your drop-in session would be an excellent place to go if you are having trouble understanding something in the written notes, or if you're still struggling on a problem sheet question after your workshop.

Microsoft Team

I have set up a Microsoft Team for the course. I propose to use the "Q and A" channel there as a discussion board. This is a good place to post questions about material from the course, and – even better! – to help answer you colleagues' questions. The idea is that you all as a group should help each other out. I will visit a couple of times a week to clarify if everybody is stumped by a question, or if there is disagreement.

Time management

It is, of course, up to you how you choose to spend your time on this module. But, if you're interested, my recommendations would be something like this:

- **Every week:** 7.5 hours per week
    - **Notes and videos:** 2 sections, 1 hour each
    - **Problem sheet:** 3.5 hours per week
    - **Lecture:** 1 hour per week
    - **Workshop:** 1 hour per week
- **When required:**
    - **Assessments 1, 2 and 4:** 2 hours each
    - **Computer worksheets:** 2 hours each
    - **Revision**: 12 hours
- **Total:** 100 hours

Exam

There will be an exam – or, rather, a final "online time-limited assessment" – after the end of the module, making up the remaining 85% of your mark. The exam will consist of four questions, and you are expected to answer all of them. You will have 48 hours to complete the exam, although the exam itself should represent half a day to a day's work. Further details to follow nearer the time.

Who should I ask about...?

- *I don't understand something in the notes or on a problem sheet*: Go to your weekly drop-in session, or post a question on the Teams Q and A board. (If you email me, I am likely to respond, "That would be an excellent question for your drop-in session or the Q and A board.")
- *I don't understand something in on a computational worksheet:* Go to your computing drop-in session in weeks 4 or 7.
- *I have an admin question about general arrangements for the module:* Email me.
- *I have an admin question about arrangements for my workshop:* Email your workshop leader.
- *I have suggestion for something to cover in the lectures:* Email me.
- *I need an extension on or exemption from an assessment:* Email the Maths Taught Students Office.

Content of MATH2750

Prerequisites

Some students have asked what background you'll be expected to know for this course.

It's essential that you're very comfortable with the basics of probability theory: events, probability, discrete and continuous random variables, expectation, variance, approximations with the normal distribution, etc. Conditional probability and independence are particularly important concepts in this course. This course will use the binomial, geometric, Poisson, normal and exponential distributions, although the notes will usually remind you about them first, in case you've forgotten.

Many students on the module will have studied these topics in MATH1710 Probability and Statistics 1; others will have covered these in different modules.

Syllabus

The course has two major parts: the first part will cover processes in discrete time and the second part processes in continuous time.

An outline plan of the topics covered is the following. (Remember that one week's work is two sections of notes.)

- **Discrete time Markov chains** [12 sections]
  - Introduction to stochastic processes [1 section]
  - Important examples: Random walk, gambler's ruin, linear difference equations, examples from actuarial science [4 sections]
  - General theory: transition probabilities, n-step transition probabilities, class structure, periodicity, hitting times, recurrence and transience, stationary distributions, long-term behaviour [6 sections]
  - Revision [1 section]
- **Continuous time Markov jump processes** [10 sections]
  - Important examples: Poisson process, counting processes, queues [5 sections]
  - General theory: holding times and jump chains, forward and backward equations, class structure, hitting times, stationary distributions, long-term behaviour [4 sections]
  - Revision [1 section]

Books

You can do well on this module by reading the notes and watching the videos, attending the lectures and workshops, and working on the problem sheets, assignments and practicals, without any further reading. However, students can benefit from optional extra background reading or an alternative view on the material.

My favourite book on Markov chains, which I used a lot while planning this course and writing these notes, is:

- J.R. Norris, *Markov Chains*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 1997. Chapters 1-3.

This a whole book just on Markov processes, including some more detailed material that goes beyond this module. Its coverage of of both discrete and continuous time Markov processes is very thorough. Chapter 1 on discrete time Markov chains is available online.

Other good books with sections on Markov processes that I have used include:

- G.R. Grimmett and D.R. Stirzaker, *Probability and Random Processes*, 4th edition, Oxford University Press, 2020. Chapter 6.
- G. Grimmet and D. Walsh, *Probability: an introduction*, 2nd edition, Oxford University Press, 2014. Chapter 12.
- D.R. Stirzaker, *Elementary Probability*, 2nd edition, Cambridge University Press, 2003. Chapter 9.

Grimmett and Stirzaker is an excellent handbook that covers most of undergraduate probability – I bought a copy when I was a second-year undergraduate and still keep it next to my desk. The whole of Stirzaker is available online.

A gentler introduction with plenty of examples is provided by:

- P.W. Jones and P. Smith, *Stochastic Processes: an introduction*, 3nd edition, Texts in Statistical Science, CRC Press, 2018. Chapters 2-7.

although it doesn't cover everything in this module. The whole book is available online.

(I've listed the newest editions of these books, but older editions will usually be fine too.)

And finally...

These notes were mostly written by Matthew Aldridge in 2018–19, and have received updates (mostly in Sections 9–11) and reformatting this year. Some of the material (especially Section 1, Section 6, and numerous diagrams) follows closely previous notes by Dr Graham Murphy, and I also benefited from reading earlier notes by Dr Robert Aykroyd and Prof Alexander Veretennikov. Dr Murphy's general help and advice was also very valuable. Many thanks to students in previous runnings of the module for spotting errors and suggesting improvements.

Part I: Discrete time Markov chains

1 Stochastic processes and the Markov property

- Stochastic processes with discrete or continuous state space and discrete or continuous time
- The Markov "memoryless" property

1.1 Deterministic and random models

A **model** is an imitation of a real-world system. For example, you might want to have a model to imitate the world's population, the level of water in a reservoir, cashflows of a pension scheme, or the price of a stock. Models allow us to try to understand and predict what might happen in the real world in a low risk, cost effective and fast way.

To design a model requires a set of assumptions about how it will work and suitable parameters need to be determined, perhaps based on past collected data.

An important distinction is between **deterministic** models and **random** models. Another word for a random model is a **stochastic** ("*sto-KASS-tik*") model. Deterministic models do not contain any random components, so the output is completely determined by the inputs and any parameters. Random models have variable outcomes to account for uncertainty and unpredictability, so they can be run many times to give a sense of the range of possible outcomes.

Consider models for:

- the future position of the Moon as it orbits the Earth,
- the future price of shares in Apple.

For the moon, the random components – for example, the effect of small meteorites striking the Moon's surface – are not very significant and a deterministic model based on physical laws is good enough for most purposes. For Apple shares, the price changes from day to day are highly uncertain, so a random model can account for the variability and unpredictability in a useful way.

In this module we will see many examples of stochastic models. Lots of the applications we will consider come from financial mathematics and actuarial science where the use of models that take into account uncertainty is very important, but the principles apply in many areas.

1.2 Stochastic processes

If we want to model, for example, the total number of claims to an insurance company in the whole of 2020, we can use a random variable X to model this – perhaps a Poisson distribution with an appropriate mean. However, if we want to track how the number of claims changes over the course of the year 2021, we will need to use a **stochastic process** (or "random process").

A stochastic process, which we will usually write as (X_n), is an indexed sequence of random variables that are (usually) dependent on each other.

Each random variable X_n takes a value in a **state space** $\mathcal{S}$ which is the set of possible values for the process. As with usual random variables, the state space $\mathcal{S}$ can be **discrete** or **continuous**. A discrete state space denotes a set of distinct possible outcomes, which can be finite or countably infinite. For example, $\mathcal{S} = \{\text{Heads}, \text{Tails}\}$ is the state space for a single coin flip, while in the case of counting insurance claims, the state space would be the nonnegative integers $\mathcal{S} = \mathbb{Z}_+ = \{0, 1, 2, \ldots\}$. A continuous state space denotes an uncountably infinite continuum of gradually varying outcomes. For example, the nonnegative real line $\mathcal{S} = \mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$ is the state space for the amount of rainfall on a given day, while some bounded subset of $\mathbb{R}^3$ would be the state space for the position of a gas particle in a box.

Further, the process has an **index set** that puts the random variables that make up the process in order. The index set is usually interpreted as a **time** variable, telling us when the process will be measured. The index set for time can also be discrete or continuous. Discrete time denotes a process sampled at distinct points, often denoted by $n = 0, 1, 2, \ldots$, while continuous time denotes a process monitored constantly over time, often denoted by $t \in \mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$. In the insurance example, we might count up the number of claims each day – then the discrete index set will be the days of the year, which we could denote $\{1, 2, \ldots, 365\}$. Alternatively, we might want to keep a constant tally that we update after every claim, requiring a continuous time index t representing time across the whole year. In discrete time, we can write down the first few steps of the process as $(X_0, X_1, X_2, \ldots)$.

This gives us four possibilities in total:

- **Discrete time, discrete space**
  - Example: Number of students attending each lecture of maths module.
  - **Markov chains** – discrete time, discrete space stochastic processes with a certain "Markov property" – are the main topic of the first half of this module.

- **Discrete time, continuous space**
  - Example: Daily maximum temperature in Leeds.
  - We will briefly mention continuous space Markov chains in the first half of the course, but these are not as important.

- **Continuous time, discrete space**
  - Example: Number of visitors to a webpage over time.
  - **Markov jump processes** – continuous time, discrete space stochastic processes with the "Markov property" – are the main topic of the second half of this module.

- **Continuous time, continuous space**
  - Example: Level of the FTSE 100 share index over time.
  - Such processes, especially the famous Brownian motion – another process with the Markov property – are very important, but outside the scope of this course. See MATH3734 Stochastic Calculus for Finance next year, for example.

1.3   Markov property

Because stochastic processes consist of a large number – even infinitely many; even uncountably infinitely many – random variables that could all be dependent on each other, they can get extremely complicated. The Markov property is a crucial property that restricts the type of dependencies in a process, to make the process easier to study, yet still leaves most of the useful and interesting examples intact. (Although particular examples of Markov processes go back further, the first general study was by the Russian mathematician Andrey Andreyevich Markov, published in 1906.)

Think of a simple board game where we roll a dice and move that many squares forward on the board. Suppose we are currently on the square X_n. Then what can we say about which square X_{n+1} we move to on our next turn?

- X_{n+1} is random, since it depends on the roll of the dice.
- X_{n+1} depends on where we are now X_n, since the score of dice will be added onto the number our current square,

- Given the square X_n we are now, X_{n+1} doesn't depend any further on which sequence of squares $X_0, X_1, \ldots, X_{n-1}$ we used to get here.

It is this third point that is the crucial property of the stochastic processes we will study in this course, and it is called the **Markov property** or **memoryless property**. We say "memoryless", because it's as if the process forgot how it got here – we only need to remember what square we've reached, not which squares we used to get here. The stochastic process before this moment has no bearing on the future, given where we are now. A mathematical way to say this is that "the past and the future are conditionally independent given the present."

To write this down formally, we need to recall **conditional probability**: the conditional probability of an event A given another event B is written $\mathbb{P}(A \mid B)$, and is the probability that A occurs *given* that B definitely occurs. You may remember the definition

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)},$$

although is often more useful to reason directly about conditional probabilities than use this formula.

(You may also remember that the definition of conditional probability requires that $\mathbb{P}(B) > 0$. Whenever we write down a conditional probability, we implicitly assume the conditioning event has strictly positive probability without explicitly saying so.)

Definition 1.1. Let $(X_n) = (X_0, X_1, X_2, \ldots)$ be a stochastic process in discrete time $n = 0, 1, 2, \ldots$ and discrete space $\mathcal{S}$. Then we say that (X_n) has the **Markov property** if, for all times n and all states $x_0, x_1, \ldots, x_n, x_{n+1} \in \mathcal{S}$ we have

$$\mathbb{P}(X_{n+1} = x_{n+1} \mid X_n = x_n, X_{n-1} = x_{n-1}, \ldots, X_0 = x_0) = \mathbb{P}(X_{n+1} = x_{n+1} \mid X_n = x_n).$$

Here, the left hand side is the probability we go to state x_{n+1} next conditioned on the entire history of the process, while the right hand side is the probability we go to state x_{n+1} next conditioned only on where we are now x_n. So the Markov property tells us that it only matters where we are now and not how we got here.

(There's also a similar definition for continuous time processes, which we'll come to later in the course.)

Stochastic processes that have the Markov property are much easier to study than general processes, as we only have to keep track of where we are now and we can forget about the entire history that came before.

In the next section, we'll see the first, and most important, example of a discrete time discrete space Markov chain: the "random walk".

2 Random walk

- Definition of the simple random walk and the exact binomial distribution
- Expectation and variance of general random walks

2.1 Simple random walk

Consider the following **simple random walk** on the integers $\mathbb{Z}$: We start at 0, then at each time step, we go up by one with probability p and down by one with probability $q = 1 - p$. When $p = q = \frac{1}{2}$, we're equally as likely to go up as down, and we call this the **simple symmetric random walk**.

The simple random walk is a simple but very useful model for lots of processes, like stock prices, sizes of populations, or positions of gas particles. (In many modern models, however, these have been replaced by more complicated continuous time and space models.) The simple random walk is sometimes called the "drunkard's walk", suggesting it could model a drunk person trying to stagger home.
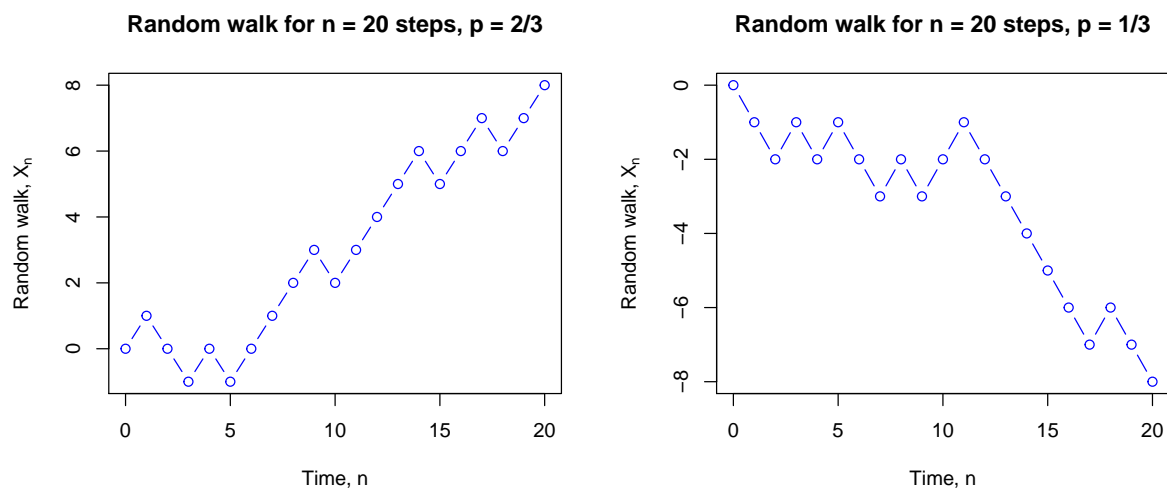

Figure 1: Two simulations of random walks.

We can write this as a stochastic process (X_n) with discrete time $n = \{0, 1, 2, \dots\} = \mathbb{Z}_+$ and discrete state space $\mathcal{S} = \mathbb{Z}$, where $X_0 = 0$ and, for $n \geq 0$, we have

$$X_{n+1} = \begin{cases} X_n + 1 & \text{with probability } p, \\ X_n - 1 & \text{with probability } q. \end{cases}$$

It's clear from this definition that X_{n+1} (the future) depends on X_n (the present), but, given X_n, does not depend on $X_{n-1}, \dots, X_1, X_0$ (the past). Thus the Markov property holds, and the simple random walk is a **discrete time Markov process** or **Markov chain**.

Example 2.1. *What is the probability that after two steps a simple random walk has reached $X_2 = 2$?*

To achieve this, the walk must go upwards in both time steps, so $\mathbb{P}(X_2 = 2) = pp = p^2$.

Example 2.2. *What is the probability that after three steps a simple random walk has reached $X_3 = -1$?*

There are three ways to reach -1 after three steps: up–down–down, down–up–down, or down–down–up. So

$$\mathbb{P}(X_3 = -1) = pqq + qpq + qqp = 3pq^2.$$

2.2 General random walks

An alternative way to write the simple random walk is to put

$$X_n = X_0 + \sum_{i=1}^{n} Z_i, \tag{1}$$

where the starting point is $X_0 = 0$ and the **increments** $Z_1, Z_2, \ldots$ are independent and identically distributed (IID) random variables with distribution given by $\mathbb{P}(Z_i = 1) = p$ and $\mathbb{P}(Z_i = -1) = q$. You can check that (1) means that $X_{n+1} = X_n + Z_{n+1}$, and that this property defines the simple random walk.

Any stochastic process with the form (1) for some X_0 and some distribution for the IID Z_is is called a **random walk** (without the word "simple").

Random walks often have state space $\mathcal{S} = \mathbb{Z}$, like the simple random walk, but they could be defined on other state spaces. We could look at higher dimensional simple random walks: in $\mathbb{Z}^2$, for example, we could step up, down, left or right each with probability $\frac{1}{4}$. We could even have a continuous state space like $\mathbb{R}$, if, for example, the Z_is had a normal distribution.

We can use this structure to calculate the expectation or variance of any random walk (including the simple random walk).

Let's start with the expectation. For a random walk (X_n) we have

$$\mathbb{E}X_n = \mathbb{E}\left(X_0 + \sum_{i=1}^{n} Z_i\right) = \mathbb{E}X_0 + \sum_{i=1}^{n} \mathbb{E}Z_i = \mathbb{E}X_0 + n\mathbb{E}Z_1,$$

where we've used the linearity of expectation, and that the Z_is are identically distributed.

In the case of the simple random walk, we have $\mathbb{E}X_0 = 0$, since we start from 0 with certainty, and

$$\mathbb{E}Z_1 = \sum_{z \in \mathbb{Z}} z\mathbb{P}(Z_1 = z) = 1 \times p + (-1) \times q = p - q.$$

Hence, for the simple random walk, $\mathbb{E}X_n = n(p - q)$.

If $p > \frac{1}{2}$, then $p > q$, so $\mathbb{E}X_n$ grows ever bigger over time, while if $p < \frac{1}{2}$, then $\mathbb{E}X_n$ grows ever smaller (that is, negative with larger absolute value) over time. If $p = \frac{1}{2} = q$, which is the case of the simple symmetric random walk, then then the expectation $\mathbb{E}X_n = 0$ is zero for all time.

Now the variance of a random walk. We have

$$\mathrm{Var}(X_n) = \mathrm{Var}\left(X_0 + \sum_{i=1}^{n} Z_i\right) = \mathrm{Var}\,X_0 + \sum_{i=1}^{n} \mathrm{Var}\,Z_i = \mathrm{Var}\,X_0 + n\,\mathrm{Var}\,Z_1,$$

where it was crucial that X_0 and all the Z_is were independent (so we had no covariance terms).

For the simple random walk we have $\mathrm{Var}\,X_0 = 0$, since we always start from 0 with certainty. To calculate the variance of the increments, we write

$$\begin{aligned}
\mathrm{Var}(Z_1) &= \mathbb{E}(Z_1 - \mathbb{E}Z_1)^2 \\
&= p\big(1 - (p - q)\big)^2 + q\big(-1 - (p - q)\big)^2 \\
&= p(2q)^2 + q(-2p)^2 \\
&= 4pq^2 + 4p^2q \\
&= 4pq(p + q) \\
&= 4pq.
\end{aligned}$$

Here we've used that $1 - p = q$, $1 - q = p$, and $p + q = 1$; you should take a few moments to check you've followed the algebra here. Hence the variance of the simple random walk is $4pqn$. We see that (unless p is 0 or 1) the variance grows over time, so it becomes harder and harder to predict where the random walk will be.

The variance of the simple symmetric random walk is $4\frac{1}{2}\frac{1}{2}n = n$.

For large n, we can use a normal approximation for a random walk. Suppose the increments process (Z_n) has mean μ and variance σ^2, and that the walk starts from $X_0 = 0$. Then we have $\mathbb{E}X_n = \mu n$ and $\text{Var}(X_n) = \sigma^2 n$, so for large n we can use the normal approximation $X_n \approx \text{N}(\mu n, \sigma^2 n)$. (Remember, of course, that the X_n are not independent.) To be more formal, the central limit theorem tells us that, as $n \to \infty$, we have

$$\frac{X_n - n\mu}{\sigma\sqrt{n}} \to \text{N}(0, 1).$$

2.3 Exact distribution of the simple random walk

We have calculated the expectation and variance of any random walk. But for the simple random walk, we can in fact give the exact distribution, by writing down an exact formula for $\mathbb{P}(X_n = i)$ for any time n and any state i.

Recall that, at each of the first n times, we independently take an upward step with probability p, and otherwise take a downward step. So if we let Y_n be the number of upward steps over the first n time periods, we see that Y_n has a binomial distribution $Y_n \sim \text{Bin}(n, p)$.

Recall that the binomial distribution has probability

$$\mathbb{P}(Y_n = k) = \binom{n}{k} p^k (1 - p)^{n-k} = \binom{n}{k} p^k q^{n-k},$$

for $k = 0, 1, \ldots, n$, where $\binom{n}{k}$ is the binomial coefficient "n choose k".

If $Y_n = k$, that means we've taken k upward steps and $n - k$ downward steps, leaving us at position $k - (n - k) = 2k - n$. Thus we have that

$$\mathbb{P}(X_n = 2k - n) = \mathbb{P}(Y_n = k) = \binom{n}{k} p^k q^{n-k}. \tag{2}$$

Note that after an odd number of time steps n we're always at an odd-numbered state, since $2k - \text{odd} = \text{odd}$, while after an even number of time steps n we're always at an even-numbered state, since $2k - \text{even} = \text{even}$.

Writing $i = 2k - n$ gives $k = (n + i)/2$ and $n - k = (n - i)/2$. So we can rearrange (2) to see that the distribution for the simple random walk is

$$\mathbb{P}(X_n = i) = \binom{n}{(n+i)/2} p^{(n+i)/2} q^{(n-i)/2},$$

when n and i have the same parity with $-n \leq i \leq n$, and is 0 otherwise.

In the special case of the simple symmetric random walk, we have

$$\mathbb{P}(X_n = i) = \binom{n}{(n+i)/2} \left(\frac{1}{2}\right)^{(n+i)/2} \left(\frac{1}{2}\right)^{(n-i)/2} = \binom{n}{(n+i)/2} 2^{-n}.$$

In the next section, we look at a gambling problem based on the simple random walk.

Problem Sheet 1

You should attempt all these questions and write up your solutions in advance of your workshop in week 2 (Monday 1 or Tuesday 2 February) where the answers will be discussed.

1. When designing a model for a quantity that changes over time, one has many decisions to make:

- Discrete or continuous state space?
- Discrete or continuous index set for time?
- Deterministic or stochastic model?
- If a stochastic model is chosen, is it reasonable to assume that the Markov property holds?

What would you decide for the following scenarios:

(a) The percentage of UK voters with a positive opinion of Boris Johnson in a weekly tracking poll.

(b) The number of points won by a football league club throughout the season.

(c) The temperature of a bowl of water placed in an oven.

(d) The number of people inside the University of Leeds library.

2. A fair six-sided dice is rolled twice, resulting in the values $X_1, X_2 \in \{1, 2, \ldots, 6\}$. Let $Y = X_1 + X_2$ be the total score. Calculate:

(a) the probability $\mathbb{P}(Y = 10)$;

(b) the conditional probability $\mathbb{P}(Y = 10 \mid X_1 = x)$ for $x = 1, 2, \ldots, 6$;

(c) the conditional probability $\mathbb{P}(X_1 = x \mid Y = 10)$ for $x = 1, 2, \ldots, 6$.

3. Let (X_n) be a simple random walk starting from $X_0 = 0$ and that at each step goes up one with probability p or down one with probability $q = 1 - p$. What are:

(a) $\mathbb{P}(X_5 = 3)$,

(b) $\mathbb{P}(X_5 = 3 \mid X_2 = 2)$,

(c) $\mathbb{P}(X_n = n - 2)$,

(d) $\mathbb{E}X_4$,

(e) $\mathbb{E}(X_6 \mid X_4 = 2)$,

4. The price X_n of a stock at the close of day n is modelled as a Gaussian random walk, where the increments (Z_n) have a normal distribution $Z_n \sim \mathrm{N}(\mu, \sigma^2)$. The model assumes a drift of $\mu = 0.7$ and a volatility of $\sigma = 2.2$. The initial price is $X_0 = 42.3$.

(a) Calculate the mean and variance of the price of the stock at the close of day 5.

(b) Give a 95% prediction interval for the price at the close of day 5. (You might find it useful to recall that, if $W \sim \mathrm{N}(0, 1)$ is a standard normal random variable, then $\mathbb{P}(W \leq 1.96) = 0.975$.)

(c) After day 4, the prices at the end of each of the first four days have been recorded as $X_1 = 44.4, X_2 = 44.0, X_3 = 47.1, X_4 = 47.8$. Update your prediction interval for the price at the close of day 5, and comment on how it differs from the earlier prediction interval.

5. A gambler decides to model her total winnings as a simple random walk starting from $X_0 = 0$ that at each time goes up one with probability p and down one with probability $1 - p$, but where p is unknown. The first 10 recordings, X_1 to X_{10}, are

$$(1, 2, 1, 2, 3, 4, 5, 6, 5, 6).$$

(a) What would you guess for the value of p, given this data?

(b) More generally, how would you estimate p from the data $X_0 = 0, X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$?

(c) Show that your estimate is in fact the maximum likelihood estimate of p.