Stochastic Modelling and Random Processes

Example sheet 1

This is an example sheet, with some problems which are meant to (a) improve your understanding of some properties of DTMCs which we looked at over the last couple of weeks and (b) get you used to simulating DTMCs. You can use any programming language you are familiar with, but Kamran will work with python.
So that you practice for the assignments, all plots must contain axis labels and a legend. Use your own judgement to find reasonable and relevant plot ranges.

1 Let $A, B, C_1, \ldots, C_n \subseteq \Omega$ be events in a probability space Ω such that $C_i \cap C_j = \emptyset$ for all $i \neq j$ and $\bigcup_{i=1}^{n} C_i = \Omega$. Using the law of total probability, show that

$$\mathbb{P}(A|B) = \sum_{i=1}^{n} \mathbb{P}(A|B \cap C_i)\, \mathbb{P}(C_i|B)\,. \qquad [2]$$

2 A dice is rolled repeatedly. Which of the following are Markov chains? For those that are, supply the state space and the transition matrix. Are they homogeneous?

(a) The largest number X_n shown up to the nth roll.

(b) The number N_n of sixes in n rolls.

(c) At time n, the time B_n since the most recent six.

3 Consider the random walk with states **s**, **c** and **r** which we used as an example in week 2, where we had:

$$\begin{array}{lll} p(s,s) = p_s, & p(c,s) = q_s, & p(r,s) = r_s \\ p(s,c) = r_c, & p(c,c) = p_c, & p(r,c) = q_c \\ p(s,r) = q_r, & p(c,r) = r_r, & p(r,r) = p_r \end{array}$$

(a) Write down the transition matrix P.

(b) Choose your favourite values for the transition probabilities $p(x,y)$ for $x, y \in \{s, c, r\}$. (make sure the resulting matrix P is stochastic). Does this Markov chain have a stationary distribution (and if yes, find it)? Is it irreducible?

(c) Simulate 500 realisations of this Markov chain with $X_0 = s$ for all of them. Plot the empirical distribution (histograms) after some days $T = 10$, 100, 1000 and compare this with your expected stationary distribution. *hint: you might want to rename* **s**, **c** *and* **r** *to* 1, 2, *and* 3.

4 Wright-Fisher model of population genetics

Consider a fixed population of L individuals. At time $t = 0$ each individuum i has a different type $X_0(i)$, for simplicity we simply put $X_0(i) = i$. Time is counted in discrete generations $t = 0, 1, \ldots$. In generation $t+1$ each individuum i picks a parent $j \sim U(\{1, \ldots, L\})$ uniformly at random, and adopts its type, i.e. $X_{t+1}(i) = X_t(j)$. This leads to a discrete-time Markov chain $(X_t : t \in \mathbb{N})$.

(a) Give the state space of the Markov chain $(X_t : t \in \mathbb{N})$. Is it irreducible? What are the stationary distributions?
(Hint: if unclear do (c) first to get an idea.)

(b) Let N_t be the number of individuals of a given species at generation t, with $N_0 = 1$. Is $(N_t : t \in \mathbb{N})$ a Markov process? Give the state space and the transition probabilities.
Is the process irreducible? What are the stationary distributions? What is the limiting distribution as $t \to \infty$ for the initial condition $N_0 = 1$?

(c) Simulate the dynamics of the full process $(X_t : t \in \mathbb{N})$ up to generation T.
Store the trajectory $(X_t : t = 1, \ldots, T)$ in a $T \times L$ matrix, with ordered types such that $X_t(1) \leq \ldots \leq X_t(L)$ for all t, and visualise the matrix with a heat map.
You may use the suggested parameter value $L = 100$ and appropriate T, or any other that make sense (it is a good idea to vary them to get a feeling for the model).
Address the following points, supported by appropriate visualisations:
- Explain the emerging patterns in a couple of sentences, what will happen when you run the simulation long enough?
- How long will it roughly take on average to reach stationarity (depending on L)?
Test your answer using (at least) three values for L, e.g. 10, 50 and 100.