Networks and Random Processes

Problem sheet 1

Sheet counts $40/100$ homework marks, [x] indicates weight of the question.
Please put solutions in my pigeon hole or give them to me by **Friday, 13.10.2017, 2pm**.

1.1 Simple random walk (SRW) [13]

(a) Consider a simple random walk on $\{1, \ldots, L\}$ with probabilities $p \in [0, 1]$ and $q = 1 - p$ to jump right and left, respectively, and consider periodic as well as closed boundary conditions. In each case, sketch the transition matrix P of the process (see lectures), decide whether the process is irreducible, and give all stationary distributions $\boldsymbol{\pi}$ and state whether they are reversible. (Hint: Use detailed balance.)
Discuss the cases $p = 1$ and $p = q = 1/2$ separately from the general case $p \in (0, 1)$.

(b) Consider the same SRW with absorbing boundary conditions, sketch the transition matrix P, decide whether the process is irreducible, and give all stationary distributions $\boldsymbol{\pi}$ and state whether they are reversible.
Let $h_k^L = \mathbb{P}[X_n = L \text{ for some } n \geq 0 | X_0 = k]$ be the absorption probability in site L. Give a recursion formula for h_k^L and solve it for $p \neq q$ and $p = q$.

(c) Consider a tree, i.e. an undirected, connected graph (G, E) without loops and double edges. $e(x, y) = e(y, x) \in \{0, 1\}$ denotes the presence of an undirected edge (x, y), and $d(x) = \sum_{y \in G} e(x, y)$ is the degree of vertex x. A simple random walk on (G, E) has transition probabilities

$$p(x, y) = e(x, y)/d(x) \quad \text{where by assumption } d(x) > 0 \text{ for all } x\ .$$

Find a formula for the stationary distribution $\boldsymbol{\pi}$.

1.2 Geometric random walk [14]

Let $X_1, X_2, \ldots$ be a sequence of iidrv's with $X_i \sim \mathcal{N}(\mu, \sigma^2)$ where $\mu \in \mathbb{R}$ and $\sigma^2 > 0$. Consider the discrete-time random walk (DTRW) on state space $\mathbb{R}$

$$(Y_n : n \geq 0) \quad \text{with} \quad Y_{n+1} = Y_n + X_{n+1} \quad \text{and} \quad Y_0 = 0\ .$$

(a) State the weak law of large numbers and the central limit theorem for this process.

(b) Using that sums of Gaussian random variables are again Gaussian, what is the distribution of Y_n for any arbitrary $n \geq 0$?

Now consider the discrete-time process $(Z_n : n \geq 0)$ on the state space $(0, \infty)$ with $Z_n = \exp(Y_n)$, which is called a **geometric random walk**.

(c) Give a recursive definition of $(Z_n : n \geq 0)$ analogous to the above.
What is the distribution of Z_n (look up log-normal on the web)?
Give the PDF, its mean, variance and median.

(d) Simulate $M = 500$ realizations of Z_n for $n = 0, \ldots, 100$ with $\mu = 0$ and $\sigma = 0.2$.
Plot the empirical average

$$\hat{\mu}_n^M := \frac{1}{M} \sum_{i=1}^{M} Z_n^i$$

as a function of time n, with error bars indicating the standard deviation.
Compare the empirical PDF at times $n = 1,\ 10$ and 100 to the theoretical prediction.

(e) For fixed $\sigma = 0.2$ pick μ such that $\mathbb{E}[Z_n] \equiv 1$ for all $n \geq 0$, and repeat the previous simulation.

(f) In the following consider the choice of parameters from (e).
Is $(Z_n : n \geq 0)$ a stationary process?
As the number of realizations $M \to \infty$, does the empirical average $\hat{\mu}_n^M$ converge for fixed $n \geq 0$, and what is the limit?
For fixed finite M, do you think that $\hat{\mu}_n^M$ converges as $n \to \infty$, and what is the limit?
Do you think the limits commute, i.e. $\lim_{n\to\infty} \lim_{M\to\infty} \hat{\mu}_n^M = \lim_{M\to\infty} \lim_{n\to\infty} \hat{\mu}_n^M$?

1.3 Wright-Fisher model of population genetics [13]

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.
Visualise the matrix with a heat map.
You may use the suggested parameter values $L = 100$, $T = 500$ 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 to reach stationarity (depending on L)? Test your answer using three values for L, e.g. 10, 50 and 100.