Stochastic Modelling and Random Processes - Class test

The class test counts $80/100$ module marks, [x] indicates weight of each question.
At the end of this paper, you can find some supporting material.
Attempt all 4 questions and justify all your answers

1. **General concepts** **[30]**

   (a) Consider a continuous time stochastic process $(X_t : t \geq 0)$ with state space S.
   When is this process a Markov chain? And when is it time-homogeneous?

   (b) Consider a discrete time Markov chain (DTMC) $(X_n : n \in \mathbb{N}_0)$ with state space S.
   What is a (i) stationary and a (ii) reversible distribution of a DTMC?
   Show that a reversible distribution is stationary.

   (c) What does it mean for a DTMC to be ergodic, and under what conditions is a DTMC ergodic?
   State the ergodic theorem and explain how to use it to justify the use of Markov Chain Monte Carlo (MCMC) to estimate, e.g., expectations.

   (d) Consider the following MCMC algorithm to sample from a given distribution $\boldsymbol{\pi}(x)$:
   Starting from $X_0 = x$, for a given $x \in \mathbb{R}$, for each $n \in \mathbb{N}_0$,

      1. Generate $Y \sim \mathcal{N}(X_n, \sigma^2)$, i.e. Y is a Gaussian random variable centered at X_n, the previous iteration, or $q(x, y) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{(y-x)^2}{2\sigma^2}}$.
      2. Set $X_{n+1} = Y$ with probability $a(X_n, Y)$, where

      $$a(x, y) = \min\left\{1, \frac{\pi(y)q(y, x)}{\pi(x)q(x, y)}\right\}$$

      3. Otherwise, reject Y and set $X_{n+1} = X_n$.

   Show that this algorithm generates a DTMC whose stationary distribution is $\boldsymbol{\pi}$, and therefore, assuming that this chain is ergodic, we can use this to sample from $\boldsymbol{\pi}$.
   Hint: Use reversibility. You do not need to show that this DTMC is ergodic.

   (e) Give the definition of a diffusion process $(X_t : t \geq 0)$ on $\mathbb{R}$ and write down its generator and the corresponding stochastic differential equation (SDE).

   (f) Consider the contact process $(\eta_t : t \geq 0)$ where an *infected* individual i infects individual j with rate $q(i, j) \geq 0$. Give the state space of the model and the transition rates $c(\eta, \eta^i)$ using the standard notation: if the state of individual i is changed, then

   $$\eta^i(k) = \begin{cases} \eta(k) & , k \neq i \\ 1 - \eta(k) & , k = i. \end{cases}$$

   Is this process ergodic (justify your answer)? Give a formula for all stationary distributions of the process, assuming that $q(i, j)$ is irreducible and the state space is finite.

   (g) Consider the following degree sequences:

   $$D_1 = (0, 1, 2, 3)\ , \quad D_2 = (3, 3, 3, 3)\ , \quad D_3 = (0, 1, 1, 2)\ ,$$
   $$D_4 = (2, 3, 3, 2)\ , \quad D_5 = (1, 1, 1, 1).$$

   For each $i = 1, \ldots, 5$, decide whether D_i is graphical. If yes, draw a graph with that degree sequence, and if not, explain why not.

2. **Markov chains** **[20]**

(a) Consider a **continuous-time Markov chain** (CTMC) $(X_t : t \geq 0)$ with state space S. State (without proof) the Chapman-Kolmogorov equations and explain how you can use them to define the generator of the CTMC and thus the transition matrix $P_t = (p_t(x, y) : x, y \in S)$, where $p_t(x, y)$ is the probability of transitioning from x to y after time t.

(b) Let $(X_t : t \geq 0)$ be a continuous time Markov chain (CTMC) with state space $S = \{0, 1\}$, $X_0 = 0$ and generator matrix

$$G = \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix}.$$

i. Give the stationary distribution $\boldsymbol{\pi}$. Is this stationary distribution reversible? And is it ergodic? Justify your answer.

ii. Give the transition matrix P^Y for the corresponding jump chain $(Y_n : n \in \mathbb{N}_0)$ and its stationary distribution $\boldsymbol{\pi}^Y$. Is this distribution reversible? And is it ergodic? Justify your answer.

iii. Compute the eigenvalues of G, and using the fact that $G = Q^{-1}\Lambda Q$ for some matrix Q give the transition matrix $P(t) = \exp(tG)$ for all $t > 0$ in terms of Q and the eigenvalues of G.
Hint: You do not need to specify Q but you need to specify Λ.

iv. Assume that $P(t) = \begin{bmatrix} p_{11}(t) & p_{12}(t) \\ p_{21}(t) & p_{22}(t) \end{bmatrix}$, where $p_{ij}(t) = a_{ij} + b_{ij}e^{\lambda t}$. Compute $p_{ij}(t)$, i.e. determine the coefficients in a_{ij} and b_{ij} for all $i,\ j = 1, 2$.
Hint: Again, it is not necessary to compute the matrix Q, instead use what you know about $p_{ij}(0)$ *and* $\frac{d}{dt}p_{ij}(t)|_{t=0}$.

3. **Diffusion processes** **[20]**

(a) Define the standard Brownian motion $(B_t : t \geq 0)$ and write down its generator $\mathcal{L}_B f(x)$.

(b) Explain how the "derivative" $\frac{dB_t}{dt}$ of a Brownian motion could be considered (if we allowed for this) as a Gaussian process with mean zero and covariance $\sigma(s, t) = \delta(|t - s|)$.

(c) State Itô's formula for $(X_t : t \geq 0)$ and a smooth function $f : \mathbb{R} \to \mathbb{R}$. Use it to derive an SDE for $Y_t = B_t^2$ and show that

$$\int_0^t B_t \, dB_t = \frac{1}{2}\left(B_t^2 - t\right).$$

Comment on the fact that we showed in lectures that $\mathbb{E}\left(\int_0^T B_t \, dB_t\right) = 0$.

(d) In mathematical finance, the Cox-Ingersoll-Ross model (below) is used to model interest rates:

$$dX_t = \alpha(\beta - X_t)\, dt + \sigma\sqrt{X_t}\, dB_t, \qquad X_0 = 0.$$

i. Use the evolution equation of expectation values of test functions $f : \mathbb{R} \to \mathbb{R}$ to derive ODEs for the mean $m(t) := \mathbb{E}(X_t)$ and the second moment $\mathbb{E}(X_t^2)$.

ii. Solve these ODEs and obtain expressions for the mean $m(t)$ and variance $v(t)$ of the solution of this SDE.

iii. What are the mean and variance of this process as $t \to \infty$?

iv. Write down the Fokker-Planck equation for the density of this diffusion process. Simplify this as much as possible.

4. **Graphs and Networks** **[30]**

(a) Choose two of the models of random graphs that we discussed in lectures. Define these graphs and discuss their main properties, including what makes them appropriate (or not) to model real-life networks.

(b) Consider an undirected graph with degree sequence $(2, 4, 3, 3, 3, 5)$.

i. Draw a graph G with this degree sequence and write its adjacency matrix A. You should label the vertices 1 to 6.

ii. Identify all cliques and draw a spanning tree of this graph.
Note: if there is no spanning tree, explain why.

iii. Write down a matrix whose entries are the vertex distances $\big(d_{ij} : i, j = 1, \ldots, 6\big)$ (you can use that $d_{ii} = 0\ \forall i$). Compute the characteristic path length $L(G)$, the diameter $\mathrm{diam}(G)$ of G and the closeness centrality for each node.

iv. Compute the degree distribution $p(k)$ and the average degree $\langle k \rangle$ of G.

v. Compute the local clustering coefficients C_i and their average $\langle C_i \rangle$. Use these to compute the global clustering coefficient C.

vi. Give all non-zero entries of the joint degree distribution $q(k, k')$ and compute the marginal distribution $q(k) = \sum_{k'} q(k, k')$. For all k' with $q(k') > 0$ compute the conditional distribution $q(k|k')$ and the corresponding expectation $k_{nn}(k')$.

Supporting material

If you need to, (and you may not need to), at any point, you can use the folowing properties of the Itô integral:

1. Additivity: $\int_0^s f(t, X_t)\, dB_t + \int_s^T f(t, X_t)\, dB_t = \int_0^T f\, dB_t$.
2. Linearity: for all $\alpha, \beta \in \mathbb{R}$ and f, g such that the integral exists, we have
$$\int_0^T (\alpha f(t, X_t) + \beta g(t, X_t))\, dB_t = \alpha \int_0^T f(t, X_t)\, dB_t + \beta \int_0^T g(t, X_t)\, dB_t.$$
3. If f is a deterministic function (i.e. it depends on time only), we have
$$\mathbb{E}\left(\int_0^T f(t)\, dB_t \right) = 0.$$
4. It verifies the Itô isometry:
$$\mathbb{E}\left(\left(\int_0^T f(s)\, dW_t \right)^2 \right) = \mathbb{E}\left(\int_0^T f^2(t)\, dt \right),$$
and, more generaly
$$\mathbb{E}\left(\int_0^t f(s)\, dB_s \int_0^u g(s)\, dB_s \right) = \mathbb{E}\left(\int_0^{\min(t,u)} f(s) g(s)\, dB_s \right).$$
5. The above two properties imply that if f is deterministic, then the stochastic integral $\mathcal{I}_T$ is a Gaussian random variable with mean zero and variance $\int_0^T f^2(s)\, ds$.