EXAMINATION PAPER

Examination Session:	Year:	Exam Code:
May/June	2023	MATH4131-WE01

Title:
Probability IV

Time:	3 hours	
Additional Material provided:		
Materials Permitted:		
Calculators Permitted:	No	Models Permitted: Use of electronic calculators is forbidden.

Instructions to Candidates:	Answer all questions. Section A is worth 20%, Section B is worth 60%, and Section C is worth 20%. Within Sections A and B, all questions carry equal marks. Students must use the mathematics specific answer book.

Revision:	

ED01/2023

SECTION A

Q1 Suppose the random variables X and Y are in L^2.

(a) Prove the Cauchy–Schwarz inequality

$$\big(\mathbb{E}(XY)\big)^2 \leq \mathbb{E}(X^2)\mathbb{E}(Y^2).$$

(b) What can you deduce about X and Y when equality holds?

(c) Define the correlation coefficient $\rho(X, Y)$ and prove that $|\rho(X, Y)| \leq 1$.

Q2 (a) Carefully define the stochastic order $\preccurlyeq$ for random variables.

(b) Let $X \sim \mathsf{Bin}(n_1, p)$ and $Y \sim \mathsf{Bin}(n_2, p)$, where n_1 and n_2 are positive integers, and $p \in (0, 1)$. What further conditions required on n_1, n_2 and p are necessary and sufficient for $X \preccurlyeq Y$ to hold?

(c) Let $U \sim \mathsf{Bin}(2, 1/4)$ and $V \sim \mathsf{Bin}(2, 1/2)$.

(i) Find the maximal coupling of U and V.

(ii) Is the maximal coupling of U and V a monotone coupling? Construct all monotone couplings of U and V.

SECTION B

Q3 Let $(S_i)_{i\geq 0}$ be a symmetric simple random walk starting at the origin (and making independent jumps ± 1 with probability $1/2$). For all $m \geq 0$ and $\ell \geq 1$ set

$$u_{2m} = \mathbb{P}(S_{2m} = 0), \quad f_{2\ell} = \mathbb{P}(S_1 \neq 0, \dots, S_{2\ell-1} \neq 0, S_{2\ell} = 0).$$

Call a time interval $(i, i+1)$ positive if at least one of S_i or S_{i+1} is positive, and call it negative otherwise.

For $n \geq 1$, consider the initial $2n$-step trajectory $(S_i)_{i=0}^{2n}$, and for $k = 0, 1, \dots, n$, let $A_{2k,2n}$ be the event that this trajectory has $2k$ positive time intervals, and $2n - 2k$ negative time intervals. Set $p_{2k,2n} = \mathbb{P}(A_{2k,2n})$.

(a) Show that $u_{2n} = \sum_{k=1}^{n} f_{2k} u_{2n-2k}$ for all $n \geq 1$.

(b) Using induction, prove the following statement for all $n \geq 1$:

$$p_{2k,2n} = u_{2k} u_{2n-2k} \text{ for all } k = 0, 1, \dots, n.$$

Hint: You may use without proof the identity $f_{2k} = u_{2k-2} - u_{2k}$ for all $k \geq 1$.

(c) Calculate a large-n approximation of $p_{4n,4n}/p_{2n,4n}$, the ratio of the probability that a $4n$-step trajectory has all positive time intervals to the probability that it has exactly half of its time intervals being positive.

Q4 (a) Let $(u_n)_{n\geq 0}$ be a sequence defined by $u_0 = 1$ and, for $n > 0$, $u_n = \sum_{k=1}^{n} f_k u_{n-k}$, where $(f_k)_{k\geq 1}$ is a real sequence with $f_k \geq 0$ and $0 < \sum_{k=1}^{\infty} f_k < \infty$. Denote by $\mathcal{U}(s) = \sum_{n\geq 0} u_n s^n$ and $\mathcal{F}(s) = \sum_{k\geq 1} f_k s^k$ the associated generating functions. Prove that $\mathcal{U}(s) = \big(1 - \mathcal{F}(s)\big)^{-1}$ for all s in an appropriate open disc $\{s \in \mathbb{C} : |s| < s_0\}$ that you should define.

A **red-blue polymer** is defined as follows. A red-blue polymer of length n is composed of monomers (blocks of length 1), and dimers (blocks of length 2), where each block may be either red or blue, and the order of the blocks matters. For example, there are 2 red-blue polymers of length 1 (namely, 1_r and 1_b) and 6 red-blue polymers of length 2 ($1_\mathsf{r} + 1_\mathsf{r}, 1_\mathsf{r} + 1_\mathsf{b}, 1_\mathsf{b} + 1_\mathsf{r}, 1_\mathsf{b} + 1_\mathsf{b}, 2_\mathsf{r}$ and 2_b).

(b) Let u_n be the number of red-blue polymers of length n. Using part (a) or otherwise, find an exact expression for u_n.

(c) Verify the asymptotic behaviour of your answer to part (b) using the renewal theorem to find constants $c > 0$ and $\alpha > 1$ such that

$$\lim_{n\to\infty} \frac{u_n}{\alpha^n} = c.$$

(d) Define p_n to be the **proportion** of red-blue polymers of length n that are monochromatic (i.e., having all blocks of the same colour). Find constants $d > 0$ and $\beta < 1$ such that

$$\lim_{n\to\infty} \frac{p_n}{\beta^n} = d.$$

Q5 Let $X_1, X_2, \ldots, X_n$ be independent $\mathcal{U}(0, 1)$ random variables, and for $i, j = 1, \ldots, n$, let $L_{i,j} := X_{(i)}/X_{(j)}$ be the ratio of the i-th order variable $X_{(i)}$ and the j-th order variable $X_{(j)}$.

(a) Calculate the conditional density $f_{X_{(i)}|X_{(j)}}(x \mid y)$ of $X_{(i)}$ given $X_{(j)}$ as a function of x and y for all $i < j$.

(b) Prove that for $i < j$ and all $0 < y < 1$ the conditional density $f_{L_{i,j}|X_{(j)}}(x \mid y)$ of $L_{i,j}$ given $X_{(j)}$ satisfies

$$f_{L_{i,j}|X_{(j)}}(x \mid y) = y f_{X_{(i)}|X_{(j)}}(xy \mid y)$$

for an appropriate range of values for x that you should find.

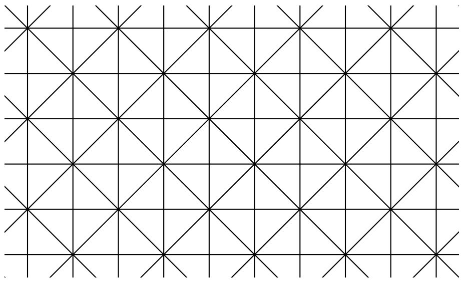
(c) For fixed $i < j$, show that $L_{i,j}$ and $X_{(j)}$ are independent and find the distribution of $L_{i,j}$.

(d) Prove that the rescaled random variable $nL_{1,n}$ converges in distribution as $n \to \infty$. What is the limiting distribution?

Q6 (a) Describe the Bernoulli bond percolation model on a fixed infinite graph G. Your answer should include a definition of the critical probability $p_{\mathrm{cr}}(G)$, and any other notation you use should be clearly defined.

(b) Suppose $G = (V(G), E(G))$ and $H = (V(H), E(H))$ are two infinite graphs with $V(G) = V(H)$ and $E(G) \subseteq E(H)$. Prove that $p_{\mathrm{cr}}(G) \geq p_{\mathrm{cr}}(H)$.

(c) Consider the Bernoulli bond percolation model on the graph T, pictured below.



Prove that $p_{\mathrm{cr}}(T) > 0$ and find an explicit value $p > 0$ with $p_{\mathrm{cr}}(T) \geq p$.

SECTION C

Q7 (a) Carefully state Cramér's theorem on large deviations for sums of independent identically distributed random variables, $X_1, X_2, \ldots$.

(b) Show that under the conditions of Cramér's theorem the rate function $I(\cdot)$ for X_1 satisfies the inequality $I(z) \geq 0$ for all $z \in \mathbb{R}$. For which values of z does $I(z) = 0$?

Hint: You may assume that $f(t) := \log \mathbb{E}\big(\exp(tX_1)\big)$ is a strictly convex function of t.

(c) Let $X_1, X_2, \ldots$ be independent $\mathcal{N}(\mu, \sigma^2)$ normal random variables. Denote $S_n = X_1 + \cdots + X_n$.

(i) Compute the moment generating function

$$\phi(t) := \mathbb{E}\big(\exp(tX_1)\big)$$

and the corresponding rate function $I(\cdot)$.

(ii) Find the limit

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \geq an)$$

for all $a > \mathbb{E}X_1$. What happens for $a \leq \mathbb{E}X_1$?