IMPERIAL

MATH50010

BSc, MSci and MSc EXAMINATIONS (MATHEMATICS)
May 2024

This paper is also taken for the relevant examination for the Associateship of the Royal College of Science

Probability for Statistics

Date: Friday, May 3, 2024

Time: 10:00 – 12:00 (BST)

Time Allowed: 2 hours

This paper has 4 Questions.

Please Answer Each Question in a Separate Answer Booklet

Candidates should start their solutions to each question on a new sheet of paper.

Supplementary books may only be used after the relevant main book(s) are full.

Any required additional material(s) will be provided.

Allow margins for marking.

Credit will be given for all questions attempted.

Each question carries equal weight.

DO NOT OPEN THIS PAPER UNTIL THE INVIGILATOR TELLS YOU TO

1. Throughout this question, let Ω be a set and let $\mathcal{F}$ be a sigma algebra on Ω. For sets $A_1, A_2, \dots$ and B you may use without proof that $\bigcup_{i=1}^{\infty}(A_i \cap B) = B \cap \bigcup_{i=1}^{\infty} A_i$ and $\bigcap_{i=1}^{\infty}(A_i \cup B) = B \cup \bigcap_{i=1}^{\infty} A_i$.

(a) State the definition of a probability measure under the Kolmogorov Axioms. (2 marks)

(b) If $|\Omega| = m$ for some finite $m \geq 1$, what is the maximal and minimal number of elements that can be in $\mathcal{F}$? You should give your answer in terms of m and justify your answer. (3 marks)

(c) Let $E \subset \Omega$, $E \neq \emptyset$, and define $\mathcal{G} = \{E \cap A : A \in \mathcal{F}\}$. Prove that $\mathcal{G}$ is a sigma algebra on the set E. (3 marks)

(d) State two equivalent conditions for the function $X : \Omega \to \mathbb{R}$ to be a random variable with respect to $\mathcal{F}$. (2 marks)

(e) Let X and Y be random variables defined on the probability space $(\Omega, \mathcal{F}, \Pr)$. Prove that $X + Y$ is a random variable. *[Hint: You may use without proof that between two distinct real numbers there exists a rational number.]* (3 marks)

(f) Let $X_1, X_2, \dots$ be a sequence of independent and identically distributed random variables with common mean 0 and variance σ^2 defined on the probability space $(\Omega, \mathcal{F}, \Pr)$. Let N be a random variable defined on the same probability space, where N takes values in the set $\{0, 1, \dots\}$ and is independent of all $X_1, X_2, \dots$. Define $S_N = X_1 + \cdots + X_N$.

(i) Using only results from this question, prove that S_N is a random variable with respect to $\mathcal{F}$. (4 marks)

(ii) Show that $\mathbb{E}[S_N^2] = \sigma^2 \mathbb{E}[N]$. (3 marks)

(Total: 20 marks)

2. For this question, it may be helpful to recall the cumulative distribution function of a standard Gumbel distribution: if $Y \sim$ Gumbel,

$$F_Y(y) = \exp\{-\exp\{-y\}\} \text{ for all } y \in \mathbb{R}.$$

Moreover, if $X \sim$ Exponential(λ) with $\lambda > 0$ then X has cumulative distribution function,

$$F_X(x) = 1 - \exp\{-\lambda x\} \text{ for } x \geq 0, \quad \text{and } F_X(x) = 0 \text{ elsewhere}$$

and moment generating function $M_X(t) = \frac{\lambda}{\lambda - t}$ for $t < \lambda$.

(a) Define convergence in distribution and convergence in probability for the sequence of random variables $X_1, X_2, \ldots$. (2 marks)

(b) Let $A_1, A_2, \ldots$ be a sequence of events and let $\lambda_n = \sum_{i=1}^{n} \Pr(A_i)$ for any $n \in \mathbf{N}$.

(i) Assume $\lim_{n\to\infty} \frac{\lambda_n}{n} = 0$ and consider the random variables $X_n = \frac{1}{n}\sum_{i=1}^{n} \mathbf{I}\{A_i\}$ where $\mathbf{I}$ is the indicator function. Determine whether X_n converges in a) probability and b) in distribution, stating the limiting random variable if so. Briefly justify your answers. (4 marks)

(ii) State sufficient conditions on λ_n and $A_1, A_2, \ldots$ with $A_n \subsetneq \Omega$ for all $n \in \mathbf{N}$ such that $\Pr(\{A_n \text{ i.o.}\}) = 1$. (2 marks)

(iii) Assume that $A_1, A_2, \ldots$ are a sequence of events with $\lim_{n\to\infty} \Pr(A_n) = 0$ and that $\sum_{n=1}^{\infty} \Pr(A_n \cap A_{n-1}^C) < \infty$. Prove that $\Pr(\{A_n \text{ i.o.}\}) = 0$. (5 marks)

(c) Let $X_1, X_2, \ldots$ be a sequence of independent Exponential(1) random variables.

(i) Find the cumulative distribution function of $\max_{1\leq i\leq n} X_i$. (3 marks)

(ii) Show that $\max_{1\leq i\leq n} X_i - \log(n) \xrightarrow{D} Y$ where $Y \sim$ Gumbel. (4 marks)

(Total: 20 marks)

3. Let X and Y be bivariate Normal random variables with means μ_X, μ_Y and variance-covariance matrix Σ. It may be helpful to recall that for any $x, y \in \mathbb{R}$, the probability density function of a bivariate Normal is given by,

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{\det(\Sigma)}} \exp\left\{-\frac{1}{2}\begin{pmatrix} x-\mu_x, & y-\mu_y \end{pmatrix} \Sigma^{-1} \begin{pmatrix} x-\mu_x \\ y-\mu_y \end{pmatrix}\right\}.$$

You may also assume that $\lim_{x\to 0} x\log(x) = 0$ and use any distributions given in the other questions.

(a) Write down $\Pr(X = Y)$, briefly justifying your answer. (2 marks)

(b) Prove that if X and Y are bivariate Normal random variables with correlation, Cor$(X,Y) = 0$ and variances $\sigma_X^2, \sigma_Y^2 > 0$, then X and Y are independent. (4 marks)

(c) Let Cor$(X,Y) = \rho$ for some $\rho \neq 0$. Find a constant $c \in \mathbb{R}$ such that X and $X - cY$ are independent. (4 marks)

(d) Let $Z = e^X$. Calculate the conditional density $f_{Z|Y}(z|y)$ for (X,Y) bivariate Normal, stating any results from lecture notes that you use. (4 marks)

(e) Let $X_1, X_2 \ldots$ be a sequence of independent Uniform$(0,1)$ random variables.

(i) Determine whether $\frac{\sum_{i=1}^n X_i^2}{n}$ converges in probability as $n \to \infty$, stating the limiting random variable if so. Briefly justify your answer, stating any results from lecture notes that you use. (2 marks)

(ii) Determine $\lim_{n\to\infty} \Pr((\frac{1}{1-X_1} \cdots \frac{1}{1-X_n})^{\frac{1}{n}} > e^1)$, stating any results from lecture notes that you use. (4 marks)

(Total: 20 marks)

4. For this question we will consider time-homogeneous Markov chains on a state space $\mathcal{E}$.

(a) Define mathematically what it means for a set of states to be irreducible, defining any relevant notation or concepts. (2 marks)

(b) Consider the time-homogeneous Markov chain with the transition diagram given in Figure 1.

(i) Write down the transition matrix for this Markov chain. (1 mark)

(ii) For each state, determine whether it is recurrent or whether it is transient. (2 marks)

(iii) Find the period of state 2. (2 marks)

(iv) Find the stationary distribution of the Markov chain. (3 marks)

(v) Show that for any starting state, the chain gets absorbed in state 5 with probability 1. (4 marks)

For parts (ii)-(v) you should justify your answer, stating any results that you use.

(c) Let $(X_n)_{n\in\mathbb{N}\cup\{0\}}$ be an irreducible time-homogeneous Markov Chain with transition matrix given by P on the state space $\mathcal{E}$ with $|\mathcal{E}| \geq 2$. Define the transition matrix Q on the same state space $\mathcal{E}$ by

$$Q_{ij} = \begin{cases} 0 & \text{if } i = j \\ (1 - P_{ii})^{-1} P_{ij} & \text{if } i \neq j \end{cases} \qquad \text{for any } i, j \in \mathcal{E}.$$

(i) Show that Q is a valid transition matrix. (2 marks)

(ii) If the Makov chain with transition matrix given by P is irreducible, is the Markov chain with transition matrix given by Q irreducible? Provide a proof or counter-example to justify your answer. (4 marks)

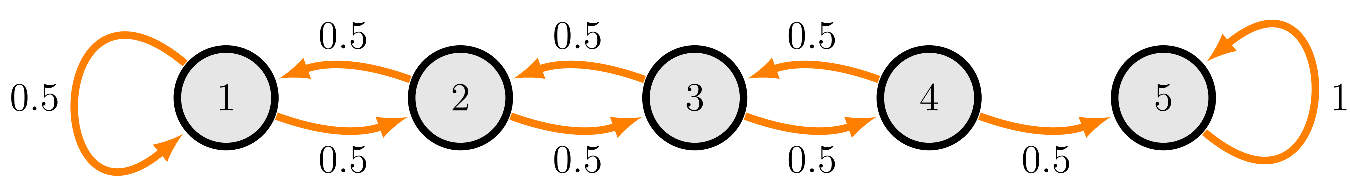

Figure 1: The five-state Markov chain in Part (b).

(Total: 20 marks)

BSc and MSci EXAMINATIONS (MATHEMATICS)

May 2024

This paper is also taken for the relevant examination for the Associateship.

XXX

XXX (Solutions)

Setter's signature	Checker's signature	Editor's signature
. . . .	. . . .	. . . .

1. (a) Given a set Ω and a sigma algebra $\mathcal{F}$ on Ω, a probability measure is a function $\Pr : \mathcal{F} \to [0,1]$ such that

   seen $\Downarrow$

   2, A

   1. $\Pr(A) \geq 0$, for all $A \in \mathcal{F}$.
   2. $\Pr(\Omega) = 1$.
   3. Countable additivity: If $A_1, A_2, \ldots \in \mathcal{F}$ are pairwise disjoint, then $\Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Pr(A_i)$.

   [Note that condition 1 is implied by $\Pr : \mathcal{F} \to [0,1]$*, so it is not necessary to include both to get marks]*

   seen $\Downarrow$

   (b) The maximum size of $\mathcal{F}$ is 2^m since the largest sigma algebra is the power set, the smallest sigma algebra on Ω is $\{\emptyset, \Omega\}$, so the minimal size is 2.

   3, A

   meth seen $\Downarrow$

   (c) We first show $\emptyset \in \mathcal{G}$. Since $\mathcal{F}$ is a sigma algebra, $\emptyset \in \mathcal{F}$ and $E \cap \emptyset = \emptyset$ so $\emptyset \in \mathcal{G}$. Now assume $B \in \mathcal{G}$. Then, there exists $A \in \mathcal{F}$ such that $B = A \cap E$ and by definition, $B \subseteq E$. In E , $B^C = E \cap B^C = E \cap (A^C \cup E^C) = (E \cap A^C) \cup (E \cup E^C) = E \cap A^C \in \mathcal{G}$ since $\mathcal{F}$ is a sigma algebra so $A^C \in \mathcal{F}$. Let $B_1, B_2, \dots \in \mathcal{G}$, then each $B_i = A_i \cap E$ for some $A_i \in \mathcal{F}$. Hence, $\bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} (A_i \cap E) = E \cap \bigcup_{i=1}^{\infty} A_i \in \mathcal{G}$ since $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$. Hence $\mathcal{G}$ satisfies the conditions to be a sigma algebra on E.

   3, C

   seen $\Downarrow$

   (d) X is a a random variable if and only if $\{X \leq x\} = \{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$ for all $x \in \mathbb{R}$. Or X is a random variable if and only if for every Borel set $B \in \mathcal{B}$, $X^{-1}(B) \in \mathcal{F}$.

   2, A

   seen $\Downarrow$

   (e) Note that $\{X + Y \leq z\} = \{X + Y > z\}^c$. Then $X + Y > z$ if and only if $X > z - Y$, which is true if and only if there exists $q \in \mathbb{Q}$ such that $X > q$ and $q > z - Y$. So then

   3, B

$$\{X + Y > z\} = \bigcup_{q \in \mathbb{Q}} \{X > q\} \cap \{Y > z - q\}.$$

   Since X and Y are random variables, $\{X > q\} = \{\omega \in \Omega : X(\omega) > q\} \in \mathcal{F}$ and similarly $\{Y > z - q\} \in \mathcal{F}$, so that $\{X + Y > z\}$ is a countable union of sets in $\mathcal{F}$, and hence also in $\mathcal{F}$.

   unseen $\Downarrow$

   (f) (i) We first prove by induction that for any $n \in \mathbb{N}$, $\sum_{i=1}^{n} X_i$ is a random variable. For the base case, $n = 1$, by Part (e) we know that $X_1 + X_2$ is a random variable. Now assume that S_n is a random variable, we prove that S_{n+1} is also a random variable. Note that $S_{n+1} = S_n + X_{n+1}$ and that S_n and X_{n+1} are both random variables. Hence, by Part (e) S_{n+1} is also a random variable. Therefore for any $n \in \mathbb{N}$, $\sum_{i=1}^{n} X_i$ is a random variable. To show that S_N is a random variable, we consider the possible values N can take, for any $s \in \mathbb{R}$,

   1, C

   3, D

$$\{S_N \leq s\} = \bigcup_{n=1}^{\infty} \{S_n \leq s\} \cap \{N = n\}$$

   Then since S_n is a random variable for any n, $\{S_n \leq s\} \in \mathcal{F}$. Additionally since N is a random variable with respect to $\mathcal{F}$, $\{N = n\} = \{N \leq n\} \cap \{N \leq n - 1\}^C \in \mathcal{F}$ since sigma algebras are closed under intersection. Since sigma algebras are also closed under countable unions, $\{S_N \leq s\} \in \mathcal{F}$ so S_N is a random variable.

unseen $\Downarrow$

3, B

(ii) We use conditional expectations to show that,

$$\begin{aligned}
\mathbb{E}[S_N^2] &= \sum_{n=0}^{\infty} \mathbb{E}[S_N^2|N=n]\Pr(N=n) = \sum_{n=0}^{\infty} \mathbb{E}[(X_1+\cdots+X_n)^2)]\Pr(N=n) \\
&= \sum_{n=0}^{\infty} (\mathbb{E}[X_1^2]+\cdots+\mathbb{E}[X_n^2])\Pr(N=n) \\
&\qquad\qquad (\text{Since } X_i \text{ are independent } \mathbb{E}[X_iX_i] = \mathbb{E}[X_i]\mathbb{E}[X_j] = 0) \\
&= \sum_{n=0}^{\infty} n\sigma^2 \Pr(N=n) \qquad\qquad (\text{Since } \mathbb{E}[X_i^2] = \sigma^2) \\
&= \sigma^2\mathbb{E}[N]
\end{aligned}$$

2. (a) The sequence $X_1, X_2, \ldots$ converges in probability to X if for all $\epsilon > 0$, $\lim_{n\to\infty} \Pr(|X_n - X| \geq \epsilon) = 0$. The sequence converges in distribution to the random variable X with CDF F_X if for all $x \in \mathbb{R}$ where F_X is continuous, $\lim_{n\to\infty} F_n(x) = F_X(x)$ where F_n is the CDF of X_n. seen ⇓ 2, A

(b) (i) We first show convergence in probability. For any $\epsilon > 0$, meth seen ⇓ 4, B

$$\begin{aligned}
0 \leq \Pr(|X_n - 0| \geq \epsilon) &= \Pr(|\sum_{i=1}^{n} \mathbf{I}\{A_i\}| \geq \epsilon n) \\
&\leq \Pr(\sum_{i=1}^{n} \mathbf{I}\{A_i\} \geq \epsilon n) \\
&\leq \frac{\mathbb{E}[\sum_{i=1}^{n} \mathbf{I}\{A_i\}]}{\epsilon n} \qquad \text{(By Markov's inequality)} \\
&= \frac{\lambda_n}{n\epsilon}
\end{aligned}$$

Then, since $\lim_{n\to\infty} \frac{\lambda_n}{n} = 0$, we have that $\lim_{n\to\infty} \Pr(|X_n - 0| > \epsilon) = 0$. Hence $X_n \xrightarrow{P} 0$. Since convergence in probability implies convergence in distribution, it also follows that $X_n \xrightarrow{D} 0$. (Note that we cannot use the weak law of large numbers here since we do not have independence).

(ii) By the second Borel-Cantelli Lemma, it is enough that $\lambda_n \to \infty$ as $n \to \infty$ and $A_1, A_2, \ldots$ to be independent for $\Pr(\{A_n \, i.o.\}) = 1$. seen ⇓ 2, A

(iii) Define $A \setminus B = A \cap B^C$ and first observe that unseen ⇓ 5, D

$$\cup_{n=N}^{\infty} A_n = A_N \cup (A_{N+1} \setminus A_N) \cup (A_{N+2} \setminus A_{N+1}) \cup \cdots = A_N \cup \bigcup_{n=N}^{\infty} (A_N \setminus A_{N-1}).$$

Define $B_N = \bigcup_{n=N}^{\infty} A_n$ and note that B_N are a decreasing sequence of events with $B_1 \supseteq B_2 \supseteq \ldots$. Hence, by the continuity property,

$$\Pr(\{A_n \text{ i.o.}\}) = \Pr(\bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} A_n) = \Pr(\bigcap_{N=1}^{\infty} B_N) = \lim_{N\to\infty} \Pr(B_N)$$

Then, note that applying the first Borel-Cantelli Lemma to $C_n = A_n \setminus A_{n-1}$, we see that $\Pr(\bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} (A_N \setminus A_{N-1})) = \Pr(\{A_n \cap A_{n-1}^C \text{ i.o.}\}) = 0$. Using this, together with the assumptions of the question and the continuity property again applied to the decreasing sequence $D_N = \bigcup_{n=N}^{\infty} C_n$, we see that,

$$\begin{aligned}
0 \leq \Pr(\{A_n \text{ i.o.}\}) &= \lim_{N\to\infty} \Pr(B_N) = \lim_{N\to\infty} \Pr(A_N \cup \bigcup_{n=N}^{\infty} (A_N \setminus A_{N-1}) \\
&\leq \lim_{N\to\infty} \Pr(A_N) + \lim_{N\to\infty} \Pr(\bigcup_{n=N}^{\infty} (A_N \setminus A_{N-1})) \\
&= \lim_{N\to\infty} \Pr(A_N) + \Pr(\bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} (A_N \setminus A_{N-1})) = 0 + 0 = 0
\end{aligned}$$

(c) (i) By independence, seen ⇓ 3, A

$$\Pr(\max_{1\leq i\leq n} X_i \leq x) = \Pr(\bigcap_{i=1}^{n} \{X_i \leq x\}) = \prod_{i=1}^{n} \Pr(X_i \leq x) = (1 - \exp\{-x\})^n$$

unseen $\Downarrow$

4, C

(ii) Let $Y_n = \max_{1 \le i \le n} X_i - \log(n)$ and derive the CDF of Y_n. For all $x \in \mathbb{R}$,

$$\begin{aligned}\Pr(Y_n \le x) &= \Pr(\max_{1 \le i \le n} X_i \le x + \log(n)) = (1 - \exp\{-(x + \log(n))\})^n \\ &= (1 - \frac{e^{-x}}{n})^n\end{aligned}$$

So $\lim_{n \to \infty} \Pr(Y_n \le x) = \lim_{n \to \infty} (1 - \frac{e^{-x}}{n})^n = \exp\{-\exp\{-x\}\}$ for all $x \in \mathbb{R}$, so the limiting random variable follows a Gumbel distribution.

3. (a) $\Pr(X = Y) = 0$ since $X - Y$ is a Gaussian random variable so continuous, and by lecture notes for any continuous random variable $\Pr(X - Y = 0) = 0$. sim. seen ⇓ 2, A

(b) If $\mathsf{Cor}(X, Y) = 0$, then it follows that $\mathsf{Cov}(X, Y) = \mathsf{Cor}(X, Y)/\sigma_X\sigma_Y = 0$, in this case the off-diagonal entries of the covariance matrix Σ are 0, and the inverse covariance matrix can be easily derived as $\Sigma^{-1} = \begin{pmatrix} 1/\sigma_X^2 & 0 \\ 0 & 1/\sigma_Y^2 \end{pmatrix}$ and $\det(\Sigma) = \sigma_X^2\sigma_Y^2$. Hence, we write the joint density as unseen ⇓ 4, A

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y} \exp\left\{ -\frac{1}{2}(x - \mu_X, y - \mu_Y) \begin{pmatrix} 1/\sigma_X^2 & 0 \\ 0 & 1/\sigma_Y^2 \end{pmatrix} \begin{pmatrix} x - \mu_X \\ y - \mu_Y \end{pmatrix} \right\}$$
$$= \frac{\exp\{-(x - \mu_X)^2/(2\sigma_X^2)\}}{\sqrt{2\pi\sigma_X^2}} \frac{\exp\{-(y - \mu_Y)^2/(2\sigma_Y^2)\}}{\sqrt{2\pi\sigma_Y^2}} = f_X(x)f_Y(y)$$

For f_X the density of a $\mathcal{N}(\mu_X, \sigma_X^2)$ and f_Y the density of $\mathcal{N}(\mu_X, \sigma_X^2)$. Hence the joint density factorizes so we know from lecture notes that the random variables must be independent.

(c) From the previous part, we need to find a c such that $\mathsf{Cov}(X, X - cY) = 0$. Note that, sim. seen ⇓ 4, B

$$\mathsf{Cov}(X, X - cY) = \mathsf{Cov}(X, X) - c\mathsf{Cov}(X, Y) = \sigma_X^2 - c\rho\sigma_X\sigma_Y$$

so setting $c = \sigma_X/(\rho\sigma_Y)$ will ensure that X and $X - cY$ are independent.

(d) By lecture notes, we know that $X|Y = y \sim \mathcal{N}(\rho y, 1 - \rho^2)$ where $\rho = \mathsf{Cor}(X, Y)$. We then note that $g(x) = e^x$ is a monotonic 1-1 transformation with $g^{-1}(z) = \log(z)$, so we can apply the transformation lemma from lecture notes to the conditional density to get that, unseen ⇓ 4, C

$$f_{Z|Y}(z|y) = f_{X|Y}(g^{-1}(z)|y) \left| \frac{dg^{-1}(z)}{dz} \right| = \frac{1}{z\sqrt{2\pi(1 - \rho^2)}} \exp\{-(\log(z) - \rho y)^2/2\}.$$

(e) (i) Note that $\mathbb{E}[X_i^2] = \int_0^1 x^2\,dx = 1/3$ and $\mathsf{Var}(X_i^2) = \mathbb{E}[X_i^4] - \mathbb{E}[X_i^2]^2 = \int_0^1 x^4\,dx - 1/9 = 1/5 - 1/9 < \infty$ so we can apply the weak law of large numbers to X_i^2. Hence, $\frac{\sum_{i=1}^n X_i^2}{n} \xrightarrow{P} \mathbb{E}[X_i^2] = 1/3$. sim. seen ⇓ 2, A

(ii) We first note that $\Pr((\frac{1}{1-X_1} \cdots \frac{1}{1-X_n})^{\frac{1}{n}} > e^{-1}) = \Pr(\frac{1}{n}\sum_{i=1}^n \log(\frac{1}{1-X_i}) > 1)$. Then we apply the central limit theorem to the random variables $Z_i = \log(\frac{1}{1-X_i})$. For this, observe that since the X_i's are i.i.d., the Z_i's are also i.i.d.. Additionally, by the probability integral transform, $Z_i \sim \mathsf{Exponential}(1)$. Therefore, Z_i have finite mean, $\mu = 1$, and variance, $\sigma^2 = 1$, and the moment generating function exists for all $t < 1$, so it exists in a region around 0. We can therefore apply the Central Limit Theorem to show, unseen ⇓ 4, D

$$\lim_{n\to\infty} \Pr((\frac{1}{1 - X_1} \cdots \frac{1}{1 - X_n})^{\frac{1}{n}} > e^1) = \lim_{n\to\infty} \Pr\left(\frac{1}{n}\sum_{i=1}^n Z_i > 1\right)$$
$$= \lim_{n\to\infty} \Pr\left(\frac{\sqrt{n}(\frac{1}{n}\sum_{i=1}^n Z_i - 1)}{\sigma} > 0\right)$$
$$= \Pr(Z > 0) = 1/2$$

where $Z \sim \mathcal{N}(0, 1)$ and we have used symmetry of the standard normal.

4. (a) A set of states C is said to be irreducible if $i \leftrightarrow j$ for all $i, j \in C$, where $i \leftrightarrow j$ means that states, i and j communicate, i.e. there exists $n, m \geq 0$ such that $p_{ij}(n) > 0, p_{ji}(m) > 0$. $\boxed{\text{seen} \Downarrow}$ $\boxed{2,\ \text{A}}$

(b) (i) $\boxed{\text{meth seen} \Downarrow}$ $\boxed{1,\ \text{A}}$

$$P = \begin{pmatrix} 0.5 & 0.5 & 0 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$\boxed{2,\ \text{A}}$

(ii) A state is recurrent if we return to it with probability 1. For all states 1,2,3,4, there is positive probability that we go from that state to state 5 then never return to our original state. Therefore, states 1,2,3,4 are all transient. There is probability 1 that we return to state 5 if we start there so state 5 is recurrent. $\boxed{2,\ \text{A}}$

(iii) We calculate the period of state 2 by considering the greatest common divisor of all times we can get back to state 2, $d(2) = gcd\{2, 3, 4, 5, ...\} = 1$. $\boxed{3,\ \text{A}}$

(iv) To find the stationary distribution, we solve the system of equations $\pi P = \pi$ and $\sum_{i=1}^{5} \pi_i = 1$,

$$\begin{pmatrix} \pi_1 & \pi_2 & \pi_3 & \pi_4 & \pi_5 \end{pmatrix} \begin{pmatrix} 0.5 & 0.5 & 0 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \pi_1 & \pi_2 & \pi_3 & \pi_4 & \pi_5 \end{pmatrix}$$

Hence,

$$\begin{aligned} 0.5\pi_1 + 0.5\pi_2 &= \pi_1 \\ 0.5\pi_1 + 0.5\pi_3 &= \pi_2 \\ 0.5\pi_2 + 0.5\pi_4 &= \pi_3 \\ 0.5\pi_3 &= \pi_4 \\ 0.5\pi_4 + \pi_5 &= \pi_5 \end{aligned} \qquad \Longrightarrow \qquad \begin{aligned} \pi_1 &= \pi_2 \\ \pi_3 &= 2\pi_2 - \pi_1 \\ \pi_4 &= 2\pi_3 - \pi_2 \\ \pi_3 &= 2\pi_4 \\ \pi_4 &= 0 \end{aligned}$$

From these equations, we see that $\pi_1 = \pi_2 = \pi_3 = \pi_4$ and since $\pi_4 = 0$, we have $\pi_i = 0$ for $i = 1, 2, 3, 4$. Then since $\sum_{i=1}^{5} \pi_1 = 1$, it follows that $\pi_5 = 1$. Therefore the stationary distribution is $\begin{pmatrix} 0 & 0 & 0 & 0 & 1 \end{pmatrix}$. [If the student immediately writes down the stationary distribution, they need to justify their answer and show $\pi P = \pi$ and that π is a distribution for full marks] $\boxed{4,\ \text{B}}$

(v) We know that once we reach state 5, we never leave it. Therefore, we need to calculate the probability of hitting state 5 from state 1, h_1^5. To do this, we know from lecture notes that we need to solve the following system of equations,

$$
\begin{aligned}
h_1^5 &= \sum_{j=1}^{5} p_{1j} h_j^5 = 0.5 h_1^5 + 0.5 h_2^5 & & & h_1^5 &= h_2^5 \\
h_2^5 &= \sum_{j=1}^{5} p_{2j} h_j^5 = 0.5 h_1^5 + 0.5 h_3^5 & & & h_2^5 &= h_3^5 \\
h_3^5 &= \sum_{j=1}^{5} p_{3j} h_j^5 = 0.5 h_2^5 + 0.5 h_4^5 & & \implies & h_3^5 &= h_4^5 \\
h_4^5 &= \sum_{j=1}^{5} p_{4j} h_j^5 = 0.5 h_3^5 + 0.5 h_5^5 & & & 0.5 h_4^5 &= 0.5 \\
h_5^5 &= 1 & & & h_5^5 &= 1
\end{aligned}
$$

Therefore we have $h_i^5 = 1$ for all $i = 1, \ldots, 5$.

unseen ⇓

2, B

(c) (i) We need to show $\sum_{j \in \mathcal{E}} Q_{ij} = 1$. Note that $\sum_{j \in \mathcal{E}} Q_{ij} = \sum_{j \in \mathcal{E}, j \neq i} (1 - P_{ii})^{-1} P_{ij} = (1 - P_{ii})^{-1} \sum_{j \in \mathcal{E}, j \neq i} P_{ij} = (1 - P_{ii})^{-1}(1 - P_{ii}) = 1$ since P is a valid transition matrix so $\sum_{j \in \mathcal{E}} P_{ij} = 1$ for all $j \in \mathcal{E}$. Moreover, all $Q_{ij} \geq 0$ since $P_{ii} < 1$ since there is positive probability of going from state i to any other state, and $P_{ij} \geq 0$.

unseen ⇓

4, D

(ii) From lecture notes, we know that irreducibility is a structural property of the chain. When going from P to Q, we observe that we are essentially removing any self-loops from the transition matrix. Therefore a lot of the structure remains the same. In particular, for any $i \neq j$, if $P_{ij} > 0$, then $Q_{i,j} = (1 - P_{ii})^{-1} P_{ii} > 0$, so the graph of the Markov chain will remain the same except for self-loops. Hence, if P is irreducible, there will exist a path from i to $j \neq i$ in n steps, where we assume n is minimal (i.e. we remove any self-loops from this path). All transitions on this path must go from one state to a different one and have positive probability under P, so they will also have positive probability under Q. Hence, there must also exist a path from i to $j \neq i$ in n steps under Q. All that remains is to show that there exists a path from i to i with positive probability under Q. By the above argument, we know that there exists a path from i to some $j \neq i$ in n steps and a path from j to i in m steps both of which have positive probability under Q. Hence, there exists a path from i to i (via j) in $n + m$ steps with positive probability under Q. Hence Q is irreducible when P is irreducible.

Review of mark distribution:
Total A marks: 32 of 32 marks
Total B marks: 20 of 20 marks
Total C marks: 12 of 12 marks
Total D marks: 16 of 16 marks
Total marks: 80 of 80 marks
Total Mastery marks: 0 of 20 marks

MATH50010 Probability for Statistics

Question	Marker's comment
1	a) generally answered wellb) Generally answered very wellc) Mostly answered well, although some students struggled to show closure under complements, and others misunderstood the definition of G and started looking at subsetsd) Mixed responses, most students got one definition but not all got bothe) Often not attempted, when attempted some students only considered one possible value of the sum rather than taking a union over all possible valuesfi) Mixed responses, generally students either proved that S_n was a random variable for any n or that S_N was a random variable for random N, but very few did both which was necessary for full marksfii) mixed responses. Several students wrote expressions along the lines of N=E[N] which is incorrect.
2	a) generally answered well, although some quantifiers missingbi) mostly answered well, although there were some attempts to apply the weak law of large numbers (we do not have independence so cannot do this) and some incorrect reasoning about probabilities of sums.nbsp;bii) Many students realised we need \lambda_n \to \infty, but not as many recognised that we also need independence.biii) Generally not answered well. Method marks were given for an application of Borel-Cantelli or the continuity property.ci) Generally answered well.nbsp;cii) mixed responses, several students got full marks, but others made mistakes in their use of the cdf in part (i). Another common answer was to attempt to use moment generating functions, but this was hindered by not knowing the mgf of a Gumbell random variable.
3	3.a) done well by those who attempted it, but attempted by a smaller proportion of students than expected. b) mostly done well by all students. Common mistake to show Cov(X, Y) = 0 and state that this implies independence, rather than showing that the joint pdf factorises.c) done well by most students. Most mistakes here were alegbraic errors. d) most students who attempted this question recognised the need for monotonic transformation but this was well justified only rarely. Only a handful of students recognised that the condtional distribution is Gaussian, with most leaving f_Z|Y as the ratio of joint and marginal distributions. e) i) not attempted by a large number of students but generally done well by those who did. ii) not attempted by a large number of students. Most of those who attempted this noticed that taking logs might be helpful but very few proceeded further. Only a couple of students made the conection to the probability integral transform and central limit theroem to get the final probablilty.
4	The students overall did well on this question with c(ii) being the toughest where the key was spotting that P and Q essentially had the same structure except for self-loops. A few also struggled to find the periodicity in b(iii) where the self-loop in state 1 means the periodicity has to be 1. The other tough question for some was b(v) where some students tried to use alternative methods to find the absorption probabilities into state 5 but these were not always rigorous or correct - a rigorous method is provided in the answers as also taught in the course.Students overall did well at the remaining questions: (a) defining an irreducible set of states, b(i) finding the transition matrix, b(ii)nbsp;the properties of different states, and b(iv) the stationary distribution - well done here!