1. (6 Marks) A Markov chain has transition matrix

$$\begin{pmatrix} 1/2 & 0 & 0 & 0 & 1/2 \\ 0 & 9/10 & 0 & 1/10 & 0 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 0 & 3/10 & 0 & 7/10 & 0 \\ 3/4 & 0 & 0 & 0 & 1/4 \end{pmatrix}.$$

Analyse the state space (reducibility, periodicity, recurrence, etc), and discuss the chain's long run behavior.

Solution: A quick look at the transition diagram shows the chain has two essential communicating classes and one non-essential class:

$$S_1^e = \{1, 5\}, S_2^e = \{2, 4\}, \quad \text{and} \quad S_1^n = \{3\}.$$

So the chain is reducible and both essential communicating classes are aperiodic because of the loops and positive recurrent since they are finite sets. For the long run behaviour, if the chain starts in state 3 it ends up in S_1^e or S_2^e with even chance due to symmetry and once in one of these essential communicating classes, the chain restricted to this class is ergodic.
In either case the long run chance the chain is in a state is given by the relevant stationary distribution. For S_1^e, the stationary distribution π satisfies

$$\pi \begin{pmatrix} 1/2 & 1/2 \\ 3/4 & 1/4 \end{pmatrix} = \pi$$

and $\pi_1 = 1 - \pi_2$. Solving yields

$$\pi = (3/5, 2/5).$$

Similarly, conditional on the chain being in S_2^e, the long run probabilities for states 2 and 4 are given by

$$(3/4, 1/4).$$

2. (12 Marks) The chance that it rains in Melbourne given that it rained both of the past two days is 3/4. If it rained yesterday but not the day before the chance it rains today is 2/3 and if it rained two days ago but not yesterday then the chance it rains today is 1/2. Finally, if it hasn't rained in the past two days then the chance it rains today is 3/7.

    (a) Model the rain in Melbourne as a Markov chain and write down its transition matrix.

(b) Analyse the state space of this Markov chain and discuss its long run behaviour.

(c) Starting from a randomly chosen day of the year, what is the chance of having two consecutive days without rain before there are two consecutive rainy days?

(d) Given it rained the last two days, what is the expected number of days before it rains twice in a row again? (If it rains tomorrow then it took one day to rain twice in a row again.)

Solution: (a) We can model the two day consecutive rain/no rain with the states $1 = RR, 2 = RN, 3 = NR, 4 = NN$ and we have transition matrix

$$P = \begin{pmatrix} a & 1-a & 0 & 0 \\ 0 & 0 & b & 1-b \\ c & 1-c & 0 & 0 \\ 0 & 0 & d & 1-d \end{pmatrix},$$

where $a = 3/4, b = 1/2, c = 2/3, d = 3/7$.

(b) This is a finite Markov chain with one communicating class and loops so its aperiodic, irreducible and positive recurrent with ergodic distribution π satisfying $\pi P = \pi$:

$$\pi = \frac{1}{cd + 2d(1-a) + (1-a)(1-b)} \left(cd, d(1-a), d(1-a), (1-a)(1-b)\right).$$

Plugging in the values above we have $\pi = (16/35, 6/35, 6/35, 7/35)$.

(c) The rain on a randomly chosen day of the year and the day before should roughly follow the stationary distribution π. If A_i is the chance of two consecutive days without rain before two consecutive rainy days given the last two days had rain as i, then the probability we want is

$$\sum_{i=1}^{4} \pi_i A_i.$$

First step analysis gives

$$\begin{aligned} A_1 &= (1-a)A_2 \\ A_2 &= bA_3 + (1-b) \\ A_3 &= (1-c)A_2 \\ A_4 &= dA_3 + (1-d). \end{aligned}$$

And solving we have

$$
\begin{aligned}
A_1 &= \frac{(1-a)(1-b)}{1-b(1-c)} \\
A_2 &= \frac{(1-b)}{1-b(1-c)} \\
A_3 &= \frac{(1-c)(1-b)}{1-b(1-c)} \\
A_4 &= \frac{d(1-c)(1-b)}{1-b(1-c)} + (1-d).
\end{aligned}
$$

Plugging in the values of the problem we have $A_1 = 3/20, A_2 = 12/20, A_3 = 4/20, A_4 = 23/35$ and mixing over π the probability we want is $2/7$.

(d) The expected return time to state i started from state i is just $1/\pi_i$.

3. (6 Marks) A certain casino game costs 180 dollars to play. The game has two rounds. In the first round a coin is tossed and if it comes up heads you get 280 dollars (that is, 100 dollars plus your 180 dollar fee is returned) and if it comes up tails you get 90 dollars (so overall you're down 90 dollars). At this point you may take your money or you can play one more round where again with even chance you gain 100 dollars or your money is halved (so after playing the second round you'll leave with one of 380, 190, 140, or 45 dollars). What strategy should you use to maximise your expected winnings in this game and what is that expectation? (Note that one potential strategy is not to play at all.)

Solution:
In general if you begin with x dollars and at each stage you either gain w dollars or your money is multiplied by a factor $0 < b < 1$, with respective probabilities p and $1-p$, then we can think of this as a Markov reward process. At the final stage your reward is what's on the table. At the second stage of the process, if you have y dollars on the table, your potential winnings are

$$V_1(y) := \max\{y, p(y+w) + (1-p)by\},$$

and you want to take the money if

$$y \geq p(y+w) + (1-p)by \iff y \geq \frac{pw}{(1-p)(1-b)}.$$

Initially your potential winnings are

$$V_2(x) := \max\{x, pV_1(w+x) + (1-p)V_1(bx)\}$$

and again you should not play if x is this maximum. With the numbers above, $p = b = 1/2$, $w = 100$ and $x = 180$, we have that $V_1(y) = y$ if $y \geq 200$, so

$$V_1(280) = 280, \text{ and } V_1(90) = 235/2$$

and the optimal strategy says to walk away from the table if you win the first round but to play on if you lose the first round with expected winnings of 117.5 dollars.
In the first round, your expected winnings by playing are

$$(V_1(280) + V_1(90))/2 = 198.75 > 180.$$

So you should play the game and take the money off the table in the first round if you win and otherwise play on. With this strategy your expected winnings will be 18.75 dollars.

4. (14 Marks) A yeast culture contains microbes that split and die according to the following rules. Each microbe lives a number of minutes having an exponential distribution with rate 1 and then splits into two microbes with probability p or dies with probability $1-p$. The lifetime and choice of split or death of a microbe is independent of that of all other microbes. In addition, independent of the process in the culture, yeast microbes from the air outside of the culture float by the culture according to a Poisson process with rate 2 per minute. Each microbe that floats by joins the population of the culture with probability p and with probability $1-p$ the microbe doesn't join the culture, this choice made independent of all else.

   (a) What is the chance that exactly four outside microbes float by in the first 3 minutes?

   (b) What is the chance that exactly four outside microbes join the culture in the first 3 minutes?

   (c) Given that 7 outside microbes have floated by the culture in first 3 minutes, what is the chance that at least two of the seven join the culture?

   (d) Given that 7 outside microbes have floated by the culture in first 3 minutes, what is the chance that exactly 3 float by in the first 1 minute?

   (e) What is the chance that in the first 3 minutes, exactly four microbes join the culture and 3 float by that don't join the culture?

   (f) Model the number of microbes in the culture as a continuous time Markov chain and define its generator.

   (g) For what values of p is the chain ergodic? Find the limiting distribution in the cases where it's ergodic.

Solution: Let N_t be the number of microbes that float by up to time to t and let M_t be the number that join the colony up to to time t. Then N_t is Poisson mean $2t$ and M_t is Poisson mean $2pt$, independent of the process $N_t - M_t$ which is Poisson mean $2(1-p)t$. Also conditional on a Poisson process equals k at time t, the distribution of the k points in the interval $(0,t)$ are the same as k iid variables that are uniform on $(0,t)$. These facts and the description of the process imply we have the following answers:

(a) $e^{-6}\frac{6^4}{4!}$ (N_t Poisson process),

(b) $e^{-6p}\frac{(6p)^4}{4!}$ (M_t Poisson process),

(c) $1 - 7p(1-p)^6 - (1-p)^7$ (description of the M_t from N_t),

(d) $\binom{7}{3}(1/3)^3(2/3)^4$ (conditional Poisson process description),

(e) $e^{-6}(6p)^4(6(1-p))^3/(4!3!)$ (Poisson and independence of M_t and $N_t - M_t$),

(f) The process is a birth-death process. If there are k microbes then the rate of births from splits is kp and the rate of births from outside immigration is $2p$. The rate of death is $k(1-p)$ so the generator A has entries $-a_{00} = a_{01} = 2p$ and for $k \geq 0$

$$a_{kk+1} = (k+2)p =: \lambda_k, \;\; a_{kk-1} = k(1-p) =: \mu_k \;\; a_{kk} = -(k+2p),$$

with zero entries elsewhere.

(g) A birth death chain as above has stationary distribution if and only if

$$\sum_{j\geq 0}\prod_{k=1}^{j}\frac{\lambda_{k-1}}{\mu_k} < \infty.$$

And here we have

$$\prod_{k=1}^{j}\frac{\lambda_{k-1}}{\mu_k} = (j+1)\left(\frac{p}{1-p}\right)^j,$$

which is summable in j if and only if $p/(1-p) < 1$ or $0 \leq p < 1/2$. And in this case, the stationary distribution is negative binomial with parameters 2 and $1/(1-p)$ having probability mass at k of

$$(k+1)\left(\frac{p}{1-p}\right)^k\left(\frac{1}{1-p}\right)^2$$

5. (12 Marks) Recall in the $M/M/2$ queue with arrival rate λ per hour and service rate μ per hour with $\lambda < \mu$ and $\rho := \lambda/\mu$, the stationary number of customers in the system has distribution

$$\pi_j = \left(\frac{2-\rho}{2+\rho}\right) \times \begin{cases} \rho^j/j! & \text{for } j = 0, 1, \\ \rho^j/2^{j-1} & \text{for } j \geq 2. \end{cases}$$

And for the same rates in the $M/M/1$ queue the analogous stationary distribution is geometric:

$$\sigma_j = (1-\rho)\rho^j \quad \text{for } j \geq 0.$$

(a) Derive the stationary expected number of customers in the queue for both the $M/M/2$ and the $M/M/1$ queues.

(b) Derive the expected waiting time of an arriving customer for these two queues in their stationary regimes.

(c) Derive the stationary expected number of busy servers in the queue for both the $M/M/2$ and the $M/M/1$ queues.

(d) A company that shreds documents is trying to determine whether to buy one or two shredders. Documents to be shredded arrive in pallets according to a Poisson process with rate 1 per hour. The number of hours it takes a shredder to shred a pallet of paper has an exponential distribution with rate 2 per hour. The cost of storing a pallet of paper is 16 dollars per pallet per hour, the cost of maintaing a shredder is 7 dollars per hour and when a shredder is in use there is an additional cost of electricity of 3 dollars per hour. Over the long run will the company's expenses be lower with one shredder or two?

(e) How would you answer the previous question (d) if the cost of storing a pallet of paper is 14 dollars per pallet per hour, with all other variables the same?

Solution: (a) If the number of customers in the system in stationary is X then the number of customers in the $M/M/a$ queue is $\max\{X-a, 0\}$. Taking the expected value of these functions against the appropriate stationary distribution we have

$$\begin{aligned}
\textbf{M/M/1:}\ Q_1 &= \sum_{k\geq 1}(k-1)\sigma_k = \frac{\rho}{1-\rho} - \rho = \frac{\rho^2}{1-\rho}.\\
\textbf{M/M/2:}\ Q_2 &= \sum_{k\geq 3}(k-2)\pi_k\\
&= \left(\frac{2-\rho}{2+\rho}\right)\sum_{k\geq 3}(k-2)\rho^k/2^{k-1}\\
&= \left(\frac{(2-\rho)\rho^3}{4(2+\rho)}\right)\sum_{k\geq 1}(k-2)\rho^{k-3}/2^{k-3}\\
&= \left(\frac{(2-\rho)\rho^3}{4(2+\rho)}\right)\frac{4}{(2-\rho)^2} = \frac{\rho^3}{4-\rho^2}.
\end{aligned}$$

(b) Little's law says $Q_i = \lambda \mathbb{E} W_i$ where W_i is the stationary waiting time of the $M/M/i$ queue.

(c) In the $M/M/a$ queue, if X is the number of customers in stationary, the number of busy servers $\min\{X, a\}$. Taking expectation against the appropriate stationary distributions:

$$\textbf{M/M/1:}\ B_1 = \sum_{k\geq 1} \sigma_k = \rho.$$

$$\textbf{M/M/2:}\ B_2 = \pi_1 + \sum_{k\geq 2} 2\pi_k = 2 - 2\pi_0 - \pi_1 = \rho.$$

(d) If there are $a = 1, 2$ shredders then the number of pallets in the system is an $M/M/a$ queue with $\lambda = 1$ and $\mu = 2$. If C_s is the cost of storing in dollars per pallet per hour, and C_m is the cost of maintaining a shredder in dollars per hour and C_e is the cost of electricity of a shredder in use in dollar per hour, then the per hour cost of shredding is given by

$$Q_a C_s + aC_m + B_a C_e.$$

When $\rho = 1/2$ this quantity is equal to

$$\textbf{a=1:}\ (1/2)16 + (1)7 + (1/2)3 = 16.5$$
$$\textbf{a=2:}\ (1/30)16 + (2)7 + (1/2)3 = 15.5 + 8/15 < 16.5,$$

so it's cheaper in the long run to have 2 shredders.

(e) Using the same formula above but changing C_s to 14 we have

$$\textbf{a=1:}\ (1/2)14 + (1)7 + (1/2)3 = 15.5$$
$$\textbf{a=2:}\ (1/30)14 + (2)7 + (1/2)3 = 15.5 + 7/15 > 15..5,$$

so now it's cheaper in the long run to have 1 shredder.

6. (12 Marks) The lifetime in hours of a saw blade in a lumber mill is a random variable with density proportional to $x(2 - x)$ on $0 < x < 2$. When the mill opened a new blade was put in the saw. Every time a blade fails it is immediately replaced with a new one. Since the mill is largely automated, it operates 24 hours a day, seven days a week, year round.

   (a) Model the number N_t of saw blades that have been replaced t hours after the mill opened as a renewal process and determine the density, mean, and variance of the random times between renewals.

   (b) About how many replacement blades for the saw will the mill need within the first 30 days of opening?

   (c) Give an interval around your estimate from (b) that will have a 95% chance of covering the true number of saws needed in the first 30 days.

(d) If you show up at the mill at the end of day 30, about what is the mean and variance of the amount of time you'll have to wait for the current blade in use to be replaced?

Solution: (a) If $X_1, \ldots$ are iid with distribution the lifetime of the saw then

$$N_t = \max \left\{ n : \sum_{i=1}^{n} X_i \leq t \right\}$$

defines a renewal process with inter-arrival times X_i. The distribution of these replacement times is $2B[2,2]$, where $B[2,2]$ has the beta distribution. So its density is

$$\frac{3}{4} x(2-x) \quad \text{for} \quad 0 < x < 2,$$

with mean $\mu = 1$ and variance $\sigma^2 = 1/5$.

(b) Renewal theory says that for large t, $N_t/t \approx 1/\mu = 1$; here $t = 30 * 24 = 720$ so we expect to need about that many saws up through hour 720.

(c) The fluctuations around the mean are given by the renewal CLT which states that for large t

$$\frac{N_t - t/\mu}{\sqrt{t\sigma^2/\mu^3}}$$

is roughly standard normal. So the fluctuations around our estimate for $t = 720$ and the parameters above are roughly normal with SD=12 and so the 95% CI is

$$720 \pm 24.$$

(d) Defining $T_k = \sum_{i=1}^{k} X_i$, we know that for large t and $Y_t := T_{N_t+1} - t$ roughly has density $(1 - F(x))/\mu$ for $0 < x < 2$. In this case this density on this support is

$$1 - 3x^2/4 + x^3/4,$$

and now its only a matter of computing the first and second moments to find the mean is $3/5$ and the variance is $13/75$. To check, we know the asymptotic distribution is UX_1^s, where U is uniform on $(0,1)$ and independent of X^s, the size-bias distribution of X. So the kth moments are

$$\frac{\mathbb{E}X_1^{k+1}}{\mu(k+1)}$$

and when $k = 1$ this is $3/5$ and when $k = 2$ this is $8/15$.

7. (8 Marks) Denote the transition probabilities for a time homogenous Markov chain on $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ by $p_{ij} := \mathbb{P}(X_1 = j | X_0 = i)$. Let $0 < p < 1$ and assume that for each i in $\mathbb{Z}$,

$$p_{ii+2} = p, \; p_{ii-1} = 1 - p \;\text{ and } p_{ij} = 0 \;\text{ for all } j \neq i+2, i-1.$$

For each state i in $\mathbb{Z}$, determine the values of p where i is recurrent and the values where i is transient. You may find it helpful to use Stirling's approximation:

$$1 \leq \frac{n! e^n}{n^n \sqrt{2\pi n}} \leq 2.$$

Solution:
The chain is irreducible and recurrence is a class property so we only need to determine recurrence and transience at 0, say. From lecture, state 0 is transient if and only if the expected number of returns to the origin is finite which decomposing into indicators is the same as

$$\sum_{n \geq 1} p_{00}(n) < \infty,$$

where $p_{ij}(n)$ are the n-step transition probabilities.

We can determine $p_{00}(n)$ as follows. Since the chain either increases by two or decrease by one at every step, if the chain takes k up steps and m down steps then at time $k+m$ the chain started from zero is at $k - 2m$. So in order for this quantity to be zero $k = 2m$ and the number of steps taken by the chain is $3m$. So the chain can only return to zero in multiples of 3 steps. Further, any path taking the chain back to zero in $3n$ steps must have n up steps and $2n$ down steps and there are $\binom{3n}{n}$ such paths (choosing the times of the up steps). Each of these paths has probability $p^{2n} q^n$ where $q = 1 - p$ and so:

$$p_{00}(3n) = \binom{3n}{n} p^{2n} q^n,$$

and

$$\sum_{n \geq 1} p_{00}(n) = \sum_{n \geq 1} p_{00}(3n). \tag{1}$$

To determine whether this series is summable, use Stirling's approximation to find

$$\frac{1}{4}\left[\frac{\sqrt{3}}{2\sqrt{\pi n}}\left(\frac{27}{4}\right)^n\right](p^2 q)^n \leq p_{00}(3n) \leq 2\left[\frac{\sqrt{3}}{2\sqrt{\pi n}}\left(\frac{27}{4}\right)^n\right](p^2 q)^n.$$

So the series (1) is summable if and only if the series

$$\sum_{n\geq 1} \frac{1}{\sqrt{n}} \left(\frac{27p^2(1-p)}{4} \right)^n,$$

is summable which occurs if and only if

$$27p^2(1-p) < 4 \iff p \neq 2/3.$$

So the chain is transient if $p \neq 2/3$ and is recurrent otherwise.