Mast30001-2019-sols

Stochastic Modelling (University of Melbourne)

THE UNIVERSITY OF MELBOURNE

Student Number

Semester 2 Assessment, 2019

School of Mathematics and Statistics

MAST30001 Stochastic Modelling

Writing time: 3 hours

Reading time: 15 minutes

This is NOT an open book exam

This paper consists of 9 pages (including this page)

Authorised Materials

- Mobile phones, smart watches and internet or communication devices are forbidden.
- Students may bring one double-sided A4 sheet of handwritten notes into the exam room.
- Hand-held electronic scientific (but not graphing) calculators may be used.

Instructions to Students

- You must NOT remove this question paper at the conclusion of the examination.
- This paper has **6 questions**. Attempt as many questions, or parts of questions, as you can. The number of marks allocated to each question is shown in the brackets after the question statement. There are **80 total marks** available for this examination. A table of **normal distribution probabilities** can be found at the end of the exam. Working and/or reasoning must be given to obtain full credit. Clarity, neatness and style count.

Instructions to Invigilators

- Students must NOT remove this question paper at the conclusion of the examination.

This paper may be held in the Baillieu Library

1. A Markov chain $(X_n)_{n\geq 0}$ with state space $S = \{1, 2, 3, 4, 5\}$ has transition matrix

$$P = \begin{pmatrix} 1/2 & 1/4 & 0 & 1/4 & 0 \\ 0 & 2/5 & 0 & 3/5 & 0 \\ 0 & 1/5 & 0 & 4/5 & 0 \\ 0 & 0 & 1/3 & 0 & 2/3 \\ 0 & 0 & 1/2 & 0 & 1/2 \end{pmatrix}.$$

(a) Find $\mathbb{P}(X_4 = 2, X_2 = 3 | X_0 = 2)$.

(b) If the initial distribution is uniform on $\{1, 2\}$, find $\mathbb{P}(X_4 = 2, X_2 = 3)$.

(c) Write down the communication classes of the chain.

(d) Find the period of each communicating class.

(e) Determine which classes are essential.

(f) Classify each essential communicating class as transient or positive recurrent or null recurrent.

(g) Describe the long run behaviour of the chain (including deriving long run probabilities where appropriate).

(h) Find the probability of reaching state 3 before state 5 given the chain starts in state 1.

[16 marks]

Ans.

(a) (2 marks)

$$p^{(2)}_{23} p^{(2)}_{32} = \left(\frac{1}{5}\right)\left(\frac{2}{25}\right) = \frac{2}{125} = 0.016.$$

(b) (2 marks)

$$\frac{1}{2}\left(p^{(2)}_{23} p^{(2)}_{32} + p^{(2)}_{13} p^{(2)}_{32}\right) = \frac{1}{2}\left(\frac{2}{125} + \left(\frac{1}{12}\right)\left(\frac{2}{25}\right)\right) = \frac{17}{1500} = 0.01133.$$

(c) (1.5 marks) Inspection of a transition diagram gives $S_1 = \{1\}$ and $S_2 = \{2, 3, 4, 5\}$.

(d) (1.5 marks) Both classes have a loop and so have period 1.

(e) (1.5 marks) S_1 is non-essential and S_2 is essential.

(f) (1.5 marks) The essential communicating class is positive recurrent since it is finite.

(g) (3 marks) The chain will eventually end up in the essential communicating class S_2 and since it is aperiodic and positive recurrent, it is ergodic. The long run probabilities are given by the stationary distribution solving $\pi P = \pi$, which is easily found to be

$$\pi = \frac{1}{11}(0, 1, 3, 3, 4).$$

(h) (3 marks) We use first step analysis. Let A be the event "reach state 3 before state 5" and $f_i = P(A|X_0 = i)$. We then have

$$f_1 = \frac{1}{2}f_1 + \frac{1}{4}f_2 + \frac{1}{4}f_4,$$
$$f_2 = \frac{2}{5}f_2 + \frac{3}{5}f_4,$$
$$f_4 = \frac{1}{3},$$

and solving yields $f_1 = 1/3$. Alternatively, note from state 1, the only way to enter states 3 or 5 is through state 4, and $p_{43} = 1/3$.

2. A renewal process $(N_t)_{t\geq 0}$ has inter-renewal distribution uniform on the interval $(3, 5)$.

   (a) Compute the mean and variance of the inter-renewal distribution.

   (b) On average, about how many renewals are there in the interval $(0, 1000)$?

   (c) Give a symmetric interval around your estimate from (b) that will have a 95% chance of covering the true number of renewals.

   (d) If T_k denotes the time of the kth renewal, $k = 1, 2, \ldots$, what would you estimate to be the mean of $\left(T_{N_{1000}+1} - T_{N_{1000}}\right)$?

   [12 marks]

Ans. We model the system as a renewal process with inter-arrival distribution as above.

(a) (2 marks) Straightforward computing give

$$\mu := E[\tau] = 4, \quad \sigma^2 := Var(\tau) = 1/3.$$

(b) (3 marks) The renewal LLN says $N_t/t \to 1/E[\tau] = 1/4$ as $t \to \infty$, and so we expect $N_{1000} \approx 1000/4 = 250$.

(c) (3 marks) The renewal CLT says that

$$N_{1000} \approx \text{Normal}(1000/\mu, 1000\sigma^2/\mu^3) = \text{Normal}(250, 1000/192),$$

and so there is approximately a 95% chance that the number of tasks the worker completes will fall in the interval

$$250 \pm (1.96)\sqrt{1000/192} \approx 250 \pm 4.473.$$

(d) (4 marks) We know that for large t, $\left(T_{N_t+1} - T_{N_t}\right)$ has approximate the size-bias distribution of τ, here having density

$$\frac{x f_\tau(x)}{\mu} = \frac{x}{8}, \;\; 3 < x < 5,$$

and thus mean

$$\frac{1}{8}\int_3^5 x^2 dx = \frac{49}{12} = 4.0833.$$

Alternatively, $\left(T_{N_t+1} - T_{N_t}\right) = Y_t + A_t$ where Y_t, A_t are the residual and age, equal in distribution, each having density on $x \geq 0$

$$\frac{1 - F_\tau(x)}{\mu} = \begin{cases} \frac{1}{4}, & 0 < x < 3, \\ \frac{1-(x-3)/2}{4}, & 3 < x < 5, \end{cases}$$

and each having mean

$$\frac{1}{4}\int_0^3 x dx + \frac{1}{4}\int_3^5 x(1 - (x-3)/2) dx = \frac{9}{8} + \frac{11}{12} = \frac{49}{24}.$$

Thus summing the two means gives $49/12$ as above.

3. Customers arrive at a shop according to Poisson process $(N_t)_{t\geq 0}$ with rate 10 per hour. Each customer independently spends an amount, rounded to the nearest dollar, that is distributed as a geometric random variable X with probability mass function

$$\mathbb{P}(X = k) = (9/10)^k (1/10), \;\; k = 0, 1, 2, \ldots$$

(a) What is the chance that no customers arrive in a given half hour?

(b) What is the probability a given customer doesn't make a purchase (spends 0 dollars)?

(c) What is the chance that in a given half hour, no customers enter the shop without making a purchase?

(d) Given that 10 customers have arrived in a given hour, what is the expected number that arrived in the first half hour of that hour?

(e) What is the probability that in a given half hour, exactly 3 customers spend 0 dollars, exactly 2 customers spend $1 - 20$ dollars (inclusive), and the remaining customers spend greater than 20 dollars?

(f) What is the mean and variance of the geometric random variable X?

(g) What is the mean and variance of the revenue for the shop over the course of 8 hours?

[18 marks]

Ans. Below, let $(M_t)_{t\geq 0}$ be the number of customers spending 0 dollars, and $(L_t)_{t\geq 0}$ be the number of customers spending between 1 and 20 dollars. Then the thinning theorem says $(M_t)_{t\geq 0}$ is a Poisson process rate $10/10 = 1$ independent of $(L_t)_{t\geq 0}$, which is a Poisson process rate $10(9/10 - (9/10)^{20}) = 9(1 - (9/10)^{19})$.

(a) (2 marks) $N_{1/2}$ is Poisson distributed with mean 5, so

$$P(N_{1/2} = 0) = e^{-5}.$$

(b) (1 mark) The chance a geometric is equal to zero is its parameter, here 1/10.

(c) (3 marks, some for thinning statement above) This is the same as

$$P(M_{1/2} = 0) = e^{-1/2}.$$

(d) (3 marks) Given the number of arrivals in an interval, their positions are i.i.d. uniformly distributed. Thus the number of the 10 that arrived in the hour that fall in the first half hour is binomially distributed with parameters 10 and 1/2, with mean 5.

(e) (3 marks, some for thinning statement above) This is the same as, for $\lambda = 9(1 - (9/10)^{19})$,

$$P(M_{1/2} = 3, L_{1/2} = 2) = P(M_{1/2} = 3)P(L_{1/2} = 2) = \frac{e^{-1/2}(1/2)^3}{3!}\frac{e^{-\lambda/2}(\lambda/2)^2}{2!}.$$

(f) (2 marks) If X is geometric with parameter 1/10, then the mean is

$$E[X] = 9,$$

and the variance is

$$Var(X) = 90.$$

(g) (4 marks) If $X_1, X_2, \ldots,$ are i.i.d. geometric with parameter 1/10, then the revenue of the shop is distributed as

$$Y = \sum_{i=1}^{N_8} X_i,$$

with mean

$$E[Y] = E[N_8]E[X_1] = (80)(9) = 720,$$

and variance

$$Var(Y) = E[N_8]Var(X_1) + E[X_1]^2 Var(N_8) = 80(90 + 81) = 80 \cdot 171.$$

4. In a certain queuing system, jobs arrive according to a Poisson process with rate 4. When jobs arrive, they have to go through two servers in sequence (meaning a job gets served by the first server and then gets served or queues for the second server). Service times are exponential, and the first server works at rate 2, and the second at rate 1. Arriving jobs are turned away if either the first server is working, or there are 2 jobs in the system.

   (a) Model the number of jobs in the system (including those being served) as a continuous time Markov chain $(X_t)_{t\geq 0}$ with appropriate state space, and specify its generator.

   (b) Find the stationary distribution of the Markov chain.

   (c) What proportion of time is the second server idle?

   (d) What is the average number of jobs in the system?

   (e) What is the probability an arriving job is turned away?

   (f) Given a job is not turned away, what is the average amount of time it spends waiting for service?

   (g) Given a job is not turned away, what is the average amount of time it spends in the system?

[17 marks]

Ans.

(a) (3 marks) We view the system as a CTMC with states $\{0, (1,0), (0,1), (1,1), (0,2)\}$, where (i,j) means there are $i = 0, 1$ jobs in the first service and $j = 0, 1, 2$ is the number of jobs in the service and queue for the second server. The generator is

$$A = \begin{array}{c} \begin{array}{ccccc} 0 & (1,0) & (0,1) & (1,1) & (0,2) \end{array} \\ \begin{pmatrix} -4 & 4 & 0 & 0 & 0 \\ 0 & -2 & 2 & 0 & 0 \\ 1 & 0 & -5 & 4 & 0 \\ 0 & 1 & 0 & -3 & 2 \\ 0 & 0 & 1 & 0 & -1 \end{pmatrix} \end{array}.$$

(b) (3 marks) To determine the stationary distribution of this system, we solve $\pi A = 0$ and find

$$\pi = (\pi_0, \pi_{(1,0)}, \pi_{(0,1)}, \pi_{(1,1)}, \pi_{(0,2)}) = \frac{1}{77}(3, 14, 12, 16, 32).$$

(c) (2 marks) From the work above, the second server is idle in states $0, (1,0)$, in which the system spends a proportion of time equal to

$$\pi_0 + \pi_{(1,0)} = \frac{17}{77} = 0.220779.$$

(d) (2 marks) The average number of jobs in the system is

$$L = \pi_{(0,1)} + \pi_{(1,0)} + 2(\pi_{(0,2)} + \pi_{(1,1)}) = \frac{122}{77} = 1.5844.$$

(e) (2 marks) Jobs are turned away when the system is in states $(1,0), (1,1), (2,0)$, and PASTA implies the proportion of jobs finding the system in those states is

$$\pi_{(1,0)} + \pi_{(1,1)} + \pi_{(2,0)} = \frac{62}{77} = 0.8052.$$

(f) (3 marks) A job is not turned away if it finds the system in states $0, (0,1)$. In the first case, the job does not wait for service. In the second case, the job waits for service only if the service of the job in the second server takes longer than its service in the first server, which happens with probability $2/3$. Given this has occurred, the job must wait for the second server to finish, which, using the memoryless property of the exponential, takes on average 1 time unit. Altogether, the average waiting time given a job is not turned away is

$$\frac{\pi_{(0,1)}}{\pi_0 + \pi_{(0,1)}}\left(\frac{2}{3}\right) = \frac{8}{15} = 0.5333.$$

(g) (2 marks) The total time in the system is the waiting time plus service time, which on average is

$$\frac{1}{2} + \frac{8}{15} + 1 = \frac{61}{30} = 2.03333.$$

Alternatively, notice that the total rate of entering is

$$\tilde{\lambda} = 4(\pi_0 + \pi_{(0,1)}) = \frac{60}{77},$$

and then Little's law implies

$$D = L/\tilde{\lambda},$$

where $L = 122/77$ is given in (d), which implies

$$D = \frac{122}{60} = 2.03333.$$

5. A Markov chain $(X_n)_{n\geq 0}$ on $\{0, 1, 2, \ldots\}$ has transition probabilities given as follows. For $i \geq 1$,

$$p_{i,i+1} = 1 - p_{i,i-1} = p,$$

and

$$p_{0,1} = 1.$$

Note that the chain is irreducible.

(a) Determine the values of p for which the chain is transient, null, and positive recurrent.

(b) For fixed $0 < p < 1$ and each state $i = 0, 1, \ldots$, find the long-run proportion of time the chain spends in state i.

[8 marks]

Ans.

(a) (6 marks) For transience/recurrence, let $f_i = P(X_n = 0, \text{ some } n \geq 1 | X_0 = i)$. Then we know that the f_i are the minimal non-negative solution to the first-step analysis equations

$$f_i = p_{i,0} + \sum_{j\geq 1} p_{ij} f_j,$$

which in this case are

$$\begin{aligned} f_0 &= f_1 \\ f_1 &= (1-p) + p f_2 \\ f_i &= (1-p) f_{i-1} + p f_{i+1}, \quad i \geq 2. \end{aligned}$$

These last two lines are the same equations for random walk with boundary from lecture (the only difference is in the first equation for f_0) and so we see that $f_1 = 1$ for $p \le 1/2$, and $f_1 < 1$ for $p > 1/2$, and so the same is true for f_0 and the chain is recurrent for $p \le 1/2$, and transient for $p > 1/2$.
The chain is positive recurrent if and only if $\pi P = \pi$ has a probability solution. These equations are

$$\begin{aligned}\pi_0 &= (1-p)\pi_1\\ \pi_1 &= \pi_0 + (1-p)\pi_2\\ \pi_i &= (1-p)\pi_{i+1} + p\pi_{i-1}, \quad i \ge 2.\end{aligned}$$

Solving recursively in the usual way, we find that for $i \ge 1$

$$\pi_i = \pi_1 \left(\frac{p}{1-p}\right)^{i-1},$$

and the first equation then implies that for $i \ge 1$,

$$\pi_i = \frac{\pi_0}{1-p}\left(\frac{p}{1-p}\right)^{i-1}.$$

We can solve for π_0 if and only if $p/(1-p) < 1$ (if an only if $p < 1/2$):

$$\begin{aligned}1 &= \pi_0 \left(1 + (1-p)^{-1}\sum_{i\ge 1}\left(\frac{p}{1-p}\right)^{i-1}\right)\\ &= \pi_0\left(1 + \frac{1}{1-2p}\right) = \pi_0\left(\frac{2(1-p)}{1-2p}\right),\end{aligned}$$

in which case

$$\begin{aligned}\pi_0 &= \frac{1-2p}{2(1-p)}\\ \pi_i &= \frac{1-2p}{2(1-p)^2}\left(\frac{p}{1-p}\right)^{i-1}, \ \ i \ge 1.\end{aligned}$$

(b) (2 marks) For $p < 1/2$, the proportion of time spent in state i is given by π_i from part (a). For $p \ge 1/2$, the proportion of time spent in any state tends to zero.

6. Let $(B_t)_{t\ge 0}$ be a Brownian motion.

(a) For $0 < t_1 < t_2 < t_3$, find constants a, b, c such that

$$B_{t_2} = aB_{t_1} + bB_{t_3} + cZ,$$

where Z is standard normal and independent of (B_{t_1}, B_{t_3}).

(b) Compute $E[B_2|B_1 = x, B_4 = y]$ and $Var(B_2|B_1 = x, B_4 = y)$.

(c) An insurance company receives 10 thousand dollars per day in payments. For $i = 1, 2, \ldots$, let X_i be the amount, in thousands of dollars, that the insurance company pays out in claims i days from now. Assume that the X_i are i.i.d. standard gamma with parameter 10, having density

$$\frac{x^9 e^{-x}}{9!}, \ \ x > 0.$$

The company currently has 100 thousand dollars in its bank account for paying out claims. Assuming payments and claims are made at the same time each day, use the approximation of random walk by Brownian motion to estimate the probability that the insurance company has a positive amount in its bank account every day for the next 1000 days. You may want to use the fact that $M_t := \min_{0\leq s\leq t}\{B_s\} \stackrel{d}{=} -|B_t|$.

[9 marks]

Ans.

(a) (3 marks) We know $Cov(B_t, B_s) = \min\{s,t\}$, which implies that a, b, c, have to satisfy

$$\begin{aligned}
t_1 &= Cov(aB_{t_1} + bB_{t_3} + cZ, B_{t_1}) = (a+b)t_1,\\
t_2 &= Cov(aB_{t_1} + bB_{t_3} + cZ, aB_{t_1} + bB_{t_3} + cZ) = (a^2+2ab)t_1 + b^2t_3 + c^2,\\
t_2 &= Cov(aB_{t_1} + bB_{t_3} + cZ, B_{t_3}) = at_1 + bt_3,
\end{aligned}$$

which implies

$$\begin{aligned}
a &= \frac{t_3-t_2}{t_3-t_1},\\
b &= \frac{t_2-t_1}{t_3-t_1},\\
c &= \sqrt{t_2 - \frac{t_3-t_2}{t_3-t_1}\left(\frac{t_3+t_2-2t_1}{t_3-t_1}\right)t_1 - \left(\frac{t_2-t_1}{t_3-t_1}\right)^2 t_3}.
\end{aligned}$$

(b) (3 marks) From the representation of the previous part, we know

$$(B_2|B_1 = x, B_4 = y) \sim \text{Normal}(ax+by, c^2),$$

where $a = 2/3, b = 1/3$ and $c = \sqrt{2/3}$. Thus the mean is $(2x+y)/3$ and the variance is $2/3$.

(c) (3 marks) The amount of profit/loss each day is i.i.d. distributed $X_i - 10$, which is a mean zero variance 10 random variable. Thus if $S_n = \sum_{i=1}^n (X_i - 10)$, then the process

$$\left(\frac{S_{\lfloor tn\rfloor}}{\sqrt{10}\sqrt{n}}\right)_{0\leq t\leq 1}$$

is well approximated by a Brownian motion for large n. In particular

$$\begin{aligned}
P\left(\min_{0\leq k\leq 1000}\{S_k\} > -100\right) &= P\left(\min_{0\leq k\leq 1000}\left\{\frac{S_k}{100}\right\} > -1\right)\\
&\approx P(M_1 > -1)\\
&= P(|B_1| \leq 1) \approx 0.682.
\end{aligned}$$

Tables of the Normal Distribution

Probability Content from -oo to Z

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817

End of Exam