EECS 126 Probability and Random Processes University of California, Berkeley: Spring 2016
Kannan Ramchandran May 11, 2016

Final Exam (Solutions)

Last name	First name	SID

Name of student on your left:
Name of student on your right:

- DO NOT open the exam until instructed to do so.
- The total number of points is **110**, but a score of $\geq$ **100** is considered perfect.
- You have 10 minutes to read this exam without writing anything and 150 minutes to work on the problems.
- Box your final answers.
- **Remember to write your name and SID on the top left corner of every sheet of paper.**
- **Do not write on the reverse sides of the pages.**
- All eletronic devices must be turned off. Textbooks, computers, calculators, etc. are prohibited.
- No form of collaboration between students is allowed. If you are caught cheating, you may fail the course and face disciplinary consequences.
- **You must include explanations to receive credit.**

Problem	Max	Points	Problem	Max	Points
1	12		7	10	
2	16		8	12	
3	8		9	12	
4	8		10	8	
5	8		11	1	
6	15				
Total				110	

Cheat sheet

1. Discrete Random Variables

1) Geometric with parameter $p \in [0,1]$:
   $P(X=n) = (1-p)^{n-1}p,\ n \geq 1$
   $E[X] = 1/p,\ \mathrm{var}(X) = (1-p)p^{-2}$

2) Binomial with parameters N and p:
   $P(X=n) = \binom{N}{n} p^n (1-p)^{N-n},\ n = 0, \ldots, N,$ where $\binom{N}{n} = \frac{N!}{(N-n)!n!}$
   $E[X] = Np,\ \mathrm{var}(X) = Np(1-p)$

3) Poission with parameter λ:
   $P(X=n) = \frac{\lambda^n}{n!} e^{-\lambda},\ n \geq 0$
   $E[X] = \lambda,\ \mathrm{var}(X) = \lambda$

2. Continuous Random Variables

1) Uniformly distributed in $[a,b]$, for some $a < b$:
   $f_X(x) = \frac{1}{b-a} \mathbf{1}\{a \leq x \leq b\}$
   $E[X] = \frac{a+b}{2},\ \mathrm{var}(X) = \frac{(b-a)^2}{12}$

2) Exponentially distributed with rate $\lambda > 0$:
   $f_X(x) = \lambda e^{-\lambda x} \mathbf{1}\{x \geq 0\}$
   $E[X] = \lambda^{-1},\ \mathrm{var}(X) = \lambda^{-2}$

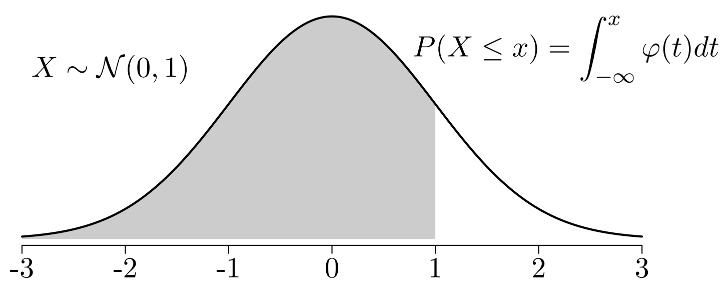
3) Gaussian, or normal, with mean μ and variance σ^2:
   $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{-\frac{(x-\mu)^2}{2\sigma^2}\}$
   $E[X] = \mu,\ \mathrm{var}(X) = \sigma^2$

4) Erlang distribution, i.e., some of k i.i.d. exponential random variables with rate λ:
   $f_X(x) = \frac{\lambda^k x^{k-1} e^{-\lambda x}}{(k-1)!} \mathbf{1}\{x \geq 0\}$
   $E[X] = \frac{k}{\lambda},\ \mathrm{var}(X) = \frac{k}{\lambda^2}$

3. Estimation

1) LLSE: Let X and Y be random variables. Then,
   $L[X|Y] = E(X) + \frac{\mathrm{cov}(X,Y)}{\mathrm{var}(Y)}(Y - E[Y])$.

4. Normal Distribution Table



	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
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Problem 1. (12 pts) Evaluate the statements with *True* or *False*. You must give brief explanations in the provided boxes to get any credit.

(a) Let $\hat{X}$ be the cubic (functions of the form $f(z) = az^3 + bz^2 + cz + d$) least squares estimate of X given the observation Y. True or False: $\mathrm{cov}(X - \hat{X}, Y^2) = 0$.

Solution: True.

(b) Let X and Y be two Jointly Gaussian random variables such that $\mathrm{cov}(X, Y) = 0$. True or False: X and Y need not be independent.

Solution: False, uncorrelated Jointly Gaussian random variables are independent.

(c) True or False: $\mathrm{var}(X) = E[\mathrm{Var}(X|Y)] + \mathrm{Var}(E[X|Y])$ for random variables X and Y.

Solution: True, this is the law of total variance.

(d) Explain in the box below why, in general, the Chernoff bound is a "better" bound than Chebyshev and why Chebyshev is a "better" bound than Markov.

Solution: The Chernoff bound uses the most information about the distribution, in particular it uses all of the moments, whereas the Chebyshev inequality uses the second moment and the Markov inequality uses only the first moment.

Problem 2. (16 pts)

(a) (4 pts) Consider an Erdos-Renyi random graph $G = G(n,p)$. Let the random variable X give the number of *isolated* nodes in the graph. Your friend Bollobas claims that $\text{Var}(X) = nq(1-q)$, where q is the probability a node is isolated. Is Bollobas correct? You must explain your answer to receive any credit.

Solution: Bollobas is not correct. This would be true if the indicators I_i, I_j of nodes i and j being isolated were independent. This is, however not true.

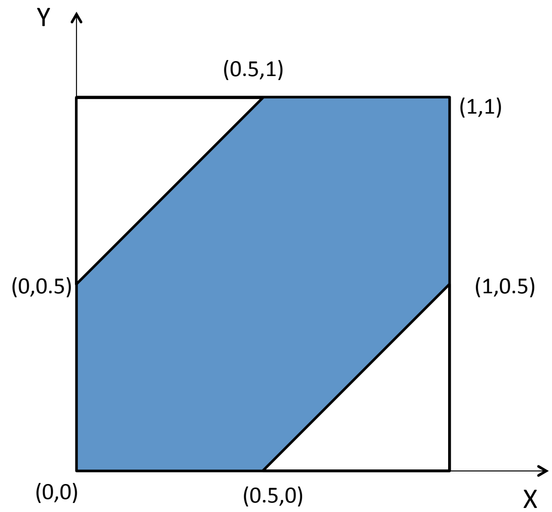
(b) (6 pts) HKN wants to take a survey of how many Berkeley EECS undergraduate students actually cheat on their exams. As this is a sensitive question and the HKN officers know that almost nobody would admit to cheating even if they did, they devise a privacy-preserving scheme as follows: Each survey-taking student privately flips a fair coin. If the coin comes up heads, he/she answers the question truthfully. If it comes up tails, he/she answers the question randomly, i.e. equally likely to be "yes" and "no". Each student cheats on his or her exam with probability p, and each student acts independently. If 200 students respond to the survey such that there are 70 respondents who said "yes" and 130 who said "no", what is the MLE of p?

Solution: We present a slightly more general solution (where there are n students, k of whom respond "yes"). Let the vote of student i by the indicator random variable X_i where $X_i = 1$ if the student votes "yes". Now, note that the probability a student answers yes is given by $P(\text{"yes"}) = \frac{1}{2}p + \frac{1}{4}$ We can see that:

$$P(X_1, X_2, \ldots, X_n | p) = \left(\frac{2p+1}{4}\right)^k \left(\frac{2(1-p)+1}{4}\right)^{n-k}$$

Taking the derivative and solving, we can see that: $p = \frac{4k-n}{2n}$. Now, note that if $n > 4k$, $\hat{p}$ is negative, and if $k > n/2$, $\hat{p} > 1$. Thus, in the former case, we set $\hat{p} = 0$, and in the latter, we set $\hat{p} = 1$. Plugging in the values for k and n, one can see that $\hat{p} = \frac{1}{5}$.

(c) (6 pts) Consider the following figure:



Suppose that (X, Y) are uniformly distributed on the shaded region. Find $L[X|Y]$.

Solution: One can see that the MMSE $E[X|Y] = \frac{1}{2}Y + \frac{1}{4}$ as shown in the figure below. As the MMSE is linear, it coincides with the LLSE.

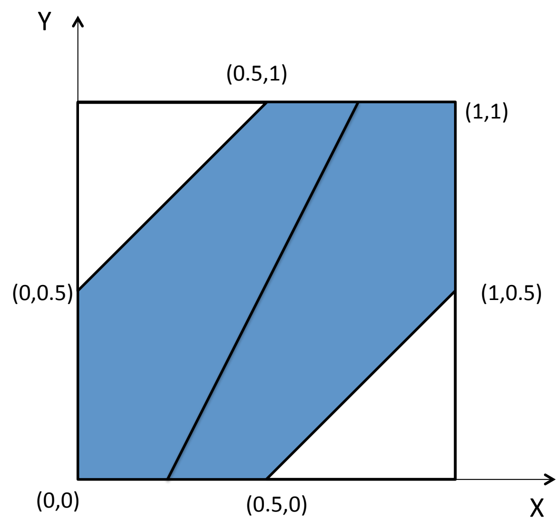


Figure 1: The black line indicates the LLSE

Problem 3. (8 pts) Consider IID random variables $X_i \sim N(0,1)$. Find the MMSE of $X_1+X_2+X_3$ given the observations X_1+X_2, X_2+X_3, and X_3+X_4.

Solution: Since the X_i are jointly Gaussian, the MMSE is linear, so we are looking for an estimator of the form:

$$a(X_1+X_2)+b(X_2+X_3)+c(X_3+X_4)+d$$

Note that the MMSE is unbiased, so $d=0$. Additionally, one may set up the following equations:

$$E[(X_1+X_2+X_3-(a(X_1+X_2)+b(X_2+X_3)+c(X_3+X_4)))(X_1+X_2)]=0$$
$$E[(X_1+X_2+X_3-(a(X_1+X_2)+b(X_2+X_3)+c(X_3+X_4)))(X_2+X_3)]=0$$
$$E[(X_1+X_2+X_3-(a(X_1+X_2)+b(X_2+X_3)+c(X_3+X_4)))(X_3+X_4)]=0$$

Simplifying gives the set of equations:

$$2-2a-b=0$$
$$2-a-2b-c=0$$
$$1-b-2c=0$$

From these linear equations, we can see that $a=\frac{3}{4}, b=\frac{1}{2}, c=\frac{1}{4}$

Problem 4. (8 pts) Suppose X, Y, Z are mutually independent random variables, each of which is uniformly distributed on $[0,1]$. Find the MMSE of $(X+Y)^2$ given the observation $Y+Z$.

Solution: We are looking for $E[(X+Y)^2|Y+Z] = E[X^2]+2E[X]E[Y|Y+Z]+E[Y^2|Y+Z]$. We are interested in $E[Y|Y+Z]$ and $E[Y^2|Y+Z]$ Now, let $Y+Z=u$. Note that since Y and Z are independently and uniformly distributed on $[0,1]$, the joint distribution of Y, Z is uniformly distributed on the line $Y+Z=u$ and in fact Y is uniformly distributed on the "valid" range. We have the following cases (shown in the figure below):

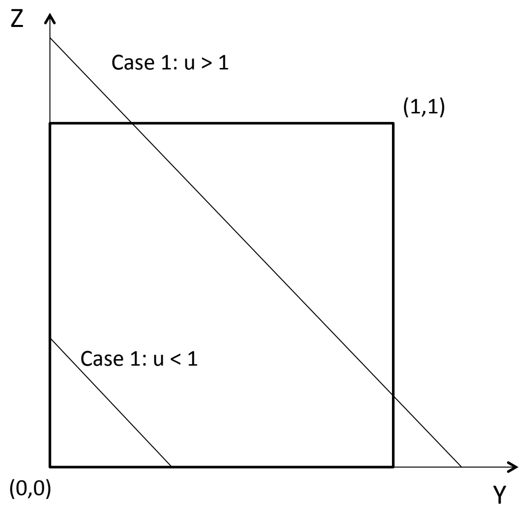


Figure 2: The two cases

Case 1: If $u \leq 1$. This implies that $Y \sim U[0,u]$. We thus have $E[Y|Y+Z] = \frac{(Y+Z)}{2}$ and $E[Y^2|Y+Z] = \frac{(Y+Z)^2}{3}$. Thus in this case:

$$E[(X+Y)^2|Y+Z] = \frac{1}{3} + \frac{(Y+Z)}{2} + \frac{(Y+Z)^2}{3}$$

Case 2: If $u > 1$, the story changes a bit. Now $Y \sim U[u-1,1]$. Thus, we can see that $E[Y|Y+Z] = \frac{(Y+Z)}{2}$ and $E[Y^2|Y+Z] = \frac{(Y+Z)+(Y+Z-1)^2}{3}$ So we have:

$$E[(X+Y)^2|Y+Z] = \frac{2}{3} + \frac{(Y+Z)}{6} + \frac{(Y+Z)^2}{3}$$

Problem 5. (8 pts) You observe one realization of a random variable Y and would like to determine which distribution it came from. Consider the hypotheses:

$$X = 0 : Y \sim N(0,1)$$
$$X = 1 : Y \sim N(0,2)$$

Design a hypothesis test that maximizes the probability of correct detection ($P(\hat{X} = 1|X = 1)$) such that the probability of false alarm $P(\hat{X} = 1|X = 0) \leq 0.05$.

Solution: We set up the likelihood ratio:

$$\frac{f_{Y|1}(y|X=1)}{f_{Y|0}(y|X=0)} = \frac{\frac{1}{\sqrt{2}\sqrt{2\pi}}e^{-y^2/2}}{\frac{1}{\sqrt{2\pi}}e^{-y^2}} = \frac{1}{\sqrt{2}}e^{y^2/2}$$

Note that this is monotonically increasing in $|y|$. So we look for a threshold τ such that $\hat{X} = 1$ if $|y| \geq \tau$. In order to find τ, we use the threshold on PFA:

$$P(\hat{X} = 1|X = 0) \leq 0.05$$

Now, conditioned on $X = 0$, $Y \sim N(0,1)$. Thus, we want $P(|Y| \geq \tau|X = 0) \leq 0.05$. Using the table, we can see that setting $\tau = 1.96$ achieves this.

Problem 6. (15 pts) Reddit is running an experiment known as "the button". The button has a t second timer. Everytime it is pressed, its clock resets to t seconds. If the clock expires, the button shuts down forever.

(a) (10 pts) Users arrive to the webpage according to a Poisson process with rate λ. Upon arriving, they decide to press the button with probability p, otherwise they leave. In addition to the users, there is an administrator for Reddit online 24 hours a day who refreshes the page and presses the button according to an independent Poisson process with rate 2λ. What is the expected amount of time until the button expires?

Solution: First note that the process that tracks when the button is clicked is a Poisson process with rate $\mu = \lambda(p+2)$. Now, we are looking for $E[T]$, the time at which the button expires. The following diagram should help guide our thinking:

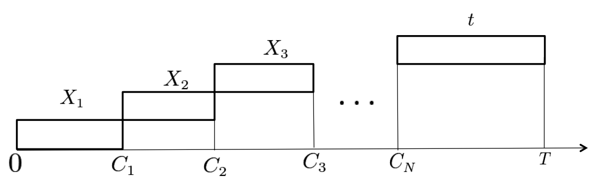


Figure 3: Here, C_i indicates the time of the ith click of the button, and X_i indicates the interarrival time between click $i-1$ and click i of the process

Suppose that N is the number of clicks which occur, then by iterated expectations we have $E[T] = E[E[T|N]]$. Now, one can see that if there were N clicks (or N arrivals in this new process), we have:

$$E[T|N] = t + nE[X_1|X_1 \le t]$$

We may find

$$\begin{aligned} E[T_1|T_1 \le t] &= \frac{\int_0^t x\mu e^{-\mu x} dx}{\int_0^t \mu e^{-\mu x} dx} \\ &= \frac{1/\mu - (t+1/\mu)e^{-\mu t}}{1-e^{-\mu t}} \end{aligned}$$

Now, we calculate $E[N]$. Let $P(X_i > t) = p$. Then, $P(N=k) = (1-p)^k p$. Thus, we have $E[N] = e^{\mu t} - 1$. Putting these together, we can see that:

$$E[T] = t + (e^{\mu t} - 1) \cdot \left(\frac{1/\mu - (t+1/\mu)e^{-\mu t}}{1-e^{-\mu t}} \right)$$

(b) (5 pts) For this part, assume no time limit on the button (that is $t \to \infty$). Suppose that you observe the third time the button was pressed and the first time the button was pressed. Find the LLSE of the time at which the button was pressed for the second time.

Solution: Keeping consistent with the notation of part (a), we are looking for $L[C_2|C_1, C_3]$. One way to solve this would be to go through the formulas. Here we present a slick way. We can rewrite the expression in terms of the interarrival times to see that:

$$L[C_1|C_1, C_3] = L[X_1 + X_2|X_1, X_1 + X_2 + X_3]$$

Now, note that by symmetry, $L[X_1+X_2|X_1, X_1+X_2+X_3] = L[X_1+X_3|X_1, X_1+X_2+X_3]$. Thus, we have:

$$\begin{aligned} L[X_1 + X_2|X_1, X_1 + X_2 + X_3] &= \frac{1}{2}L[X_1 + X_2|X_1, X_1 + X_2 + X_3] + \frac{1}{2}L[X_1 + X_3|X_1, X_1 + X_2 + X_3] \\ &= \frac{1}{2}L[X_1|X_1, X_1 + X_2 + X_3] + \frac{1}{2}L[X_1 + X_2 + X_3|X_1, X_1 + X_2 + X_3] \\ &= \frac{1}{2}(X_1 + X_1 + X_2 + X_3) \\ &= \frac{1}{2}(C_1 + C_3) \end{aligned}$$

Another way to see this is that conditioned on the third arrival and the first arrival, the second arrival is uniform between the two.

Problem 7. (10 pts) A particle is moving randomly according to the following update equations: at time $i = 1, 2, \dots,$ the particle's position is given by:

$$X(i) = 2X(i-1) + V(i).$$

with initial position $X(0)$. At each time $j \geq 0$, you have access only to a noisy measurement of the position::

$$Y(j) = X(j) + W(j).$$

Assume that $X(0), V(i), W(i)$ are IID $N(0,1)$ for all $i \geq 0$.

(a) (4 pts) Find $\hat{X}(0) = E[X(0)|Y(0)]$

Solution: $\frac{1}{2}Y(0)$

(b) (6 pts) Your friend Rudolf wants to estimate $X(1)$ from the new observation $Y(1)$ that comes in, and the estimate $\hat{X}(0)$ from part (i). In other words, he wants to form the estimate

$$\hat{X}(1) = \alpha\hat{X}(0) + \beta Y(1).$$

Find α and β and explain your reasoning using a diagram.

Solution: This can be done by updating the LLSE in the following way:

$$L[X(1)|Y(0), Y(1)] = L[X(1)|Y(0)] + L[X(1)|Y(1) - L[Y(1)|Y(0)]]$$

Now, we can see that:

$$L[X(1)|Y(0)] = 2L[X(0)|Y(0)] = 2\hat{X}(0)$$

Also,

$$L[Y(1)|Y(0)] = L[X(1)|Y(0)] = L[2X(0)|Y(0)] = 2\hat{X}(0)$$

Note that since $\hat{X}(0) = \frac{1}{2}Y(0)$, we haveL

$$L[X(1)|Y(1) - L[Y(1)|Y(0)]] = L[X(1)|Y(1) - Y(0)]$$

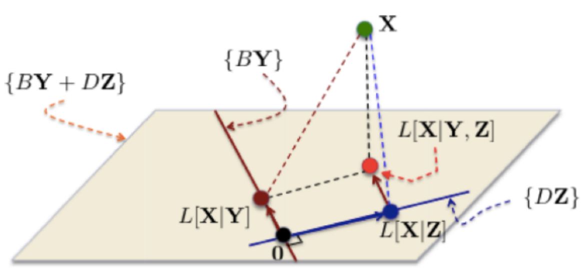
Now, we can calculate this using the formula. We can see that:

$$\begin{aligned}
\text{cov}(X(1), Y(1) - Y(0)) &= E[(X(1))(Y(1) - Y(0)))] \\
&= E[X(1)Y(1) - X(1)Y(0)] \\
&= E[(2X(0 + V(1))(2X(0) + V(1) + W(1)))] - \\
&\quad E[(2X(0) + V(1))(X(0) + W(0))] \\
&= 3
\end{aligned}$$

Similar calculations yield $\text{Var}(Y(1) - Y(0)) = 4$. Thus, we have:

$$\hat{X}(1) = 2\hat{X}(0) + \frac{3}{4}(Y(1) - 2\hat{X}(0))$$

Thus, we have $\alpha = \frac{1}{2}$ and $\beta = \frac{3}{4}$. The picture is the following:



Problem 8. (12 pts) Consider an infinite continuous-time Markov chain X_t with $t \geq 0$, depicted in Fig. 4. Let $\mathbf{X_0} = (\mathbf{0}, \mathbf{2})$ and T be the first time the chain visits state $(0,0)$. Find $E[T]$ in terms of λ, μ and α.

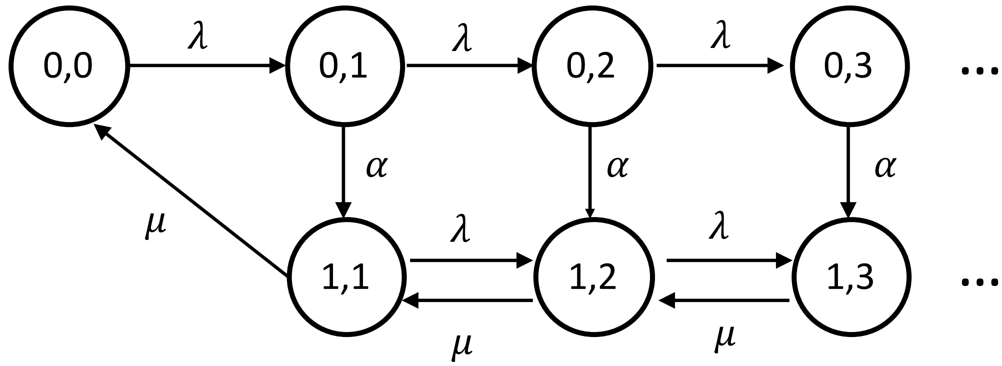


Figure 4: Infinite continuous-time Markov chain

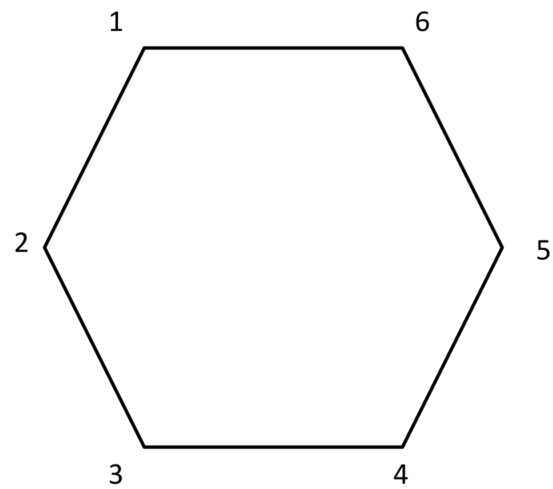
Solution: Define the hitting time to the one-level-left column starting from some state in the top row as p. Similarly, define the hitting time to the one-level-left column starting from some state in the bottom row as q. Using symmetry, we can find the following equations:

$$
\begin{aligned}
p &= \frac{1}{\lambda+\alpha} + \frac{\lambda}{\lambda+\alpha}(p+q) + \frac{\alpha}{\lambda+\alpha}q, \\
q &= \frac{1}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu}(2q).
\end{aligned}
$$

From the second equation, $q = \frac{1}{\mu-\lambda}$. Then, $p = \frac{1}{\alpha} + \frac{\lambda+\alpha}{\alpha(\mu-\lambda)}$. Now, our answer is

$$
p + q = \frac{1}{\alpha} + \frac{\lambda+\alpha}{\alpha(\mu-\lambda)} + \frac{1}{\mu-\lambda}
$$

Problem 9. (12 pts) An ant is performing a random walk on the vertices $\{1,2,3,4,5,6\}$ of the following hexagon:



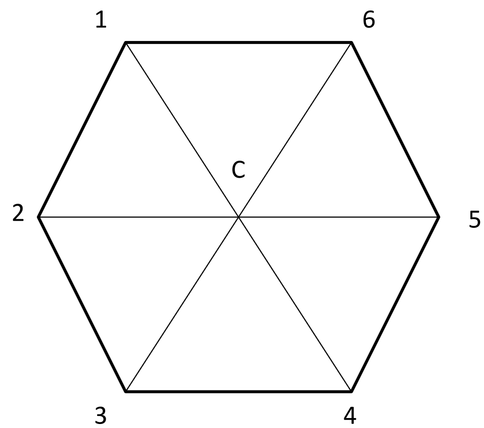
(a) (6 pts) If the ant starts at vertex 1, what is the expected amount of time until it returns to vertex 1?

Solution: Let $\beta(i)$ be the hitting time from state i to state 1. Note that states 2 and 6 are equivalent, and states 3 and 5 are equivalent. We are interested in $\beta(1) = 1 + \beta(2)$. We have the following equations:

$$\begin{aligned}
\beta(1) &= 1 + \beta(2) \\
\beta(2) &= 1 + \frac{1}{2}\beta(3) \\
\beta(3) &= 1 + \frac{1}{2}\beta(2) + \frac{1}{2}\beta(4) \\
\beta(4) &= 1 + \beta(2)
\end{aligned}$$

Solving these gives: $\beta(1) = 6$.

(b) (6 pts) An exterminator puts diagonal lines on the hexagon and places a trap at point C. Supposing that the ant performs a random walk on the vertices $\{1, 2, 3, 4, 5, 6, C\}$ and starts the walk at point 1, what is the probability that the ant returns to point 1 before hitting point C?



Solution: Let $T_i = \min\{n \geq 1 : X_n = i\}$. Let us define $p_i := P(T_1 < T_C | X_0 = i)$. We are interested in p_1. We can set up the following equations:

$$p_1 = \frac{1}{3}p_6 + \frac{1}{3}p_2 = \frac{2}{3}p_2$$
$$p_2 = \frac{1}{3} + \frac{1}{3}p_3$$
$$p_3 = \frac{1}{3}p_2 + \frac{1}{3}p_4$$
$$p_4 = \frac{1}{3}p_3 + \frac{1}{3}p_5 = \frac{2}{3}p_3$$

Solving these equations gives $p_1 = \frac{7}{27}$

Problem 10. (8 pts) The final exam scores of 10 students taking a graduate EECS class are: $34, 45, 50, 56, 60, 74, 80, 81, 95, 100$. Suppose everyone gets either an A or a B, with A grades being distributed as a Normal with mean μ_A and variance σ^2, and B grades as Normal with mean μ_B and variance σ^2. If we were to use a Hard-EM algorithm to cluster the scores into A and B grade clusters and we initialize the algorithm with $\mu_A = 80$ and $\mu_B = 50$, what will be estimates of μ_A and μ_B produced once the algorithm converges?

Solution: We find the boundary at $\frac{\mu_A+\mu_B}{2} = 65$. Thus, we have the following estimates:

$$A : 74, 80, 81, 95, 100$$
$$B : 34, 45, 50, 56, 60$$

Note that we update $\mu_A \leftarrow \frac{74+80+81+95+100}{5} = 86$ and $\mu_B \leftarrow \frac{34+45+50+56+60}{5} = 49$. Then, in the next iteration, we set the boundary at $\frac{49+86}{2} = 67.5$. This separates the two scores into the same groups, and we will get the same values for μ_A, μ_B. Thus, after converging, $\mu_A = 86$ and $\mu_B = 49$.

Problem 11. (1 point) Please leave any feedback for the course staff here. What did you like and dislike about the course? What can we improve upon?

END OF THE EXAM.

Please check whether you have written your name and SID on every page.

Hope you enjoyed the class! You learned a lot!