Massachusetts Institute of Technology

Department of Electrical Engineering & Computer Science

6.041/6.431: Probabilistic Systems Analysis

(Spring 2009)

Problem Set 10: Solutions

Due: May 6, 2009

1. **(a)** The probability the person is working on project B on the third day given that he was working on project A on the first day is the same as the probability that in two transitions the Markov chain ends up in state B if we start in state A since each transition corresponds to a day. So, we are looking for the 2-step transition probability from A to B.

$$\begin{aligned} p_{AB}(2) &= p_{AA}(1)p_{AB}(1) + p_{AB}(1)p_{BB}(1) \\ &= (0.5)(0.4) + (0.4)(0.6) \\ &= \boxed{0.44} \end{aligned}$$

Note that it is unnecessary to include transitions to $Florida$, $Hawaii$, and $Fired$ because once we end up in those states we can never get back to B.

(b) We are looking for the absorption time to a recurrent state starting from state A. Define

$$T_A - \text{time until absorption starting from state } A$$

$$T_B - \text{time until absorption starting from state } B$$

We are looking for $E[T_A]$. The reason we define T_B is to aid us in our calculation of $E[T_A]$. Our solution approach is to consider the expected absorption time *conditioned on taking a particular path*. Then the *uncondtional* expected absorption time is gotten by averaging over these conditioned expectations by weighting them with the probability that that particular path is taken. In mathematical terms,

$$\begin{aligned} E[T_A] = \; & E[T_A \mid \text{take path A to A}] \cdot P(\text{take path A to A}) + \\ & E[T_A \mid \text{take path A to B}] \cdot P(\text{take path A to B}) + \\ & E[T_A \mid \text{take path A to Florida}] \cdot P(\text{take path A to Florida}) \end{aligned}$$

Note that the above is nothing more than the law of iterated expectations. To find

$$E[T_A \mid \text{take path A to A}] \ ,$$

think of the event, T_A given that we take the path from A to A, in the following manner. First, we take *one* step in taking the path from A to A. Then since we ended up at A, we must take more additional steps to get absorbed into a recurrent state. We have defined a variable for this which turns out to be T_A itself. Thus,

$$\begin{aligned} E[T_A \mid \text{take path A to A}] &= E[1 + T_A] \\ E[T_A \mid \text{take path A to B}] &= E[1 + T_B] \\ E[T_A \mid \text{take path A to Florida}] &= E[1 + 0] = 1 \end{aligned}$$

Although it may seem that we are running in circles by using variables we are trying to find, the rest of this solution should enlighten how such an approach will lead to an answer. Now, plugging in numbers and putting everything in terms of T_A or T_B, we get

$$E[T_A] = E[1 + T_A] \cdot 0.5 + E[1 + T_B] \cdot 0.4 + 1 \cdot 0.1$$

Similarly, for T_B, we get

$$E[T_B] = E[1 + T_A] \cdot 0.2 + E[1 + T_B] \cdot 0.6 + 1 \cdot 0.2$$

We have two equations in two unknowns, the unknowns being $E[T_A]$ and $E[T_B]$, we have that
$\boxed{E[T_A] = \frac{20}{3}}$

(c) The probability of being in Florida on day 1000000 is asking for the steady state probability of being in Florida. Since we begin in transient state A and there are two recurrent classes of states, we must first find the probability of being absorbed into the recurrent class consisting of states $Florida$ and $Hawaii$. Then, assuming that we are in this recurrent class, we must find the steady state probability of being in $Florida$. The product of these two probabilities is the desired answer.

Finding the absorption probability
Define

$P(A, F)$ - prob of being absorbed by $Florida$ and $Hawaii$ starting in state A
$P(B, F)$ - prob of being absorbed by $Florida$ and $Hawaii$ starting in state B

Then, looking at the chain we can immediately write

$$P(A, F) = 0.1 + 0.5P(A, F) + 0.4P(B, F)$$

$$P(B, F) = 0.1 + 0.6P(B, F) + 0.2P(A, F)$$

Solving the above two equations in two unknowns, we get

$$P(A, F) = \frac{2}{3}$$

Finding the steady state probability assuming we are in that recurrent class
Since the probability of all transitions *out* of a state equals the probability of all transitions *into* a state, we have

$$0.6P_{Florida} = 0.3P_{Hawaii}$$

where $P_{Florida}$ and P_{Hawaii} are the steady-state probabilities of being in states $Florida$ and $Hawaii$ respectively. Furthermore, we know that

$$P_{Florida} + P_{Hawaii} = 1$$

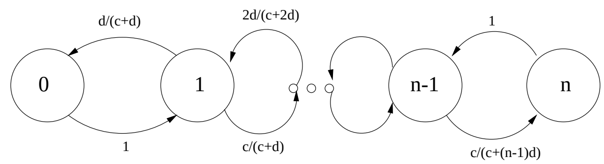
So, solving these two equations in two unknowns, we get

$$P_{Florida} = \frac{1}{3}$$

Thus,

$$P(\text{being in Florida}) = \frac{2}{3} \cdot \frac{1}{3} = \boxed{\frac{2}{9}}$$

2. (a) The corresponding Markov Chain is shown in the figure, where the states are the number of calls on the system. Transitions on this chain correspond to change of state.



We first find the steady state equations for the chain. Note that the chain has one recurrent class, so the solution to the set of equations is unique. However, the chain is periodic. Hence, the probability of being in state i at time n does not converge. However, the solution π_i still corresponds to the long-run proportion of time that the chain is in state i:

$$\begin{aligned}
&\frac{c}{c+id}\pi_i = \frac{(i+1)d}{c+(i+1)d}\pi_{i+1},\ i = 0\ldots n-2 \\
&\frac{c}{c+(n-1)d}\pi_{n-1} = \pi_n \\
&\pi_0 + \pi_1 + \ldots + \pi_n = 1
\end{aligned}$$

We may rewrite the first two equations as follows:

$$\begin{aligned}
\pi_i &= \frac{c^{i-1}(c+id)}{i!d^i}\pi_0,\ i = 1\ldots n-1 \\
\pi_n &= \frac{c^{n-1}}{(n-1)!d^{n-1}}\pi_0
\end{aligned}$$

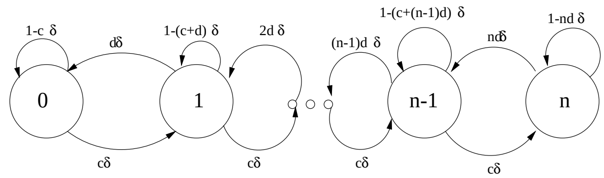
Therefore,

$$\begin{aligned}
\pi_0 &= [\sum_{i=0}^{n-1}(\frac{c^{i-1}(c+id)}{i!d^i}) + \frac{c^{n-1}}{(n-1)!d^{n-1}}]^{-1} \\
\pi_i &= \frac{c^{i-1}(c+id)}{i!d^i}\pi_0,\ i = 1\ldots n-1 \\
\pi_n &= \frac{c^{n-1}}{(n-1)!d^{n-1}}\pi_0
\end{aligned}$$

Let K be the number of calls on the system given that the number of calls just changed.

$$p_K(k) = \begin{cases} \pi_k, & k = 0\ldots n \\ 0, & \text{otherwise} \end{cases}$$

We may also use a time-sampled approximation for very small time increments δ. The corresponding Markov Chain is shown in the figure, where the states are the number of calls on the system.



We first find the steady state equations for the chain. Note that the chain has one recurrent, aperiodic class, so the steady-state probabilities exist and are uniqe. Observing the local balance equations around states i and $i+1$, we arrive at:

$$(i+1)d\delta\pi_{i+1} = c\delta\pi_i$$
$$\pi_0 + \pi_1 + \ldots + \pi_n = 1$$

We may rewrite this as follows:

$$\pi_{i+1} = \frac{c}{(i+1)d}\pi_i \tag{1}$$
$$\pi_i = \frac{c^i}{i!d^i}\pi_0 \tag{2}$$

Therefore,

$$\begin{aligned}
\pi_0\Big(\sum_{i=0}^{n}\frac{c^i}{i!d^i}\Big) &= 1 \\
\pi_0 &= \Big[\sum_{i=0}^{n}\frac{c^i}{i!d^i}\Big]^{-1} \\
\pi_i &= \frac{c^i}{i!d^i}\pi_0,\ i = 1\ldots n-1
\end{aligned}$$

Let us denote the event
{i calls on the system immediately after state change} as i_{ASC},
the event
{k calls on the system before state change} as k_{ASC},
and the event
{the number of calls on the system just changed} as SC. Then

$$\begin{aligned}
P(i_{ASC}|SC) &= \frac{P(i_{ASC}, SC)}{P(SC)} \\
&= \frac{P(i_{ASC}, SC|(i-1)_{ASC})P((i-1)_{ASC}) + P(i_{ASC}, SC|(i+1)_{ASC})P((i+1)_{ASC})}{P(SC)} \\
&= \frac{P(\text{birth}|(i-1)_{ASC})\pi_{i-1} + P(\text{death}|(i+1)_{ASC})\pi_{i+1}}{P(SC)} \\
&= \frac{(i+1)d\delta\pi_{i+1} + c\delta\pi_{i-1}}{P(SC)} \\
&= \frac{c\delta\pi_i + id\delta\pi_i}{P(SC)} \text{ via local balance eqns (1)} \\
&= \frac{(c+id)\delta c^i\pi_0}{i!d^iP(SC)} \\
P(SC) &= \sum_{k=0}^{n} P(SC|k_{ASC})P(k_{ASC}) \\
&= \sum_{k=0}^{n}[P(\text{birth}|k_{ASC}) + P(\text{death}|k_{ASC})]\pi_k
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=0}^{n-1}(c\delta + kd\delta)\pi_k + nd\delta\pi_n \\
&= \sum_{k=0}^{n-1}\frac{c^k}{k!d^k}(c\delta + kd\delta)\pi_0 + \frac{c^n\delta}{(n-1)!d^{n-1}}\pi_0 \text{ via (2)}
\end{aligned}
$$

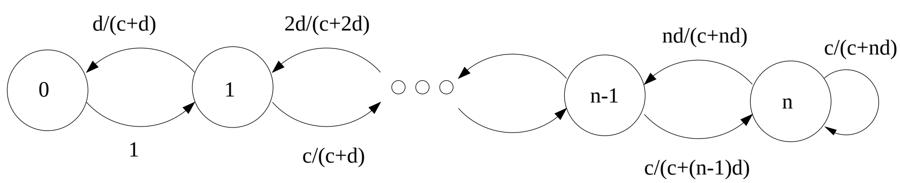
So the final result is

$$
P(i_{ASC}|SC) = \begin{cases} \frac{(c+id)c^{i-1}}{i!d^i}\left[\sum_{k=0}^{n-1}\frac{c^{k-1}}{k!d^k}(c+kd) + \frac{c^{n-1}}{(n-1)!d^{n-1}}\right]^{-1}, & i = 0,\ldots,n \\ 0, & otherwise \end{cases}
$$

which is the same as what was found using the first chain.

(b) Using the Markov chain from part (a), we need to find the probability of a death on any given transition. Since there are no self-transitions, $P(\text{birth}) + P(\text{death}) = 1$. In a finite state birth-death Markov Chain, $P(\text{birth}) = P(\text{death})$. Therefore, $P(\text{death}) = 1/2$.

(c) For this part, we also need to model blocked calls in our Markov chain. Blocked calls only occur if the system is full, i.e. when there are n calls on the system. So the Markov chain for this part is the same as the chain we found in part (a) except for the transition probabilities out of state n.



We first find the steady-state equations.

$$
\begin{aligned}
\frac{c}{c+id}\pi_i &= \frac{(i+1)d}{c+(i+1)d}\pi_{i+1} \; i = 0\ldots n-1 \\
&\pi_0 + \pi_1 + \ldots + \pi_n = 1
\end{aligned}
$$

Let K be the number of calls on the system given that a light just flashed. Solving the above equations, we have:

$$
\begin{aligned}
\pi_0 &= \left(\sum_{i=0}^{n}\frac{c^{i-1}(c+id)}{i!d^i}\right)^{-1} \\
\pi_i &= \frac{c^{i-1}(c+id)}{i!d^i}\pi_0 \; i = 1\ldots n \\
p_K(k) &= \begin{cases} \pi_k, & k = 0\ldots n \\ 0, & \text{otherwise} \end{cases}
\end{aligned}
$$

(d) $P(\text{birth}) + P(\text{death}) + P(\text{self-transition}) = 1$
$P(\text{self-transition}) = \pi_n\frac{c}{c+nd}$, ($\pi_n$:from part(c))
In a finite state birth-death Markov chain, $P(\text{birth}) = P(\text{death})$
Therefore, $P(\text{light was green} \mid \text{a light just flashed}) = \frac{1-\pi_n\frac{c}{c+nd}}{2}$.

(e) Using the π_i from part (c),

$$\text{P(k calls on the system)} = \begin{cases} \frac{\pi_{k-1}\frac{c}{c+(k-1)d}}{\sum_{i=0}^{n}\pi_i\frac{c}{c+id}}, & k = 1\ldots n-1 \\ \frac{\pi_{n-1}\frac{c}{c+(k-1)d}+\pi_n\frac{c}{c+nd}}{\sum_{i=0}^{n}\pi_i\frac{c}{c+id}}, & k = n \\ 0, & \text{otherwise} \end{cases}$$

3. (a) Use the Markov inequality on the random variable e^{sX} where $s \geq 0$, and one can have

$$\mathbf{P}(X \geq b) = \mathbf{P}(e^{sX} \geq e^{sb}) \leq \frac{\mathbf{E}[e^{sX}]}{e^{sb}} = e^{-sb}M_X(s).$$

This inequality is valid for all $s \geq 0$, and hence the tightest upper bound can be written as

$$\mathbf{P}(X \geq b) \leq \min_{s\geq 0}\left[e^{-sb}M_X(s)\right].$$

(b) Since $M_Y(s) = [M_X(s)]^n = \exp(s(nm) + \frac{s^2n}{2})$. For $s \geq 0$:

$$\begin{aligned} P(Y \geq \alpha) &\leq e^{-s\alpha}M_Y(s) \\ &= e^{-s\alpha}\left[e^{s(nm)+\frac{s^2n}{2}}\right] \end{aligned}$$

By differentiating the right-hand side and setting to 0, we obtain

$$\begin{aligned} 0 &= \frac{d}{ds}e^{-s\alpha}\left[e^{snm+\frac{s^2n}{2}}\right] \\ 0 &= (nm + s^*n - \alpha)e^{-s^*\alpha}\left[e^{s^*nm+\frac{(s^*)^2n}{2}}\right] \end{aligned}$$

So we have that $s^* = \frac{\alpha - nm}{n}$, and substituting back we get

$$P(Y \geq \alpha) \leq e^{-\frac{(\alpha-nm)^2}{2n}}$$

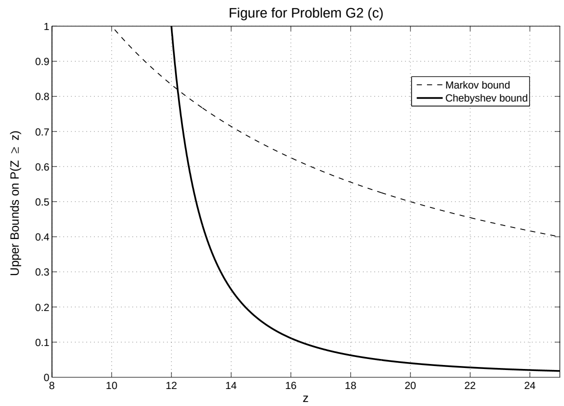
and is valid when $s^* \geq 0$, which is when $\alpha \geq nm$.

4. (a) Consider the derivation of the Markov inequality on page 265 of the text. For the inequality to be tight, we require that $E[Y_a] = E[X]$. Since $Y_a \leq X$, this can happen only if $Y_a = X$. (To see this, consider $E[X - Y_a]$. It is a weighted sum of the values that $X - Y_a$ can take, where the weights are the corresponding probabilities, and therefore at least one weight is positive. But $X - Y_a$ can take only non-negative values, since $Y_a \leq X$. If this weighted sum has to be 0, then $X - Y_a$ cannot take any positive value, and must be 0.)
This implies that

$$X = Y_a = \begin{cases} 0, & \text{if } X < a, \\ a, & \text{if } X \geq a. \end{cases}$$

This simply means X takes only two values: 0 and a. Now, $E[X] = aP(X = a) = \mu_X$. Hence, $P(X = a) = \frac{\mu_X}{a}$. This gives the following PMF for X:

$$p_X(x) = \begin{cases} 1 - \frac{\mu_X}{a}, & x = 0 \\ \frac{\mu_X}{a}, & x = a \end{cases}$$



(b) As shown in pages 266-267 of the textbook, the Chebyshev bound for a random variable Y is derived by applying the Markov bound for the random variable $Z = (Y - \mu_Y)^2$. Hence, for the Chebyshev bound to be tight for Y, the Markov bound has to be tight for Z. We now use the result of part (a). Since $E[Z] = E[(Y - \mu_Y)^2]$ is given to be σ_Y^2, this means Z must have the following PMF:

$$p_Z(z) = \begin{cases} 1 - \frac{\sigma_Y^2}{b^2}, & z = 0 \\ \frac{\sigma_Y^2}{b^2}, & z = b^2 \end{cases}$$

If $Z = 0$, then $Y = \mu_Y$ and this happens with probability $1 - \frac{\sigma_Y^2}{b^2}$. However, if $Z = b^2$, then Y can take the value of $(\mu_Y + b)$ or $(\mu_Y - b)$. For the mean to be μ_Y, both these values have to be equally likely. This gives the following PMF for Y:

$$p_Y(y) = \begin{cases} 1 - \frac{\sigma_Y^2}{b^2} & y = \mu_Y \\ \frac{1}{2}\frac{\sigma_Y^2}{b^2}, & y = (\mu_Y - b) \\ \frac{1}{2}\frac{\sigma_Y^2}{b^2}, & y = (\mu_Y + b) \end{cases}$$

(c) **Markov upper bound:** For $z > E[Z] = 10$,

$$P(Z \geq z) \leq \frac{E[Z]}{z} = \frac{10}{z}$$

Chebyshev upper bound:

$$P(Z \geq z) = P(Z - 10 \geq z - 10) \leq P(|Z - 10| \geq z - 10) \leq \frac{\sigma_Z^2}{(z-10)^2} = \frac{4}{(z-10)^2}$$

where the first inequality follows from the fact that the event $\{Z - 10 \geq z - 10\}$ is a subset of the event $\{|Z - 10| \geq z - 10\}$, and the second inequality is an application of the Chebyshev bound.

The figure shows the plot of both bounds as a function of z.

The two curves cross each other when

$$\frac{10}{z} = \frac{4}{(z-10)^2}$$

which gives $5z^2 - 102z - 500 = 0$. Solving for z, we get, $z =$8.1 9 or 12.21. Since $z > 10$, the solution is $z = 12.21$. Thus, for $z < 12.21$, the Markov bound is tighter, while for $z > 12.21$, the Chebyshev bound is tighter.

5. (a) Let $C = 0$. Hence, for every ϵ,

$$\lim_{n\to\infty} \mathbf{P}(|X_n| > \epsilon) = \lim_{n\to\infty} \frac{1}{n} = 0,$$

implying that X_n converges in probability to $C = 0$.

(b) Note that $\mathbf{E}[X_n] = 1$, and $\text{var}(X_n) = \mathbf{E}[X_n^2] - (\mathbf{E}[X_n])^2 = n-1$. Hence, $\lim_{n\to\infty} \text{var}(X_n) \to \infty$.

G1†. (a) Suppose that X_n converges to c in the mean square. Using the Markov inequality, we have

$$P(|X_n - c| \geq \epsilon) = P(|X_n - c|^2 \geq \epsilon^2) \leq \frac{E[(X_n - c)^2]}{\epsilon^2}$$

Taking the limit as $n \to \infty$, we obtain

$$\lim_{n\to\infty} P(|X_n - c| \geq \epsilon) = 0,$$

which establishes convergence in probability.

(b) Any example where the variance of the sequence X_n goes to infinity but the probability that $X_n \neq c$ goes to zero will work, such as the example from lecture. Let X_n be equal to n with probability $\frac{1}{n}$ and 0 with probability $1 - \frac{1}{n}$.

We show X_n converges to 0 in probability. For all $\epsilon > 0$, we have:

$$\lim_{n\to\infty} P(|X_n| \geq \epsilon) = \lim_{n\to\infty} P(X_n = n) = \lim_{n\to\infty} \frac{1}{n} = 0$$

We show X_n does not converge to 0 in the mean square.

$$\lim_{n\to\infty} E[X_n^2] = \lim_{n\to\infty} (0^2(1 - \frac{1}{n}) + n^2 \frac{1}{n}) = \lim_{n\to\infty} n = \infty$$