ACM 216, Problem Set 1

Ethan Wilk

January 20, 2023

NOTE: *All transition diagrams are attached at the end of the file and indexed by problem; also, $\mathbb{N}$ is assumed to include 0 throughout this set.*

1 Problem 1

Consider the HMC $(X_n)_{n\in\mathbb{N}}$ with state space $E = \{0, 1, 2\}$ and transition matrix P given by

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{pmatrix} 0.25 & 0.5 & 0.25 \\ 0.5 & 0.25 & 0.25 \\ 0.25 & 0.5 & 0.25 \end{pmatrix} \end{array}$$

Conditioned on $X_1 \in \{0, 1\}, X_0 = 0$, we know $(X_2, X_1, X_0) \in \{(0, 0, 0), (0, 1, 0), (1, 0, 0), (1, 1, 0), (2, 0, 0), (2, 1, 0)\} = A$. Using this knowledge, we can compute

$$\mathbb{P}[X_2 = 1|X_1 \in \{0, 1\}, X_0 = 0] =$$

$$\frac{\mathbb{P}[(X_2, X_1, X_0) \in \{(1, 0, 0), (1, 1, 0)\}]}{\mathbb{P}[(X_2, X_1, X_0) \in A]} = \frac{0.25}{0.75} = \boxed{\frac{1}{3}}.$$

However, we have also that

$$\mathbb{P}[X_2 = 1|X_1 \in \{0, 1\}, X_0 = 1] =$$

$$\frac{\mathbb{P}[(X_2, X_1, X_0) \in \{(1, 0, 1), (1, 1, 1)\}]}{\mathbb{P}[(X_2, X_1, X_0) \in A]} = \frac{0.3125}{0.75} = \boxed{\frac{5}{12}}.$$

Thus

$$\mathbb{P}[X_2 = 1|X_1 \in \{0, 1\}, X_0 = 0] \neq \mathbb{P}[X_2 = 1|X_1 \in \{0, 1\}, X_0 = 1].$$

2 Problem 2

Let R_t, C_t be the positions of Rat and Cat, respectively, at time t. Then the state of the system at any such time t can be captured by the tuple (R_t, C_t). Now define the stopping time

$$\tau = \min\{n : R_n = C_n\},$$

so we seek $\mathbb{E}[\tau|(R_0, C_0) = (2,1)]$. We define, $\forall\ t \geq 0$,

$$u_t(r,c) = \mathbb{E}[\tau|(R_t, C_t) = (r,c)],$$

$$p_t(r,c) = \mathbb{P}[(R_{t+1}, C_{t+1}) = (r,c)|(R_t, C_t)].$$

Of course, since we are given transition matrices for the system, and transition matrices are representations of Markov chains, we must have the following Markov property for the system:

$$\mathbb{P}[(R_{t+1}, C_{t+1}) = (r,c)|(R_t, C_t), ..., (R_0, C_0)]$$

$$= \mathbb{P}[(R_{t+1}, C_{t+1}) = (r,c)|(R_t, C_t)] = p_t(r,c).$$

Additionally, since the processes are homogeneous with respect to their transition matrices (the entries of the transition matrices are not written as functions of time), we have that

$$\mathbb{P}[(R_{i+1}, C_{i+1}) = (r,c)|(R_i, C_i) = (r,c)] = \mathbb{P}[(R_{j+1}, C_{j+1}) = (r,c)|(R_j, C_j) = (r,c)],$$

$$u_t(r,c) = \mathbb{E}[\tau|(R_{t+1} = r, C_{t+1} = c)] - 1 = \mathbb{E}[\tau|(R_t = r, C_t = c)]$$

$\forall\ i, j \geq 0$. (The expected time to reach a given state k from another state l will be independent of time, but since τ is defined by the time index on the system, we must shift the LHS by 1 to account for its 1-unit latency in time relative to the RHS.) We will thus suppress the subscript in $u_t(r,c)$ such that $u(r,c) = u_0(r,c)$ in all forthcoming computations.

We will now use these results to compute the desired value. From the initial state of the system, $(R_0, C_0) = (2,1)$, we know the system can take on four possible states at time $t = 1 : (R_1, C_1) \in \{(2,1), (1,2), (1,1), (2,2)\}$. We thus have that, using first-step analysis,

$$u(2,1) = (u(2,1)+1)p_0(2,1) + (u(1,2)+1)p_0(1,2)$$

$$+(u(1,1)+1)p_0(1,1) + (u(2,2)+1)p_0(2,2),$$

and similarly,

$$u(1,2) = (u(1,2)+1)p_1(1,2) + (u(2,1)+1)p_1(2,1)$$

$$+(u(1,1)+1)p_1(1,1) + (u(2,2)+1)p_1(2,2)$$

$$\implies u(1,2) = \frac{25}{23} + \frac{12}{23}u(2,1)$$

$$\implies u(2,1) = (u(2,1)+1)\left(\frac{1}{5}\times\frac{3}{10}\right) + \left(\frac{48}{23}+\frac{12}{23}u(2,1)\right)\left(\frac{4}{5}\times\frac{7}{10}\right)$$
$$+\left(\frac{4}{5}\times\frac{3}{10}\right)+\left(\frac{1}{5}\times\frac{7}{10}\right)$$
$$\implies \boxed{u(2,1) = \frac{370}{149}}.$$

3 Problem 3

FALSE. Let $(X_n)_{n\in\mathbb{N}} = (U_n)_{n\in\mathbb{N}}$, where

$$U_n = \begin{cases} 1, & \text{with probability } 0.5 \\ 0, & \text{with probability } 0.5. \end{cases}$$

It is obvious that $\mathbb{P}[U_n = u_n | U_{n-1} = u_{n-1}, ..., U_0 = u_0] = \mathbb{P}[U_n = u_n | U_{n-1} = u_{n-1}]$, since U_n is in fact constructed to be completely independent of all previous variables in its sequence (such is the case if it models the outcome of a coin toss in a sequences of tosses, for instance). Thus $(U_n)_{n\in\mathbb{N}}$ is a Markov chain, so $(U_{n+1})_{n+1\in\mathbb{N}}$ is also a Markov chain by Problem 4a. We will define $(Y_n)_{n\in\mathbb{N}} = (U_{n+1})_{n+1\in\mathbb{N}}$, and $(Z_n)_{n\in\mathbb{N}} = (X_n + Y_n)_{n\in\mathbb{N}}$.

Then we may compute

$$\mathbb{P}[Z_2 = 2 | Z_1 = 1] = \mathbb{P}[(U_3, U_2, U_1) = (1,1,0) | (U_2, U_1) \in \{(0,1),(1,0)\}] = \frac{1}{4}.$$

The preceding computations are justified by the fact that, conditioned on $Z_1 = 1$, the tuple $(U_2, U_1) \in \{(0,1),(1,0)\}$, so we may ask what the chance is that $Z_2 = U_3 + U_2 = 2$ given this information. Since $(U_2, U_1) \in \{(0,1),(1,0)\}$, it is clear that $(U_3, U_2, U_1) \in \{(1,1,0),(0,1,0),(1,0,1),(0,0,1)\}$, and so our desired probability is simply the chance that $(U_3, U_2, U_1) = (1,1,0)$, or $\frac{1}{4}$. Now, to complete the counterexample, we use a similar methodology to compute

$$\mathbb{P}[Z_2 = 2 | Z_1 = 1, Z_0 = 0] = \mathbb{P}[(U_3, U_2, U_1, U_0) = (1,1,0,0) | (U_2, U_1, U_0) = (1,0,0)]$$
$$= \frac{1}{2} \neq \frac{1}{4},$$

and so our counterexample is complete.

4 Problem 4

Let $X = (X_m)_{m\geq 0}$ be a Markov chain with state space E, so we have that, for $m \in \mathbb{N}, y, x_{m-1}, ..., x_0 \in E$,

$$\mathbb{P}[X_m = y | X_{m-1} = x_{m-1}, ..., X_0 = x_0] = \mathbb{P}[X_m = y | X_{m-1} = x_{m-1}]. \tag{1}$$

(a) TRUE. (For $r \in \mathbb{N}$). For $m, r \in \mathbb{N}$, we have that $m + r \in \mathbb{N}$, so letting $k = m+r$, we must have that from (1), replacing m with k, $(X_k)_{k\in\mathbb{N}}$ is a Markov chain. Of course, if $r \geq 0$ but $r \notin \mathbb{N}$, then $m + r \notin \mathbb{N}$, so in this case, $(X_k)_{k\in\mathbb{N}}$ will not be defined.

(b) TRUE. (TA Brandon said we may assume the chain is homogeneous.) As shown in lecture, if $(X_m)_{m\in\mathbb{N}}$ is an HMC with transition matrix P, then $(X_{2m})_{m\in\mathbb{N}}$ is an HMC with transition matrix P^2.

(c) TRUE. Let $Y_n = (X_n, X_{n+1})$ for $n \geq 0$. Then we have that

$$\mathbb{P}[Y_{n+1} = y_{n+1} | Y_n = y_n, ..., Y_0 = y_0] =$$

$$\mathbb{P}[(X_{n+1}, X_{n+2}) = (x_{n+1}, x_{n+2}) | (X_n, X_{n+1}) = (x_n, x_{n+1}), ..., (X_0, X_1) = (x_0, x_1)]$$
$$= \mathbb{P}[(X_{n+1}, X_{n+2}) = (x_{n+1}, x_{n+2}) | (X_n, X_{n+1}) = (x_n, x_{n+1})],$$

since X_{n+1} and X_n depend only on X_n and X_{n-1}, respectively, so Y_{n+1} depends only on Y_n as shown.

5 Problem 5

We will denote the process described in the problem as $X = (X_t)_{t\in\mathbb{N}}$. It is clear that X is a Markov chain because, according to the problem description, $\mathbb{P}[X_{t+1} = 2 | X_t = 1] = 1$, and for $i \geq 2, \mathbb{P}[X_{t+1} = i + 1 | X_t = i] = \mathbb{P}[X_{t+1} = i - 1 | X_t = i] = 1/2$, and we may therefore create a transition matrix P for X where for $i \geq 2, P(1, 2) = 1, \mathbb{P}(i, i+1) = \mathbb{P}(i, i-1) = 1/2$. Since these transition probabilities are independent of n, we have also that X is in fact an HMC with transition matrix $P(x, y) = P(i, i + 1), i \geq 1$.

Now, we seek the expected number of steps $N = \min\{t : X_t = n\}$ for the chain starting at 1 with a barrier at 1 to reach n. We define

$$E_{i,j} = \mathbb{E}[\text{number of steps to reach } j \text{ from } i],$$

so that

$$E_{1,2} = 1,$$
$$\mathbb{E}[N] = \sum_{i=1}^{n-1} E_{i,i+1}.$$

Now, for every step starting at $i \geq 2$, the chances are even-money that we move to either $i - 1$ or $i + 1$ in the next immediate state. If we move to $i + 1$, then clearly $E_{i,i+1} = 0$, but if we move to $i - 1$, then we must sum the expected time to first return to i and then to arrive at $i + 1$ from i. Formally,

$$E_{i-1,i+1} = E_{i-1,i} + E_{i,i+1}.$$

Putting these observations together, we compute

$$E_{i,i+1} = \mathbb{P}[X_{t+1} = i-1 | X_t = i](E_{i-1,i} + E_{i,i+1} + 1) + \mathbb{P}[X_{t+1} = i+1 | X_t = i](0+1),$$

$$E_{i,i+1} = \frac{1}{2}(E_{i-1,i} + E_{i,i+1} + 1) + \frac{1}{2}(1)$$
$$\implies \frac{1}{2}E_{i,i+1} = 1 + \frac{1}{2}E_{i-1,i}, E_{i,i+1} = 2 + E_{i-1,i}.$$

Since $E_{1,2} = 1$, we have $E_{2,3} = 3, E_{3,4} = 5, ..., E_{i,i+1} = 2i + 1, i \geq 2$. It follows that we must then compute

$$\mathbb{E}[N] = \sum_{i=1}^{n-1} E_{i,i+1} = \sum_{i=1}^{n-1}(2i - 1) = n(n-1) - (n-1) = \boxed{(n-1)^2}.$$

6 Problem 6

The HMC shown is clearly irreducible, since every entry in the transition matrix is positive. It follows that, taking any states $i, j \in E$ we have that $\mathbb{P}[X_n = j|X_{n-1} = i] > 0$, so all states in the chain communicate with one another.

Furthermore, the HMC is aperiodic, since as proved in lecture, an irreducible Markov chain will have the property that all its states will have the same period. Therefore, since $\mathbb{P}[X_{i+1} = 0|X_i = 0] > 0$, we have that state 0 has period 1, and so all states have period 1, and the chain is therefore aperiodic.

The transition diagram for this problem is attached at the end of this document.

We compute that P^3 is given by

$$P^3 = \begin{array}{c c} & \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \end{array} & \begin{pmatrix} 0.344 & 0.251 & 0.405 \\ 0.283 & 0.307 & 0.410 \\ 0.287 & 0.248 & 0.465 \end{pmatrix} \end{array}$$

From lecture, we know that if the HMC X_n has transition matrix P, then X_{3n} must have transition matrix P^3, so we may use this fact to compute

$$\mathbb{P}[X_3 = 1|X_0 = 1] = \boxed{0.307},$$

$$\mathbb{P}[X_7 = 2|X_4 = 0] = \boxed{0.465}.$$

The second computation is justified by the fact that the chain is homogeneous, and thus $\mathbb{P}[X_7 = 2|X_4 = 0] = \mathbb{P}[X_3 = 2|X_0 = 0]$.

7 Problem 7

A Markov chain $X = (X_n)_{n\in\mathbb{N}}$ is homogeneous if $\mathbb{P}[X_n = y|X_{n-1} = x] = P(x, y)$, for P the transition matrix of X, is independent of n.

For our system, if the maximum observation after $n \geq 0$ observations is $X_n = x$, it is clear that the probability we transition to some $y < x$ is 0. Indeed, to have positive transition probability, we require that $y \geq x$, and then we have that, for $y > x$, $\mathbb{P}[X_{n+1} = y|X_n = x] = P[Z_{n+1} = y]$, and for $y = x$, $\mathbb{P}[X_{n+1} = y|X_n = x] = P[Z_{n+1} \in \{0, 1, 2, ..., x\}]$. These follow because we know $X_n = x < y \implies$ we have not yet seen y, and since the Z_n terms are independent of one another (so conditioning on their own past will not change their realization probabilities), we simply ask for the probability that we observe $Z_{n+1} = y$, so our max at time $n + 1$, X_{n+1}, may be updated to be exactly y. Furthermore, for $y = x$, we are simply asking for the probability that Z_{n+1} takes on a value that is less than or equal to x, the current maximum observation thus far. Putting these conclusions together, we may write

$$\mathbb{P}[X_{n+1} = y|X_n = x] = \begin{cases} \mathbb{P}[Z_{n+1} = y] & \text{if } y > x \\ \mathbb{P}[Z_{n+1} \in \{0, 1, 2, ..., x\}] & \text{if } y = x \\ 0 & \text{else} \end{cases},$$

where for $y = x, \mathbb{P}[X_{n+1} = y|X_n = x] = \mathbb{P}[Z_{n+1} \in \{0, 1, 2, ..., x\}] = (1 - p)^0 p + (1 - p)^1 p + ... + (1 - p)^x p$ by mutual exclusion (Z_{n+1} may only take one value). It is obvious, then, that since we are told $\mathbb{P}[Z_{n+1} = k] = (1 - p)^k p$, we see that the transition probabilities of the system are independent of n, and so $(X_n)_n$ is indeed homogeneous. This function also fully describes our transition matrix for the system, a snippet of which can be written:

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ ... \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & ... \end{array} \\ \begin{pmatrix} f(0) & (1-p)p & (1-p)^2 p & ... \\ 0 & f(1) & (1-p)^2 p & ... \\ 0 & 0 & f(2) & ... \\ ... & ... & ... & ... \end{pmatrix} \end{array}$$

for

$$f(k) = \sum_{i=0}^{k} (1 - p)^i p.$$

8 Problem 8

We may capture the state of the system via the tuple (A, B, C), where A, B, and C take values in $\{0, 1\}$ and denote whether delinquents A, B, and C are alive, respectively. (A 1 will indicate that the individual is alive.)

To simplify notation, we will adopt the following state labels for each tuple:

State tuple (A, B, C)	State label
(1, 1, 1)	1
(1, 0, 1)	2
(1, 0, 0)	3
(0, 1, 1)	4
(0, 1, 0)	5
(0, 0, 1)	6
(0, 0, 0)	7

We will define our process for this system to be $X = (X_t)_{t\in\mathbb{N}}$ so that $X_t \in \{1, 2, 3, 4, 5, 6, 7\}\ \forall\ t \geq 0$. Since the probabilities a, b, c are constant in time and the system is captured entirely by (A, B, C), it is clear $(X_t)_{t\in\mathbb{N}}$ is an HMC. For this reason, we will suppress the subscript to 0 when conditioning on X_t. We define $\tau = \min\{n : X_n \in \{3, 6, 5\}\}$. Using first-step analysis, we can then compute

$$\mathbb{P}[X_\tau = 3|X_0 = 1] = (1-a)(1-b)(1-c)\mathbb{P}[X_\tau = 3|X_0 = 1]$$

$$+a(1-b)(1-c)\mathbb{P}[X_\tau = 3|X_0 = 2]+$$

$$[c(1-b)(1-a) + bc(1-a) + b(1-c)(1-a)]\mathbb{P}[X_\tau = 3|X_0 = 4]$$

$$+[ca(1-b) + bca + ab(1-c)]\mathbb{P}[X_\tau = 3|X_0 = 6]$$

$$\implies [1-(1-a)(1-b)(1-c)]\mathbb{P}[X_\tau = 3|X_0 = 1] = a(1-b)(1-c)\mathbb{P}[X_\tau = 3|X_0 = 2],$$

and we have also that

$$\mathbb{P}[X_\tau = 3|X_0 = 2] = (1-a)(1-c)\mathbb{P}[X_\tau = 3|X_0 = 2] + ac\mathbb{P}[X_\tau = 3|X_0 = 7]$$

$$+a(1-c)\mathbb{P}[X_\tau = 3|X_0 = 3] + c(1-a)\mathbb{P}[X_\tau = 3|X_0 = 6]$$

$$\implies [1-(1-a)(1-c)]\mathbb{P}[X_\tau = 3|X_0 = 2] = a(1-c)$$

$$\implies \mathbb{P}[X_\tau = 3|X_0 = 2] = \frac{a(1-c)}{1-(1-a)(1-c)},$$

and so

$$\mathbb{P}[X_\tau = 3|X_0 = 1] = \boxed{\frac{a(1-b)(1-c)}{1-(1-a)(1-b)(1-c)}\left(\frac{a(1-c)}{1-(1-a)(1-c)}\right)}$$

is the probability that A is the last one standing.

Similarly, we may now compute

$$\mathbb{P}[X_\tau = 6|X_0 = 1] = [c(1-b)(1-a)+bc(1-a)+b(1-c)(1-a)]\mathbb{P}[X_\tau = 6|X_0 = 4]$$

$$+(1-a)(1-b)(1-c)\mathbb{P}[X_\tau = 6|X_0 = 1] + a(1-c)(1-b)\mathbb{P}[X_\tau = 6|X_0 = 2]$$

$$+[ca(1-b) + bca + ab(1-c)]\mathbb{P}[X_\tau = 6|X_0 = 6],$$

$$\mathbb{P}[X_\tau = 6|X_0 = 4] = b(1-c)\mathbb{P}[X_\tau = 6|X_0 = 5] + c(1-b)\mathbb{P}[X_\tau = 6|X_0 = 6]$$

$$+(1-b)(1-c)\mathbb{P}[X_\tau = 6|X_0 = 4]$$

$$\implies \mathbb{P}[X_\tau = 6|X_0 = 4] = \frac{c(1-b)}{1-(1-b)(1-c)},$$

$$\mathbb{P}[X_\tau = 6|X_0 = 2] = (1-a)(1-c)\mathbb{P}[X_\tau = 6|X_0 = 2] + a(1-c)\mathbb{P}[X_\tau = 6|X_0 = 3]$$

$$+(1-a)c\mathbb{P}[X_\tau = 6|X_0 = 6]$$

$$\implies \mathbb{P}[X_\tau = 6|X_0 = 2] = \frac{(1-a)c}{1-(1-a)(1-c)},$$

and so

$$\mathbb{P}[X_\tau = 6|X_0 = 1] =$$

$$\boxed{\frac{a^2bc^2 - a^2b - a^2c^2 + a^2c - 2abc^2 + abc + 2ac^2 - 2ac + bc^2 - c^2}{(ac-a-c)(abc-ba-ac+a-bc+b+c)}}$$

is the probability that C is the last individual standing.

Finally, we see that

$$\mathbb{P}[X_\tau = 5|X_0 = 1] = (1-c)(1-b)(1-a)\mathbb{P}[X_\tau = 5|X_0 = 1]$$

$$+[c(1-a)b + c(1-b)(1-a) + b(1-c)(1-a)]\mathbb{P}[X_\tau = 5|X_0 = 4]$$

$$+[ab(1-c) + ac(1-b) + abc]\mathbb{P}[X_\tau = 5|X_0 = 6]$$

$$+a(1-b)(1-c)\mathbb{P}[X_\tau = 5|X_0 = 2],$$

and we have that

$$\mathbb{P}[X_\tau = 5|X_0 = 4] = (1-b)(1-c)\mathbb{P}[X_\tau = 5|X_0 = 4]$$

$$+c(1-b)\mathbb{P}[X_\tau = 5|X_0 = 6]$$

$$+cb\mathbb{P}[X_\tau = 5|X_0 = 7]$$

$$+b(1-c)\mathbb{P}[X_\tau = 5|X_0 = 5]$$

$$\implies \mathbb{P}[X_\tau = 5|X_0 = 4] = \frac{b(1-c)}{1-(1-b)(1-c)}$$

$$\implies \mathbb{P}[X_\tau = 5|X_0 = 1] = \boxed{\frac{b(1-a)(1-c)}{(1-(1-a)(1-b)(1-c))}}$$

is the probability that B is the last one standing.

9 Problem 9

We will define X_t to be the number of unique faces seen by time t and $\tau = \min\{t : X_t = 6\}$, so we seek $\mathbb{E}[\tau|X_0 = 0]$.

Since the probability of seeing any one of the six possible outcomes of a given roll is 1/6 and since we have parameterized our state space by the current count of unique numbers seen, it is clear we are working with a homogeneous process. This is further reflected by the fact that the probabilities in the transition diagram (attached at the end of the document) are not written as functions of time.

We may also confirm that, due to the homogeneity of the chain, $\mathbb{E}[\tau|X_t = n] + s = \mathbb{E}[\tau|X_{t+s} = n], s \geq 0, 0 \leq n \leq 6$. (The expected number of rolls to reach $X_t = 6$ given $X_t = n$ is time-invariant.) However, although the number of expected rolls to reach 6 from a given state will remain constant with time, τ is defined by the time index on X_t, so we will have to compensate for the additional transition (or return to) a given state by adding s if the chain has been shifted by s points in time to the expectation conditioned on the X_t of lower time index.

We now suppress the time subscript on X_t as before to define $m_n = \mathbb{E}[\tau|X_0 = n], 0 \leq n \leq 6$. Then, using the probabilities in the transition diagram as well as the time-shift property confirmed in the paragraph above, we may write

$$
\begin{aligned}
m_0 &= m_1 + 1, \\
m_1 &= \frac{1}{6}(m_1 + 1) + \frac{5}{6}(m_2 + 1), \\
m_2 &= \frac{2}{6}(m_2 + 1) + \frac{4}{6}(m_3 + 1), \\
m_3 &= \frac{3}{6}(m_3 + 1) + \frac{3}{6}(m_3 + 1), \\
m_4 &= \frac{4}{6}(m_4 + 1) + \frac{2}{6}(m_5 + 1), \\
m_5 &= \frac{5}{6}(m_5 + 1) + \frac{1}{6}(m_6 + 1), \\
m_6 &= 0.
\end{aligned}
$$

Solving this linear system of equations yields the solutions

$$
m_5 = 6, m_4 = 9, m_3 = 11, m_2 = 12.5, m_1 = 13.7
$$

$$
\implies m_0 = \boxed{14.7}.
$$

In general, if our Markov process is governed by some arbitrary transition matrix P that at least agrees with our original matrix in which probabilities are positive and which are 0, we have that, using N as defined in the problem statement,

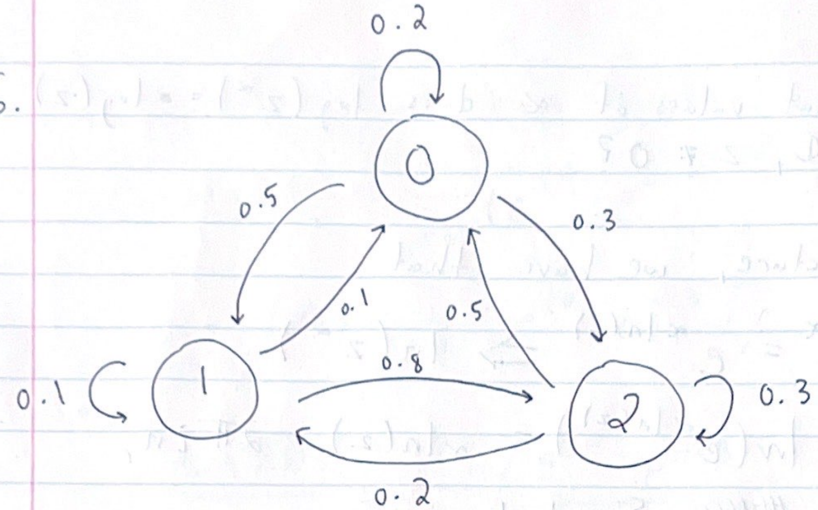
$$
\mathbb{P}[N = k] = \mathbb{P}[X_k = 6, X_{k-1} = 5|X_0 = 0]
$$

$$= \boxed{P^{k-1}(0,5)P(5,6)},$$

since we care only that two events occur for us to have that $N = k$: (1) By the kth roll, we have indeed seen all 6 numbers, and (2) The $(k-1)$st roll has only seen 5 unique numbers, implying that it is in fact the kth roll that yields all our 6 unique numbers.

Ethan Wilk

6.



9.