

# ECS 132 Homework 3

Joseph Lewis, Aaron Kaloti, Sean Malloy

due: Tuesday, February 14th, 2017, 11:59pm

## 1 Problem 1 - Bus Ridership Markov Chain

For this question, we need to prove that (6.1) from our book for Markov chains does not hold for people who alight from the bus at each stop,  $A_i$ , with (6.1):

$$P(X_{t-1} = s_{t-1} | X_t = s_t, X_{t-1} = s_{t-1}, \dots, X_0 = s_0) = P(X_{t+1} = s_{t+1} | X_t = s_t) \quad (1.1)$$

Which is saying that the probabilities of future states, given the present state and the the past state, depends only on the present state; the past is irrelevant. So for our problem, we need to prove that:

$$P(A_4 = 1 | A_3 = 0) \neq P(A_4 = 1 | A_3 = 0, A_2 = 1) \quad (1.2)$$

or any similar equation is true for our bus ridership problem. In other words, prove that the bus ridership problem of people alighting at each stop is not a Markov chain problem. We will do so by counter-example.

Let us define the states of the  $A_i$  so that we can try to match it to the Markov model. We will define the countably infinite state space of  $0, 1, 2, \dots$  to represent the number of people who get off the bus at a given stop  $i$ . Note that no-one can get off at the first stop, so  $A_1$  does not exist, and our chain starts at  $A_2$ .

Let us provide the counter-example. We have a bus which picks up 2 people at the first stop. From this point on, no one else is picked up by the bus. Note that the equation given above is inherently difficult to solve, as it depends on a lot of different factors. So instead, let us prove an equivalent statement for the sake of the Markov chain: the probability of moving from one state to the next is the same, no matter how we got to the current state. In mathematical sense, we will try to prove (1.5) for our bus which only picks up 2 passengers at the first stop (for simplicity's sake):

$$LHS = P(A_4 = 1 | A_3 = 0, A_2 = 2, B_1 = 2, B_2 = 0, B_3 = 0) \quad (1.3)$$

$$RHS = P(A_4 = 1 | A_3 = 0, A_2 = 1, B_1 = 2, B_2 = 0, B_3 = 0) \quad (1.4)$$

$$LHS \neq RHS \quad (1.5)$$

Then we can continue by the following (using the idea of “how can it happen” and the Bus Ridership example in Sec 2.12):

Begin with the *LHS*, which we know to be the probability that one person leaves the bus at the fourth stop given that two people leave at the second stop. We are already given that only 2 people get on the bus in both the LHS and RHS, so that is factored out of the equation by simple algebra. Then we know that there is no one on the bus at stop 4, so

$$LHS = P(A_4 = 1 | A_3 = 0, A_2 = 2, B_1 = 2, B_2 = 0, B_3 = 0) = 0 \quad (1.6)$$

and it follows that if only one person gets off the bus at stop 2, then 1 person remains on the bus at stop 4, so

$$RHS = P(A_4 = 1 | A_3 = 0, A_2 = 1, B_1 = 2, B_2 = 0, B_3 = 0) = 0.2 \quad (1.7)$$

Then

$$(LHS \neq RHS) \implies A_i \text{ is not properly modeled as a Markov chain} \quad (1.8)$$

And we know that equation (1.5) is true ( $0 \neq 0.2$ ), so we know that the implication of (1.8) is also true. Note that we can make this claim only because the given information about the number of passengers boarding the bus in each case is the same, as well as that the state of the system in  $A_3$  is the same. Basically, given that the states match in  $A_3$ , we are asking the same probability of the state of  $A_4$ , but we know that this state also depends on the state of  $A_2$ , and thus is not only dependent on the present ( $A_3$ ), but also the past ( $A_2$ ).

## 2 Problem 2 - 3-Heads-Total Game

(2a) For this question, we need to find the long-run proportion of games we win for the 3-heads-total game.

While the 3-heads-total game is based on the 3-heads-in-a-row game, we can no longer use the discrete Markov chain model for it, because we end the game after a set number of tosses. This means that our chances (the probability) of advancing from, say, state 2 (2 total heads) to state 0 (win the game and restart, or lose the game and restart) vary dramatically based on the prior states and how we got to them. For example:

We reach 2 total heads after our first two tosses, then we know that we either flip another head and win, advancing from state 2 to state 0, or we flip a tails and keep playing, remaining in state 2. Then,

$$P_{20} = \frac{1}{2} \quad (2.1)$$

$$P_{22} = \frac{1}{2} \quad (2.2)$$

But let us say that have 2 total heads after our first 9 tosses. There is only one toss remaining in the game, so no matter what we flip, we always advance to state 0 (restart the game with 0 total heads). So

$$P_{20} = 1 \quad (2.3)$$

$$P_{22} = 0 \quad (2.4)$$

The Markov property no longer holds - the probability of advancing from one state to another depends now on how we got to the current state, not just the state that we are in. So we cannot use the same Markov chain to model our new game, despite its seeming similarity to the three-heads-in-a-row game. We could define a Markov chain of states as tuples of (heads, tails), but this would result in the equivalent of the *easier* and *more comprehensible* method below.

*Note* : We **can** *almost* perfectly model our game as a negative binomial family among discrete distribution families. This is because in each game we are basically counting the number of trials required to achieve the 3rd head toss, which we mark as a “success”. Here we see the variable  $N$  defined as the number of trials needed to achieve the 3rd success. We must use (4.36) from our books:

$$P(N = K) = \binom{K-1}{r-1} (1-p)^{K-r} p^r \quad (2.5)$$

We use this equation so we can discover our chances of winning the game by calculating the following:

$$P(\text{win the game}) = P(N \leq 10) \quad (2.6)$$

Which is (because the events are disjoint):

$$P(N \leq 10) = \sum_{i=2}^{10} P(N = i) = 0.9453 \quad (2.7)$$

Each of these events is defined below:

$$P(N=10) = \binom{9}{2}(1 - 0.5)^{(10-3)}0.5^3 = 0.03516$$

$$P(N=9) = \binom{8}{2}(1 - 0.5)^{(9-3)}0.5^3 = 0.05469$$

$$P(N=8) = \binom{7}{2}(1 - 0.5)^{(8-3)}0.5^3 = 0.08203$$

$$P(N=7) = \binom{6}{2}(1 - 0.5)^{(7-3)}0.5^3 = 0.11719$$

$$P(N=6) = \binom{5}{2}(1 - 0.5)^{(6-3)}0.5^3 = 0.15625$$

$$P(N=5) = \binom{4}{2}(1 - 0.5)^{(5-3)}0.5^3 = 0.18750$$

$$P(N=4) = \binom{3}{2}(1 - 0.5)^{(4-3)}0.5^3 = 0.18750$$

$$P(N=3) = \binom{2}{2}(1 - 0.5)^{(3-3)}0.5^3 = 0.12500$$

We can also use the cumulative distribution function defined in R to calculate this value, but we must remember that the negative binomial distribution is defined in R as the number of failures before the k-th success, so we must subtract the number of successes we are looking for to get the total number of failures when using the R defined functions. Thus

$$P(N \leq 10) = \text{pnbinom}(10 - 3, 3, 0.5) = 0.9453 \quad (2.8)$$

**(2b)** For this part, we need to find the long-run average of our winnings from playing the game many times. Let us model our winnings by defining the random variable  $W$  as follows:  $W = 11 - N$  if we win the game, and  $W = 0$  if we lose. Then we can find the expected value of winnings as follows (we will calculate the expected value of  $N$  when we win the game in part **2c**):

$$\begin{aligned} E(W) &= P(\text{win the game}) \cdot E(11 - N) + P(\text{lose the game}) \cdot E(0) \quad (2.9) \\ &= 0.9453 \cdot (11 - E(N)) \\ &= 0.9453 \cdot (11 - 5.628) \\ &= 0.9453 \cdot (5.372) \\ &= 5.08 \end{aligned}$$

So from this, we are expected to win \$5.08 per game! *Note* : We cannot actually expect to get this number in any one game, but rather it is the long-run per-game average of winnings from a series of infinitely played games. The value itself, like many other expected values in other problems, is unachievable.

(2c) For this part, we need to find the expected value of the number of tosses required to win. We will use the following value for our calculations in part 2b as well, so that math that follows verifies that section.

*Note* : We cannot simply use the expected value of a standard negative binomial distribution because our game ends after 10 tosses, whether we win or not, so we must consider the expected value of the number of tosses required to win to be relative to the possible ways in which we can win. This is why we say that our game *almost* perfectly is modeled by the negative binomial distribution.

$$\begin{aligned}
E(N) &= P(N = 3 | \text{win the game}) \cdot 3 + P(N = 4 | \text{win the game}) \cdot 4 + \dots \quad (2.10) \\
&= \sum_{i=3}^{10} \frac{P(N=i) \cdot i}{P(\text{win the game})} \\
&= \frac{\sum_{i=3}^{10} \text{dnbinom}(i - 3, 3, 0.5) \cdot i}{P(\text{win the game})} \\
&= \frac{5.320}{0.9453} \\
&= 5.628
\end{aligned}$$

*Note* : We can make the leap that  $E(N)$  goes straight to our summation in (2.10) by noting that:

From the book (2.8):

$$P(N = 3 | \text{win the game}) = \frac{P(\text{win the game and } N = 3)}{P(\text{win the game})} \quad (2.11)$$

From the book (2.7)

$$P(\text{win the game and } N = 3) = P(N = 3) \cdot P(\text{win the game} | N=3) \quad (2.12)$$

*Note* : We cannot lose if  $N = 3$ , so:

$$P(\text{win the game} | N = 3) = 1 \quad (2.13)$$

$$(2.13) \implies P(\text{win the game and } N = 3) = P(N = 3) \cdot 1 = P(N = 3) \quad (2.14)$$

Then we have the solution:

$$P(N = 3 | \text{win the game}) = \frac{P(N = 3)}{P(\text{win the game})} \quad (2.15)$$

Then the distributive property (of standard mathematics) allows us to simply divide the sum of all these  $P(N = i)$  by the  $P(\text{win})$ , as long as each  $i$  is a guaranteed winning case (Which they are from 3 to 10)!