Week 4 HW, Stochastic Process 25-26

Srikanth Pai,Asst. Professor, MSE

August 10, 2025

1. Given a DTMC on a state space $\mathcal{S}$, let $j \in \mathcal{S}$. Define M_j as the number of visits to state j. Let f_{jj} denote the probability that the process will return to state j in the future given that the process starts from j. Prove that for $m \geq 0$,
$$\Pr\{M_j = m \mid X_0 = j\} = f_{jj}^m(1 - f_{jj})$$
[We have already proved this in class for $m = 0, 1$. Prove by induction for the rest.]

2. Consider an finite DTMC $\{X_n\}$ and fix a state i. Define:
$$T_1 = \min\{n \geq 1 : X_n = i | X_0 = i\} \tag{1}$$
$$T_2 = \min\{n > T_1 : X_n = i | X_0 = i\} \tag{2}$$
$$T_3 = \min\{n > T_2 : X_n = i | X_0 = i\} \tag{3}$$
Let $W_1 = T_1$, $W_2 = T_2 - T_1$, $W_3 = T_3 - T_2$.
   (a) Show that W_1 and W_2 have the same distribution.
   (b) Prove that W_1 and W_2 are independent. [Hint: Strong Markov property.]
   (c) Conclude W_1, W_2, W_3 are i.i.d.

3. Manish and Medha bet 1 hundred rupees in each round. In other words, Medha wins 1 hundred rupees each round with probability p and loses 1 hundred rupees with probability $1 - p$. Assume that Manish starts the game with 1 hundred rupees and Medha starts with 2 hundred rupees. The game ends when one of them goes bankrupt. Let M_n denote Medha's fortune (in units of hundred rupees) at the end of round n and let M_0 denote Medha's initial fortune.
   (a) What is the range of the variable M_n for any n?
   (b) Show that M_n is a DTMC. Draw the state diagram of the Markov chain.
   (c) List the equivalence classes of communicating states with proof.
   (d) In terms of the notations in the class, what is f_{33}?
   (e) Compute f_{23}, f_{32} using recursive methods.
   (f) What is the probability that Medha bankrupts Manish if she starts with 2 hundred rupees and Manish starts with 1 hundred rupees?
   (g) Let T be the number of rounds played. Let $u_1 = \mathbb{E}(T \mid X_0 = 1)$ and $u_2 = \mathbb{E}(T \mid X_0 = 2)$. Using the technique employed in part (e), show that
$$u_1 = pu_2 + 1, \;\; u_2 = (1 - p)u_1 + 1.$$
   Now compute $\mathbb{E}(T)$ using u_1 and u_2.As an alternative, directly compute $\mathbb{E}(T)$ by analyzing the chain as well.

4. **Background:** Using data from the National Sample Survey Office (NSSO) spanning 1983-2012, Reddy (2015) studied intergenerational occupational mobility in India.

   Occupational Classification: Based on the National Classification of Occupations (NCO-2004), individuals were grouped into four categories by skill level and socio-economic status:

(a) **White Collar:** Legislators, senior officials, managers, professionals, and associate professionals (highest skill/income)

(b) **Skilled Workers:** Clerks, service workers, sales workers, craft workers, machine operators (moderate skill)

(c) **Farmers:** Self-employed in agriculture (heterogeneous group, but distinct from agricultural wage workers)

(d) **Unskilled Workers:** Elementary occupations requiring minimal training (lowest skill/income)

While not strictly hierarchical, Table 3 in Reddy (2015) confirms the relative economic positions: white collar workers had the highest median income and education, while unskilled workers had the lowest.

Data and Estimation: The transition matrix below was estimated from approximately 25,000 father-son pairs in the 2011-12 NSSO Employment & Unemployment Survey. Each entry p_{ij} represents the empirical probability that a son enters occupation j given that his father was in occupation i, calculated as the proportion of sons in each destination category for each father's occupation group.

The following transition matrix shows the probability that a son enters a particular occupation given his father's occupation (based on Table 5 in the paper):

Father's → Son's	**White Collar**	**Skilled**	**Farmer**	**Unskilled**
White Collar	0.603	0.279	0.054	0.064
Skilled	0.090	0.750	0.053	0.107
Farmer	0.057	0.126	0.705	0.112
Unskilled	0.028	0.212	0.022	0.737

1. Since this is a real world problem, you may use a calculator/computer to calculate.

   (a) Verify this is a valid transition matrix. List the communicating classes.

   (b) Classify each state as transient or recurrent. Justify your answer.

   (c) What does the classification tell us about Indian society?

   (d) Calculate P^2 and P^3. What is the probability that a farmer's grandson becomes a white-collar worker?

   (e) Starting from the unskilled category, what is the probability of reaching white-collar status within exactly 3 generations?

   (f) Compute P^{10} and P^{20}. What do you observe about the rows as n increases? What might this represent in terms of long-term occupational distribution?

   (g) An invariant distribution π is a pmf on the states with the property that it does not change over time. Compute the invariant distribution from the tpm and compare it to your observations in the previous problem.

   (h) Reddy's study shows that Scheduled Castes and Scheduled Tribes (SC/ST) face lower mobility. Suppose their transition probabilities from the unskilled category are:

   $$P^{SC/ST}_{unskilled \to unskilled} = 0.85, \quad P^{SC/ST}_{unskilled \to white} = 0.01$$

   while keeping other transition probabilities proportionally adjusted.

      i. Construct the modified transition matrix for SC/ST populations.

      ii. Compare P^{10} for both populations. What does this reveal about inequality?

   (i) The government introduces a skill development program that increases $P_{unskilled \to skilled}$ from 0.212 to 0.318 (50% increase), with a corresponding decrease in $P_{unskilled \to unskilled}$. Calculate the new transition matrix.

   (j) Compare the long-term behavior (large powers of P) before and after this policy intervention. Is this policy effective?