§2.1, 2.2 Markov chains introduction

Law of Total Probability:

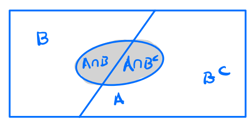
Given events A and B, with $P(B) \neq 0$,

$$P(A) = P(A \mid B)P(B) + P(A \mid B^c)P(B^c).$$

More generally, if $B_1, B_2, B_3, \ldots, B_n$ are disjoint events that make up the whole sample space, with $P(B_1), \ldots, P(B_n) \neq 0$,

$$P(A) = P(A \mid B_1)P(B_1) + \cdots + P(A \mid B_n)P(B_n).$$

Idea of proof

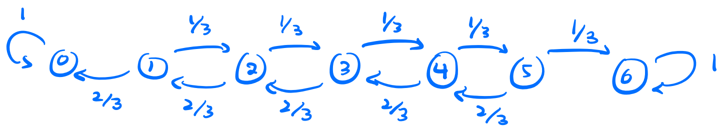


$$P(A) = P(A \cap B) + P(A \cap B^c)$$

$$= P(A \mid B)P(B) + P(A \mid B^c)P(B^c) \quad \text{since } P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

Try the general version for yourself.

Example A gambler plays a game where they win \$1 each round with probability $1/3$ and lose \$1 with probability $2/3$. They stop playing if their net winnings ever reach \$0 or \$6. They start the game with \$1, 2, 3, 4, or 5 uniformly at random. Draw the transition state diagram and find the probability their net winnings are \$3 after 1 round.



$$P(X_1=3) = \sum_{k=1}^{5} P(X_1=3 \mid X_0=k)\,P(X_0=k)$$

$$= P(X_1=3 \mid X_0=4)\,P(X_0=4) + P(X_1=3 \mid X_0=2)\,P(X_0=2)$$

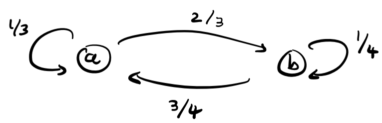
$$= (2/3)(1/5) + (1/3)(1/5) = 1/5$$

Conditional Law of Total Probability

$$P(A \mid C) = P(A \mid BC)\,P(B \mid C) + P(A \mid B^c C)\,P(B^c \mid C)$$

More generally, $P(A \mid C) = P(A \mid B_1 C)\,P(B_1 \mid C) + \dots + P(A \mid B_n C)\,P(B_n \mid C)$

Example Consider the Markov chain with transition state diagram below.



Find $P(X_2 = a \mid X_0 = a)$

for a Markov chain, older part of history is irrelevant

$$= P(X_2=a \mid X_1=a, \cancel{X_0=a})\,P(X_1=a \mid X_0=a)$$

$$+ P(X_2=a \mid X_1=b, \cancel{X_0=a})\,P(X_1=b \mid X_0=a)$$

$$= (1/3)^2 + (3/4)(2/3)$$

Def Let S be a discrete set (so S is finite o countably infinite). A <u>discrete-time</u>, <u>discrete-state</u> <u>Markov chain</u> is a sequence of random variables $(X_0, X_1, X_2, \dots)$ with the property

$$P(X_{n+1} = j \mid X_0 = x_0, X_1 = x_1, \dots, X_{n-1} = x_{n-1}, X_n = i)$$

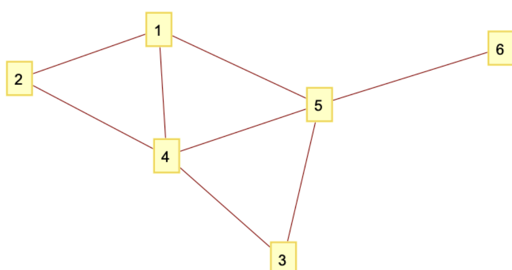
$$= P(X_{n+1} = j \mid X_n = i)$$

for any time $n \geq 0$ and any states $i, j, x_0, \dots, x_{n-1} \in S$.

This is called the <u>Markov property</u>. A Markov chain is called <u>time-homogeneous</u> if

$$P(X_{n+1} = j \mid X_n = i) = P(X_1 = j \mid X_0 = i)$$

for all times $n \geq 0$.



Problem 1. Try the following exercises using the graph shown above.

a. Find the transition matrix P.

b. What do the following conditional probabilities mean in words? How many time units elapse? Between which states do you transition? What does the notion of *time-homogeneity* tell us about parts b and c?

1. $P(X_2 = 3 \mid X_0 = 1)$
2. $P(X_7 = 5 \mid X_4 = 4)$
3. $P(X_{50} = 5 \mid X_{40} = 2)$

c. Try computing the probability in part 1 above using the Conditional Law of Total Probability. Wait until after doing Problem 2 before you think about the other probabilities.

(a)

$$P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \begin{bmatrix} 0 & 1/3 & 0 & 1/3 & 1/3 & 0 \\ 1/2 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 \\ 1/4 & 1/4 & 1/4 & 0 & 1/4 & 0 \\ 1/4 & 0 & 1/4 & 1/4 & 0 & 1/4 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

(b) (1) prob. of transitioning from 1 to 3 in 2 time units

(2) time-homogeneity $\Rightarrow$ $P(X_7 = 5 \mid X_4 = 4) = P(X_3 = 5 \mid X_0 = 4)$

prob. of transitioning from 4 to 5 in 3 time units

(3) time-homogeneity $\Rightarrow$ $P(X_{50} = 5 \mid X_{40} = 2) = P(X_{10} = 5 \mid X_0 = 2)$

prob. of transitioning from 2 to 5 in 10 time units

(c) $$P(X_2 = 3 \mid X_0 = 1) = \sum_{k=1}^{6} P(X_2 = 3 \mid X_1 = k, X_0 = 1) P(X_1 = k \mid X_0 = 1)$$

$$= \sum_{k=1}^{6} P(X_2 = 3 \mid X_1 = k) P(X_1 = k \mid X_0 = 1)$$

$$= \sum_{k=1}^{6} P_{k3} P_{1k}$$

$$= \sum_{k=1}^{6} P_{1k} P_{k3}$$

= dot product of row 1 and column 3 of the transition matrix P

$= (P^2)_{13}$ (ie. the (1,3) entry of P^2)

$= (0)(0) + (1/3)(0) + (0)(0) + (1/3)(1/4) + (1/3)(1/4) + (0)(0)$

$= 1/6$

Problem 2. Let's think more generally now. Suppose we're working with a Markov chain whose state space is $\mathcal{S} = \{1, 2, \ldots, m\}$. Its transition matrix P is an $m \times m$ matrix and its ij-entry is $P_{ij} = P(X_1 = j \mid X_0 = i)$. This is all given to us. Try using the Conditional Law of Total Probability to write an expression for

$$P(X_2 = j \mid X_0 = i),$$

called the 2-step transition probability, in terms of the entries of the transition matrix P. Your expression should be a summation and its terms should be written using entries of the matrix P. Do you recognize this expression as something related to matrix algebra? If you figure these questions out, try thinking about $P(X_3 = j \mid X_0 = i)$, called the 3-step transition probability.

$$P(X_2 = j \mid X_0 = i) = (P^2)_{ij}$$

$$P(X_3 = j \mid X_0 = i) = (P^3)_{ij}$$

(prove this)