Introduction to DTMC

STA 624, Spring 2016

Ruriko Yoshida
Dept. of Statistics University of Kentucky

polytopes.net

Some example...

There are your favorite bars, $S = \{A, C, G, T\}$, in Lexington. You and your friends, Arne, Connie, Dave, decide to go to bar hopping one night. Since you are already drunk, you and your friends decide where you and your friends are going next by rollong a four faced die (tetrahedron).

Each of you and your friends has a four faced die. You will roll your die if you and your friends are currently at Bar A, Arne rolls his die if you and your friends are currently at Bar C, Connie rolls her die if you and your friends are currently at Bar G, and Dave rolls his die if you and your friends are currently at Bar T.

Each die has different weights on each face.

Some example...

The probability of obtaining each letter differs depending on which die you are rolling:

	A	C	G	T
Your die	P_{AA}	P_{AC}	P_{AG}	P_{AT}
Arne's die	P_{CA}	P_{CC}	P_{CG}	P_{CT}
Connie's die	P_{GA}	P_{GC}	P_{GG}	P_{GT}
Dave's die	P_{TA}	P_{TC}	P_{TG}	P_{TT}

Where P_{xy} for any x, y in S, $P_{AA} + P_{AC} + P_{AG} + P_{AT} = 1$, $P_{CA} + P_{CC} + P_{CG} + P_{CT} = 1$, $P_{GA} + P_{GC} + P_{GG} + P_{GT} = 1$, and $P_{TA} + P_{TC} + P_{TG} + P_{TT} = 1$.

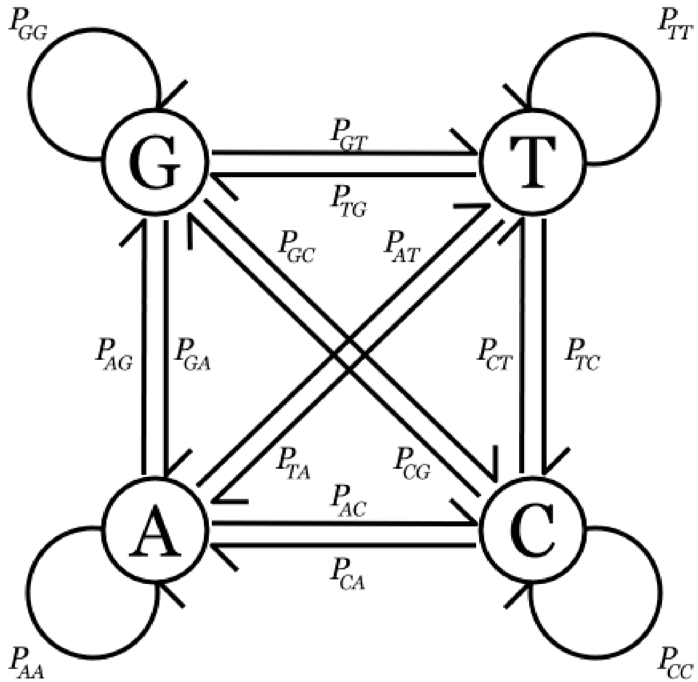
Some example...

Here is a specific example:

	A	C	G	T
Your die	$1/4$	$1/4$	$1/4$	$1/4$
Arne's die	$1/5$	$1/5$	$2/5$	$1/5$
Connie's die	$1/3$	$1/3$	$1/6$	$1/6$
Dave's die T	$1/6$	$1/3$	$1/3$	$1/6$

We can describe this process by drawing a picture...

Some example...



This is an example of **Discrete Time Markov process**.

Some definitions on MC

Definition A **discrete time stochastic process** is a collection of random variables $\{X_0, X_1, X_2, \ldots\}$ defined on a common sample space and state space S which depends on time $n = 0, 1, 2, \cdots$.

Definition A **discrete time Markov process** is a discrete time stochastic process $\{X_n\}_{n=0}^{\infty}$ which satisfies the **Markov property**, that is, for all $n \in \{0, 1, \ldots\}$ and any states $x_0, x_1, \cdots x_n, y \in S$,

$$P(X_{n+1} = y | X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n) = P(X_{n+1} = y | X_n = x_n).$$

Definition A **time homogeneous Markov process** $\{X_n\}_{n=0}^{\infty}$ is a stochastic process such that for all $n \in \{0, 1, \ldots\}$ and any states $x, y \in S$,

$$P(X_{n+1} = y | X_n = x) = P(X_1 = y | X_0 = x).$$

Finite State MC

Definition Given a Markov chain with finite state space S, a **transition matrix** is a matrix whose entry in the ith row and jth column is

$$P(X_{t+1} = j | X_t = i).$$

Definition Given a Markov chain with finite state space S, the **transition graph** has vertex set S, and has directed edges (i, j) with weight

$$P(X_{t+1} = j | X_t = i)$$

whenever this weight is positive.

Definition We call a Markov process **Markov chain** if we can describe the process as the transition graph.

Go back to our example...

Since the probability where you are going next depends only on the place where you are currently at, this process satisfies Markov property.

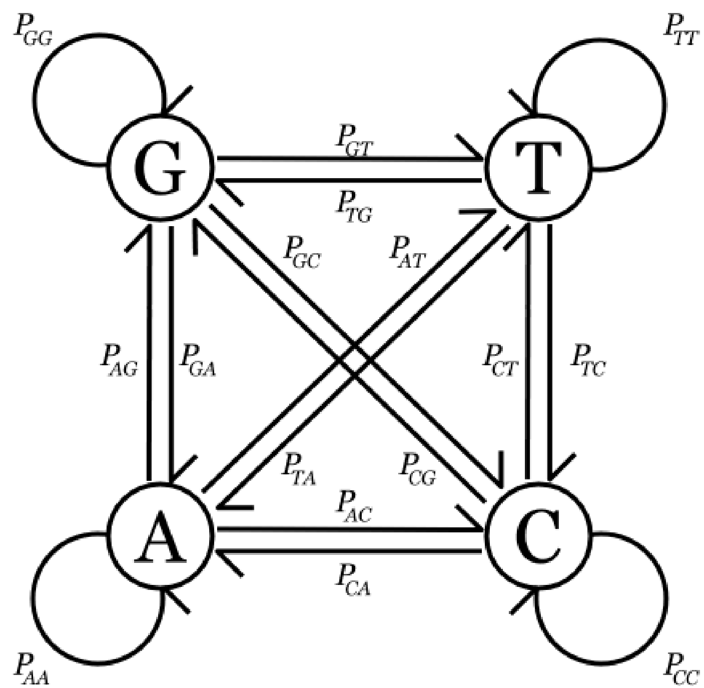
Since the probability where you are going next is not affected by the time this process is time homogeneous.

The transition martix of our Markov chain is:

	A	C	G	T
Your die	$1/4$	$1/4$	$1/4$	$1/4$
Arne's die	$1/5$	$1/5$	$2/5$	$1/5$
Connie's die	$1/3$	$1/3$	$1/6$	$1/6$
Dave's die T	$1/6$	$1/3$	$1/3$	$1/6$

Some example...

The transition graph is:



This is an example of **Discrete Time Markov chain**.

Question

Question You and your friends decide the initial bar by rolling a fair die (X_0). Then You and your friends move twice ($n = 2$). What is the probability that you and your friends have been at Bar C at $n = 0$, at Bar T at $n = 1$, and at Bar A at $n = 2$?

Review Suppose A_1, A_2 are events such that $P(A_2) \neq 0$. Then the **conditional probability** of event A_1 given A_2 is:

$$P(A_1|A_2) = \frac{P(A_1A_2)}{P(A_2)}.$$

Thus we have

$$P(A_1|A_2)P(A_2) = P(A_1A_2).$$

Computing probability

By this way we can compute the probability of a path, $x_0 x_1 x_2 \cdots x_n x_{n+1}$, in a graph by:

$$\begin{array}{rl} & P(X_{n+1} = x_{n+1}, X_n = x_n, \cdots, X_1 = x_1, X_0 = x_0) \\ = & P(X_{n+1} = x_{n+1} | X_n = x_n, \cdots, X_1 = x_1, X_0 = x_0) \\ & \times P(X_n = x_n | X_{n-1} = x_{n-1}, \cdots, X_1 = x_1, X_0 = x_0) \\ & \times \cdots \times P(X_1 = x_1 | X_0 = x_0) \times P(X_0). \end{array}$$

Note that this is a discrete time Markov chain so it satisfies Markov property, thus we have:

$$\begin{array}{rl} & P(X_{n+1} = x_{n+1}, X_n = x_n, \cdots, X_1 = x_1, X_0 = x_0) \\ = & P(X_{n+1} = x_{n+1} | X_n = x_n) \times P(X_n = x_n | X_{n-1} = x_{n-1}) \\ & \times \cdots \times P(X_1 = x_1 | X_0 = x_0) \times P(X_0). \end{array}$$

Go back to the example...

We want to compute the probability that you and your friends have been at Bar C at $n = 0$, at Bar T at $n = 1$, and at Bar A at $n = 2$. So we want to compute

$$P(X_2 = A, X_1 = T, X_0 = C).$$

By using conditional probability and Markov property we have

$$\begin{aligned} & P(X_2 = A, X_1 = T, X_0 = C) \\ = \; & P(X_2 = A | X_1 = T) P(X_1 = T | X_0 = C) P(X_0 = C) \end{aligned}$$

Now use the transition martix of our Markov chain:

	A	C	G	T
Your die A	$1/4$	$1/4$	$1/4$	$1/4$
Arne's die C	$1/5$	$1/5$	$2/5$	$1/5$
Connie's die G	$1/3$	$1/3$	$1/6$	$1/6$
Dave's die T	$1/6$	$1/3$	$1/3$	$1/6$

Then we have:

$$\begin{aligned} & P(X_2 = A, X_1 = T, X_0 = C) \\ = \; & P(X_2 = A | X_1 = T) P(X_1 = T | X_0 = C) P(X_0 = C) \\ = \; & P_{TA} \times P_{CT} \times P(X_0) \\ = \; & 1/6 \times 1/5 \times 1/4 \end{aligned}$$

How about...

Question You and your friends decide the initial bar by rolling a fair die (X_0). Then You and your friends move five times ($n = 5$). What is the probability that you and your friends have been at Bar A at $n = 5$?

Fact $P(X_n = y|X_0 = x)$ for states x, y is the (x, y)th element of the matrix P^n where P is the transition matrix.

P^5 is:

0.24352	0.27621	0.28410	0.19617
0.24346	0.27611	0.28421	0.19621
0.24357	0.27630	0.28400	0.19613
0.24352	0.27622	0.28410	0.19616

Thus the answer is

$$
\begin{array}{rl}
 & P(X_5 = A) \\
= & \sum_{x \in S} P(X_5 = A | X_0 = x) P(X_0 = x) \\
= & 0.24352 * 1/4 + 0.24346 * 1/4 + 0.24357 * 1/4 + 0.24352 * 1/4 \\
= & 0.24352
\end{array}
$$

Definition π is a **stationary** or **invariant** distribution for a Markov chain if $X_t \sim \pi$ implies that $X_{t+1} \sim \pi$ (i.e. $\sum_{y \in S} \pi_y p(y, x) = \pi_x$).

Definition For countable state Markov chains, if

$$\sum_{x \in S} \pi_x P(X_{t+1} = y | X_t = x) = \pi_y$$

then π is a stationary distribution. These are called the **balance equations.**

Definition π is a **limiting** distribution for a countable state Markov chain if

$$\lim_{t \to \infty} P(X_t = i | X_0 = j) = \pi_i,$$

for all states i and j.

Note. The limiting distribution is a stationary distribution but not always true that a stationary distribution is the limiting distribution.

I will show a proof for a fact that the limiting distribution is a stationary distribution.

Example. $S = \{1, 2, 3\}$. If the transition matrix $P = I_3$, then any distribution is stationary but it does not have to be the limiting distribution.

Exercise

Question 1. Go back to the bar hopping example. Suppose we have the transition matrix

	A	C	G	T
Your die A	$1/4$	$1/4$	$1/4$	$1/4$
Arne's die C	$1/5$	$1/5$	$2/5$	$1/5$
Connie's die G	$1/3$	$1/3$	$1/6$	$1/6$
Dave's die T	$1/6$	$1/3$	$1/3$	$1/6$

and suppose you choose the initial bar at random (i.e., probability $1/4$ for going to each bar). If you and your friends go to the bar and record which bar you go, you observe the sequence $ACGTTTCGA$. What is the probability to observe the sequence.

Question 2. Compute $P(X_{25} = A)$? Please note that you do not want to compute by hand....

Exercise

Question 3. Also compute a stationary distribution π. Is it unique and the same as the limiting distribution π?