UPPSALA UNIVERSITY
Department of Mathematics

Markov Processes, 1MS012
Spring semester 2024

Lecture 14 Markov Processes, 1MS012

1 Some applications

1.1 Card shuffling

$X_n =$ the order of a deck of cards after shuffling n times. State space S the set of permutations of $1, 2, ..., N$ where N is the number of cards.
$|S| = N!$

Example: Shuffling dynamics: take top card and put it at a random position. (X_n) is clearly irreducible and aperiodic since if $x = (x_1, ..., x_N)$ and $y = (y_1, ..., y_N)$ are 2 permutations, then it is possible to go from x to y in N shuffles, and it is also possible to go from a state to itself if the upper most card is chosen.

Observe that all cards under the original bottom card are equally likely to be in any order. The deck is thus in a completely random order when the original bottom card is first reinserted.
Let T be the time when this happens.
$T = T_1 + ... + T_N$ where T_i is the time it takes for the original bottom card to move from position i from below one step upwards for $i = 1, ..., N-1$, and $T_N = 1$. Since T_i is geometrically distributed with parameter i/N it follows that $E(T_i) = N/i$. Thus

$$E(T) = \sum_{i=1}^{N} E(T_i) = N \sum_{i=1}^{N} \frac{1}{i} \sim N \log N.$$

1.2 Google PageRank: Decide the order search results are displayed.

In 1998 Larry Page and Sergey Brin at Stanford university founded Google Inc., the company behind the Google search engine.

Let W be the set of webpages that can be reached by Google. W is stored in a database of size $n = |W|$.

Each of the n pages are regularly ranked with respect to popularity measured by the degree of links from other pages.

To each query submitted to the search engine, Google displays those pages that match the query in order of page ranking.

The basic idea behind the Google search engine is roughly the following: Consider an imaginary internet surfer randomly clicking on links. We rank webpages according to the probability distribution on W describing the location of the surfer after a huge number of clicks.

Basic Markov model:

The set W can be regarded as the set of nodes in a directed graph where we put a directed edge between two nodes (web-pages) v_1 and v_2 if there is a hyperlink from v_1 to v_2.

Consider a Markov chain, (X_n), with state space W obtained by the following dynamics; If $X_n = v \in W$, then, with probability $p = 0.85$ we pick one of the outgoing links of v at random with equal probabilities, and with probability $1-p$ we move to a randomly chosen page (among the n pages). (If there are no outgoing links of v, i.e. if v is a **dangling node** then one remedy to make this a well defined Markov chain is to then always choose a random page.)

This Markov chain has finite state space and is irreducible and aperiodic and thus there exists a unique stationary distribution.

Pages are ranked according to the stationary probabilities for this Markov chain.

Other choices of parameter values p are of course also possible. The natural choice of $p = 1$ (corresponding to the rough idea described above) leads to many problems, since the long run behavior of the Markov chain will then depend on the initial state. A small value of p gives quick convergence, but makes it hard to distinguish differences in rankings.

Some major advantages/disadvantages of the method:
Advantage: Rather hard to manipulate ranking of a single website.
(A way to manipulate it is e.g. to build *link farms* i.e. a big set of pages where every page have links to every other page in the set.)
Disadvantage: Favors older webpages.

1.3 Calculating integrals by using a coin

Suppose we want to calculate

$$\int_a^b f(x)dx,$$

for some continuous function f where a and b are real numbers.

Our tool: A coin.

Assume w.l.o.g. that $a = 0$ and $b = 1$. Consider random iterations with $w_0(x) = x/2$, and $w_1(x) = x/2 + 1/2$ chosen with probability $1/2$ each, i .e. let

$$X_{n+1} = w_{I_{n+1}}(X_n),$$

where (I_n) is i.i.d. with $P(I_n = 0) = P(I_n = 1) = 1/2$. Let $X_0 = x_0$ be an arbitrary point in $[0, 1]$.

(X_n) is a Markov chain.

We claim that the average of f along a trajectory, $\frac{\sum_{k=0}^{N-1} f(X_k)}{N}$, converges to $\int_0^1 f(x)dx$, as $N \to \infty$, (with probability one).

In order to motivate why the claim holds, divide the interval according to the base 2-expansion; Let

$$\Delta_{i_1,\ldots,i_k} = (i_1 2^{-1} + i_2 2^{-2} + \ldots + i_k 2^{-k}, i_1 2^{-1} + i_2 2^{-2} + \ldots + (i_k + 1)2^{-k}).$$

If for any fixed k, $x \in \Delta_{i_1,\dots,i_k}$, then $w_0(x) \in \Delta_{0i_1,\dots,i_{k-1}}$, and $w_1(x) \in \Delta_{1i_1,\dots,i_{k-1}}$.

Thus if Y_n is the interval where X_n belongs then (Y_n) is an irreducible aperiodic Markov chain with state space $S = \{\Delta_{i_1,\dots,i_k}; i_j \in \{0,1\}, 1 \le j \le k\}$. By symmetry Y_n will be uniformly distributed on S as $n \to \infty$.

We have

$$\frac{\sum_{n=0}^{N-1} f_k(Y_n)}{N} \to \sum f_k(\Delta_{i_1,\dots,i_k})|\Delta_{i_1,\dots,i_k}| = \int_0^1 f_k(x)dx$$

as $N \to \infty$ with probability one, if f_k is a constant function on each interval $\Delta_{i_1,\dots,i_k}$.

Thus also

$$\frac{\sum_{n=0}^{N-1} f_k(X_n)}{N} \to \sum f_k(\Delta_{i_1,\dots,i_k})|\Delta_{i_1,\dots,i_k}| = \int_0^1 f_k(x)dx.$$

By letting $k \to \infty$ (if $f_k \to f$), we get

$$\frac{\sum_{n=0}^{N-1} f(X_n)}{N} \to \int_0^1 f(x)dx,$$

as $N \to \infty$ with probability one.

1.4 Yahtzee

Yahtzee is a popular dice game. The object of the game is to score the most points by rolling five dice to make certain combinations. On each turn, a player gets up to three rolls of the dice. He or she can save any dice that are wanted to complete a combination and then re-roll the other dice. Suppose a player wants as many "sixes" as possible. We can model the number of "sixes" after n rolls with a Markov chain.

Let X_n denote the number of sixes the player has after n rolls. We are interested in finding the distribution of X_3 given that $X_0 = 0$. If we have j sixes after n rolls then the number of new sixes after $n+1$ rolls will be binomially distributed with parameters $5 - j$ and $1/6$. Thus (X_n) is a Markov chain with state space $S = (0, 1, 2, 3, 4, 5)$ with $(X_{n+1} - j | X_n = j) \sim \mathrm{Bin}(5-j, 1/6)$, i.e.

$$P(X_{n+1} = i + j | X_n = j) = \binom{5-j}{i} (\frac{1}{6})^i (\frac{5}{6})^{5-i-j}, \quad 0 \le i \le 5 - j,$$

so

$$p_{jk} = P(X_{n+1} = k | X_n = j) = \binom{5-j}{k-j} (\frac{1}{6})^{k-j} (\frac{5}{6})^{5-k}, \quad j \leq k \leq 5,$$

i.e. (X_n) has transition matrix

$$\mathbf{P} = \begin{pmatrix} \binom{5}{0}(\frac{1}{6})^0(\frac{5}{6})^5 & \binom{5}{1}(\frac{1}{6})^1(\frac{5}{6})^4 & \binom{5}{2}(\frac{1}{6})^2(\frac{5}{6})^3 & \binom{5}{3}(\frac{1}{6})^3(\frac{5}{6})^2 & \binom{5}{4}(\frac{1}{6})^4(\frac{5}{6})^1 & \binom{5}{5}(\frac{1}{6})^5(\frac{5}{6})^0) \\ 0 & \binom{4}{0}(\frac{1}{6})^0(\frac{5}{6})^4 & \binom{4}{1}(\frac{1}{6})^1(\frac{5}{6})^3 & \binom{4}{2}(\frac{1}{6})^2(\frac{5}{6})^2 & \binom{4}{3}(\frac{1}{6})^3(\frac{5}{6})^1 & \binom{4}{4}(\frac{1}{6})^4(\frac{5}{6}^0) \\ 0 & 0 & \binom{3}{0}(\frac{1}{6})^0(\frac{5}{6})^3 & \binom{3}{1}(\frac{1}{6})^1(\frac{5}{6})^2 & \binom{3}{2}(\frac{1}{6})^2(\frac{5}{6})^1 & \binom{3}{3}(\frac{1}{6})^3(\frac{5}{6})^0 \\ 0 & 0 & 0 & \binom{2}{0}(\frac{1}{6})^0(\frac{5}{6})^2 & \binom{2}{1}(\frac{1}{6})^1(\frac{5}{6})^1 & \binom{2}{2}(\frac{1}{6})^2(\frac{5}{6})^0 \\ 0 & 0 & 0 & 0 & \binom{1}{0}(\frac{1}{6})^0(\frac{5}{6})^1 & \binom{1}{1}(\frac{1}{6})^1(\frac{5}{6})^0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\approx \begin{pmatrix} 0.4019 & 0.4019 & 0.1608 & 0.0322 & 0.0032 & 0.0001 \\ 0 & 0.4823 & 0.3858 & 0.1157 & 0.0154 & 0.0008 \\ 0 & 0 & 0.5787 & 0.3472 & 0.0694 & 0.0046 \\ 0 & 0 & 0 & 0.6944 & 0.2778 & 0.0278 \\ 0 & 0 & 0 & 0 & 0.8333 & 0.1667 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Thus

$$\mathbf{P}^3 \approx \begin{pmatrix} 0.0649 & 0.2363 & 0.3440 & 0.2504 & 0.0912 & 0.0133 \\ 0 & 0.1122 & 0.3266 & 0.3566 & 0.1731 & 0.0315 \\ 0 & 0 & 0.1938 & 0.4233 & 0.3081 & 0.0748 \\ 0 & 0 & 0 & 0.3349 & 0.4876 & 0.1775 \\ 0 & 0 & 0 & 0 & 0.5787 & 0.4213 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

so in particular

$$P(X_3 = k | X_0 = 0) \approx \begin{cases} 0.0649, & k = 0 \\ 0.2363, & k = 1 \\ 0.3440, & k = 2 \\ 0.2504, & k = 3 \\ 0.0912 & k = 4 \\ 0.0133 & k = 5 \end{cases}.$$

(This can also be seen without involving Markov chains in the following way: The probability of not getting a six on a given dice in 3 turns is $(5/6)^3$ independently of the other dice. Thus $(X_3 | X_0 = 0) \sim \mathrm{Bin}(5, 1 - (5/6)^3) = \mathrm{Bin}(5, \frac{91}{216})$.)