Exercise Sheet 12

July 6th 2023

Due to the delay in the upload submission of the homework assignments until July 15th, 11:30 am online in TUM-moodle in groups of two. Please put the *full names* and *student IDs* of you *and* your partner on all parts of your submission. The solution will be discussed in the classes and published on moodle after some time.

Homework

Problem H 49 - A biased coin **[4 pts.]**

You bought a biased coin that does not equally show heads and tails but shows heads by the probability $p \neq \frac{1}{2}$ (independently over several trials). The vendor did not tell you the exact value of p but assures that it either is equal to $\frac{1}{4}$ (hypothesis H_0) or to $\frac{3}{4}$ (hypothesis H_1). In order to test these two hypotheses, you toss the coin n times and keep track of the number of heads, X.

a) Let $K \subseteq \{0, \ldots, n\}$ be the region of rejection of the hypothesis H_0 with respect to the test statistic X. How should you choose K in dependence of n such that at the same time the type-I and the type-II error are as small as possible (i.e. their sum is minimal)?

b) Assume that you will reject H_0 if you observe $X \geq \frac{n}{2}$. Find by means of the Chernoff bounds a value for n of manageable order of magnitude such that the type-I error will be less than 0.05. You may use, without a further proof, that $(e/4)^8 < 0.05$ holds.

Solution:

a) Denote the type-I and type-II error by α and β, respectively. By instruction, the test statistic X is a binomially distributed random variable with parameters n and p. Thus, we have

$$\alpha = \sup_{p \in H_0} Pr_p(X \in K) = \sum_{k \in K} Pr_{1/4}(X = k) = \sum_{k \in K} \binom{n}{k} \cdot \left(\frac{1}{4}\right)^k \cdot \left(\frac{3}{4}\right)^{n-k}$$

as well as

$$\beta = \sup_{p \in H_1} Pr_p(X \notin K) = \sum_{k \notin K} Pr_{3/4}(X = k) = \sum_{k \notin K} \binom{n}{k} \cdot \left(\frac{3}{4}\right)^k \cdot \left(\frac{1}{4}\right)^{n-k}.$$

The sum of the two errors is

$$\alpha + \beta = \left(\sum_{k \in K} \binom{n}{k} \cdot \left(\frac{1}{4}\right)^k \cdot \left(\frac{3}{4}\right)^{n-k} \right) + \left(\sum_{k \notin K} \binom{n}{k} \cdot \left(\frac{3}{4}\right)^k \cdot \left(\frac{1}{4}\right)^{n-k} \right). \checkmark$$

We notice that any value of $k \in \{0, \dots, n\}$ contributes exactly once to this term $\alpha + \beta$, either to the left or the right sum. Clearly, by the choice of the area of rejection $K \subset \{0, \dots, n\}$ we may determine to which one of the sum some k will give its contribution. We aim at minimizing $\alpha + \beta$, hence, $k \in \{0, \dots, n\}$ should be in K if and only if

$$\binom{n}{k} \cdot \left(\frac{1}{4}\right)^k \cdot \left(\frac{3}{4}\right)^{n-k} \leq \binom{n}{k} \cdot \left(\frac{3}{4}\right)^k \cdot \left(\frac{1}{4}\right)^{n-k}$$

which is equivalent to

$$3^{n-k} \leq 3^k.$$

The last condition is met for all $k \geq n/2$. In conclusion, we choose the area of rejection $K = \{\lceil n/2 \rceil, \dots, n\}$. ✓

b) From the first part, we know already the type-I error α. As X is binomially distributed we may find an upper bound for α by means of the Chernoff bound that was presented in lecture 4:

$$\alpha = Pr_{\frac{1}{4}}\left(X \geq \frac{n}{2}\right) = Pr_{\frac{1}{4}}\left(X \geq (1+1) \cdot \frac{n}{4}\right) \leq \left(\frac{e^1}{(1+1)^{1+1}}\right)^{n/4} = \left(\frac{e}{4}\right)^{n/4}. \checkmark$$

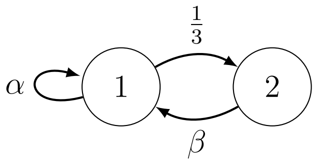
Here, we used that $\mathbb{E}(X) = n/4$ holds in the case of $p = 1/4$ and then applied, after a clever transformation, the Chernoff bound with the parameters $\mu = n/4$ and $\delta = 1$. From the instruction, we further know that $(e/4)^8 < 0.05$ holds. Thus, if we choose $n = 32$ we will find for the type-I error

$$\alpha \leq \left(\frac{e}{4}\right)^{32/4} = \left(\frac{e}{4}\right)^8 < 0.05$$

which is what we were asked for. ✓

Problem H 50 - Jumping in Markov Chains **[3 pts.]**

Consider the discrete Markov chain given by the graph below.



a) Determine the two transition probabilities $\alpha = p_{1,1}$ and $\beta = p_{2,1}$. Prove your result by showing that the transition matrix $P \in [0,1]^{2\times 2}$ is a stochastic matrix.

b) Assume the initial state distribution $\pi_0 = (1, 0)$ at $t = 0$. Explicitly write down the time evolution to time steps $t = 1, 2, 3$.

Solution:

a) Given a state i the sum of all transition probabilities from i (including the transition back to i) have to add up to 1. For the state 1 this means $1/3 + \alpha = 1$, yielding $\alpha = 2/3$, for 2 we directly get $\beta = 1$. The transition matrix hence is

$$P = \begin{pmatrix} 2/3 & 1/3 \\ 1 & 0 \end{pmatrix}.$$

Obviously, both rows of P sum up to 1, so P in fact is a stochastic matrix. ✓

b) The time evolution from a distribution π_t to π_{t+1} in the next time step can be calculated by multiplying from right by the transition matrix:

$$\pi_{t+1} = \pi_t \cdot P.$$

So, we have

$$\pi_1 = \pi_0 \cdot P = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 2/3 & 1/3 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \end{pmatrix},$$

$$\pi_2 = \pi_1 \cdot P = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 2/3 & 1/3 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{4}{9} + \frac{1}{3} & \frac{2}{9} + 0 \end{pmatrix} = \begin{pmatrix} \frac{7}{9} & \frac{2}{9} \end{pmatrix}$$

and

$$\pi_3 = \pi_2 \cdot P = \begin{pmatrix} \frac{7}{9} & \frac{2}{9} \end{pmatrix} \begin{pmatrix} 2/3 & 1/3 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{14}{27} + \frac{2}{9} & \frac{7}{27} + 0 \end{pmatrix} = \begin{pmatrix} \frac{20}{27} & \frac{7}{27} \end{pmatrix}. ✓✓$$

Alternatively, we could calculate π_2 and π_3 by

$$\pi_{t+n} = \pi_0 \cdot P^n,\ n \in \mathbb{N},$$

i.e.

$$\pi_2 = \pi_0 \cdot P^2 = \begin{pmatrix} 1 & 0 \end{pmatrix} \left[\begin{pmatrix} 2/3 & 1/3 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2/3 & 1/3 \\ 1 & 0 \end{pmatrix} \right] = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 7/9 & 2/9 \\ 2/3 & 1/3 \end{pmatrix} = \begin{pmatrix} \frac{7}{9} & \frac{2}{9} \end{pmatrix}$$

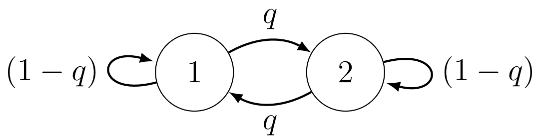
and

$$\pi_3 = \pi_0 \cdot P^3 = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 20/27 & 7/27 \\ 7/9 & 2/9 \end{pmatrix} = \begin{pmatrix} \frac{20}{27} & \frac{7}{27} \end{pmatrix}.$$

Note, that all found state distributions π_i are normalized (their entries sum up to 1) and that all P^n are stochastic matrices - right as it should be.

Problem H 51 - Properties of Markov Chains **[4 pts.]**

The graph below corresponds to a Markov Chain with parameter $0 \leq q \leq 1$.



a) Fix the parameter q such that the Markov chain is aperiodic, irreducible or ergodic, respectively.

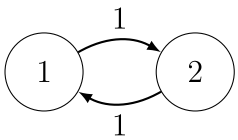
b) For which values of q exists a vector π such that $\lim_{t\to\infty} p_{i,j}^{(t)} = \pi_j$ holds for all states $i, j \in \{1, 2\}$?

Solution:

a) (i) By a lemma from the lecture, a *state* (i) is aperiodic if and only if there is a natural number k_0 such that

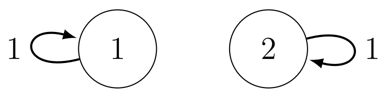
$$p_{i,i}^{(k)} > 0 \quad \forall k \geq k_0.$$

A *Markov chain* is aperiodic if all of its states are aperiodic. In the case $q < 1$ both states (1) and (2) have self-loops so the Markov chain certainly is aperiodic. In the case $q = 1$, however, the graph of the chain reduces to



from which we see that the k-step transition probability $p_{i,i}^{(k)}$ from a state (i) to itself is only positive if k is even. So, any state has period 2 (because the greatest common divisor of all k with $p_{i,i}^{(k)} > 0$ is 2) and the chain is not aperiodic.

(ii) A Markov chain is irreducible if each state is reachable from any state which is the case if and only if its graph is strongly connected. It is easy to see that this is exactly the case if $q > 0$. In the case of $q = 0$ we have the graph



.

This is certainly not strongly connected - the Markov chain is reducible in that case.

(iii) We call a Markov chain ergodic if it is aperiodic and irreducible. Putting together what we have found in (i) and (ii) this is the case for $0 < q < 1$. This allows us to get an intuition of ergodicity. ✓✓

b) Let us assume that the Markov chain is ergodic, i.e. $0 < q < 1$. The fundamental theorem for ergodic Markov chains then tells us that the distribution x_t will converge to a unique stationary distribution π independently from the initial distribution x_0, i.e. $\lim_{t\to\infty} x_t = \pi$. If we chose the initial distribution $(x_0)_i = 1$ for some i and $(x_0)_k = 0$ for all $k \neq i$ we may transform the t-step transition matrix to

$$p_{i,j}^{(t)} = \left(x_0 \cdot P^t\right)_j = (x_t)_j,$$

effectively picking the ith row of P and then taking the jth component. By means of the fundamental theorem for ergodic Markov chains we conclude that

$$\lim_{t\to\infty} p_{i,j}^{(t)} = \lim_{t\to\infty} (x_t)_j = \pi_j.$$

Next, we consider the case $q = 0$. Since now the Markov chain is aperiodic but not irreducible the fundamental theorem of ergodic Markov chains cannot be applied. In particular, the transition matrix P equals to the identity matrix. Thus, for arbitrary t we get

$$P^t = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^t = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

So, if $i = j$, we have $\lim_{t\to\infty} p_{i,j}^{(t)} = 1$. Elsewise, i.e. if $i \neq j$, it holds that $\lim_{t\to\infty} p_{i,j}^{(t)} = 0$. This implies that it is impossible that $\lim_{t\to\infty} p_{i,j}^{(t)}$ converges to the same value π_j for all starting points i.

Finally, we consider the case $q = 1$ in which the Markov chain is irreducible but not aperiodic. For even t, i.e. $t = 2k$, it follows for the transition matrix P that

$$P^t = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{2k} = \left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^2\right)^k = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^k = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

For odd t, i.e. $t = 2k + 1$, we have

$$P^t = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{2k+1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{2k} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

As the results for even and odd numbers differ, the limit $\lim_{t\to\infty} p_{i,j}^{(t)}$ obviously cannot exist.

To sum up, we have shown that the property $\lim_{t\to\infty} p_{i,j}^{(t)} = \pi_j$ with $i, j \in \{1, 2\}$ is true for all $0 < q < 1$. For the limiting cases that $q = 0$ or $q = 1$, respectively, it is false. In particular we have shown that neither irreducibility nor aperiodicity alone are sufficient to imply that $\lim_{t\to\infty} p_{i,j}^{(t)} = \pi_j$. ✓✓

Problem H 52 - It's raining, man! **[6 pts.]**

Professor Luttenberger has three umbrellas that he distributes to his office and his home. Every morning, he goes from his home to the office, and every evening he goes home. Whenever he leaves one of the places while it is raining he takes one of the umbrellas - if there is one at this place. If he does not have one at hand he will become wet. When it is not raining, he will not take an umbrella. Assume that it is raining at any of the walks independently and with the same probability $p = 2/3$.

We model this scenario by a Markov chain $(X_t)_{t\geq 0}$. Here, the random variable X_t describes how many umbrellas are available at the professor's current place (at home or in the office). The set of states hence is $S = \{0, 1, 2, 3\}$.

a) Draw the transition graph of the Markov chain $(X_t)_{t\geq 0}$.

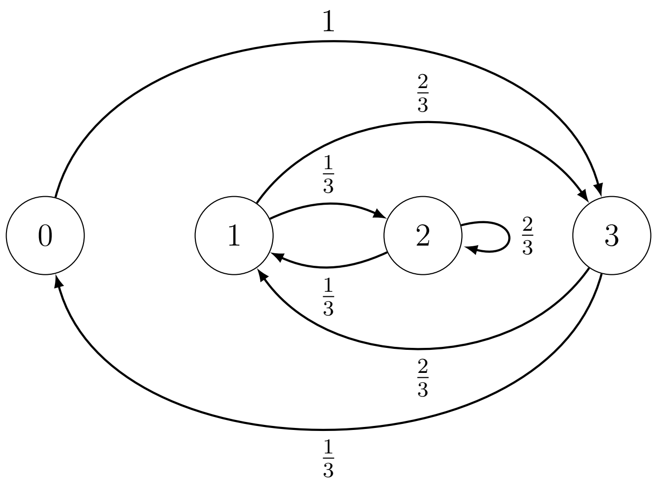
b) Prove or disprove that $(X_t)_{t\geq 0}$ is ergodic.

c) Professor Luttenberger follows this scheme to distribute the umbrellas for many years. Given this, what is the probability that he will become wet in rain at a walk?

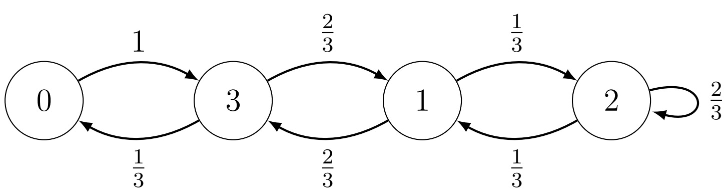
d) Some evening the professor notices that all umbrellas are at his home. After how many walks will he be at a place without any umbrella, on expectancy?

Solution:

a) In order to find the transition graph it is important to realize that one may infer the number of umbrellas at the destination of a walk from the number of umbrellas that are available for the next walk at the Professor's current place. For instance, if he has exactly one umbrella at hand at his current stay (e.g. at home), the other two umbrellas have to be at the other place (e.g. in the office). Now, if it is raining (which happens by $p = 2/3$), he takes the last umbrella and thus will have three umbrellas at the destination place available for his next walk. If it is not raining (occurring by $1 - p = 1/3$), he does not take the last umbrella and thus will have two umbrellas at hand at the destination. This leads to the transition graph below (that takes a simpler form in the alternative arrangement). ✓✓



or



b) As apparent from the graph, the Markov chain is strongly connected and hence irreducible. By the self-loop at the state 2, each state has two different paths of return that have coprime length. So, all states are aperiodic and therefore the Markov chain is aperiodic as a whole. Taken together, the Markov chain is ergodic. ✓

c) Denote by π the unique stationary distribution of the ergodic Markov chain. The fundamental theorem for ergodic Markov chains tells us that the Markov chain converges to π independently from the initial state. Professor Luttenberger will get wet only if he does not have any umbrellas at hand (meaning that he is in state 0) and, at the same time, it is raining. The probability sought for hence is given by $(\pi)_0 \cdot 2/3$ (here $(\pi)_0$ is the component of π referring to 0). Solving the equation system $\pi \cdot P = \pi$ of the stationary distribution with the transition matrix

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1/3 & 2/3 \\ 0 & 1/3 & 2/3 & 0 \\ 1/3 & 2/3 & 0 & 0 \end{pmatrix}$$

yields

$$\begin{aligned} (\pi)_3 \cdot \frac{1}{3} &= (\pi)_0 \\ (\pi)_2 \cdot \frac{1}{3} + (\pi)_3 \cdot \frac{2}{3} &= (\pi)_1 \\ (\pi)_1 \cdot \frac{1}{3} + (\pi)_2 \cdot \frac{2}{3} &= (\pi)_2 \\ (\pi)_0 \cdot 1 + (\pi)_1 \cdot \frac{2}{3} &= (\pi)_3. \end{aligned}$$

From the third equation it directly follows that

$$(\pi)_1 = (\pi)_2.$$

Putting this in the second equation, we similarly obtain

$$(\pi)_3 = (\pi)_1.$$

By the additional requirement that π is normalized, i.e.

$$(\pi)_0 + (\pi)_1 + (\pi)_2 + (\pi)_3 = 1$$

we get

$$(\pi)_1 = (\pi)_2 = (\pi)_3 = \frac{3}{10} \quad \text{and} \quad \pi_0 = \frac{1}{10}.$$

The desired probability hence is equal to 1/15. ✓✓

d) It is asked for the expected hitting time from state 3 to state 0 which we denote by $h_{3,0} = \mathbb{E}(T_{3,0})$. (Following the lecture, this would be called $x_s = x_3$ where the second index is omitted because the more general situation of reaching a set of target states G was considered. Clearly, here $G = \{0\}$ contains only one state.) For this the following equations have to hold:

$$h_{3,0} = 1 + \frac{2}{3} \cdot h_{1,0}$$
$$h_{1,0} = 1 + \frac{2}{3} \cdot h_{3,0} + \frac{1}{3} \cdot h_{2,0}$$
$$h_{2,0} = 1 + \frac{2}{3} \cdot h_{2,0} + \frac{1}{3} \cdot h_{1,0}.$$

The last equation yields

$$\frac{1}{3} \cdot h_{2,0} = 1 + \frac{1}{3} \cdot h_{1,0}.$$

Putting this into the second one we find

$$\frac{2}{3} \cdot h_{1,0} = 2 + \frac{2}{3} \cdot h_{3,0}.$$

Finally, putting this in the first equation leads to

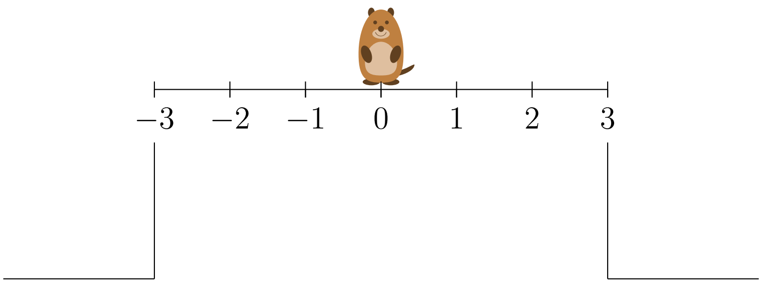
$$h_{3,0} = 9. \checkmark$$

Problem H 53 - Random Walk on a Finite Line **[8 pts.]**

A lemming is a very dumb animal that might unintentionally fall down a great height. Let us model the walk of a lemming on a one-dimensional plateau as a symmetric one-dimensional random walk of step size 1:

$$Pr(X_{t+1} = k + 1 | X_t = k) = \frac{1}{2},$$
$$Pr(X_{t+1} = k - 1 | X_t = k) = \frac{1}{2}$$

for all $t \in \mathbb{N}$. Instead of $X_i \in \mathbb{Z}$, we restrict the walk to $X_i \in \{-3, \ldots, 3\}$ corresponding to a plateau of 7 step-sizes length. If the lemming would enter -4 and 4 in the infinite case, it will fall down of the plateau and its walk will end (see the sketch below).



a) Is this stochastic process a Markov chain? If so draw the underlying transition graph.

b) Write a program that simulates this stochastic process with the initial condition $X_0 = 0$. Use a random number from Unif($[0, 1]$) to determine whether a step is taken to the right ($+1$) or to the left (-1).

c) Perform 10 independent simulations. Visualize the paths $(X_t)_{t\in\mathbb{N}}$ until time $t = 30$ in a plot with t on the x-axis and the position on the y-axis.

d) How would you call the states at positions $-3, \dots, 3$ on the plateau and how the states that the lemming has fallen down on the left or right end, respectively?

e) Determine the expected value of its position $\mathbb{E}(X_t)$ for an arbitrary point in time. Additionally, write down the stationary distribution for the limit $t \to \infty$. Based on this, will the lemming be alive or dead in the long run?

Solution:

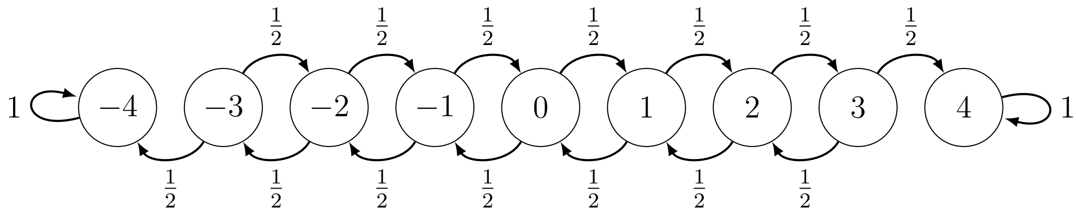
a) Clearly, for $k \in \{-3, \dots, 3\}$ the transition probabilites are given by the random walk

$$Pr(X_{t+1} = k+1 | X_t = k) = \frac{1}{2} \quad \text{and} \quad Pr(X_{t+1} = k-1 | X_t = k) = \frac{1}{2},$$

and for $k = \pm 4$ by

$$Pr(X_{t+1} = -4 | X_t = -4) = 1 \quad \text{and} \quad Pr(X_{t+1} = 4 | X_t = 4) = 1.$$

In any case, the probabilities for a state at time step $t+1$ depend solely on the present state at t which means that this discrete-state and discrete-time process satisfies the Markov property and hence is a Markov chain. ✓ Its transition graph takes he following form: ✓



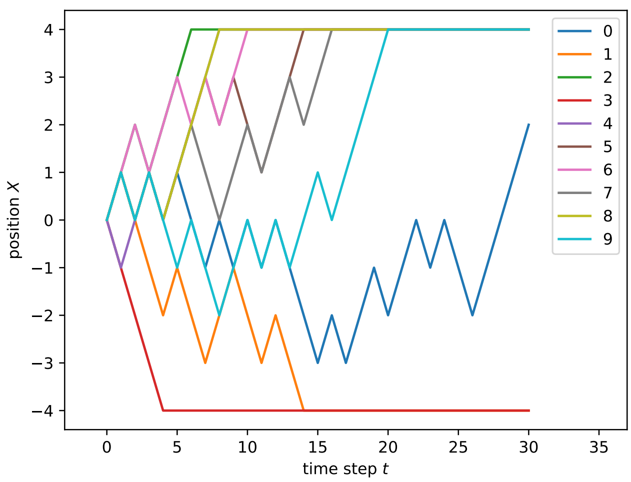
.

b) Here is some sample implementation in Python. ✓

```
#!/usr/bin/python3

import numpy as np
import matplotlib.pyplot as plt

#Function to find state at t+1 from state at t
def updater(state):
  if state in [-4,4]:
    return state
  else:
    r = np.random.rand()  #Random number
    if r <= 0.5:
      return state-1
    else:
      return state+1

```

c) The code continues with the actual 10 independent paths and their visualization. From the plot below we see that at $t = 30$ most of the paths are either in state -4 or 4. ✓✓

```python
#Do the time evolution 10 times
paths = []
for counter in range(10):
  path = [0]
  for t in np.arange(1,31,1): #Time steps until 30
    path.append(updater(path[-1]))
  paths.append(path)
paths = np.array(paths)

#Plot the 10 simulated paths
for counter in range(10):
  plt.plot(np.arange(0,31,1),paths[counter],label=str(counter))
plt.xlabel("time step $t$")
plt.ylabel("position $X$")
plt.xlim(-3,37)
plt.legend()
plt.show()
```



d) The states at positions $-3, \ldots, 3$ on the plateau are *transient*. Formally this means that there is positive chance to never return to each of that states. In other words, starting in one of these states, the probability to return the first time to that state before infinite time is less than 1, i.e.

$$Pr(T_j < \infty | X_0 = j) < 1 \quad \text{where} T_i = \min\{n \geq 1 | X_n = i\}.$$

This certainly is true because we always may reach one of the states -4 and 4 from which it is impossible to return. This property - that -4 and 4 cannot be left once entered - makes them *absorbing*. ✓Since all states are connected to an absorbing state we call the Markov chain absorbing as a whole, by the way. Note, that it follows that a non-absorbing state of an absorbing Markov chain is transient.

e) Given the initial state $X_0 = 0$, it is clear from symmetry that $\mathbb{E}(X_t) = 0$ for any t. More formally, we could consider all paths ω_t of length t, i.e. all possible walks with t steps. Clearly, for any path ω_t there is exactly one path with exactly the opposite sequence of steps. Since all of these paths are equally likely, they cancel out in the calculation and the expected value of X_t is equal to 0. ✓

From the fact that the states -4 and 4 and the symmetry of the set-up we suppose that the stationary distribution is $\pi = (1/2, 0, \dots, 0, 1/2)$. This can be verified by the equation system $\pi \cdot P = \pi$. Using the probability matrix of the graph we indeed get

$$(1/2, 0, \dots, 0, 1/2) \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1/2 & 0 & 1/2 & 0 & \dots \\ & \vdots & \vdots & \vdots & \\ \dots & 0 & 1/2 & 0 & 1/2 \\ 0 & \dots & 0 & 0 & 1 \end{pmatrix} = (1/2, 0, \dots, 0, 1/2). \checkmark$$

Importantly note that this stationary distribution is *not* independent from the initial distribution. In particular, the above solution of $\pi \cdot P = \pi$ is not unique.

It might seem paradox that in the long run the lemming will be at 0, i.e. at a position where it is alive, but almost sure dead (it is in -4 or 4 by probability of 1). However, the "mean" of having died at the left or the right end is at 0, so we expect the lemming to be dead in the long run.