Home Assignment 1

Markov Chains (273023), Return by 19 February. The purpose of this assignment is to estimate numerically the behaviour of the simple random walk on $\mathbb{Z}$. To complete the assignment, please return a written report (max. 2 pages + Appendix) by 19 February, as a pdf file by email to the lecturer. Include the code used as an appendix in the same pdf file.

NB: Only one pdf file for each group (1-2 persons). The pdf should be named: $surname_1 surname_2 tau.pdf$.

The main goal

Let (X_t) be the simple random walk on $\mathbb{Z}$. Suppose the walk starts at zero, $X_0 = 0$, and let τ_k be the first time the simple random walk is at state k, i.e.

$$\tau_k := \min\{t > 0 : X_t = k\},$$

The goal of the assignment is to estimate numerically the constant $C > 0$ in the inequality

$$\mathbb{P}(\tau_k > t) \leq C\frac{k}{\sqrt{t}}.$$

The following paragraphs describe one possible way of finding the C.

A suggestion for implementation

Simulate the random walk of fixed length. Write a program that generates a path (e.g. $(0, 1, 0, -1, -2, \ldots, 17)$) of length t starting from the state $0 \in \mathbb{Z}$ and visualize the path by plotting it. The path is a vector of length t consisting of the states where the chain visited, i.e. $(X_0, X_1, \ldots, X_n)$.

Write another program (function) that gives the first moment that the random walk is at state k. This function could for example take in a path vector from the previous paragraph. Do not forget to consider the case where the path does not reach the state k.

Estimate the time steps needed. Observe that

$$\mathbb{P}(\tau_k > t) \leq C\frac{k}{\sqrt{t}}$$

for some $C > 0$ using the course material (Hint: This is the Theorem 2.17 in Levin, Peres, and Wilmer, p. 30). Find $N(k) > 0$ such that

$$\mathbb{P}(\tau_k > N(k)) < 1/3.$$

Simulating the constant. For the constant C you can proceed as follows. For every $k = 10, 20, 40, 80, 160$ simulate $M = 1000$ independent paths of length $N(k)$ (where $N(k)$ is as calculated in the previous paragraph). Let $\tau_{k,j}$ be the value of τ_k at jth simulation. Take the median of the M simulated values $\tau_{k,j}$ and denote it by $\widehat{\tau_k}$. We can assume (why?) that

$$\mathbb{P}(\tau_k > \widehat{\tau_k}) \approx \frac{1}{2}.$$

Hence an estimate for the C is obtained by

$$\widehat{C} = \frac{\sqrt{\widehat{\tau_k}}}{2k}.$$

Report these estimates for C. Report also how many times the chain did not reach τ_k at $N(k)$ steps. In addition, specify how much computing time your simulations took.

Conclusions

Discuss the goodness of the estimate for the constant C. Are there weaknesses in the method? How could the method be improved? Were there particularly slow parts in the simulation? What happens to C and the computation time if the number of simulations is increased from 1000 to 10.000? How do you expect the estimate for C and the computation time to change if $k = 320$?