Discrete random processes with memory: Models and applications

Tomáš Kouřim[1,2], Petr Volf[2]
[1] Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague
[2]Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic

November 29, 2019

Abstract

The contribution focuses on non-Markov discrete-time random processes, i.e. processes with memory, and in particular, on Bernoulli-like random walks where the past events affect significantly the walk's future development. The main concern of the paper is therefore the formulation of models describing the dependence of transition probabilities on the process history. Such an impact can be incorporated explicitly and transition probabilities modulated using a few parameters reflecting the current state of the walk as well as the information about the past path. The behavior of proposed random walks, as well as the task of their parameters estimation, are studied both theoretically and with the aid of simulations.

Key words: Random walk, history dependent transition probabilities, non-Markov process, success punishing/rewarding walk

1 Introduction

Random process is one of the most important object of mathematics. It is well described theoretically and has real life representations in almost every aspect of human life, from physics and biology to economy and social sciences. The random process itself is merely a series of realizations of random variables. Depending on the type of the random variables and their mutual interactions random processes can be split into a large number of different categories. The most common type of a discrete random process is a random walk, a mathematical object first introduced by K.Pearson in 1905 [6]. Similarly to random processes in general, there exist many well described variations of a random walk with various applications to real life problems [9, 8]. Yet there are still new possibilities and options how to alter and improve the classical random walk and present yet another model representing different real life events. One of such modifications is the random walk with varying step size introduced in 2010 by Turban [9] which together with the idea of *self-exciting point processes* [2] and the perspective of model's applications in reliability analysis and also in sports statistics, served as an inspiration to the random walk with varying transition probabilities introduced by Kouřim [3, 4]. In the present paper, the theoretical properties of the model are described and further examined, numerical procedures of model parameters estimation are specified and the results are tested on generated data.

The rest of the paper is organized as follows. Sections 2 and 3 describe the properties of different versions of the model, section 4 provides results from simulated model testing and finally section 5 concludes the work.

2 Random walk with varying probabilities

The random walk with varying probabilities is based on a standard Bernoulli random walk [1] with some starting transition probability p_0. This probability is then altered after each step of the walk using a coefficient λ so that the repetition of the same step becomes less probable. Formally, it can be defined as

Definition 1. Let $\{X_n\}_{n=1}^{\infty}$ and $\{P_n\}_{n=1}^{\infty}$ be sequences of discrete random variables, and $p_0 \in [0, 1]$ and $\lambda \in (0, 1)$ constant parameters, such that the first random variable X_1 is given by

$$P(X_1 = 1) = p_0$$

$$P(X_1 = -1) = 1 - p_0.$$

Further

$$P_1 = \lambda p_0 + \frac{1}{2}(1-\lambda)(1-X_1) \tag{1}$$

and for $i \geq 2$

$$P(X_i = 1 | P_{i-1} = p_{i-1}) = p_{i-1}$$

$$P(X_i = -1 | P_{i-1} = p_{i-1}) = 1 - p_{i-1}$$

$$P_i = \lambda P_{i-1} + \frac{1}{2}(1-\lambda)(1-X_i). \tag{2}$$

The sequence $\{S_n\}_{n=0}^{\infty}$, $S_N = S_0 + \sum_{i=1}^{N} X_i$ for $n \in \mathbb{N}$, with $S_0 \in \mathbb{R}$ some given starting position, is called a *random walk with varying probabilities*, with $\{X_n\}_{n=1}^{\infty}$ being the steps of the walker and $\{P_n\}_{n=1}^{\infty}$ transition probabilities.

From [4], it can be further derived that at each step $t + k$, $t, k > 0$ the value of a transition probability P_{t+k} can be computed from the knowledge of transition probability P_t and the realization of the walk $X_{t+1}, \ldots, X_{t+k}$ using formula

$$P_{t+k} = P_t \lambda^{t+k} + \frac{1}{2}(1-\lambda) \sum_{i=t+1}^{t+k} \lambda^{t+k-i}(1 - X_i).$$

2.1 Properties

Basic properties of the random walk with varying probabilities are described in [4], namely that

$$EP_t = (2\lambda - 1)^t p_0 + \frac{1 - (2\lambda - 1)^t}{2} \tag{3}$$

and

$$ES_t = S_0 + (2p_0 - 1)\frac{1 - (2\lambda - 1)^t}{2(1 - \lambda)} \tag{4}$$

for $\forall t \geq 1$. This further yields $EP_t \to \frac{1}{2}$ and $ES_t \to S_0 + \frac{2p_0 - 1}{2(1-\lambda)}$ for $t \to +\infty$.

Now to describe the variance of the transition probability, let us prove the following proposition.

Proposition 1. *For $\forall t \geq 1$, it holds that*

$$Var(P_t) = (3\lambda^2 - 2\lambda)^t p_0^2 + \sum_{i=1}^{t} K(i-1)(3\lambda^2 - 2\lambda)^{t-i} - k(t)^2, \tag{5}$$

where

$$k(t) = EP_t = (2\lambda - 1)^t p_0 + \frac{1 - (2\lambda - 1)^t}{2}$$

and

$$K(t) = k(t) \cdot (-3\lambda^2 + 4\lambda - 1) + (1 - \lambda)^2.$$

Proof. To prove the proposition several support formulas has to be derived first. From the definition of variance follows

$$Var(P_t) = E(P_t^2) - E(P_t)^2. \tag{6}$$

$E(P_t)$ is given by (3), therefore in order to prove the proposition it is sufficient to prove the following statement

$$E(P_t^2) = (3\lambda^2 - 2\lambda)^t p_0^2 + \sum_{i=1}^{t} K(i-1)(3\lambda^2 - 2\lambda)^{t-i}. \tag{7}$$

To do so, let us first express the relation between $E(P_t^2)$ and $E(P_{t-1}^2)$ and $E(P_{t-1})$. From the definition of the expected value and the definition of the walk (2) follows

$$E(P_t^2) = E[E(P_t^2|P_{t-1}^2)] = E[E(P_t^2|P_{t-1})] =$$

$$= E[E(\lambda P_{t-1} + \frac{1}{2}(1-\lambda)(1 - X_t))^2 | P_{t-1}]. \tag{8}$$

Using that $E(X_t|P_{t-1}) = 2P_{t-1} - 1$,$E(X_t^2) = 1$ and further that

$$E[(1 - X_t)^2|P_{t-1}] = E[(1 - 2X_t + X_t^2)|P_{t-1}] = E[(2 - 2X_t)|P_{t-1}] =$$

$$= 4(1 - P_{t-1}).$$

Equation (8) further yields

$$E(P_t^2) = E[\lambda^2 P_{t-1}^2 + \lambda P_{t-1}(1-\lambda)E(1-X_t|P_{t-1}) + \frac{1}{4}(1-\lambda)^2 E((1-X_t)^2|P_{t-1})] =$$

$$= E[\lambda^2 P_{t-1}^2 + 2\lambda P_{t-1}(1-\lambda)(1-P_{t-1}) + (1-\lambda)^2(1-P_{t-1})]$$

and finally

$$E(P_t^2) = E(P_{t-1}^2)(3\lambda^2 - 2\lambda) + EP_{t-1}(-3\lambda^2 + 4\lambda - 1) + (1-\lambda)^2. \qquad (9)$$

Statement (7) can be proved using induction. Based on the trivial fact that $Ep_0 = p_0$ and $E(p_0)^2 = p_0^2$, for $t = 1$ we get

$$E(P_1^2) = (3\lambda^2 - 2\lambda)^1 p_0^2 + \sum_{i=1}^{1} K(i-1)(3\lambda^2 - 2\lambda)^{1-i} = (3\lambda^2 - 2\lambda)p_0^2 + K(0) =$$

$$= (3\lambda^2 - 2\lambda)p_0^2 + ((2\lambda - 1)^0 p_0 + \frac{1 - (2\lambda - 1)^0}{2}) \cdot (-3\lambda^2 + 4\lambda - 1) + (1-\lambda)^2 =$$

$$= (3\lambda^2 - 2\lambda)p_0^2 + p_0(-3\lambda^2 + 4\lambda - 1) + (1-\lambda)^2,$$

and from (9) follows that the induction assumption holds. Now for the induction step $t \to t+1$ we get by substituting (7) into (9)

$$E(P_{t+1}^2) = E(P_t^2)(3\lambda^2 - 2\lambda) + EP_t(-3\lambda^2 + 4\lambda - 1) + (1-\lambda)^2 =$$

$$= ((3\lambda^2 - 2\lambda)^t p_0^2 + \sum_{i=1}^{t} K(i-1)(3\lambda^2 - 2\lambda)^{t-i}) \cdot (3\lambda^2 - 2\lambda) + K(t) =$$

$$= (3\lambda^2 - 2\lambda)^{t+1} p_0^2 + \sum_{i=1}^{t} K(i-1)(3\lambda^2 - 2\lambda)^{t+1-i} + K(t) =$$

$$= (3\lambda^2 - 2\lambda)^{t+1} p_0^2 + \sum_{i=1}^{t+1} K(i-1)(3\lambda^2 - 2\lambda)^{t+1-i}$$

and the formula thus holds. Now substituting (3) and (7) into (6) yields (5) and proves the Proposition. □

From Proposition 1 the limit behavior of $Var(P_t)$ can be derived, namely

Corollary 1. *For* $t \to +\infty$,

$$Var(P_t) \to \frac{\frac{1}{2}(1-\lambda^2)}{(1 - 3\lambda^2 + 2\lambda)} - \frac{1}{4}.$$

Figure 1 shows the comparison of computed theoretical values of transition probability variance and its expected value and the actual observed values of average transition probability and variance for different starting probabilities p_0 and memory coefficients λ.

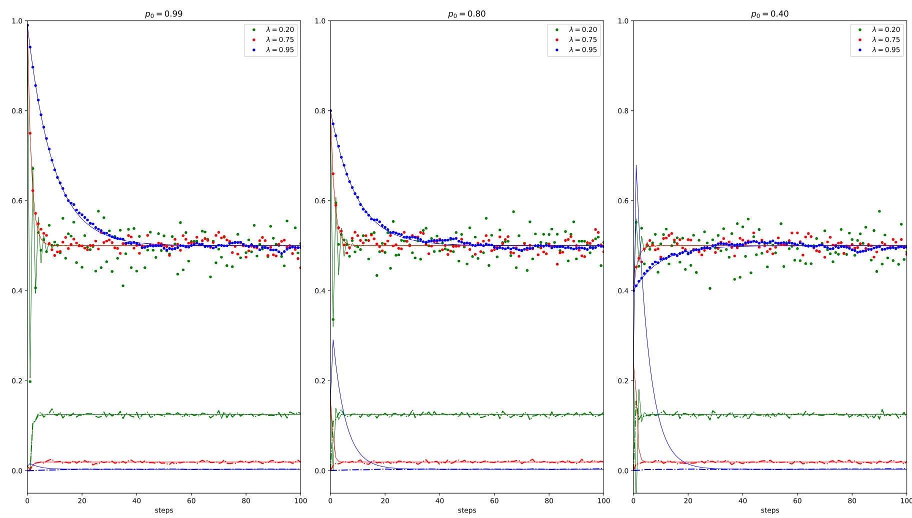

Figure 1: The development of the observed average transition probability (dotted) of a *success punished* version of the random walk and its observed variance (dot-dashed lines) compared to the theoretical values computed using from (3) and Proposition 1 (solid lines). The values were aggregated from 100 simulated realizations of each parameter combination.

3 Random walk with varying transition probability - alternatives

3.1 Success rewarded model

The basic definition of the random walk (Definition 1) presents a *success punished* model, meaning the probability of an event is decreased every time that event occurs. Opposite situation can be considered, where the probability of an event is increased every time that event occurs. Formally, such a random walk is defined in a following manner [4].

Definition 2. Let $\{X_n\}_{n=1}^{\infty}$, p_0 and λ be as in Definition 1. Further let $\{P_n\}_{n=1}^{\infty}$ be a sequence of discrete random variables given by

$$P_1 = \lambda p_0 + \frac{1}{2}(1-\lambda)(1+X_1) \tag{10}$$

$$P_i = \lambda P_{i-1} + \frac{1}{2}(1-\lambda)(1+X_i) \;\; \forall i \geq 2. \tag{11}$$

The sequence $\{S_n\}_{n=0}^{\infty}$, $S_N = S_0 + \sum_{i=1}^{N} X_i$ for $n \in \mathbb{N}$, with $S_0 \in \mathbb{R}$ some given starting position, is a random walk with varying probabilities - *success rewarded*.

In this section, all variables are considered to be related to the *success rewarded* model, whereas the variables with the same notations (P, X, S) from previous Section 2 are considered to be related to the model from Definition 1.

The *success rewarded* version of the model behaves differently than the *success punished* version, which can be observed with the help of the following propositions.

Proposition 2. *For* $\forall t \geq 2$,

$$P_t = p_0\lambda^t + \frac{1}{2}(1-\lambda)\sum_{i=1}^{t}\lambda^{t-i}(1+X_i) \tag{12}$$

Proof. The proposition is proved using induction. For $t = 2$ using (10) and (11) it holds that

$$P_2 = \lambda P_1 + \frac{1}{2}(1-\lambda)(1+X_2) = \lambda(\lambda p_0 + \frac{1}{2}(1-\lambda)(1+X_1)) + \frac{1}{2}(1-\lambda)(1+X_2) =$$

$$= p_0\lambda^2 + \frac{1}{2}(1-\lambda)\sum_{i=1}^{2}\lambda^{2-i}(1+X_i),$$

which is in accordance with (12). Now for the induction step $t \to t+1$ we obtain from (11) and the induction assumption

$$P_{t+1} = \lambda P_t + \frac{1}{2}(1-\lambda)(1+X_{t+1}) =$$

$$= \lambda(p_0\lambda^t + \frac{1}{2}(1-\lambda)\sum_{i=1}^{t}\lambda^{t-i}(1+X_i)) + \frac{1}{2}(1-\lambda)(1+X_{t+1}) =$$

$$= p_0\lambda^{t+1} + \frac{1}{2}(1-\lambda)\sum_{i=1}^{t}\lambda^{t-i+1}(1+X_i) + \frac{1}{2}(1-\lambda)(1+X_{t+1}) =$$

$$= p_0\lambda^{t+1} + \frac{1}{2}(1-\lambda)\sum_{i=1}^{t+1}\lambda^{t+1-i}(1+X_i).$$

□

Proposition 3. *For* $\forall t \geq 1$, $E(P_t) = p_0$.

Proof. Using $E(X_t|P_{t-1}) = 2P_{t-1} - 1$ and (11) we obtain

$$EP_t = E[E(P_t|P_{t-1})] = E[E(\lambda P_{t-1} + \frac{1}{2}(1-\lambda)(1+X_t)|P_{t-1})] =$$

$$= E[\lambda P_{t-1} + \frac{1}{2}(1-\lambda)(1+2P_{t-1}-1)] = E[\lambda P_{t-1} + (1-\lambda)P_{t-1}) =$$

$$= E(P_{t-1}).$$

Recursively we get

$$E(P_t) = E(p_0) = p_0. \tag{13}$$

□

Now to calculate the expected position of the walker at a given step $t \geq 1$, it is easy to see that $E(S_t) = S_{t-1} + 2P_{t-1} - 1$. From this, we can prove the following statement about the expected position of the walker after step t just from the knowledge of the input parameters.

Proposition 4. *For* $\forall t \geq 1$,

$$E(S_t) = S_0 + t(2p_0 - 1).$$

Proof. Using the result of Proposition 3 we get

$$E(S_{t+1}) = E[E(S_{t+1}|S_t)] = E[S_t + (2P_{t-1} - 1)] =$$

$$= ES_t + (2p_0 - 1)$$

which then recursively proves the statement. □

Corollary 2. *For* $t \to +\infty$,

$$\lim_{t\to+\infty} E(S_t) = \begin{cases} +\infty & p_0 > \frac{1}{2} \\ 0 & p_0 = \frac{1}{2} \\ -\infty & p_0 < \frac{1}{2} \end{cases}.$$

Proposition 5. *For* $\forall t \geq 1$,

$$Var(P_t) = (2\lambda - \lambda^2)^t p_0^2 + p_0(1-\lambda)^2 \sum_{i=1}^{t} (2\lambda - \lambda^2)^{t-i} - p_0^2. \tag{14}$$

Proof. The proof will be done in several steps similar as in Proposition 1. It is based on the definition of variance

$$Var(P_t) = E(P_t^2) - E(P_t)^2. \tag{15}$$

From Proposition 3 follows $E(P_t) = p_0$ and it is thus sufficient to prove that

$$E(P_t^2) = (2\lambda - \lambda^2)^t p_0^2 + p_0(1-\lambda)^2 \sum_{i=1}^{t} (2\lambda - \lambda^2)^{t-i}. \tag{16}$$

The proof will be done using induction again. First observe that

$$E(P_t^2) = E[E(P_t^2|P_{t-1}^2)] = E[E(P_t^2|P_{t-1})] =$$

$$= E[E(\lambda P_{t-1} + \frac{1}{2}(1-\lambda)(1+X_t))^2|P_{t-1}] =$$

$$= EP_{t-1}^2(2\lambda - \lambda^2) + p_0(1-\lambda)^2, \tag{17}$$

where the facts that $E[(1+X_t)^2|P_{t-1}] = 4P_{t-1}$, $E[(1+X_t)|P_{t-1}] = 2P_{t-1}$ and Proposition 3 were used. Now for $t = 1$ we get

$$EP_1 = p_0^2(2\lambda - \lambda^2) + p_0(1-\lambda)^2 = (2\lambda - \lambda^2)^1 p_0^2 + p_0(1-\lambda)^2 \sum_{i=1}^{1} (2\lambda - \lambda^2)^{1-i}$$

and the induction assumption holds. For the induction step $t \to t+1$ we get from the induction assumption and (17)

$$E(P_{t+1}^2) = EP_t^2(2\lambda - \lambda^2) + p_0(1-\lambda)^2 =$$

$$= ((2\lambda - \lambda^2)^t p_0^2 + p_0(1-\lambda)^2 \sum_{i=1}^{t} (2\lambda - \lambda^2)^{t-i}) \cdot (2\lambda - \lambda^2) + p_0(1-\lambda)^2 =$$

$$= (2\lambda - \lambda^2)^{t+1} p_0^2 + p_0(1-\lambda)^2 \sum_{i=1}^{t} (2\lambda - \lambda^2)^{t-i+1} + p_0(1-\lambda)^2 =$$

$$= (2\lambda - \lambda^2)^{t+1} p_0^2 + p_0(1-\lambda)^2 \sum_{i=1}^{t+1} (2\lambda - \lambda^2)^{t+1-i}.$$

The Proposition statement is then obtained by substituting (13) and (16) into (15). $\square$

Notice that the last sum in (14), after re-indexing by $j = t - i$, yields $\sum_{j=0}^{t-1} (2\lambda - \lambda^2)^j = \frac{1-(2\lambda-\lambda^2)^t}{1-2\lambda+\lambda^2}$. Hence the limit follows immediately

Corollary 3. *For $t \to +\infty$,*

$$\lim_{t \to +\infty} Var(P_t) = p_0(1 - p_0).$$

3.2 Model with two λ parameters

Another level of complexity can be added by using separate λ parameters for each direction of the walk. Again, two ways of handling success are available. The *success punished* version is defined as follows.

Definition 3. Let $\{X_n\}_{n=1}^{\infty}$ and p_0 be as in Definition 1. Further let λ_0, $\lambda_1 \in (0, 1)$ be constant coefficients and $\{P_n\}_{n=1}^{\infty}$ be a sequence of discrete random variables given by

$$P_1 = \frac{1}{2}[(1 + X_1)\lambda_0 p_0 + (1 - X_1)(1 - \lambda_1(1 - p_0))] \tag{18}$$

$$P_i = \frac{1}{2}[(1 + X_i)\lambda_0 P_{i-1} + (1 - X_i)(1 - \lambda_1(1 - P_{i-1}))] \;\; \forall i \geq 2. \tag{19}$$

The sequence $\{S_n\}_{n=0}^{\infty}$, $S_N = S_0 + \sum_{i=1}^{N} X_i$ for $n \in \mathbb{N}$, with $S_0 \in \mathbb{R}$ some given starting position, is a random walk with varying probabilities - *success punished two λ*.

And similarly the *success rewarded* version.

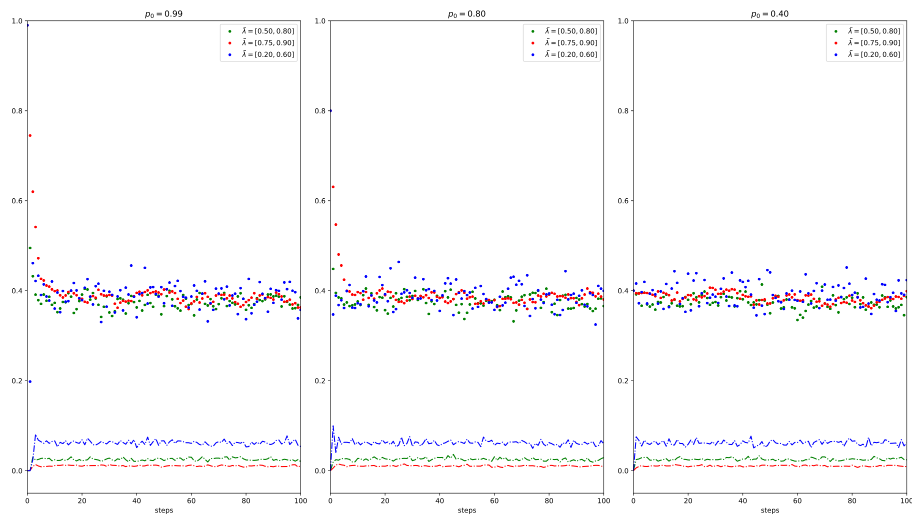

Figure 2: The development of the observed average transition probability (dotted) of a *success punished two* λ version of the random walk and its observed variance (dot-dashed lines). The values were aggregated from 100 simulated realizations of each parameter combination.

Definition 4. Let $\{X_n\}_{n=1}^{\infty}$ and p_0 be as in Definition 1. Further let $\lambda_0, \lambda_1 \in (0, 1)$ be constant coefficients and $\{P_n\}_{n=1}^{\infty}$ be a sequence of discrete random variables given by

$$P_1 = \frac{1}{2}[(1 - X_1)\lambda_0 p_0 + (1 + X_1)(1 - \lambda_1(1 - p_0))]$$

$$P_i = \frac{1}{2}[(1 - X_i)\lambda_0 P_{i-1} + (1 + X_i)(1 - \lambda_1(1 - P_{i-1}))] \;\; \forall i \geq 2.$$

The sequence $\{S_n\}_{n=0}^{\infty}$, $S_N = S_0 + \sum_{i=1}^{N} X_i$ for $n \in \mathbb{N}$, with $S_0 \in \mathbb{R}$ some given starting position, is a random walk with varying probabilities - *success rewarded two* λ.

Derivation of model properties is not so straightforward. The development of transition probability and its variance for different starting probabilities p_0 and memory coefficients pairs $\bar{\lambda}$ for the *success punished two* λ version of the model is shown on Figure 2. Similarly as in the single λ version of the model, the variance seems to depend on the $\bar{\lambda}$ pair only. The expected transition probability seems to converge to a constant value independently on both the starting probability p_0 and memory coefficients $\bar{\lambda}$. This interesting property of the walk will be subject to further study.

3.3 Other alternatives

The presented model of a random walk can be further developed and more versions can be derived and described. These variants include but are not limited to multidimensional walk (with either one or multiple λ parameters, again with *success rewarded* or *success punished*), a walk with the transition probability explicitly dependent on more than the last step, i.e. $P_t(k) \sim P_t(X_t, X_{t-1}, \ldots, X_{t-(k-1)})$, or the walk with λ parameter not constant, but a function of the time t, i.e. $P_t(\lambda(t))$. Detailed properties of such walks together with their possible applications on real life problems will by subject of a further study.

4 Simulations

Testing dataset was generated in order to validate the quality of the model and its ability to be fitted on a real life problem. The data generation was performed using the Python programming language and its package NumPy. Following values of input parameters were chosen. The memory coefficient values varied in $\lambda \in \{0.5, 0.8, 0.9, 0.99\}$ and similarly the pair of memory coefficients $[\lambda_0, \lambda_1] \in \{[0.5, 0.8], [0.5, 0.99], [0.99, 0.9]\}$. The starting transition probability p_0 was chosen from the set $P_0 = \{0.5, 0.8, 0.9, 0.99\}$ and the length of the walk was $n \in \{5, 10, 50, 100\}$. For each permutation of the parameters 100 walks were generated.

Four different fitting tasks were performed on the generated dataset. Using the maximum likelihood estimate (MLE) [7] and Python language with SciPy package the fitting tasks were

- Find $\overrightarrow{\lambda}$ with known p_0 and model type
- Find p_0 with known $\overrightarrow{\lambda}$ and model type
- Find p_0, $\overrightarrow{\lambda}$ with known model type
- Find model type without any prior knowledge

Table 1 shows the results of the model & parameter fitting algorithms using the MLE method. The data shows the model can be fitted with high accuracy. The only exception is finding the correct model type when the original model was based on a single λ parameter. The maximum-likelihood estimate almost always prefers the model with two λ parameters. To improve the results the Akaike Information Criterion $AIC = 2k - 2ln(\hat{L})$, which helps to correctly identify models with smaller number of parameters, was used. Here k is the number of model parameters and $\hat{L}$ is the maximal likelihood. This approach shows much better results. The performance can be observed in Table 2.

	SP - 1λ	SR - 1λ	SP - 2λ	SR - 2λ
Find $\overrightarrow{\lambda}$	96.9 %	34.4 %	80.2 %	77.1 %
Find p_o	92.2 %	82.8 %	89.6 %	93.8 %
Find $\overrightarrow{\lambda}$, p_0	91.4 %	84.4 %	83.3 %	79.9 %
Find model type	1.6 %	1.6 %	87.5 %	89.6 %

Table 1: Fitting results. *SP* stands for *success punished*, *SR* for *success rewarded*. 1λ vs. 2λ distinguish between the basic model with a single λ parameter and the more advanced model with two λ parameters.

	SP - 1λ	SR - 1λ	SP - 2λ	SR - 2λ
Find model type	79.7%	79.7%	83.3%	83.3%

Table 2: Finding the optimal model type using the Akaike Information Criterion.

5 Conclusion

This work follows up on the recent results on random walks with varying probabilities. It describes and proves certain properties of such a walk, other properties have been studied with the help of numerical methods. The study also shows the results of the maximum likelihood and *AIC* based estimations of model parameters and types using optimization procedures. The method has been tested successfully on a set of randomly generated data. The presented model has also many possible uses in real life application. Such a type of random walk describes especially well processes where either a single or just a small number of events can significantly affect the future development of the process. Such processes can be found in reliability analysis, medical as well as econometric studies, and very often in sports modeling. The authors recently published a study where the success rewarded model was applied to predict the *in-play* development of a Grand Slam tennis matches with compelling results when used for live betting against a bookmaker [5].

The source code containing all functionality mentioned in this article is freely available as open source at GitHub (https://github.com/tomaskourim/amistat2019).

Acknowledgment

The research was supported by the grant No. 18-02739S of the Grant Agency of the Czech Republic.

References

[1] William Feller. An introduction to probability theory and its applications. *1957*.

[2] Alan G Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90, 1971.

[3] Tomáš Kouřim. Random walks with varying transition probabilities. *Doktorandské dny FJFI*, 2017. Available at http://kmwww.fjfi.cvut.cz/ddny/historie/17-sbornik.pdf.

[4] Tomáš Kouřim. Statistical analysis, modeling and applications of random processes with memory. *PhD Thesis Study, ČVUT FJFI*, 2019.

[5] Tomáš Kouřim and Petr Volf. Tennis match as random walk with memory: Application to grand slam matches modelling. *Submitted to IMA Journal of Management Mathematics*, 2019.

[6] Karl Pearson. The problem of the random walk. *Nature*, 72(1865):294, 1905.

[7] Richard J Rossi. *Mathematical Statistics: An Introduction to Likelihood Based Inference*. John Wiley & Sons, 2018.

[8] Gunter M Schütz and Steffen Trimper. Elephants can always remember: Exact long-range memory effects in a non-markovian random walk. *Physical Review E*, 70(4):045101, 2004.

[9] Loïc Turban. On a random walk with memory and its relation with markovian processes. *Journal of Physics A: Mathematical and Theoretical*, 43(28):285006, 2010.