DISCRETE RANDOM PROCESSES WITH MEMORY: MODELS AND APPLICATIONS

Tomáš Kouřim, Petr Volf, Prague

Abstract. The contribution focuses 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.

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

MSC 2010: 60G50, 62F10

1. Introduction

One of the most common types of a discrete random process is a random walk, first introduced by K.Pearson in 1905 [8]. There exist many variations of a random walk with various applications to real life problems [10, 11]. 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 [11] which, together with the idea of *self-exciting point processes* [3] and the perspective of model applications in reliability analysis and also in sports statistics, served as an inspiration to the random walk with varying transition probabilities introduced by Kouřim [4, 6]. The definition of the walk falls into a rather broad class of processes

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

described for instance in the paper of Davis and Liu [1]. However, other assumptions, e.g. the condition of contraction, are not fulfilled by the walk and thus the conclusions from [1] cannot be applied.

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 evaluation 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 [2] 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 2.1. 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

$$P_1 = \lambda p_0 + \frac{1}{2}(1-\lambda)(1-X_1) \tag{2.1}$$

and for $i \geq 2$

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

$$P_i = \lambda P_{i-1} + \frac{1}{2}(1-\lambda)(1-X_i). \tag{2.2}$$

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.

2.1. **Properties.** The random walk with varying probabilities was first introduced in [4] and further elaborated in [6]. Following properties of the walk were described in these previous papers.

The value of a transition probability P_{t+k} at each step $t+k,\ t, k > 0$ 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

$$P_{t+k} = P_t \lambda^k + \frac{1}{2}(1-\lambda) \sum_{i=t+1}^{t+k} \lambda^{t+k-i}(1-X_i). \tag{2.3}$$

To compute the expected value of transition probability and position of the walker following formula can be used

$$EP_t = (2\lambda - 1)^t p_0 + \frac{1-(2\lambda-1)^t}{2} \tag{2.4}$$

and

$$ES_t = S_0 + (2p_0 - 1)\frac{1-(2\lambda-1)^t}{2(1-\lambda)} \tag{2.5}$$

for all $t \geq 1$. 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 walk in more detail, let us prove the following propositions.

Proposition 2.2. *For all $t \geq 1$, it holds that*

$$E(X_t) = (2\lambda - 1)^{t-1}(2p_0 - 1). \tag{2.6}$$

Proof. Using that $E(X_t|P_{t-1}) = 2P_{t-1} - 1$ the proposition can be proved directly using (2.4) as

$$E(X_t) = E(E(X_t)|P_{t-1}) = E(2P_{t-1} - 1) = 2E(P_{t-1}) - 1 =$$
$$= 2((2\lambda-1)^{t-1}p_0 + \frac{1-(2\lambda-1)^{t-1}}{2}) - 1 = (2\lambda-1)^{t-1}(2p_0-1).$$

□

Corollary 2.3. *The distribution of X_t converges to the Bernoulli $(1, -1)$ distribution with $p = \frac{1}{2}$. This Bernoulli distribution is simultaneously the stationary distribution of the random sequence X_t.*

Proof. As X_t are Bernoulli $(1, -1)$, their distributions are fully characterized by their expectations EX_t, and it holds that $EX_t = 2 \cdot EP_{t-1} - 1$. Then the first statement of the Corollary follows from the fact that $EP_t \to \frac{1}{2}$.

Further, let $EP_{t-1} = \frac{1}{2}$ be the characteristics of X_t, i.e. $EX_t = 0$. As then $EP_t = EP_{t-1}\lambda + (1-\lambda)/2(1 - EX_t) = \frac{1}{2}$, therefore $EX_{t+1} = 0$ again.

□

Remark 2.4. *For $p_0 = \frac{1}{2}$ and $t \geq 1$ or $\lambda = \frac{1}{2}$ and $t \geq 2$ it hods that X_t is the stationary random sequence with the distribution given by Corollary 2.3.*

Proposition 2.5. *For all $t \geq 1$, it holds that*

$$Var(P_t) = (3\lambda^2 - 2\lambda)^t p_0^2 + \sum_{i=1}^{t} K(i-1)(3\lambda^2 - 2\lambda)^{t-i} - k(t)^2, \tag{2.7}$$

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 it follows

$$Var(P_t) = E(P_t^2) - E(P_t)^2. \tag{2.8}$$

$E(P_t)$ is given by (2.4), therefore in order to prove the proposition it is sufficient to prove the following statement

$$E(P_t^2) = (3\lambda^2 - 2\lambda)^t p_0^2 + \sum_{i=1}^{t} K(i-1)(3\lambda^2 - 2\lambda)^{t-i}. \tag{2.9}$$

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 (2.2) it follows

$$E(P_t^2) = E[E(P_t^2|P_{t-1})] = E[E(\lambda P_{t-1} + \frac{1}{2}(1-\lambda)(1-X_t))^2|P_{t-1}]. \tag{2.10}$$

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 (2.10) 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

$$E(P_t^2) = E(P_{t-1}^2)(3\lambda^2 - 2\lambda) + EP_{t-1}(-3\lambda^2 + 4\lambda - 1) + (1-\lambda)^2. \tag{2.11}$$

Statement (2.9) can be proved using mathematical 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 (2.11) it follows that (2.9) holds for $t = 1$. Now for the induction step $t \to t+1$ we get by substituting (2.9) into (2.11)

$$E(P_{t+1}^2) = E(P_t^2)(3\lambda^2 - 2\lambda) + EP_t(-3\lambda^2 + 4\lambda - 1) + (1-\lambda)^2 =$$
$$= ((3\lambda^2 - 2\lambda)^t p_0^2 + \sum_{i=1}^{t} K(i-1)(3\lambda^2 - 2\lambda)^{t-i}) \cdot (3\lambda^2 - 2\lambda) + K(t) =$$
$$= (3\lambda^2 - 2\lambda)^{t+1} p_0^2 + \sum_{i=1}^{t} K(i-1)(3\lambda^2 - 2\lambda)^{t+1-i} + K(t) =$$
$$= (3\lambda^2 - 2\lambda)^{t+1} p_0^2 + \sum_{i=1}^{t+1} K(i-1)(3\lambda^2 - 2\lambda)^{t+1-i}$$

and the formula thus holds. Now substituting (2.4) and (2.9) into (2.8) yields (2.7) and proves the Proposition. □

From Proposition 2.5 the limit behavior of $Var(P_t)$ can be derived easily:

Corollary 2.6. *For* $t \to +\infty$,

$$\lim_{t \to +\infty} Var(P_t) = \frac{\frac{1}{2}(1-\lambda^2)}{1 - 3\lambda^2 + 2\lambda} - \frac{1}{4}. \tag{2.12}$$

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

Proposition 2.7. *For all* $t \geq 1$, *it holds that*

$$Var(X_t) = 1 - (2\lambda - 1)^{2(t-1)}(2p_0 - 1)^2. \tag{2.13}$$

Proof. The fact that X_t are Bernoulli $(1, -1)$ implies $E(X_t^2) = 1$. The statement then follows directly from the definition of variance and Proposition 2.2. □

Corollary 2.8. *For* $t \to +\infty$,

$$\lim_{t \to +\infty} Var(X_t) = 1. \tag{2.14}$$

The variance of the position of the walker was studied with the help of computer simulations, presented in Figure 2. The simulations show that the variance grows to infinity with $t \to \infty$ depending on both p_0 and λ. The derivation of an exact formula will be subject of further studies.

3. Random walk with varying transition probability - alternatives

3.1. **Success rewarding model.** The basic definition of the random walk (Definition 2.1) presents a *success punishing* model, meaning that 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 with each event's occurrence. Formally, such a random walk is defined in a following manner [6]:

Definition 3.1. Let $\{X_n\}_{n=1}^{\infty}$, p_0 and λ be as in Definition 2.1. Further let $\{P_n\}_{n=1}^{\infty}$ be a sequence of discrete random variables given by

$$P_1 = \lambda p_0 + \frac{1}{2}(1 - \lambda)(1 + X_1), \tag{3.1}$$

$$P_i = \lambda P_{i-1} + \frac{1}{2}(1 - \lambda)(1 + X_i) \;\; \forall i \geq 2. \tag{3.2}$$

The sequence $\{S_n\}_{n=0}^{\infty}$, given as in Definition 2.1, is a random walk with varying probabilities - *success rewarding*.

In this section, all variables (P, X, S) are related to the *success rewarding* model. This version of the model behaves differently than the *success punishing* version, which can be observed with the help of the following propositions.

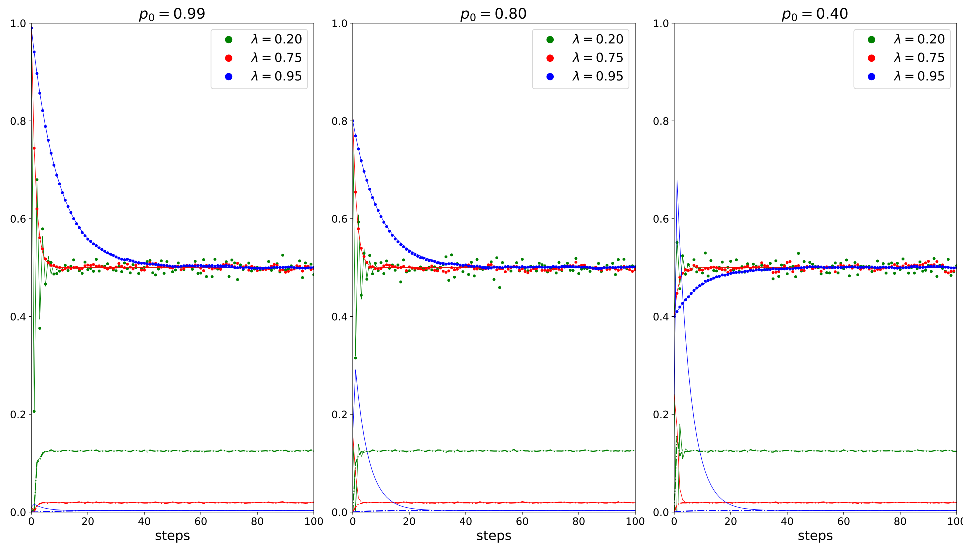

Figure 1. The observed average transition probability (dotted, upper part of the figure) of a *success punishing* version of the random walk and its observed variance (dot-dashed lines, lower part of the figure) compared to the theoretical values computed using (2.4) and Proposition 2.5 (same colors, solid lines). The values were computed from 1000 simulated realizations of each parameter combination.

Proposition 3.2. *For all* $t \geq 2$,

$$P_t = p_0\lambda^t + \frac{1}{2}(1-\lambda)\sum_{i=1}^{t}\lambda^{t-i}(1+X_i). \tag{3.3}$$

Proof. The Proposition is proved using mathematical induction. For $t = 2$ using (3.1) and (3.2) it holds that

$$P_2 = \lambda P_1 + \frac{1}{2}(1-\lambda)(1+X_2) = \lambda(\lambda p_0 + \frac{1}{2}(1-\lambda)(1+X_1)) + \frac{1}{2}(1-\lambda)(1+X_2) =$$
$$= p_0\lambda^2 + \frac{1}{2}(1-\lambda)\sum_{i=1}^{2}\lambda^{2-i}(1+X_i),$$

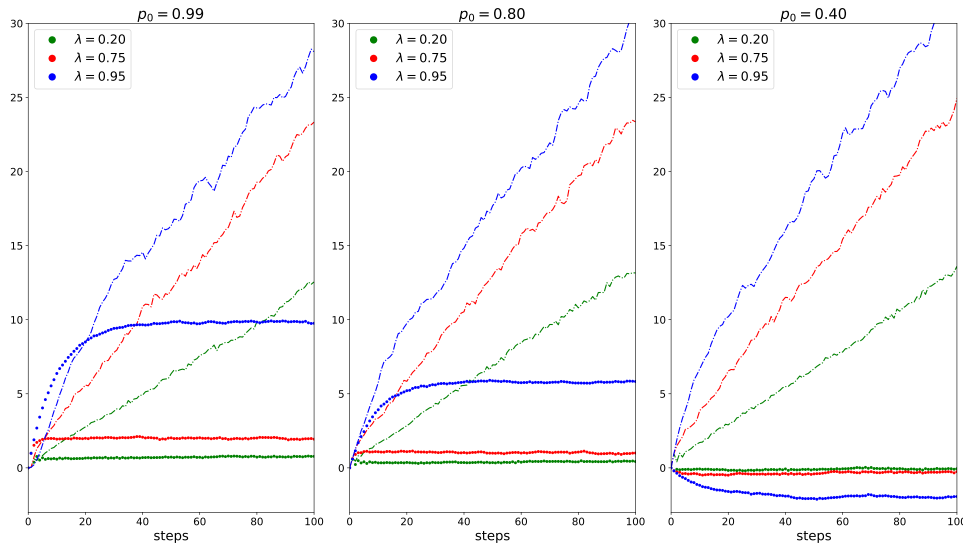

FIGURE 2. The observed average position of the walker (dotted, "thicker") of a *success punishing* version of the random walk and its variance (dot-dashed lines, "thinner"). The values were computed from 1000 simulated realizations of each parameter combination.

which is in accordance with (3.3). Now for the induction step $t \to t+1$ we obtain from (3.2) and the induction assumption

$$\begin{aligned} P_{t+1} &= \lambda P_t + \frac{1}{2}(1-\lambda)(1+X_{t+1}) = \\ &= \lambda\Big(p_0\lambda^t + \frac{1}{2}(1-\lambda)\sum_{i=1}^{t}\lambda^{t-i}(1+X_i)\Big) + \frac{1}{2}(1-\lambda)(1+X_{t+1}) = \\ &= p_0\lambda^{t+1} + \frac{1}{2}(1-\lambda)\sum_{i=1}^{t}\lambda^{t-i+1}(1+X_i) + \frac{1}{2}(1-\lambda)(1+X_{t+1}) = \\ &= p_0\lambda^{t+1} + \frac{1}{2}(1-\lambda)\sum_{i=1}^{t+1}\lambda^{t+1-i}(1+X_i). \end{aligned}$$

$\square$

Proposition 3.3. *For all $t \geq 1$, $E(P_t) = p_0$.*

Proof. Using $E(X_t|P_{t-1}) = 2P_{t-1} - 1$ and (3.2) 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

$$E(P_t) = E(p_0) = p_0. \tag{3.4}$$

□

Proposition 3.4. *The sequence X_t is a stationary sequence of Bernoulli random variables with values* 1, −1 *and with* $P(X_t = 1) = p_0$.

Proof. As the distribution of X_t is fully given by $E(P_{t-1})$, the statement follows directly from Proposition 3.3. □

Further we can calculate the expected position of the walker at a given step t just from the knowledge of the input parameters.

Proposition 3.5. *For all* $t \geq 1$,

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

Proof. As $EX_{t+1} = E[E(X_{t+1}|P_t)] = E(2P_t - 1)$, using the result of Proposition 3.3 we get

$$E(S_{t+1}) = E(S_t + X_{t+1}) = ES_t + E(2P_t - 1) = ES_t + (2p_0 - 1)$$

which then recursively proves the statement. □

Corollary 3.6. *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 3.7. *For all* $t \geq 1$,

$$Var(P_t) = (2\lambda - \lambda^2)^t p_0^2 + p_0(1-\lambda)^2 \sum_{i=1}^{t} (2\lambda - \lambda^2)^{t-i} - p_0^2. \tag{3.5}$$

Proof. The proof will be done in several steps similar as in Proposition 2.5. It is based on the definition of variance (2.8). From Proposition 3.3 it follows $E(P_t) = p_0$ and it is thus sufficient to prove that

$$E(P_t^2) = (2\lambda - \lambda^2)^t p_0^2 + p_0(1-\lambda)^2 \sum_{i=1}^{t} (2\lambda - \lambda^2)^{t-i}. \tag{3.6}$$

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

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

$$= EP_{t-1}^2(2\lambda - \lambda^2) + p_0(1-\lambda)^2, \tag{3.7}$$

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 3.3 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 thus (3.6) holds for $t = 1$. For the induction step $t \to t+1$ we get from the induction assumption and (3.7)

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

The Proposition is then proved by substituting (3.4) and (3.6) into (2.8). □

Notice that the last sum in (3.5), 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 3.8. *For $t \to +\infty$,*

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

Proposition 3.9. *For all $t \geq 1$, it holds that*

$$Var(X_t) = 4p_0(1-p_0).$$

Proof. As $E(X_t) = 2p_0 - 1$ and $E(X_t^2) = 1$ the proof follows similarly as in Proposition 2.7 directly from the definition of variance. □

3.2. **Two-parameter models.** 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.

Definition 3.10. Let $\{X_n\}_{n=1}^{\infty}$ and p_0 be as in Definition 2.1. 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

$$P_1 = \frac{1}{2}[(1+X_1)\lambda_0 p_0 + (1-X_1)(1-\lambda_1(1-p_0))] \tag{3.8}$$

$$P_i = \frac{1}{2}[(1+X_i)\lambda_0 P_{i-1} + (1-X_i)(1-\lambda_1(1-P_{i-1}))] \;\; \forall i \geq 2. \tag{3.9}$$

The sequence $\{S_n\}_{n=0}^{\infty}$, given as in Definition 2.1, is a random walk with varying probabilities - *two-parameter success punishing*.

Definition 3.11. Let $\{X_n\}_{n=1}^{\infty}$ and p_0 be as in Definition 2.1, λ_0, λ_1 as in Definition 3.10 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 $\{S_n\}_{n=0}^{\infty}$, given as in Definition 2.1, is a random walk with varying probabilities - *two-parameter success rewarding*.

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 $[\lambda_0,\, \lambda_1] = \bar{\lambda}$ for the *two-parameter success punishing* version of the

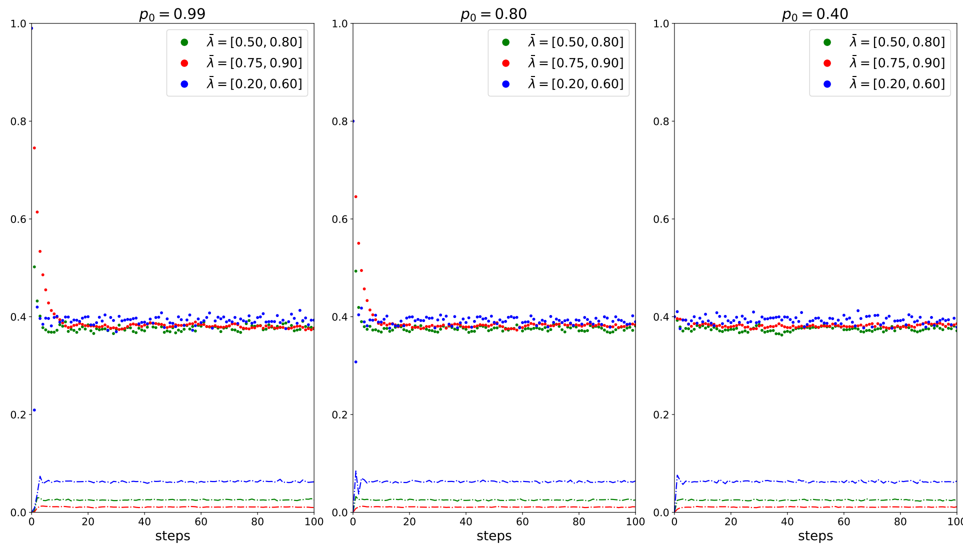

FIGURE 3. The development of the observed average transition probability (dotted, upper part of the figure) of a *two-parameter success punishing* version of the random walk and its variance (dot-dashed lines, lower part of the figure). The values were computed from 1000 simulated realizations of each parameter combination.

model is shown on Figure 3. 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 of a 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, a multidimensional walk (with either one or multiple λ parameters, *success rewarding* or *success punishing*), 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 a 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

A simulation study was performed in order to verify the possible usage of the presented model in real life situations, namely on processes with relatively few events (i.e. short walks), which can be (under certain assumptions) observed repetitively. Such processes include for example the recurrence of diseases (few recurrences but many patients), reliability of machines (few failures but multiple same machines) or the modelling of sports (few significant events in a match but multiple matches). The experiment consisted of generating K random walks of length n, of the same walk type and parameter configuration, and estimating the walk type and parameter values from the generated data. Four different task were considered:

(1) find $\tilde{\lambda}$ with known p_0 and model type,
(2) find p_0 with known $\tilde{\lambda}$ and model type,
(3) find p_0 and $\tilde{\lambda}$ with known model type,
(4) find model type without any prior knowledge.

Following parameter values were considered for data generation: $K \in \{5, 100\}$, $n \in \{5, 10, 50, 100\}$, $p_0 \in \{0.5, 0.8, 0.9, 0.99\}$, $\lambda \in \{0.5, 0.8, 0.9, 0.99\}$ and $[\lambda_0, \lambda_1] \in \{[0.5, 0.8], [0.1, 0.5], [0.5, 0.99], [0.99, 0.9]\}$ (further let $\tilde{\lambda}$ denote either λ or $[\lambda_0, \lambda_1]$ depending on the walk type).

Tasks 1–3 were solved using the maximum likelihood estimate (MLE) [9]. The derivation of the theoretical likelihood values is rather complicated, therefore numerical approach using the Python programming language and its scientific package SciPy was applied. The Akaike Information Criterion $AIC = 2k - 2ln(\hat{L})$, where k is the number of model parameters and $\hat{L}$ is the maximal likelihood, was then used for the last task.

As mentioned above, we were not able to use standard approach to confidence intervals. Each experiment was therefore repeated independently $N = 100$ times for each parameter combination and sample characteristics were computed from the 100 parameter estimates. To assess the quality of the parameter estimation (tasks 1–3) four different evaluation criteria were tested.

(1) true parameter value lies within the standard $(1 - \alpha)$ two-sided confidence interval around the mean,
(2) true parameter value lies within the "percentile" interval, i.e. between the $\frac{100\alpha}{2} - th$ and $100\left(1 - \frac{\alpha}{2}\right) - th$ percentile,
(3) the mean fitted parameter value lies within the "proximity" interval around the true parameter value ω, computed as $[\omega - \frac{\alpha}{2}\omega, \omega + \frac{\alpha}{2}\omega]$,
(4) the median fitted parameter value lies within the "proximity" interval.

For task 4, the model was estimated using the AIC on each set of K walks. To evaluate the quality of such estimation the proportion of correctly chosen model types for the given walk configuration was computed as the proportion of the number of correctly chosen models to the number of analyzed walk sets N.

The above mentioned criteria serve only as an approximate tool to evaluate the estimate quality, however the results show that the model can be successfully fitted on empirical data. For $K = 100$ and $\alpha = 0.1$, 89% of all evaluation criteria (for tasks 1–3) were successful and the correct model was found in 85% of cases (task 4). As expected, the results are less convincing for $K = 5$, with only 71% of all evaluation criteria being successful and 70% of correctly found models. Longer walks show generally better results when finding the coefficients $\tilde{\lambda}$ while the performance of finding correct p_0 seems independent on the walk's length. This is not surprising as the parameter p_0 affects mostly the first few steps of the walk, while $\tilde{\lambda}$ play their role thorough the entire course of the walk. As expected, tasks 1–2 show better results than task 3, as there are less parameters to estimate. An example of the parameter estimation evaluation can be seen in Table 1 (there just task 1), an example of the model type identification results (task 4) can be observed in Table 2. Both tables contain only a brief illustration of results due to space limitations. Full results of all evaluation setups as well as several values of parameter α can be found in the GitHub repository (see the last paragraph of the paper).

A drawback of the presented simulation experiment is the discrepancy between the theoretical model (where $0 < P_t < 1$, for all t) and its representation in computer simulation, where the limited precision rounds values very close to 1. This rounding causes the optimization algorithm to produce inaccurate estimates or to not converge at all. It affects only the *success rewarding* version of the model when estimating the λ parameter and only cases with $K = 100$ and $n \geq 50$ (63% not converged estimations), independent of the choice of λ or p_0 parameters. An example can seen in Table 1 (row 6). The effect of this phenomenon on other model configurations is negligible. Handling of such unwanted behavior will be subject of further research.

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 successfully tested on a set of randomly generated data. The presented model has also many possible uses in real life applications. Such a type of random walk describes especially well processes

Table 1. The table shows an example of task 1 evaluation results, with true parameter value $\lambda = 0.5$ or $\lambda_0 = 0.5$ (and corresponding $\lambda_1 = 0.8$), and $p_0 = 0.5$, $\alpha = 0.1$. The mean estimated parameter value and its standard deviation and the median estimated parameter value and the corresponding "percentile" interval are presented. *SP* stands for *success punishing*, *SR* for *success rewarding*, the number 2 denotes the model with two λ parameters.

		Type	mean	st. dev.	median	percentile
$K = 100$	$n = 5$	SP	0.505	0.043	0.504	$[0.439; 0.576]$
		SR	0.502	0.033	0.503	$[0.451; 0.549]$
		SP2	0.505	0.060	0.501	$[0.399; 0.606]$
		SR2	0.491	0.047	0.490	$[0.415; 0.564]$
	$n = 100$	SP	0.499	0.008	0.499	$[0.484; 0.511]$
		SR	-	-	-	-
		SP2	0.502	0.012	0.502	$[0.478; 0.521]$
		SR2	0.498	0.026	0.501	$[0.452; 0.535]$
$K = 5$	$n = 5$	SP	0.468	0.155	0.495	$[0.214; 0.690]$
		SR	0.489	0.214	0.499	$[0.000; 0.830]$
		SP2	0.462	0.209	0.464	$[0.139; 0.780]$
		SR2	0.521	0.211	0.527	$[0.123; 0.923]$
	$n = 100$	SP	0.493	0.037	0.494	$[0.419; 0.554]$
		SR	0.376	0.056	0.382	$[0.382; 0.382]$
		SP2	0.497	0.056	0.498	$[0.405; 0.586]$
		SR2	0.461	0.173	0.513	$[0.001; 0.655]$

Table 2. The table shows model estimation success rate. Notation and parameter configuration is the same as in Table 1.

K	**n**	**SP**	**SR**	**SP2**	**SR2**
100	5	83%	80%	100%	100%
	100	86%	88%	100%	100%
5	5	84%	85%	42%	34%
	100	82%	80%	100%	93%

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 presented a study where the *success rewarding* 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, 7].

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

References

[1] Richard A Davis and Heng Liu. Theory and inference for a class of observation-driven models with application to time series of counts. *Statistica Sinica*, 26:1673–1707, 2016.

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

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

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

[5] Tomáš Kouřim. Random walks with memory applied to grand slam tennis matches modeling. In *Proceedings of MathSport International 2019 Conference (e-book)*, pages 220–227. Propobos Publications, 2019.

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

[7] 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.

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

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

[10] Gunter M Sch ü tz 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.

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

Authors' addresses: *Tomáš Kouřim*, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Czech Republic e-mail: `kourim@outlook.com`. *Petr Volf*, Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Prague e-mail: `volf@utia.cas.cz`.