DISCRETE RANDOM PROCESSES WITH MEMORY: MODELS AND APPLICATIONS

Tomáš Kouřim, Petr Volf, Praha

Abstract. The contribution focuses on Bernoulli-like random walks, where the past events significantly affect 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 parameter estimation, are studied both theoretically and with the aid of simulations.

Keywords: random walk; history dependent transition probability; non-Markov process; success punishing walk; success rewarding walk

MSC 2020: 60G50, 62F10

1. Introduction

One of the most common types of a discrete random process is a random walk, first introduced by Pearson in 1905, see [7]. There exist many variations of a random walk with various applications to real-life problems [9], [10]. Yet there are still new possibilities and options regarding how to alter and improve the classical random walk and present yet another model representing different real-life events. One such modification is the random walk with varying step size introduced in 2010 by Turban [10] 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 for 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 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 a 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

$$P(X_1 = 1) = p_0, \quad P(X_1 = -1) = 1 - p_0.$$

Further,

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

and for $i \geqslant 2$

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

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

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 the formula

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

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

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

and

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

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

Now, to describe the walk in more detail, let us prove the following propositions.

Proposition 2.2. *For all $t \geqslant 1$ we have*

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

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

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

$\square$

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.*

P r o o f.   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. $\qquad\square$

R e m a r k 2.4.   For $p_0 = \frac{1}{2}$ and $t \geqslant 1$ or $\lambda = \frac{1}{2}$ and $t \geqslant 2$ it holds that X_t is the stationary random sequence with the distribution given by Corollary 2.3.

Proposition 2.5. *For all $t \geqslant 1$ we have*

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

where

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

and

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

P r o o f.   To prove the proposition, several support formulas have to be derived first. From the definition of variance it follows that

$$(2.8) \qquad\qquad\qquad \mathrm{Var}(P_t) = E(P_t^2) - E(P_t)^2,$$

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

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

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

$$(2.10) \qquad E(P_t^2) = E[E(P_t^2 | P_{t-1})] = E\left[E(\lambda P_{t-1} + \tfrac{1}{2}(1 - \lambda)(1 - X_t))^2 | P_{t-1}\right].$$

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

$$E[(1 - X_t)^2 | P_{t-1}] = E[(1 - 2X_t + X_t^2)|P_{t-1}] = E[(2 - 2X_t)|P_{t-1}] = 4(1 - P_{t-1}),$$

274

equation (2.10) then yields

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

and finally

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

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

$$E(P_1^2) = (3\lambda^2 - 2\lambda)^1 p_0^2 + \sum_{i=1}^{1} K(i-1)(3\lambda^2 - 2\lambda)^{1-i} = (3\lambda^2 - 2\lambda)p_0^2 + K(0)$$
$$= (3\lambda^2 - 2\lambda)p_0^2 + \left((2\lambda - 1)^0 p_0 + \frac{1 - (2\lambda - 1)^0}{2}\right) \cdot (-3\lambda^2 + 4\lambda - 1) + (1-\lambda)^2$$
$$= (3\lambda^2 - 2\lambda)p_0^2 + p_0(-3\lambda^2 + 4\lambda - 1) + (1-\lambda)^2,$$

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)

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

and the formula thus holds. Now substituting (2.4) and (2.9) into (2.8) yields (2.7) and proves the Proposition. $\qquad\square$

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

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

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

275

Figure 1 shows the comparison of computed theoretical values of the 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.
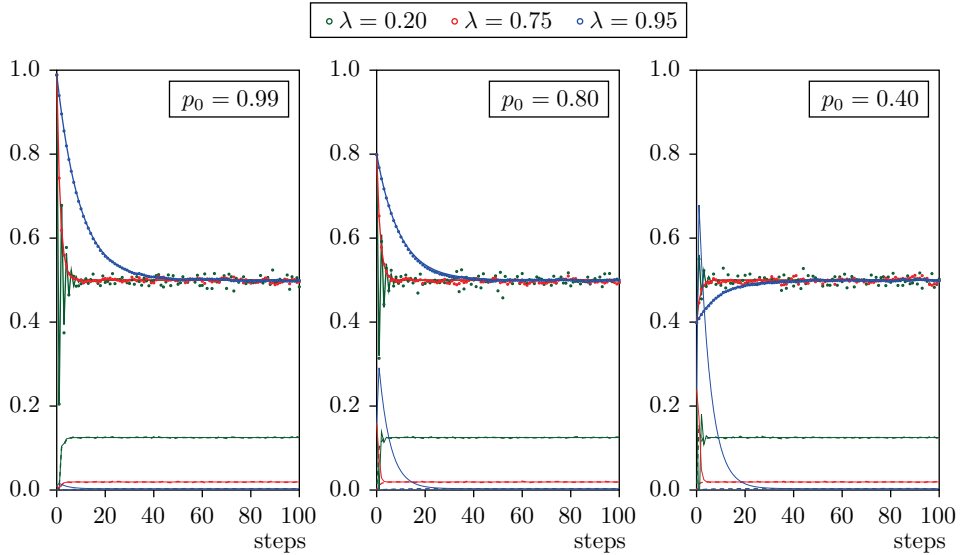

Figure 1. The observed average transition probability (dotted, upper part of the figure) of a *success punishing* version of the random walk and its observed variance (dashed lines, lower part of the figure) compared to the theoretical values computed using (2.4) and Proposition 2.5 (same colors, solid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

Proposition 2.7. *For all $t \geqslant 1$ we have*

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

P r o o f. 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 $t \to \infty$,*

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

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 the subject of further studies.
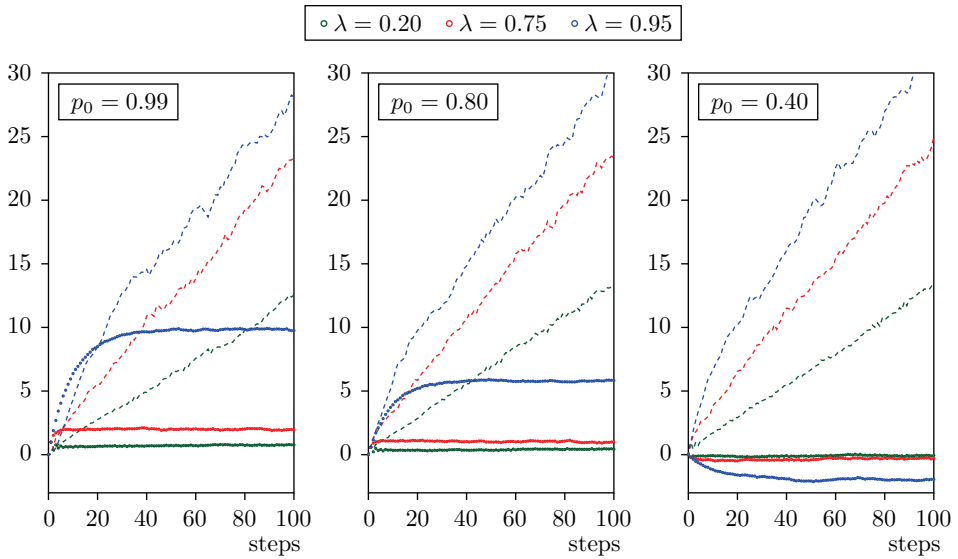

Figure 2. The observed average position of the walker (dotted, "thicker") of a *success punishing* version of the random walk and its variance (dashed lines, "thinner"). The values were computed from 1000 simulated realizations of each parameter combination.

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. The 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

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

$$(3.2) \qquad P_i = \lambda P_{i-1} + \frac{1}{2}(1 - \lambda)(1 + X_i) \quad \forall\, i \geqslant 2.$$

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.

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

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

P r o o f.   The Proposition is proved using mathematical induction. For $t = 2$ using (3.1) and (3.2), we find that

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

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

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

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

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

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

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

$\square$

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

P r o o f.   Using $E(X_t|P_{t-1}) = 2P_{t-1} - 1$ and (3.2), we obtain

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

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

$$= E[\lambda P_{t-1} + (1-\lambda)P_{t-1}) = E(P_{t-1}).$$

Recursively we get

$$E(P_t) = E(p_0) = p_0.$$

$\square$

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

P r o o f. As the distribution of X_t is fully given by $E(P_{t-1})$, the statement follows directly from Proposition 3.3. $\square$

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 $t \geqslant 1$,*

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

P r o o f. 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. $\square$

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 $t \geqslant 1$,*

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

P r o o f. The proof will be done in several steps, similarly to the proof of Proposition 2.5. It is based on the definition of variance (2.8). From Proposition 3.3 it

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

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

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

$$(3.7) \qquad E(P_t^2) = E[E(P_t^2 | P_{t-1})] = E\left[E\left(\lambda P_{t-1} + \frac{1}{2}(1-\lambda)(1+X_t)\right)^2 | P_{t-1}\right]$$
$$= EP_{t-1}^2(2\lambda - \lambda^2) + p_0(1-\lambda)^2,$$

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

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

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). $\qquad \square$

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 $t \to \infty$,*

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

Proposition 3.9. *For all $t \geqslant 1$ we have*

$$\mathrm{Var}(X_t) = 4p_0(1 - p_0).$$

P r o o f. 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

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

$$(3.9) \qquad P_i = \frac{1}{2}[(1 + X_i)\lambda_0 P_{i-1} + (1 - X_i)(1 - \lambda_1(1 - P_{i-1}))] \quad \forall i \geqslant 2.$$

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

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

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

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 coefficient pairs $[\lambda_0, \lambda_1] = \bar{\lambda}$ for the *two-parameter success punishing* version of the model is shown in 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 the memory coefficients $\bar{\lambda}$. This interesting property of the walk will be the subject of a further study.
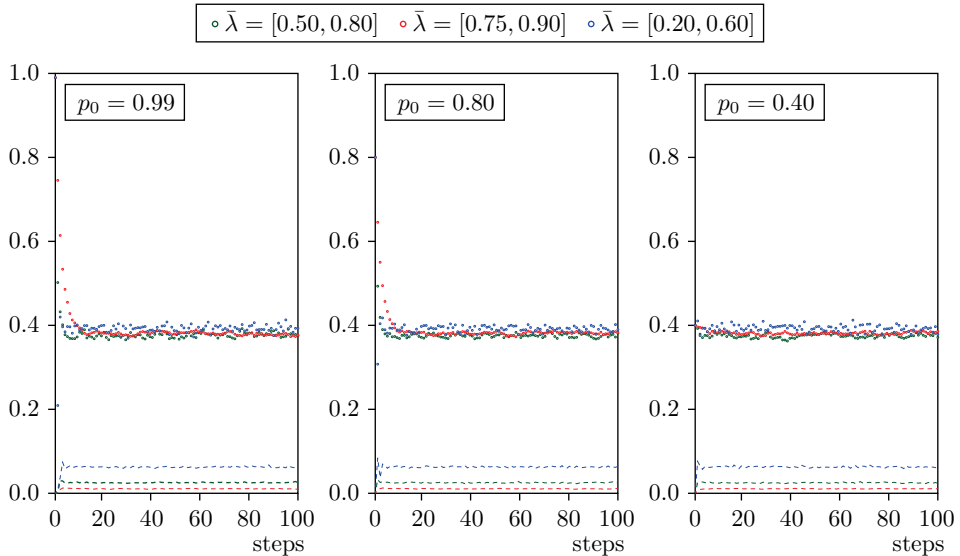
Figure 3. The development of the observed average transition probability (dotted, upper part of the figure) of a *two-parameter success punishing* version of the random walk and its variance (dashed lines, lower part of the figure). The values were computed from 1000 simulated realizations of each parameter combination.

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 to real life problems will be the 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 multiple processes with relatively few events (i.e. multiple short walks of the same kind). 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 tasks were considered:

(1) find λ or $[\lambda_0, \lambda_1]$ (further denoted as $\tilde{\lambda}$) 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.

The 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 $[\lambda_0, \lambda_1] \in \{[0.5, 0.8], [0.1, 0.5], [0.5, 0.99], [0.99, 0.9]\}$.

Tasks 1–3 were solved using the maximum likelihood estimate (MLE), see [8]. The derivation of the theoretical likelihood values is rather complicated; therefore a numerical approach using the Python programming language and its scientific package SciPy was applied. The Akaike Information Criterion AIC $= 2k - 2\ln(\hat{L})$, where k is the number of model parameters and $\hat{L}$ is the maximal likelihood, was then used for the last task.

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

To evaluate task 4, all four presented models were fitted on the set of K walks and the AIC was computed for each one of them. The model with the lowest AIC value was then selected. The quality of such estimation for the given walk configuration was then assessed using 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's quality, however the results show that the model can be successfully fitted to empirical data. For $K = 100$, $\alpha = 0.1$ and all combinations of input parameters and walk lengths $\tilde{\lambda}$, p_0 and n, 92% 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 73% of all evaluation criteria being successful and 70% of correctly found models. An example of the parameter

estimation evaluation can be seen in Table 1 (there just task 1), and 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, i.e. combinations of input parameters $\tilde{\lambda}$ and p_0, number of observed walks K and their different lengths n as well as several values of parameter α can be found in the GitHub repository (see the last paragraph of the paper).

		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	0.502	0.022	0.502	$[0.460, 0.535]$
		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	0.485	0.102	0.496	$[0.391, 0.624]$
		SP2	0.497	0.056	0.498	$[0.405, 0.586]$
		SR2	0.461	0.173	0.513	$[0.001, 0.655]$

Table 1. The table shows an example of task 1 evaluation results, with true parameter values $\lambda = 0.5$ or $\lambda_0 = 0.5$ (and corresponding $\lambda_1 = 0.8$), and $p_0 = 0.5$, $\alpha = 0.1$. The mean of parameter estimates and its standard deviation and the median of parameter estimates 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.

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%

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

Longer walks show generally better results when finding the coefficients $\tilde{\lambda}$ especially for the *success rewarding* version of the model (as seen for example in row 10 in Table 1), 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.

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, while 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 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].

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] *R. A. Davis, H. Liu*: Theory and inference for a class of nonlinear models with application to time series of counts. Stat. Sin. *26* (2016), 1673–1707. zbl MR doi

[2] *W. Feller*: An Introduction to Probability Theory and Its Applications. A Wiley Publication in Mathematical Statistics. John Wiley & Sons, New York, 1957. zbl MR

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

[4] *T. Kouřim*: Random walks with varying transition probabilities. Doktorandské dny 2017 (P. Ambrož, Z. Masáková, eds.). ČVUT, FJFI, Praha, 2017, pp. 141–150.

[5] *T. Kouřim*: Random walks with memory applied to grand slam tennis matches modeling. MathSport International 2019: Conference Proceedings. Propobos Publications, Athens, 2019, pp. 220–227.

[6] *T. Kouřim*: Statistical Analysis, Modeling and Applications of Random Processes with Memory: PhD Thesis Study. ČVUT FJFI, Praha, 2019.

[7] *K. Pearson*: The problem of the random walk. Nature *72* (1905), 342. zbl doi

[8] *R. J. Rossi*: Mathematical Statistics: An Introduction to Likelihood Based Inference. John Wiley & Sons, Hoboken, 2018.   zbl doi

[9] *G. M. Schütz, S. Trimper*: Elephants can always remember: Exact long-range memory effects in a non-Markovian random walk. Phys. Rev. E *70* (2004), Article ID 045101.   doi

[10] *L. Turban*: On a random walk with memory and its relation with Markovian processes. J. Phys. A, Math. Theor. *43* (2010), Article ID 285006, 9 pages.   zbl MR doi

Authors' addresses: *Tomáš Kouřim*, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Břehová 7, 115 19 Praha 1, Czech Republic, e-mail: `kourim@outlook.com`; *Petr Volf*, Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 182 00 Praha 8, Czech Republic, e-mail: `volf@utia.cas.cz`.