A model of random walk with varying transition probabilities

Tomáš Kouřim[1] and Petr Volf[2]

[1] Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Czech Republic
(E-mail: kourim@outlook.com)

[2] Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Prague
(E-mail: volf@utia.cas.cz)

Abstract. This paper considers a model of the one-dimensional discrete time random walk in which the position of the walker is controlled by a varying transition probabilities. Namely, these probabilities depend explicitly on the previous move of the walker and implicitly on the entire walk history. Hence, transition probabilities evolve in time making the walk a non-Markovian stochastic process. The paper follows on the recent work of the authors. Two basic version of the model are introduced, some of their properties are recalled and new theoretical results derived. Then, more complex variants of models are presented. Development of walks themselves as well as the properties of connected sequences of transition probabilities are illustrated also with the aid of simulation. Possible application of the model in real life situations are discussed and briefly described, too.

Keywords: Random walk, history dependent transition probabilities, success punishing/rewarding walk.

1 Introduction

Stochastic processes and the corresponding mathematical theory represent a significant part of mathematics. One of the most prominent of such processes is the random walk, introduced by K. Pearson over hundred years ago [6]. This concept has been then further elaborated by many authors creating a number of different versions of a random walk [7] and there are still new possibilities and options how the classical random walk can be altered and adapted to specific application field. The model discussed in this paper follows on the work of Turban [8] and represents yet another version of a random walk, walk with varying transition probabilities. The model falls within a rather broad class of processes presented in recent work of Davis and Liu [1], but not all assumptions from [1] are met.

The original inspiration for the model comes from one of its applications - modelling of sports events. Many types of sport, such as tennis or volleyball, are played in a strictly discrete matter with steps divided by individual *points*,

6^{th} *SMTDA Conference Proceedings, 2-5 June 2020, Barcelona, Spain*

games or *sets*. One sport match can be thus viewed as a random walk with individual parts of the match representing the steps of the walk. Success, i.e. scoring a point or winning a set, then significantly affects further development of the entire walk by changing the transition probability. Other real life situations with similar properties can be found everywhere in areas where both "successes" or "failures" occur. In fact, also discrete time recurrent counts data occurrence can be often modelled in a similar way, when the event probability is affected by the process recent history. Such cases include the recurrence of diseases, recidivism in crime, as well as repeated defects and maintenance of a technical device.

The present contribution continues on the recent authors' exploration of the models of random walk with varying probabilities. Selected properties of the model are presented and possible real life implementations of the model are discussed. The rest of the paper is organized as follows. Next section introduces theoretical properties of the model and describes in detail two main variants of the model. Section (3) presents possible applications of the model and the last section concludes this work.

2 Theoretical properties

The original motivation of the model comes from modelling of sport events where the probability of a success (i.e. scoring a goal, achieving a point etc.) is at the center of interest. After each occurrence of such success this probability either decreases or increases, and thus two basic model alternatives exist - *success punishing* and *success rewarding*. The basic version of the model operates with starting success probability p_0 and a memory coefficient λ affecting the severity of probability change after a success as input parameters. Formally the walk is defined as follows:

Definition 1. Let $p_0 \in (0, 1)$, $\lambda \in (0, 1)$ be constant parameters, $\{P_n\}_{n=0}^{\infty}$ and $\{X_n\}_{n=1}^{\infty}$ sequences of discrete random variables with $P_0 = p_0$ and for $t \geq 1$

$$P(X_t = 1 | P_{t-1} = p_{t-1}) = p_{t-1}, \quad P(X_t = -1 | P_{t-1} = p_{t-1}) = 1 - p_{t-1},$$

and (*success punishing)*

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

or (*success rewarding)*

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

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 $\{P_n\}_{n=0}^{\infty}$ transition probabilities. Depending on the chosen formula to calculate P_i the walk type is either *success punishing* (1) or *success rewarding* (2).

The model was first introduced in [2] and a more thorough description was provided in [4]. Selected properties were then presented in [5] and a practical implementation of the model in modelling tennis matches was presented in [3]. The basic results are recalled in the following sections and then variance of S_t is described in more detail.

2.1 Success punishing model

The basic properties of the *success punishing* version of the walk are presented in this section. Previous results are presented as a set of expressions only, the reader is kindly asked to see referred papers for full proves of those expressions. Newly described properties are then provided with full proves and all necessary details.

For the expected value and variance of the step size for the $t \geq 1$ iteration of the walk X_t it holds [5]

$$EX_t = (2\lambda - 1)^{t-1}(2p_0 - 1), \tag{3}$$

$$Var\, X_t = 1 - (2\lambda - 1)^{2(t-1)}(2p_0 - 1)^2. \tag{4}$$

For the expected value and variance of the transition probability or the $t \geq 1$ iteration of the walk P_t it holds [2,5]

$$EP_t = (2\lambda - 1)^t p_0 + \frac{1 - (2\lambda - 1)^t}{2}, \tag{5}$$

$$Var\, P_t = (3\lambda^2 - 2\lambda)^t p_0^2 + \sum_{i=0}^{t-1} K(i; p_0, \lambda)(3\lambda^2 - 2\lambda)^{t-1-i} - k(t; p_o, \lambda)^2, \tag{6}$$

where

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

and

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

Finally, the expected position of the walker S_t after $t \geq 1$ iterations can be expressed as [2]

$$ES_t = S_0 + (2p_0 - 1)\frac{1 - (2\lambda - 1)^t}{2(1 - \lambda)}. \tag{7}$$

The last formula missing is the one expressing the variance of the position of the walker $Var\, S_t$. Before it is presented, let us first prove a support proposition.

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

$$E(P_t S_t) = (2\lambda - 1)^t p_0 S_0 + \sum_{i=0}^{t-1} (2\lambda - 1)^i q(t - 1 - i; p_0, S_0, \lambda), \tag{8}$$

where

$$q(j;p_0,S_0,\lambda) = (1-\lambda)s(j;p_0,S_0,\lambda) + 2\lambda\pi(j;p_0,\lambda) + (1-2\lambda)p(j;p_0,\lambda) + \lambda - 1$$

and $p(j;p_0,\lambda) = EP_j$ *given by* (5), $s(j;p_0,S_0,\lambda) = ES_j$ *given by* (7) *and* $\pi(j;p_0,\lambda) = EP_j^2$ *given as*

$$EP_t^2 = (3\lambda^2-2\lambda)^t p_0^2 + \frac{1-(3\lambda^2-2\lambda)^t}{3\lambda+1}\frac{(\lambda+1)}{2} - (p_0-\frac{1}{2})(3\lambda^2-4\lambda+1)M(t-1;\lambda), \tag{9}$$

where

$$M(t;\lambda) = \sum_{i=0}^{t-1}(3\lambda^2-2\lambda)^{t-1-i}(2\lambda-1)^i.$$

Proof. The formula for EP_t^2 follows from the proof of Proposition 2.5 in [5]. To prove (8) let us start with expressing the value of $E(P_tS_t)$ from the knowledge of past steps as

$$\begin{aligned} E(P_tS_t) &= E[E(P_tS_t|P_{t-1})] = \\ &= E[E((\lambda P_{t-1} + \frac{1}{2}(1-\lambda)(1-X_t))(S_{t-1}+X_t)|P_{t-1})] = \\ &= E[E(\lambda P_{t-1}S_{t-1} + \frac{1-\lambda}{2}S_{t-1} - \frac{1-\lambda}{2}X_tS_{t-1} + \\ &+ \lambda X_tP_{t-1} + \frac{1-\lambda}{2}X_t - \frac{1-\lambda}{2}X_t^2|P_{t-1})] \end{aligned}$$

and using $E(X_t|P_{t-1}) = 2P_{t-1}-1$ and $EX_t^2 = 1$ finally

$$E(P_tS_t) = (2\lambda-1)E(P_{t-1}S_{t-1}) + (1-\lambda)ES_{t-1} + 2\lambda EP_{t-1}^2 + (2\lambda-1)EP_{t-1} + \lambda - 1. \tag{10}$$

Further we will continue using mathematical induction. For $t=1$ using the definition of the walk it holds that

$$\begin{aligned} E(P_1S_1) &= p_0(p_0\lambda(S_0+1)) + (1-p_0)[(1-(1-p_0)\lambda)(S_0-1)] = \\ &= (2\lambda-1)p_0S_0 + (1-\lambda)S_0 + 2\lambda p_0^2 - (2\lambda-1)p_0 + \lambda - 1. \end{aligned}$$

When substituting $t=1$ into (8) we obtain

$$\begin{aligned} E(P_1S_1) &= (2\lambda-1)p_0S_0 + \sum_{i=0}^{0}(2\lambda-1)^i q(0-i;p_0,S_0,\lambda) = \\ &= (2\lambda-1)p_0S_0 + (1-\lambda)s(0;p_0,\lambda) + \\ &+ 2\lambda\pi(0;p_0,\lambda) + (1-2\lambda)p(0;p_0,\lambda) + \lambda - 1 \end{aligned}$$

and finally using (5), (7) and (9)

$$E(P_1S_1) = (2\lambda - 1)p_0S_0 + (1-\lambda)S_0 + 2\lambda p_0^2 + (1-2\lambda)p_0 + \lambda - 1.$$

Equation (8) thus holds for $t = 1$. Now for the induction step $t \to t+1$, after substituting(8) into (10) we get $E(P_{t+1}S_{t+1}) =$

$$\begin{aligned}
&= (2\lambda-1)E(P_tS_t) + (1-\lambda)ES_t + 2\lambda EP_t^2 + (2\lambda-1)EP_t + \lambda - 1 = \\
&= (2\lambda-1)[(2\lambda-1)^t p_0S_0 + \sum_{i=0}^{t-1}(2\lambda-1)^i q(t-1-i)] + (1-\lambda)s(t) + \\
&\quad +2\lambda\pi(t) + (2\lambda-1)p(t) + \lambda - 1 = \\
&= (2\lambda-1)^{t+1}p_0S_0 + \sum_{i=0}^{t}(2\lambda-1)^i q(t-i).
\end{aligned}$$

Theorem 1. *For all $t \geq 1$*

$$Var\,S_t = t + 4\sum_{i=0}^{t-1}\sigma(i;p_0,0,\lambda) - a(t;p_0,\lambda),$$

with $\sigma(i;p_0,S_0,\lambda) = E(P_tS_t)$ given by (8) and

$$a(t;p_0,\lambda) = (2p_0-1)\sum_{i=0}^{t-1}\frac{1-(2\lambda-1)^i}{1-\lambda} + (2p_0-1)^2\frac{(1-(2\lambda-1)^t)^2}{4(1-\lambda)^2}.$$

Proof. As clearly the value S_0 does not affect the variance, let us from now assume $S_0 = 0$. From the definition of variance

$$Var\,S_t = ES_t^2 - E^2S_t \tag{11}$$

and (7) follows that to prove the proposition it is enough to prove that

$$ES_t^2 = t + 4\sum_{i=0}^{t-1}\sigma(i;p_0,0,\lambda) - (2p_0-1)\sum_{i=0}^{t-1}\frac{1-(2\lambda-1)^i}{1-\lambda}. \tag{12}$$

First of all, let us express ES_t^2 from the knowledge of past walk development. From the definition of the expected value and the definition of the walk (Definition 1) it follows

$$\begin{aligned}
ES_t^2 &= E[E(S_t^2|P_{t-1})] = E[E((S_{t-1}+X_t)^2|P_{t-1})] = \\
&= E(S_{t-1}^2 + 2(2P_{t-1}-1)S_{t-1} + 1) = \\
&= ES_{t-1}^2 + 4E(P_{t-1}S_{t-1}) - 2ES_{t-1} + 1,
\end{aligned} \tag{13}$$

where the fact that $EX_t^2 = 1$ was used. The theorem will be proved using mathematical induction again . For $t = 1$ we get using the definition of the walk (Definition 1)

$$ES_1^2 = p_0(S_0+1)^2 + (1-p_0)(S_0-1)^2 = 1.$$

Substituting $t = 1$ into (12) we obtain

$$ES_1^2 = 1 + 4\sigma(0; p_0, 0, \lambda) - (2p_0 - 1)\frac{1 - (2\lambda - 1)^0}{1 - \lambda} = 1$$

and (12) thus holds for $t = 1$. Now for the induction step $t \to t + 1$ we get by substituting (12) into (13)

$$\begin{gathered}
ES_{t+1}^2 = ES_t^2 + 4E(P_t S_t) - 2ES_t + 1 = \\
= t + 4\sum_{i=0}^{t-1} \sigma(i; p_0, 0, \lambda) - (2p_0 - 1)\sum_{i=0}^{t-1} \frac{1 - (2\lambda - 1)^i}{1 - \lambda} + \\
+4\sigma(t; p_0, 0, \lambda) - 2(2p_0 - 1)\frac{1 - (2\lambda - 1)^t}{2(1 - \lambda)} + 1 = \\
(t + 1) + 4\sum_{i=0}^{t} \sigma(i; p_0, 0, \lambda) - (2p_0 - 1)\sum_{i=0}^{t} \frac{1 - (2\lambda - 1)^i}{1 - \lambda},
\end{gathered}$$

which proves (12). Substituting (12) and (7) into (11) then proves the theorem.

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

$$Var\, S_t = +\infty$$

and

$$\lim_{t \to +\infty} (Var\, S_t - (c_1(p_0, \lambda)t + c_2(p_0, \lambda)) = 0,$$

where $c_i(p_0, \lambda)$ are some t-independent constants.

Corrolary 1 shows that with $t \to +\infty$ $Var\, S_t$ behaves as a linear function with respect to t. This can be seen also on Figure (1) together with a comparison of observed and theoretical values of $Var\, S_t$.

2.2 Success rewarding model

Similar formulas can be derived for the *success rewarding* model. Once again for previous results only the formulas are presented with proofs in the referred literature, new properties are derived with full complexity. For the sake of clarity the set of expressions is presented in the same manner as in the previous section.

For the expected value and variance of the step size for the $t \geq 1$ iteration of the walk X_t it holds [5]

$$EX_t = 2p_0 - 1, \tag{14}$$

$$Var\, X_t = 4p_0(1 - p_0). \tag{15}$$

For the expected value and variance of the transition probability or the $t \geq 1$ iteration of the walk P_t it holds [5]

$$EP_t = p_0, \tag{16}$$

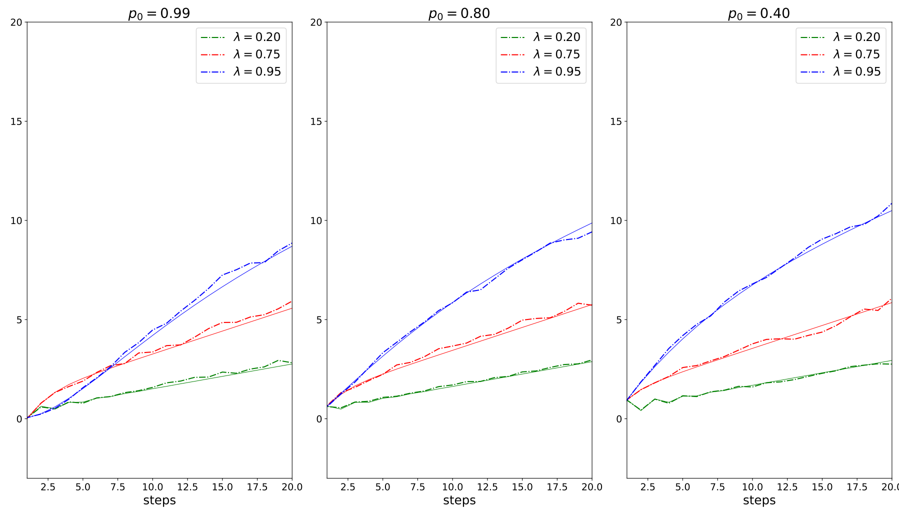

Fig. 1. Observed (dash-dotted) and theoretical (solid lines) values of $Var\,S_t$ - *success punishing* model. The data were obtained from 1000 walks generated with given parameters.

$$Var\,P_t = (2\lambda - \lambda^2)^t p_0^2 + p_0(1-\lambda)^2 \sum_{i=0}^{t-1}(2\lambda - \lambda^2)^i - p_0^2. \tag{17}$$

As the sum in the formula equals $\frac{1-(2\lambda-\lambda^2)^t}{1-(2\lambda-\lambda^2)}$, it can be further simplified as

$$\begin{aligned} Var\,P_t = (2\lambda - \lambda^2)^t p_0^2 + p_0(1-\lambda)^2 \frac{1-(2\lambda-\lambda^2)^t}{(1-\lambda)^2} - p_0^2 = \\ = p_0[(2\lambda-\lambda^2)^t(p_0-1)+1] - p_0^2, \\ Var\,P_t = p_0[(1-p_0)(1-(2\lambda-\lambda^2)^t)]. \end{aligned}$$

Finally, the expected position of the walker S_t after $t \geq 1$ iterations can be expressed as [5]

$$ES_t = S_0 + t(2p_0 - 1). \tag{18}$$

To prove a formula allowing to compute the variance of the position of the walker, let us again start with a support proposition.

Proposition 2. *For all* $t \geq 1$

$$E(P_t S_t) = p_0 S_0 + p_0 t + 2\lambda p_0(p_0-1)\frac{1-(2\lambda-\lambda^2)^t}{(1-\lambda)^2}. \tag{19}$$

Proof. We will once again start with expressing $E(P_tS_t)$ from the knowledge of the past step.

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

and using $E(X_t|P_{t-1}) = 2P_{t-1} - 1$ and $EX_t^2 = 1$ finally

$$E(P_tS_t) = E(P_{t-1}S_{t-1}) + 2\lambda EP_{t-1}^2 - (2\lambda - 1)EP_{t-1}. \tag{20}$$

Further we will continue using mathematical induction. For $t = 1$ using the definition of the walk it holds that

$$E(P_1S_1) = p_0(1-(1-p_0)\lambda)(S_0+1) + (1-p_0)\lambda p_0(S_0-1) =$$
$$= p_0S_0 + 2\lambda p_0^2 - (2\lambda-1)p_0.$$

When substituting $t = 1$ into (19) we obtain

$$E(P_1S_1) = p_0S_0 + p_0 + 2\lambda p_0(p_0-1)\frac{1-(2\lambda-\lambda^2)^0}{(1-\lambda)^2} =$$
$$= p_0S_0 + p_0 + 2\lambda p_0(p_0-1)$$

and finally

$$E(P_1S_1) = p_0S_0 + 2\lambda p_0^2 - (2\lambda-1)p_0.$$

Equation (19) thus holds for $t = 1$. Now for the induction step $t \to t+1$ we get by substituting(19) into (20)

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

and further using

$$EP_t^2 = p_0((2\lambda-\lambda^2)^t(p_0-1)+1),$$

which follows from the proof of Proposition 3.7 in [5], and (16)

$$E(P_{t+1}S_{t+1}) = p_0S_0 + p_0t + 2\lambda p_0(p_0-1)\frac{1-(2\lambda-\lambda^2)^t}{(1-\lambda)^2} +$$
$$+2\lambda p_0((2\lambda-\lambda^2)^t(p_0-1)+1) - (2\lambda-1)p_0 =$$
$$= p_0S_0 + p_0(t+1) + 2\lambda p_0(p_0-1)[\frac{1-(2\lambda-\lambda^2)^t}{(1-\lambda)^2} + (2\lambda-\lambda^2)^t] =$$
$$= p_0S_0 + p_0(t+1) + 2\lambda p_0(p_0-1)\frac{1-(2\lambda-\lambda^2)^{t+1}}{(1-\lambda)^2}.$$

Theorem 2. *For all $t \geq 1$ holds*

$$Var\, S_t = 4p_0(1-p_0)t^2 + a(p_0,\lambda)t - a(p_0,\lambda)\frac{1-(2\lambda-\lambda^2)^t}{(1-\lambda)^2},$$

where

$$a(p_0,\lambda) = \frac{8p_0(1-p_0)}{(1-\lambda)^2}.$$

Proof. As clearly the value S_0 does not affect the variance, let us from now assume $S_0 = 0$. From the definition of variance and (18) follows that to prove the theorem it is enough to prove that

$$ES_t^2 = t^2 + a(p_0,\lambda)t - a(p_0,\lambda)\frac{1-(2\lambda-\lambda^2)^t}{(1-\lambda)^2}. \tag{21}$$

First of all let us recall that formula (13) holds for the *success rewarding* type of the model as well. The theorem will be once again proved using mathematical induction. For $t = 1$ the definition of the walk yields the same result as in the proof of Theorem (1). By substituting $t = 1$ into (21) we obtain

$$ES_1^2 = 1 + a(p_0,\lambda)t - a(p_0,\lambda) = 1$$

and (21) thus holds for $t = 1$. Now for the induction step $t \to t+1$ we get by substituting (21), (19) and (18) into (13)

$$\begin{aligned}
ES_{t+1}^2 &= ES_t^2 + 4E(P_tS_t) - 2ES_t + 1 = \\
&= t^2 + a(p_0,\lambda)t - a(p_0,\lambda)\frac{1-(2\lambda-\lambda^2)^t}{(1-\lambda)^2} + \\
&+4(p_0t + 2\lambda p_0(p_0-1)\frac{1-(2\lambda-\lambda^2)^t}{(1-\lambda)^2}) - 2t(2p_0-1) + 1 = \\
&= (t+1)^2 + a(p_0,\lambda)(t+1) - a(p_0,\lambda)\frac{1-(2\lambda-\lambda^2)^{t+1}}{(1-\lambda)^2}.
\end{aligned}$$

Substituting (21) and (18) into the definition of variance then proves the theorem.

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

$$\lim_{t\to+\infty} Var\, S_t = +\infty$$

and

$$\lim_{t\to+\infty}\left(Var\, S_t - \left(4p_0(1-p_0)t^2 + a(p_0,\lambda)t - \frac{a(p_0,\lambda)}{(1-\lambda)^2}\right)\right) = 0,$$

with $a(p_0,\lambda)$ as in Theorem (2), i.e. $Var\, S_t$ behaves as a quadratic function with $t \to +\infty$.

Corollary 2 shows that with $t \to +\infty$ $Var\, S_t$ behaves as a quadratic function with respect to t. Similarly as with the *success punishing* model, such behavior is illustrated on Figure (2), which also shows the comparision of the theoretical value of position variance and an empirical one obtained using simulated data.

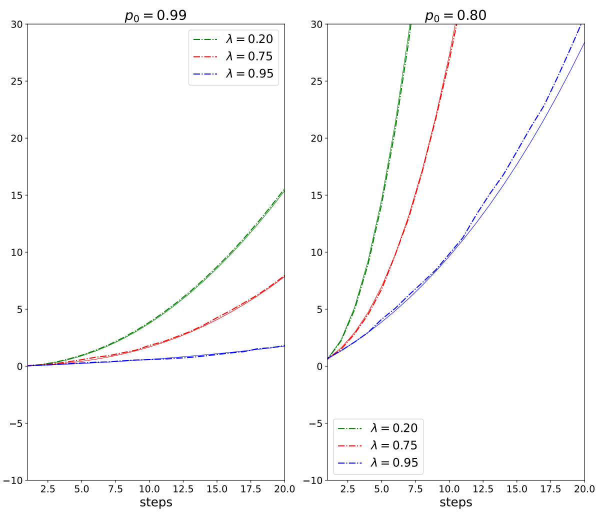

Fig. 2. Observed (dash-dotted) and theoretical (solid lines) values of $Var\, S_t$ - *success rewarding* model. The data were obtained from 1000 walks generated with given parameters.

2.3 Two-parameter models

The presented model can be further extended by adding additional levels of complexity. The first option is to use two separate λ parameters for each direction of the walk. Maintaining the two basic options - *success punishing* and *success rewarding* models, this level of complexity can defined as follows [5]

Definition 2. Let $p_0, \lambda_0, \lambda_1 \in (0, 1)$ be constant parameters, $\{P_n\}_{n=0}^{\infty}$ and $\{X_n\}_{n=1}^{\infty}$ sequences of discrete random variables with $P_0 = p_0$ and for $t \geq 1$

$$P(X_t = 1 | P_{t-1} = p_{t-1}) = p_{t-1}, \;\; P(X_t = -1 | P_{t-1} = p_{t-1}) = 1 - p_{t-1},$$

and (*success punishing)*

$$P_t = \frac{1}{2}[(1 + X_t)\lambda_0 P_{t-1} + (1 - X_t)(1 - \lambda_1(1 - P_{t-1}))] \qquad (22)$$

or (*success rewarding)*

$$P_t = \frac{1}{2}[(1 - X_t)\lambda_0 P_{t-1} + (1 + X_t)(1 - \lambda_1(1 - P_{t-1}))]. \qquad (23)$$

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 - two-parameter model*, with $\{X_n\}_{n=1}^{\infty}$ being the steps of the walker and $\{P_n\}_{n=0}^{\infty}$ transition probabilities. Depending on the chosen formula to calculate P_i the walk type is either *success punishing* (22) or *success rewarding* (23).

The derivation of exact model properties is not so straightforward as in the case with single lambda. The properties were thus studied with the help of simulations. Figures (3) and (4) present again the variance of S_t. It seems that the position variance of the *success punishing* model goes to infinity linearly and of the *success rewarding* model quadratically, similarly as in the corresponding single parameter scenarios.

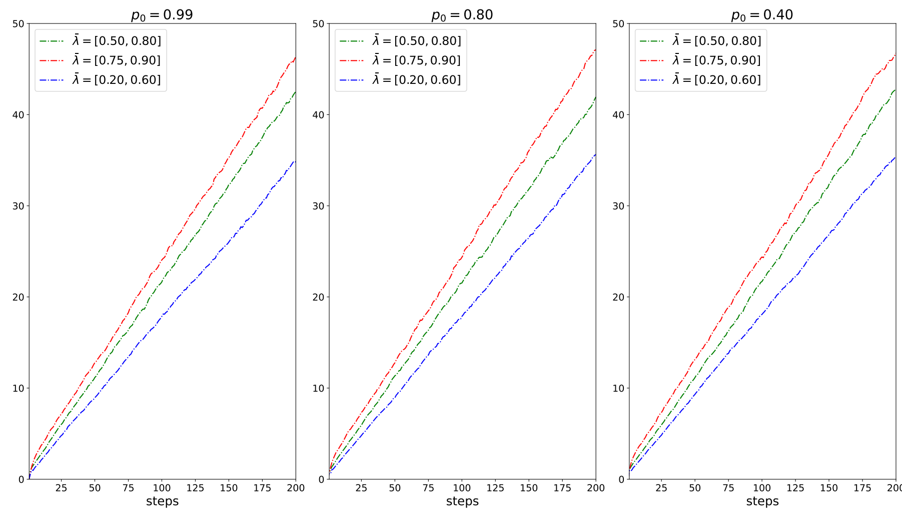

Fig. 3. The observed position variance in the *two-parameter success punishing* model. The data were gathered from 10000 simulations with given parameters.

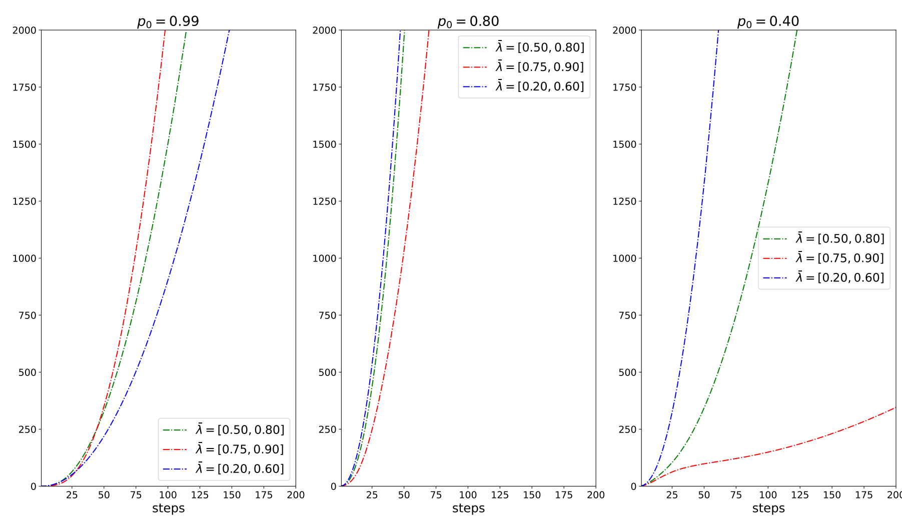

Fig. 4. The observed position variance in the *two-parameter success rewarding* model. The data were gathered from 10000 simulations with given parameters.

2.4 Other model alternatives

There are many possibilities how to further enhance the model. The model can be combined with a varying step size model from [8], the λ parameter can be handled as a function of time and position or a combination of varying probability model with regression part (e.g. logistic) can be considered, which seems especially promising from application point of view. These variations of the model will be subject of further study.

3 Model application

The model is especially well suited for simulation of random processes where a single or just a few events significantly affect the process's future development. An example such a process can be found in sports modelling. In such applications rather short walks occur, but they can be observed repetitively. For example in modelling tennis sets, the longest walk has only 5 steps (occurring in men Grand Slam or Davis Cup matches), but there are many matches played each year, which can be (under some assumptions) considered as multiple observations of the same walk. The authors recently presented a study where the model was used for modelling the men tennis Grand Slam matches with results suggesting the model might provide precious insights when modelling tennis. Here is a brief summary of the modelling approach.

The *success rewarding* version of the model was selected as the historical results show that the development of a tennis match follows such pattern. The p_0 was obtained using input bookmaker odds from *Pinnacle Sports*, an industry leading bookmaker. The appropriate λ parameter was then found from historical data using the maximal likelihood estimate. The model was tested on a database consisting of 4255 tennis matches that took place between 2009 and 2018 and the results suggest that such a model could be used for *in-play* odds prediction. For full details of the model derivation and testing, see the original paper [3].

The quality of such *in-play* predictions was tested on a small study in real life setting with active betting against a bookmaker. The model from [3] was implemented into an automatic odds scraping and betting tool. Whenever the odds provided by bookmaker a_i were higher than the model implied odds, i.e. $a_i > \frac{1}{p}$, a bet was made. This test was carried over the 2019 men tennis US Open and resulted in 128 placed bets with the total amount 59.85 units bet. As the bets were not placed simultaneously, but rather consecutively, the theoretical total bankroll needed for the betting was only 0.52 units - the minimal actual account balance over the entire US Open. The actual number of wins was 57, slightly below the number of expected wins, but thanks to the average winning odds of 2.3 the final balance was plus 2.24 units, creating a theoretical win $\frac{2.24}{0.52} = 4.3$ times the investment, which is an outstanding performance.

This study just briefly shows the possibilities of the presented model and presents rather encouraging results. The testing dataset, consisting of only 128 bets, is however too small to provide a strong evidence favoring the model over bookmaker's odds. A more conclusive test on a bigger dataset will be subject of further study.

4 Conclusion

The present paper continues in the research on one specific set of models of Bernoulli-like random walks, the models where the transition probabilities depend on the walk history. After reviewing basic models types and the results of previous studies of their properties, new properties of the model characterizing the variability of the walk were proved. These properties were explored additionally with the aid of simulations in order to compare derived theoretical results with empirical ones based on simulated data. The problem of parameters estimation was not addressed here, the properties of the maximum likelihood estimate of both λ and p_0 were studied in depth in [5] and utilized in the [3], a study devoted to one of possible model applications, namely the modelling tennis matches development. This kind of application was briefly recalled also in the present contribution and partial results of model testing on a new dataset were reported.

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. MR3586234.
2. Tomáš Kouřim. Random walks with varying transition probabilities. *Doktorandské dny FJFI*, 2017. Available at http://kmwww.fjfi.cvut.cz/ddny/historie/17-sbornik.pdf.
3. 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.
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. Discrete random processes with memory: Models and applications. *Applications of Mathematics, accepted for publishing*, 2020.
6. Karl Pearson. The problem of the random walk. *Nature*, 72(1865):294, 1905.
7. Frank Spitzer. *Principles of random walk*, volume 34. Springer Science & Business Media, 2013.
8. 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. MR2658904.