Discrete random processes with memory: Models and applications

Tomáš Kouřim

November 24, 2019

Abstract

The contribution focuses on non-Markov random processes, i.e. processes with memory, and is especially concerned with random processes where either one or just a small number of events significantly affects its future development. Reliability analysis, medical studies or sport statistics provide many real-life examples of such processes. This paper focuses on statistical event-history analysis with discretized time, where the role of the hazard rate can be substituted by different variants of transition probabilities. In the simplest case the Bernoulli scheme is used.

The main concern of the paper is the formulation of models describing the dependence of transition probabilities on the process development, as well as on exogenous factors – covariates. 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. In more complicated cases, as well as in the presence of exogenous covariates, the changes of probabilities are modeled via a regression model, for instance the logistic one. The behavior of proposed random walks is studied both theoretically and with the aid of simulations. Finally, the approach is illustrated on several real data examples.

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 types of the random variables and their mutual interactions random processes can be splitted into a large nuber of different categories. The processes vary in the way of development, their memory, underlying domain and many other aspects. One example of a random process can be found in reliability analysis, where complex, non-homogeneous non-Markov random processes play an important role [5]. Such processes help to model failure times in complex reliability systems, incorporating their corrective maintenance and preventive maintenance and their predictions and simulations [1, 4, 3, 8]. A most prominent model in reliability analysis is the Cox's model using a so called *hazard rate* or *intensity of failure* as a key parameter describing the process [2].

Most common type of a discrete random process is a random walk, a mathematical object first introduced by K.Pearson in 1905 [9]. Similarly to random processes in general, there exist many well described variations of a random walk with various applications to real life problems [11, 10]. Yet there are still new possibilities and options how to alter and improve the classical random walk and present yet another new 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 togehter with the Cox's model and the reliability analysis in general served as an inspiration to the random walk with varying probabilities introduced by Kouřim [6, 7]. This novel variant of a random walk presents a discrete alternative to a Cox's model and has possible aplications in reliability analysis and many other real life problems.

In this paper, the theoretical properties of the model are described and further examined and the results are tested on generated data. The rest of the paper is organized as follows....

2 Random walk with varying probabilities

The random walk with varying probabilities is based on a standard Bernoulli (zdroj) random walk 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 [7]

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 [7], 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 are described in [7], namely that

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

and

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

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 firstprove the following suport propositions.

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

$$Var(P_t) = (3\lambda^2 - 2\lambda)^t p_0^2 + \sum_{i=1}^{t} K(i-1)(3\lambda^2 - 2\lambda)^{t-i} - k(t)^2, \tag{4}$$

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

$$Var(P_t) = E(P_t^2) - E(P_t)^2. \tag{5}$$

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

$$E(P_t^2) = (3\lambda^2 - 2\lambda)^t p_0^2 + \sum_{i=1}^{t} K(i-1)(3\lambda^2 - 2\lambda)^{t-i}. \tag{6}$$

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

$$E(P_t^2) = E[E(P_t^2|P_{t-1}^2)] = E[E(P_t^2|P_{t-1})] =$$
$$= E[E(\lambda P_{t-1} + \frac{1}{2}(1-\lambda)(1-X_t))^2|P_{t-1}]. \tag{7}$$

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 7 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

$$E(P_t^2) = E(P_{t-1}^2)(3\lambda^2 - 2\lambda) + EP_{t-1}(-3\lambda^2 + 4\lambda - 1) + (1-\lambda)^2. \tag{8}$$

Now 6 can be proved using induction. Based on the trivial fact that $Ep_0 = p_0$ and $E(p_0)^2 = p_0^2$, for $t = 1$

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

and from 8 follows that the induction assumption holds. Now for the induction step $t \to t+1$ we get by substituing 6 into 8

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

and the formula thus holds. Now substituing 3 and 6 into 5 yields 4 and proves the Proposition. $\square$

And similarly for the variance of the position of the walker, following statements can be proved.

Proposition 2. *For* $\forall t \geq 1$, *it holds that*

$$Var(S_t) = TODO$$

Proof. TODO □

3 Random walk with varying transition probability - alternatives

3.1 Success rewarded

The basic definition of the random walk (1) presents a model, where the "success is punished", 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 [7].

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

$$P_1 = \lambda p_0 + \frac{1}{2}(1-\lambda)(1+X_1) \tag{9}$$

$$P_i = \lambda P_{i-1} + \frac{1}{2}(1-\lambda)(1+X_i) \;\; \forall i \geq 2. \tag{10}$$

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.

This version of the random walk presents a model where "success is rewarded". For the sake of clarity let us denote all variables connected with the "success punished" model from Definition 1 with the subscript P (as ${}_PP$, ${}_PS$) and the "success rewarded" version from Definition 2 with subscript R (as ${}_RP$, ${}_RS$). For the reward version it is easy to prove similar proposition as for the punish version.

Proposition 3. *For* $\forall t \geq 2$,

$${}_RP_t = p_0\lambda^t + \frac{1}{2}(1-\lambda)\sum_{i=1}^{t}\lambda^{t-i}(1+X_i) \tag{11}$$

Proof. The proposition is proved using induction. For $t = 2$ using 1 it holds that

$${}_RP_2 = \lambda {}_RP_1 + \frac{1}{2}(1-\lambda)(1+{}_RX_2) = \lambda(\lambda p_0 + \frac{1}{2}(1-\lambda)(1+X_1)) + \frac{1}{2}(1-\lambda)(1+{}_RX_2) =$$

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

which is in accordance with 11. Now for the induction step $t \to t+1$ we obtain from 2 and the induction assumption

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

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

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

□

Proposition 4. *For* $\forall t \geq 1$, $E({}_RP_t) = p_0$.

Proof. Using $E({}_RX_t|{}_RP_{t-1}) = 2{}_RP_{t-1} - 1$ and 10 we obtain

$$E_RP_t = E[E({}_RP_t|{}_RP_{t-1})] = E[E(\lambda_RP_{t-1} + \frac{1}{2}(1-\lambda)(1+X_t)|{}_RP_{t-1})] =$$

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

$$= E_RP_{t-1}.$$

Recursively we get

$$E_RP_t = Ep_0 = p_0.$$

□

Now to calculate the position of the walker at a given step $t \geq 1$, it is easy to see that $E({}_RS_t) = {}_RS_{t-1} + 2{}_RP_{t-1} - 1$. From this, we can prove the following statement about the expected position of the walker.

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

$$E({}_RS_t) = {}_RS_0 + t(2p_0 - 1).$$

Proof. Using the result of Proposition 4 we get

$$E({}_RS_{t+1}) = E[E({}_RS_{t+1}|{}_RS_t)] = E[{}_RS_t + (2{}_RP_{t-1} - 1)] =$$

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

which recursively proves the statement. □

Proposition 6. *For* $\forall t \geq 1$,

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

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

$$Var(P_t) = E(P_t^2) - E(P_t)^2. \tag{13}$$

From Proposition 4 follows $E(P_t) = p_0$ and it is thus sufficient to prove that

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

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

$$\begin{aligned} E(P_t^2) &= E[E(P_t^2|P_{t-1}^2)] = E[E(P_t^2|P_{t-1})] = \\ &= E[E(\lambda P_{t-1} + \frac{1}{2}(1-\lambda)(1+X_t))^2|P_{t-1}] = \\ &= EP_{t-1}^2(2\lambda - \lambda^2) + p_0(1-\lambda)^2, \end{aligned} \tag{15}$$

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 4 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 15 and the induction assumption

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

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

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

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

The Proposition statement is then obtained by substituin 4 and 14 into 13. □

Proposition 7. *For* $\forall t \geq 1$, $Var({}_R S_t) = 0$.

Proof. Tipuju □

vytvorit nove simulace a obrazky ukazani jednotlivych hodnot, protovnani - pozor na floating point zaokrouhlovani

3.2 Two lambdas

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

$$P_1 = \frac{1}{2}[(1 + X_1)\lambda_0 p_0 + (1 - X_1)(1 - \lambda_1(1 - p_0))] \tag{16}$$

$$P_i = \frac{1}{2}[(1 + X_i)\lambda_0 P_{i-1} + (1 - X_i)(1 - \lambda_1(1 - P_{i-1}))] \ \ \forall i \geq 2. \tag{17}$$

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.

And the success rewarded version as

Definition 4. 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

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

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

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.

TODO mel bych to rovnou nejak znacit ${}^1_R P$ a ${}^2_S P$ jako reward s 1 lambda a success s 2 lambda? Abych se zas neuindexoval

Let us prove the following propositions describing the properties of a random walk affected by two coefficients lambda.

TODO stejne veci opet dokazat pro tuhle variantu

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.

	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.

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 $\bar{\lambda} = \{[0.5, 0.8], [0.5, 0.99], [0.99, 0.9]\}$. The starting transition probability was chosen from the set $p_0 = \{0.5, 0.8, 0.9, 0.99\}$ and the length of the walk was $n = \{5, 10, 50, 100\}$. 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 and again Python language with Numpy 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. Ctvrty ukol bych mohl udela pomoci AIC, tam by to treba dopadlo lepe. Uvidim, jak budu stihat.

Nasimulovani ruznych druhu prochazky s ruznymi parametry. Nasledne pokus o zpetne odhaleni druhu prochazky a jejich paramteru. Budu to delat zrejme podle MLE, vyhodnocovat asi podle te nejvetsi verohodnosti, pripadne podle goodness-of-fit

5 Conclusion

This work follows up on the recent results on random walks with varying probabilities. It describes and proves more properties of such a walk, enhancing the theoretical mathematical device available to the study of the novel type of a random walk. The study also show the quality of the described model by successfully fitting it on a set of randomly generated data. The presented model has also many possible uses in real life application. It is a discrete alternative of a non-Markov random process, i.e. a process with memory. The modified random walk describes especially well processes where either a single or just a small number of events can significantly affect future development of the process. Such processes can be found in reliability analysis, medical 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 (zdroj Ateny).

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] AH Christer and WM Waller. An operational research approach to planned maintenance: modelling pm for a vehicle fleet. *Journal of the Operational Research Society*, 35(11):967–984, 1984.

[2] David R Cox. Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2):187–202, 1972.

[3] DR Cox and PAWL Lewis. The statistical analysis of series of events. 1966.

[4] Nat Jack and DNP Murthy. A new preventive maintenance strategy for items sold under warranty. *IMA Journal of Management Mathematics*, 13(2):121–129, 2002.

[5] John D Kalbfleisch and Ross L Prentice. *The statistical analysis of failure time data*, volume 360. John Wiley & Sons, 2011.

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

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

[8] William Q Meeker and Luis A Escobar. *Statistical methods for reliability data*. John Wiley & Sons, 2014.

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

[10] 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.

[11] 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.