Non-markov discrete random walks - selected properties

April 17, 2020

Abstract

This paper elaborates the newly introduced model of discrete random walks with memory implemented through a memory coefficient λ. It follows on the recent work of the authors and further examines the theoretical properties of the model.

1 Introduction

Modelling sporting events and reliability analysis are two seemingly very different areas of mathematics, yet they offer many similarities. The key factor in both domains is a certain probability measure determining the development of observed phenomenon. This probability measure is called "hazard rate" in reliability analysis and simply "probability of scoring" in modelling of sport events. It is altered either continuously with time according to some underlying probability function or suddenly as a reaction to an event. Breakage or repair of a given machine in reliability analysis or scoring (goal, point, basket etc.) in sports modelling. The most prominent tool in reliability analysis is the Cox's model ZDROJ expressing the changes of a hazard rate with a base hazard rate and a covariate vector. This model, operating with continuous time, is often used for sports modelling as well. Many types of sport such as tennis or volleyball, however, are played in a strictly discrete matter with steps divided by individual *points*, *games* or *sets*. Other real life situations with similar properties can be found as well, including the recurrence of diseases or recidivism in crime. Thus authors thus recently presented a novel discrete time random process model - discrete random walk with varying probabilities - which aims to serve as a discrete alternative to the standard Cox's model ZDROJ. In this paper 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.... The last section concludes this work.

2 Random walk with varying probability - overview

Shrnuti dosavadnich vysledku a vypoctenych vzorcu

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 has 2 parameters, starting success probability p_0 and a memory coefficient λ affecting the severity of probability change after a success. Formally the walk is defined as follows:

Definition 1. Let $p_0, \lambda \in (0,\, 1)$ 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 $i \geq 1$

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

and (*success punishing)*

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

or (*success rewarding)*

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

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

The model was first introduced in DDNY and a more thorough description was provided in TEZE. More properties were then presented in AMISTAT and a practical implementation of the model in modelling tennis matches was presented in IMAMAN+ProceedingsAteny. Next let us recall the results of the previous work.

2.1 Success punishing model

The basic properties of the *success rewarding* version of the walk are presented in this section. They are presented as a set of expressions, the reader is kindly asked to see reffered papers for full proves of those expressions.

For the expected value and variance of the step size in the $t \geq 1$ iteration of the walk X_t it holds

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

3 Random walk with varying probability - properties

Dukazy novych veci, hlavni cast clanku

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

$$Var\, S_t = neco.$$

Proof. From the definition of variance

$$Var\, S_t = ES_t^2 - E^2 S_t$$

and XXXXES_t vzorecXXX follows that to prove the proposition it is enough to prove that

$$ES_t^2 = neco.$$

First of all, let us express $E(S_t^2)$ from the knowledge of past walk development. From the definition of the expected value and the definition of the walk XXXXXX it follows

$$ES_t^2 = E[E(S_t^2|P_{t-1})] = E[E((S_{t-1}+X_t)^2|P_{t-1})] = E(S_{t-1}^2+2(2P_{t-1}-1)S_{t-1}+1) =$$
$$= E(S_{t-1}^2 + 4P_{t-1}S_{t-1} - 2S_{t-1} + 1).$$

The values of ES_t and EP_t are given by XXXXXXX. The only unknown element (clen rovnice - jak se to rekne anglicky?) is the mixed one. Let us again express it from the knowledge of past steps as

$$E(P_t S_t) = E[E(P_t S_t|P_{t-1})] = E[E((\lambda P_{t-1}+\frac{1}{2}(1-\lambda)(1-X_t))(S_{t-1}+X_t)|P_{t-1})] =$$
$$= E[(\lambda P_{t-1} + \frac{1}{2}(1-\lambda)(1-(2P_{t-1}-1)))(S_{t-1}+(2P_{t-1}-1))] =$$
$$= E[(\lambda P_{t-1} + (1-\lambda)(1-P_{t-1}))(S_{t-1}+2P_{t-1}-1)] =$$
$$= E[(\lambda P_{t-1} + 1 - P_{t-1} - \lambda + \lambda P_{t-1})(S_{t-1}+2P_{t-1}-1)] =$$
$$= E[(\lambda P_{t-1}S_{t-1}+S_{t-1}-P_{t-1}S_{t-1}-\lambda S_{t-1}+\lambda P_{t-1}S_{t-1}+2\lambda P_{t-1}^2+2P_{t-1}-2P_{t-1}^2-2\lambda P_{t-1}+2\lambda P_{t-1}^2-$$
$$-\lambda P_{t-1} - 1 + P_{t-1} + \lambda - \lambda P_{t-1})] =$$
$$= E((2\lambda-1)P_{t-1}S_{t-1} + (1-\lambda)S_{t-1} + 2(2\lambda-1)P_{t-1}^2 + (3-4\lambda)P_{t-1} - 1 + \lambda).$$

Together we get

$$ES_t^2 = E(S_{t-1}^2+4((2\lambda-1)P_{t-2}S_{t-2}+(1-\lambda)S_{t-2}+2(2\lambda-1)P_{t-2}^2+(3-4\lambda)P_{t-2}-1+\lambda)-2S_{t-1}+1)$$
$$= E(S_{t-1}^2+4((2\lambda-1)P_{t-2}S_{t-2}+(1-\lambda)S_{t-2}+2(2\lambda-1)P_{t-2}^2+(3-4\lambda)P_{t-2}))-2S_{t-1})-3+4\lambda.$$

□

4 Conclusion