Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁRSKA PRÁCA

Alena Babiaková

Náhodná prechádzka

Katedra pravděpodobnosti a matematické statistiky
Vedúci bakalárskej práce: Prof. RNDr. Josef Štěpán, Dr. Sc

Študijný program: Matematika
Študijný obor: Obecná matematika

2006

Moje poďakovanie patrí predovšetkým Prof. RNDr. Josefovi Štěpánovi. A potom taktiež Tomášovi Maradovi.

Prehlasujem, že som svoju bakalársku prácu napísala samostatne a výhradne s použitím citovaných prameňov. Súhlasím so zapožičiavaním práce a jej zverejňovaním.

V Prahe dňa 16.4.2006 Alena Babiaková

Obsah

Názov práce: Náhodná prechádzka
Autor: Alena Babiaková
Katedra (ústav): Katedra pravděpodobnosti a matematické statistiky
Vedúcí bakalárskej práce: Prof. RNDr. Josef Štěpán, Dr. Sc
e-mail vedúceho: `Josef.Stepan@mff.cuni.cz`

Abstrakt: V tejto práci sa zaoberáme symetrickou náhodnou prechádzkou po $\mathbb{Z}$. Jeden zo spôsobov, ako si ju možno predstaviť, je náhodný pohyb častice, ktorá sa v čase 0 nachádzala v počiatku celočíselnej priamky, v každom časovom okamihu častica vykoná pohyb (nezávisle na pohybe v predchádzajúcich časových okamihoch) o jednotku doľava alebo doprava. Ak sú pravdepodobnosti pohybu doprava a doľava rovnaké, potom tento pohyb nazývame symetrickou náhodnou prechádzkou. Symetrickú náhodnú prechádzku skúmame kombinatoricky (princíp zrkadlenia) a jej asymptotické vlastnosti metódami teórie pravdepodobnosti. Práca vrcholí podrobným dôkazom zákona arkusínu.

Kľúčové slová: Symetrická náhodná prechádzka po celočíselnej priamke v diskrétnom čase, prvý návrat do nuly, maximum náhodnej prechádzky, princíp zrkadlenia, zákon arkusínu.

Title: Random Walk
Author: Alena Babiaková
Department: Department of Probability and Mathematical Statistics
Supervisor: Prof. RNDr. Josef Štěpán, Dr. Sc
Supervisor's e-mail address: `Josef.Stepan@mff.cuni.cz`

Abstract: Our topic is the symmetric random walk on $\mathbb{Z}$. A possible way how to visualize the motion is to consider the life of a particle placed in the origin $k = 0$ at the time $n = 0$. Each unit of time the particle makes a random step to the left or to the right, independently of its previous motions. If the particle is in the position k at the time n seconds, it is to be found either in the position $k + 1$ or the position $k - 1$ at the time $n + 1$, with a given pair probabilities p and $1 - p$. If both these probabilities equal $\frac{1}{2}$, then the movement of the particle is called the symmetric random walk. We study the symmetric random walk using both combinatorial methods (the reflection principle) and the methods of more advanced probability theory. Our work culminates in the detailed proof of the arcsin law.

Keywords: the symmetric random walk on $\mathbb{Z}$ in discrete time, first returns, the maximum of the random walk, the reflection principle, the arcsin law.

Kapitola 1

Základné pojmy a vlastnosti

Definícia 1.1. **Náhodná prechádzka** je postupnosť čiastočných súčtov

$$S_n = \sum_{j=1}^{n} X_j, \qquad n \in \mathbb{N},$$

kde $X_1, X_2, \ldots$ sú nezávislé rovnako rozdelené náhodné veličiny.
Ak sú X_k, $k \in \mathbb{N}$, náhodné veličiny s hodnotami v $(\mathbb{R}^m, \mathcal{B}^m)$, potom hovoríme, že $\{S_n\}_{n=1}^{\infty}$ je **náhodná prechádzka v** $\mathbb{R}^m$. $\triangle$

Existuje niekoľko spôsobov, ako si možno náhodnú prechádzku predstaviť. Jeden z nich je napríklad náhodný pohyb častice v $\mathbb{R}^m$, ktorá sa v čase $n = 0$ nachádza v počiatku a ktorá sa môže pohybovat o určitú vzdialenosť v d rôznych smeroch, pravdepodobnosti týchto pohybov sú dané rozdelením X_k, $k = 1, \ldots, d$. S_n potom udáva pozíciu častice v čase n.

Definícia 1.2. Ak pre náhodnú prechádzku platí

$$P[X_1 = -1] = P[X_1 = 1] = \frac{1}{2},$$

potom hovoríme o **diskrétnej symetrickej náhodnej prechádzke**. $\triangle$

Diskrétnu symetrickú náhodnú prechádzku si môžeme predstaviť ako pohyb častice (v čase $n = 0$ umiestnenej v počiatku) po celočíselnej priamke v diskrétnom čase, v každom časovom okamihu častica vykoná pohyb buď o jednotku doprava alebo doľava, pravdepodobnobnosť pohybu doľava a doprava je rovnaká, tj. rovná $\frac{1}{2}$, a pohyb častice v j-tom časovom okamihu nezávisí na jej pohyboch v predchádzajúcich $j - 1$ časových okamihoch.

Ďalším modelom diskrétnej symetrickej náhodnej prechádzky je hra dvoch hráčov pozostávajúca z postupnosti partií, ktoré sú nezávislé a ich výsledok (tj.pravdepodobnosť, že v danej partii vyhrá prvý, respektíve druhý hráč) závisí na rozdelení náhodných veličín X_1, respektíve X_2. Predpokladáme, že partia nemôže skončiť remízou. Hráč, ktorý v danom kole prehral, odovzdá protihráčovi 1 peniaz. S_n predstavuje výhru prvého hráča po n partiách, výhra druhého hráča je potom $-S_n$.

Definícia 1.3. Pre $n \in \mathbb{N}$ konečné možno na klasickom pravdepodobnostnom priestore s $\Omega = \{0,1\}^n$ **náhodnú prechádzku** definovať ako konečnú n-člennú postupnosť $(i_1, i_2, \ldots, i_n)$, kde $i_j = 1$, respektíve $i_j = 0$, znamená, že v čase $j-1$ častica vykonala pohyb o jednotku doprava, respektíve doľava. △

Podľa predchádzajúcej definície môžeme náhodnú prechádzku generovať napríklad hádzaním mincou: ak padne rub, pošleme časticu doprava, ak líc, pošleme ju doľava.

Definícia 1.4. Ekvivalentne možno pre konečné $n \in \mathbb{N}$ náhodnú prechádzku modelovať klasickým pravdepodobnostným priestorom (predpokladáme, že $S_0 = 0$) s

$$\Omega = \{(S_0, S_1, S_2, \ldots, S_n), |S_{j+1} - S_j| = 1\},$$

kde S_j predstavuje polohu častice v čase j, a teda $(S_0, S_1, S_2, \ldots, S_n)$ je zápis trajektórie, ktorú častica prešla za n časových okamihov. △

Niekedy je vhodná reprezentácia náhodnej prechádzky pomocou po častiach lineárnej lomenej čiary, ktorá graficky znázorňuje trajektóriu častice: vodorovná os predstavuje čas n a na zvislú os nanášame hodnoty S_n. Lomenú čiaru získame tak, že úsečkami pospájame body (S_k, k) a $(S_k+1, k+1)$.

Náhodná veličina S_n určuje polohu častice v čase n, táto poloha je jednoznačne určená tým, koľko z n pohybov vykonala častica smerom doprava. Označme tento počet pohybov doprava k, $0 \leq k \leq n$. Potom počet pohybov doľava je rovný $n-k$ a konečná poloha častice je $k-(n-k) = 2k-n$. S_n je súčtom n nezávislých náhodných veličín s alternatívnym rozdelením $Alt(\frac{1}{2})$, teda má binomické rozdelenie $Bi(n, \frac{1}{2})$ a $P[S_n = 2k-n] = \binom{n}{k} 2^{-n}$ pre $0 \leq k \leq n$.

Veta 1.5. *Pre náhodnú prechádzku* $\{S_n\}_{n=1}^{\infty}$ *platí:* $ES_n = 0, var\, S_n = n$.

Dôkaz. Zavedieme si transformáciu $Y_j = \frac{X_j+1}{2}$, $j \in \mathbb{N}$. Platí, že

$$\begin{array}{rcccl} P[Y_j = 0] & = & P[X_j = -1] & = & \frac{1}{2}, \\ P[Y_j = 1] & = & P[X_j = 1] & = & \frac{1}{2}. \end{array}$$

Teda $Y_j \sim Alt(\frac{1}{2})$. Podľa pravidiel pre počítanie so strednou hodnotou (a s využitím toho, ze pre náhodnú veličinu $U \sim Alt(p)$ je $EU = p$):

$$\begin{aligned} EX_j &= E[2Y_j - 1] = 2E[Y_j] - 1 = 2\frac{1}{2} - 1 = 0, \\ ES_n &= E\sum_{j=1}^{n} X_j = \sum_{j=1}^{n} EX_j = 0. \end{aligned}$$

Podobne pre rozptyl (využívame, že ak $U \sim Alt(p)$, potom $var\, U = p(1-p)$):

$$\begin{aligned} var\, X_j &= var\, (2Y_j - 1) = 4 var\, Y_j = 4\frac{1}{4} = 1, \\ var\, S_n &= var \sum_{j=1}^{n} X_j = \sum_{j=1}^{n} var\, X_j = n, \end{aligned}$$

kde druhá rovnosť je dôsledok nezávislosti veličín $X_1, X_2, \dots$ □

Ak $S_n = 0$, potom hovoríme, že v čase n nastal návrat do počiatku. Zrejme to môže nastať, iba ak je n párne číslo.
Preto pre nepárne n je pravdepodobnosť návratu do počiatku v čase n rovná nule. Pravdepodobnosť návratu do nuly v čase $2m$ spočítame ako podiel počtu všetkých trajektórií dĺžky $2m$, ktoré začínajú a končia v počiatku, a počtu všetkých trajektórií dĺžky $2m$. Teda,

$$P[S_{2m} = 0] = \binom{2m}{m} 2^{-2m}$$

a po použití Stirlingovej formule

$$n! = n^n \sqrt{2\pi n}\, e^{-n}(1 + \epsilon_n), \text{ kde } \lim_{n\to\infty} \epsilon_n = 0$$

dostaneme:

$$P[S_{2m} = 0] = \frac{1}{\sqrt{\pi m}}(1 + \delta_m), \text{ kde } \lim_{m\to\infty} \delta_m = 0. \tag{1.1}$$

Teda

$$P[S_{2m} = 0] \doteq \frac{1}{\sqrt{\pi m}}.$$

Pre porovnanie presnosti tejto aproximácie odkazujeme na [2], strana 17, tabuľka 1.5 b).

Ak $S_{2m} = 0$, $m > 0$ a $S_{2k} \neq 0$ pre všetky $k < n$, potom hovoríme, ze v čase n nastal prvý návrat do počiatku.

Definícia 1.6. Zobrazenie $t^* : \Omega \to \mathbb{N} \cup \{\infty\}$ nazývame **markovský čas** náhodnej prechádzky $\{S_n\}$, ak platí

$$[t^* = n] \in \sigma(S_1, \dots, S_n), \quad \text{kde} \quad n = 1, 2, \dots.$$

△

Kapitola 2

Ďalšie tvrdenia o náhodnej prechádzke

Veta 2.1. *Označme*

$$T_k = \{S = (S_1, S_2, \dots, S_k) : |S_1| = 1, |S_j - S_{j-1}| = 1, 2 \leq j \leq k\}.$$

Pre $S \in T_k$ *nech je* $m(S) = \min\{1 \leq j \leq k : S_j = a\}$, *kde* a *je celé číslo,* $-k \leq a \leq k$, *definujme:*

$$\begin{aligned} Z_a^+(S) &= (S_1, S_2, \dots, S_{m(S)}, a - (S_{m(S)+1} - a), \dots, a - (S_k - a)), \\ Z_a^-(S) &= (a - (S_1 - a)), \dots, a - (S_{m(S)} - a)), S_{m(S)+1}, \dots, S_k). \end{aligned}$$

Potom sú $Z_a^+(S)$ *a* $Z_a^-(S)$ *vzájomne jednoznačné zobrazenia množiny* $D_a = \{S \in T_k : S_j = a$ *pre niektoré* $j\}$ *samej na seba.*

Práve uvedený poznatok sa označuje ako princíp reflexie alebo princíp zrkadlenia.

Nasledujúcu vetu môžeme interpretovať ako postavenie bariéry do cesty častice v nejakom celočíselnom bode a, kde $1 \leq a \leq n$, a zaujíma nás pravdepodobnosť toho, že častica v niektorom časovom okamihu $j = 1, 2, \dots, n$ prekročí túto bariéru.

Veta 2.2. *Na postupnosti* $\{S_n\}_{n=1}^{\infty}$ *definujeme náhodnú veličinu*

$$M_n = \max\{S_j : 0 \leq j \leq n\}. \tag{2.1}$$

Nech a *je celé číslo,* $a \geq 0$. *Potom*

$$P[M_n \geq a] = 2P[S_n > a] + P[S_n = a]. \tag{2.2}$$

Dôkaz. Náhodný jav $[M_n \geq a]$ si napíšeme ako zjednotenie troch disjunktných javov:

$$[M_n \geq a] = [M_n \geq a, S_n > a] \cup [M_n \geq a, S_n < a] \cup [M_n \geq a, S_n = a].$$

Preto

$$P[M_n \geq a] = P[M_n \geq a, S_n > a] + P[M_n \geq a, S_n < a] + P[M_n \geq a, S_n = a].$$

Definujeme si zobrazenie množiny trajektórií

$$\{(S_0, S_1, \ldots, S_n) : M_n \geq a, S_n > a)\} \tag{2.3}$$

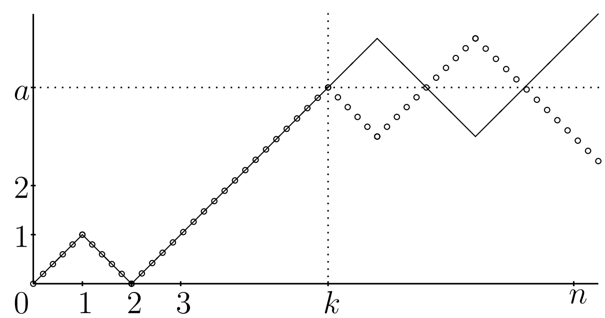
na množinu

$$\{(S_0, S_1, \ldots, S_n) : M_n \geq a, S_n < a)\}, \tag{2.4}$$

a to nasledujúcim spôsobom:
Označme pre dané $(S_0, S_1, \ldots, S_n)$ (ktoré je prvkom množiny (2.3))

$$k = \min\{j : 1 \leq j \leq n-1, S_j = a, M_n \geq a, S_n > a\}.$$

Takéto k musí existovať, keďže $S_0 = 0$ a $M_n \geq a$.



Obrázok 2.1: Dvojica trajektórií z množín (2.3) a (2.4)

Zobrazenie pre časť trajektórie

$$(S_0, S_1, \ldots, S_k)$$

definujeme ako identitu (viz. obrázok 2.1), teda pre každú takúto k-ticu ide o vzájomne jednoznačné zobrazenie. Pre

$$(S_{k+1}, \ldots, S_n)$$

zobrazenie definujeme ako osovú súmernosť podľa osi $y = a$ (viz. obrázok 2.1), a teda podľa vety 2.1 ide o vzájomne jednoznačné zobrazenie. Preto práve definované zobrazenie množiny (2.3) na množinu (2.4) je vzájomne jednoznačné zobrazenie množiny (2.3) na množinu (2.4). A teda dostávame:

$$P[M_n \geq a, S_n > a] = P[M_n \geq a, S_n < a].$$

Keďže z $S_n > a$, respektíve $S_n = a$, plynie, že $M_n = \max\{S_j : 0 \leq j \leq n\} \geq a$, je

$$P[M_n \geq a] = 2P[M_n \geq a, S_n > a] + P[M_n \geq a, S_n = a] = 2P[S_n > a] + P[S_n = a],$$

čo je tvrdenie vety. $\square$

Veta 2.3. *Pre rozdelenie náhodnej veličiny $\frac{M_n}{\sqrt{n}}$ platí:*

$$P\left[\frac{M_n}{\sqrt{n}} \leq x\right] \xrightarrow{n\to\infty} 2\Phi(x) - 1 \quad \text{pre} \quad x \geq 0,$$

kde $\Phi(x)$ je distribučná funkcia $N(0,1)$.

Dôkaz. S využitím vzťahu (2.2) máme

$$P\left[\frac{M_n}{\sqrt{n}} > x\right] = P\left[M_n \geq \lfloor x\sqrt{n}\rfloor + 1\right] = 2P\left[S_n > \lfloor x\sqrt{n}\rfloor + 1\right] + P\left[S_n = \lfloor x\sqrt{n}\rfloor + 1\right].$$

Z Lokálnej limitnej vety pre náhodné veličiny s rešetovitým rozdelením (viz. [1], V.2.7) plynie

$$\lim_{n\to\infty} P\left[S_n = \lfloor x\sqrt{n}\rfloor + 1\right] = 0.$$

Použitím Berry-Essenovej nerovnosti (viz. [1], V.4.5) a toho, že

$$\lim_{n\to\infty} \frac{\lfloor x\sqrt{n}\rfloor + 1}{\sqrt{n}} = x,$$

dospejeme k tomu, že

$$P\left[S_n > \lfloor x\sqrt{n}\rfloor + 1\right] = P\left[\frac{S_n}{\sqrt{n}} > \frac{\lfloor x\sqrt{n}\rfloor + 1}{\sqrt{n}}\right] \xrightarrow{n\to\infty} \quad 1 - \Phi(x).$$

A teda

$$P\left[\frac{M_n}{\sqrt{n}} \leq x\right] \xrightarrow{n\to\infty} 1 - 2(1 - \Phi(x)) = 2\Phi(x) - 1.$$

□

Lemma 2.4 (Bertran)**.** *Nech $x \in \mathbb{N}$, $y \in \mathbb{Z}$. Označme*

$$\begin{aligned} N_{(x,y)} &= card\{(S_0, S_1, \ldots, S_x) : S_0 = 0, S_x = y\} \\ N^+_{(x,y)} &= card\{(S_0, S_1, \ldots, S_x) : S_0 = 0, S_1 > 0, S_2 > 0, \ldots, S_x = y > 0\}. \end{aligned}$$

Potom platí

$$N^+_{(x,y)} = \frac{y}{x} N_{(x,y)}.$$

Dôkaz. Budeme dokazovať pre $N_{(x,y)} \neq 0$, pre $N_{(x,y)} = 0$ tvrdenie platí triviálne.
$N_{(x,y)}$ predstavuje počet trajektórií častice takých, že v čase x sa častica nachádzala v polohe y.
$N_{(x,y)}$ predstavuje počet trajektórií častice takých, že v čase x sa častica nachádzala v polohe y a

pohybovala sa nad osou x. Ak označíme p, respektíve q, počet krokov, ktorě častica pri svojom pohybe vykonala smerom doprava, respektíve doľava, potom $x = p + q$ a $y = p - q$ a

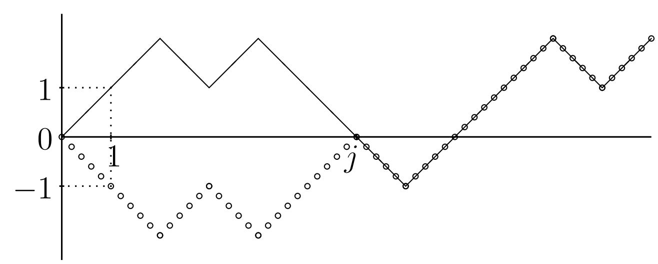
$$N_{x,y} = \binom{p+q}{p} = \binom{x}{\frac{x+y}{2}}. \tag{2.5}$$

$$\begin{aligned}
N^+_{x,y} &= card\{(S_1, S_2, \ldots, S_x) : S_1 = 1, S_2 > 0, \ldots, S_x = y > 0\} \\
&= card\{(S_1, S_2, \ldots, S_x) : S_1 = 1, S_x = y\} \\
&\quad - card\{(S_1, S_2, \ldots, S_x) : S_1 = 1, S_x = y > 0, \exists\, j, 1 < j < x : S_j = 0\} \\
&= N_{x-1,y-1} - card\{(S_1, S_2, \ldots, S_x) : S_1 = -1, S_x = y\} \\
&= N_{x-1,y-1} - N_{x-1,y+1}.
\end{aligned}$$

Pričom rovnosť

$$\begin{aligned}
card\ &\{(S_1, S_2, \ldots, S_x) : S_1 = 1, S_x = y > 0, \exists\, j, 1 < j < x : S_j = 0\} = \\
card\ &\{(S_1, S_2, \ldots, S_x) : S_1 = -1, S_x = y\}
\end{aligned}$$

plynie z toho, že ku každému prvku množiny na ľavej strane rovnosti existuje jednoznačne určený prvok množiny na pravej strane rovnosti, a to pomocou zobrazenia, ktoré je znázornené na obrázku 2.2 (trajektóriu $(S_1, \ldots, S_j)$ zobrazujeme osovo súmerne podľa osi, využívame vetu 2.1, a pre $(S_{j+1}, \ldots, S_x)$ ide o identické zobrazenie.).



Obrázok 2.2: Ukážka bijekcie z dôkazu vety 2.4 na jednej z trajektórií.

Takže podľa práve spočítaného a poznatku (2.5) platí

$$\begin{aligned}
N^+_{x,y} &= \binom{x-1}{\frac{x+y-2}{2}} - \binom{x-1}{\frac{x+y}{2}} = \binom{p+q-1}{p-1} - \binom{p+q-1}{q-1} \\
&= \binom{p+q}{p}\left(\frac{p}{p+q} - \frac{q}{p+q}\right) = N_{x,y}\frac{p-q}{p+q} = N_{x,y}\frac{y}{x}.
\end{aligned}$$

□

Veta 2.5. *Nech t^* je markovský čas prvého vstupu do nuly. Potom*

$$P[t^* = 2n] = \binom{2n-2}{n-1} n^{-1}\, 2^{-(2n-1)}. \tag{2.6}$$

Dôkaz. Jav $[t^* = 2n]$ si napíšeme ako disjunktné zjednotenie dvoch javov (častica sa pred prvým návratom do nuly pohybovala napravo alebo naľavo od nuly)

$$\begin{aligned}[t^* = 2n] &= [S_1 \neq 0, \ldots, S_{2n-1} \neq 0, S_{2n} = 0] \\ &= [S_1 = 1, S_2 > 0, \ldots, S_{2n-1} > 0, S_{2n} = 0] \\ &\quad \cup [S_1 = -1, S_2 < 0, \ldots, S_{2n-1} < 0, S_{2n} = 0].\end{aligned}$$

Preto

$$\begin{aligned}P[t^* = 2n] &= P[S_1 = 1, S_2 > 0, \ldots, S_{2n-1} = 1, S_{2n} = 0] \\ &\quad + P[S_1 = -1, S_2 < 0, \ldots, S_{2n-1} = -1, S_{2n} = 0]\end{aligned}$$

S využitím viet 2.1 a 2.4 a vzťahu (2.5):

$$\begin{aligned}P[t^* = 2n] &= 2P[S_1 = 1, S_2 > 0, \ldots, S_{2n-1} = 1, S_{2n} = 0] \\ &= 2\,\frac{1}{2}\,P[S_1 = 1, S_2 > 0, \ldots, S_{2n-1} = 1] \\ &= N^+_{2n-1,1}\, 2^{-(2n-1)} = \frac{1}{2n-1} N_{2n-1,1}\, 2^{-(2n-1)} \\ &= \frac{1}{2n-1}\binom{2n-1}{n} 2^{-(2n-1)} = \binom{2n-2}{n-1} n^{-1}\, 2^{-(2n-1)}.\end{aligned}$$

□

Veta 2.6. *Nech t^* je markovský čas prvého vstupu do nuly. Potom*

$$Et^* = +\infty.$$

Dôkaz.

$$Et^* = \sum_{n=1}^{\infty} 2n \binom{2n-2}{n-1} n^{-1}\, 2^{-(2n-1)} = 2\sum_{n=0}^{\infty} \binom{2n}{2} 2^{-2n}.$$

Označme

$$a_n = \binom{2n}{2} 2^{-2n} \geq 0.$$

Na vyšetrenie konvergencie tohoto radu použijeme limitné Raabeovo kritérium:

$$\begin{aligned}
\lim_{n\to\infty} n(\frac{a_n}{a_{n+1}} - 1) &= \lim_{n\to\infty} n\left(\frac{\binom{2n}{n}2^{-2n}}{\binom{2n+2}{n+1}2^{-(2n+2)}} - 1\right) \\
&= \lim_{n\to\infty} n\left(4\,\frac{(n+1)^2}{(2n+2)(2n+1)} - 1\right) \\
&= \lim_{n\to\infty} \frac{2n^2 - n}{4n^2+6n+2} = \frac{1}{2} < 1
\end{aligned}$$

a teda $\sum_{n=0}^{\infty} a_n$ diverguje a $Et^* = \infty$. □

Počítajme

$$\begin{aligned}
P[S_{2n-2}=0] - P[S_{2n}=0] &= \binom{2n-2}{n-1}2^{-(2n-2)} - \binom{2n}{n}2^{-2n} \\
&= \binom{2n-2}{n-1}2^{-(2n-1)}2 - \binom{2n-2}{n-1}2^{-(2n-1)}\frac{(2n-1)}{n} \\
&= \binom{2n-2}{n-1} n^{-1}\, 2^{-(2n-1)} = P[t^* = 2n].
\end{aligned}$$

Teda platí, že

$$P[t^* = 2n] = P[S_{2n-2}=0] - P[S_{2n}=0]. \tag{2.7}$$

S využitím vzťahu (2.7) dostávame

$$1 - \sum_{k=1}^{n} P[t^* = 2k] = 1 - \sum_{k=1}^{n}(P[S_{2k-2}=0] - P[S_{2k}=0]) = P[S_{2n}=0].$$

S ohľadom na (1.1) je

$$P[S_{2n}=0] = \frac{1}{\sqrt{\pi n}}(1+\delta_n) \xrightarrow{n\to\infty} 0. \tag{2.8}$$

Teda

$$\sum_{k=1}^{n} P[t^* = 2k] \xrightarrow{n\to\infty} 1 \quad \text{a} \quad P[t^* = +\infty] = 0.$$

Tento poznatok interpretujeme tak, že k návratu do nuly iste dôjde v konečnom markovskom čase, avšak (s ohľadom na Vetu 2.6) je stredná doba čakania na túto udalosť nekonečná.

Veta 2.7. *Nech t^* je markovský čas prvého výstupu S_n z intervalu (u, v), kde $-\infty < u < 0 < v < +\infty$ a u, v sú celé čísla. Potom platí:*

$$\begin{aligned}
&(1) \quad ES_{t^*} = 0 \\
&(2) \quad var\, S_{t^*} = Et^* \\
&(3) \quad P[S_{t^*} = u] = \frac{v}{v-u} \\
&(4) \quad P[S_{t^*} = v] = \frac{-u}{v-u} \\
&(5) \quad Et^* = -uv.
\end{aligned}$$

Dôkaz. (1) Podľa vety 1.5 a rozdelením integračného oboru dostávame

$$0 = ES_n = \int_\Omega S_n \, dP = \int_{[t^*>n]} S_n \, dP + \int_{[t^* \leq n]} S_n \, dP.$$

A teda

$$\begin{aligned}
-\int_{[t^*>n]} S_n \, dP &= \int_{[t^* \leq n]} S_n \, dP = \sum_{k=1}^{n} \int_{[t^*=k]} S_n \, dP \\
&= \sum_{k=1}^{n} \int_{[t^*=k]} (S_n - S_k) \, dP + \sum_{k=1}^{n} \int_{[t^*=k]} S_k \, dP \\
&= \sum_{k=1}^{n} \int_{[t^*=k]} (S_n - S_k) \, dP + \sum_{k=1}^{n} \int_{[t^*=k]} S_{t^*} \, dP \\
&= \sum_{k=1}^{n} E[I_{[t^*=k]}(S_n - S_k)] + \int_{[t^* \leq n]} S_{t^*} \, dP.
\end{aligned}$$

Náhodná veličina $S_n - S_k$ je funkciou $X_{k+1}, \ldots, X_n$ a náhodná veličina $I_{[t^*=k]}$ je funkciou $X_1, \ldots, X_k$, preto sú veličiny $S_n - S_k$ a $I_{[t^*=k]}$ nezávislé a stredná hodnota ich súčinu je rovná súčinu stredných hodnôt:

$$\sum_{k=1}^{n} E[I_{[t^*=k]}(S_n - S_k)] = \sum_{k=1}^{n} (P[t^* = k] E[S_n - S_k]) = 0,$$

pretože

$$E[S_n - S_k] = 0.$$

Takže platí:

$$-\int_{[t^*>n]} S_n \, dP = \int_{[t^* \leq n]} S_{t^*} \, dP. \tag{2.9}$$

Limitným prechodom $n \to \infty$ na pravej strane rovnosti (2.9) dostaneme ES_{t^*}. Pri odhade ľavej strany rovnosti (2.9) s využitím

$$[t^* > n] \;\Rightarrow\; [|S_n| \leq \max(-u, v)] = C \text{ kde } C \text{ je konštanta} \tag{2.10}$$

získame:

$$\left| \int_{[t^*>n]} S_n \, dP \right| \leq \int_{[t^*>n]} |S_n| \, dP \leq C \, P[t^* > n].$$

Platí

$$[t^* > n] \searrow [t^* = \infty] \,\text{pre}\, n \to \infty \tag{2.11}$$

a

$$P[t^* = \infty] = 0 \quad . \tag{2.12}$$

Rovnosť (2.12) plynie z toho, že

$$t^* \in L_r, \forall r \in \mathbb{N} \quad \text{(viz. [1], V.1.7)}, \tag{2.13}$$

teda špeciálne

$$t^* \in L_1, \text{tj. } Et^* < \infty,$$

z čoho vyplyva (2.12).
A teda

$$P[t^* > n] \to 0 \,\text{pre}\, n \to \infty. \tag{2.14}$$

A preto ľavá strana rovnosti konverguje k 0 pre $n \to \infty$ a dostávame $ES_{t^*} = 0$.

(2) Dokazujeme analogicky ako v (1), preto budeme postupovať rýchlejšie. Využívame vetu 1.5:

$$n = var\, S_n = \int_\Omega (S_n - ES_n)^2 \, dP = \int_\Omega S_n^2 \, dP = \int_{[t^*>n]} S_n^2 \, dP + \int_{[t^* \leq n]} S_n^2 \, dP$$

Teda

$$\begin{aligned}
n \quad &- \int_{[t^*>n]} S_n^2 \, dP = \\
= \quad &\int_{[t^* \leq n]} S_n^2 \, dP = \sum_{k=1}^n \int_{[t^*=k]} S_n^2 \, dP = \sum_{k=1}^n \int_{[t^*=k]} ((S_n - S_k) + S_k)^2 \, dP \\
= \quad &\sum_{k=1}^n \int_{[t^*=k]} (S_n - S_k)^2 \, dP + 2 \sum_{k=1}^n \int_{[t^*=k]} S_k (S_n - S_k)^2 \, dP + \sum_{k=1}^n \int_{[t^*=k]} S_k^2 \, dP \\
= \quad &\sum_{k=1}^n E[I_{[t^*=k]}(S_n - S_k)]^2 + 2 \sum_{k=1}^n E[I_{[t^*=k]} S_k (S_n - S_k)] + \int_{[t^* \leq n]} S_{t^*}^2 \, dP
\end{aligned}$$

Teraz opäť využijeme nezávislosť náhodných veličín $S_n - S_k$ a $I_{[t^*=k]}$, resp. $S_n - S_k$ a $I_{[t^*=k]}S_k$:

$$\begin{aligned}
\sum_{k=1}^{n} E[I_{[t^*=k]}(S_n - S_k)]^2 &= \sum_{k=1}^{n} E[I_{[t^*=k]}]^2 E[S_n - S_k]^2 = \sum_{k=1}^{n} E[I_{[t^*=k]}]\, E[S_n - S_k]^2 \\
&= \sum_{k=1}^{n} P[t^* = k]\, E[S_n - S_k]^2 = \sum_{k=1}^{n} P[t^* = k](n-k) \\
\sum_{k=1}^{n} E[I_{[t^*=k]}S_k(S_n - S_k)] &= \sum_{k=1}^{n} E[I_{[t^*=k]}S_k]\, E[S_n - S_k] = \sum_{k=1}^{n} ES_{t^*}\, E[S_n - S_k] = 0.
\end{aligned}$$

A tým sme dospeli k rovnosti:

$$n - \int_{[t^*>n]} S_n^2\, dP = \sum_{k=1}^{n} P[t^* = k](n-k) + \sum_{k=1}^{n} \int_{[t^*=k]} S_n^2\, dP. \tag{2.15}$$

Postupnými úpravami odvodíme:

$$\begin{aligned}
n &- \int_{[t^*>n]} S_n^2\, dP = n\sum_{k=1}^{n} P[t^* = k] - \sum_{k=1}^{n} kP[t^* = k] + \sum_{k=1}^{n} \int_{[t^*=k]} S_n^2\, dP \\
n &- \int_{[t^*>n]} S_n^2\, dP = nP[t^* \le n] - \sum_{k=1}^{n} kP[t^* = k] + \sum_{k=1}^{n} \int_{[t^*=k]} S_n^2\, dP \\
n &\;(1 - P[t^* \le n]) - \int_{[t^*>n]} S_n^2\, dP = -\sum_{k=1}^{n} kP[t^* = k] + \sum_{k=1}^{n} \int_{[t^*=k]} S_n^2\, dP \\
n &\;P[t^* > n] - \int_{[t^*>n]} S_n^2\, dP = -\sum_{k=1}^{n} kP[t^* = k] + \sum_{k=1}^{n} \int_{[t^*=k]} S_n^2\, dP. \qquad (2.16)
\end{aligned}$$

Limitným prechodom $n \to \infty$ na pravej strane rovnosti (2.16) dostaneme $-Et^* + var\, S_n$. Teraz už len potrebujeme dokázať, že ľavá strana rovnosti (2.16) konverguje k 0 pre $n \to \infty$. S využitím (2.10) a (2.14) odhadneme

$$\left| \int_{[t^*>n]} S_n^2\, dP \right| \le \int_{[t^*>n]} |S_n|^2\, dP \le C^2 P[t^* > n] \to 0 \text{ pre } n \to \infty.$$

Taktiež podľa (2.13) a s využitím implikácie

$$Et^* < \infty \Rightarrow nP[t^* > n] \to 0$$

platí

$$nP[t^* > n] \to 0 \quad \text{pre} \quad n \to \infty.$$

A ľavá strana rovnosti (2.16) naozaj konverguje k 0 a $var\, S_t^* = Et^*$.

(3) Označme $P = P[S_{t^*} = u]$, potom $P[S_{t^*} = v] = 1 - P$. S využitím (1) je:

$$0 = ES_{t^*} = Pu + (1-P)v.$$

Odťiaľ:

$$\begin{aligned} P &= \frac{v}{v-u} \\ 1-P &= \frac{-u}{v-u}. \end{aligned}$$

Tým sme dokázali (3) aj (4).

(5) Podľa (2) platí:

$$Et^* = var\, S_{t^*} = ES_{t^*}^2 = u^2 P[S_{t^*} = u] + v^2 P[S_{t^*} = v]$$

Dosadením za $P[S_{t^*} = u]$ a $P[S_{t^*} = v]$ dostávame:

$$Et^* = u^2 \frac{v}{v-u} + v^2 \frac{-u}{v-u} = \frac{1}{v-u}(u^2 v - v^2 u) = \frac{uv(u-v)}{v-u} = -uv.$$

$\square$

Kapitola 3

Zákon arkusínu

V tejto kapitole sa budeme zaoberať hlbším tvrdením, ktoré sa nazýva zákon arkusínu. Najprv si však vyslovíme a dokážeme pomocné tvrdenia, pomocou ktorých zákon arkusínu dokážeme.

Veta 3.1 (Markovská vlastnosť). *Nech $\{S_n\}$ je náhodná prechádzka a t^* je jej skoro iste konečný markovský čas. Potom náhodná prechádzka $\{S_n\}^*$, kde ${S_n}^* = S_{t^*+n} - S_{t^*}$, je náhodná prechádzka s rovnakým rozdelením ako náhodná prechádzka $\{S_n\}$ a je nezávislá s každým náhodným javom A, pre ktorý platí:*

$$A \in \sigma(S_1, S_2, \ldots), \qquad A \cap [t^* = n] \in \sigma(S_1, S_2, \ldots, S_n), \qquad n \in \mathbb{N}.$$

Dôkaz. Viz. [1], Veta 1.1. □

Práve vyslovená veta hovorí, že chovanie náhodnej prechádzku môžeme začať študovať až po uplynutí prvých t^* časových okamihov a dostaneme náhodnú prechádzku s rovnakými vlastnosťami, ako mala pôvodná náhodná prechádzka, a jej chovanie nezávisí na chovaní pôvodnej náhodnej prechádzky do času t^*.

Na postupnosti $\{S_n\}$ definujme náhodné veličiny

$$T_n = card\,\{1 \leq j \leq n : S_j > 0\}, \tag{3.1}$$
$$U_n = card\,\{1 \leq k \leq n : S_{k-1} \geq 0, S_k \geq 0\}. \tag{3.2}$$

Lemma 3.2. *Platí:*

$$\frac{|T_n - U_n|}{n} \xrightarrow{P} 0 \quad pre \quad n \to \infty.$$

Dôkaz. Platí, že:

$$\{1 \leq k \leq n : S_k \geq 0, S_{k-1} \geq 0\} = \{1 \leq k \leq n : S_k > 0\} \cup \{1 \leq k \leq n : S_k = 0, S_{k-1} > 0\}.$$

Preto

$$U_n \leq T_n + card\,\{1 \leq k \leq n : S_k = 0, S_{k-1} > 0\}.$$

Ak si označíme

$$K_n = card\,\{1 \leq k \leq n : S_k = 0\},$$

potom

$$U_n \leq T_n + K_n.$$

A keďže

$$\{1 \leq k \leq n : S_k > 0\} \subset \{1 \leq k \leq n : S_k \geq 0, S_{k-1} \geq 0\},$$

platí

$$0 \leq U_n - T_n \leq K_n.$$

K dokončeniu dôkazu stačí ukázať

$$\frac{K_n}{n} \xrightarrow{\text{P}} 0.$$

Ďalej počítame

$$\begin{aligned}
\lim_{n\to\infty} E(n^{-1}K_n) &= \lim_{n\to\infty} n^{-1}EK_n = \lim_{n\to\infty} n^{-1}\sum_{k=1}^{n} P[S_k = 0] \\
&= (\lim_{n\to\infty} n^{-1})(\sum_{k=1}^{n}\lim_{n\to\infty} P[S_k = 0]) = 0,
\end{aligned}$$

pretože $P[S_k = 0] = 0$ pre k nepárne a $\lim_{n\to\infty} P[S_k = 0] = 0$ pre k párne podľa (2.8).
Takže platí

$$\frac{K_n}{n} \xrightarrow{L^1} 0 \quad \text{z čoho plynie} \quad \frac{K_n}{n} \xrightarrow{P} 0.$$

□

Lemma 3.3. *Pre* $0 \leq k \leq n$, $n \in \mathbb{N}$ *platí:*

$$P[U_{2n} = 2k] = 2^{-2n}\binom{2k}{k}\binom{2(n-k)}{n-k}.$$

Dôkaz. Budeme postupovať indukciou podľa n.
Pre $n = 1$ je:

$$P[U_2 = 0] = P[U_2 = 2] = \frac{1}{2},$$

takže pre $n = 1$ tvrdenie platí.
Nech vzorec platí pre $\forall k : 0 \leq k \leq n$, kde $n = 1, 2, \ldots, m-1; m \in \mathbb{N}$. Chceme ukázať, že platí pre $\forall k : 0 \leq k \leq n$, kde $n = 1, 2, \ldots, m$. Pre $k = 0$ a $k = m$ plynie vzorec triviálne z

$$P[U_{2m} = 0] = P[U_{2m} = m] = \binom{2m}{m}2^{-2m}.$$

Nech teda k, $1 \leq k \leq m-1$, je pevne zvolené a t^* je markovský čas prvého návratu do nuly, náhodný jav $[U_{2m} = 2k]$ si môžeme vyjadriť ako zjednotenie $m+1$ disjunktných javov, potom

$$P[U_{2m} = 2k] = \sum_{j=1}^{m} P[U_{2m} = 2k, t^* = 2j] + P[U_{2m} = 2k, t^* > 2m].$$

Z $[t^* > 2m]$ plynie buď $[U_{2m} = 2m]$ alebo $[U_{2m} = 0]$, avšak my uvažujeme $1 \leq k \leq m-1$, a preto

$$P[U_{2m} = 2k, t^* > 2m] = 0.$$

Teraz sa budeme na chvíľu zaoberať chovaním náhodnej prechádzky po prvom návtrate do nuly, nebude nás zaujímať jej chovanie do času t^*, túto posunutú nahodnú prechádzku budeme označíme S_n^*. Takže:

$$S_n^* = S_{t^*+n} - S_{t^*}$$

Označme:

$$U_n^* = card\,\{1 \leq j \leq n : S_{j-1}^* \geq 0, S_j^* \geq 0\}.$$

Potom pre $1 \leq j \leq m$ platí

$$\begin{aligned} P[U_{2m} = 2k, t^* = 2j, S_1 = 1] &= P[U_{2(m-j)}^* = 2(k-j), t^* = 2j, S_1 = 1] \\ &= P[U_{2(m-j)} = 2(k-j)]P[t^* = 2j, S_1 = 1], \end{aligned}$$

kde prvá rovnosť plynie z implikácie ($[t^* = 2j, S_1 = 1] \Rightarrow [S_r \geq 0, 1 \leq r \leq 2j]$) a druhá je dôsledok markovskej vlastnosti náhodnej prechádzky. (Pre $j > k$ sú obe strany rovnosti nulové, preto môžeme posunút hornú medz v sume z m na k.)
Analogicky pre $1 \leq j \leq m$ platí:

$$P[U_{2m} = 2k, t^* = 2j, S_1 = -1] = P[U_{2(m-j)} = 2k]P[t^* = 2j, S_1 = -1].$$

(Pre $j > k$ sú opäť obe strany rovnosti nulové.)
S využitím toho, že

$$P[t^* = 2j, S_1 = 1] = P[t^* = 2j, S_1 = -1] = \frac{1}{2}P[t^* = 2j]$$

platí:

$$\begin{aligned}
P[U_{2m}=2k] &= \sum_{j=1}^{m} P[U_{2m}=2k, t^*=2j] \\
&= \sum_{j=1}^{m} P[U_{2m}=2k, t^*=2j, S_1=1] + \sum_{j=1}^{m} P[U_{2m}=2k, t^*=2j, S_1=-1] \\
&= \sum_{j=1}^{k} P[U_{2(m-j)}=2(k-j)] P[t^*=2j, S_1=1] \\
&\quad + \sum_{j=1}^{k} P[U_{2(m-j)}=2k] P[t^*=2j, S_1=-1] \\
&= \sum_{j=1}^{k} P[U_{2(m-j)}=2(k-j)] \, \frac{1}{2} P[t^*=2j] \\
&\quad + \sum_{j=1}^{k} P[U_{2(m-j)}=2k)] \, \frac{1}{2} P[t^*=2j].
\end{aligned}$$

Pomocou vzťahu (2.7) je možné vyjadriť $\frac{1}{2}P[t^*=2j]$ v tvare:

$$\begin{aligned}
\frac{1}{2}P[t^*=2j] &= \frac{1}{2}\left(\binom{2j-2}{j-1}2^{-(2j-2)} - \binom{2j}{j}2^{-2j}\right) \\
&= \binom{2j-2}{j-1}2^{-(2j-1)} - \binom{2j}{j}2^{-(2j+1)}
\end{aligned}$$

S využitím indukčného predpokladu a dosadením za $\frac{1}{2}P[t^*=2j]$ podľa predchádzajúceho výpočtu dostávame:

$$\begin{aligned}
P[U_{2m}=2k] &= \sum_{j=1}^{k} 2^{-2(m-j)} \binom{2(k-j)}{k-j} \binom{2(m-k)}{m-k} \frac{1}{2} P[t^*=2j] \\
&\quad + \sum_{j=1}^{k} 2^{-2(m-j)} \binom{2k}{k} \binom{2(m-j-k)}{m-j-k} \frac{1}{2} P[t^*=2j]
\end{aligned}$$

A vzhľadom k vlastnostiam kombinačných čísel a platnosti vzťahu (viz. [1], Veta 4.23)

$$\sum_{j=0}^{k} \binom{2(k-j)}{k-j} \binom{2j}{j} = 2^{2k},$$

dostávame

$$P[U_{2m}=2k] = 2^{-2m} \binom{2k}{k} \binom{2(m-k)}{m-k}.$$

□

Veta 3.4 (Zákon arkusínu). *Pre rozdelenie náhodnej veličiny $\frac{T_n}{n}$ platí:*

$$P\left[\frac{T_n}{n} < x\right] \xrightarrow{n\to\infty} \Psi(x),$$

kde $x \in \mathbb{R}$ a $\Psi(x)$ je distibučná funkcia rozdelenia arkusinu:

$$\begin{array}{lll} \Psi(x) = & 0 & \textit{pre} \quad x \leq 0, \\ \Psi(x) = & \frac{2}{\pi}\arcsin\sqrt{x} & \textit{pre} \quad x \in (0,1), \\ \Psi(x) = & 1 & \textit{pre} \quad x \geq 1. \end{array}$$

Dôkaz. Lemma 3.2 hovorí, že náhodné veličiny U_n a T_n sú si asymptoticky blízke, a preto môžeme v dôkaze tejto vety nahradiť náhodnú veličinu T_n náhodnou veličinou U_n. Rozdelenie U_n sme síce odvodili len pre n párne (viz. lemma 3.3), ale

$$\left|\frac{U_{2n}}{2n} - \frac{U_{2n+1}}{2n+1}\right| \xrightarrow{\text{P}} 0 \quad \text{pre} \quad n \to \infty.$$

(Plynie to zo vzťahu $U_{2n} \leq U_{2n+1} \leq U_{2n} + 1$.)
Nech $0 < a < b < 1$, potom s využitím lemmy 3.3 máme:

$$P\left[a \leq \frac{U_{2n}}{2n} < b\right] = \sum_{k=[na]+1}^{[nb]} 2^{-2n}\binom{2k}{k}\binom{2(n-k)}{n-k}.$$

Pre $b \leq 0$ je $P\left[a \leq \frac{U_{2n}}{2n} < b\right] = 0$, pretože $U_{2n} \geq 0$ a $n > 0$.
Pre $b \geq 1$ je $P\left[a \leq \frac{U_{2n}}{2n} < b\right] = 1$, keďže $U_{2n} \leq n$, $\forall n \in \mathbb{N}$.

Použitím Stirlingovho vzorca v tvare

$$n! = n^n e^{-n}\sqrt{2\pi n}\, exp\left\{\frac{\delta_n}{4n}\right\}, \quad \text{kde} \quad 0 < \delta_n < 1,$$

dostávame

$$P\left[a \leq \frac{U_{2n}}{2n} < b\right] = \frac{1}{\pi}\sum_{k=[na]+1}^{[nb]} \frac{1}{\sqrt{k(n-k)}} + d_n,$$

kde

$$\lim_{n\to\infty} d_n = 0.$$

(podľa [1], V.4.24) Teda:

$$\lim_{n\to\infty} P\left[a \le \frac{U_{2n}}{2n} < b\right] = \lim_{n\to\infty} \frac{1}{\pi}\frac{1}{n} \sum_{k=[na]+1}^{[nb]} \frac{1}{\sqrt{\frac{k}{n}(1-\frac{k}{n})}}$$
$$= \frac{1}{\pi}\int_a^b \frac{1}{\sqrt{x(1-x)}}\,dx = \frac{2}{\pi}(\arcsin\sqrt{b} - \arcsin\sqrt{a}).$$

□

Funkcia $\Psi(x)$ má minimum v bode $x = \frac{1}{2}$. Vzhľadom k symetrii by sme očakávali, že častica sa bude práve v polovici časových okamihov nachádzať k kladnej časti celočíselnej priamky a práve v polovici časových okamihov v zápornej časti, avšak zákon arkusínu ukazuje opak. Hovorí, že častica vykazuje tendenciu zotrvávať v tej časti celočíselnej priamky, do ktorej vstúpila.

Literatúra

[1] Štěpán, J.: *Teorie pravděpodobnosti: Matematické základy*, Academia, Praha, 1987.

[2] Zvára, K., Štěpán, J.: *Pravděpodobnost a matematická statistika*, Matfyzpress, Praha, 2001.