3. Náhodné procházky

- druhý klasický příklad náhodného procesu
- obecná definice: Buď $X_1, X_2, \dots$ posloupnost nezávislých náhodných veličin s hodnotami v $\mathbb{R}^d$, $d \geq 1$.
  Pak $S_n = X_1 + \dots + X_n$, $n \geq 1$ a $S_0 = 0$ je d-dimenzionální náhodná procházka s kroky $X_1, X_2, \dots$

Definice: Jednoduchá náhodná procházka na $\mathbb{Z}$ je náhodný proces s $S_0 = 0$, $S_n = X_1 + \dots + X_n$, kde $X_1 \dots$ jsou nezávislé stejně rozdělené s $\mathbb{P}(X_i = 1) = p$ & $\mathbb{P}(X_i = -1) = q = 1 - p$.

Pozn. Symetrická náh. procházka: $p = q = 1/2$.

Některé parametry lze spočítat přímo z definice:

- $\mathbb{E} X_n = -1 \cdot q + 1 \cdot p = p - q$
- $\mathrm{Var}[X_n] = \mathbb{E}[X_n - \mathbb{E}X_n]^2 = \mathbb{E}[X_n]^2 - 2(\mathbb{E}X_n)^2 + (\mathbb{E}X_n)^2 = \mathbb{E}[X_n]^2 - (\mathbb{E}X_n)^2$

$$= 1 - (p-q)^2 = 1 - p^2 - q^2 + 2pq = 4pq.$$

$\boxed{X_n^2 = 1}$

$$p + q = 1$$
$$(p+q)^2 = 1$$
$$p^2 + q^2 + 2pq = 1$$

- $\mathbb{E}[S_n] = n(p - q)$
- $\mathrm{Var}[S_n] = \mathbb{E}\left[(X_1 + \dots + X_n) - \mathbb{E}(X_1 + \dots + X_n)\right]^2 = \mathbb{E}\left[\sum_{j=1}^{n}(X_j - \mathbb{E}X_j)\right]^2 =$

$$= \mathbb{E}\sum_{j=1}^{n}(X_j-\mathbb{E}X_j)^2 + \mathbb{E}\sum_{j\neq k}{}'(X_j-\mathbb{E}X_j)(X_k-\mathbb{E}X_k)$$

$$= n\operatorname{Var}(X_j) + 0 = 4npq.$$

<u>Rozložení S_n</u>:

- Po sudém počtu kroků je S_{2k} sudé číslo, po lichém počtu kroků je S_{2k+1} liché číslo, tedy
  $$\mathbb{P}(S_{2n}=2k+1)=0=\mathbb{P}(S_{2n+1}=2k),\quad k\in\mathbb{Z},\ n\in\mathbb{N}$$
- Po ~~k~~ n krocích je $|S_n|\leq k$ protože $S_0=0$ & $|X_n|=1$, tedy $\mathbb{P}(S_n=k)=0,\ |k|>n$
- Uvažujme sudý ($=2n$) počet kroků a pravděpodobnost $S_{2n}=2k$, kde $-n\leq k\leq n$. To nastane, pokud provedeme $n+k$ kroků $(+1)$ a $n-k$ kroků (-1).

  Toto lze provést následujícím způsobem:
  - Nejprve vybereme $\binom{2n}{n+k}$ kroků z $X_1,\dots,X_{2n}$, které budou rovny $+1$
  - Pravděpodobnost, že opravdu tyto kroky jsou $=+1$ a zbývající $=-1$:
    $$p^{n+k}q^{n-k}$$

Celkem je tedy $\mathbb{P}(S_{2n}=2k)=\binom{2n}{n+k}p^{n+k}q^{n-k};\ -n\leq k\leq n.$

Cvičení: Dokažte, že $\mathbb{P}(S_{2n+1}=2k+1)=\binom{2n+1}{n+k+1}p^{n+k+1}q^{n-k},\ -n\leq k\leq n.$

I přesto, že S_n je možno přesně popsat (= udat jeho rozložení), je tento popis technický a nehodí se k výpočtu zajímavých parametrů náhodné procházky.

První návrat do počátku

Označme $T_0^R = \inf\{n \geq 1 : S_n = 0\}$ … první čas návratu do počátku (pozn. $\inf \emptyset = +\infty$ … pokud tedy $S_n \neq 0$ pro všechna $n \geq 1$, pak $T_0^R = +\infty$)

Chceme spočítat $g(n) = \mathbb{P}(T_0^R = n)$, $n \geq 1$

- $n = 1$ … $\mathbb{P}(T_0^R = 1) = 0$, $n = 2$: $\mathbb{P}(T_0^R = 2) = 2pq$
- $\mathbb{P}(T_0^R = 2k+1) = 0$; $\mathbb{P}(T_0^R = 4) = 2p^2q^2$ … 2x nahoru, 2x dolů … popoř.
- Je zřejmé, že podobně pro $n = 2k$, $k \geq 3$ by bylo velice technické.

Položme $h(n) = \mathbb{P}(S_n = 0)$ … pak $h(2n) = \binom{2n}{n} p^n q^n$, $h(2n+1) = 0$, $n \in \mathbb{N}$.

Pak platí:

Lemma: Funkce $g(n) = \mathbb{P}(T_0^R = n)$, $n \geq 1$ splňuje konvoluční identitu

$$h(n) = \sum_{k=0}^{n-2} g(n-k)\, h(k), \quad n \geq 1$$

a počáteční podmínku $g(1) = 0$.

Důkaz: Událost $\{S_n = 0\}$ rozložme (disjunktně) jako

$$\{S_n = 0\} = \bigcup_{k=0}^{n-2} \{S_k = 0, S_{k+1} \neq 0, \dots, S_{n-1} \neq 0, S_n = 0\}$$

podle posledního návratu do nuly před n.

Pak platí $h(n)=\mathbb{P}(S_n=0)=\sum_{k=0}^{n-2}\mathbb{P}(S_k=0, S_{k+1}\neq 0, \dots, S_{n-1}\neq 0, S_n=0)$

$$=\sum_{k=0}^{n-2}\mathbb{P}(S_{k+1}\neq 0, \dots, S_{n-1}\neq 0, S_n=0 \mid S_k=0)\cdot\mathbb{P}(S_k=0)$$

$$S'_u = S_{m+k}$$

$$=\sum_{k=0}^{n-2}\mathbb{P}(S'_1\neq 0, \dots, S'_{n-k-1}\neq 0, S'_{n-k}=0 \mid S'_0=0)\,\mathbb{P}(S_k=0)$$

$$=\sum_{k=0}^{n-2}\mathbb{P}(T_0^R=n-k)\,\mathbb{P}(S_k=0)=\sum_{k=0}^{n-2}h(k)g(n-k),\ n\geq 1.$$

Konvoluční identitu s okrajovou podmínkou vyřešíme pomocí vytvořující funkce: $G(s):=\sum_{n=0}^{\infty}s^n g(n),\ -1\leq s\leq 1.$

Stejně definujeme $H(s)=\sum_{k=0}^{\infty}s^k h(k)=\sum_{k=0}^{\infty}s^k\,\mathbb{P}(S_k=0),\ -1\leq s\leq 1.$

Lemma: Funkce $H(s)$ splňuje $H(s)=(1-4pqs^2)^{-1/2}$; $4pqs^2<1$

a splňuje $G(s)H(s)=H(s)-1.$

<u>Důkaz:</u>

$$H(s) = \sum_{k=0}^{\infty}{}' s^k \mathbb{P}(S_k = 0) \overset{S_{2k+1}\neq 0}{=} \sum_{k=0}^{\infty}{}' s^{2k} \mathbb{P}(S_{2k} = 0)$$

$$= \sum_{k=0}^{\infty}{}' s^{2k} \binom{2k}{k} p^k q^k = \sum_{k=0}^{\infty}{}' (pqs^2)^k \frac{(2k)(2k-1)\dots\cdot 4\cdot 3\cdot 2\cdot 1}{[k!]^2}$$

$$= \sum_{k=0}^{\infty}{}' (pqs^2)^k \cdot 4^k \frac{k(k-1/2)(k-1)(k-3/2)\dots 2\cdot 3/2\cdot 1\cdot 1/2}{[k!]^2}$$

$$= \sum_{k=0}^{\infty}{}' (4pqs^2)^k \cdot \frac{(k-1/2)(k-3/2)\dots 3/2\cdot 1/2}{k!}$$

$$= \sum_{k=0}^{\infty}{}' (-4pqs^2)^k \cdot \frac{(-1/2)(-3/2)\dots(3/2-k)\overbrace{(1/2-k)}^{=(-1/2)-(k-1)}}{k!} = \frac{1}{\sqrt{1-4pqs^2}}.$$

$$\dots |x|<1:\ (1+x)^\alpha = \sum_{n=0}^{\infty}{}' \underbrace{\binom{\alpha}{n}}_{=\frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}} x^n$$

Dále užijeme konvoluční identitu

$$G(s)H(s) = \left(\sum_{n=1}^{\infty}{}' s^n g(n)\right)\left(\sum_{k=0}^{\infty}{}' s^k h(k)\right) = \sum_{n=1}^{\infty}{}' \sum_{k=0}^{\infty}{}' s^{n+k} g(n) h(k)$$

$\boxed{g(1)=0}$
$\boxed{g(0):=0}$

$$= \sum_{k=0}^{\infty}{}' \sum_{n=2}^{\infty}{}' s^{n+k} g(n) h(k) = \sum_{k=0}^{\infty}{}' \sum_{l=k+2}^{\infty}{}' s^l g(l-k) h(k)$$

k pevné, $l = n+k$

$$= \sum_{l=2}^{\infty}{}' s^l \underbrace{\sum_{k=0}^{l-2} g(l-k) h(k)}_{=h(l)} = \sum_{l=2}^{\infty}{}' s^l h(l) = \sum_{l=1}^{\infty}{}' s^l h(l) = -1 + H(s)$$

Platí tedy $G(s) = 1 - \frac{1}{H(s)} = 1 - \sqrt{1-(4pqs^2)}$, $4pqs^2 < 1$

Pravděpodobnost prvního návratu je potom dána:

$$G(s) = 1-(1-4pqs^2)^{1/2} = 1 - \sum_{k=0}^{\infty} \frac{1}{k!}(-4pqs^2)^k \left(\tfrac{1}{2}-0\right)\left(\tfrac{1}{2}-1\right)\dots\left(\tfrac{1}{2}-(k-1)\right)$$

$$= -\sum_{k=1}^{\infty} \frac{1}{k!}(-4pqs^2)^k \left(\tfrac{1}{2}-0\right)\dots\left(\tfrac{1}{2}-(k-1)\right)$$

$$= -\sum_{k=1}^{\infty} \frac{(4pqs^2)^k}{k!}\left(0-\tfrac{1}{2}\right)\left(1-\tfrac{1}{2}\right)\dots\left((k-1)-\tfrac{1}{2}\right)$$

$$= \frac{1}{2}\sum_{k=1}^{\infty} s^{2k}\frac{(4pq)^k}{k!}\left(1-\tfrac{1}{2}\right)\dots\left((k-1)-\tfrac{1}{2}\right)$$ ... srovnáním koeficientů s $G(s) = \sum_{n=0}^{\infty} s^n g(n)$ plyne

$$g(2k+1) = 0,\ k \in N$$

$$g(2k) = \frac{1}{2}\cdot\frac{(4pq)^k}{k!}\left(1-\tfrac{1}{2}\right)\dots\left((k-1)-\tfrac{1}{2}\right) = \frac{1}{2k-1}\binom{2k}{k}p^k q^k = \frac{h(k)}{2k-1}.$$

Pravděpodobnost návratu do nuly v konečném čase

$$\mathbb{P}(T_0^n < +\infty) = \sum_{k=0}^{\infty}\mathbb{P}(T_0^n = k) = \sum_{k=0}^{\infty} g(k) = G(1) = 1-(1-4pq)^{1/2}$$

$$= 1-|p-q| = \begin{cases} 2q, & p \geq 1/2 \\ 2p, & p \leq 1/2 \end{cases} = 2\min(p,q).$$

.... pro formuli pro $G(s)$ potřebujeme $4pqs^2 < 1$ ... odvození výše platí tedy pro $4pq < 1$ .. $0 < 1-4pq = |p-q|^2$ ... $p \neq q$.

Pro $p=q$ plyne z Abelovy věty pro mocninné řady.

Pro $p=q$ je tedy $\mathbb{P}(T_0^R < +\infty) = 1$ a (tedy) $\mathbb{P}(T_0^R = +\infty) = 0$

Pro $p \neq q$ jsou obě tyto pravděpodobnosti $\in (0;1)$.

Střední doba návratu do počátku:

- Pro $p \neq q$ je $\mathbb{E}[T_0^R] = \sum_{k=1}^{\infty} k \cdot \mathbb{P}(T_0^R = k) + \infty \cdot \overbrace{\mathbb{P}(T_0^R = +\infty)}^{>0} = +\infty.$

  Dále je $G'(1) = \sum_{n=0}^{\infty} n s^n g(n) \Big|_{s=1} = \sum_{n=0}^{\infty} n g(n) = \sum_{n=0}^{\infty} n \mathbb{P}(T_0^R = n)$

  $$= \cancel{\mathbb{E}[T_0^R | T_0^R < +\infty]} = \frac{4pqs}{\sqrt{1-4pqs^2}}\Big|_{s=1} = \frac{4pq}{\sqrt{1-4pq}} = \frac{4pq}{|p-q|}.$$

  A tedy $\mathbb{E}[T_0^R | T_0^R < +\infty] = \dfrac{G'(1)}{\mathbb{P}(T_0^R < +\infty)} = \dfrac{1}{2\min(p,q)} \cdot \dfrac{4pq}{|p-q|} = \dfrac{2\max(p,q)}{|p-q|}.$

- Pro $p = q = 1/2$: $\mathbb{P}(T_0^R = +\infty) = 0$ & $\mathbb{E}[T_0^R] = \sum_{n=0}^{\infty} n g(n) =$

  $$= \sum_{n=1}^{\infty} (2n) g(2n) = \sum_{n=1}^{\infty} \frac{2n}{2n-1} \binom{2n}{n} \left(\frac{1}{2}\right)^n \left(\frac{1}{2}\right)^n \geq \sum_{n=1}^{\infty} \binom{2n}{n} \cdot \frac{1}{4^n}$$

Tato řada diverguje: Stirling: $n! \sim \sqrt{2\pi n} \left(n/e\right)^n$

$$\binom{2n}{n} \frac{1}{4^n} = \frac{(2n)!}{[n!]^2} \cdot \frac{1}{4^n} \sim \frac{\sqrt{4\pi n}(2n)^{2n} e^{-2n}}{[\sqrt{2\pi n}\, n^n e^{-n}]^2 \cdot 4^n} \sim \frac{1}{\sqrt{n}}.$$

$\Rightarrow \mathbb{E}[T_0^R] = +\infty.$