Generující funkce a náhodná procházka

Uvažujme opět náhodnou procházku

$$S_n = \sum_{i=1}^{n} X_i,$$

kde $P(X_i = 1) = p$ a $P(X_i = -1) = q = 1 - p$. Přitom X_i jsou nezávislé a $S_0 = 0$.

Problém rekurence:

– Jak často se náhodná procházka vrací do počátku?

– Jaké je pravděpodobnostní rozdělení prvního návratu do počátku?

K zodpovězení těchto otázek využijeme generující funkce.

Označme

$$p_0(n) = P(S_n = 0)$$

pravděpodobnost, že náhodná procházka je v 0 v čase n, a

$$f_0(n) = P(S_1 \neq 0, S_2 \neq 0, \ldots, S_{n-1} \neq 0, S_n = 0)$$

pravděpodobnost, že první návrat do počátku nastal v čase n.

Uvažujme generující funkce těchto dvou posloupností,

$$P_0(s) = \sum_{n=0}^{\infty} p_0(n) s^n$$

a

$$F_0(s) = \sum_{n=0}^{\infty} f_0(n) s^n.$$

Máme $p_0(0) = 1$ (neboť $S_0 = 0$) a $f_0(0) = 0$.

Lemma 4.1. *Platí*

$$P_0(s) = \left(1 - 4pqs^2\right)^{-\frac{1}{2}}.$$

Důkaz: Víme, že $S_n = 0 \Leftrightarrow$ počet kroků doprava $=$ počet kroků doleva, tedy $r = \frac{n}{2} = l$. Počet takových cest je $\binom{n}{\frac{n}{2}}$ pro n sudé a 0 pro n liché. Označme $k = \frac{n}{2}$, tj. $n = 2k$. Máme

$$p_0(2k) = \binom{2k}{k} p^k q^k$$

a

$$P_0(s) = \sum_{k=0}^{\infty} \binom{2k}{k} (pqs^2)^k. \tag{1}$$

Tvrdíme, že $P_0(s) = \frac{1}{\sqrt{1-4pqs^2}}$.

Využijeme obecného binomického rozvoje

$$(1+x)^a = \sum_{n=0}^{\infty} \binom{a}{n} x^n,$$

kde

$$\binom{a}{n} = \frac{a\,(a-1)\ldots(a-n+1)}{n!}.$$

Pro $a \in \mathbb{N}$ je rozvoj konečný, pro $a \in \mathbb{R} \wedge a \notin \mathbb{N}$ je rozvoj nekonečný. Dosazením $a = -\frac{1}{2}$ a $x = -4pqs^2$ dostaneme

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

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

Porovnáním 1 a 2 tedy stačí dokázat, že

$$\binom{-\frac{1}{2}}{k}(-4)^k = \binom{2k}{k}.$$

Pro levou stranu dostaneme

$$\frac{-\frac{1}{2}\left(-\frac{3}{2}\right)\left(-\frac{5}{2}\right)\ldots\left(-\frac{2k-1}{2}\right)}{k!}(-1)^k 2^{2k} =$$

$$2^k(-1)^{2k}\frac{1\cdot 3\cdot 5\cdot \ldots(2k-1)}{k!} = 2^k\frac{1\cdot 3\cdot 5\cdot \ldots\cdot(2k-1)}{k!}.$$

Pro pravou stranu

$$\binom{2k}{k} = \frac{2k\,(2k-1)\,(2k-2)\ldots 1}{k!k!} = 2^k k!\,\frac{(2k-1)\,(2k-3)\ldots 3\cdot 1}{k!k!} =$$

$$2^k\,\frac{(2k-1)\,(2k-3)\ldots 3\cdot 1}{k!}.$$

Tedy obě strany se rovnají. Tím je lemma dokázáno.

Věta 4.2. *Platí*

$$1)\ P_0(s) = 1 + P_0(s)F_0(s)$$

a

$$2)\ F_0(s) = 1 - \left(1 - 4pqs^2\right)^{\frac{1}{2}}\ .$$

Důkaz: Označme A jev, že $S_n = 0$ a nechť B_k jsou jevy "první návrat do počátku nastal v k-tém kroku" ($k = 1, 2, ..., n$).

1) B_k jsou disjunktní a dávají rozklad, tedy

$$P(A) = \sum_{k=1}^{n} P(A \mid B_k) P(B_k).$$

Máme $P(B_k) = f_0(k)$ a $P(A \mid B_k) = p_0(n-k)$, z časové homogenity. Tedy

$$p_0(n) = \sum_{k=1}^{n} p_0(n-k) f_0(k).$$

Vynásobíme s^n a sečteme,

$$\sum_{n=1}^{\infty} p_0(n)s^n = \sum_{n=1}^{\infty}\sum_{k=1}^{n} p_0(n-k)f_0(k)s^n =$$

$$\left(\sum_{k=1}^{\infty} f_0(k)s^k\right)\left(\sum_{n=k}^{\infty} p_0(n-k)s^{n-k}\right) = P_0(s)F_0(s).$$

Protože

$$\sum_{n=1}^{\infty} p_0(n)s^n = P_0 - 1,$$

dostáváme $P_0 - 1 = P_0F_0$, tedy $P_0 = 1 + P_0F_0$.

Pro důkaz 2) dostáváme z 1)

$$F_0 = \frac{P_0 - 1}{P_0} = 1 - \frac{1}{P_0} = 1 - \sqrt{1 - 4pqs^2}.$$

Důsledek 4.3. Pravděpodobnost toho, že se částice někdy vrátí do počátku, je rovna

$$\sum_{n=1}^{\infty} f_0(n) = F_0(1) = 1 - \mid p - q \mid .$$

Speciálně, pro symetrickou náhodnou procházku ($p = q$) je návrat jistý.

Důsledek 4.4. (**Pólyova věta v dimenzi 1**) Symetrická náhodná procházka se s pravděpodobností 1 vrátí do počátku.

Důsledek 4.5. Je-li $p = q = \frac{1}{2}$, tedy návrat je jistý, pak očekávání času T_0 prvního návratu do počátku je

$$E(T_0) = \sum_{n=1}^{\infty} n f_0(n) = F_0'(1) = \infty.$$

Důkaz: Máme

$$F_0(1) = 1 - \sqrt{1-4pq} = 1 - \sqrt{1-4p(1-p)} = 1 - \sqrt{1-4p+4p^2} =$$

$$= 1 - \sqrt{(1-2p)^2} = 1 - \mid 1-2p \mid = 1 - \mid 1-p-p \mid = 1 - \mid q-p \mid .$$

Pro $q = p = \frac{1}{2}$ je $F_0(1) = 1 - 0 = 1$.

Je-li návrat jistý, tedy $p = \frac{1}{2}$, pak $P(T_0 = \infty) = 0$, a tedy $E(T_0) = F_0'(1)$, kde $F_0 = 1 - \sqrt{1 - 4pqs^2}$. Máme

$$F_0' = -\frac{1}{2}\frac{1}{\sqrt{1-4pqs^2}}(-4pq2s) = \frac{4pqs}{\sqrt{1-4pqs^2}},$$

tedy $\lim\limits_{s \to 1_-} F_0'(s) = +\infty$.

Časy navštívení bodu r

Označme

$$f_r(n) = P(S_1 \neq r, S_2 \neq r, ..., S_{n-1} \neq r, S_n = r)$$

pravděpodobnost, že se náhodná procházka dostane poprvé do bodu r v čase n, s generující funkcí

$$F_r(s) = \sum_{n=0}^{\infty} f_r(n) s^n.$$

Věta 4.6. *Platí*

$$F_r(s) = [F_1(s)]^r$$

pro $r \geq 1$,

$$F_1(s) = \frac{1 - \left(1 - 4pqs^2\right)^{\frac{1}{2}}}{2qs}$$

Důkaz: Označme $T_r = min\{n;\ S_n = r\}$ počet kroků potřebných k dosažení hodnoty r poprvé. Nechť $r > 0$. Abychom dosáhli r, musíme se dostat nejdříve do bodu 1, a potom z bodu 1 do bodu r. To je z prostorové homogenity totéž jako dostat se z 0 do $r - 1$. Odtud plyne $T_r = T_1 + T_{r-1}$. Z nezávislosti tedy dostaneme $F_r = F_1 F_{r-1}$.

Máme pro $n > 1$

$$f_1(n) = P(S_1 \neq 1,\, S_2 \neq 1,\, ...,\, S_{n-1} \neq 1,\, S_n = 1) =: P(A) =$$

$$= P(A \mid S_1 = 1)P(S_1 = 1) + P(A \mid S_1 = -1)P(S_1 = -1) =$$

$$= P(A \mid S_1 = 1)p + P(A \mid S_1 = -1)q = 0p + f_2(n-1)q.$$

Odtud

$$f_1(n) = f_2(n-1)q.$$

Vynásobíme s^n

$$f_1(n)s^n = qf_2(n-1)s^n$$

a sečteme přes $n > 1$.

Tedy

$$\sum_{n=2} f_1(n)s^n = \sum_{n=2} qf_2(n-1)s^n =$$

$$\sum_{n=2} sqf_2(n-1)s^{n-1} = sq\sum_{n=1} f_2(n)s^n.$$

Odtud dostáváme

$$F_1 - ps = F_2 qs.$$

Víme, že $F_2 = F_1^2$, tedy

$$F_1 - ps = F_1^2 qs,$$

což vede ke kvadratické rovnici

$$qsF_1^2 - F_1 + ps = 0.$$

Řešením dostaneme dva kořeny

$$F_1 = \begin{cases} \frac{1-\sqrt{1-4qps^2}}{2qs} \\ \frac{1+\sqrt{1-4qps^2}}{2qs} \end{cases} .$$

Kořen $\frac{1+\sqrt{1-4qps^2}}{2qs}$ ale nevyhovuje zadání, protože má v bodě 0 limitu ∞. Tím je tvrzení dokázáno.

Důsledek 4.7. Pravděpodobnost, že náhodná procházka někdy navštíví kladnou část reálné osy, je rovna $min\left\{1, \frac{p}{q}\right\}$.

Důkaz: Hledaná pravděpodobnost je rovna pravděpodobnosti, že náhodná procházka navštíví bod 1. Ta je rovna

$$\sum_{n=1}^{\infty} f_1(n),$$

součtu pravděpodobností, že se dostaneme do bodu 1 v nějakém čase n. Víme, že

$$F_1(s) = \sum_{n=0}^{\infty} f_1(n) s^n.$$

Dosazením $s = 1$ dostaneme

$$\sum_{n=1}^{\infty} f_1(n) = F_1(1) = \frac{1 - (1 - 4pq)^{\frac{1}{2}}}{2q} = \frac{1 - \sqrt{1 - 4p(1-p)}}{2(1-p)} =$$

$$\frac{1 - \sqrt{(1-2p)^2}}{2(1-p)} = \frac{1 - \mid 1 - 2p \mid}{2(1-p)} = \frac{1 - \mid q - p \mid}{2q} =$$

$$= \begin{cases} \frac{1-(-(q-p))}{2q} = \frac{1-p+q}{2q} = \frac{2q}{2q} = 1 & \text{pro } q < p \\ \frac{1-q+p}{2q} = \frac{2p}{2q} = \frac{p}{q} & \text{pro } q > p \end{cases}.$$

Příklad 4.8. Ruleta má 37 čísel (včetně 0). Budeme sázet stále na lichá čísla, tedy $p = \frac{18}{37}$ a $q = \frac{19}{37}$. Pravděpodobnost, že budeme někdy vyhrávat je $\frac{p}{q} = \frac{18}{19}$.