2. technika – počítání trajektorií

Uvažujme náhodnou procházku vycházející z bodu a. Máme tedy

$$S_0 = a, \quad P(X_i = 1) = p, \quad P(X_i = -1) = q,$$

a

$$S_n = a + \sum_{i=1}^{n} X_i.$$

Zvolme pevně danou cestu.

Pravděpodobnost, že prvních n kroků bude sledovat právě tuto cestu, je rovna

$$p^r q^l,$$

kde r je počet kroků doprava a l je počet kroků doleva.

Tedy

$$r = \mid \{i : S_{i+1} - S_i = 1, i \leq n-1\} \mid,$$

kde $\mid . \mid$ značí velikost množiny.

Každý jev můžeme vyjádřit pomocí vhodné množiny trajektorií (které jsou s ním v souladu).

Jeho pravděpodobnost je součet pravděpodobností těchto trajektorií.

Máme

$$P(S_n = b) = \sum_r M_n^r(a, b) p^r q^{n-r},$$

kde $M_n^r(a, b)$ je počet cest (S_0, S_1, ..., S_n) takových, že $S_0 = a$, $S_n = b$, a majících přesně r kroků doprava.

Víme ale, že $r + l = n$ a $r - l = b - a$.

Odtud $2r = n + b - a$, tedy

$$r = \frac{n + b - a}{2}$$

a

$$l = \frac{n - b + a}{2}.$$

r je tedy určeno hodnotami a, b, n a v označení je *nadbytečné*

Tedy

$$P(S_n = b) = \binom{n}{r} p^{\frac{n+b-a}{2}} q^{\frac{n-b+a}{2}} = \binom{n}{\frac{1}{2}(n+b-a)} p^{\frac{n+b-a}{2}} q^{\frac{n-b+a}{2}},$$

pokud $\frac{1}{2}(n+b-a) \in \mathbb{N}$ (jinak je pravděpodobnost rovna 0).

Princip reflexe

Označme $N_n(a,b)$ počet všech cest z bodu $(0,a)$ do bodu (n,b).

Víme, že

$$N_n(a,b) = M_n^r(a,b) = \binom{n}{\frac{1}{2}(n+b-a)}.$$

Nechť $N_n^0(a,b)$ je počet všech cest z bodu $(0,a)$ do bodu (n,b), které obsahují nějaký bod $(k,0)$ na ose x, tedy navštíví bod 0.

Věta 3.1. ***(princip reflexe)****: Je-li a, $b > 0$, pak platí*

$$N_n^0(a,b) = N_n(-a,b).$$

Důkaz: Každá cesta z bodu $(0, -a)$ do (n, b) protne osu x poprvé v nějakém bodě $(k, 0)$.

Reflexí této cesty okolo osy x dostaneme cestu z bodu $(0, a)$ do (n, b), která navštíví osu x.

Tato operace dává vzájemně jednoznačnou korespondenci mezi

– cestami z $(0, -a)$ do (n, b)

– cestami $(0, a)$ do (n, b), které navštíví osu x.

Tedy $N_n^0(a, b) = N_n(-a, b)$.

Věta 3.2. *(o volbách): Je-li $b > 0$, pak počet cest z bodu $(0, 0)$ do bodu (n, b), které se nevrátí do bodu 0, je*

$$\frac{b}{n} N_n(0, b),$$

kde $N_n(0, b)$ je počet všech cest z bodu $(0, 0)$ do bodu (n, b).

Důkaz: Pro všechny takové cesty je první krok bod $(1, 1)$ (jinak se nutně dostaneme do nuly), tedy jejich počet je roven (z časové homogenity):

$$N_{n-1}(1, b) - N^0_{n-1}(1, b) = N_{n-1}(1, b) - N_{n-1}(-1, b)$$

Dále máme

$$N_{n-1}(1,b) - N_{n-1}(-1,b) = \binom{n-1}{\frac{1}{2}(n-1+b-1)} - \binom{n-1}{\frac{1}{2}(n+b)}$$

$$= \binom{n-1}{m-1} - \binom{n-1}{m} = \frac{b}{n}\binom{n}{m} = \frac{b}{n}N_n(0,b),$$

kde jsme označili $m = \frac{1}{2}(n+b)$.

Využili jsme identitu

$$\binom{n-1}{m-1} - \binom{n-1}{m} = \frac{b}{n}\binom{n}{m}$$

Její důkaz je za domácí úkol

Příklad 3.3. (*Úloha o volbách*) Kandidát A má α hlasů; kandidát B dostal β hlasů, kde $\alpha > \beta$, tj. kandidát A zvítězil. Jaká je pravděpodobnost, že během voleb byl A celou dobu před B?

Označme $X_i = 1$, je-li i-tý hlas pro A, $X_i = -1$, je-li i-tý hlas pro kandidáta B. Je tedy $n = \alpha + \beta$. Součet

$$S_k = \sum_{i=1}^{k} X_i$$

popisuje, o kolik vede A nad B v čase k (případně prohrává, je-li $S_k < 0$).

Podle věty o volbách je trajektorií z bodu $(0,0)$ do bodu $(\alpha+\beta, \alpha-\beta)$, které se nedostanou do 0, přesně

$$\frac{\alpha-\beta}{\alpha+\beta} N_{\alpha+\beta}\left(0, \alpha-\beta\right).$$

Hledaná pravděpodobnost je tedy

$$\frac{\frac{\alpha-\beta}{\alpha+\beta} N_{\alpha+\beta}\left(0, \alpha-\beta\right)}{N_{\alpha+\beta}\left(0, \alpha-\beta\right)} = \frac{\alpha-\beta}{\alpha+\beta}.$$

Nyní se budeme zajímat o to, jaká je pravděpodobnost, že se náhodná procházka nevrátí do 0. Máme $S_0 = 0$ a chceme $S_1 \neq 0$, ..., $S_n \neq 0$, jinak řečeno $S_1 S_2 \ldots S_n \neq 0$.

Věta 3.4. *Je-li $S_0 = 0$, pak pro $n \geq 1$ platí*

$$P(S_1 S_2 \ldots S_n \neq 0 \mid S_n = b) = \frac{\mid b \mid}{n} P(S_n = b),$$

tedy

$$P(S_1 S_2 \ldots S_n \neq 0) = \frac{1}{n} E(|S_n|).$$

Důkaz: Nechť $S_n = b > 0$.

Podle věty o volbách je počet cest z bodu $(0,0)$ do bodu (n,b),

které nenavštíví počátek celkem

$$\frac{b}{n} \cdot N_n(0,b).$$

Tedy

$$\mathrm{P}(S_n = b \,\&\, S_1 \cdot S_2 \cdots S_n \neq 0) = \frac{b}{n} \cdot N_n(0,b) \cdot p^{\frac{n+b}{2}} \cdot q^{\frac{n-b}{2}} = \frac{b}{n} \cdot \mathrm{P}(S_n = b).$$

Podobně pro $b < 0$.

Generující funkce

Uvažujeme posloupnost reálných čísel

$$a = \{a_n;\ n = 0,\ 1,\ 2,\ ..\}\,.$$

Taková posloupnost obsahuje velké množství informace, kterou můžeme výhodně "zakódovat" do jediného objektu (funkce).

S ním budeme moci lépe pracovat.

Získáme možnost použít operace (např. derivaci), které pro posloupnosti nemají smysl.

Generující funkce posloupnosti a je funkce daná součtem mocninné řady

$$G_a(s) = \sum_{n=0}^{\infty} a_n s^n$$

pro $s \in \mathbb{R}$, pro která řada konverguje.

Posloupnost a dostaneme z generující funkce G_a zpět vztahem

$$a_n = \frac{G_a^{(n)}(0)}{n!},$$

kde $G_a^{(n)}(0)$ je n-tá derivace G_a v bodě 0.

Příklad: Nechť $a = \{0, 1, 0, -1, 0, 1, 0, ...\}$. Pak

$$G_a = s - s^3 + s^5 - s^7 + \ ... \ ,$$

což je geometrická řada s prvním členem s a s kvocientem $q = -s^2$. Tedy

$$G_a(s) = \frac{s}{1 + s^2}$$

pro $\mid s \mid < 1$ (obor konvergence).

Dále budeme definovat generující funkci diskrétní náhodné veličiny.

Nechť X je diskrétní náhodná veličina, která nabývá hodnoty v množině $\{0, 1, 2, ...\}$, s pravděpodobnostní funkcí

$$f(i) = P(X = i).$$

Generující funkce náhodné veličiny X je definovaná jako generující funkce její pravděpodobnostní funkce, tedy

$$G_X(s) = \sum_{i=0}^{\infty} f(i)s^i = \sum_{i=0}^{\infty} P(X = i)s^i.$$

Platí zřejmě

$$G_X(s) = E(s^X).$$

Základní vlastnosti generujících funkcí:

– Existuje nezáporné číslo R (poloměr konvergence) takové, že $G(s)$ konverguje pro $\mid s \mid < R$ a diverguje pro $\mid s \mid > R$.

– $G(s)$ můžeme derivovat nebo integrovat člen po členu, libovolně mnohokrát, pro $\mid s \mid < R$.

–Jednoznačnost: Je-li $G_a(s) = G_b(s)$ pro $\mid s \mid < R'$, kde $0 < R' \leq R$, pak $a_n = b_n$ pro všechna n.

Příklady generujících funkcí náhodných veličin:

1. Konstantní náhodná veličina. $P(N=k)=1$, kde $k \in \mathbb{N} \cup \{0\}$.

Máme

$$G_N(s) = 1s^k = s^k.$$

2. Bernoulliho náhodná veličina. $P(N=1)=p$ a

$P(N=0)=1-p$. Tedy

$$G_N(s) = ps^1 + (1-p)s^0 = 1-p+ps.$$

3. Geometrické rozdělení. $P(N = k) = p(1-p)^k$ pro $k \in \mathbb{N}_0$.

Počet neúspěchů před prvním úspěchem

Dostaneme

$$G_N(s) = \sum_{i=0}^{\infty} p_N(n) s^n = \sum_{n=0}^{\infty} p(1-p)^n s^n = \sum_{n=0}^{\infty} p\left[(1-p)s\right]^n =$$

$$= \frac{p}{1-(1-p)s} = \frac{p}{1-s+sp}.$$

4. Poissonovo rozdělení s parametrem λ.

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Tedy

$$G_X(s) = \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} e^{-\lambda} s^i = e^{-\lambda} \sum_{i=0}^{\infty} \frac{(\lambda s)^i}{i!} = e^{-\lambda} e^{\lambda s} = e^{\lambda s - \lambda} = e^{\lambda(s-1)},$$

s využitím $\sum_n \frac{x^n}{n!} = e^x$.

Charakteristiky náhodných veličin a generující funkce

Základní charakteristiky n.v., $E(X)$ a $Var(X)$, lze snadno spočítat pomocí $G_X(s)$.

Věta: Nechť X je náhodná veličina s generující funkcí $G_X(s)$. Pak platí:

$E(X) = G'_X(1)$.

Obecně,

$$E(X(X-1)...(X-k+1)) = G_X^{(k)}(1)$$

(tzv. k-tý faktoriální moment).

Důkaz: První tvrzení je speciální případ druhého. Máme

$$G_X^{(k)}(s) = \sum_i s^{i-k} i(i-1)...(i-k+1) p_X(i) =$$

$$= E(s^{X-k} X(X-1)...(X-k+1)).$$

Pro $s = 1$ dostaneme

$$G_X^{(k)}(1) = E(X(X-1)...(X-k+1)).$$

Pro rozptyl dostaneme speciálně vztah

$$Var(X) = E(X^2) - E(X)^2 = E(X(X-1) + X) - E(X)^2 =$$

$$= E(X(X-1)) + E(X) - E(X)^2 = G_X''(1) + G_X'(1) - \left[G_X'(1)\right]^2.$$

Součty náhodných veličin Nechť $a = \{a_i,\ i \geq 0\}$ a $b = \{b_i,\ i \geq 0\}$ jsou dvě posloupnosti, pak konvoluce $c = a \star b$ je posloupnost definovaná vztahem

$$c_n = a_0 b_n + a_1 b_{n-1} + \ldots + a_n b_0.$$

Jsou-li G_a, G_b generující funkce posloupností a a b, pak generující funkce posloupnosti c je

$$G_c(s) = \sum_{n=0}^{\infty} c_n s^n = \sum_{n=0}^{\infty} \left(\sum_{i=0}^{n} a_i b_{n-i} \right) s^n =$$

$$= \left(\sum_{i=0}^{\infty} a_i s^i\right) \left(\sum_{n=i}^{\infty} b_{n-i} s^{n-i}\right) = \left(\sum_{i=0}^{\infty} a_i s^i\right) \left(\sum_{n=0}^{\infty} b_n s^n\right) = G_a(s) G_b(s).$$

Tím jsme dokázali následující tvrzení.

Věta 3.5. *Generující funkce konvoluce dvou posloupností je součinem generujících funkcí těchto posloupností.*

Věta Nechť X a Y jsou nezávislé náhodné veličiny. Pak

$$G_{X+Y}(s) = G_X(s)G_Y(s).$$

Důkaz:

$$E(s^{X+Y}) = E(s^X s^Y) = E(s^X)E(s^Y)$$

první rovnost plyne z vlastností exponenciály, druhá z nezávislosti X a Y. $\square$

– Jiný důkaz dostaneme využitím předchozí věty.

Obecně, pro součet více nezávislých náhodných veličin dostaneme:

Je-li $S = X_1 + X_2 + ... + X_n$, kde X_i jsou nezávislé, pak z předchozí věty plyne

$$G_S = G_{X_1} G_{X_2} ... G_{X_n}.$$

Definice 3.6. **Sdružená pravděpodobnostní generující funkce náhodných veličin**, nabývajících hodnot v $\mathbb{N} \cup \{0\}$, je definovaná jako

$$G_{X_1,X_2}(s_1, s_2) = \sum P(X_1 = i \wedge X_2 = j)\, s_1^i s_2^j = E(s_1^{X_1} s_2^{X_2})$$

Analogicky je možné definovat sdruženou pravděpodobnostní generující funkci pro více náhodných veličin.