Náhodné cesty

ve 2D a ve 3D

Traduje se, že japonský matematik Kakutani Shizuo řekl v 70. letech,

„Opilý muž se vždy vrátí domů, ale opilý pták může bloudit navždy.“

Za tímhle rčením je schována pravděpodobnostní úloha. Řekněme, že opilý muž vyjde z domu a, bo je opilý, chodí náhodně. Slovo ‘náhodně’ tu chápeme v tom smyslu, že po každém kroku, který je pro jednoduchost vždy stejně dlouhý (i když to na výsledku úlohy nic nemění), pokračuje s pravděpodobností 1/4, nebo 25 %, v jednom ze 4 směrů. Opilý pták, bo létá, volí po každém kroku z 6 různých směrů.

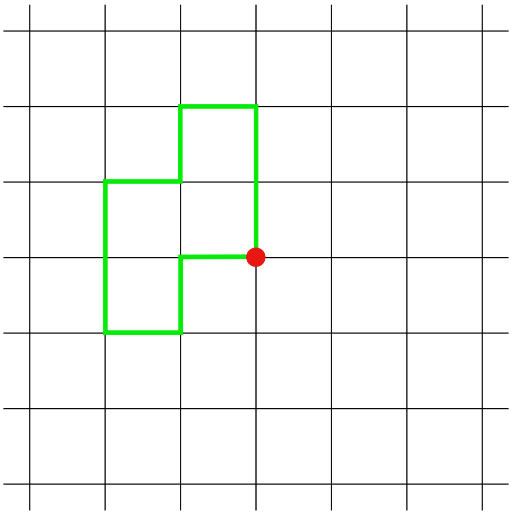
Tvrdíme, že opilý muž se po nějakém počtu učiněných kroků vždycky vrátí zpět do svého domu, ale ptákovi se může stát, že už cestu do hnízda zpátky nikdy nenajde.

Myslím si, že tohle je jedna z úloh, kdy naše přirozená lidská intuice selhává. Je to pravděpodobně proto, že nám lidem obecně schází intuice pro ‘nekonečné’ věci. Cílem tohoto (snad krátkého) textu je spočítat, že pronesené tvrzení je pravdivé.

Trocha formalismu. Budeme předpokládat, že opilá zvířata (ano, toto slovo zahrnuje i druh Člověka moudrého) se pohybují v reálném prostoru (buď 2D nebo 3D) a jedním *krokem* je vždy posun přesně o 1 ve směru souřadnicových os. *Cestou* délky k ve 2D budeme myslet **konečnou** posloupnost přesně k učiněných kroků. Cesta je pak *náhodná*, pokud má náhodnou délku a směr každého dalšího kroku je volen náhodně (ve smyslu prvního odstavce).

Teď můžeme úlohu přeformulovat tak, že pravděpodobnost toho, že se opilý muž vrátí, když chodí náhodně, je 1, nebo 100 %, a pravděpodobnost toho, že se opilý pták vrátí, když bude létat náhodně, je **ostře** menší než 100 %.

Pravděpodobnost toho, že se stane nějaká skutečnost/jev X, budu psát jako $\mathbf{P}(X)$. Symbolicky, pravděpodobnost návratu můžu psát třeba jako $\mathbf{P}(\text{návrat})$.



Obrázek 1: Kus náhodné cesty ve 2D se začátkem v počátku.

Pamatujte, že pravděpodobnost návratu **nezávisí** na volbě cesty! Naopak, pravděpodobnost návratu vyjadřuje, jakou šanci mám, že se vrátím, když zvolím nějakou náhodnou cestu.

Tuhle pravděpodobnost potřebujeme nějak hezky vyjádřit.

V tom nám brání následující pozorování. Hodnota $\mathbf{P}(\text{návrat})$ nerozlišuje mezi cestami, podle toho, kolikrát se vrátí. Cesta, která projde počátkem 50krát je v tomto smyslu stejná jako ta, která se vrátí jednou. Protože cesta, která se vrátí dvakrát, lze rozdělit do dvou cest, každáže z nich se vrátí jednou, budeme chtít úlohu převést na počítání s cestami, které se vrací přesně jednou.

Jako první krok k tomu si definujme náhodnou veličinu (to je pravděpodobnostní termín, stačí si to představovat jako něco, co má při každém 'experimentu' náhodnou hodnotu) V jako 'počet návratů'. Tedy například $\mathbf{P}(V = 4)$ představuje pravděpodobnost, že se vrátím do počátku přesně 4krát.

Jako druhý rok uvážíme tzv. 'střední hodnotu' V, což je velmi nevhodný český název pro údaj, kolikrát očekáváme, že se vrátíme. Pro intuitivní představu slouží následující příklad.

Příklad. Řekněme, že házím kostkou a výsledek hodu sleduje náhodná veličina H. Pak samozřejmě $\mathbf{P}(H = k) = 1/6$ pro každé číslo k od 1 do 6. Je rozumné očekávat, že hodím 1 s pravděpodobností 1/6, 2 s pravděpodobností 1/6 atd. Takže moje střední hodnota bodu bude

$$\begin{aligned}\mathbf{E}[H] &= 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} \\ &= 1 \cdot \mathbf{P}(H = 1) + 2 \cdot \mathbf{P}(H = 2) + 3 \cdot \mathbf{P}(H = 3) \\ &+ 4 \cdot \mathbf{P}(H = 4) + 5 \cdot \mathbf{P}(H = 5) + 6 \cdot \mathbf{P}(H = 6).\end{aligned}$$

Když to spočtete, vyjde vám $\mathbf{E}[H] = 3.5$. Všimněte si, že to je přesně aritmetický průměr čísel 1 až 6. To není náhoda. Když je pravděpodobnost každé možnosti stejná, je střední hodnota vždycky aritmetický průměr.

Intuitivně, střední hodnota je číslo, ke kterému se náhodná veličina 'nejvíce kloní'.

Znetvořme herní kostku tak, aby pravděpodobnost toho, že hodím 2 byla 0.4, že hodím 3 taky 0.4, a u všech ostatních čísel 0.05. Tvrdím, že teď bude střední hodnota hodu někde mezi 2 a 3, ale blíž 3, protože 2 a 3 jsou nejpravděpodobnější, ale 3 má napravo od sebe ještě tři další hodnoty, zatímco 2 má nalevo od sebe jen jednu. Schválně, máme

$$\begin{aligned}\mathbf{E}[H] &= 1 \cdot \mathbf{P}(H = 1) + 2 \cdot \mathbf{P}(H = 2) + 3 \cdot \mathbf{P}(H = 3) \\ &\quad + 4 \cdot \mathbf{P}(H = 4) + 5 \cdot \mathbf{P}(H = 5) + 6 \cdot \mathbf{P}(H = 6) \\ &= 1 \cdot 0.05 + 2 \cdot 0.4 + 3 \cdot 0.4 + 4 \cdot 0.05 + 5 \cdot 0.05 + 6 \cdot 0.05 = 2.8.\end{aligned}$$

Zpět k náhodným cestám. Z příkladu s kostkou je doufám aspoň částečně jasné, že očekávám, že se vrátím přesně k-krát s pravděpodobností toho, že počet návratů je k pro každé přirozené číslo $k \in \mathbb{N}$. To znamená, že

$$\mathbf{E}[V] = 1 \cdot \mathbf{P}(V = 1) + 2 \cdot \mathbf{P}(V = 2) + 3 \cdot \mathbf{P}(V = 3) + \dots,$$

kdy sčítám nekonečně mnoho členů, protože jsem nedal žádné omezení na délku cesty ani počet návratů. Pomocí symbolu Σ to můžu taky napsat jako

$$\mathbf{E}[V] = \sum_{k=1}^{\infty} k \cdot \mathbf{P}(V = k),$$

což přesně znamená, že spolu sčítám všechny výrazy $k \cdot \mathbf{P}(V = k)$, kde k (což je počet návratů) postupně probíhá všechna přirozená čísla.

Jak z tohohle součtu vykutat pravděpodobnost návratu $\mathbf{P}(\text{návrat})$, která nás reálně zajímá. Ten součet nahoře si můžeme napsat trochu jinak. Konkrétně, protože sčítanec $\mathbf{P}(V = k)$ se objevuje přesně k-krát, rozdělím si součet na

$$\begin{array}{ccccccccc}
\mathbf{P}(V=1) & + & & & & & & & \\
\mathbf{P}(V=2) & + & \mathbf{P}(V=2) & + & & & & & \\
\mathbf{P}(V=3) & + & \mathbf{P}(V=3) & + & \mathbf{P}(V=3) & + & & & \\
\mathbf{P}(V=4) & + & \mathbf{P}(V=4) & + & \mathbf{P}(V=4) & + & \mathbf{P}(V=4) & & \\
\vdots & & \vdots & & \vdots & & \vdots & & \ddots
\end{array}$$

Všimněte si, že v prvním sloupci mám součet

$$\mathbf{P}(V = 1) + \mathbf{P}(V = 2) + \mathbf{P}(V = 3) + \dots,$$

neboli pravděpodobnost, že se vrátím jednou nebo se vrátím dvakrát nebo se vrátím třikrát atd. To je ale přesně pravděpodobnost toho, že se vrátím **aspoň jednou**. Neboli, ten součet v prvním sloupci je **přesně** $\mathbf{P}$(návrat), protože to, že se vrátím, znamená, že se vrátím aspoň jednou.

Hodnotu $\mathbf{P}$(návrat) si označím pro strohost zápisu jako r.

Jak souvisí součet v druhém sloupci s r? No, úplně stejnou úvahou dojdeme k tomu, že součet v druhém sloupci mi dává pravděpodobnost, že se vrátím **aspoň dvakrát**.