

1. cvičení z STP

18. února 2021

Uvažujme výběr z n různých předmětů, které vybíráme k -krát. Počet všech jednotlivých možností pro různé způsoby výběru uvádí následující tabulka:

Výběr	bez opakování (vracení)	s opakováním (vracením)
uspořádaný	$n(n-1)\cdots(n-k+1)$	n^k
neuspořádaný	$\binom{n}{k}$	$\binom{n+k-1}{k}$

Jako úvod při počítání s pravděpodobnostmi budeme uvažovat situaci, že

- Ω bude nějaká (konečná) množina všech možných výsledků (které můžeme získat) a které budeme považovat za "rovnocenné"
- jevem budeme rozumět (zatím jakoukoliv) podmnožinu A množiny Ω (jev tedy představuje např. nějakou vlastnost, kterou dané výsledky obsažené v A sdílí)
- jevu A pak přiřadíme pravděpodobnost předpisem $P(A) = \frac{|A|}{|\Omega|}$, čímž dáváme vlastně najevo, tu "rovnocennost" výsledků, kdy pravděpodobnost daného jevu A je úměrná pouze velikosti $|A|$ a nikoliv tomu, které konkrétní výsledky tento jev obsahuje.

Takovéto počítání pravděpodobnosti se nazývá *Laplaceův model pravděpodobnosti* (nebo někdy jen *Laplaceova pravděpodobnost*).

Příklad 1.1 *Házíme dvěma kostkami. Stanovte pravděpodobnost jevu*

$A =$ "na kostkách padne součet menší než 5".

Řešení:

Výsledky pokusu jsou uspořádané dvojice, kde první člen dvojice odpovídá hodu 1. kostkou a druhý člen odpovídá hodu 2. kostkou. Tedy množina všech možných výsledků je $\Omega = \{(i, j) \mid i, j = 1, \dots, 6\}$ a vypsáno konkrétně je to:

(1,1) (1,2) (1,3) (1,4) (1,5) (1,6),
 (2,1) (2,2) (2,3) (2,4) (2,5) (2,6),
 (3,1) (3,6),
 (4,1) (4,6),
 (5,1) (5,6),
 (6,1) (6,2) (6,3) (6,4) (6,5) (6,6),

tzn. počet všech možných výsledků je $|\Omega| = 36$. Jev A pak představuje podmnožinu $A = \{(i, j) \in \Omega \mid i + j < 5\}$ a konkrétně je to

$$A = \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (3, 1)\}$$

Počet výsledků příznivých jevu A je tak $|A| = 6$. Všechny možné výsledky považujeme za stejně pravděpodobné, proto je hledaná pravděpodobnost jevu A rovna $P(A) = \frac{|A|}{|\Omega|} = \frac{6}{36} = \frac{1}{6}$.

Poznámka: Předpokládejme, že v předchozím příkladu máme dvě nerozlišitelné kostky. Jestliže pak např. padnou současně čísla 1 a 2, neumíme poznat, co padlo na jaké kostce, a jako výsledek máme tedy jen *neuspořádanou* dvojici $\{1, 2\}$ (a podobně u dalších výsledků). Jestliže bychom nyní chtěli postupovat jako v předchozím příkladu, měli bychom množinu výsledků

$$\Omega' = \{\{i, j\} \mid 1 \leq i \leq j \leq 6\},$$

kde zápisem $\{i, j\}$ nyní (na chvíli) myslíme neuspořádanou dvojici, tj. Ω je tvořeno prvky

$$\begin{aligned} &\{1, 1\} \{1, 2\} \{1, 3\} \{1, 4\} \{1, 5\} \{1, 6\}, \\ &\quad \{2, 2\} \{2, 3\} \{2, 4\} \{2, 5\} \{2, 6\}, \\ &\quad \quad \{3, 3\} \{3, 4\} \{3, 5\} \{3, 6\}, \\ &\quad \quad \quad \{4, 4\} \{4, 5\} \{4, 6\}, \\ &\quad \quad \quad \quad \{5, 5\} \{5, 6\}, \\ &\quad \quad \quad \quad \quad \{6, 6\}, \end{aligned}$$

kterých je $|\Omega'| = 1 + 2 + 3 + 4 + 5 + 6 = 21$. Jev "součet je menší než 5" pak bude odpovídat množině

$$A' = \{\{1, 1\}, \{1, 2\}, \{2, 2\}, \{1, 3\}\}$$

která má $|A'| = 4$ prvky. Pokud bychom nyní chtěli určit pravděpodobnost opět způsobem $P(A') = \frac{|A'|}{|\Omega'|} = \frac{4}{21}$, zjistíme, že jsme dostali jinou (dokonce větší) hodnotu než $\frac{1}{6}$ z předchozího příkladu. Který postup byl tedy správně a proč?

Problém je v druhém postupu, a sice v tom, že jsme mlčky opět předpokládali, že výsledky množiny Ω' jsou rovnocenné. To ale už nemůžeme, protože např. výsledek $\{1, 2\}$ vzniká ze dvou případů $(1, 2)$ a $(2, 1)$, zatímco výsledek $\{1, 1\}$ pouze z jednoho případu $(1, 1)$. I když tedy kostky třeba rozlišit neumíme, jsou stále dvě a každá má své výsledky (bez ohledu na naši schopnost je od sebe odlišit). Proto je také výsledek $\{1, 2\}$ dvakrát častější při skutečných fyzických hodech kostkami než výsledek $\{1, 1\}$ a měl by proto mít také dvojnásobnou pravděpodobnost (oproti výsledku $\{1, 1\}$).

Proto neuspořádané dvojice (obecněji: neuspořádané výběry) v tomto případě nemůžou sloužit jako základ pro Laplaceovu pravděpodobnost a my si nutně musíme vzít uspořádané dvojice (obecněji: uspořádané výběry). Na druhé straně i výsledky popisované neuspořádanými dvojicemi lze použít, ovšem s tím, že daná neuspořádaná dvojice bude mít pravděpodobnost úměrnou počtu jejích uspořádaných verzí, tj. např.

- $P(\{1, 1\}) = \frac{1}{36}$
- $P(\{1, 2\}) = 2 \times \frac{1}{36} = \frac{1}{18}$.

To už je obecnější model pravděpodobnosti než ten Laplaceův. Ovšem, jak je vidět, stejně se i v tomto případě vlastně z Laplaceova modelu odvozuje a z hlediska počítání je tak stejně výhodnější přejít k uspořádaným výběrům.

Příklad 1.2 V balíčku máme 32 karet, z toho 4 esa. Dvakrát za sebou vytáhneme náhodně jednu kartu. Stanovte pravděpodobnost jevu

$$A = \text{"alespoň jedna z vytažených karet je eso"},$$

jestliže po prvním tahu kartu

- (1) vrátíme,
- (2) nevrátíme

zpět do balíčku.

Řešení:

(1) Výsledky pokusu jsou opět uspořádané dvojice (množina Ω_1). První člen dvojice odpovídá kartě vytažené v prvním tahu a druhý člen kartě vytažené v druhém tahu. V prvním tahu můžeme kartu vytáhnout 32 způsoby. Protože vytaženou kartu vracíme zpět do urny, i v druhém tahu máme 32 možností. Počet všech možných případů je tedy $|\Omega_1| = 32^2$. Příznivým případům odpovídají tahy (libovolná karta - eso), (eso - libovolná karta), (eso - eso). Počet příznivých případů je $|A_1| = 28 \cdot 4 + 4 \cdot 28 + 4 \cdot 4$. Hledaná pravděpodobnost je rovna

$$P(A_1) = \frac{|A_1|}{|\Omega_1|} = \frac{28 \cdot 4 + 4 \cdot 28 + 4 \cdot 4}{32^2} = \frac{15}{64} \doteq 0.2344.$$

Jednodušeji se k výsledku můžeme dostat přes doplňkový jev

$$A_1^c = \Omega_1 \setminus A_1 = \text{“žádná z vytažených karet není eso”},$$

Pro pravděpodobnost pak platí, že

$$P(A_1^c) = \frac{|\Omega_1 \setminus A_1|}{|\Omega_1|} = \frac{|\Omega_1| - |A_1|}{|\Omega_1|} = 1 - P(A_1)$$

Počet prvků A_1^c je pak (analogicky jako u Ω_1) roven 28^2 , takže

$$P(A_1) = 1 - P(A_1^c) = 1 - \frac{28 \cdot 28}{32 \cdot 32} = 1 - \frac{7 \cdot 7}{8 \cdot 8} = \frac{15}{64}$$

(2) Tentokrát z uspořádaných dvojic karet musíme vyloučit ty, kde první i druhá karta jsou stejné (dostaneme tak množinu Ω_2). Počet možných případů je vzhledem k tomu, že po prvním tahu kartu nevrátíme, $|\Omega_2| = 32 \cdot 31$. Příznivým případům odpovídají opět tahy (libovolná karta - eso), (eso - libovolná karta), (eso - eso). Počet příznivých případů je nyní $|A_2| = 28 \cdot 4 + 4 \cdot 28 + 4 \cdot 3$. Hledaná pravděpodobnost je rovna

$$P(A_2) = \frac{|A_2|}{|\Omega_2|} = \frac{2 \cdot 4 \cdot 28 + 4 \cdot 3}{32 \cdot 31} = \frac{59}{248} \doteq 0.2379.$$

Přes doplňkový jev je to opět jednodušší:

$$A_2^c = \text{“žádná z vytažených karet není eso”},$$

$$|A_2^c| = 28 \cdot 27$$

$$P(A_2) = 1 - P(A_2^c) = 1 - \frac{28 \cdot 27}{32 \cdot 31} = \frac{59}{248}$$

Kromě toho je vidět, že pravděpodobnosti se v případech vrácení (0.2344) i nevrácení (0.2379) příliš neliší. Je to proto, že pokud máme velké množství N (zde $N = 32$), ze kterého taháme pouze k -krát (zde $k = 2$), tj. ”několikrát” v porovnání s tím, jak velké množství máme, neboli $k \ll N$, tak pravděpodobnost, že bychom opakovaně vytáhli znovu tentýž předmět je zanedbatelná. Tudíž obě pravděpodobnosti se budou téměř shodovat.

Příklad 1.3 Na party se sešlo 14 studentů, z toho 8 vysokoškoláků a 6 středoškoláků. Jaká je pravděpodobnost, že v náhodně vybrané čtveřici budou

- (1) všichni čtyři středoškoláci (jev A),
- (2) právě jeden vysokoškolák (jev B),
- (3) aspoň jeden vysokoškolák (jev C)?

Řešení:

Prostor všech možných rovnocenných výsledků je

$$\Omega = \text{“všechny neuspořádané 4-ce utvořené ze 14 osob”}$$

tedy všechny 4-prvkové podmnožiny 14-ti prvkové množiny (neboli *neuspořádaný vyber 4 prvku ze 14 prvků bez opakování*). Jejich počet je kombinační číslo $|\Omega| = \binom{14}{4} = \frac{14!}{4!(14-4)!}$.

(1) Počet všech čtveřic vytvořených pouze ze středoškoláků je $\binom{6}{4}$, tudíž hledaná pravděpodobnost je

$$P(A) = \frac{\binom{6}{4}}{\binom{14}{4}} \doteq 0.015.$$

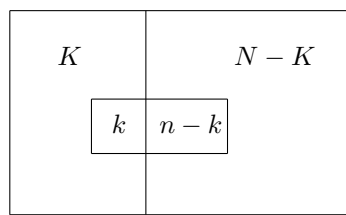
(2) Počet všech trojic vytvořených ze středoškoláků je $\binom{6}{3}$, počet všech "jednic" vytvořených z vysokoškoláků je $\binom{8}{1} = 8$. K vytvoření příznivých čtveřic je třeba všechny trojice zkombinovat se všemi "jednicemi", tudíž hledaná pravděpodobnost je

$$P(B) = \frac{\binom{8}{1} \cdot \binom{6}{3}}{\binom{14}{4}} \doteq 0.16.$$

(3) Jedná se o doplňkový jev k jevu A , tj. $C = A^c$, tudíž

$$P(C) = 1 - P(A) \doteq 0.985.$$

Poznámka: V tomto příkladu jsme setkali s tzv. *hypergeometrickým* rozdělením.



To se objevuje tehdy, když

- z množiny, která má N prvků (zde $N = 14$),
- z nichž právě K má nějakou vlastnost \mathcal{V} (zde $K = 8$ a \mathcal{V} = "být vysokoškolák"),
- chceme vybrat právě n prvků (zde $n = 4$)

a ptáme se, s jakou pravděpodobností bude právě k z nich mít vlastnost \mathcal{V} .

Tato pravděpodobnost je dána podílem příznivých možností ku všem. Příznivé jsou dány počtem způsobů jak vybrat k prvků (s \mathcal{V}) z K prvků (s \mathcal{V}) násobeno počtem způsobů jak vybrat zbytek, tj. $n - k$ prvků (bez \mathcal{V}) z $N - K$ prvků (bez \mathcal{V}). Celkem tedy

$$\frac{\binom{K}{k} \cdot \binom{N-K}{n-k}}{\binom{N}{n}}$$

přičemž rozsah proměnné k je určen pomocí

$$\max\{0, n + K - N\} \leq k \leq \min\{n, K\}$$

tedy

$$0 = \max\{0, 4 + 8 - 14\} \leq k \leq \min\{4, 8\} = 4 .$$

Tuto podmínku dostaneme ihned z nerovností

$$k \leq K, \quad n - k \leq N - K,$$

$$k \leq n, \quad n - k \leq n .$$

Příklad 1.4 Jaká je pravděpodobnost, že ve skupině n lidí mají alespoň dva narozeniny ve stejný den? (Neuvažujte přestupné roky a předpokládejte, že se během celého roku děti rodí rovnoměrně.)

Řešení:

První část z předpokladů znamená, že počet všech dnů v roce pro nás bude vždy 365. Druhá část předpokladů pak znamená, že všechny dny v roce považujeme za rovnocenné. Tedy pravděpodobnost, že se daný člověk narodí v daný den v roce bude pro všechny dny stejná. To, že pracujeme se skupinou n lidí si také můžeme ekvivalentně představit tak, že máme urnu s lístky s čísly od 1 do 365 (představujícími očíslované dny v roce) a my z ní n -krát opakovaně budeme losovat čísla s tím, že lístky vždy vrátíme zpět.

Výsledky pokusu jsou tudíž uspořádané n -tice s hodnotami od 1 do 365. Tedy $\Omega = \{1, \dots, 365\}^n$ (tj. kartézský součin množiny $\{1, \dots, 365\}$ a to celkem n -krát, podobně jako třeba zapisujeme \mathbb{R}^n).

Nechť A je jev, že ve skupině n lidí mají alespoň dva narozeniny ve stejný den. Pak jev A^c znamená, že ve skupině n lidí má každý člověk narozeniny v jiný den. Jde tedy o variace bez opakování třídy n z 365 prvků.

$$|\Omega| = 365^n$$

$$|A^c| = 365 \cdot 364 \cdot \dots \cdot (365 - n + 1)$$

a tedy

$$P(A^c) = \frac{365 \cdot 364 \cdot \dots \cdot (365 - n + 1)}{365^n},$$

a tudíž

$$P(A) = 1 - P(A^c) = 1 - \frac{365 \cdot 364 \cdot \dots \cdot (365 - n + 1)}{365^n}.$$

Pro zajímavost se ještě podíváme na přibližnou hodnotu této pravděpodobnosti. Pro jednoduchost označme $H = 365$. Pak máme

$$\begin{aligned} P(A) &= 1 - \frac{H \cdot (H - 1) \cdot \dots \cdot (H - (n - 1))}{H^n} = 1 - \left(1 - \frac{1}{H}\right) \dots \left(1 - \frac{n - 1}{H}\right) = \\ &= 1 - e^{\ln\left(1 - \frac{1}{H}\right) \dots \left(1 - \frac{n - 1}{H}\right)} = 1 - e^{\sum_{k=1}^{n-1} \ln\left(1 - \frac{k}{H}\right)} \end{aligned}$$

Použijeme teď lineární aproximaci funkce

$$f(x) = \sum_{k=1}^{n-1} \ln(1 - kx)$$

v bodě $x_0 = 0$, kterou pak vyčíslíme pro $x = \frac{1}{H}$, která je blízka nule, a to jako $f(x) \approx f(0) + f'(0) \cdot (x - 0)$. Tedy

$$f'(x) = \left(\sum_{k=1}^{n-1} \ln(1 - kx) \right)' = - \sum_{k=1}^{n-1} \frac{k}{1 - kx}$$

$$f'(0) = - \sum_{k=1}^{n-1} k = - \frac{n(n-1)}{2}$$

a

$$\sum_{k=1}^{n-1} \ln\left(1 - \frac{k}{H}\right) = f\left(\frac{1}{H}\right) \approx 0 + f'(0) \cdot \frac{1}{H} = - \frac{n(n-1)}{2H}.$$

Odsud máme, že

$$P(A) \doteq 1 - e^{-\frac{n(n-1)}{2 \cdot 365}}$$

a ukázkou některých hodnot v tabulce:

n	23	24	25	26	27	50
$P(A)$	50%	53.05%	56.04%	58.95%	61.77%	96.51%