2. cvičení z PRA

26. února - 1. března 2023

Příklad 2.1 *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

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

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)
$$A = \text{“\textit{všechny neuspořádané} 4\textit{-ce utvořené ze} 6 \textit{ osob (středoškoláků)}”}$$

Počet všech takových čtveřic je $\binom{6}{4}$, tudíž hledaná pravděpodobnost je

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

(2) Požadované čtveřice pro B můžeme popsat pomocí množiny všech ”jednic” vytvořených z vysokoškoláků (označme ji jako J) a trojic vytvořených ze středoškoláků (označme ji jako T). K vytvoření příznivých čtveřic je třeba všechny trojice zkombinovat se všemi ”jednicemi”, tedy ekvivalentně pomocí kartézského součinu:

$$B = J \times T$$

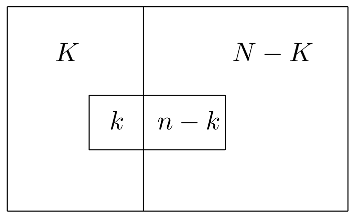
Zřejme $|T| = \binom{6}{3}$ a $|J| = \binom{8}{1} = 8$. Tudíž hledaná pravděpodobnost je

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

(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 2.2 *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. Situaci, která popisuje, jaké kdo má narozeniny ve skupině n lidí vyjádříme jako uspořádanou n-tici nad 365 prvky (i-tá hodnota v n-tici odpovídá datu narození i-tého člověka).

Všechny možné výsledky tedy představují množinu $\Omega = \{1, \ldots, 365\}^n$ (tj. kartézský součin množiny $\{1, \ldots, 365\}$ a to celkem n-krát, podobně jako třeba zapisujeme $\mathbb{R}^n$). Protože všechny dny považujeme za rovnocenné, tak i n-ticemi z nich vytvořené budeme považovat za rovnocenné. Také si to můžeme 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.

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

$$P(A) = 1 - \frac{H \cdot (H-1) \cdot \dots \cdot (H-(n-1))}{H^n} = 1 - \left(1 - \frac{1}{H}\right) \dots \left(1 - \frac{n-1}{H}\right) =$$

$$= 1 - e^{\ln\left(1-\frac{1}{H}\right)\dots\left(1-\frac{n-1}{H}\right)} = 1 - e^{\sum_{k=1}^{n-1} \ln\left(1-\frac{k}{H}\right)}$$

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 blízká 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 ukázku některých hodnot v tabulce:

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

Jevy $A_1, \dots, A_n$ jsou nezávislé právě když pro každou indexovou podmnožinu $\emptyset \neq K \subseteq \{1, \dots, n\}$ platí

$$P\left(\bigcap_{i \in K} A_i\right) = \prod_{i \in K} P(A_i) \ .$$

Tj. pravděpodobnost libovolných průniků je součin pravděpodobností příslušných jevů.

Poznámka: Mějme n nezávislých opakování daného pokusu, jehož úspěšnost je $0 < p < 1$. Jev

$$B_k = \text{“}\textit{nastane právě } k \textit{ úspěchů v } n \textit{ pokusech}\text{”}$$

pro $k = 0, 1, \ldots, n$ má pravděpodobnost

$$P(B_k) = \binom{n}{k} p^k \cdot (1-p)^{n-k} \ .$$

Důvod, proč tu máme kombinační číslo, je v tom, že daných k úspěchů je rozmístěno mezi n pokusů právě $\binom{n}{k}$ způsoby.

Odvození: Vezmeme si jevy

$$A_i = \text{“}\textit{úspěch v } i\textit{-tém pokusu}\text{”}$$

Ty jsou nezávislé a mají pravděpodobnost $P(A_i) = p$. Pak jev B_k vyjádříme jako sjednocení disjunktních jevů (ty v hranaté závorce)

$$B_k = \bigcup_{\substack{K \subseteq \{1,\ldots,n\} \\ |K|=k}} \left[\left(\bigcap_{i \in K} A_i \right) \cap \left(\bigcap_{j \in \{1,\ldots,n\} \setminus K} A_j^c \right) \right]$$

odkud máme ihned

$$P(B_k) = \sum_{\substack{K \subseteq \{1,\ldots,n\} \\ |K|=k}} \left[\underbrace{\left(\prod_{i \in K} P(A_i) \right)}_{p^k} \cdot \underbrace{\left(\prod_{j \in \{1,\ldots,n\} \setminus K} P(A_j^c) \right)}_{(1-p)^{n-k}} \right] =$$

$$= \sum_{\substack{K \subseteq \{1,\ldots,n\} \\ |K|=k}} p^k \cdot (1-p)^{n-k} = \binom{n}{k} p^k \cdot (1-p)^{n-k} \ .$$

Příklad 2.3 *Pravděpodobnost výhry* 1.*hráče nad* 2.*hráčem je* 0.7*. Jaká je pravděpodobnost, že během deseti po sobě jdoucích zápasů*

(1) alespoň jednou vyhrál 2.*hráč (jev H),*

(2) maximálně dvakrát vyhrál 1.*hráč (jev I)?*

Řešení:
Jevy

$$B_k = \text{“ 1.}\textit{hráč vyhrál právě } k \textit{ z } 10 \textit{ zápasů}\text{”}$$

pro $k = 0, 1, \ldots, 10$ jsou evidentně navzájem disjunktní (tj. mají prázdné průniky neboli žádné dva nemůžou nastat současně) a mají pravděpodobnost

$$P(B_k) = \binom{10}{k} 0.7^k \cdot 0.3^{10-k} \ .$$

(a) Zde bude vhodnější přejít k doplňku a pak je

$$H^c = \text{“ 1.}\textit{hráč vyhrál všech } 10 \textit{ zápasů}\text{”} = B_{10}$$

$$P(H) = 1 - P(H^c) = 1 - P(B_{10}) = 1 - 0.7^{10} \doteq \mathbf{0.9718}.$$

(b) Jev I znamená, že

- buď vše vyhrál 2.hráč

- nebo devět her vyhrál 2.hráč a jednu 1.hráč (na pořadí hry nezáleží)
- nebo osm her vyhrál 2.hráč a dvě 1.hráč (na pořadí rovněž nezáleží).

Tedy je

$$I = B_0 \cup B_1 \cup B_2$$

a z jejich disjunktnosti máme

$$P(I) = P(B_0) + P(B_1) + P(B_2) = 0.7^{10} + \binom{10}{1} 0.3 \cdot 0.7^9 + \binom{10}{2} 0.3^2 \cdot 0.7^8 \doteq \mathbf{0.3828}.$$