1. cvičení z PRA

19. - 23. února 2024

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 výsledku (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 konkretní 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

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

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

$$\begin{array}{rrrrrr} \{1,1\} & \{1,2\} & \{1,3\} & \{1,4\} & \{1,5\} & \{1,6\}, \\ & \{2,2\} & \{2,3\} & \{2,4\} & \{2,5\} & \{2,6\}, \\ & & \{3,3\} & \{3,4\} & \{3,5\} & \{3,6\}, \\ & & & \{4,4\} & \{4,5\} & \{4,6\}, \\ & & & & \{5,5\} & \{5,6\}, \\ & & & & & \{6,6\}, \end{array}$$

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

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

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 vracení (0.2344) i nevracení (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 $k << N$, tak pravděpodobnost, že bychom opakovaně vytáhli znovu tentýž předmět je zanedbatelná. Tudíž obě pravděpodobnosti se budou téměř shodovat.

Jak přirozeně definovat nezávislost jevů: Nejdříve si zavedeme podmíněnou pravděpodobnost $P(A|B)$, tj. pravděpodobnost, že nastane jev A za předpokladu, že výsledky se budou omezovat jen na jev B (také to můžeme chápat tak, že nastal jev B a my se *zpětně* ptáme, jaká byla za tohoto předpokladu pravděpodobnost jevu A). Přirozeně to bude $P(A|B) = \frac{P(A \cap B)}{P(B)}$, pokud $P(B) \neq 0$.

To, že jev A nebude záviset na jevu B, si pak přirozeně určíme podmínkou $P(A|B) = P(A)$ a podobně B nebude záviset na jevu A pokud $P(B) = P(B|A)$. Takže jevy A a B budou nezávislé, pokud platí podmínky $P(A|B) = P(A)$ a $P(B) = P(B|A)$ (a také bychom ještě mohli uvažovat i nezávislost na doplňcích $P(A) = P(A|B^c)$ atd.).

Jak je ale vidět, všechny tyto podmínky odpovídají jediné rovnici $P(A \cap B) = P(A) \cdot P(B)$, samozřejmě za předpokladu, že $P(A) \neq 0$ a $P(B) \neq 0$.

Proto se nezávislost jevů A a B definuje jako $P(A \cap B) = P(A) \cdot P(B)$ (ať už jsou $P(A)$ nebo $P(B)$ nulové nebo ne).

Podobným způsobem dojdeme k definici pro více jevů jako:

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.

Příklad 1.3 *Cyril bude hrát s Adamem a Bohdanem tenis. Bude hrát tři sety. Může si vybrat pořadí protihráčů Adam-Bohdan-Adam nebo Bohdan-Adam-Bohdan. Jestliže vyhraje dva sety po sobě, získá prémii. Jaké pořadí má Cyril zvolit, aby měl větší šanci získat prémii, jestliže Adam je lepší hráč než Bohdan?*

Řešení:

V tomto případě budeme posuzovat pravděpodobnosti ve dvou různých pravděpodobnostních prostorech daných výběrem pořadí protihráčů. Jde tedy vlastně o dva různé modely, ve kterých spočítáme pravděpodobnosti. Postup bude v obou případech stejný, pouze hodnoty pravděpodobností jednotlivých jevů se budou lišit.

Budeme tedy uvažovat jevy:

$$V_i = \text{“}\textit{výhra Cyrila v i-tém setu}\text{”} \ \text{ pro } \ i = 1, 2, 3$$

které budou nezávislé a dále jev

$$Z = \text{“}\textit{Cyril získá prémii}\text{”} \ .$$

Protože na prémii je potřeba vyhrát alespoň dva sety po sobě, tak dostaneme, že

$$Z = (V_1 \cap V_2 \cap V_3) \cup (V_1 \cap V_2 \cap V_3^c) \cup (V_1^c \cap V_2 \cap V_3)$$

kde v závorkách jsou evidentně navzájem disjunktní jevy. Takže z tohoto a z nezávislosti dostaneme

$$P(Z) = P(V_1 \cap V_2 \cap V_3) + P(V_1 \cap V_2 \cap V_3^c) + P(V_1^c \cap V_2 \cap V_3) =$$

$$= P(V_1) \cdot P(V_2) \cdot P(V_3) + P(V_1) \cdot P(V_2) \cdot P(V_3^c) + P(V_1^c) \cdot P(V_2) \cdot P(V_3)$$

Také to můžeme o něco jednodušeji dostat pomocí doplňku Z^c jako

$$Z^c = V_2^c \cup (V_1^c \cap V_2 \cap V_3^c)$$

což je opět sjednocení disjunktních jevů, takže

$$P(Z^c) = P(V_2^c) + P(V_1^c \cap V_2 \cap V_3^c) = P(V_2^c) + P(V_1^c) \cdot P(V_2) \cdot P(V_3^c)$$

Pravděpodobnost výhry nad Adamem v daném setu označíme jako $a(\neq 0)$ a nad Bohdanem v daném setu jako $b(\neq 0)$. Přitom máme, že $a < b$.

A teď v závislosti na zvoleném modelu dosadíme:

- pro Adam-Bohdan-Adam to bude $P(V_1) = a$, $P(V_2) = b$, $P(V_3) = a$:

$$P(Z) = 1 - P(Z^c) = 1 - P(V_2^c) - P(V_1^c) \cdot P(V_2) \cdot P(V_3^c) =$$

$$= 1 - (1 - b) - (1 - a)b(1 - a) = b(1 - (1 - a)^2) = ab(2 - a)$$

- pro Bohdan-Adam-Bohdan to bude $P(V_1) = b$, $P(V_2) = a$, $P(V_3) = b$, což znamená, že jen prohodíme a a b:

$$P(Z) = ba(2 - b)$$

Protože $a < b$, tak lépe vychází první model, kdy horší hráč je uprostřed.