Pravděpodobnost a statistika

Podmíněná pravděpodobnost a nezávislost jevů

Vilém Vychodil

KMI/PRAS, Přednáška 4

Vytvořeno v rámci projektu 2963/2011 FRVŠ

Přednáška 4: Přehled

1. Podmíněná pravděpodobnost:
    - definice a vlastnosti podmíněné pravděpodobnosti,
    - pravidlo o násobení pravděpodobností.
2. Bayesova věta:
    - věta o celkové pravděpodobnosti,
    - Bayesova věta,
    - apriorní a aposteriorní pravděpodobnost.
3. Nezávislost náhodných jevů:
    - nezávislost dvou náhodných jevů,
    - vzájemná nezávislost více náhodných jevů,
    - vlastnosti nezávislých náhodných jevů,
    - nezávislost opakovaných experimentů.

Příklad (Motivace pro podmíněnou pravděpodobnost)

Máme 20 počítačů se stejnou konfigurací a máme informaci, že 8 z nich běží na Linuxu, 12 na Free BSD. Dále víme, že na 13 z nich byl nainstalován Apache a na zbylých 7 byl nainstalován Lighttpd, přitom počty jsou dány tabulkou:

	Apache (A)	Lightppd (B)	celkem
Linux (C)	5	8	13
Free BSD (D)	3	4	7
celkem	8	12	20

- Jeden počítač je náhodně zvolen a zapnut. Pak $P(C) = \frac{13}{20} = 0.65$.
- Pokud (nějak) zjistíme, že daný počítač má nainstalován Apache, pak: $P(C|A) = \frac{5}{8} = 0.625$.

  $=$ pravděpodobnost výskytu jevu C za předpokladu, že nastal A .

Příklad (Motivace pro podmíněnou pravděpodobnost)

Počítače jsou vybírány se stejnou šancí výběru (klasický pravděpodobnostní prostor).

Důsledek:

$$P(C|A) = \frac{5}{8} = \frac{|C \cap A|}{|A|} = \frac{\frac{|C \cap A|}{|\Omega|}}{\frac{|A|}{|\Omega|}} = \boxed{\frac{P(C \cap A)}{P(A)}}.$$

- V některých případech se zajímáme o pravděpodobnosti jevů, za předpokladu, že $A \in \mathcal{F} \subseteq \Omega$ nastal;
- A lze chápat jako *nový prostor elementárních jevů*;
- pro daný náhodný jev B chceme zavést pravděpodobnost $P(B|A)$, která je spočtena pouze na základě elementárních jevů z A.

Příklad (Motivace pro podmíněnou pravděpodobnost)

Nechť $P\colon \mathcal{B} \cap 2^{[0,1]} \to \mathbb{R}$ je Lebesgueova pravděpodobnostní míra na σ-algebře Borelovských podmnožin reálného intervalu $[0,1]$.

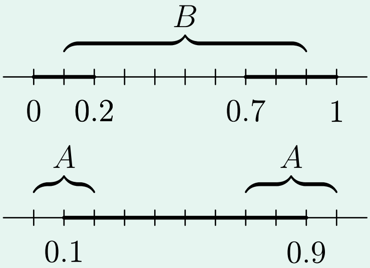
Připomeňme: P zobrazuje každý interval na jeho délku, tj. $P([a,b]) = b - a$.

Pro náhodné jevy $A = (0, 0.2] \cup (0.7, 1)$, $B = [0.1, 0.9)$ **platí:**

- $P(A) = (0.2 - 0) + (1 - 0.7) = 0.5$, $P(B) = 0.9 - 0.1 = 0.8$,
- $P(A \cap B) = P\big([0.1, 0.2] \cup (0.7, 0.9)\big) = (0.2 - 0.1) + (0.9 - 0.7) = 0.3$,

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{0.3}{0.5} = 0.6,$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.3}{0.8} = 0.375.$$



Podmíněná pravděpodobnost

Definice (Podmíněná pravděpodobnost, angl.: *conditional probability*)

Mějme pravděpodobnostní prostor $\langle\Omega, \mathcal{F}, P\rangle$ a náhodný jev $B \in \mathcal{F}$ takový, že $P(B) > 0$. Pak **podmíněná pravděpodobnost $P(A|B)$** výskytu náhodného jevu $A \in \mathcal{F}$ za podmínky, že nastal náhodný jev B, je definovaná vztahem

$$P(A|B) = \frac{P(A \cap B)}{P(B)},$$

Funkce $P(\cdot|B)\colon \mathcal{F} \to \mathbb{R}$ se nazývá **podmíněná pravděpodobnostní míra**.

Poznámky:

- $P(A|B)$ čteme „pravděpodobnost A za předpokladu (výskytu) B“;
- $P(\cdot|B)$ je funkce, která každému A přiřazuje $P(A|B) \in \mathbb{R}$;
- dále prokážeme, že $P(\cdot|B)$ je vskutku pravděpodobnostní míra.

Věta (Základní vlastnosti podmíněné pravděpodobnosti)

Mějme pravděpodobnostní prostor $\langle\Omega, \mathcal{F}, P\rangle$ *a* $B \in \mathcal{F}$ *takový, že* $P(B) > 0$. *Potom*

1. $P(A|B) \geq 0$ *pro každou* $A \in \mathcal{F}$;
2. $P(C|B) = 1$ *pro každou* $C \in \mathcal{F}$ *takovou, že* $B \subseteq C$;
3. *pokud* $A \cap B = \emptyset$, *pak* $P(A|B) = 0$; *speciálně:*
4. *pokud* $\{x\} \in \mathcal{F}$ *a* $x \notin B$, *pak* $P(\{x\}|B) = 0$.

Důkaz.

První tvrzení je zřejmé: $P(A|B) = \dfrac{P(A \cap B)}{P(B)} \geq 0$.

Mějme $B \subseteq C$, to jest $B \cap C = B$. Odtud $P(C|B) = \dfrac{P(C \cap B)}{P(B)} = \dfrac{P(B)}{P(B)} = 1$.

Pokud $A \cap B = \emptyset$, pak $P(A \cap B) = 0$, to jest $P(A|B) = \dfrac{P(A \cap B)}{P(B)} = 0$. $\square$

Věta (O podmíněné pravděpodobnosti)

Mějme pravděpodobnostní prostor $\langle \Omega, \mathcal{F}, P \rangle$ *a* $B \in \mathcal{F}$ *takový, že* $P(B) > 0$*. Potom funkce* $P(\cdot|B)$ *je pravděpodobnostní míra na* $\mathcal{F}$*.*

Důkaz.

Z předchozího dostáváme, že $P(\cdot|B)$ splňuje Kolmogorovy axiomy (P1) a (P2), protože $B \subseteq \Omega$, to jest $P(\Omega|B) = 1$. Stačí tedy ověřit σ-aditivitu (P3).

Pro každou spočetnou $\{A_i \in \mathcal{F} \,|\, i \in I\}$, kde $A_i \cap A_j = \emptyset$ $(i \neq j)$, platí

$$P\Big(\bigcup\nolimits_{i \in I} A_i \Big| B\Big) = \frac{P\Big(\Big(\bigcup_{i \in I} A_i\Big) \cap B\Big)}{P(B)} = \frac{P\Big(\bigcup_{i \in I} (A_i \cap B)\Big)}{P(B)}$$

$$= \frac{\sum_{i \in I} P(A_i \cap B)}{P(B)} = \sum\nolimits_{i \in I} \frac{P(A_i \cap B)}{P(B)} = \sum\nolimits_{i \in I} P(A_i|B),$$

to jest $P(\cdot|B)$ je σ-aditivní na $\mathcal{F}$. $\square$

Věta (O podmíněných pravděpodobnostních prostorech)

Mějme pravděpodobnostní prostor $\langle \Omega, \mathcal{F}, P \rangle$ *a* $B \in \mathcal{F}$ *takový, že* $P(B) > 0$. *Potom následující jsou pravděpodobnostní prostory:*

1. $\langle \Omega, \mathcal{F}, P(\cdot|B) \rangle$,
2. $\langle B, \mathcal{F} \cap 2^B, P'(\cdot|B) \rangle$, *kde* $P'(\cdot|B)$ *je funkce* $P(\cdot|B)$ *zúžená na* $\mathcal{F} \cap 2^B$.

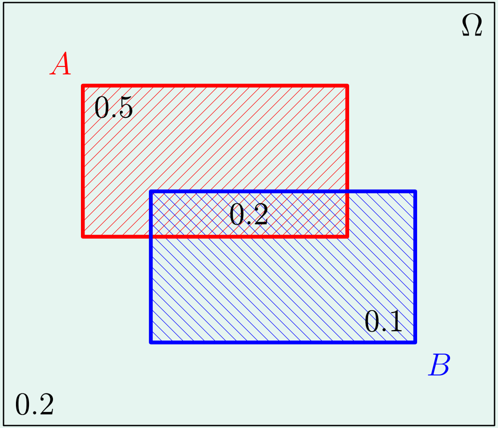
Důkaz.

První tvrzení plyne z předchozích dvou vět.

Druhé tvrzení je důsledkem prvního a následujících faktů: $B \neq \emptyset$, protože $P(B) > 0$. Dále $\mathcal{F} \cap 2^B$ je σ-algebra, protože se jedná o průnik dvou σ-algeber. Funkce $P'(\cdot|B) \colon \mathcal{F} \cap 2^B \to \mathbb{R}$ definovaná $P'(A|B) = P(A|B)$ pro každé $A \in \mathcal{F} \cap 2^B$ splňuje všechny axiomy (P1), (P2) i (P3), protože i výchozí $P(\cdot|B)$ je splňovala, viz předchozí věty (speciálně $P(B|B) = 1$). □

Příklad (Grafická ilustrace významu podmíněné pravděpodobnosti)

Mějme $P(A) = 0.7$, $P(B) = 0.3$ a $P(A \cap B) = 0.2$.



$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.2}{0.3} = \frac{2}{3},$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{0.2}{0.7} = \frac{2}{7},$$

$$P(A|B) \neq P(B|A).$$

Příklad (Součet 3 před součtem 5, *začátek*)

Vrháme párem čtyřstěnných kostek a počítáme součet teček. Kostkami házíme tak dlouho, dokud nepadne součet 3 nebo součet 5.

Úkol: *Jaká je pravděpodobnost, že součet 3 padne před součtem 5?*

Příklady (série hodů):

- $2+2=4$, $2+4=6$, $3+1=4$, $1+2=3$, (součet 3 padl před součtem 5).
- $2+2=4$, $2+4=6$, $3+1=4$, $2+3=5$, (součet 5 padl před součtem 3).

Řešení: Označme následující náhodné jevy:

- A ... náhodný jev „výsledek hodu je součet 3“ (možnosti: $1+2$ nebo $2+1$), $P(A) = \dfrac{2}{16}$;
- B ... náhodný jev „výsledek hodu je součet 3 nebo 5“ (možnosti: $1+2$, $2+1$, $1+4$, $2+3$, $3+2$ nebo $4+1$), $P(B) = \dfrac{6}{16}$.

Příklad (Součet 3 před součtem 5, *dokončení*)

Uvažujeme následující náhodné jevy:

- A . . . náhodný jev „výsledek hodu je součet 3“

  (možnosti: $1+2$ nebo $2+1$), $P(A) = \dfrac{2}{16}$;
- B . . . náhodný jev „výsledek hodu je součet 3 nebo 5“

  (možnosti: $1+2$, $2+1$, $1+4$, $2+3$, $3+2$ nebo $4+1$), $P(B) = \dfrac{6}{16}$.

Řešením je *podmíněná pravděpodobnost* $P(A|B)$:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)} = \frac{\frac{2}{16}}{\frac{6}{16}} = \frac{2}{6} = \frac{1}{3}.$$

Jediné elementární jevy, o které se zajímáme, jsou jevy vedoucí na součet 3 nebo 5.

Příklad (Případ heliových balonků)

Reklamní agent v supermarketu má 25 heliových balonků:
10 je žlutých, 8 je červených a 7 je zelených.

Balonky jsou postupně rozdávány kolemjdoucím.

Úkol: *Za předpokladu, že balonek daný prvnímu kolemjdoucímu je žlutý, jaká je pravděpodobnost že balonek daný druhému kolemjdoucímu je opět žlutý?*

Řešení:

- 24 balonků zbývá,
- 9 z nich je žlutých.

Výsledná podmíněná pravděpodobnost je $\boxed{\frac{9}{24} = 0.375}$.

Příklad (Modifikovaný případ heliových balonků)

Reklamní agent v supermarketu má 25 heliových balonků:
10 je žlutých, 8 je červených a 7 je zelených.

Balonky jsou postupně rozdávány kolemjdoucím.

Úkol: *Jaká je pravděpodobnost, že první dva rozdané balonky jsou žluté?*

Řešení: Hledáme $P(A \cap B)$ pro následující náhodné jevy:

- A ... první rozdaný balonek je žlutý,
- B ... druhý rozdaný balonek je žlutý.

Z předchozího příkladu: $P(B|A) = \dfrac{P(A \cap B)}{P(A)} = \dfrac{9}{24} = 0.375$, odtud:

$$P(A \cap B) = \frac{P(A) \cdot P(A \cap B)}{P(A)} = P(A) \cdot P(B|A) = P(A) \cdot \frac{9}{24} = \frac{10}{25} \cdot \frac{9}{24} = 0.15.$$

Pravidlo o násobení pravděpodobností

Ekvivalentním vyjádřením

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \qquad \text{pokud } P(B) > 0$$

dostáváme:

Důsledek zavedení podmíněné pravděpodobnosti

Mějme pravděpodobnostní prostor $\langle \Omega, \mathcal{F}, P \rangle$ a $A, B \in \mathcal{F}$. Pak platí:

$$P(A \cap B) = P(A) \cdot P(B|A), \qquad \text{pokud } P(A) > 0,$$

$$P(A \cap B) = P(B) \cdot P(A|B), \qquad \text{pokud } P(B) > 0.$$

Proč se používá:

- v řadě situací je jednodušší přiřadit $P(A)$ a $P(B|A)$ než $P(A \cap B)$;
- nezáleží na tom, který z A a B považujeme za první, a který za následující.

Příklad (Tahání lístků z urny)

Urna obsahuje sedm modrých lístků a tři červené lístky. Náhodně vytáhneme dva lístky bez toho, aniž bychom první vraceli zpět.

Úkol: *Vypočítejte pravděpodobnost, že v prvním tahu bude vybrán červený lístek a v druhém tahu modrý lístek.*

Řešení: Uvažujeme následující náhodné jevy a jejich pravděpodobnosti:

- A ... v prvním tahu je vybrán červený lístek, $P(A) = \frac{3}{10}$;
- B ... v druhém tahy je vybrán modrý lístek, $P(B|A) = \frac{7}{9}$.

Použitím pravidla o násobení pravděpodobností:

$$P(A \cap B) = P(A) \cdot P(B|A) = \frac{3}{10} \cdot \frac{7}{9} = \frac{7}{30}.$$

Příklad (Modifikace problému tahání lístků z urny)

Urna obsahuje sedm modrých lístků a tři červené lístky. Náhodně vytáhneme dva lístky bez toho, aniž bychom první vraceli zpět.

Úkol: *Vypočítejte pravděpodobnost, že první v obou dvou tazích budou vylosovány červené lístky.*

Řešení založené na podmíněné pravděpodobnosti: $\dfrac{3}{10} \cdot \dfrac{2}{9} = \dfrac{1}{15}$.

Kombinatorické řešení: $\dfrac{P_{2,3}}{P_{2,10}} = \dfrac{6}{90} = \dfrac{C_{2,3}}{C_{2,10}} = \dfrac{\binom{3}{2}}{\binom{10}{2}} = \dfrac{3}{45}$.

Obecně platí: $\boxed{\dfrac{P_{k,n_1}}{P_{k,n_2}} = \dfrac{C_{k,n_1}}{C_{k,n_2}}}$.

Příklad (Třetí křížová karta tažená šestá v pořadí)

Vybíráme karty ze standardního baíčku 52 karet.

Úkol: *Vypočtěte pravděpodobnost, že třetí vytažená křížová karta bude tažena jako šestá v celkovém pořadí.*

Řešení:

- A . . . právě dvě křížové karty byly vybrány mezi prvními pěti kartami

$$P(A) = \frac{\binom{13}{2} \cdot \binom{39}{3}}{\binom{52}{5}} = 0.274.$$

- B . . . křížová karta je vybrána jako šestá v pořadí; odtud $P(B|A) = \dfrac{11}{47} = 0.234.$
- Hledaná pravděpodobnost: $P(A \cap B) = P(A) \cdot P(B|A) = 0.064.$

Věta (Obecné pravidlo o násobení pravděpodobností)

Mějme pravděpodobnostní prostor $\langle \Omega, \mathcal{F}, P \rangle$ *a posloupnost náhodných jevů* $A_1, \ldots, A_n \in \mathcal{F}$ *takovou, že* $P\left(\bigcap_{j=1}^{i-1} A_j\right) > 0$ *pro každé* $i = 1, \ldots, n$*. Potom*

$$P\Big(\bigcap_{i=1}^{n} A_i\Big) = \prod_{i=1}^{n} P\Big(A_i \,\Big|\, \bigcap_{j=1}^{i-1} A_j\Big).$$

Důkaz.

Pro $n = 1$ je triviální. Indukcí přes n dostáváme

$$\begin{aligned} P\Big(\bigcap_{i=1}^{n} A_i\Big) &= P\Big(\bigcap_{i=1}^{n-1} A_i \cap A_n\Big) = P\Big(\bigcap_{i=1}^{n-1} A_i\Big) \cdot P\Big(A_n \,\Big|\, \bigcap_{i=1}^{n-1} A_i\Big) \\ &= \prod_{i=1}^{n-1} P\Big(A_i \,\Big|\, \bigcap_{j=1}^{i-1} A_j\Big) \cdot P\Big(A_n \,\Big|\, \bigcap_{i=1}^{n-1} A_i\Big) = \prod_{i=1}^{n} P\Big(A_i \,\Big|\, \bigcap_{j=1}^{i-1} A_j\Big), \end{aligned}$$

což je hledaná rovnost. $\square$

Příklad (Násobení pravděpodobností čtyř náhodných jevů)

Vybíráme čtyři karty ze standardního baíčku 52 karet.

Úkol: *Spočítejte pravděpodobnost, že dostaneme $\heartsuit$, $\clubsuit$, $\diamondsuit$ a $\spadesuit$ po sobě.*

Řešení:

$$P(A_\heartsuit) = \frac{13}{52}, \qquad P(A_\clubsuit | A_\heartsuit) = \frac{13}{51},$$

$$P(A_\diamondsuit | A_\heartsuit \cap A_\clubsuit) = \frac{13}{50}, \qquad P(A_\spadesuit | A_\heartsuit \cap A_\clubsuit \cap A_\diamondsuit) = \frac{13}{49}.$$

Použitím pravidla násobení pravděpodobností:

$$\begin{aligned} & P(A_\heartsuit \cap A_\clubsuit \cap A_\diamondsuit \cap A_\spadesuit) \\ & = P(A_\heartsuit) \cdot P(A_\clubsuit | A_\heartsuit) \cdot P(A_\diamondsuit | A_\heartsuit \cap A_\clubsuit) \cdot P(A_\spadesuit | A_\heartsuit \cap A_\clubsuit \cap A_\diamondsuit) \\ & = \frac{2\,197}{499\,800} \approx 0.004395758. \end{aligned}$$

Příklad (Problém s barevnými kuličkami, *začátek*)

Chlapec má kalhoty a v kapsách má barevné kuličky:

- levá kapsa: pět modrých kuliček a čtyři bílé kuličky,
- pravá kapsa: čtyři modré kuličky a pět bílých kuliček.

Za této situace chlapec přesune jednu kuličku z levé kapsy do pravé.

Úkol: *Jaká je pravděpodobnost, že potom vytáhne modrou kuličku z pravé kapsy?*

Řešení: Uvažujme následující náhodné jevy:

- M_L ... modrá kulička je přesunuta z levé kapsy do pravé,
- B_L ... bílá kulička je přesunuta z levé kapsy do pravé,
- M_R ... modrá kulička je vytažena z pravé kapsy (výsledný pokus).

Zřejmě platí:

- M_L a B_L tvoří *úplný systém neslučitelných jevů*, to jest
- $M_L \cap M_R$ a $B_L \cap M_R$ jsou *neslučitelné*.

Příklad (Problém s barevnými kuličkami, *dokončení*)

Počáteční stav: levá kapsa: $5\text{M} + 4\text{B}$; pravá kapsa: $4\text{M} + 5\text{B}$, dále:

- M_L ... modrá kulička je přesunuta z levé kapsy do pravé,
- B_L ... bílá kulička je přesunuta z levé kapsy do pravé,
- M_R ... modrá kulička je vytažena z pravé kapsy (výsledný pokus).

Jelikož M_L a B_L tvoří úplný systém jevů, platí:

$$P(M_R) = P(\Omega \cap M_R) = P((M_L \cup B_L) \cap M_R) = P((M_L \cap M_R) \cup (B_L \cap M_R)).$$

Dále, $M_L \cap M_R$ a $B_L \cap M_R$ jsou neslučitelné, to jest platí:

$$\begin{aligned} P(M_R) &= P((M_L \cap M_R) \cup (B_L \cap M_R)) = P(M_L \cap M_R) + P(B_L \cap M_R) \\ &= P(M_L) \cdot P(M_R|M_L) + P(B_L) \cdot P(M_R|B_L) \\ &= \frac{5}{9} \cdot \frac{5}{10} + \frac{4}{9} \cdot \frac{4}{10} = \frac{41}{90}. \end{aligned}$$

Věta o celkové pravděpodobnosti

Zobecněním principu z predchozího příkladu dostáváme:

Věta (Věta o celkové pravděpodobnosti)

Mějme pravděpodobnostní prostor $\langle \Omega, \mathcal{F}, P \rangle$ *a nejvýš spočetnou* $\{A_i \in \mathcal{F} \,|\, i \in I\}$ *tvořící úplný systém neslučitelných jevů tak, že* $P(A_i) > 0$ *pro každé* $i \in I$*. Pak*

$$P(B) = \sum\nolimits_{i \in I} P(A_i \cap B) = \sum\nolimits_{i \in I} P(A_i) \cdot P(B|A_i)$$

platí pro libovolný $B \in \mathcal{F}$*.*

Důkaz.

Z faktu, že $\{A_i \in \mathcal{F} \,|\, i \in I\}$ je úplný systém jevů platí, že $B = \bigcup_{i \in I}(A_i \cap B)$. Z neslučitelnosti jevů v $\{A_i \in \mathcal{F} \,|\, i \in I\}$ navíc plyne, že i všechny $A_i \cap B$ $(i \in I)$ jsou neslučitelné. Tvrzení je proto přímý důsledek σ-aditivity míry P a pravidla o násobení pravděpodobností. $\square$

Příklad (Použití věty o celkové pravděpodobnosti)

Mějme tři urny obsahující barevné lístky:

- A_1: dva modré a čtyři bílé lístky;
- A_2: jeden modrý a dva bílé lístky; a
- A_3: pět modrých a čtyři bílé lístky.

Předpokládejme následující *pravděpodobnosti zvolení urny*:

$$P(A_1) = \frac{1}{3}, \qquad P(A_2) = \frac{1}{6}, \qquad P(A_3) = \frac{1}{2}.$$

Otázka: *Jaká je pravděpodobnost, že náhodně vytažený lístek z urny, která je před tím náhodně zvolena podle P bude modrý?*

Řešení:

$$P(M) = \sum_{i=1}^{3} P(A_i) \cdot P(M|A_i) = \frac{1}{3} \cdot \frac{2}{6} + \frac{1}{6} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{5}{9} = \frac{4}{9}.$$

Příklad (Motivace pro Bayesovu větu)

Vyjdeme z předchozího příkladu a změníme zadání otázky.

- A_1: dva modré a čtyři bílé lístky; $P(A_1) = \frac{1}{3}$,
- A_2: jeden modrý a dva bílé lístky; $P(A_2) = \frac{1}{6}$,
- A_3: pět modrých a čtyři bílé lístky; $P(A_3) = \frac{1}{2}$.

Otázka: *Předpokládejme, že výsledkem výběru je modrý lístek. Jaké jsou pravděpodobnosti, že byl vytažen z urny A_1, A_2 a A_3?*

Řešení: Užitím věty o celkové pravděpodobnosti:

$$P(A_1|M) = \frac{P(A_1 \cap M)}{P(M)} = \frac{P(A_1)P(M|A_1)}{P(A_1)P(M|A_1) + P(A_2)P(M|A_2) + P(A_3)P(M|A_3)}$$

$$= \frac{\frac{1}{3} \cdot \frac{2}{6}}{\frac{1}{3} \cdot \frac{2}{6} + \frac{1}{6} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{5}{9}} = \frac{2}{8} = \frac{1}{4}. \qquad = \text{pravděp. } A_1 \text{ za předpokladu } M$$

Příklad (Apriorní a aposteriorní pravděpodobnosti)

Problém tahání lístků ze tří uren:

- A_1: dva modré a čtyři bílé lístky; $P(A_1) = \frac{1}{3}$,
- A_2: jeden modrý a dva bílé lístky; $P(A_2) = \frac{1}{6}$,
- A_3: pět modrých a čtyři bílé lístky; $P(A_3) = \frac{1}{2}$.

Apriorní pravděpodobnosti: $P(A_1) = \dfrac{1}{3}$, $P(A_2) = \dfrac{1}{6}$, $P(A_3) = \dfrac{1}{2}$.

Aposteriorní pravděpodobnosti: $P(A_1|M) = \dfrac{1}{4}$, $P(A_2|M) = \dfrac{1}{8}$, $P(A_3|M) = \dfrac{5}{8}$.

Intuitivní interpretace:

- aposteriorní pravděpodobnost = aktualizovaná pravděpodobnost (aktualizace je provedena na základě znalosti výskytu určitého jevu);
- v tomto případě: výskyt M zvyšuje pravděpodobnost zvolení A_3.

Bayesova věta

Zobecněním principu z predchozího příkladu dostáváme:

Věta (Bayesova věta)

Mějme pravděpodobnostní prostor $\langle\Omega, \mathcal{F}, P\rangle$ *a nejvýše spočetnou* $\{A_i \in \mathcal{F} \mid i \in I\}$ *tvořící úplný systém neslučitelných jevů tak, že* $P(A_i) > 0$ *pro každé* $i \in I$*. Pak*

$$P(A_i|B) = \frac{P(A_i) \cdot P(B|A_i)}{\sum_{j\in I} P(A_j) \cdot P(B|A_j)}$$

platí pro libovolné $i \in I$ *a* $B \in \mathcal{F}$ *takové, že* $P(B) > 0$.

Důkaz.

Užitím věty o celkové pravděpodobnosti a pravidla násobení pravděpodobností:

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i \cap B)}{\sum_{j\in I} P(B \cap A_j)} = \frac{P(A_i) \cdot P(B|A_i)}{\sum_{j\in I} P(A_j) \cdot P(B|A_j)}. \qquad \square$$

Příklad (Analýza pravděpodobnosti výroby vadných součástek)

Výrobce má tři stroje vyrábějící tytéž součástky. Některé vyráběné součástky jsou vadné: první stroj vyrobí 2% vadných součástek ze své celkové produkce; druhý a třetí stroj 1% a 3%. Celková produkce součástek je rozdělena mezi stroje následovně: 35% (první stroj), 25% (druhý stroj) a 40% (třetí stroj).

Úkol: *Jaká je pravděpodobnost, že vadný výrobek byl vyroben na třetím stroji?*

Řešení: Pro náhodné jevy A (výrobek byl vyroben na 1. stroji), B (výrobek byl vyroben na 2. stroji), C (výrobek byl vyroben na 3. stroji) a D (výrobek je vadný):

$$\begin{aligned} P(D) &= P(A) \cdot P(D|A) + P(B) \cdot P(D|B) + P(C) \cdot P(D|C) \\ &= \frac{35}{100} \cdot \frac{2}{100} + \frac{25}{100} \cdot \frac{1}{100} + \frac{40}{100} \cdot \frac{3}{100} = \frac{215}{10\,000}. \end{aligned}$$

$$P(C|D) = \frac{P(C) \cdot P(D|C)}{P(D)} = \frac{\frac{40}{100} \cdot \frac{3}{100}}{\frac{215}{10\,000}} = \frac{120}{215} \approx 0.558\,.$$

Příklad (Analýza účinnosti tuberkulínového testu, *začátek*)

Test kožní tuberkulinové přecitlivělosti:

- provádí se intrakutánní injekcí 2 TU (jednotek) čistěného tuberkulinu PPD;
- výsledek se odčítá po 24–48 hodinách.

Statistická fakta (ověřená na základě dlouholeté zkušenosti):

- $8\,\%$ TBC pozitivních lidí má negativní test („falešně negativní“);
- $4\,\%$ TBC negativních lidí mají pozitivní test („falešně pozitivní“).

Úkol: *Osoba se podrobí testu a test vyjde pozitivní. Jaká je pravděpodobnost, že osoba je TBC pozitivní?*

Řešení:

Zajímáme se o pravděpodobnost $P(T|P)$, to jest pravděpodobnost, že osoba je TBC pozitivní (T) za předpokladu, že test vyjde pozitivně (P).

Poznámka: T a T' tvoří úplný systém neslučitelných jevů.

Příklad (Analýza účinnosti tuberkulínového testu, *pokračovnání*)

Z předchozích dat získáváme:

- $P(P|T) = 1 - 0.08 = 0.92$ (pravděp. TBC pozitivní má pozitivní test)
- $P(P|T') = 0.04$ (pravděp. TBC negativní má pozitivní test).

Užitím Bayesovy věty:

$$P(T|P) = \frac{P(T) \cdot P(P|T)}{P(T) \cdot P(P|T) + P(T') \cdot P(P|T')} = \frac{P(T) \cdot 0.92}{P(T) \cdot 0.92 + (1 - P(T)) \cdot 0.04}.$$

Musíme dodat (neznámou) hodnotu $P(T)$, to jest pravděpodobnost, že náhodně vybraná osoba z populace je TBC pozitivní (pro USA je $P(T) = 0.0075$).

$$P(T|P) = \frac{0.0075 \cdot 0.92}{0.0075 \cdot 0.92 + 0.9925 \cdot 0.04} \approx 0.148\,.$$

Překvapivě nízká hodnota.

Příklad (Analýza účinnosti tuberkulínového testu, *dokončení*)

Jak interpretovat hodnotu $P(T|P) = 0.148$?

- Pokud je náš test pozitivní, máme (pouze) $3:17$ šanci, že jsme TBC pozitivní.

Význam testu:

- Test je levný (významný ekonomický faktor).
- Test má spolehlivě prokazovat, že osoba *není TBC pozitivní* (nikoliv opak), pro tento účel je test *velmi efektivní*:

$$P(T'|N) = \frac{P(T') \cdot P(N|T')}{P(T') \cdot P(N|T') + P(T) \cdot P(N|T)} = \frac{0.9925 \cdot 0.96}{0.9925 \cdot 0.96 + 0.0075 \cdot 0.08}.$$

  $P(T'|N) \approx 0.9994$ (pravděp. negativně testovaná osoba je TBC negativní),
  $P(T|N) \approx 0.0006$ (pravděp. negativně testovaná osoba je TBC pozitivní).

- V případě pozitivního testu lze použít efektivnější (a dražší) metodu.

Příklad (Informatická aplikace: Bayesovský filtr na spam)

Paul Graham (*A Plan for Spam*), `http://www.paulgraham.com/spam.html`.

Základní princip:

$Spam$ je jev „dopis je spam“; $Spam'$ je jev „dopis není spam“ (tak zvaný Ham):

$$P(Spam|W) = \frac{P(Spam) \cdot P(W|Spam)}{P(Spam) \cdot P(W|Spam) + P(Spam') \cdot P(W|Spam')}.$$

Vysvětlivky:

$P(Spam)$ apriorní pravděpodobnost, že příchozí dopis je spam ($P(Spam) = 0.5$);

$P(W|Spam)$ je pravděpodobnost, že spam obsahuje slovo W;

$P(W|Spam')$ je pravděpodobnost, že nespam obsahuje slovo W.

Poznámky:

- $P(W|Spam)$ a $P(W|Spam')$ se stanovují z uložené pošty a spamu;
- $P(Spam)$ je globálně nastavená (0.5 až 0.8).

Příklad (Motivace pro závislost a nezávislost náhodných jevů)

Hodíme dvakrát mincí a pozorujeme výsledky hodů: $\Omega = \{HH, HT, TH, TT\}$; $\mathcal{F} = 2^{\Omega}$. Každý z elementárních jevů nastává se stejnou pravděpodobností 0.25.

Uvažujme následující náhodné jevy:

- $A = \{HH, HT\}$ ... výsledkem prvního hodu je hlava, $P(A) = 0.5$;
- $B = \{HT, TT\}$ ... výsledkem druhého hodu je orel, $P(B) = 0.5$;
- $C = \{TT\}$ ... výsledkem obou hodů je orel, $P(C) = 0.25$.

Platí:

$$P(B|C) = \frac{P(B \cap C)}{P(C)} = \frac{P(C)}{P(C)} = 1, \qquad P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{0.25}{0.5} = 0.5.$$

Pozorování:

- výskyt C **ovlivnil** pravděpodobnost výskytu B: $P(B) = 0.5 \neq 1 = P(B|C)$.
- výskyt A **neovlivnil** pravděpodobnost výskytu B: $P(B) = 0.5 = P(B|A)$.

Ekvivalentní vyjádření podmínek nezávislosti

Poznámka:

- budeme se snažit zavést definici nezávislosti náhodných jevů,
- nezávislost ve smyslu „pravděpodobnosti jejich výskytu“,
- existuje několik (užitečných) ekvivalentních formulací.

Ekvivalence podmínek pro nezávislost bude patrná z následující věty:

Věta

Mějme pravděpodobnostní prostor $\langle \Omega, \mathcal{F}, P \rangle$ a náhodné jevy $A, B \in \mathcal{F}$. Následující podmínky jsou ekvivalentní:

1. $P(A \cap B) = P(A) \cdot P(B)$,
2. *pokud* $P(A) > 0$, *pak* $P(B|A) = P(B)$,
3. *pokud* $P(B) > 0$, *pak* $P(A|B) = P(A)$.

Důkaz.

„(1) ⇒ (2)“: Předpokládejme, že $P(A \cap B) = P(A) \cdot P(B)$. Pokud $P(A) > 0$, pak

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A) \cdot P(B)}{P(A)} = P(B).$$

„(2) ⇒ (3)“: Předpokládejme, že pokud $P(A) > 0$, pak $P(B|A) = P(B)$. Dále předpokládejme, že $P(B) > 0$. Vyšetříme dva případy. Buď $P(A) > 0$ a tedy $P(B|A) = P(B)$ a můžeme psát:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \cdot P(B|A)}{P(B)} = \frac{P(A) \cdot P(B)}{P(B)} = P(A);$$

nebo $P(A) = 0$, ale potom z monotonie dostáváme:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \leq \frac{P(A)}{P(B)} = \frac{0}{P(B)} = 0 = P(A).$$

„(3) ⇒ (1)“: Pokud $P(B) > 0$, pak přímo $P(A \cap B) = P(A) \cdot P(B)$.
Pokud $P(B) = 0$, pak z monotonie dostáváme $P(A \cap B) = 0 = P(A) \cdot P(B)$. $\square$

Nezávislé náhodné jevy

Nezávislost *dvou náhodných jevů* formazilujeme následovně:

Definice (Nezávislé náhodné jevy)

Mějme pravděpodobnostní prostor $\langle \Omega, \mathcal{F}, P \rangle$. Náhodné jevy $A, B \in \mathcal{F}$ nazveme **nezávislé** (angl.: *independent*), pokud

$$P(A \cap B) = P(A) \cdot P(B)$$

V opačném případě nazýváme A a B **závislé** (angl.: *dependent*).

Terminologie: Existuje několik různých pojmů pro totéž, například

- statisticky nezávislé jevy, angl.: *statistically independent events*,
- stochasticky nezávislé jevy, angl.: *stochastically independent events*,
- jevy nezávislé ve smyslu pravděpodobností výskytu, ...

Příklad (Nezávislé náhodné jevy)

Házíme červenou a bílou šestistrannou kostkou. Uvažujeme následující náhodné jevy:

- $A \ldots 4$ padne na červené kostce,
- $B \ldots$ součet teček na obou kostkách je lichý.

Máme 36 možných elementárních jevů (se stejnou pravděpodobností výskytu).

- $P(A) = \frac{6}{36} = \frac{1}{6}$
- $P(B) = \frac{18}{36} = \frac{1}{2}$ (protože $2 + 4 + 6 + 4 + 2 = 18$)
- $P(A \cap B) = \frac{3}{36} = \frac{1}{12}$ (elementární jevy $\langle 4, 1\rangle$, $\langle 4, 3\rangle$ a $\langle 4, 5\rangle$)

$P(A) \cdot P(B) = \frac{1}{6} \cdot \frac{1}{2} = \frac{1}{12} = P(A \cap B)$ **A a B jsou nezávislé**

Příklad (Závislé náhodné jevy)

Házíme červenou a bílou šestistrannou kostkou. Uvažujeme následující náhodné jevy:

- C ... 5 padne na červené kostce,
- D ... součet teček na obou kostkách je 11.

Máme 36 možných elementárních jevů (se stejnou pravděpodobností výskytu).

- $P(C) = \frac{6}{36} = \frac{1}{6}$
- $P(D) = \frac{2}{36} = \frac{1}{18}$ (elementární jevy $\langle 5, 6\rangle$ a $\langle 6, 5\rangle$)
- $P(C \cap D) = \frac{1}{36}$ (elementární jev $\langle 5, 6\rangle$)

$P(C) \cdot P(D) = \frac{1}{6} \cdot \frac{1}{18} = \frac{1}{108} \neq \frac{1}{36} = P(C \cap D)$ **C a D jsou závislé**

Vlastnosti nezávislých jevů

Nezávislost jevů se přenáší i na jejich doplňky:

Věta (Vlastnosti nezávislých jevů)

Mějme pravděpodobnostní prostor $\langle \Omega, \mathcal{F}, P \rangle$ *a libovolné náhodné jevy* $A, B \in \mathcal{F}$*. Pokud jsou* A *a* B *nezávislé, pak jsou i následující dvojice náhodných jevů nezávislé:*

1. A a B',
2. A' a B,
3. A' a B'.

Poznámka: Druhé a třetí tvrzení plyne z prvního, protože:

- druhé tvrzení dostáváme z prvního záměnou A a B;
- pokud jsou A a B nezávislé, pak i A a B' jsou nezávislé (z prvního tvrzení) a tedy i A' a B' jsou nezávislé (z druhého tvrzení).

Důkaz.

Předpokládejme, že A a B jsou nezávislé jevy. Pokud $P(A) = 0$, pak ihned dostáváme $P(A \cap B') \leq P(A) = 0$, to jest $P(A \cap B') = 0$. Jako důsledek máme

$$P(A \cap B') = 0 = 0 \cdot P(B') = P(A) \cdot P(B').$$

Tvrzení tedy stačí prokázat pro A splňující $P(A) > 0$. Jelikož $\langle \Omega, \mathcal{F}, P(\cdot|A) \rangle$ je pravděpodobnostní prostor, platí $P(B'|A) = 1 - P(B|A)$.

Odtud dostáváme

$$\begin{aligned} P(A \cap B') &= P(A) \cdot P(B'|A) = P(A) \cdot (1 - P(B|A)) \\ &= P(A) - P(A) \cdot P(B|A) = P(A) - P(A \cap B) \\ &= P(A) - P(A) \cdot P(B) = P(A) \cdot (1 - P(B)) = P(A) \cdot P(B'). \end{aligned}$$

To jest A a B' jsou nezávislé náhodné jevy. Druhé a třetí tvrzení plyne z prvního. $\square$

Věta (Triviálně nezávislé náhodné jevy)

Mějme pravděpodobnostní prostor $\langle \Omega, \mathcal{F}, P \rangle$ *a libovolný náhodný jev* $A \in \mathcal{F}$. *Pak následující dvojice náhodných jevů jsou nezávislé:*

1. A *a* $\emptyset$,
2. A *a* Ω.

Důkaz.

Platí $A \cap \emptyset = \emptyset$, to jest

$$P(A \cap \emptyset) = P(\emptyset) = 0 = P(A) \cdot 0 = P(A) \cdot P(\emptyset).$$

Dále platí $A \cap \Omega = A$, to jest

$$P(A \cap \Omega) = P(A) = P(A) \cdot 1 = P(A) \cdot P(\Omega).$$

$\square$

Věta (Kritérium nezávislosti komplementárních jevů)

Mějme pravděpodobnostní prostor $\langle \Omega, \mathcal{F}, P \rangle$. *Pak pro každý náhodný jev* $A \in \mathcal{F}$ *platí, že* A *a* A' *jsou nezávislé, právě když* $P(A) \in \{0, 1\}$.

Důkaz.

Podle definice nezávislosti jevů, A a A' jsou nezávislé, právě když $P(A \cap A') = P(A) \cdot P(A')$, to jest, právě když $0 = P(\emptyset) = P(A) \cdot P(A')$. Poslední rovnost platí, právě když $0 = P(A) \cdot (1 - P(A))$, což znamená, právě když $0 = P(A) - P(A)^2$, to jest, právě když $P(A) = P(A)^2$. Jelikož $P(A) > P(A)^2$ pokud $0 < P(A) < 1$, poslední rovnost platí, právě když $P(A) \in \{0, 1\}$. □

Poznámka:

- Náhodné jevy s pravděpodobností výskytu jinou než 0 a 1 a jejich komplementy jsou *vždy závislé.*
- Neformálně: A a A' jsou závislé (pro $0 < P(A) < 1$) protože se vylučují.

Příklad (Motivace pro nezávislost tří náhodných jevů)

Krabice obsahuje čtyři míčky číslované 1, 2, 3 a 4. Jeden z míčků je náhodně vybrán z krabice. Uvažujme náhodné jevy: $A = \{1, 2\}$ (vybrán míček 1 a/nebo 2), $B = \{1, 3\}$ (vybrán míček 1 a/nebo 3), $C = \{1, 4\}$ (vybrán míček 1 a/nebo 4) a předpokládejme $P(A) = P(B) = P(C) = 0.5$.

Dostáváme:

$$\begin{aligned} P(A \cap B) &= P(\{1\}) = 0.25 = P(A) \cdot P(B), \\ P(A \cap C) &= P(\{1\}) = 0.25 = P(A) \cdot P(C), \\ P(B \cap C) &= P(\{1\}) = 0.25 = P(B) \cdot P(C). \end{aligned}$$

Každé dva různé náhodné jevy z A, B, a C jsou **nezávislé.**

$$P(A \cap B \cap C) = P(\{1\}) = 0.25 \neq 0.125 = 0.25^3 = P(A) \cdot P(B) \cdot P(C).$$

Pokud A a B nastanou současně:

$$P(C|A \cap B) = \frac{P((A \cap B) \cap C)}{P(A \cap B)} = \frac{0.25}{0.25} = 1 \neq P(C).$$

Náhodné jevy $A \cap B$ a C jsou **závislé.**

Vzájemně nezávislé náhodné jevy

Definice (Vzájemně nezávislé náhodné jevy)

Mějme pravděpodobnostní prostor $\langle\Omega, \mathcal{F}, P\rangle$. Náhodné jevy $A_1, A_2, \ldots, A_n \in \mathcal{F}$ nazýváme **vzájemně nezávislé** pokud pro každou neprázdnou podmnožinu $I \subseteq \{1, 2, \ldots, n\}$ platí

$$P\left(\bigcap\nolimits_{i \in I} A_i\right) = \prod\nolimits_{i \in I} P(A_i).$$

Jako speciální případ pro $n = 3$ dostáváme:

Příklad (Kritérium vzájemné nezávislosti tří náhodných jevů)

Náhodné jevy A, B a C jsou vzájemně nezávislé, když

- $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$ a
- $P(A \cap B) = P(A) \cdot P(B)$, $P(A \cap C) = P(A) \cdot P(C)$, $P(B \cap C) = P(B) \cdot P(C)$.

Poznámka: Počet podmínek nutných otestovat roste exponenciálně!

Příklad (Nezávislost redundantních komponent systému)

Server má tři redundantní napájecí zdroje. Pokud zdroj Z_1 selže, je nahrazen zdrojem Z_2. Pokud zdroj Z_2 selže, je nahrazen zdrojem Z_3. Předpokládejme, že pravděpodobnost selhání každého ze zdrojů (během určité doby provozu) je 0.15 a předpokládejme, že možnosti selhání jednotlivých zdrojů jsou vzájemně nezávislé.

Úkol: *Vypočítejte pravděpodobnost, že server neselže kvůli zdrojům.*

Řešení: A_i … náhodný jev: „i-tý zdroj selže"

$$
\begin{aligned}
P((A_1 \cap A_2 \cap A_3)') &= 1 - P(A_1 \cap A_2 \cap A_3) \\
&= 1 - (P(A_1) \cdot P(A_2) \cdot P(A_3)) = 1 - (0.15)^3 = 0.996625.
\end{aligned}
$$

Ačkoliv se pravděpodobnost selhání jednoho zdroje může zdát velká, pravděpodobnost selhání všech zdrojů zároveň je zanedbatelně malá.

Princip: Přidání redundantních komponent $\Longmapsto$ zvýšení spolehlivosti.

Věta (Vlastnosti tří vzájemně nezávislých jevů)

Mějme pravděpodobnostní prostor $\langle \Omega, \mathcal{F}, P \rangle$ *a uvažujme tři náhodné jevy* $A, B, C \in \mathcal{F}$, *které jsou vzájemně nezávislé. Pak následující dvojice náhodných jevů jsou nezávislé:*

1. A *a* $(B \cap C)$,
2. A *a* $(B \cup C)$,
3. A' *a* $(B \cap C)$.

Navíc platí, že A', B' *a* C' *jsou vzájemně nezávislé.*

Důkaz (začátek).

Předpokládejme, že A, B a C jsou vzájemně nezávislé.

První tvrzení je zřejmé:

$$\begin{aligned} P(A \cap (B \cap C)) = P(A \cap B \cap C) &= P(A) \cdot P(B) \cdot P(C) \\ &= P(A) \cdot (P(B) \cdot P(C)) = P(A) \cdot P(B \cap C). \end{aligned}$$

Důkaz (pokračování).

Druhé tvrzení plyne z následujícího:

$$\begin{aligned}
P(A \cap (B \cup C)) &= P((A \cap B) \cup (A \cap C)) \\
&= P(A \cap B) + P(A \cap C) - P((A \cap B) \cap (A \cap C)) \\
&= P(A \cap B) + P(A \cap C) - P(A \cap B \cap C) \\
&= P(A \cap B) + P(A \cap C) - (P(A) \cdot P(B) \cdot P(C)) \\
&= P(A) \cdot P(B) + P(A) \cdot P(C) - (P(A) \cdot P(B) \cdot P(C)) \\
&= P(A) \cdot (P(B) + P(C) - (P(B) \cdot P(C))) \\
&= P(A) \cdot (P(B) + P(C) - P(B \cap C)) = P(A) \cdot P(B \cup C).
\end{aligned}$$

Třetí tvrzení je důsledkem:

$$\begin{aligned}
&P(A' \cap (B \cap C)) = P((B \cap C) - A) = P(B \cap C) - P((B \cap C) \cap A) \\
&= P(B \cap C) - P(A \cap B \cap C) = P(B \cap C) - (P(A) \cdot P(B) \cdot P(C)) \\
&= P(B \cap C) - (P(A) \cdot P(B \cap C)) = P(B \cap C) \cdot (1 - P(A)) = P(A') \cdot P(B \cap C).
\end{aligned}$$

Důkaz (dokončení).

Zbývá prokázat poslední tvrzení. Z De Morganových zákomů platí, že

$$P(A' \cap B' \cap C') = P(A' \cap (B' \cap C')) = P(A' \cap (B \cup C)').$$

Jelikož A, B a C jsou vzájemně nezávislé, jedním z předchozích tvrzení dostáváme, že A a $B \cup C$ jsou nezávislé. Dále dostáváme, že A' a $(B \cup C)'$ jsou nezávislé (viz předchozí větu o nezávislosti dvou náhodných jevů). Dohromady z toho plyne

$$\begin{aligned} P(A' \cap B' \cap C') &= P(A' \cap (B \cup C)') \\ &= P(A') \cdot P((B \cup C)') = P(A') \cdot P(B' \cap C') \end{aligned}$$

Z nezávislosti náhodných jevů B a C plyne i nezávilsosti B' a C', odtud $P(B' \cap C') = P(B') \cdot P(C')$. Z toho dostáváme

$$P(A' \cap B' \cap C') = P(A') \cdot P(B' \cap C') = P(A') \cdot P(B') \cdot P(C').$$

Zbytek plyne z věty o nezávislosti dvou náhodných jevů. $\square$

Příklad (Série hodů šestistrannou kostkou, Přednáška 3)

Šestistraná kostka je šestkrát vržena. Pokud na i-tý pokus hodíme i, pak nazeveme výsledek shodou. Náhodný pokus považujeme za úspěch, pokud padne během šesti hodů alespoň jedna shoda.

Úkol: *Vypočítejte pravděpodobnost, že experiment skončí úspěchem.*

Řešení: B_i ... náhodný jev: „shoda na i-tý pokus".

Zřejmě $P(B_i) = \frac{1}{6}$, to jest $P(B_i') = 1 - P(B_i) = \frac{5}{6}$.

Označme A náhodný jev: „nastala aspoň jedna shoda".
Potom A' je náhodný jev: „nenastala ani jedna shoda" a platí:

$$\begin{aligned} P(A) &= 1 - P(A') = 1 - P(B_1' \cap B_2' \cap \cdots \cap B_6') \\ &= 1 - \left(\frac{5}{6} \cdot \frac{5}{6} \cdot \cdots \cdot \frac{5}{6}\right) = 1 - \frac{15625}{46656} = \frac{31031}{46656} \approx 0.665102. \end{aligned}$$

Nezávislost opakovaných experimentů

Složený náhodný pokus:

- náhodný pokus sestávající ze sekvence n náhodných pokusů (stejného typu);
- jednotlivé pokusy jsou (obvykle) vzájemně nezávislé
  (výsledek jednoho pokusu nemá vliv na výsledek dalšího pokusu).

Důsledek: $\boxed{P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1) \cdot P(A_2) \cdot \cdots \cdot P(A_n)}$.

Poznámka:

Prostor elementárních jevů experimentu složeného z n nezávislých pokusů téhož typu je množina Ω^n, kde Ω je prostor elementárních jevů jednotlivého pokusu.

Příklad (Posloupnosti nezávislých náhodných pokusů)

Hrací kostka je vržena třikrát: $\Omega^3 = \{\langle x_1, x_2, x_3 \rangle \mid x_1, x_2, x_3 \in \{1, 2, 3, 4, 5, 6\}\}$.
Mince je hozena čtyřikrát: $\Omega^4 = \{\langle x_1, x_2, x_3, x_4 \rangle \mid x_1, x_2, x_3, x_4 \in \{H, T\}\}, \ldots$

Příklad (Nezávislost a závilost opakovaných experimentů)

Krabice obsahuje tři červené míčky (A), dva bílé míčky (B) a čtyři žluté míčky (C).

Z krabice jsou taženy tři míčky. Máme dva možné způsoby losování:

Míček je po vytažení vrácen $\Longmapsto$ **nezávislé náhodné pokusy**, například:

$$P(\{\langle A, B, C\rangle\}) = \frac{3}{9} \cdot \frac{2}{9} \cdot \frac{4}{9} = \frac{8}{243},$$

$$P(\{\langle C, C, A\rangle, \langle A, B, B\rangle\}) = \frac{4}{9} \cdot \frac{4}{9} \cdot \frac{3}{9} + \frac{3}{9} \cdot \frac{2}{9} \cdot \frac{2}{9} = \frac{20}{243}.$$

Míček není po vytažení vrácen $\Longmapsto$ **závislé náhodné pokusy**, například:

$$P(\{\langle A, B, C\rangle\}) = \frac{3}{9} \cdot \frac{2}{8} \cdot \frac{4}{7} = \frac{1}{21} > \frac{8}{243},$$

$$P(\{\langle C, C, A\rangle, \langle A, B, B\rangle\}) = \frac{4}{9} \cdot \frac{4}{8} \cdot \frac{3}{7} + \frac{3}{9} \cdot \frac{2}{8} \cdot \frac{2}{7} = \frac{7}{84} > \frac{20}{243}.$$

Příklad (Pravděpodobnost výhry ve stírací loterii)

Předpokládejme, že koupíme pět stíracích losů po sobě. Dále předpokládejme, že při každém zakoupení losu máme pravděpodobnost výhry 0.2.

Za předpokladu nezávislosti jednotlivých pokusů,

$$P(\{\langle V, V, N, N, N\rangle\}) = \left(\frac{1}{5}\right)^2 \cdot \left(\frac{4}{5}\right)^3 = 0.04 \cdot 0.512,$$

$$P(\{\langle N, V, N, V, N\rangle\}) = \left(\frac{1}{5}\right)^2 \cdot \left(\frac{4}{5}\right)^3 = 0.04 \cdot 0.512, \ \ldots$$

Obecně platí, že pravděpodobnost zakoupení dvou výherních losů a tří nevýherních losů je rovna $0.04 \cdot 0.512$ násobena počtem možností, jako koupit dva výherní losy během pěti různých pokusů. To jest:

$$\binom{5}{2} \cdot \left(\frac{1}{5}\right)^2 \cdot \left(\frac{4}{5}\right)^3 = \frac{5!}{2! \cdot 3!} \cdot \left(\frac{1}{5}\right)^2 \cdot \left(\frac{4}{5}\right)^3 = 10 \cdot 0.04 \cdot 0.512 = 0.2048.$$

Příklad (Problém Montyho Halla)

- Tři zavřené garaže obsahují dvě kozy a jedno auto (každá garáž obsahuje právě jednu položku).
- Můžeme si vzít auto, pokud uhádneme, ve které je garáži.

Průběh:

- Na počátku si zvolíme garáž.
- Monty Hall, který ví, kde je auto, otevře některou garáž, ve které je koza, ale ne tu, kterou jsme si na začátku zvolili.
- Potom máme ještě jednu šanci opravit naši volbu:
  - buď budeme trvat na naší výchozí volbě, nebo
  - zvolíme druhou dosud zavřenou garáž.

Otázka: *Která strategie je lepší?*

Příklad (Řešení problému Montyho Halla)

První strategie: Pokud trváme na výchozí volbě, pravděpodobnost úspěchu je $\frac{1}{3}$.

Druhá strategie: Zvolíme si druhou (zavřenou) garáž.

1. **Postup 1:** Náhodné jevy A (výhra při první strategii) a B (výhra při druhé strategii) tvoří *úplný systém neslučitelných jevů*, z $P(A \cup B) = P(A) + P(B)$ tedy dostáváme $1 = P(A) + P(B)$, to jest
$$P(B) = 1 - P(A) = 1 - \frac{1}{3} = \frac{2}{3}.$$

2. **Postup 2:** Pro náhodné jevy B (výhra při druhé strategii) a C (na počátku ukazujeme na garáž s autem) můžeme s využitím věty o celkové pravděpodobnosti psát
$$P(B) = P(C) \cdot P(B|C) + P(C') \cdot P(B|C') = \frac{1}{3} \cdot 0 + \frac{2}{3} \cdot 1 = \frac{2}{3}.$$

Závěr: *Druhá strategie je lepší, protože zvyšuje naši šanci na zisk auta.*

Přednáška 4: Závěr

Pojmy:

- podmíněné pravděpodobnost, pravidlo o násobení pravděpodobností
- věta o celkové pravděpodobnosti, Bayesova věta
- apriorní a aposteriorní pravděpodobnost
- nezávislost dvou jevů, vzájemná nezávislost více náhodných jevů

Použité zdroje:

- Capinski M., Zastawniak T. J.: *Probability Through Problems* Springer 2001, ISBN 978–0–387–95063–1.
- Hogg R. V., Tanis E. A.: *Probability and Statistical Inference* Prentice Hall; 7. vydání 2005, ISBN 978–0–13–146413–1.
- Johnson J. L.: *Probability and Statistics for Computer Science* Wiley-Interscience 2008, ISBN 978–0–470–38342–1.