

Podmíněná pravděpodobnost

... vyjadřuje změnu pravděpodobnosti, pokud se vyskytne nějaká nová informace, která zúží počet možností.

Například, pravděpodobnost, že mezi 12 kartami je srdcová hláška (král a svršek) je $\binom{30}{10} \binom{32}{12}^{-1} = \frac{30! 20! 12!}{20! 10! 32!} = \frac{12 \cdot 11}{32 \cdot 31} = \frac{33}{248} \doteq 13\%$

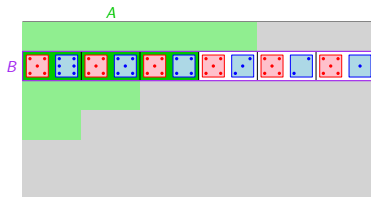
Pokud je mezi prvními pěti kartami srdcový král, pravděpodobnost se zvýší téměř na dvojnásobek $\binom{26}{6} \binom{27}{7}^{-1} = \frac{26! 20! 7!}{20! 6! 27!} = \frac{7}{27} \doteq 26\%$

Definice: *Podmíněná pravděpodobnost* jevu A za podmínky B , kde $P[B] \neq 0$, je pravděpodobnost $P[A|B] = \frac{P[A \cap B]}{P[B]}$.

Ukázka:

Pravděpodobnost hodu alespoň 9 (jev A) za podmínky, že na červené kostce je 5 (jev B) je

$$P[A|B] = \frac{P[A \cap B]}{P[B]} = \frac{\frac{3}{36}}{\frac{6}{36}} = \frac{1}{2}.$$



Bayesův vzorec

Věta o úplné pravděpodobnosti:

Jsou-li $B_1, \dots, B_k$ disjunktní jevy pokrývající celé Ω , pak platí

$$P[A] = P[A|B_1]P[B_1] + P[A|B_2]P[B_2] + \dots + P[A|B_k]P[B_k]$$

... jev A lze pokrýt průniky

$A \cap B_1, \dots, A \cap B_k$, přičemž

$$P[A \cap B_i] = P[A|B_i]P[B_i], \text{ i}$$

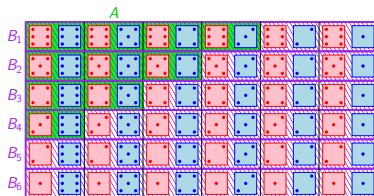
$$P[A \cap B_j] = P[B_j|A]P[A]$$

Ukázka:

$$P[A] = \frac{2}{3} \cdot \frac{1}{6} + \frac{1}{2} \cdot \frac{1}{6} + \frac{1}{3} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{6} + 0 \cdot \frac{1}{6} + 0 \cdot \frac{1}{6} = \frac{5}{18}$$

Důsledek (Bayesův vzorec): Za stejných podmínek platí i

$$P[B_j|A] = \frac{P[A \cap B_j]}{P[A]} = \frac{P[A|B_j]P[B_j]}{P[A]} = \frac{P[A|B_j]P[B_j]}{\sum_{i=1}^k P[A|B_i]P[B_i]}$$



Použití Bayesova vzorce

... víme, jak se nějaký test (jev A) chová na ověřených případech (pravděpodobnosti $P[A|B_i]$) a chceme usoudit, který případ (jev B_i) nejspíše nastal, když se test A zdaří (čili $P[B_i|A]$).

Aplikace ve zdravotnictví, testování kvality, umělé inteligenci apod.

Ukázka: Zkoumaná choroba postihuje 0,5 % populace (*prevalence*).
Její test ukáže pozitivní výsledky u 99,9 % nemocných pacientů (*senzitivita*), a negativní u 99 % zdravých (*specifita*).*

S jakou pravděpodobností je pacient **nemocný**, má-li test **pozitivní**?

S jakou pravděpodobností je pacient **zdravý**, má-li test **negativní**?

Jevy: A — pacienti s pozitivním testem, B — nemocní pacienti,

$$\begin{aligned} P[A] &= P[A|B]P[B] + P[A|\bar{B}]P[\bar{B}] \\ &= 0,014\,945 \end{aligned}$$

$$P[B|A] = \frac{P[A|B]P[B]}{P[A]} \doteq 0,334\,225$$

$$P[\bar{B}|\bar{A}] = \frac{P[\bar{A}|\bar{B}]P[\bar{B}]}{P[\bar{A}]} \doteq 0,999\,995$$

B_i	$P[B_i]$	$P[A B_i]$	$P[\bar{A} B_i]$
B	0,005	0,999	0,001
$\bar{B}$	0,995	0,01	0,99

* Štěpán, Zvára: Pravděpodobnost a matematická statistika, Příklad 2.7

Použití Bayesova vzorce

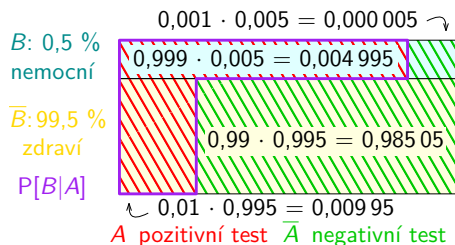
... víme, jak se nějaký test (jev A) chová na ověřených případech (pravděpodobnosti $P[A|B_i]$) a chceme usoudit, který případ (jev B_i) nejspíše nastal, když se test A zdaří (čili $P[B_i|A]$).

Aplikace ve zdravotnictví, testování kvality, umělé inteligenci apod.

Ukázka: Zkoumaná choroba postihuje 0,5 % populace (*prevalence*).
Její test ukáže pozitivní výsledky u 99,9 % nemocných pacientů (*senzitivita*), a negativní u 99 % zdravých (*specifita*)*.

S jakou pravděpodobností je pacient **nemocný**, má-li test **pozitivní**?

S jakou pravděpodobností je pacient **zdravý**, má-li test **negativní**?



Pozitivní výsledek testu jen málo odpovídá správné diagnóze $P[B|A] \doteq 33\%$ protože jevy $A \cap B$ a $A \cap \bar{B}$ jsou podobně pravděpodobné. Ovšem pro vyvrácení diagnózy je test dostatečně spolehlivý.

* Štěpán, Zvára: Pravděpodobnost a matematická statistika, Příklad 2.7

Použití Bayesova vzorce

... pro případ více než jen dvou navzájem doplňkových jevů B a $\bar{B}$

Ukázka: Říká se: „Večerní červánky věští pěkný den.“ Dejme tomu, že polovina dní je slunečných a červánky jim předcházejí v 10 %.

Dvě pětiny jsou zatažené, jim předcházejí červánky v 5 % případů.

Zbylé dny jsou deštivé, a po červáncích nastanou jen v 1 %.

Do jaké míry by šlo pranostice věřit? (Data nejsou skutečná.*)

Jevy: A — červánky, B_1, B_2, B_3 — slunečno, zataženo, deštivo.

Pravděpodobnosti: $P[A] = \frac{1}{2} \cdot \frac{1}{10} + \frac{2}{5} \cdot \frac{1}{20} + \frac{1}{10} \cdot \frac{1}{100} = \frac{71}{1000}$

	$P[B_i]$	$P[A B_i]$	$P[B_i A]$
B_1	$\frac{1}{2}$	$\frac{1}{10}$	$\frac{1}{2} \cdot \frac{1}{10} \cdot \frac{1000}{71} = \frac{50}{71} \doteq 71 \%$
B_2	$\frac{2}{5}$	$\frac{1}{20}$	$\frac{2}{5} \cdot \frac{1}{20} \cdot \frac{1000}{71} = \frac{20}{71} \doteq 28 \%$
B_3	$\frac{1}{10}$	$\frac{1}{100}$	$\frac{1}{10} \cdot \frac{1}{100} \cdot \frac{1000}{71} = \frac{1}{71} \doteq 1 \%$

Odpověď: Asi ve $\frac{3}{4}$ případů by po červáncích mělo být slunečno.

* Červánky coby subjektivní úkaz nepatří mezi sledované meteorologické jevy.

Nezávislé jevy

... výsledek jednoho jevu neovlivní pravděpodobnosti ostatních.

Definice: Jevy A a B jsou *nezávislé*, pokud $P[A \cap B] = P[A]P[B]$.

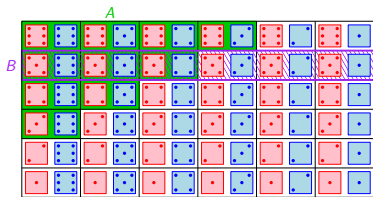
... nezávislé A a B splňují: $P[A|B] = \frac{P[A \cap B]}{P[B]} = \frac{P[A]P[B]}{P[B]} = P[A]$.

Obecně, $A_1, A_2, \dots, A_k$ jsou nezávislé, když pro každou neprázdnou $I \subseteq \{1, \dots, k\}$ platí: $P\left[\bigcap_{i \in I} A_i\right] = \prod_{i \in I} P[A_i]$.

Ukázka: Závislé jevy:

A „součet je alespoň 9“ a

B „na první kostce padla 5“.

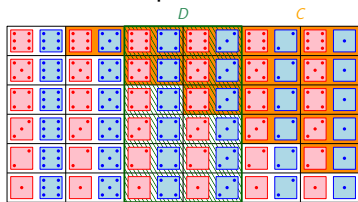


$$P[A \cap B] = \frac{3}{36}$$

$$\neq \frac{5}{108} = \frac{5}{18} \cdot \frac{1}{6} = P[A]P[B]$$

Nezávislé jevy: C „na červené padlo víc než na modré“ a

D „na modré padlo 3 nebo 4“.



$$P[C \cap D] = \frac{5}{36}$$

$$= \frac{5}{12} \cdot \frac{1}{3} = P[C]P[D]$$

Nezávislé jevy

... výsledek jednoho jevu neovlivní pravděpodobnosti ostatních.

Definice: Jevy A a B jsou *nezávislé*, pokud $P[A \cap B] = P[A]P[B]$.

... nezávislé A a B splňují: $P[A|B] = \frac{P[A \cap B]}{P[B]} = \frac{P[A]P[B]}{P[B]} = P[A]$.

Obecně, $A_1, A_2, \dots, A_k$ jsou nezávislé, když pro každou neprázdnou $I \subseteq \{1, \dots, k\}$ platí: $P\left[\bigcap_{i \in I} A_i\right] = \prod_{i \in I} P[A_i]$.

Ukázka: Jev

E „na červené padlo 3 nebo 4“
je nezávislý na C i na D .

$$P[C \cap E] = \frac{5}{36} = P[C]P[E]$$

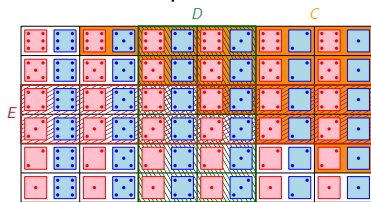
$$P[D \cap E] = \frac{1}{9} = P[D]P[E]$$

Pozor! Nezávislost dvojic
nezaručuje nezávislost všech.

Všechny tři nezávislé nejsou:

$$\begin{aligned} P[C \cap D \cap E] &= \frac{1}{36} \neq \frac{5}{108} \\ &= \frac{5}{12} \cdot \frac{1}{3} \cdot \frac{1}{3} = P[C]P[D]P[E] \end{aligned}$$

Nezávislé jevy: C „na červené padlo víc než na modré“ a
 D „na modré padlo 3 nebo 4“.



$$\begin{aligned} P[C \cap D] &= \frac{5}{36} \\ &= \frac{5}{12} \cdot \frac{1}{3} = P[C]P[D] \end{aligned}$$

Kvíz — řešení

Je-li u některých otázek více možností správných, vyberte všechny.

1. Pravděpodobnost, že při třech hodech spravedlivou mincí padla alespoň dvakrát panna, za podmínky, že v prvních dvou hodech padl nejvýše jeden orel, je:
a) 0, b) $\frac{1}{6}$, c) $\frac{1}{4}$, d) $\frac{3}{8}$, e) $\frac{1}{3}$, f) $\frac{1}{2}$, g) $\frac{2}{3}$, h) $\frac{5}{8}$, i) $\frac{3}{4}$, j) $\frac{5}{6}$, k) 1.
2. Pokud $P[A|B] = P[B|A]$, potom platí také: a) $P[A] = 1$,
b) $P[A] = P[B]$, c) $P[B] = 1$, d) ani jedno z uvedených.
3. Pravda nebo lež? Pokud je průnik jevů A , B a C prázdný, ale každá dvojice má neprázdný průnik, pak $P[A|B \cap C]$ je nulová.
4. Pravda nebo lež?
Hodíme-li pětkrát mincí, pak jevy „padl lichý počet orlů“ a „v obou prvních dvou hodech padla panna“ jsou nezávislé.
5. Jsou-li jevy A a B závislé, pak platí
a) $P[A] > P[A|B]$, b) $P[A] \neq P[A|B]$, c) $P[A] < P[A|B]$.

Otázky k porozumění tématu přednášky

- ▶ Co by se stalo, kdyby ve větě o úplné pravděpodobnosti nebyly jevy $B_1, \dots, B_k$ disjunktní, ale přesto pokrývaly Ω ?
- ▶ Je nějaká souvislost mezi množinovou inkluzí $A \subseteq B$ a (ne)závislostí těchto jevů?
- ▶ Dostali bychom ekvivalentní definici nezávislosti, pokud bychom podmínku omezili jen na množinu $\{1, \dots, k\}$ a vynechali všechny její vlastní podmnožiny $I \subsetneq \{1, \dots, k\}$?

Poznámky k pojmosloví a značení

U podmíněné pravděpodobnosti se jevy $B_1, \dots, B_k$, které se navzájem vylučují a vyčerpávají všechny možnosti, nazývají *hypotézy* a pravděpodobnostem $P[B_i]$ se říká *apriorní*.
Pravděpodobnosti, které berou v úvahu, že nastal jev A , čili hledané $P[B_i|A]$, se nazývají *aposteriorní*.