Lineární modely - 2. přednáška

Konečná pravděpodobnost

Ondřej Klíma

28. 2. 2013

Osnova přednášky

- Klasická pravděpodobnost
- Geometrická pravděpodobnost
- Podmíněná pravděpodobnost
- Princip inkluze a exkluze

Pozn.: Dlouhé slovo „pravděpodobnost“ budeme občas nahrazovat zkratkou „prst“.

Házení kostkami

Příklad (2 poctivé kostky)

Házíme červenou a modrou poctivou kostkou. Jaká je pravděpodobnost, že padne součet 6?

Pravděpodobnost je podíl počtu příznivých možností ku počtu všech možností. Máme 36 možností (uspořádané dvojice).
Příznivých je 5: $(1,5),(2,4),(3,3),(4,2),(5,1)$.
Pravděpodobnost je tedy $\frac{5}{36}$, tj. cca $13,9\%$.

součet	2	3	4	5	6	7	8	9	10	11	12
prst $\frac{1}{36}\cdot$	1	2	3	4	5	6	5	4	3	2	1

Další příklady:
Pravděpodobnost, že padne sudý součet, je $\frac{1}{2}$.
Pravděpodobnost, že na červené kostce padne 1, je $\frac{1}{6}$.
Prst, že padne sudý součet a na červené kostce 1, je $\frac{3}{36} = \frac{1}{12}$.

Hádání čísla

Příklad (Hádání čísla)

Myslím si číslo od 1 do 50. Každý student z 50 si tipne, jaké je to číslo. Jaká je pravděpodobnost, že se někdo trefí?

Pravděpodobnost, že se trefí jeden konkrétní student, je $\frac{1}{50}$.

Pravděpodobnost, že se tento student netrefí, je $\frac{49}{50}$.

Prst, že se netrefí padesát studentů, je $(\frac{49}{50})^{50} \sim 0,364$.

Prst, že se trefí aspoň jeden student, je $1 - (\frac{49}{50})^{50}$.

Odpověď je:

$$\text{cca} \quad 63,6\% \,.$$

Kontrolní otázka: Jak se změní odpověď, pokud může každý student volit 2 čísla?

Další příklad:

Prst, že studenti zvolí různá čísla je $\frac{50!}{50^{50}} \sim 3,42 \cdot 10^{-21}$.

Volba předsedy

Příklad (Volba předsedy)

Pětice delegátů volí ze tří kandidátů A, B a C předsedu strany. Volba probíhá tak, že každý napíše na lístek jedno jméno a kadidát je zvolen pokud se objeví aspoň na 3 lístcích. Jaká je pravděpodobnost, že bude zvolen některý kandidát?

Vyřešte samostatně.

$$\frac{153}{243} \sim 63\%$$

Příklad (Sportka - minulá přednáška)

Jaká je pravděpodobnost výhry prvního pořadí ve sportce? (Na tiketu tipujeme pouze jednu šestici.)

Musí se uhodnout všech 6 čísel. Počet všech možností $\binom{49}{6} = 13\,983\,816$. Prst je $\frac{1}{13\,983\,816} \sim 0{,}000\,007\,151\%$.

Nejjednodušší model pravděpodobnosti

Ω množina všech možných výsledků, tzv. základní prostor.

V předchozím $\Omega = \{\omega_1, \ldots, \omega_n\}$ byla vždy konečná množina.

Prvky ω_i jsou tzv. možné výsledky.

Každá podmnožina $A \subseteq \Omega$ představuje možný jev.

Přičemž pravděpodobnost jevu A je

$$P(A) = \frac{|A|}{|\Omega|} \,.$$

Vlastnosti pravděpodobnosti

Takto definovaná skalární funkce $P : \mathcal{P}(\Omega) \to \mathbb{R}$ má vlastnosti:

- je nezáporná, tj. $P(A) \geq 0$ pro všechny jevy A,
- je aditivní, tj. $P(A \cup B) = P(A) + P(B)$, kdykoliv je $A \cap B = \emptyset$, (additivnost platí pro jakýkoliv konečný počet po dvou disjunktních jevů $A_i \subseteq \Omega$, $i \in I$),
- $P(\Omega) = 1$,
- pro všechny jevy A platí $P(A^c) = 1 - P(A)$.

Složitější model pravděpodobnosti

Základní prostor Ω nemusí být konečný.
Místo všech podmnožin $A \subseteq \Omega$ uvažujeme jen některé. Tzn. ne všechny „podmnožiny“ mohou nastat. Hovoříme pak o jevovém poli $\mathcal{A}$. Zde každý jev $A \in \mathcal{A}$ je podmnožina Ω. Požadujeme přitom:

- $\Omega \in \mathcal{A}$, tzn. základní prostor je jevem,
- pokud $A, B \in \mathcal{A}$, pak $A \cup B \in \mathcal{A}$, $A \cap B \in \mathcal{A}$, $A^c = \Omega \setminus A \in \mathcal{A}$.

Tzv. Booleova algebra.

Terminologie

- Celý základní prostor Ω se nazývá jistý jev.
- Prázdná podmnožina $\emptyset = \Omega^c \in \mathcal{A}$ se nazývá nemožný jev.
- Jednoprvkové podmnožiny $\{\omega\} \in \Omega$ se nazývají elementární jevy.
- Společné nastoupení jevů A_i, $i \in I$, odpovídá jevu $\cap_{i \in I} A_i$.
- Nastoupení alespoň jednoho z jevů A_i, $i \in I$, odpovídá jevu $\cup_{i \in I} A_i$.
- Jevy $A, B \in \mathcal{A}$ jsou neslučitelné, je-li $A \cap B = \emptyset$.
- Jev A má za důsledek jev B, jestliže $A \subseteq B$.

Pravděpodobnostní funkce

Mějme základní prostor Ω a konečné jevové pole $\mathcal{A}$.
Pravděpodobnostní funkce $P : \mathcal{A} \to \mathbb{R}$ je funkce, která má následující vlastnosti:

- nezápornost, tj. $P(A) \geq 0$ pro všechny jevy A,
- aditivnost pro jakýkoliv konečný počet neslučitelných jevů $A_i \subseteq \Omega$, $i \in I$, tj.

$$P(\cup_{i \in I} A_i) = \sum_{i \in I} P(A_i), \text{ kdykoliv je } A_i \cap A_j = \emptyset,\ i \neq j,\ i, j \in I.$$

- $P(\Omega) = 1$.

Falešné kostky

Příklad (Falešné kostky)

Mějme dvě falešné kostky, na kterých padá číslo 6 s větší pravděpodobností než číslo 1. Přesněji $p(6) = \frac{1}{4}$, $p(5) = p(4) = p(3) = p(2) = \frac{1}{6}$, $p(1) = \frac{1}{12}$. Jaká je pravděpodobnost, že padne součet 6?

$P(1,5) = P(5,1) = \frac{1}{12} \cdot \frac{1}{6}$,

$P(2,4) = P(3,3) = P(4,2) = \frac{1}{6} \cdot \frac{1}{6}$,

$P(suma\ 6) = \frac{1}{72}(1+2+2+2+1) = \frac{8}{72} = \frac{1}{9}$.

Pozn.: Fromálně správně P({(1,5)}).

součet	2	3	4	5	6	7	8	9	10	11	12
prst $\frac{1}{144}\cdot$	1	4	8	12	16	22	24	20	16	12	9

Další příklady:

Pravděpodobnost, že padne sudý součet, je $\frac{74}{144}$.

Pravděpodobnost, že na první kostce padne 1, je $\frac{1}{12}$.

Pravděpodobnost, že padne sudý součet a na první 1, je $\frac{5}{144}$.

Nekonečný základní prostor - příklad

Příklad (Náhodné číslo)

Volíme náhodně přirozené číslo, přičemž pravděpodobnost, že číslo bude větší nebo rovno 1 000 000 je $10^{-6} = 0.000\,001$ a ostatní čísla mají stejnou pravděpodobnost, tj. také 10^{-6}. Jaká je pravděpodobnost, že dvě náhodně vybraná čísla jsou menší než 1 000 000 a budou mít stejný počet cifer?

$\Omega = \mathbb{N}$. Jevy z $\mathcal{A}$ jsou konečná sjednocení elementárních jevů $\{n\}$, kde $n < 1\,000\,000$, a $\{x \in \mathbb{N} \mid x \geq 1\,000\,000\}$.

Vyřešte samostatně.

Ještě obecněji nemusí být jevové pole $\mathcal{A}$ konečné. Pak ale musíme mít další rozumné předpoklady ohledně nekonečně mnoha disjunktních jevů a sčítání jejich pravděpodobností. Předně musíme umět sčítat nekonečné součty.

Geometrická pravděpodobnost – příklad

Příklad (Setkání v menze)

Petr s Pavlem se dohodli, že se potkají v menze v době 12.00 až 14.30. Oba obědvají 30 minut a přijdou náhodně v intervalu 12.00 až 14.00. Jaká je pravděpodobnost, že se potkají?

- Základní prostor $\Omega = (12, 14) \times (12, 14)$ je čtverec v rovině – prvky jsou uspořádané dvojice časů příchodů.
- Jev, že se Petr s Pavlem potkají, je reprezentován jistým útvarem A ve čtverci Ω, těch bodů (x, y), že $|x - y| \leq 0.5$.
- Podíl obsahů $\frac{\operatorname{vol} A}{\operatorname{vol} \Omega}$ pak dává $P(A) = \frac{7}{16}$.

Další příklady:

Prst, že Petr přijde v době 12.00-13.00 a potkají se, je $\frac{1}{2} \cdot \frac{7}{16}$.

Prst setkání za předpokladu, že Petr přijde 12.00-13.00, je $\frac{7}{16}$.

Prst setkání za předpokladu, že oba přijdou 12.00-13.00, je $\frac{3}{4}$.

Geometrická pravděpodobnost

- $\Omega \subseteq \mathbb{R}^2$ se známým obsahem.
- Jevové pole $\mathcal{A}$ je vhodný systém podmnožin, u kterých umíme určit jejich obsah.
- Pravděpodobnostní funkce $P : \mathcal{A} \to \mathbb{R}$ je dána vztahem

$$P(A) = \frac{\operatorname{vol} A}{\operatorname{vol} \Omega}.$$

Příklad (Lámání úsečky)

Je dána tyč délky 2, na které náhodně zvolíme dva body v nichž tyč přeřízneme. Jaká je pravděpodobnost, že tři vzniklé úsečky tvoří strany trojúhelníku?

Vyřešte samostatně.

Podmíněná pravděpodobnost a nezávislé jevy

Definice

Nechť H je jev s nenulovou pravděpodobností. Podmíněná pravděpodobnost jevu A za předpokladu, že nastane jev H je definována vztahem

$$P(A \mid H) = \frac{P(A \cap H)}{P(H)} .$$

Pokud $P(A \mid H) = P(A)$, říkáme, že jev A je nezávislý na H.

Vztah $P(A) = P(A \mid H)$ znamená $P(A \cap H) = P(A) \cdot P(H)$ a tudíž i $P(H) = P(H \mid A)$ (za předpokladu $P(A) \neq 0$). Častěji proto říkáme, že dvojice jevů A a H je (stochasticky) nezávislá.

Pozn.: Ještě lépe je nezávislost definovat přímo formulí $P(A \cap H) = P(A) \cdot P(H)$.

Pozn.: Všimněte si, že dva neslučitelné jevy nemohou být nezávislé.

Podmíněná pravděpodobnost – revize příkladů

Příklad (Setkání v menze – nezávislost jevů)

Jev A … potkají se … $P(A) = \frac{7}{16}$.
Jev B … Petr přijde před jednou hodinou … $P(B) = \frac{1}{2}$.
$P(A \cap B) = \frac{1}{2} \cdot \frac{7}{16} = P(B) \cdot P(A)$, jevy A a B jsou nezávislé.
Jinak též $P(A \mid B) = \frac{7}{16} = P(A)$.
Jev C …oba přijdou před jednou hodinou …$P(C) = \frac{1}{4}$.
$P(A \mid C) = \frac{3}{4} \neq P(A)$, jevy A a C nejsou nezávislé.

Kontrolní otázka: Je jev A nezávislý na jevu „Petr přijde v době 12.00–12.30“?

Podmíněná pravděpodobnost – revize příkladů - II

Příklad (Hádání čísla – nezávislost jevů)

Jev A ... někdo se trefí ... $P(A) \sim 0,636$.
Jev B ... studenti zvolí různá čísla ... $P(B) \sim 3,42 \cdot 10^{-21}$.
$P(A \mid B) = 1 \neq P(A)$, jevy A a B nejsou nezávislé.

Příklad (Poctivé a falešné kostky)

Jev padnutí sudého součtu a jev padnutí 1 na první kostce, jsou jevy stochasticky nezávislé pro poctivou kostku, ale nejsou stochasticky nezávislé pro falešnou kostku.

Stochasticky nezávislé jevy

Definice

Jevy $A_1, \dots, A_k$ jsou stochasticky nezávislé, jestliže pro libovolné z nich vybrané jevy $A_{i_1}, \dots, A_{i_\ell}$, $1 \le \ell \le k$ platí $P(A_{i_1} \cap \dots \cap A_{i_\ell}) = P(A_{i_1}) \cdot \ldots \cdot P(A_{i_\ell})$.

Příklad (Dvanáctistěnná kostka)

Základní prostor $\Omega = \{1, 2, \dots, 12\}$.
Jev $A = \{2, 4, 6, 8, 10, 12\}$ – padne sudá – $P(A) = \frac{1}{2}$.
Jev $B = \{3, 6, 9, 12\}$ – padne číslo dělitelné 3 – $P(B) = \frac{1}{3}$.
$A \cap B = \{6, 12\}$; $P(A \cap B) = \frac{1}{6} = P(A) \cdot P(B)$.
Jev $C = \{1, 2, 3, 4, 8, 9\}$; $P(C) = \frac{1}{2}$.
$A \cap C = \{2, 4, 8\}$; $P(A \cap C) = \frac{3}{12} = \frac{1}{4} = P(A) \cdot P(C)$.
$B \cap C = \{3, 9\}$; $P(B \cap C) = \frac{2}{12} = \frac{1}{6} = P(B) \cdot P(C)$.
$A \cap B \cap C = \emptyset$; $P(A \cap B \cap C) = 0 \neq \frac{1}{12} = \frac{1}{2}\frac{1}{3}\frac{1}{2} = P(A)P(B)P(C)$.
Jevy A, B a C jsou po dvou nezávislé, ale nejsou nezávislé.

Stochasticky nezávislé jevy – vlastnosti

- Podsystém nezávislých jevů je opět nezávislý.
- Pro dva stochasticky nezávislé jevy A, B platí

$$P(A \cap B^c) = P(A \setminus B) = P(A) - P(A \cap B)$$

$$= P(A)(1 - P(B)) = P(A)P(B^c).$$

- Odtud lze odvodit, postupnou volbou $A = A_{i_1} \cap \cdots \cap A_{i_{\ell-1}}$, že záměnou jednoho (nebo více) nezávislých jevů za jejich opačné jevy obdržíme opět nezávislé jevy.

Sjednocení nezávislých jevů

Často se hledá pravděpodobnost, že nastane alespoň jeden ze stochasticky nezávislých jevů, tzn. hledáme $P(A_1 \cup \cdots \cup A_k)$.

Protože

$$A_1 \cup \cdots \cup A_k = (A_1^c \cap \cdots \cap A_k^c)^c,$$

dostáváme

$$P(A_1 \cup \cdots \cup A_k) = 1 - P(A_1^c \cap \cdots \cap A_k^c) = 1 - (1 - P(A_1)) \ldots (1 - P(A_k)) .$$

Pozn.: Vzpomeňme příklad Hádání čísla.

Pravděpodobnost sjednocení jevů

- Stochasticky nezávislé – předchozí strana.
- Neslučitelné jevy
  $P(A_1 \cup A_2 \cup \cdots \cup A_k) = P(A_1) + P(A_2) + \cdots + P(A_k)$.
- Obecně – složitější, ale zvládnutelné.
  - $A \cup B = (A \setminus B) \cup (A \cap B) \cup (B \setminus A)$,
    kde $P(A \setminus B) = P(A) - P(A \cap B)$.
    Odtud $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
  - $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$.

Princip inkluze a exkluze.

Příklad (Dvanáctistěnná kostka – sjednocení)

$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C) = \frac{1}{2} + \frac{1}{3} + \frac{1}{2} - \frac{1}{6} - \frac{1}{4} - \frac{1}{6} + 0 = \frac{1}{12}(6 + 4 + 6 - 2 - 3 - 2) = \frac{9}{12} = \frac{3}{4}$

Zde $A \cup B \cup C = \{1, 2, 3, 4, 6, 8, 9, 10, 12\}$.

Příklad na princip inkluze a exkluze

Příklad (Rozesílání dopisů)

Máme poslat čtyři dopisy na čtyři adresáty. Napíšeme dopisy, vložíme do obálky a zalepíme. Nyní nevíme, na kterou obálku která adresa patří, a proto napíšeme adresy náhodně. Jaká je pravděpodobnost, že nikdo nedostane svůj dopis?

- Jev A_i ... i-tý dopis bude v i-té obálce.
- Ω ... všechna pořadí čísel $1, 2, 3, 4$. Tedy $|\Omega| = 4! = 24$
- Chceme určit pravděpodobnost jevu $(A_1 \cup A_2 \cup A_3 \cup A_4)^c$ – použijeme princip inkluze a exkluze.
- $A_1 = \{1234, 1243, 1324, 1342, 1423, 1432\}$, $|A_1| = 3! = 6$. Podobně $|A_i| = 3!$ pro všechna i.
- $A_1 \cap A_2 = \{1234, 1243\}$, $|A_i \cap A_j| = 2!$ pro všechna $i \neq j$.
- $A_1 \cap A_2 \cap A_3$: pokud dáme tři dopisy do správných obálek, pak dáme i čtvrtý. Tedy pro trojice máme $|A_i \cap A_j \cap A_k| = 1$.

Příklad na princip inkluze a exkluze – pokračování

$$\begin{aligned}
|(A_1 \cup A_2 \cup A_3 \cup A_4)^c| &= |\Omega| - |A_1| - |A_2| - |A_3| - |A_4| \\
&\quad + |A_1 \cap A_2| + \cdots + |A_3 \cap A_4| \\
&\quad - |A_1 \cap A_2 \cap A_3| - \ldots - |A_2 \cap A_3 \cap A_4| \\
&\quad + |A_1 \cap A_2 \cap A_3 \cap A_4| \\
&= 4! - 4 \cdot 3! + 6 \cdot 2! - 4 \cdot 1 + 1 \\
&= 24 - 24 + 12 - 4 + 1 = 9
\end{aligned}$$

Pravděpodobnost, že nikdo nedostane svůj dopis je
$\frac{9}{24} = \frac{3}{8} = 37,5\%$.

Kontrolní otázka: Ověřte výpisem všech devíti možností.

Příklad (Rozesílání dopisů - obecně)

Řešme stejnou úlohu, kde rozesíláme n dopisů.

Princip inkluze a exkluze – finále

Pro $A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}$ máme $|A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}| = (n-k)!$
Počet všech k-prvkových podmnožin je $\binom{n}{k}$.
Proto je hledaná velikost množiny $(A_1 \cup A_2 \cup \dots \cup A_n)^c$ rovna

$$n! - n \cdot (n-1)! + \binom{n}{2}(n-2)! - \binom{n}{3}(n-3)! + \dots (-1)^n \binom{n}{0} 0!$$

$$= \sum_{k=0}^{n} (-1)^k \binom{n}{k} (n-k)! = \sum_{k=0}^{n} (-1)^k \frac{n!}{k!} = n! \cdot \sum_{k=0}^{n} (-1)^k \frac{(-1)^k}{k!}$$

Následně pravděpodobnost, že nikdo nedostane dopis je

$$1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} \dots = \frac{1}{2} - \frac{1}{6} + \frac{1}{24} - \frac{1}{120} + \dots \sim 36,8\%$$

Konverguje k hodnotě $\frac{1}{e}$.

Shrnutí, požadavky

- Pravděpodobnost: elementární příklady.
- Podmíněná a geometrická pravděpodobnost: obtížnost příkladů z přednášky.
- Princip inkluze a exkluze: příklady typu rozesílání dopisů ($n \leq 5$).

Princip inkluze a exkluze – přídavek

Příklad (Studenti u zkoušky)

Ke zkouškám ze tří předmětů přišlo 100 stejných studentů, přičemž známe následující statistiku. U předmětu A uspělo 60 studentů, u předmětu B uspělo 50 studentů a u předmětu C 70. Dále 35 studentů uspělo zároveň u předmětů A i B, 30 zvládlo B i C, a 40 A i C. Konečně 20 studentů uspělo u všech tří zkoušek. Kolik je studentů, kteří neuspěli ani u jedné zkoušky?

Vyřešte samostatně.

Příklad (Rozesazování s omezením)

Kolika způsoby můžeme posadit do řady 3 Němce, 3 Francouze a 3 Angličany tak, aby žádní tři krajané neseděli vedle sebe?

Rozesazování s omezením – řešení

Vlastnosti podle kterých dělíme všechna rozesazení jsou tři: 3 Angličani (resp. Francouzi, resp. Němci) sedí vedle sebe.

- Celkový počet všech rozesazení je $9!$.
- Počet všech rozesazení, kdy sedí vedle sebe 3 Angličani: (i) pro pořadí Angličanů máme $3!$ možností, (ii) pro rozesazení 6 osob a jednoho bloku (trojice Angličanů) máme $7!$ možností. Tzn. celkem $3! \cdot 7!$.
  Stejně pro trojici Francouzů (resp. Němců).
- Počet všech rozesazení, kdy sedí vedle sebe 3 Angličani a zároveň 3 Francouzi: počet pořadí Angličanů, krát počet pořadí Francouzů, krát počet rozesazení 3 Němců a dvou trojic (tedy 5 osob/bloků); celkem $(3!)^2 \cdot 5!$.
- Rozesazení, že sedí vedle sebe 3 Angličani, zároveň 3 Francouzi a zároveň 3 Němci je $(3!)^4$.
- Celkem podle principu inkluze a exkluze dostáváme $9! - \binom{3}{1} \cdot 3!7! + \binom{3}{2} \cdot (3!)^2 \cdot 5! - (3!)^4 = 283824$.