Pravděpodobnost, proč je to tak těžké?

Tomáš Svoboda and Petr Pošík

Vision for Robots and Autonomous Systems, Center for Machine Perception
Department of Cybernetics
Faculty of Electrical Engineering, Czech Technical University in Prague

11. března 2026

Notes

Specializovaný předmět na pravděpodobnost a statistiku teprve přijde.

Pravděpodobnost (nejistota) je všude

- Pravděpodobnost srážek zítra je 70%.
- Jakou mám šanci vyhrát v loterii?
- Mám pozitivní test na nemoc X, jsem opravdu nemocný?
- Svědectví jsou X, Y, a Z, je obviněný vinen?
- Nezaměstnanost se změnila o X, jaká bude inflace?
- Jak se bude vyvíjet cena akcií?
- Vybraná akce je X, o kolik se robot pohne?
- Jaká je pravděpodobnost, že na fotce je osoba XY?
- Jak dlouho mi bude trvat cesta, když pojedu tramvají?
- ...

Potřebujeme matematický popis ...

Notes

Co se naučíme

- Náhodný vs. deterministický pokus
- Náhodný jev jako jeho výsledek; prostor jevů
- Náhodná proměnná a její očekávaná hodnota

Notes

(Náhodný) pokus, jev

Pokus :

- Zjednodušeně: pozorování nějaké vlastnosti světa
- Děj či procedura, která
  - může být libovolněkrát konzistentně opakována a
  - má jednoznačně definovanou množinu možných výsledků.
- Deterministický pokus má jediný možný výsledek.
- Náhodný pokus má více možných výsledků.
- Neznáme výsledek *před* vykonáním náhodného pokusu. Po pokusu, je vše již jasné, nejistota mizí.

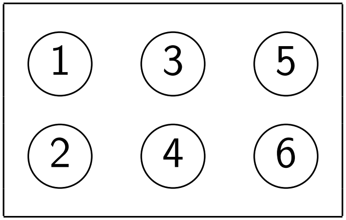
⚠ Náhodný jev je výsledek náhodného pokusu.

Jaká je pravděpodobnost náhodného jevu?

Notes

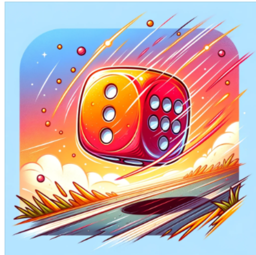
Náhodný pokus: Jeden hod hrací kostkou

Prostor (množina!) $\mathcal{S}$ všech elementárních jevů :



Náhodné jevy (výsledky pokusů) :

- A - číslo 6, $P(A) = ?$
- B - liché číslo, $P(B) = ?$
- C - číslo větší než 2, $P(C) = ?$
- D - číslo větší než 6, $P(D) = ?$



Think twice, DALL-E!

Notes

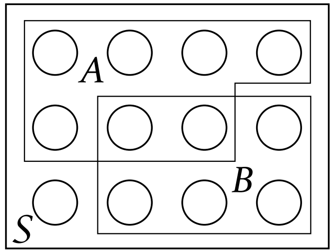
(Náhodné) jevy

Elementární jevy jsou *všechny možné, vzájemně se vylučující výsledky* nějakého experimentu.

Množinu elementárních jevů označme $\mathcal{S}$.

Náhodný jev A je jakákoli podmnožina množiny elementárních jevů, $A \subseteq \mathcal{S}$.

- Jev A nastal, pokud výsledek experimentu patří do A.
- Jevem je jakýkoli výrok o výsledku experimentu, pro který můžeme rozhodnout, zda nastal či nikoli.



Notes

Náhodný pokus 1: Házíme třikrát mincí (Head/Tails)

(Náhodný) Jev – výsledek náhodného pokusu

Prostor (množina!) $\mathcal{S}$ všech elementárních jevů. Jak je velká?

A 3^2

B 2^3

C $2 \cdot 3$

D ∞

(Náhodné) Jevy :

- A - $3\times$ panna (hlava), $P(A) = ?$
- B - $3\times$ stejný symbol, $P(B) = ?$
- C - alespoň jeden orel (tail), $P(C) = ?\ \ldots$

Notes

Pro výpočet velikosti prostoru všech elementárních jevů někdy poslouží vhodný model. Tady např. to může být $n-$bitové binární číslo.

Vyjděme z množiny elementárních jevů - $HHH, HHT, \ldots, TTT$.

Je pravděpodobnost některého elementárního jevu větší nebo menší než u ostatních? Je to tak vždy?

Jak definovat jevy A, B, C? Nešlo by jev C definovat snáze, pomocí množinových operací a jiných elementárních jevů?

Jaká je jejich pravděpodobnost? Jak by se dala spočítat $P(C)$ pomocí již známých pstí?

Proč nám nestačí pracovat přímo s jevy?

Pokus: hodíme třikrát mincí. Prostor elementárních jevů:

$$\mathcal{S} = \{\underbrace{HHH,\ HHT,\ HTH,\ HTT,\ THH,\ THT,\ TTH,}_{\text{alespoň jedna hlava – 7 prvků}}\ TTT\}$$

Chceme popsat jevy jako:

- *Padly právě 2 hlavy* $\Rightarrow$ jev $\{HHT, HTH, THH\}$
- *Padl sudý počet hlav* $\Rightarrow$ jev $\{TTT, HHT, HTH, THH\}$
- *Padlo více hlav než orlů* $\Rightarrow$ jev $\{HHH, HHT, HTH, THH\}$

Problém

Každý takový jev musíme definovat jako *výčet prvků* $\mathcal{S}$. Pro složitější pokusy (hod 10 kostkami, 1000 měření senzoru ...) to je **krajně nepraktické**.

Potřebujeme způsob, jak jevy popsat **číslem**.

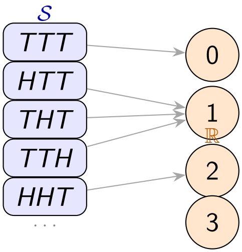
Notes

Náhodná veličina je funkce – ale proč se jmenuje veličina?

Definujme funkci X:

$$X(s) = \text{počet hlav v } s$$

$s \in \mathcal{S}$	$X(s)$
TTT	0
HTT, THT, TTH	1
HHT, HTH, THH	2
HHH	3



Formální definice

Náhodná veličina X na pravděpodobnostním prostoru $(\mathcal{S}, P)$ je **funkce**

$$X : \mathcal{S} \to \mathbb{R}.$$

Proč se tedy říká „veličina"?

Historický důvod: X se chová jako číslo – můžeme ji sčítat, násobit, počítat její průměr ...

Ale *před* provedením pokusu nevíme, jakou hodnotu X nabyde. Proto je **„náhodná"**.

Zdroj záměny:
X sama o sobě je deterministická funkce.
Náhodnost přichází ze *vstupu* – z výsledku pokusu s.

Notes

Toto je jádro celého konceptu. Napište na papír velké: X : S -> R a řekněte: „X je stroj, který každému výsledku pokusu přiřadí číslo. Stroj je deterministický. Náhodné je to, co do něj vložíme."
Analogie: výška člověka je deterministická funkce člověka. Ale pokud vybereme náhodného člověka z populace, výška se stane náhodnou veličinou.

Pojem „náhodná veličina" pochází z doby, kdy nebylo běžné myslet v pojmech funkcí a zobrazení. Dnes bychom řekli „náhodná funkce" nebo „měřitelná funkce", ale historický název zůstal.

Tři věci, které X může znamenat

Zápis	Co to je	Příklad
X	Celá funkce $X : \mathcal{S} \to \mathbb{R}$ (před pokusem)	„Počet hlav při hodu 3 mincemi"
$X(s)$	Konkrétní hodnota funkce pro výsledek s (po pokusu)	$X(HTH) = 2$
$P(X = k)$	Pravděpodobnost jevu $\{s \in \mathcal{S} : X(s) = k\}$	$P(X = 2) = \frac{3}{8}$

Správně

„Náhodná veličina X nabývá hodnoty 2 s pravděpodobností $\frac{3}{8}$."

Zápis: $P(X = 2) = \frac{3}{8}$

Pozor – časté chyby

$X = 2$ **není** pravda/nepravda – je to *jev* (množina výsledků).

$X(s)$ je číslo, X je funkce – nejsou to totéž!

Notes

Klíčový bod: $X = 2$ se tváří jako rovnice, ale ve skutečnosti je to zkrácený zápis za jev – množinu všech s, pro která $X(s) = 2$. Je to totéž jako $P(A)$ kde $A = \{HTH, HHT, THH\}$.

Cvičení na místě: jaká je množina elementárních jevů pro $P(X \geq 2)$ Pak ukažte, že P(X >= 2) = P(X=2) + P(X=3) = 3/8 + 1/8 = 1/2.

Příklady náhodných veličin

- Doba jízdy z domova do školy.
- Čekací doba na autobus.
- Počet zákazníků v obchodě daný den.
- Srážkový úhrn.
- Počet chyb v programu.
- ...

Notes

Diskrétní náhodná veličina

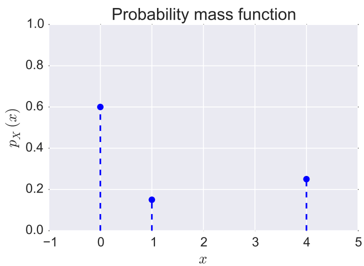
Náhodná veličina X je diskrétní, pokud hodnoty $X(s)$ přes všechny $s \in \mathcal{S}$ tvoří buď

- konečnou množinu hodnot $a_1, a_2, \ldots, a_n$, nebo
- nekonečnou, ale spočetnou množinu $a_1, a_2, \ldots$

Podpora (Support) n.v. X: $\mathcal{S}_X = \{x \in \mathbb{R} : P(X = x) > 0\} = \{a_1, a_2, \ldots\}$

Pravděpodobností funkce (rozdělení pravděpodobnosti) diskrétní n.v. X je funkce p_X definovaná jako

$$p_X(x) = P(X = x) = P(\{s \in \mathcal{S} : X(s) = x\}).$$



Notes

Nakreslete si jak to vypadá pro experiment hodu třemi mincemi a $X(s) =$ počet hlav.

Spojitá náhodná veličina – přehled

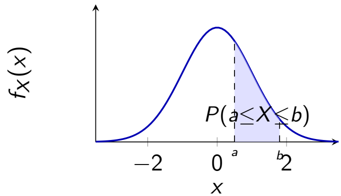
Náhodná veličina X je **spojitá**, pokud existuje funkce $f_X : \mathbb{R} \to \mathbb{R}_{\geq 0}$ taková, že pro každý interval $\langle a, b \rangle$:

$$P(a \leq X \leq b) \;=\; \int_a^b f_X(x)\,\mathrm{d}x.$$

f_X se nazývá **hustota pravděpodobnosti** (probability density function, PDF).

Klíčové vlastnosti:

- $f_X(x) \geq 0$ pro všechna x
- $\displaystyle\int_{-\infty}^{\infty} f_X(x)\,\mathrm{d}x = 1$
- $P(X = x) = 0$ pro každé konkrétní x



Příklady: doba jízdy tramvají, naměřená vzdálenost, šum senzoru.

Notes

Je dobré si uvědomit, že ne vše je diskrétní. Nemusíme rozumět integrálům do hloubky – stačí intuice, že PDF je „hustota" a plocha pod křivkou je pravděpodobnost. Uvažmě: $P(X = x) = 0$ je překvapivé, ale logické. Proč? Pro MDP a RL v tabulkovém případě nám postačí diskrétní případ, ale v robotice se spojitým stavovým prostorem se setkáme velmi brzy.

Střední (očekávaná) hodnota náhodné veličiny

$$\mathrm{E}[X] = \sum_{x \in \mathcal{S}_X} x \cdot P(X = x)$$

Příklad: Hod kostkou, X = číslo na kostce.

$$\mathrm{E}[X] = 1 \cdot \tfrac{1}{6} + 2 \cdot \tfrac{1}{6} + \cdots + 6 \cdot \tfrac{1}{6} = 3{,}5$$

Vlastnosti:

- $\mathrm{E}[r] = r, \quad \mathrm{E}[rX + s] = r\,\mathrm{E}[X] + s$
- $\mathrm{E}[X + Y] = \mathrm{E}[X] + \mathrm{E}[Y]$ (linearita, vždy platí)
- $\mathrm{E}[X \cdot Y] = \mathrm{E}[X] \cdot \mathrm{E}[Y]$ (pouze pro *nezávislé* X, Y)

⚠ Proč nás střední hodnota zajímá?
Střední hodnota je **nejlepší jednočíselný odhad** výsledku před provedením pokusu – minimalizuje střední kvadratickou chybu.

Notes

Klíčová věta pro expectimax: „maximalizujeme očekávanou odměnu." Linearita expectace je jeden z nejužitečnějších nástrojů v ML/RL

Střední hodnota – součet dvou kostek

Součet obou kostek: $S = X_1 + X_2$

$X_1 \backslash X_2$	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Přímý výpočet (zdlouhavý):

$$\mathrm{E}[S] = 2 \cdot \tfrac{1}{36} + 3 \cdot \tfrac{2}{36} + \cdots + 12 \cdot \tfrac{1}{36} = \frac{252}{36} = 7$$

Linearita – mnohem rychlejší!

$$\mathrm{E}[S] = \mathrm{E}[X_1 + X_2] = \mathrm{E}[X_1] + \mathrm{E}[X_2] = 3{,}5 + 3{,}5 = 7$$

Linearita platí **vždy** – X_1, X_2 nemusí být nezávislé.

Notes

Literatura, další zdroje

Klasické české čtení [2], na domácí stránce https://cmp.felk.cvut.cz/~navara/stat/ mnoho dalších zajímavých studijních materiálů.

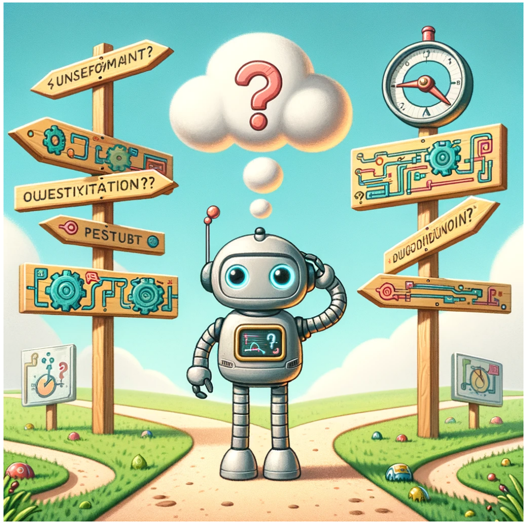
[1] Joseph K. Blitzstein and Jessica Hwang.
Introduction to Probability.
CRC Press, 2nd edition, 2019.
http://probabilitybook.net/.

[2] Mirko Navara.
Pravděpodobnost a matematická statistika.
ČVUT v Praze, Praha, 2007.
https://cmp.felk.cvut.cz/~navara/stat/.

[3] N. Silver.
Signál a šum.
Paseka, 2014.

Notes

Zbytek prezentace použijte dle potřeby . . .



Notes

Jevy a jejich kombinace

Důležité jevy:

- Jistý jev : $\mathcal{S}$, 1
- Nemožný jev : $\emptyset$, 0

Kombinace jevů:

- Konjunkce (A a B): $A \cap B$
- Disjunkce (A nebo B): $A \cup B$
- Jev opačný k A: $A^c = \mathcal{S} \setminus A$
- $A \Rightarrow B$: $A \subseteq B$
- Jevy neslučitelné : $A_1, \ldots, A_n : \bigcap\limits_{i \leq n} A_i = \emptyset$
- Jevy po dvou neslučitelné = vzájemně se vylučující : $A_1, \ldots, A_n : \forall i, j \in \{1, \ldots, n\},\ i \neq j : A_i \cap A_j = \emptyset$

Notes

Je to velmi podobné jako výroková logika: „Při hodu kostkou hodím číslo 5 nebo 6“, „Nehodím liché číslo, ani číslo 6.“

Výrokovou logiku lze k popisu jevů použít místo množin, je to ekvivalentní popis. Je dobré ale obojí nemixovat.

Úplný systém jevů

Úplný systém jevů je každá množina jevů $B_1, \ldots, B_n$, které se vzájemně vylučují a současně $\bigcup_{i=1}^{n} B_i = \mathcal{S}$.

- Množina elemenárních jevů $\mathcal{S}$ je z definice úplný systém jevů.
- Jevy $\{C, C^c\}$ tvoří úplný systém jevů: $C \cap C^c = \emptyset$ and $C \cup C^c = \mathcal{S}$.

Proč je úplný systém jevů důležitý koncept?

Notes

Proč je úplný systém jevů důležitý koncept?

Protože víme, že výsledkem experimentu je právě jeden z jevů v úplném systému. Proč?

Interpretace pravděpodobnosti

Frekventistická :

- Relativní četnost jevu po mnoha opakováních náhodného pokusu.

Bayesovská :

- Stupeň důvěry v to, že jev nastane.
- Umožňuje nám přiřadit pravděpodobnosti k výrokům typu "kandidát A vyhraje volby" nebo "podezřelý X je vinen", ačkoli nemůžeme opakovat stejné volby nebo stejný zločin stále dokola.

Notes

Vlastnosti pravděpodobnostní funkce

Pro jakoukoli platnou pravděpodobnostní funkci musí platit:

- $P(A) \in \langle 0, 1 \rangle$ (z definice)
- $P(\emptyset) = 0$; $P(\mathcal{S}) = 1$ (axiomy)
- $P(A^c) = 1 - P(A)$
- Pokud $A \subseteq B$, pak $P(A) \leq P(B)$
- Pokud $A \subseteq B$, pak $P(B \setminus A) = P(B) - P(A)$
- Pokud $A \cap B = \emptyset$, pak $P(A \cup B) = P(A) + P(B)$ (*aditivita*)
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Notes

Důkazy ponecháme specializovanějšímu předmětu. Intutivní náhled se snadno získá kreslením různých variant oblázkových světů, či Vennových diagramů.

Příklad 3: Pravděpodobnost částí

Vybereme-li z populace náhodně jednoho člověka, pak pro něj platí:

- Trpí nemocí X a je mladší 18 let s pravděpodobností 0.01.
- Trpí nemocí X a je mu/jí mezi 18 a 65 lety s pravděpodobností 0.05.
- Trpí nemocí X a je starší 65 let s pravděpodobností 0.09.

Jaká je pravděpodobnost, že náhodně vybraný člověk má nemoc X?

Notes

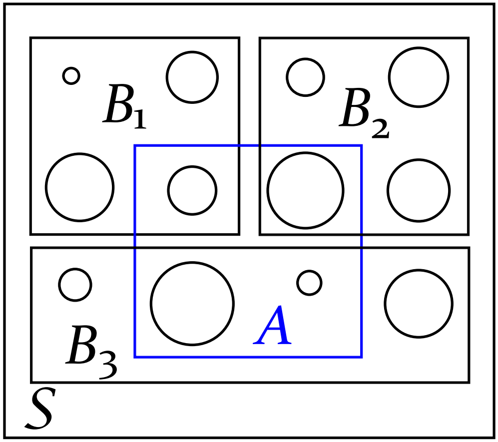
Vlastnostni pravděpodobnostní funkce (pokr.)

Pokud je $\{B_1, \ldots, B_n\}$ úplný systém jevů, pak pro jakýkoli jev $A \subseteq \mathcal{S}$

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) \, .$$

Konkrétně, pro úplný systém jevů $\{C, C^c\}$

$$P(A) = P(A \cap C) + P(A \cap C^c) \, .$$



Notes

Nezávislé jevy

Vzpomínáte na čerpadla u elektrárny? Chtěli bychom, aby byla nezávislá!

Jevy A a B jsou **nezávislé** právě tehdy, když

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

Jsou-li A a B nezávislé, pak

- $P(A \cup B) = P(A) + P(B) - P(A) \cdot P(B)$,
- a páry A, B^c a A^c, B a A^c, B^c jsou také nezávislé.

Notes

U čerpadel bychom například chtěli, aby byla připojena k různým zdrojům elektřiny. Ideálně třeba i různým zdrojem energie.

Nezávislost jevů: hod dvěma mincemi

- A - na první minci padla hlava
- B - na druhé minci padla hlava
- C - na mincích padly různé symboly

Které skupiny jevů jsou nezávislé?

A žádná skupina jevů

B dvojice (A, B), (B, C), (A, C)

C dvojice (A, B), (B, C), (A, C) i trojice (A, B, C)

D pouze trojice (A, B, C)

Notes

Každá dvojice jevů z A, B, C je nezávislá. Ale trojice (A, B, C) nezávislá není.
Mohou být dvojice (A, A) a (A, A^c) nezávislé?
Kdy pro (A, A) platí, že $P(A \cap A) = P(A)P(A) = P(A)$?
Kdy pro (A, A^c) platí, že $P(A \cap A^c) = P(A)P(A^c) = P(A)(1 - P(A))$?

Vlastnosti podmíněné pravděpodobnosti

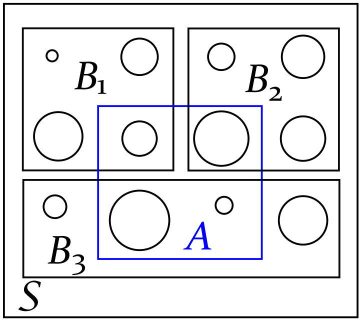
- $P(\mathcal{S}|B) = 1$, $P(\emptyset|B) = 0$.
- $P(A|A) = 1$, $P(A^c|A) = 0$.
- Pokud $B \subseteq A$, pak $P(A|B) = 1$.
- Pokud $P(A \cap B) = 0$, pak $P(A|B) = 0$.
- Pokud $A_1, \ldots, A_n$ jsou vzájemně se vylučující jevy, pak $P\left(\bigcup_{i=1}^{n} A_i \middle| B\right) = \sum_{i=1}^{n} P(A_i|B)$.
- Jevy A, B jsou nezávislé právě tehdy, když $P(A|B) = P(A)$ (pokud je $P(A|B)$ definována).

Notes

Věta o úplné pravděpodobnosti

Nechť $B_1, \ldots, B_n$ je úplný systém jevů vzhledem k $\mathcal{S}$ (tj., B_i jsou vzájemně neslučitelné a jejich sjednocením je $\mathcal{S}$) s $P(B_i) > 0$ pro všechna i.
Pak

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)$$



Notes