Pravděpodobnost je pravítko, které nekoupíte v obchodě

Jiří Dvořák
Katedra pravděpodobnosti a matematické statistiky, MFF UK

Celostátní konference učitelů matematiky SŠ
18. 9. 2023

Jaká je pravděpodobnost, že mi padne šestka?

Házení kostkou: šest jevů, každý stejně pravděpodobný $\Rightarrow$ 1/6

Definice – klasická pravděpodobnost

Nejjemnější výsledek náhodného pokusu označme *elementární jev*.

Předpokládejme, že existuje konečně mnoho elementárních jevů a všechny jsou stejně pravděpodobné.

Pravděpodobnost jevu A je rovna počtu příznivých elementárních jevů ku počtu všech elementárních jevů.

Příklad: P(liché číslo) = 3/6 = 1/2

Námitky: definice kruhem, konečně mnoho elementárních jevů

Axiomatická definice pravděpodobnosti (1933)

Andrej Nikolajevič Kolmogorov (1903–1987)

základ moderní teorie pravděpodobnosti

Mějme množinu Ω, její prvky budou *elementární jevy*. Může být nekonečná.

σ-algebra *(množiny, které chceme umět měřit)*

Definice: σ*-algebra* $\mathcal{A}$ je neprázdný systém podmnožin Ω, který je uzavřený vzhledem ke konečným a spočetným sjednocením a doplňkům, neboli

- $A_1, A_2, \dots \in \mathcal{A} \Rightarrow \cup_i A_i \in \mathcal{A}$,
- $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$.

Vybrané vlastnosti:

(i) $\emptyset, \Omega \in \mathcal{A}$

(ii) $A_1, \mathcal{A}_2, \ldots \in \mathcal{A} \Rightarrow \cap_i A_i \in \mathcal{A}$

Pravděpodobnost měří množiny

Definice: *Pravděpodobnost* je funkce P definovaná na $\mathcal{A}$, pro kterou platí:

- $P(A) \geq 0$ pro každé $A \in \mathcal{A}$,
- $P(\Omega) = 1$,
- $P(\cup_i A_i) = \sum_i P(A_i)$ pro $\{A_i\}$ po dvou disjunktní.

Vybrané vlastnosti:

(i) $P(\emptyset) = 0$

(ii) $P(A) + P(A^c) = 1$

(iii) $P(A) \leq P(B)$, pokud $A \subseteq B$

(iv) $0 \leq P(A) \leq 1$ pro každé $A \in \mathcal{A}$

Vyrábíme pravítko P

Otázka: Náhodně vybereme číslo X od 1 do 12. Jaká je pravděpodobnost, že bude $X \leq 6$?

Množina elementárních jevů: $\Omega = \{1, 2, \ldots, 12\}$

σ-algebra: $\mathcal{A} = 2^\Omega$

Pravítko poskládáme tak, že $P(\{1\}) = 1/12 = P(\{2\}) = \ldots = P(\{12\})$

Už můžeme měřit i jiné množiny: $P(\{1,2,3\}) = 3/12$ apod.

Odpověď: $P(X \leq 6) = P(\{1,2,3,4,5,6\}) = 6/12 = 1/2$

Vyrábíme jiné pravítko

Otázka: Náhodně vybereme číslo X od 1 do 12. Jaká je pravděpodobnost, že bude $X \leq 6$?

Množina elementárních jevů: $\Omega = \{1, 2, \ldots, 12\}$

σ-algebra: $\mathcal{A} = 2^{\Omega}$

Pravítko poskládáme tak, že platí

k	1	2	3	4	5	6
$P(\{k\})$	0	1/36	2/36	3/36	4/36	5/36
k	7	8	9	10	11	12
$P(\{k\})$	6/36	5/36	4/36	3/36	2/36	1/36

Odpověď: $P(X \leq 6) = P(\{1,2,3,4,5,6\}) = 15/36 \neq 1/2$

Poznámka: Náhodnost nemusí znamenat rovnoměrnost!

Nejdůležitější slide této přednášky

Pravděpodobnost P (jako množinová funkce) popisuje chování uvažovaného experimentu, dává jeho *model*.

Model určuje hřiště, na kterém hrajeme, a pravidla, která tam platí. Tedy určuje odpověď na naši otázku.

Nalezená odpověď ale platí jen na tom konkrétním hřišti. Na jiném hřišti (v jiném modelu) může platit jiná odpověď.

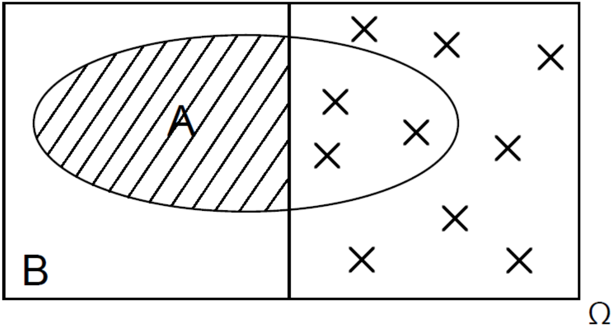
Jestli je odpověď *správná* nebo *špatná* není absolutní fakt, ale záleží na použitém modelu.

Úplně jinou otázkou ale je, který model dobře popisuje tu reálnou situaci, a který model tedy dává *užitečné* odpovědi.

Podmíněná pravděpodobnost – upravujeme pravítko

Definice: Pokud B je jev splňující $P(B) > 0$, pak *podmíněnou pravděpodobnost* jevu A za podmínky jevu B definujeme jako

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



Podmíněná pravděpodobnost je pořád pravděpodobnost

Staré pravítko: $P(A)$

Nové pravítko: $P_1(A) = P(A \mid B)$

Vybrané vlastnosti:

(i) $P_1(\emptyset) = 0$

(ii) $P_1(A) + P_1(A^c) = 1$

(iii) $P_1(A) \leq P_1(B)$, pokud $A \subseteq B$

(iv) $0 \leq P_1(A) \leq 1$ pro každé $A \in \mathcal{A}$

Dostihy

Favority dostihu jsou koně Amarant a Baklažán. Odborníci tipují, že Amarant zvítězí s pravděpodobností $5/10$ a Baklažán s pravděpodobností $3/10$.

Amarant ztratil na startu tolik, že je již jisté, že nezvítězí.

Jaká je nyní pravděpodobnost, že zvítězí Baklažán?

Odpověď: P(zvítězí Baklažán $\mid$ Amarant nezvítězí) = 3/5, protože
P(zvítězí Baklažán, Amarant nezvítězí) = 3/10,
P(Amarant nezvítězí) = 5/10.

Opatrnosti je třeba! $P(A \mid B) \neq P(B \mid A)$

U viníků dopravních nehod bylo v 10 % případů zjištěno požití alkoholických nápojů. Znamená to, že střízliví řidiči jsou více nebezpeční?

Když N bude značit jev, že došlo k nehodě, a O jev, že řidič pil alkohol, pak zadání říká $P(O \mid N) = 0{,}1$, neboli $P(O^c \mid N) = 0{,}9$.

To ale nic nevypovídá o $P(N \mid O)$ nebo $P(N \mid O^c)$. K tomu bychom potřebovali mít nějakou dodatečnou znalost.

Předpokládejme, že $P(O) = 0{,}005$, potom dostáváme
$P(N \mid O) = P(N \cap O)/P(O) = P(O \mid N)P(N)/P(O) = 20P(N)$.

$P(N \mid O^c) = P(O^c \mid N)P(N)/P(O^c) = \frac{0{,}9}{0{,}995}P(N) \doteq 0{,}905P(N)$,
tedy požití alkoholu je daleko nebezpečnější.

Zmatky jsou bohužel časté

Vybráno ze zpráv:

- pozor na české turisty v Tatrách (nejvíc zahraničních turistů zapletených do nehody pochází z Česka),
- chlapci jsou ve větším ohrožení při jízdě na kole (většina dětí při nehodách na kole jsou chlapci),
- fotbal je nejnebezpečnější sport (průzkum nehod při sportu),
- domov je nebezpečné místo (ke třetině nehod dochází ve vlastním bydlení).

Z toho, že $P(A \mid B)$ je velká, je nesprávně usuzováno, že $P(B \mid A) > P(B)$. To ale nelze učinit bez informace o $P(A)$.

Například A = český turista, B = nehoda.

$P(A \mid B) > P(A \mid C)$ neznamená $P(A \mid B \cap D) > P(A \mid C \cap D)$

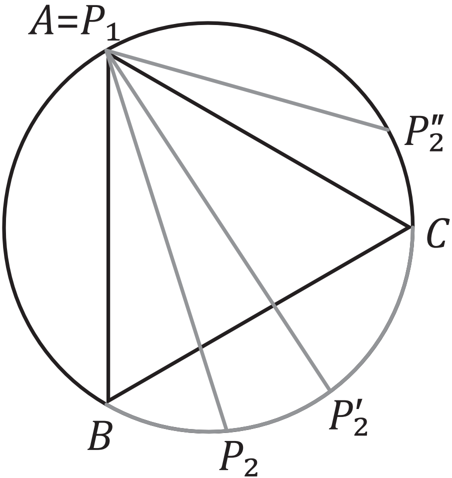
V době španělsko–americké války (1898) byla v námořnictvu USA úmrtnost zhruba 9 promile, zatímco u civilistů v New Yorku byla 16 promile. Armáda použila tyto statistiky při náboru nováčků.

V armádě jsou zdraví a silní lidé, zatímco mezi civilisty jsou zahrnuti také všichni staří nebo nemocní. Označme jako jev A úmrtí daného člověka, jev B, že jde o obyvatele New Yorku, a jev C, že jde o člena námořnictva. Pak data udávají $P(A \mid B) > P(A \mid C)$.

Kdybychom uvažovali stejně staré lidi, např. ve věku 18–30 let (jev D), pak můžeme čekat, že $P(A \mid B \cap D) < P(A \mid C \cap D)$.

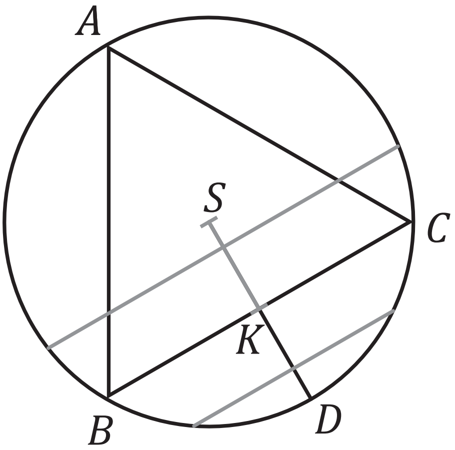
Bertrandův paradox poprvé

Do kružnice je vepsán rovnostranný trojúhelník. Náhodně zvolíme tětivu, která kružnici protíná. Jaká je pravděpodobnost, že délka tětivy uvnitř kružnice je větší než délka strany vepsaného trojúhelníku?



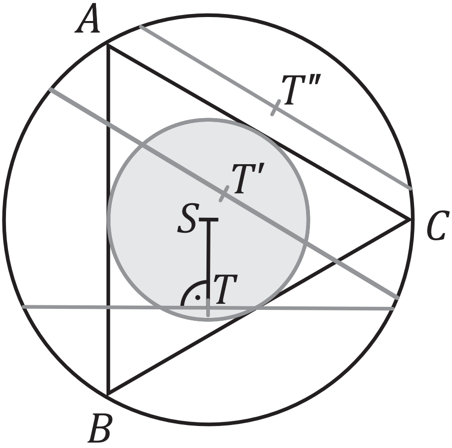
Bertrandův paradox podruhé

Do kružnice je vepsán rovnostranný trojúhelník. Náhodně zvolíme tětivu, která kružnici protíná. Jaká je pravděpodobnost, že délka tětivy uvnitř kružnice je větší než délka strany vepsaného trojúhelníku?



Bertrandův paradox potřetí

Do kružnice je vepsán rovnostranný trojúhelník. Náhodně zvolíme tětivu, která kružnici protíná. Jaká je pravděpodobnost, že délka tětivy uvnitř kružnice je větší než délka strany vepsaného trojúhelníku?



Bertrandův paradox

Obrázky převzaty z článku:

J. Dvořák, M. Snětinová (2019): Bertrandův paradox aneb není náhoda jako náhoda. Rozhledy matematicko-fyzikální 94(2), 12-17.

(dostupný na https://dml.cz/handle/10338.dmlcz/147999)