Matematika III – 9. přednáška
Pravděpodobnost – opakování a zobecnění pojmů

Michal Bulant

Masarykova univerzita
Fakulta informatiky

13. 11. 2013

Obsah přednášky

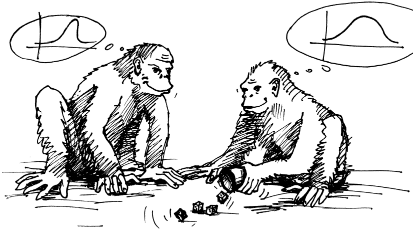
Doporučené zdroje

- Martin Panák, Jan Slovák, **Drsná matematika**, *e*-text.
- Karel Zvára, Josef Štěpán, **Pravděpodobnost a matematická statistika**, Matfyzpress, 4. vydání, 2006, 230 stran, ISBN 80-867-3271-1.
- Marie Budíková, Štěpán Mikoláš, Pavel Osecký, **Teorie pravděpodobnosti a matematická statistika (sbírka příkladů)**, Masarykova univerzita, 3. vydání, 2004, 117 stran, ISBN 80-210-3313-4.
- Marie Budíková, **Statistika**, Masarykova univerzita, 2004, distanční studijní opora ESF, `http://www.math.muni.cz/~budikova/esf/Statistika.zip`.
- Marie Budíková, Tomáš Lerch, Štěpán Mikoláš, **Základní statistické metody**, Masarykova univerzita, 2005, 170 stran, ISBN 80-210-3886-1.

Motto:

42,35 procenta všech statistik je nesmyslných.

Statistika v širším slova smyslu je jakékoliv zpracování číselných dat o nějakém souboru objektů a jejich více či méně přehledná prezentace.



Podstatou **matematické statistiky** je pro daná data zjišťovat, jaké vlastnosti mají objekty, které jsou daty popisovány. Zpravidla jde o sběr dat o části souboru objektů, jejich následnou analýzu a konečně o vyslovení důsledků pozorování pro celý soubor. Výsledkem práce matematického statistika je sdělení o velkém souboru objektů na základě studia malé (zpravidla náhodně vybrané) části z nich **společně s kvalitativním odhadem věrohodnosti výsledného sdělení.**
Teorie pravděpodobnosti studuje modely popisující chování abstraktních souborů (pravděpodobnost jevů z jevového pole), statistika studuje skutečné náhodné výběry z nějakého základního souboru a zdůvodňuje výběr teoretického pravděpodobnostního modelu, resp. kvalitativní informace o jeho parametrech.

Příklad

Za soubor objektů vezměme všechny studenty přednášky Matematika III (podzim 2007), jako číselný údaj můžeme uvažovat

1. průměrné bodové hodnocení studenta u zkoušky,
2. průměrnou známku u zkoušky z tohoto (2,92) a z jiných pevně vybraných předmětů (`IB000` – 2,95; `IB102` – 2,89) ,
3. nejčastější známku (resp. úspěšnou známku) z tohoto předmětu (`F` – 92 krát, `E` – 91 krát), nejméně častou známku (`B` – 15 krát),
4. průměrný počet bodů dosažených na jednotlivých termínech zkoušky (1. – 16,8; 2. – 8,9; 3. – 8,1; příklad, za nějž bylo uděleno nejvíce (nejméně) procent možných bodů – min. kostra (1B, 82,5%), resp. rekurence (2A, 3,6%)
5. počet pracovních hodin týdně odpracovaných mimo fakultu,
6. číselná data vypovídající o historii dřívějšího studia

a mnoho dalších údajů.

Zastavme se u prvního údaje. Samotný aritmetický průměr bodů nám mnoho neřekne nejen o kvalitě přednášky a o kvalitě přednášejícího, ale ani o samotném hodnocení. Zajímá nás také hodnota, která bude „uprostřed souboru“, tj. počet bodů, pro které je stejně studentů pod ní a nad ní.
Obdobně první a poslední čtvrtina, desetina apod. Všem takovým údajům říkáme **statistiky** posuzované veličiny. V uvedených příkladech se jim říká **medián**, **kvartil**, **decil** apod.
Z obecné zkušenosti nebo jako výsledek úvah mimo matematiku víme, že rozumné hodnocení by mělo mít tzv. **normální rozdělení** (odpovídá tzv. *Gaussově křivce*). Tento pojem patří do teorie pravděpodobnosti a k jeho zavedení budeme potřebovat poměrně dost matematiky.
Porovnáním výsledku třeba i docela malého náhodného výběru studentů s teoretickým modelem můžeme zjistit odhad parametrů takového rozdělení a činit závěry, zda je hodnocení „rozumné“. Zároveň lze popsat věrohodnost našich závěrů.

Daleko zajímavější vývody ovšem můžeme činit, když porovnáním statistik pro různé veličiny budeme moci dovozovat informace o souvislostech (*korelace*, *závislost*). Pokud např. neexistuje žádná doložitelná souvislost mezi historií předchozího studia a výsledky v dané přednášce, je jedním z možných vysvětlení závěr, že je přednáška (nebo její hodnocení) prostě špatná.

Závěr úvodních úvah:

- V matematice pracujeme s abstraktním matematickým popisem pravděpodobnosti.
- Vývody pro konktrétní soubory dat, pro které je zvolený model relevantní, dává matematická statistika.
- To, zda je takový popis adekvátní pro konkrétní výběr dat, je také možné podpořit nebo zavrhnout pomocí metod matematické statistiky.

Připomeneme (a trochu zobecníme) pojmy a výsledky z prvního semestru.

Definice (Náhodné jevy)

Budeme pracovat s neprázdnou pevně zvolenou množinou Ω všech možných výsledků, kterou nazýváme **základní prostor**.
Prvky $\omega \in \Omega$ představují jednotlivé **možné výsledky**, též **elementární jevy**[a].
Systém (ne nutně všech) podmnožin $\mathcal{A}$ základního prostoru se nazývá **jevové pole** a jeho prvky se nazývají **jevy**, jestliže

- $\Omega \in \mathcal{A}$, tj. základní prostor, je jevem,
- je-li $A, B \in \mathcal{A}$, pak $A \setminus B \in \mathcal{A}$, tj. pro každé dva jevy je jevem i jejich množinový rozdíl,
- je-li $A_i \in \mathcal{A}$, $i \in I$ nejvýše spočetný systém jevů, pak také jejich sjednocení je jevem, tj. $\cup_{i \in I} A_i \in \mathcal{A}$.

[a] Přesněji: elementárním jevem je $\{\omega\}$.

Důsledek

- *Komplement $A^c = \Omega \setminus A$ jevu A je jevem, který nazýváme* opačný jev *k jevu A.*
- *Průnik dvou jevů opět jevem, protože pro každé dvě podmnožiny $A, B \subset \Omega$ platí*

$$A \setminus (\Omega \setminus B) = A \cap B.$$

Takový systém množin $\mathcal{A}$ se pak nazývá σ-algebra.
Jevové pole je tedy systém podmnožin základního prostoru uzavřený na konečné průniky, spočetná sjednocení a množinové rozdíly. Jednotlivé množiny $A \in \mathcal{A}$ nazýváme **náhodné jevy** (vzhledem k $\mathcal{A}$).

Terminologie připomíná souvislosti s popisem skutečných jevů a jejich statistickým popisem:

- celý základní prostor Ω se nazývá **jistý jev**, prázdná podmnožina $\emptyset \in \mathcal{A}$ 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 $\cup_{i \in I} A_i$,
- $A, B \in \mathcal{A}$ jsou **neslučitelné jevy**, je-li $A \cap B = \emptyset$,
- jev A má za **důsledek** jev B, když $A \subset B$,
- je-li $A \in \mathcal{A}$, pak se jev $B = \Omega \setminus A$ nazývá **opačný jev k jevu** A, píšeme $B = A^c$.

Definice (Kolmogorovova definice pravděpodobnosti)

Pravděpodobnostní prostor je jevové pole $\mathcal{A}$ podmnožin (konečného) základního prostoru Ω, na kterém je definována funkce $P : \mathcal{A} \to \mathbb{R}$ s následujícími vlastnosti:

- je nezáporná, tj. $P(A) \geq 0$ pro všechny jevy A,
- je aditivní, tj. $P(\cup_{i \in I} A_i) = \sum_{i \in I} P(A_i)$, pro každý nejvýše spočetný systém po dvou neslučitelných jevů,
- pravděpodobnost jistého jevu je 1.

Funkci P nazýváme **pravděpodobností** na jevovém poli $(\Omega, \mathcal{A})$.

Důsledek

Pro všechny jevy $A, B \in \mathcal{A}$ platí

- $P(\emptyset) = 0,\ 0 \leq P(A) \leq 1,$
- $P(A^c) = 1 - P(A),$
- $A \subseteq B \implies P(A) \leq P(B),\ P(B \setminus A) = P(B) - P(A),$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Podobná tvrzení platí i pro nekonečné posloupnosti jevů:

Tvrzení

Pro libovolnou nejvýše spočetnou množinu jevů $(A_i)_{i=1}^{\infty}$ *platí:*

- *Je-li* $A_1 \subseteq A_2 \subseteq \cdots$, *pak*

$$P(\bigcup_{i=1}^{\infty} A_i) = \lim_{i\to\infty} P(A_i),$$

- *Je-li* $A_1 \supseteq A_2 \supseteq \cdots$, *pak*

$$P(\bigcap_{i=1}^{\infty} A_i) = \lim_{i\to\infty} P(A_i),$$

- $P(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i)$,
- $P(\bigcap_{i=1}^{\infty} A_i) \geq 1 - \sum_{i=1}^{\infty}(1 - P(A_i))$.

Příklad

Náhodný pokus spočívá v hodu kostkou. Jev A znamená, že padne liché číslo, jev B, padne-li prvočíslo.

a) Určete základní prostor Ω.

b) Uveďte všechny možné výsledky příznivé nastoupení jevů A, B.

c) Pomocí A, B a operací s jevy vyjádřete:
   - padne sudé číslo,
   - padne číslo 2,
   - padne číslo 2 nebo 3

d) Určete nejmenší jevové pole $(\Omega, \mathcal{A})$, obsahující jevy A i B.

Klasická pravděpodobnost

Připomeňme si klasickou konečnou pravděpodobnost.

Definice

Nechť Ω je konečný základní prostor a nechť jevové pole $\mathcal{A}$ je právě systém všech podmnožin v Ω. **Klasická pravděpodobnost** je pravděpodobnostní prostor $(\Omega, \mathcal{A}, P)$ s pravděpodobnostní funkcí $P : \mathcal{A} \to \mathbb{R}$,

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

Zjevně takto zadaná funkce skutečně definuje pravděpodobnost, kdy všem elementárním jevům přiřazujeme stejnou pravděpodobnost.

Že s klasickou pravděpodobností nevystačíme, ukazují následující příklady:

Příklad

- Cestou z Kotlářské na Botanickou jsem ztratil zadání písemky. Určete pravděpodobnost jevu ω_X slovně vyjádřeného: *ztracená písemka se nachází nejblíže k zastávce trolejbusu* X.
- Určete pravděpodobnost, jevu ω_k: *při opakovaném hodu mincí padne hlava poprvé při k-tém pokusu*.

V prvním případě je třeba pracovat s nekonečně mnoha stejně pravděpodobnými elementárními jevy: *písemku jsem ztratil v bodě* (x, y) , ve druhém pak musíme připustit teoretickou možnost, že hlava nepadne nikdy, a prostorem jevů tedy bude $\mathbb{N} \cup \{\infty\}$.

Geometrická pravděpodobnost

V praktických problémech se často setkáváme s daleko složitějšími modely, kde základní prostor není konečnou množinou. Uvažme rovinu $\mathbb{R}^2$ dvojic reálných čísel a v ní podmnožinu Ω se známým obsahem $\operatorname{vol}\Omega$ (tedy např. takovou, že Ω je Riemannovsky měřitelná)). Náhodné jevy budou reprezentovány podmnožinami $A \subseteq \Omega$ a za jevové pole $\mathcal{A}$ bereme nějaký vhodný systém podmnožin, u kterých umíme určit jejich obsah. Nastoupení nebo nenastoupení jevu je dáno výběrem bodu v Ω, kterým se trefíme nebo netrefíme do množiny reprezentující jev A.
Podobně jako u klasické pravděpodobnosti definujeme pravděpodobnostní funkci
$P : \mathcal{A} \to \mathbb{R}$ vztahem

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

kde A jsou podmnožiny v rovině, které odpovídají námi vybraným jevům.

Příklad

Jaká je pravděpodobnost, že dvě náhodně zvolená čísla z intervalu $(0,1)$ budou mít součet menší než 1 a součin větší než $2/9$?

Příklad (Buffonova úloha)

Rovina je rozdělena rovnoběžkami umístěnými rovnoměrně ve vzdálenosti d. Do roviny je náhodně umístěna jehla délky $l < d$. Jaká je pravděpodobnost, že jehla protne některou rovnoběžku?

Podmíněná pravděpodobnost a nezávislost

Motto:

Je dokázáno, že slavení narozenin je zdraví prospěšné. Statistika ukazuje, že lidé, kteří oslavili nejvíce narozenin, se dožívají nejvyššího věku.

Vtip – co je na něm špatně?

Statistik procházel bezpečnostní kontrolou na letišti, když byla v jeho kufru nalezena bomba. Vysvětloval: „Podle statistik je pravděpodobnost přítomnosti bomby v letadle 1:1000. Takže šance, že na palubě budou dvě bomby, je 1:1000000. Tím pádem jsem mnohem více v bezpečí . . . “

Obvyklé je také klást dotazy s dodatečnou podmínkou. Např.

- Jaká je pravděpodobnost, že při hodu dvěma kostkami padly dvě pětky, je-li součet hodnot deset?
- Mějme urnu s 10 koulemi. Desetkrát jsem vytáhl kouli, zkontroloval její barvu a vrátil do urny. Jestliže byla vždy bílé barvy, s jakou pravděpodobností jsou všechny koule v urně bílé?
- Na dostizích jsou známy pravděpodobnosti vítězství jednotlivých koní. Jak se tyto pravděpodobnosti změní, pokud uprostřed závodu spadne jezdec jednoho z koní ze sedla?

Připomeňme, že formalizovat takové úvahy umíme následovně.

Definice

Nechť H je jev s nenulovou pravděpodobností v jevovém poli $\mathcal{A}$ v pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$. **Podmíněná pravděpodobnost** $P(A|H)$ jevu $A \in \mathcal{A}$ vzhledem k jevu H je definována vztahem

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

Přirozená definice nezávislosti je, že hypotéza H a jev A jsou nezávislé tehdy, je-li $P(A) = P(A|H)$.
Z výše uvedeného snadno vyplývá *symetričtější* definice:

Definice

Říkáme, že jevy A a B jsou nezávislé, jestliže

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

Definice

Říkáme, že jevy $A_1, A_2, \dots$ jsou nezávislé, jestliže pro každou k-tici $A_{i_1}, \cdots, A_{i_k}$ z nich platí

$$P\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{j=1}^{k} P(A_{i_j}).$$

Příklad

V urně jsou 4 lístky označené 000, 110, 101, 011. Uvažujme pro $i = 1, 2, 3$ náhodné jevy
$A_i = \{$náhodně vytažený lístek má na i-tém místě 1$\}$.
Snadno se vidí, že $P(A_1) = P(A_2) = P(A_3) = \frac{1}{2}$, dále, že $P(A_1 \cap A_2) = P(A_1 \cap A_3) = P(A_2 \cap A_3) = \frac{1}{4}$ a že $P(A_1 \cap A_2 \cap A_3) = 0$. Jevy A_1, A_2, A_3 jsou tedy po dvou nezávislé, ale nejsou nezávislé.

Bayesovy věty

Přepsáním formule pro podmíněnou pravděpodobnost dostáváme

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

Věta (Bayesovy věty)

Pro pravděpodobnost jevů A a B platí

1. $P(A|B) = \frac{P(A)P(B|A)}{P(B)}$.
2. $P(A|B) = \frac{P(A)P(B|A)}{P(A)P(B|A)+P(A^c)P(B|A^c)}$.

Důkaz.

První tvrzení je přepsáním předchozí formule, druhé z prvého plyne dosazením $P(B) = P(A)P(B|A) + P(A^c)P(B|A^c)$. $\square$