Pravděpodobnost a statistika

Pravděpodobnost

- předmětem zkoumání jsou zákonitosti náhody
- vznik asi v polovině 17.století jako „teorie hazardních her"
- deduktivní charakter
- na základě úvahy o podstatě jevu usuzujeme o výsledku pokusu

Statistika

- induktivní uvažování (statistická interference)
- na základě souboru výsledků pokusu usuzujeme o podstatě jevu
- teorie odhadů, testování hypotéz, regresní statistika, ekonomická statistika, ...

Náhodný pokus

- každý proces, jehož výsledek je při jinak stejných počátečních podmínkách nejistý
- výsledek závisí na náhodě
- výsledek nejsme schopni s jistotou předpovědět
- výsledek je ovlivněn řadou drobných ne úplně zjistitelných nebo nezjistitelných činitelů
- množinu všech možných výsledků náhodného pokusu označujeme Ω

Příklad
— hod kostkou
— sledování doby bezporuchového chodu stroje
— sledování počtu výskytu události během definovaného časového intervalu
— stanovení chyby měřícího přístroje,
— vyrobení výrobku,
— podání léku pacientovi, ...

$\triangle$

Náhodný jev

- výsledek náhodného pokusu
- náhodné jevy značíme velkými latinskými písmeny z počátku abecedy A,B,C, ...
- jev A je podmnožina množiny Ω ($A \subset \Omega$)
- celá množina Ω je jev jistý
- prázdná množina $\emptyset$ je jev nemožný

Příklad

— na kostce padl počet bodu menší než 3
— na kostce padlo liché číslo
— k poruše stroje došlo mezi 3 a 4 hodinou provozu
— během časového intervalu se vyskytlo více než 10 sledovaných událostí
— vyrobený výrobek odpovídal všech požadavkům na kvalitu, ...

$\triangle$

Elementární jev

- prvky prostoru Ω značíme ω a mluvíme o elementárních jevech
- $\omega_i, \; i = 1, 2, \ldots$ jsou možné výsledky náhodného pokusu
- ω_i jsou minimální jevy různé od jevu nemožného
  (ω je elementární jev: $\forall A \subset \omega \Rightarrow (A \equiv \omega)$ nebo $(A \equiv \emptyset)$)
- elementární jevy jsou párově neslučitelné
  (ω_1, ω_2 různé elemetární jevy, pak $\omega_1 \cap \omega_2 \equiv \emptyset$)
- elementární jevy tvoří úplný systém neslučitelných jevů
- každý jev A lze vyjádřit jako množinu elementárních jevů
  ($A \equiv \{\omega_1, \omega_2, \ldots\}$)

Příklad

— hodíme kostkou a sledujeme počet padlých bodů: $\Omega = \{\omega_i = i, i = 1, 2, 3, 4, 5, 6\}$

— jev A - padne sudé číslo: $A = \{\omega_2, \omega_4, \omega_6\}$

— jev B - padne číslo větší než 3: $B = \{\omega_4, \omega_5, \omega_6\}$

$\triangle$

Operace s jevy

$A \equiv B$	rovnocenné jevy
$\bar{A}$ (nebo A^c nebo A') $\equiv \Omega \backslash A$	jev opačný doplněk jevu
$A \subset B$	jev A je podjevem jevu B
$A \cap B$	průnik jevů jev A a zároveň jev B
$A \cup B$	sjednocení jevů jev A nebo jev B (nebo oba jevy)
$A \backslash B$	rozdíl jevů jev A, ale nikoliv jev B
$A \cap B \equiv \emptyset$	jevy disjunktní jevy neslučitelné
$\bigcup A_i \equiv \Omega$	úplný systém jevů

---

zákon jedinečnosti	$\forall A, B \;\; \exists!\; A \cap B$ a $\exists!\; A \cup B$
zákon komutativní	$A \cup B \equiv B \cup A$ $A \cap B \equiv B \cap A$
zákon asociativní	$(A \cup B) \cup C \equiv A \cup (B \cup C)$ $(A \cap B) \cap C \equiv A \cap (B \cap C)$

Operace s jevy - pokračování

zákon identity $\quad A \cup \emptyset \equiv A \qquad A \cap \emptyset \equiv \emptyset$

$A \cup \Omega \equiv \Omega \qquad A \cap \Omega \equiv A$

zákon komplementu $\quad A \cup \bar{A} \equiv \Omega \qquad A \cap \bar{A} \equiv \emptyset$

zákon distributivní $\quad (A \cap B) \cup C = (A \cup C) \cap (B \cup C)$

$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$

de Morganovy vzorce $\quad \overline{A \cup B} \equiv \overline{A} \cap \overline{B}$

$\overline{A \cap B} \equiv \overline{A} \cup \overline{B}$

obecněji nechť $A_i, i = 1, 2, \ldots, A_n$ jsou jevy

$$\overline{\bigcup_{i=1}^{n} A_i} \equiv \bigcap_{i=1}^{n} \overline{A_i}$$

$$\overline{\bigcap_{i=1}^{n} A_i} \equiv \bigcup_{i=1}^{n} \overline{A_i}$$

Pravděpodobnost jevu

- každému jevu A přiřazujeme reálné číslo $\mathcal{P}(A)$
- pravděpodobnost (ppst) lze chápat jako předpověď poměrných četností výsledků při mnohonásobném opakování daného pokusu
- ppst lze chápat jako kvantitativní ohodnocení stupně jistoty

  $\dots$

- klasická definice ppsti
- statistická definice ppsti
- geometrická deifnice ppsti
- axiomatická definice ppsti

Klasická definice pravděpodobnosti

Nechť

- $\Omega = \{\omega_1, \omega_2, \omega_3, \ldots, \omega_N\}$ množina možných výsledků pokusu je konečná a neprázdná ($0 < N < \infty$)
- $p_1, p_2, p_3, \ldots, p_N$ jsou nezáporná čísla splňující $\sum_{i=1}^{N} p_i = 1$ charakterizující poměrnou četnost výskytu ω_i
- $p_1 = p_2 = \ldots = p_N = \frac{1}{N}$ všechny výsledky pokusu jsou stejně možné
- každý jev A lze popsat množinou jevů $\{\omega_{i_1}, \omega_{i_2}, \ldots, \omega_{i_k}\}$ kde ω_i jsou výsledky pokusu příznivé jevu A

pak

$$\mathcal{P}(A) = \frac{N_A}{N}$$

kde

N_A je počet výsledků příznivých jevu A
N je počet všech možných výsledků

Kombinatorické vzorce

Nechť $M = \{a_1, a_2, \ldots, a_n\}$ je skupina n různých prvků

- permutace n prvků: upořádání prvků skupiny M v daném pořadí
  - počet permutací $P_n = n!$
  - pokud M se skládá z $i_1, i_2, \ldots, i_k$ stejných prvků, je počet permutací $P_n = \frac{n!}{i_1! i_2! \ldots i_k!}$
  - počet permutací s opakováním $P_n = n^n$

  kolika způsoby lze uspořádat n-tici prvků

- variace n prvků k-té třídy: ze skupiny M vybereme k prvků a ty permutujeme
  - počet variací $V_n^k = \frac{n!}{(n-k)!} = n(n-1)\ldots(n-k+1)$
  - počet variací s opakováním $\overline{V_n^k} = n^k$

  kolika způsoby lze z n-tici prvků vybrat k-tici, přičemž záleží na pořadí výběru

Kombinatorické vzorce - pokračování

- kombinace n prvků k-té třídy: ze skupiny M vybereme k prvků
  - počet kombinací $C_n^k = \binom{n}{k} = \frac{n!}{(n-k)!k!}$
  - počet kombinací s opakováním $\overline{C_n^k} = \binom{n+k-1}{k}$

  kolika způsoby lze z n-tici prvků vybrat k-tici, přičemž nezáleží na pořadí výběru

- binomické číslo lze přibližně určit za použití Stirlingovy formule pro určení hodnoty $n!$

$$\log n! \approx \log \sqrt{2\pi k} + k(\log k - \log e)$$

- vlastnosti kombinačních čísel
  - $\binom{n}{k} = \binom{n}{n-k}$
  - $\binom{n}{0} = \binom{n}{n} = 1$
  - $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$
  - $\binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{n} = 2^n$
  - $\binom{n}{0} - \binom{n}{1} + \ldots + (-1)^n \binom{n}{n} = 0$

Geometrická definice pravděpodobnosti

Nechť

- Ω jsme schopni vyjádřit jako neprázdnou omezenou oblast v $\mathbb{R}^n$ (například pomocí omezené přímky v $\mathbb{R}^1$, omezené plochy v $\mathbb{R}^2$, omezeného tělesa v $\mathbb{R}^3$)
- jev A jsme schopni vyjádřit jako podoblast oblasti Ω

pak

$$\mathcal{P}(A) = \frac{\lambda(A)}{\lambda(\Omega)}$$

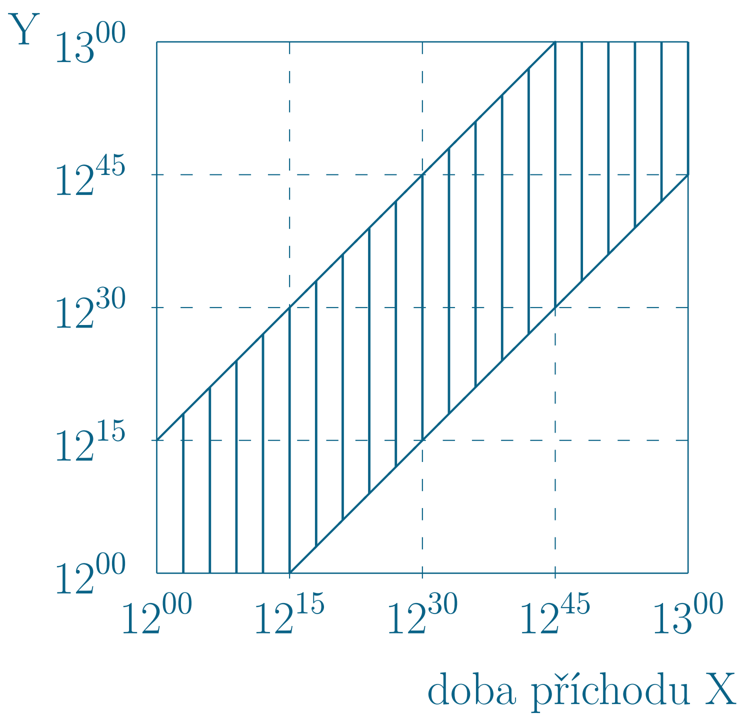
kde
$\lambda(A)$ je „míra“ oblasti A (délka, obsah plochy, objem tělesa)
$\lambda(\Omega)$ je „míra“ oblasti Ω (délka, obsah plochy, objem tělesa)

. . . .

- uvažujeme vybrané nespočetné prostory Ω
- předpokládáme, že všechny elementy Ω jsou stejně možné,
- obecněji pokud $\Omega \in \mathcal{B}^n$, mohu uvažovat λ Lebesqueovu míru

Geometrická definice pravděpodobnosti-příklad

Úloha o setkání
Dva přátelé (X a Y) se domluvili, že přijdou na určité místo v době mezi polednem a jednou hodinou odpoledne. Na místo přijde v tomto časovém intervalu každý z nich zcela náhodně a nezávisle na příchodu toho druhého. Bude čekat patnáct minut na příchod druhého, ne déle než do jedné hodiny odpoledne.
Určete ppst., že se za těchto podmínek sejdou.



Pravděpodobnost setkání odpovídá podílu obsahu vyšrafované plochy vzhledem k celkové ploše a je $\mathcal{P} = \frac{7}{16}$.

Statistická definice pravděpodobnosti

Nechť

- n je počet opakování pokusu
- $n(A)$ je počet opakování pokusu, při nichž nastal jev A

pak

$$\mathcal{P}(A) = \lim_{n \to \infty} \frac{n(A)}{n}$$

. . . .

- nepožadujeme konečnost možných výsledků pokusu
- nepožadujeme, aby všechny výsledky pokusu byly stejně možné

Vlastnosti pravděpodobnosti

Pro všechny jevy A_i, $i = 1, 2, 3, \ldots$ platí

- $0 \leq \mathcal{P}(A_i) \leq 1$
- jsou-li A_i a A_j neslučitelné, potom $\mathcal{P}(A_i \cup A_j) = \mathcal{P}(A_i) + \mathcal{P}(A_j)$
  resp. obecněji jsou-li $A_i, i = 1, 2, \ldots$ neslučitelné, potom $\mathcal{P}(\bigcup_i A_i) = \sum_i \mathcal{P}(A_i)$
- $\mathcal{P}(\Omega) = 1,\ \mathcal{P}(\emptyset) = 0$
- $A_i \subset A_j \Rightarrow \mathcal{P}(A_i) \leq \mathcal{P}(A_j)$
- $\mathcal{P}(\overline{A_i}) = 1 - \mathcal{P}(A_i)$
- $A_i \subset A_j \Rightarrow \mathcal{P}(A_j \backslash A_i) = \mathcal{P}(A_j) - \mathcal{P}(A_i)$
- $\mathcal{P}(A_i \cup A_j) \leq \mathcal{P}(A_i) + \mathcal{P}(A_j)$
- $\mathcal{P}(A_i \cup A_j) = \mathcal{P}(A_i) + \mathcal{P}(A_j) - \mathcal{P}(A_i \cap A_j)$

Axiomatická definice pravděpodobnosti

Pravděpodobnostní prostor $(\Omega, \mathcal{A}, \mathcal{P})$

- Ω je neprázdná množina

- $\mathcal{A}$ je σ-algebra podmnožin množiny Ω
  tj. $\mathcal{A}$ je systém podmnožin, pro které platí
  - $\Omega \in \mathcal{A}$
  - je-li $A \in \mathcal{A}$, pak i $\overline{A} \in \mathcal{A}$
  - jsou-li A_i, $i = 1, 2, \ldots \in \mathcal{A}$, pak i $\bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{A}$

- $\mathcal{P}$ je míra definována na $\mathcal{A}$,
  tj. $\mathcal{P}$ je funkce $\mathcal{P} : \mathcal{A} \longrightarrow \mathbb{R}$ s těmito vlastnosti
  - $\mathcal{P}(A) \geq 0$, pro $A \in \mathcal{A}$
  - $\mathcal{P}(\Omega) = 1$
  - je-li $\{A_i\}$ posloupnost (konečná nebo nekonečná) po dvou disjunktních jevů, pak $\mathcal{P}(\bigcup\limits_{i=1}^{\infty} A_i) = \sum\limits_{i=1}^{\infty} \mathcal{P}(A_i)$,

Axiomatická definice pravděpodobnosti - poznámky

- klasická definice ppsti je speciálním případem axiomatické definice, kdy $\Omega = \{\omega_i, \ldots, \omega_n\}$,$\mathcal{A}$ jsou podmnožiny množiny Ω a $\mathcal{P}(\omega_1) = \ldots = \mathcal{P}(\omega_n) = 1/n$

- ppst definovaná axiomaticky má všechny vlastnosti uváděné výše a dále platí

- je-li $A_1 \subset A_2 \ldots \subset A_n \subset \ldots$, pak $\lim\limits_{n \to \infty} \mathcal{P}(A_n) = \mathcal{P}(\bigcup\limits_{n=1}^{\infty} A_n)$

- je-li $A_1 \supset A_2 \ldots \supset A_n \supset \ldots$, pak $\lim\limits_{n \to \infty} \mathcal{P}(A_n) = \mathcal{P}(\bigcap\limits_{n=1}^{\infty} A_n)$