Pravděpodobnost a statistika

1 Náhodné pokusy a náhodné jevy

Činnostem, jejichž výsledek není jednoznačně určen podmínkami, za kterých probíhají, a které jsou (alespoň teoreticky) neomezeně opakovatelné, říkáme *náhodné pokusy.* Klasickými příklady jsou hod hrací kostkou nebo hod mincí. Náhodným pokusem však může být i sledování doby životnosti žárovky, měření hodnoty určité fyzikální veličiny, vyrobení výrobku s určitými technickými parametry, apod.

Náhodným jevem nazýváme jakékoliv tvrzení o výsledku náhodného pokusu, o kterém lze po provedení pokusu rozhodnout, zda je či není pravdivé. Například, při hodu hrací kostkou může jít o tvrzení: padne stěna se sudým počtem teček. Nebo jiný příklad, pokud náhodný pokus spočívá ve výběru pěti výrobků z celkových dvanácti, mezi nimiž jsou čtyři vadné, potom náhodným jevem může být tvrzení: mezi vybranými pěti výrobky budou právě dva vadné.

Jednotlivé výsledky náhodného pokusu obvykle označujeme jako *elementární jevy.* Množinu všech elementárních jevů nazveme *prostor elementárních jevů* a budeme ji označovat Ω. Náhodný jev je pak podmnožina (ne nutně každá) prostoru elementárních jevů. Náhodné jevy budeme označovat velkými písmeny latinské abecedy. Provádíme s nimi stejné operace jako s množinami (definujeme *sjednocení* a *průnik náhodných jevů*, *disjunktní jevy*, *doplňkový jev*, *rozdíl jevů*). *Nemožnému jevu*, který nemůže v daném pokusu nikdy nastat, odpovídá prázdná množina, *jistému jevu* celá množina Ω.

2 Pravděpodobnost

2.1 Pravděpodobnost a relativní četnost

Představme si posloupnost velkého počtu n realizací náhodného pokusu. *Relativní četnost* náhodného jevu (stručně jevu) A definujeme vztahem

$$h(A) = \frac{n(A)}{n},$$

kde $n(A)$ je počet těch realizací náhodného pokusu, při kterých nastal jev A. Pro relativní četnost náhodného jevu A platí, že

$$0 \leq h(A) \leq 1.$$

Dále, jsou-li A_1, A_2, ...,A_m navzájem disjunktní jevy a A jejich sjednocení, potom

$$h(A) = \sum_{i=1}^{m} h(A_i).$$

Na relativních četnostech je založena tzv. *statistická definice pravděpodobnosti*: *Pravděpodobnost jevu* A je číslo $P(A)$ přiřazené jevu A, které má tu vlastnost, že relativní četnost $h(A)$ jevu A se s rostoucím počtem realizací pokusu blíží číslu $P(A)$.

Protože pravděpodobnosti jsou čísla, jejichž empirickým protějškem jsou relativní četnosti, musí přiřazení pravděpodobností jevům splňovat obdobné podmínky, tzv. *axiomy pravděpodobnosti*:

(i) $0 \leq P(A) \leq 1$,

(ii) je-li A_1, A_2, ... konečný nebo spočetný systém navzájem disjunktních jevů, pak pro pravděpodobnost jeho sjednocení platí

$$P\Big(\bigcup_i A_i\Big) = \sum_i P(A_i),$$

(iii) pravděpodobnost jistého jevu je rovna jedné, tj. $P(\Omega) = 1$.

Z uvedených axiomů lze odvodit další vlastnosti pravděpodobnosti, např.

(iv) $P(\overline{A}) = 1 - P(A)$, kde $\overline{A}$ je doplňkový jev k jevu A,

(v) $P(\emptyset) = 0$,

(vi) pravděpodobnost sjednocení dvou (obecně nedisjunktních) jevů A a B je rovna

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

2.2 Pravděpodobnost v konečných prostorech

Uvažujme nyní náhodné pokusy, kterým přísluší prostor elementárních jevů s konečným počtem N prvků. Příkladem je hod hrací kostkou, kdy

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_6\}$$

a kde ω_i značí padnutí čísla i. Je-li kostka naprosto pravidelná, očekáváme, že v dlouhé řadě opakování pokusu se každý z elementárních jevů $\omega_1, \omega_2, \dots, \omega_6$ bude vyskytovat přibližně stejně často, tj. budou mít přibližně stejné relativní četnosti. Je tedy přirozené přiřadit jim stejné pravděpodobnosti. Protože pravděpodonost celého prostoru musí být rovna 1, je pak

$$P(\{\omega_i\}) = \frac{1}{6}$$

pro všechna $i = 1, \dots, 6$.

Zobecněním této úvahy získáme *klasickou definici pravděpodobnosti*: Nechť náhodnému pokusu přísluší prostor elementárních jevů s N prvky. Jestliže není důvod předpokládat, že některý elementární jev nastane spíše než jiný, pak pravděpodobnost jevu A, který je sjednocením N_A elementárních jevů, je rovna

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

Příklad. V sérii 100 výrobků je 5 vadných. Vybereme náhodně 10 výrobků. Jaká je pravděpodobnost, že mezi nimi budou právě 3 vadné?

Výsledkem náhodného pokusu je vybrání 10 výrobků z celkových 100. Počet všech možných výsledků pokusu je dán počtem všech kombinací 10 prvků ze 100, tj. číslem $\binom{100}{10}$. Počet těch výsledků, kdy právě 3 výrobky jsou vadné (ozn. jako jev A), získáme následující úvahou: Vzhledem k tomu, že série obsahuje 5 vadných výrobků, počet možností pro vybrání trojice vadných výrobků je dán číslem $\binom{5}{3}$. Ke každé takové trojici existuje $\binom{95}{7}$ možností, jak vybrat zbývajících 7 výrobků, které vadu nemají, tedy celkově $\binom{5}{3}\binom{95}{7}$ výsledků vede k tomu, že nastává jev A. Podle klasické definice pravděpodobnosti je

$$P(A) = \frac{\binom{5}{3}\binom{95}{7}}{\binom{100}{10}}.$$

2.3 Podmíněná pravděpodobnost a nezávislost jevů

Představme si nyní n realizací nějakého náhodného pokusu a uvažujme dvě množiny (náhodné jevy) A a B v prostoru elementárních jevů Ω. Z posloupnosti realizací vyberme nyní ty, při kterých nastal jev B. Takových realizací je $n(B)$. *Relativní četnost jevu A podmíněnou jevem B* potom definujeme vztahem

$$h(A|B) = \frac{n(A \cap B)}{n(B)},$$

kde $n(A \cap B)$ je počet realizací pokusu, při kterých nastaly současně oba jevy A a B. Podmíněné relativní četnosti dávají důležitou informaci o vzájemné souvislosti jevů. Například, je-li (přibližně) $h(A|B) = h(A)$, jevy A a B považujeme za nezávislé. Relativní četnost jevu A podmíněnou jevem B lze zapsat také takto:

$$h(A|B) = \frac{n(A \cap B)/n}{n(B)/n} = \frac{h(A \cap B)}{h(B)}.$$

S rostoucím počtem realizací náhodného pokusu se relativní četnosti blíží k odpovídajícím pravděpodobnostem, definujeme tedy (za předpokladu $P(B) \neq 0$) *pravděpodobnost jevu A podmíněnou jevem B* jako podíl

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

Uvažujme nyní speciální případ, kdy prostor elementárních jevů je konečný s N prvky. Omezíme-li množinu všech možných výsledků na neprázdnou množinu B, potom pravděpodobnost toho, že nastane A (za podmínky, že B nastal) je (podle klasické definice pravděpodobnosti)

$$P(A|B) = \frac{N_{A \cap B}}{N_B},$$

kde $N_{A \cap B}$ je počet výsledků, při kterých nastávají současně jevy A i B, N_B je počet všech výsledků, při kterých nastává jev B. Jednoduchou úpravou dojdeme opět k závěru, že

$$P(A|B) = \frac{N_{A \cap B}}{N_B} = \frac{N_{A \cap B}/N}{N_B/N} = \frac{P(A \cap B)}{P(B)}.$$

Pokud platí (za předpokladu $P(B) \neq 0$)

$$P(A|B) = P(A),$$

jevy A a B jsou *nezávislé*. Rozepíšeme-li tento vztah, dostaneme

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

a následně

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

Právě tímto posledním vztahem se v teorii pravděpodobnosti definuje nezávislost jevů A a B.

Pro více než dva náhodné jevy definujeme tzv. *skupinovou nezávislost náhodných jevů* následujícím způsobem: Jevy A_1, $A_2, \dots$, A_n se nazývají *navzájem nezávislé*, jestliže pro libovolnou skupinu indexů $1 \leq i_1 < i_2 \dots i_r \leq n$ platí

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_r}) = P(A_{i_1})P(A_{i_2}) \dots P(A_{i_r}).$$

Znamená to, že žádný z jevů A_1, $A_2, \dots$, A_n nezávisí na libovolném průniku jevů ostatních. Slabší je tzv. *nezávislost po dvou* (pro $r = 2$). Mohou nastat situace, že libovolné dva jevy ve skupině jsou nezávislé, ale skupina jako celek není skupinou navzájem nezávislých jevů. To lze ilustrovat na následujícím příkladě.

Příklad. V osudí jsou 4 lístky označené 4, 9, 25, 30. Nechť jev A znamená, že náhodně vytažené číslo je dělitelné 2, jev B, že náhodně vytažené číslo je dělitelné 3 a jev C, že vytažené číslo je dělitelné 5.

Platí, že $A = \{4, 30\}$, $B = \{9, 30\}$, $C = \{25, 30\}$, tedy podle klasické definice pravděpodobnosti je

$$P(A) = P(B) = P(C) = \frac{1}{2}.$$

Protože $A \cap B = B \cap C = A \cap C = \{30\}$, platí

$$\begin{aligned} P(A \cap B) = P(B \cap C) = P(A \cap C)& \\ = \frac{1}{4} = P(A)P(B) = P(B)P(C) &= P(A)P(C), \end{aligned}$$

jevy jsou tedy po dvou nezávislé. Naproti tomu,

$$P(A \cap B \cap C) = P(\{30\}) = \frac{1}{4} \neq P(A)P(B)P(C) = \frac{1}{8},$$

jevy A, B, C tedy nejsou skupinou navzájem nezávislých jevů.

2.4 Věta o úplné pravděpodobnosti

Při výpočtu pravděpodobnosti bývá někdy užitečné následující pravidlo, označované jako *věta o úplné pravděpodobnosti*:

Necht' $A \subset \bigcup_{i=1}^{n} B_i$, $P(B_i) > 0$ pro $i = 1, \ldots, n$, B_i jsou navzájem disjunktní jevy. Potom

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

Důkaz: Protože $A \subset \bigcup_{i=1}^{n}(A \cap B_i)$ a $A \cap B_i$ jsou vzájemně disjunktní, platí

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

Příklad. Uvažujme 10 osudí, v každém je 10 koulí, v i-tém osudí je i koulí černých a $10 - i$ koulí bílých. Náhodně vybereme osudí a potom z vybraného osudí náhodně jednu kouli. Jaká je pravděpodobnost, že vybraná koule je černá?

Jako A označme jev, že vytažená koule je černá , B_i necht' je jev, že bylo vybráno osudí s i černými koulemi. Podle klasické definice pravděpodobnosti lze spočítat, že pro $i = 1, \ldots, 10$ je

$$P(B_i) = \frac{1}{10}$$

a

$$P(A|B_i) = \frac{i}{10}.$$

Podle věty o úplné pravděpodobnosti platí

$$P(A) = \sum_{i=1}^{10} P(B_i)P(A|B_i) = \sum_{i=1}^{10} \frac{1}{10}\frac{i}{10} = 0,55.$$

2.5 Bayesova formule

Dalším důležitým vztahem v teorii pravděpodobnosti je tzv. *Bayesova formule*: Necht' $A \subset \bigcup_{i=1}^{n} B_i$, $P(A) > 0$, $P(B_i) > 0$ pro $i = 1, \ldots, n$, B_i jsou navzájem disjunktní jevy. Potom pro každé j platí

$$P(B_j|A) = \frac{P(B_j)P(A|B_j)}{\sum_{i=1}^{n} P(B_i)P(A|B_i)}.$$

Důkaz: Stačí použít definici podmíněné pravděpodobnosti a uvědomit si, že v čitateli je pravděpodobnost $P(B_j \cap A)$ průniku jevů A a B_j a ve jmenovateli (podle předchozí věty o úplné pravděpodobnosti) pravděpodobnost P(A) jevu A. □

Příklad. Mějme 3 osudí, v prvním jsou 4 červené a 3 bílé koule, v druhém 5 bílých a 2 modré koule a v třetím je 7 modrých a 3 červené koule. Z některého osudí jsme vybrali bílou kouli. Jaká je pravděpodobnost, že byla vybrána z prvního osudí, když vybrání koule z kteréhokoli osudí je stejně pravděpodobné?

Jako A označme jev, že byla vytažena bílá koule, B_i necht' označuje vytažení koule z i-tého osudí. Podle Bayesovy formule je

$$P(B_1|A) = \frac{P(B_1)P(A|B_1)}{\sum_{i=1}^{3} P(B_i)P(A|B_i)} = \frac{\frac{1}{3}.\frac{3}{7}}{\frac{1}{3}.\frac{3}{7} + \frac{1}{3}.\frac{5}{7} + \frac{1}{3}.0} = \frac{3}{8}.$$

2.6 Bernoulliho schéma

Uvažujme n-násobné nezávislé opakování nějakého náhodného pokusu, např. n-násobné házení hrací kostkou, mincí apod. To znamená, že libovolný jev A, který může při realizaci pokusu nastat, nezávisí na nastoupení libovolného průniku jevů při jiných realizacích (lze tedy použít pravidlo o násobení pravděpodobností).

Necht' tedy $A \subset \Omega$ je náhodný jev, který může nastat při uvažovaném náhodném pokusu, $P(A) = p$. Jako B označme jev, že při n-násobném nezávislém opakování pokusu nastal jev A právě k-krát. Potom platí

$$P(B) = \binom{n}{k} p^k (1-p)^{n-k}.$$

2.7 Axiomatická teorie pravděpodobnosti

V této kapitole zavedeme pravděpodobnost jako abstraktní matematický pojem, který je vymezen pouze svými základními vlastnostmi, tzv. axiomy pravděpodobnosti (Kolmogorov, 1933). Pojem pravděpodobnosti tak přestane být vázán na praktické experimenty. Rovněž předpoklad konečného prostoru elementárních jevů, tak jak to vyžadovala klasická definice pravděpodobnosti, nebude potřebný. Předchozí definice pravděpodobnosti, tedy statistická a klasická definice, slouží spíš k praktickému určení pravděpodobnosti sledovaného jevu.

Pravděpodobnost se v axiomatické teorii formálně definuje jako zobrazení, které podmnožinám základního prostoru Ω (náhodným jevům) přiřazuje číslo z intervalu $[0,1]$, splňující určité základní podmínky. Přitom není nutné, abychom každou podmnožinu množiny Ω považovali za náhodný jev, kterému přísluší určitá pravděpodobnost. Důležité je, abychom s danými podmnožinami (jevy), které jsou v dané situaci důležité a jejichž pravděpodobnosti nás zajímají (což může být i otázka subjektivní volby), brali v úvahu a uměli přiřadit pravděpodobnost i těm podmnožinám (jevům), které mohou z již uvažovaných vzniknout použitím nějaké poslupnosti základních množinových (logických) operací (doplňky, sjednocení, apod.).

Definice. Nechť Ω je neprázdná množina a $\mathcal{S}$ je systém podmnožin množiny Ω. Řekneme, že $\mathcal{S}$ je *σ-algebra*, jestliže

(i) $\Omega \in \mathcal{S}$

(ii) $A \in \mathcal{S} \Rightarrow \overline{A} = \Omega \backslash A \in \mathcal{S}$

(iii) $A_n \in \mathcal{S},\ n = 1, 2, \ldots \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{S}$

Z uvedených axiomů lze odvodit další vlastnosti systému $\mathcal{S}$.

Tvrzení. Nechť $\mathcal{S}$ je σ-algebra podmnožin množiny Ω. Potom platí:

(i) $\emptyset \in \mathcal{S}$

(ii) $A, B \in \mathcal{S} \Rightarrow A \cup B, A \cap B, A \backslash B \in \mathcal{S}$

(iii) $A_n \in \mathcal{S},\ n = 1, 2, \ldots \Rightarrow \bigcap_{n=1}^{\infty} A_n \in \mathcal{S}$

Definice. *Pravděpodobnost* je zobrazení $P : \mathcal{S} \to [0,1]$ definované na σ-algebře $\mathcal{S}$ podmnožin množiny Ω, pro které platí:

(i) $P(\Omega) = 1$

(ii) $A_n \in \mathcal{S}$, $n = 1, 2, \ldots$ a $A_n \cap A_m = \emptyset$ pro $n \neq m \Rightarrow P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$

Trojice $(\Omega, \mathcal{S}, P)$ se nazývá *pravděpodobnostní prostor.* Uvedeme zde dva jednoduché příklady pravděpodobnostních prostorů:

1. Necht' $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$ a $\mathcal{S}$ je systém všech podmnožin množiny Ω. Pro $A \in \mathcal{S}$ položíme

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

kde N_A je počet prvků množiny A. Trojice $(\Omega, \mathcal{S}, P)$ je pravděpodobnostní prostor.

2. Necht' $\Omega = [0, 2] \times [0, 2]$ je čtverec v rovině, definujme množinu

$$A = \left\{ (x, y) \in \Omega | 0 \leq y \leq \frac{x^2}{2} \right\}.$$

Jaká je pravděpodobnost, že náhodně zvolený bod čtverce padne do množiny A?

V daném příkladě stačí vzít $\mathcal{S} = \{\emptyset, A, \overline{A}, \Omega\}$ a pro $B \in \mathcal{S}$ položit

$$P(B) = \frac{m(B)}{m(\Omega)},$$

kde $m(B)$ a $m(\Omega)$ jsou plošné obsahy množin B a Ω. Snadným výpočtem zjistíme, že $m(B) = \frac{4}{3}$. Protože $m(\Omega) = 4$, dostáváme

$$P(A) = \frac{1}{3}.$$

Poznámka. V uvedeném příkladě jsme použili tzv. *geometrickou definici pravděpodobnosti.*

Z uvedených axiomů σ-algebry a pravděpodobnosti lze odvodit některé další vlastnosti:

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

(ii) je-li $A_1, \ldots, A_n$ konečná posloupnost navzájem disjunktních množin z $\mathcal{S}$, potom $P(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$

(iii) je-li $A \in \mathcal{S}$, potom $P(\overline{A}) = 1 - P(A)$

(iv) jsou-li $A, B \in \mathcal{S}$, potom $P(A \backslash B) = P(A) - P(A \cap B)$

(v) jsou-li $A, B \in \mathcal{S}$, potom $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(vi) jsou-li $A, B \in \mathcal{S}$ a $A \subset B$, potom $P(A) \leq P(B)$

(vii) jsou-li $A, B \in \mathcal{S}$ a $A \subset B$, potom $P(B \backslash A) = P(B) - P(A)$

(viii) jestliže $A_n \in \mathcal{S}$ a $A_n \subset A_{n+1}$ pro $n = 1, 2, \ldots,$ potom

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n \to \infty} P(A_n)$$

(ix) jestliže $A_n \in \mathcal{S}$ a $A_{n+1} \subset A_n$ pro $n = 1, 2, \ldots,$ potom

$$P\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n \to \infty} P(A_n)$$

Důkazy uvedených tvrzení lze najít v doporučené studijní literatuře (Z. Riečanová a kolektív: Numerické metódy a matematická štatistika).

Poznámka. V kapitole 2.1 jsme do systému axiomů zahrnuli i konečnou aditivitu (v souladu s tím, jak se to někdy dělává). Tento předpoklad nám umožňuje bezprostředně snadno odvodit zde uvedenou vlastnost (iii) a následně vlastnost (i). Nicméně, stačí předpokládat pouze spočetnou aditivitu, konečná aditivita je pak jejím důsledkem (vlastnost (ii)). Napřed je ale nutno dokázat, že $P(\emptyset) = 0$ (viz Z. Riečanová a kolektív: Numerické metódy a matematická štatistika).