Pravděpodobnost

Kombinatorika pro bioinformatiky

4.11.2020

1 Pravděpodobnostní prostory

Definice 1 *Konečný pravděpodobnostní prostor je trojice* $(\Omega, \mathcal{F}, P)$*, kde* Ω *je konečná množina,* $\mathcal{F} = 2^{\Omega}$ *a* $P : \mathcal{F} \mapsto [0, 1]$ *splňující pro každou* $A \subseteq \Omega$

$$P(A) = \sum_{\omega \in \Omega} P(\{\omega\})$$

a $P(\Omega) = 1$.

Příklady pravděpodobnostních prostorů:

Hod (spravedlivou) mincí: $\Omega = \{R, L\}$, $P(R) = P(L) = \frac{1}{2}$.

Hod kostkou: $\Omega = \{1, 2, 3, 4, 5, 6\}$, $P(k) = \frac{1}{6}$ pro $k = 1, 2, 3, 4, 5, 6$.

Opakovaný hod mincí: $\Omega = \{R, L\}^n$, $P(\omega) = \frac{1}{2^n}$ pro $\forall \omega \in \{R, L\}^n$.

Náhodná permutace: $\Omega = S_n$, $P(\pi) = \frac{1}{n!}$ pro všechny $\pi \in S_n$.

Definice 2 *Pravděpodonostní prostor* $(\Omega, 2^{\Omega}, P)$ *se nazývá uniformní (též klasický) pravděpodobnostní prostor, pokud* $P(A) = \frac{|A|}{|\Omega|}$ *pro všechny jevy* $A \subseteq \Omega$.

2 Podmíněná pravděpodobnost

Definice 3 *Jevy A a B se nazývají nezávislé, pokud $P(A \cap B) = P(A) \cdot P(B)$. Obecně, jevy $A_1, A_2, \ldots, A_n$ jsou nezávislé, pokud pro každou (neprázdnou) podmnožinu $I \subseteq \{1, 2, \ldots, n\}$ platí $P(\bigcap_{i \in I} A_i) = \prod_{i \in I} P(A_i)$.*

Definice 4 *Podmíněná pravděpodobnost jevu A za předpokladu, že nastal jev B, se definuje jako*

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

(je def. pouze pro $P(B) > 0$).

Pozorování 5 *Jsou-li jevy A, B nezávislé a $P(B) > 0$, pak $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$. A naopak, pokud $P(A|B) = P(A)$, jedná se o nezávislé jevy.*

Příklad: Uvažujme, že v nějaké populaci je četnost bezejmenné choroby 1 osoba z 1000. K dispozici máme test, jehož spolehlivost je 95% (u skutečně nemocné osoby vyjde test s pravděpodobností 95% pozitivně, a taktéž u zdravé osoby vyjde test s pravděpodobností 95% negativně). Jaké je pravděpodobnost, že testovaná osoba je skutečně nemocná, pokud jí vyšel pozitivní test?

Zavedeme si značení: N bude označovat jev, že daná osoba je nemocná, a $\overline{N}$ jev, že nemocná není. Analogicky, T značí pozitivní test a $\overline{T}$ negativní test. Zadání pak znamená $P(N) = \frac{1}{1000}, P(T|N) = P(\overline{T}|\overline{N}) = \frac{19}{20}$ a zajímá nás $P(N|T)$.

Podle definice podmíněné pravděpodobnosti je

$$P(N|T) = \frac{P(N \cap T)}{P(T)} \ .$$

Čitatele lze vyjádřit z

$$P(T|N) = \frac{P(T \cap N)}{P(N)}$$

jako

$$P(N \cap T) = P(T|N) \cdot P(N) = \frac{19}{20} \cdot \frac{1}{1000} = \frac{19}{20000} \ .$$

Jmenovatele v řešeném výrazu spočteme

$$\begin{aligned} P(T) &= P(T|N) \cdot P(N) + P(T|\overline{N}) \cdot P(\overline{N}) = \\ &= P(T|N) \cdot P(N) + (1 - P(\overline{T})\overline{N})) \cdot P(\overline{N}) = \frac{19}{20} \cdot \frac{1}{1000} + \frac{1}{20} \cdot \frac{999}{1000} = \frac{1018}{20000} \ . \end{aligned}$$

Nyní již jen dosadíme

$$P(N|T) = \frac{P(N \cap T)}{P(T)} = \frac{\frac{19}{20000}}{\frac{1018}{20000}} = \frac{19}{1018} < \frac{20}{1000} = 0.02 \ .$$

Věta 6 (Věta o úplné pravděpodobnosti) *Nechť A je nějaký jev a $B_1, \ldots, B_n$ jsou disjunktní jevy pokrývající celé Ω, navíc $P(B_i) > 0$ pro $i = 1, \ldots, n$. Pak*

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

Důkaz: Protože jevy $B_1, \ldots, B_n$ pokrývají Ω, platí pro každé A

$$A = \bigcup_{i=1}^{n} (A \cap B_i) \ .$$

Z jejich disjunktnosti plyne, že i členy $A \cap B_i$ jsou disjunktní a tedy $P(A) = \sum_{i=1}^{n} P(A \cap B_i)$, což dále upravujeme s využitím vztahu

$$P(A \cap B_i) = P(A|B_i) \cdot P(B_i) \ .$$

A dotáváme

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

□

Věta 7 (Bayesova věta) *Nechť A je nějaký jev a $B_1, \ldots, B_n$ jsou disjunktní jevy pokrývající celé Ω, navíc $P(B_i) > 0$ pro $i = 1, \ldots, n$. Pak*

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

Důkaz: Z def. podmíněné pravděpodonosti vyjádříme pravděpodobnost průniku dvou jevů oběma možnými způsoby

$$P(A \cap B_k) = P(A|B_k) \cdot P(B_k) = P(B_k|A) \cdot P(A)$$

Z toho

$$P(B_k|A) = \frac{P(A|B_k) \cdot P(B_k)}{P(A)}$$

a za jmenovatele zlomku dosadíme z Věty o úplné pravděpodobnosti.

□

Definice 8 *Součin pravděpodobnostních prostorů $(\Omega_1, 2^{\Omega_1}, P_1)$ a $(\Omega_2, 2^{\Omega_2}, P_2)$ je prostor $(\Omega, 2^{\Omega}, P)$, kde $\Omega = \Omega_1 \times \Omega_2$ a pro $A \subseteq \Omega$ je*

$$P(A) = \sum_{(\omega_1, \omega_2) \in A} P_1(\{\omega_1\}) \cdot P_2(\{\omega_2\}) \ .$$

3 Náhodné veličiny

Definice 9 *(Reálná) náhodná veličina na konečném pravděpodobnostním prostoru* $(\Omega, 2^\Omega, P)$ *je libovolná funkce* $X : \Omega \mapsto \mathbb{R}$*.*

Příklad: Pro $n \in \mathbb{N}$ definujeme náhodnou veličinu (používá se též pojmenování náhodná proměnná) X_n rovnu počtu líců při n hodech spravedlivou mincí. To zamená, že $\Omega_n = \{L, R\}^n$ a X_n nabývá hodnot $0, 1, \ldots, n$, přičemž

$$P(X_n = k) = \frac{\binom{n}{k}}{2^n} \ .$$

Definice 10 *Buď* $X : \Omega \mapsto \mathbb{R}$ *náhodná veličina na konečném pravděpodobnostním prostoru* $(\Omega, 2^\Omega, P)$*. Střední hodnota* X *je definována jako*

$$E(X) = \sum_{\omega \in \Omega} P(\{\omega\}) \cdot X(\omega) \ .$$

Věta 11 (Linearita střední hodnoty) *Nechť* X, Y *jsou náhodné reálné veličiny na (konečném) pravděpodobnostním prostoru* $(\Omega, 2^\Omega, P)$*,* $\alpha \in \mathbb{R}$*. Potom*

$$\begin{aligned} E(\alpha X) &= \alpha E(X) \\ E(X+Y) &= E(X) + E(Y) \end{aligned}$$

Důkaz: Plyne přímo rozepsáním definice:

$$\begin{aligned} E(\alpha X) &= \sum_{\omega \in \Omega} P(\{\omega\}) \cdot (\alpha X)(\omega) = \sum_{\omega \in \Omega} \alpha P(\{\omega\}) \cdot X(\omega) = \\ &= \alpha \sum_{\omega \in \Omega} P(\{\omega\}) \cdot X(\omega) = \alpha E(X) \end{aligned}$$

a

$$\begin{aligned} E(X+Y) &= \sum_{\omega \in \Omega} P(\{\omega\}) \cdot (X+Y)(\omega) = \sum_{\omega \in \Omega} P(\{\omega\}) \cdot (X(\omega) + Y(\omega)) = \\ &= \sum_{\omega \in \Omega} P(\{\omega\}) \cdot X(\omega) + \sum_{\omega \in \Omega} P(\{\omega\}) \cdot Y(\omega) = E(X) + E(Y) \ . \end{aligned}$$

$\square$

Definice 12 *Indikátor jevu* $A \subseteq \Omega$ *je náhodná veličina* $I_A : \Omega \mapsto \{0, 1\}$ *definovaná*

$$I_A(\omega) = \begin{cases} 1 & \text{pro } \omega \in A \\ 0 & \text{pro } \omega \notin A \end{cases} \ .$$

Pozorování 13 *Pro střední hodnotu indikátoru platí*

$$E(I_A) = P(A) \ .$$

Příklad: Použití indikátorů: Pro náhodnou veličinu z předchozího příkladu (počet líců při n hodech mincí) spočtěte její středn hodnotu.

Definujme jevy A_i: při i-tém hody padl líc, pro $i = 1, \ldots, n$. Potom $P(A_i) = \frac{1}{2}$ pro každé i a dále $X_n = \sum_{i=1}^{n} I_i$ (Pro přehlednější zápis označujeme indikátor jevu A_i jako I_i).

Z toho

$$E(X_n) = E(\sum_{i=1}^{n} I_i) = \sum_{i=1}^{n} E(I_i) = \sum_{i=1}^{n} P(A_i) = \frac{n}{2} \ .$$

Poznamejme, že ke stejnému výsledku lze dospět i přímo z definice střední hodnoty, takto je to ale jednodušší.

Věta 14 (Markovova nerovnost) *Pro nézapornou náhodnou veličinu X splňující $E(X) > 0$ a reálné $t > 1$ platí*

$$P(X \geq tE(X)) \leq \frac{1}{t} \ .$$

Důkaz: Nejprve ukážeme, že $E(X) \geq aP(X \geq a)$ a pak zvolíme vhodnou hodnotu paramtru a.

Označme $A \subseteq \Omega$ jev, že hodnota X je větší rovna a, tedy $A = \{\omega \in \Omega | X(\omega) \geq a\}$. Pak

$$\begin{aligned} E(X) &= \sum_{\omega\in\Omega} P(\{\omega\}) \cdot X(\omega) = \sum_{\omega\in A} P(\{\omega\}) \cdot X(\omega) + \sum_{\omega\in\Omega\setminus A} P(\{\omega\}) \cdot X(\omega) \geq \\ &\geq \sum_{\omega\in A} P(\{\omega\}) \cdot a + 0 = a \cdot \sum_{\omega\in A} P(\{\omega\}) = a \cdot P(A) = aP(X \geq a) \ . \end{aligned}$$

Nyní dosadíme $a = tE(X)$ a upravíme

$$E(X) \geq (tE(X))P(X \geq tE(X))$$

na požadovaný výsledek

$$P(X \geq tE(X)) \leq \frac{1}{t} \ .$$

$\square$

Definice 15 *Rozptyl náhoné veličiny X je definován jako*

$$Var(X) = E((X - E(X))^2) \ .$$

Věta 16 (Čebyševova nerovnost) *Pro náhodnou veličinu X a reálné číslo $a \geq \sqrt{Var(X)}$ platí*

$$P(|X - E(X)| \geq a) \leq \frac{Var(X)}{a^2} \ .$$

Důkaz: Uvažme novou náhodnou veličinu $Y = (X - E(X))^2$. Dále si uvědomíme, že podmínka $|X - E(X)| \geq a$ je splněna právě tehdy, když $(X - E(X))^2 \geq a^2$, neboli $Y \geq a^2$.

Protože Y je nezáporná náhodná veličina, lze použít Markovovu nerovnost s volbou $t = \frac{a^2}{Var(X)}$. Případ, kdy je $Var(X) = 0$, tedy X je konstatně rovna $E(X)$, si musíme rozmyslet zvlášť, ale Čebyševova nerovnost pak jistě platí.

Markovova nerovnost nám pro Y dává

$$P(Y \geq tE(Y)) \leq \frac{1}{t} \ ,$$

kde $E(Y) = E((X - E(X))^2) = Var(X)$ a při výše uvedené volbě t je levá strana

$$\begin{aligned} P(Y \geq tE(Y)) &= P((X - E(X))^2 \geq \frac{a^2}{Var(X)} \cdot Var(X)) = P((X - E(X))^2 \geq a^2) = \\ &= P(|X - E(X)| \geq a) \end{aligned}$$

a pravá strana je rovna

$$\frac{1}{t} = \frac{Var(X)}{a^2} \ .$$

Tím je důkaz hotov.

$\square$