Společná část: 19 - PSI

Náhodná veličina a náhodný vektor. Distribuční funkce, hustota a pravděpodobnostní funkce náhodné veličiny. Střední hodnota a rozptyl náhodné veličiny a jejich odhady. Sdružené charakteristiky náhodného vektoru. Korelace a nezávislost náhodných veličin. Metoda maximální věrohodnosti. Základní principy statistického testování hypotéz. Markovské řetězce, klasifikace stavů.

1 Základní pojmy pravděpodobnosti

1.1 Laplaceova (klasická) pravděpodobnost

- **Náhodný pokus** má $n \in \mathbb{N}$ různých, vzájemně se vylučujících výsledků, které jsou stejně možné.
- **Elementární jevy** = výsledky nahodného pokusu
- **Množina všech elementárńıch jev̊u**: Ω
- **Jev** je podmnožina všech elementárních jevů ($A \subseteq \Omega$)
- **Pravděpodobnost jevu** A :
$$P(A) = \frac{\mid A \mid}{\mid \Omega \mid}$$
- **Jevové pole**: všechny jevy pozorovatelné v náhodném pokusu, zde expΩ (=množina všech podmnožin množiny Ω)

1.2 Kolmogorovova pravděpododobnost

- **Elementárních jevů** (=prvků množiny Ω) může být nekonečně mnoho, nemusí být stejně pravděpodobné
- **Jevy** jsou podmnožiny množiny Ω, ale ne nutně všechny. Tvoří podmnožinu $\mathcal{A} \subseteq exp\Omega$, která splňuje podmínky σ-algebry (viz. 1.3).
- **Pravděpodobnost** není určena strukturou jevů jako u Laplaceova modelu, je to funkce $P : \mathcal{A} \to \langle 0, 1\rangle$, splňující podmínky:
(P1) $P(\mathbf{1}) = 1$,
(P2)$P\left(\bigcup\limits_{n\in\mathbb{N}} A_n\right) = \sum\limits_{n\in\mathbb{N}} P(A_n)$, pokud jsou množiny (=jevy) A_n, $n \in \mathbb{N}$, po dvou neslučitelné
- **Pravděpodobnostní prostor** je trojice $(\Omega, \mathcal{A}, P)$, kde Ω je neprázdná množina, $\mathcal{A}$ je σ-algebra podmnožin množiny Ω a P je pravděpodobnost.

1.3 σ-algebra

σ-algebra je teoretický koncept výběru jistých podmnožin dané množiny, který splňuje pevně definované podmínky. Koncept σ-algebry umožňuje například zavést míru, čehož se dále využívá zejména v matematické analýze k budování pojmu integrál a právě v teorii pravděpodobnosti [wikipedia]. Systém podmnožin $\mathcal{A}$ nějáké množiny Ω musí splňovat podmínky:

1. $\emptyset \in \mathcal{A}$
2. $A \in \mathcal{A} \Rightarrow \bar{A} \in \mathcal{A}$ (uzavřenost vůči doplňku)

3. $(\forall n \in \mathbb{N} : A_n \in \mathcal{A}) \Rightarrow \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A})$ (uzavřenost vůči sjednocení)

Nejmenší σ-algebra podmnožin $\mathbb{R}$, která obsahuje všechny intervaly, se nazývá **Borelova σ-algebra**. Obsahuje všechny intervaly otevřené, uzavřené i polouzavřené, i jejich spočetná sjednocení, a některé další množiny, ale je menší´než exp$\mathbb{R}$. Její prvky nazýváme borelovské množiny.

2 Náhodná veličina a náhodný vektor

2.1 Náhodná veličina

Je na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ měřitelná funkce $X : \Omega \to \mathbb{R}$ (přiřazuje každému jevu jevového pole reálné číslo [wikipedia]).
Náhodné veličiny lze rozdělit na nespojité (diskrétní) a spojité. Diskrétní veličiny mohou nabývat pouze spočetného počtu hodnot (konečného i nekonečného), zatímco spojité veličiny nabývají hodnoty z nějakého intervalu (konečného nebo nekonečného) [wikipedia].
Příklad: Havárii auta označíme cenou škody. Nebo strany hrací kostky označíme čísly 1 až 6.
Pro každý interval I platí

$$X^{-1}(I) = \{\omega \in \Omega \mid X(\omega) \in I\} \in \mathcal{A}$$

Je popsána **pravděpodobnostní funkcí**:

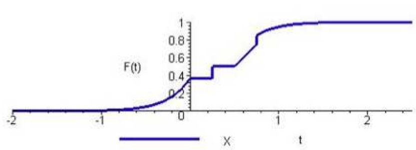
$$P_X(I) = P[X \in I] = P(\{\omega \in \Omega \mid X(\omega) \in I\})$$

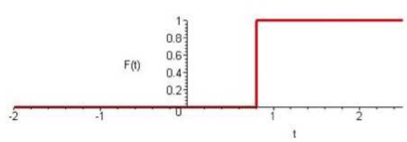
Místo pravděpodobnostní funkce, můžeme použít úspornější **Distribuční funkci** (F_X), která se omezuje na intervaly tvaru $I = (-\infty, t\rangle, t \in \mathbb{R}$

$$P[X \in (-\infty, t\rangle] = P[X \leq t] = P_X((-\infty, t\rangle) = F_X(t)$$

Různými kombinacemi distribuční funkce ($F_X : \mathbb{R} \to \langle 0, 1\rangle$) můžeme plně nahradit pravděpodobnostní funkci.
Vlastnosti distribuční funkce:

- neklesající
- zprava spojitá
- $\lim_{t\to-\infty} F_x(t) = 0, \lim_{t\to\infty} F_x(t) = 1$





2.2 Náhodný vektor (n-rozměrná náhodná veličina)

Je na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ měřitelná funkce $\boldsymbol{X} : \Omega \to \mathbb{R}^n$. Používáme ho v případech, kdy je k popisu výsledku náhodného pokusu nutné použít více čísel [wikipedia].
Pro každý n-rozměrný interval I platí

$$\boldsymbol{X}^{-1}(I) = \{\omega \in \Omega \mid \boldsymbol{X}(\omega) \in I\} \in \mathcal{A}$$

Lze psát

$$\boldsymbol{X}(\omega) = (X_1(\omega), \ldots, X_n(\omega)),$$

kde zobrazení $X_k : \Omega \to \mathbb{R}, k = 1, \ldots, n$ jsou náhodné veličiny.
Náhodný vektor lze považovat za vektor náhodných veličin $\boldsymbol{X} = (X_1, \ldots, X_n)$.
Je popsán **pravděpodobnostmi**:

$$P_{\boldsymbol{X}}(I_1 \times \ldots \times I_n) = P[X_1 \in I_1, \ldots, X_n \in I_n] = P(\{\omega \in \Omega \mid X_1(\omega) \in I_1, \ldots, X_n(\omega) \in I_n\}),$$

kde $I_1, \ldots, I_n$ jsou intervaly v $\mathbb{R}$. Z toho výplývá pravděpodobnost pro libovolnou borelovskou množinu I v $\mathbb{R}^n$

$$P_{\boldsymbol{X}}(I) = P[\boldsymbol{X} \in I] = P(\{\omega \in \Omega \mid \boldsymbol{X}(\omega) \in I\})$$

Opět můžeme použít úspornější **distribuční funkci ($F_{\boldsymbol{X}}$)**

$$P[X_1 \in (-\infty, t_1\rangle, \dots, X_n \in (-\infty, t_n\rangle] = P[X_1 \leq t_1, \dots, X_n \leq t_n] = P_{\boldsymbol{X}}((-\infty, t_1\rangle \times \dots \times (-\infty, t_n\rangle) = F_{\boldsymbol{X}}(t_1, \dots, t_n)$$

Vlastnosti distribuční funkce:

- neklesající (ve všech proměnných)
- zprava spojitá (ve všech proměnných)
- $\lim\limits_{t_1 \to \infty, \dots, t_n \to \infty} F_{\boldsymbol{X}}(t_1, \dots, t_n) = 1$
- $\lim\limits_{t_1 \to -\infty, \dots, t_n \to -\infty} F_{\boldsymbol{X}}(t_1, \dots, t_n) = 0$

2.3 Obecné náhodné veličiny

Náhodné veličiny nemusí být reprezentovány pouze reálným čísly, ale třeba i čísly komplexními. V některých případech se používájí i jiné něž numerické hodnoty, například “rub”, “líc”, ”kámen”, ”papír” atp.