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 $P : \mathcal{A} \rightarrow \langle 0, 1 \rangle$, 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. $(\forall n \in \mathbb{N} : A_n \in \mathcal{A}) \Rightarrow \bigcup\limits_{n\in\mathbb{N}} A_n \in \mathcal{A})$ (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

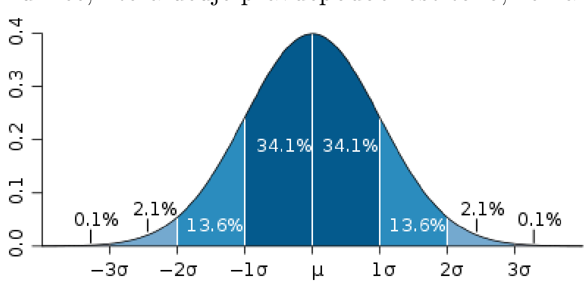
Je na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ měřitelná funkce $X : \Omega \to \mathbb{R}$ (přiřazuje každému jevu jevového pole reálné číslo [wikipedia]).
Náhodné veličiny lze rozdělit na nespojité (diskrétní) a spojité. Diskrétní veličiny mohou nabývat pouze spočetného počtu hodnot (konečného i nekonečného), zatímco spojité veličiny nabývají hodnoty z nějakého intervalu (konečného nebo nekonečného) [wikipedia].
Příklad: Havárie aut označíme cenou škody a můžeme se ptát, jak je pravděpodobné, že havárie dosáhne určité škody.
Pro každý interval I platí

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

Popisuje ji **Rozdělení pravděpodobnosti** náhodné veličiny X:

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

to je funkce, která učuje pravděpodobnost toho, že náhodná veličina nabyde určité hodnoty.

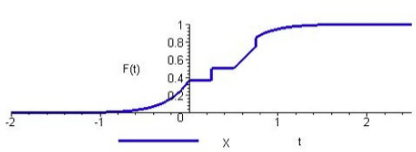


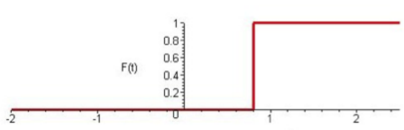
Místo pravděpodobnostní funkce, můžeme použít úspornější **Distribuční funkci** (F_X), která se omezuje na intervaly tvaru $I = (-\infty, t\rangle, t \in \mathbb{R}$

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

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

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





Distribuční funkce pro absolutně spojitou veličinu:

$$F_X(t) = \int_{-\infty}^{t} f_X(u)du,$$

kde f_X je tzv. **hustota náhodné veličiny**. Je to nezáporná funkce ($f_X : \mathbb{R} \to \langle 0, \infty))$ a splňuje $\int\limits_{-\infty}^{\infty} f_X(u)du = 1$

Náhodné veličiny $X_1, \ldots, X_n$ jsou **nezávislé**, pokud pro všechny intervaly $I_1, \ldots, I_n$ jsou jevy $X_1 \in I_1, \ldots, X_n \in I_n$ nezávislé, tj.

$$P[X_1 \in I_1, \ldots, X_n \in I_n] = \prod_{i=1}^{n} P[X_i \in I_i]$$

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].
Příklad: Chceme popsat vztah např. mezi výškou a váhou osob. K tomu potřebujeme více informací nežjen popis jednotlivých náhodných veličin.
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 **Sdruženým rozdělením pravděpodobnosti**:

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

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 **Sdruženou distribuční funkci ($F_{\boldsymbol{X}}$)**

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

Vlastnosti distribuční funkce:

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

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.

3 Základní charakteristiky náhodných veličin a náhodných vektorů

3.1 Střední hodnota

Značení E nebo μ. Jedná se o tzv. charakteristiku polohy (angl. measures of central tendency)

3.1.1 Střední hodnota náhodné veličiny

Je definována zvlášť pro:

- **diskrétní** náhodnou veličinu U s oborem hodnot $\mathbb{R}$:

$$EU = \mu_U = \sum_{t \in \mathbb{R}} t \cdot p_U(t)$$

- **spojitou** náhodnou veličinu V:

$$EV = \mu_V = \int_{-\infty}^{\infty} t \cdot f_V(t) dt$$

3.1.2 Střední hodnota náhodného vektoru

$$E\boldsymbol{X} = (EX_1, \ldots, EX_n)$$

3.2 Rozptyl (disperze)

Značení σ^2 , D , var. Jedná se o charakteristiku variability (angl. measures of central tendency).

3.2.1 Rozptyl náhodné veličiny

Je to vlastně střední hodnota kvadrátu odchylky od střední hodnoty.

$$DX = E\left((X - EX)^2\right) = E\left(X^2\right) - (EX)^2$$

3.2.2 Rozptyl náhodného vektory

$$D\boldsymbol{X} = (DX_1, \ldots, DX_n)$$

3.3 Další číselné charakteristiky náhodného vektoru

Pro jednorozměrné náhodné veličiny střední hodnota a rozptyl dávají **dostatečnou** informaci pro výpočet rozptylu jeho lineárních funkcí (lineární kombinace různých náhodných veličin):

$$E(X + Y) = EX + EY$$
$$E(X - Y) = EX - EY$$
$$D(X + Y) = DX + DY$$
$$D(X - Y) = DX + DY$$

Pro náhodný vektor to ale nestačí, a proto zavádíme další charakteristiky:

$$E(X + Y) = EX + EY$$

$$D(X + Y) = DX + DY + 2cov(X, Y),$$

kde $cov(X, Y)$ je kovariance náhodných veličin X, Y.

3.3.1 Kovariace a korelace

Kavariace určuje míru statistické závislosti mezi náhodnými veličinami. Je definována jako střední hodnota součinu odchylek obou náhodných veličin X, Y od jejich středních hodnot (druhý vzorec není tak srozumitelný, ale je jednodušší pro výpočet)

$$cov(X,Y) = E((X-EX)(Y-EY)) = E(XY) - EXEY$$

Vlastnosti kovariace:

$$cov(X,X) = DX,$$
$$cov(Y,X) = cov(X,Y)$$
$$cov(aX+b, cY+d) = a\,c\,cov(X,Y)$$
$$cov(X,Y) = 0 - pro\ nezávislé\ veličiny$$

Při výpočtech je místo kovariace výhodnější používat **korelaci** (což je kovariace pro normované náhodné veličiny)

$$\varrho(X,Y) = cov(norm\,X, norm\,Y) = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = E(norm\,X \cdot norm\,Y)$$

Korelace nabývá hodnot $\langle$-1,-1$\rangle$, pro $\varrho = 1$ je mezi X, Y přímá lineární závislost, pro $\varrho = -1$ nepřímá lineární závislost. Pro $\varrho = 0$ říkáme, že jsou veličiny **nekorelované**. Zároveň to znamená, že jsou lineárně nezávislé, nikoliv obecně nezávislé (je to nutná podmínka pro obecnou nezávislost, ale nikoliv postačující).
Pro náhodný vektor $\boldsymbol{X} = (X_1, \ldots, X_2)$ definujeme **kovariační matici:**

$$\Sigma_{\boldsymbol{X}} = \begin{bmatrix} cov(X_1,X_1) & cov(X_1,X_2) & \ldots & cov(X_1,X_n) \\ cov(X_2,X_1) & cov(X_2,X_2) & \ldots & cov(X_2,X_n) \\ \vdots & \vdots & \ddots & \vdots \\ cov(X_n,X_1) & cov(X_n,X_2) & \ldots & cov(X_n,X_n) \end{bmatrix} = \begin{bmatrix} DX_1 & cov(X_1,X_2) & \ldots & cov(X_1,X_n) \\ cov(X_2,X_1) & DX_2 & \ldots & cov(X_2,X_n) \\ \vdots & \vdots & \ddots & \vdots \\ cov(X_n,X_1) & cov(X_n,X_2) & \ldots & DX_n \end{bmatrix}$$

a **korelační matici:**

$$\varrho_{\boldsymbol{X}} = \begin{bmatrix} 1 & \varrho(X_1,X_2) & \ldots & \varrho(X_1,X_n) \\ \varrho(X_2,X_1) & 1 & \ldots & \varrho(X_2,X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \varrho(X_n,X_1) & \varrho(X_n,X_2) & \ldots & 1 \end{bmatrix}$$