

Použití Bayesova vzorce pro odhad parametru polohy diskrétní náhodné veličiny

Dáno:

- Fyzikální veličina x , která je přístupná přímému měření. Nazýváme ji **měřená veličina**. Předpokládáme, že hodnoty veličiny x nabývají jen celočíselných hodnot, tj. $x \in \mathbb{Z}$. Kdyby tento předpoklad nebyl splněn, zvolíme jinou jednotku pro měření veličiny x tak, aby $x \in \mathbb{Z}$.
- Očekávaný rozsah měřených hodnot veličiny x a pravděpodobnosti jejich výskytu. Tím je dána tzv. **apriorní pravděpodobnostní funkce**. Označíme ji p_{aprior} .
- Několik naměřených hodnot veličiny x . Označíme je $x_1, x_2, \dots, x_n, n \in \mathbb{N}, x_i \in \mathbb{Z}$ pro $\forall i \in \{1, 2, \dots, n\}$. Měření hodnot x_i bylo provedeno vzájemně nezávisle.

Měření veličiny x je vždy ovlivněno nevyhnutelnými měřickými chybami $\varepsilon_i \in \mathbb{Z}, i \in \{1, 2, \dots, n\}$. Platí pro ně

$$x_i = x^* + \varepsilon_i . \quad (1)$$

Symbol x^* představuje **skutečnou hodnotu** měřené veličiny x . Hodnota x^* je neznámá, v podstatě nepoznatelná, protože měření veličiny x je vždy ovlivněno nevyhnutelnými měřickými chybami. Předpokládáme $x^* \in \mathbb{Z}$.

- Diskrétní náhodná veličina \mathcal{E}_X příslušná chybám měření veličiny x . Všechny možné chyby měření veličiny x tvoří tzv. **základní soubor** náhodné veličiny \mathcal{E}_X . Označíme ho $\mathcal{H}(\mathcal{E}_X)$. Platí pro něj $\mathcal{H}(\mathcal{E}_X) \subset \mathbb{Z}$, neboť o chybách měření předpokládáme, že nabývají jen celočíselných hodnot. Kdyby tento předpoklad nebyl splněn, zvolíme jinou jednotku pro měření veličiny x tak, aby $x \in \mathbb{Z}$.

Pravděpodobnostní funkci náhodné veličiny \mathcal{E}_X označíme $p_{\mathcal{E}_X}$. Je definována následovně:

$$p_{\mathcal{E}_X} : \mathcal{H}(\mathcal{E}_X) \rightarrow \mathbb{R} : \varepsilon \mapsto p_{\mathcal{E}_X}(\varepsilon) := P(\mathcal{E}_X = \varepsilon) . \quad (2)$$

K náhodným měřickým chybám \mathcal{E}_X můžeme definovat náhodnou veličinu X vztahem

$$X := x^* + \mathcal{E}_X . \quad (3)$$

Dané měřené hodnoty $x_i, \forall i \in \{1, 2, \dots, n\}$ jsou prvky **základního souboru** náhodné veličiny X , tzn. $x_i \in \mathcal{H}(X)$. Neznámá hodnota x^* se nazývá **parametr**

polohy náhodné veličiny X . Můžeme ji poznat pouze přibližně, prostřednictvím měřených hodnot x_1, x_2, \dots, x_n . Lze ji proto považovat za prvek základního souboru nějaké náhodné veličiny.

Hledá se:

- Diskrétní náhodná veličina Y příslušná **parametru polohy** náhodné veličiny X . Její pravděpodobnostní funkci označíme p_Y .

Řešení

Rozdělení pravděpodobnosti náhodné veličiny Y určíme na základě provedených měření $x_1, x_2, \dots, x_n \in \mathcal{H}(X)$. Hodnoty pravděpodobnostní funkce p_Y jsou proto podmíněné pravděpodobnosti.

$$p_Y(y) = P(Y = y | (X = x_1) \wedge (X = x_2) \wedge \dots \wedge (X = x_n)) . \quad (4)$$

Tato podmíněná pravděpodobnost není známa, lze však vypočítat opačně podmíněnou podmíněnou pravděpodobnost

$$P((X = x_1) \wedge (X = x_2) \wedge \dots \wedge (X = x_n) | Y = y) = \prod_{i=1}^n P(X = x_i | Y = y) , \quad (5)$$

protože jednotlivá měření jsou podle předpokladu vzájemně nezávislá.

Konjunkci jevů $(X = x_i)$, která se vyskytuje v rovnostech (4) a (5), označíme písmenem A .

$$A := (X = x_1) \wedge (X = x_2) \wedge \dots \wedge (X = x_n) . \quad (6)$$

Rovnosti (4), (5) tak dostanou přehlednější tvar

$$P(Y = y | A) = p_Y(y) , \quad (7)$$

$$P(A | Y = y) = \prod_{i=1}^n P(X = x_i | Y = y) . \quad (8)$$

Vztah mezi opačně podmíněnými podmíněnými pravděpodobnostmi popisuje Bayesův vzorec:

$$P(B_y | A) = \frac{P(A | B_y) P(B_y)}{\sum_{t=y_{\min}}^{y_{\max}} P(A | B_t) P(B_t)}$$

Abychom mohli podmíněné pravděpodobnosti (8), (7) využít v Bayesově vzorci, položíme

$$B_y := (Y = y) \in \{B_{y_{\min}}, B_{y_{\min}+1}, \dots, B_{y_{\max}}\} = \{B_t \mid t \in \{y_{\min}, y_{\min} + 1, \dots, y_{\max}\}\} . \quad (9)$$

Přitom

$$\begin{aligned} y_{\min} &:= \min \{y \in \mathbb{Z} \mid \forall i \in \{1, 2, \dots, n\} : P(X = x_i \mid Y = y) > 0\} , \\ y_{\max} &:= \max \{y \in \mathbb{Z} \mid \forall i \in \{1, 2, \dots, n\} : P(X = x_i \mid Y = y) > 0\} . \end{aligned} \quad (10)$$

Náhodný jev $X = x_i$ je vzhledem k rovnostem (3) a (1) totožný s náhodným jevem $\mathcal{E}_X = \varepsilon_i$. Pro pravděpodobnosti $P(X = x_i \mid Y = y)$ tudíž platí

$$P(X = x_i \mid Y = y) = P(\mathcal{E}_X = \varepsilon_i \mid Y = y) = P(\mathcal{E}_X = x_i - y) . \quad (11)$$

Tento vztah převádí podmíněné pravděpodobnosti $P(X = x_i \mid Y = y)$ na nepodmíněné pravděpodobnosti $P(\mathcal{E}_X = x_i - y)$. Využijeme ho pro substituci v součinu na pravé straně rovnosti (8).

Apriorní pravděpodobnosti $P(Y = y)$ určíme pomocí **apriorní pravděpodobnostní funkce**, která je dána.

$$P(Y = y) = p_{\text{aprior}}(y) . \quad (12)$$

Po dosazení substitucí (2), (7), (8), (9), (11), (12) do Bayesova vzorce dostáváme řešení zadaného problému ve tvaru:

$$p_Y(y) = \frac{\prod_{i=1}^n p_{\mathcal{E}_X}(x_i - y) p_{\text{aprior}}(y)}{\sum_{t=y_{\min}}^{y_{\max}} \prod_{i=1}^n p_{\mathcal{E}_X}(x_i - t) p_{\text{aprior}}(t)} .$$

6. listopadu 2017

Lubomír Soukup

soukup@utia.cas.cz