Použití Bayesovy věty
pro odhad parametru polohy
spojité náhodné veličiny

Dáno:

- Fyzikální veličina x, která je přístupná přímému měření. Nazýváme ji **měřená veličina**. Předpokládáme, že hodnoty veličiny x jsou reálná čísla, tj. $x \in \mathbb{R}$.

- Očekávaný rozsah měřených hodnot veličiny x a rozdělení pravděpodobnosti jejich výskytu založené na přibližné předběžné (apriorní) informaci o **měřené veličině**. Tato apriorní informace je dána prostřednictvím tzv. **apriorní hustoty pravděpodobnosti**. Označíme ji f_{aprior}.

- Několik naměřených hodnot veličiny x. Označíme je x_1, x_2, ..., x_n, $n \in \mathbb{N}$, $x_i \in \mathbb{R}$ pro $\forall i \ \in \{1, 2, \ldots, n\}$. Měření hodnot x_i bylo provedeno vzájemně nezávisle.

  Měření veličiny x je vždy ovlivněno nevyhnutelnými měřickými chybami $\varepsilon_i \in \mathbb{R}$, $i \in \{1, 2, \ldots, n\}$. Platí pro ně
  $$x_i = x^* + \varepsilon_i \ . \tag{1}$$
  Symbol x^* představuje **skutečnou hodnotu** měřené veličiny x. Hodnota x^* je neznámá, vpodstatě nepoznatelná, protože měření veličiny x je vždy ovlivněno nevyhnutelnými měřickými chybami. Předpokládáme $x^* \in \mathbb{R}$.

- Spojitá 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 $\mathcal{H}(\mathcal{E}_X) \subset \mathbb{R}$. Hustotu pravděpodobnosti náhodné veličiny $\mathcal{E}_X$ označíme $f_{\mathcal{E}_X}$. Je definována následovně:
  $$f_{\mathcal{E}_X} : \mathcal{H}(\mathcal{E}_X) \to \mathbb{R} : \varepsilon \mapsto f_{\mathcal{E}_X}(\varepsilon) \ . \tag{2}$$
  Pomocí náhodné veličiny $\mathcal{E}_X$ definujeme náhodnou veličinu X vztahem
  $$X := x^* + \mathcal{E}_X \ . \tag{3}$$
  Dané měřené hodnoty x_i, $\forall i \ \in \{1, 2, \ldots, n\}$ 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 $x_1, x_2, \ldots, x_n$. Lze ji proto považovat za prvek základního souboru nějaké náhodné veličiny.

Hledá se:

- Spojitá náhodná veličina Y příslušná **parametru polohy** náhodné veličiny X a naměřeným hodnotám x_1, x_2, …, x_n. Její hustota pravděpodobnosti se nazývá **aposteriorní hustota**.

Řešení

Rozdělení pravděpodobnosti náhodné veličiny Y určíme na základě provedených měření $x_1, x_2, \ldots, x_n \in \mathcal{H}(X)$. Tyto naměřené hodnoty uspořádáme do vektoru

$$\mathbf{x} := [\,x_1, x_2, \ldots, x_n\,]\,. \tag{4}$$

Skutečnost, že náhodná veličina X může při konstantní hodnotě x^* nabývat různých hodnot x_1, x_2, …, x_n současně, tj.

$$(X = x_1) \wedge (X = x_2) \wedge \cdots \wedge (X = x_n))$$

lze vyjádřit též pomocí náhodného jevu

$$\mathbf{X} = \mathbf{x}\,, \tag{5}$$

kde $\mathbf{X}$ je náhodný vektor

$$\mathbf{X} := \quad [\,\underbrace{X, X, \ldots, X}_{n}\,]\quad .$$

Hledaná hustota pravděpodobnosti náhodné veličiny Y bude tedy podmíněná hustota pravděpodobnosti daná podmínkou (5). Takto podmíněnou aposteriorní hustotu pravděpodobnosti označíme $f_Y(y\,|\,\mathbf{x})$.

Podmíněná hustota pravděpodobnosti $f_Y(y\,|\,\mathbf{x})$ není známa, lze však vypočítat opačně podmíněnou podmíněnou hustotu náhodného vektoru $\mathbf{X}$.

$$f_{\mathbf{X}}(\mathbf{x}\,|\,y) = \prod_{i=1}^{n} f_X(x_i\,|\,y)\,, \tag{6}$$

protože jednotlivá měření jsou podle předpokladu vzájemně nezávislá.

Vztah mezi opačně podmíněnými podmíněnými hustotami popisuje Bayesova věta:

$$\boxed{f_Y(y\,|\,\mathbf{x}) = \frac{f_{\mathbf{X}}(\mathbf{x}\,|\,y)\, f_{\text{aprior}}(y)}{\displaystyle\int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x}\,|\,t)\, f_{\text{aprior}}(t)\, dt}}$$

Náhodný jev $X = x_i$ je vzhledem k rovnostem (3) a (1) totožný s náhodným jevem $\mathcal{E}_X = \varepsilon_i$. Pro hustotu $f_X(x_i\,|\,y)$ z pravé strany rovnosti (6) tudíž platí:

$$f_X(x_i\,|\,y) = f_{\mathcal{E}_X}(x_i - y)\,. \tag{7}$$

Tento vztah převádí podmíněnou hustotu $f_X(.\,|\,y)$ na nepodmíněnou hustotu $f_{\mathcal{E}_X}$. Využijeme ho pro substituci v (6) a poté k vyjádření hustoty pravděpodobnosti náhodného vektoru $\mathbf{X}$ pomocí součinu jednorozměrných hustot $f_{\mathcal{E}_X}$.

$$f_{\mathbf{X}}(\mathbf{x}\,|\,y) = \prod_{i=1}^{n} f_X(x_i\,|\,y) = \prod_{i=1}^{n} f_{\mathcal{E}_X}(x_i - y)\ . \tag{8}$$

Po substituci (8) v Bayesově větě dostáváme řešení zadaného problému ve tvaru:

$$\boxed{f_Y(y\,|\,\mathbf{x}) = \frac{\displaystyle\prod_{i=1}^{n} f_{\mathcal{E}_X}(x_i - y)\ f_{\text{aprior}}(y)}{\displaystyle\int_{-\infty}^{\infty}\prod_{i=1}^{n} f_{\mathcal{E}_X}(x_i - t)\ f_{\text{aprior}}(t)\,dt}}\ .$$

Za předpokladu, že náhodná veličina $\mathcal{E}_X$ má normální rozdělení pravděpodobnosti $\mathcal{N}(0,\,\sigma_X^2)$ a apriorní rozdělení pravděpodobnosti je neinformativní (rovnoměrné na $\mathbb{R}$), je aposteriorní rozdělení rovněž normální a platí

$$Y \sim \mathcal{N}\left(\frac{\displaystyle\sum_{i=1}^{n} x_i}{n},\ \frac{\sigma_X^2}{n}\right)\ .$$

11. prosince 2017

Lubomír Soukup

soukup@utia.cas.cz