Bayesovské metody 3 -Stat. rozh.fce

5. října 2020

Outline

Outline

Statistické rozhodovací funkce

Formulace problému

$\boldsymbol{X} = (X_1, \ldots, X_n)^\top$ – náhodný vektor s hustotou $r(\boldsymbol{x}|\boldsymbol{\theta})$ vzhl. k σ -konečné míře ν_n na $(\mathcal{R}^n, \mathcal{B}^n)$, parametr $\boldsymbol{\theta} \in \Theta$

$\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)^\top$ – vektor konstant (klasický přístup) nebo náhodný vektor (bayesovský přístup)

Množina možných rozhodnuti o parametru $\boldsymbol{\theta}$ – $\mathcal{D}$

d- prvek množiny $\mathcal{D}$, rozhodnuti o parametru $\boldsymbol{\theta}$

$L(\boldsymbol{\theta}, d)$ – ztráta, je-li skutečná hodnota parametru $\boldsymbol{\theta}$ a naše rozhodnutí d; ztrátová funkce def. na $\Theta \times \mathcal{D}$, měřitelná fce, obvykle nezáporná

$\delta(\boldsymbol{x})$ –rozhodovaci funkce, proměnné $\boldsymbol{x}$, $\delta(\boldsymbol{x}) \in \mathcal{D}$ pros. $\boldsymbol{x}$

$R(\theta, \delta) = EL(\theta, \delta(\boldsymbol{X})|\theta)$ – riziko, riziková funkce

Δ – množina možných rozhodovacich funkci

(Θ, Δ, R) – Statistický rozhodovaci problém

Příklad $X_1, \ldots, X_n$ –náhodný výběr z alternativního rozdělení s parametrem $\theta \in [0,1]$, $\Theta = [0,1]$,
množina možných rozhodnutí $\mathcal{D} = [0,1]$, Δ– množina možných odhadů,

$L(\theta, \delta(\boldsymbol{X})) = (\theta - \delta(\boldsymbol{X}))^2$– ztrátová fce

$R(\theta, \delta(\boldsymbol{X})) = E\Big((\theta - \delta(\boldsymbol{X}))^2 | \theta\Big)$ – riziková fce

$H_0 : \theta \in \Theta_0$ versus $H_1 : \theta \in \Theta_1 = \Theta - \Theta_0$,

$\mathcal{D} = \{d_0, d_1\}$, $d_i = \{H_i \; plati\}$

$L(\theta, d_i) = 0, \quad \theta \in \Theta_i, \; i = 0, 1, \quad L(\theta, d_i) = a > 0 \quad \theta \notin \Theta_i$

Optimalni rozhodovaci funkce – minimalizuje rizikovou funkci $R(\theta, \delta(\boldsymbol{X}))$ vzhledem k $\delta(\boldsymbol{X})$ v nejakem smyslu

V přednášce budeme uvažovat minimalizaci tzv. bayeovskou rizikovou fci

$$\rho(q, \delta) = E\Big(R(\theta, \delta)|\theta)\Big) = EL(\theta, \delta)$$

bayesovská rozhodovaci funkce $\delta_*(\boldsymbol{x})$ – minimalizuje $\rho(q, \delta)$

Jiná možnost – minimax – minimalizuje se $\max_\theta R(\theta, \delta)|\theta)$

Bayesovské metody 3-Stat. rozh.fce

Outline

Statistické rozhodovací funkce

Podíváme se na bayesovskou rizikovou fci s apriorní hustotou $q(.)$

$$\rho(q,\delta) = EL(\theta,\delta(\boldsymbol{X}) = \int_\Theta \int_{\mathcal{R}_{\boldsymbol{k}}} L(\boldsymbol{\theta},\delta(\boldsymbol{x}))\frac{r(\boldsymbol{x}|\boldsymbol{\theta})q(\boldsymbol{\theta})}{r(\boldsymbol{x})} r(\boldsymbol{x}) d\lambda(\boldsymbol{\theta}) d\nu_{\boldsymbol{n}}(\boldsymbol{x})$$

$$= \int_{\mathcal{R}_{\boldsymbol{n}}} \Big[\int_\Theta L(\boldsymbol{\theta},\delta(\boldsymbol{x}))\pi(\boldsymbol{\theta}|\boldsymbol{x}) d\lambda(\boldsymbol{\theta}) \Big] r(\boldsymbol{x}|\boldsymbol{\theta})q(\boldsymbol{\theta}) d\nu_{\boldsymbol{n}}(\boldsymbol{x})$$

Tedy pro zvolené $q(.)$ a $\boldsymbol{x}$ stačí minimalizovat

$$\int_\Theta L(\boldsymbol{\theta},\delta(\boldsymbol{x}))\pi(\boldsymbol{\theta}|\boldsymbol{x}) d\lambda(\boldsymbol{\theta}) = E\Big(L(\boldsymbol{\theta},\delta(\boldsymbol{X}))|\boldsymbol{X} = \boldsymbol{x}\Big)$$

vzhledem k $\delta(\boldsymbol{x})$. Toto pak vede k definici rozhodovací funkce, která je bayesovská rozhodovaci funkce $\delta_*(\boldsymbol{x})$ pro skoro vs. $\boldsymbol{x}$ vzhledem k míře $\nu_{\boldsymbol{n}}$.

Toto dává návod, jak určit $\delta_*(\boldsymbol{x})$, kde za $\boldsymbol{x}$ dosadíme realizaci $\boldsymbol{X}$

Příklad 1, pokračování. Problém určení IQ sledovaného dítěte.
Výsledek testu X má rozdělení $N(\theta, 100)$, θ je IQ sledovaného dítěte. θ má rozd. $N(100,\ 225)$ – to vyjadřuje informaci o možných hodnotách θ před provedením testu. Aposteriorní rozdělení θ při daném X je normální

$$N\Big(X\frac{225}{225+100} + 100\frac{100}{225+100}, 69,23\Big)$$

Pro ztratovou funkci $L(\theta,\delta(x)) = (\theta-\delta(x))^2$ dostaneme bayesovskou rozhodovací fci $X\frac{225}{225+100} + 100\frac{100}{225+100}$

Příklad. Předpokládejme $\Theta = \{\theta_1, \theta_2\}$ a $\mathcal{D} = \{d_1, d_2\}$ a uvažujme ztrátovou funkci

$$\begin{aligned} L(\theta_i, d_i) &= 0, \quad i = 1, 2 \\ L(\theta_1, d_2) &= a_1 \\ L(\theta_2, d_1) &= a_2 \end{aligned}$$

a

$$P(\theta = \theta_1) = \xi, \quad P(\theta = \theta_2) = 1 - \xi, \quad \xi \in (0, 1)$$

Pak pro bayesovskou rizikovou funkci máme

$$\rho(\xi, \delta) = a_1 \xi P(\delta(X) = d_2 | \theta = \theta_1) + a_2(1 - \xi) P(\delta(X) = d_1 | \theta = \theta_2),$$

kde $P(\delta(X) = d_2|\theta = \theta_1)$ a $P(\delta(X = d_1|\theta = \theta_2)$ jsou pravděpodobnosti podmíněných rozdělení. Bayesovská rozhodovací funkce δ_* minimalizuje $\rho(\xi, \delta)$ vzhledem k δ.

Definujme rozhodovací funkci δ_0 předpisem

$$\begin{aligned} \delta_0(X) = & d_1 \qquad a_1\xi \cdot r(x|\theta_1) > a_2(1 - \xi) \cdot r(x|\theta_2) \\ = & d_2 \qquad a_1\xi \cdot r(x|\theta_1) < a_2(1 - \xi) \cdot r(x|\theta_2) \\ = & lib. \qquad a_1\xi \cdot r(x|\theta_1) = a_2(1 - \xi) \cdot r(x|\theta_2) \end{aligned}$$

Pro libovolnou rozhodovací funkci δ platí

$$\begin{aligned} & a_1\xi P(\delta_0(X) = d_2|\theta = \theta_1) + a_2(1 - \xi)P(\delta_0(X = d_1|\theta = \theta_2) \\ \leq \ & a_1\xi P(\delta(X) = d_2|\theta = \theta_1) + b a_2(1 - \xi)P(\delta(X = d_1|\theta = \theta_2) \end{aligned}$$

Odvození $a = a_1\xi$ a $b = a_2(1-\xi)$

$$
\begin{aligned}
&aP(\delta(X) = d_2|\theta = \theta_1) \;+\; bP(\delta(X = d_1|\theta = \theta_2) \\
&= a\int_{\delta(X)=d_2} r(x|\theta_1)d\nu(x) + b\int_{\delta(X)=d_1} r(x|\theta_2)d\nu(x) \\
&= a + \int_{\mathcal{R}_k} I\{\delta(X) = d_1,\, \delta_0(X) = d_1\}(-ar(x|\theta_1) + br(x)|\theta_2)d\nu(x) \\
&+ \int_{\mathcal{R}_k} I\{\delta(X) = d_1,\, \delta_0(X) = d_2\}(-ar(x|\theta_1) + br(x|\theta_2)d\nu(x)
\end{aligned}
$$

Platí

$$
\begin{aligned}
&\int_{\mathcal{R}_k} I\{\delta(X) = d_1,\, \delta_0(X) = d_1\}(-ar(x|\theta_1) + br(x) = \theta_2)d\nu(x) \\
&\qquad \geq \int_{\mathcal{R}_k} I\{\delta_0(X) = d_1\}(-ar(x|\theta_1) + br(x)|\theta_2)d\nu(x),
\end{aligned}
$$

$$
\int_{\mathcal{R}_k} I\{\delta(X) = d_1,\, \delta_0(X) = d_2\}(-ar(x|\theta_1) + br(x)|\theta_2)d\nu(x) \geq 0
$$

Odtud již plyne tvrzení.

Diskuse