Bayesovské metody, část I

Marie Hušková

přednáška 2020

Outline

Úvod

Částečný text pro první část přednášky Bayesovké metody založen na

skripta: Hušková : Bayesovské metody (na webu doc. Komárka)

kniha: Christian P. Robert: The Bayesian Choice, Springer, 2007

Něco málo z historie

Thomas Bayes IPA byl anglický duchovní,
narozen:1702, Londýn, Velká Británie
zemřel 7. dubna 1761, Royal Tunbridge Wells, Velká Británie
vzdělání: Edinburská univerzita
význačná vlastnost: Bayesova věta
Knihy: An Introduction to the Doctrine of Fluxions,: And Defence of the Mathematicians Against the Objections of the Author of the Analyst
ovlivněn prací A. de Moivera (1718)
rodiče: Joshua Bayes, Anne Carpenter

Ze základních kurzů je připomeňme tzv. Bayesův vzorec

$$P(A_k|B_i) = \frac{P(A_k)P(B_i|A_k)}{\sum_{j=1}^m P(A_j)P(B_i|A_j)}$$

kde $A_1, \dots, A_m$ jsou jevy s vlastností

$$P(A_j \cap A_s) = 0, \quad j \neq s, \quad P(\cup_{j=1}^m A_j) = 1,$$

$P(A_k|B_i)$ je podmíněná pravděpodobnost, ze nastane jev A_k za podmínky, že nastal jev B_i

$$P(A_k|B_i) = \frac{P(A_k \cap B_i)}{P(B_i)}, \quad P(B_i) > 0$$

$P(A_k)$– apriorní pravděpodobnost jevu A_k

$P(A_k|B_i)$ – aposteriorní pravděpodobnost jevu A_k, nastal-li jev B_i

Bayes odvodil vzorec pro $P(A_k) = \frac{1}{m}$ pro vš. k

Vyšlo v r. 1763, zobecněno P.C. Laplacem 1773.

Další vývoj až de Finneti ve 30.letech 20. století.

Pak po 2. světové válce v souvislosti s teorií statistického rozhodování.

Další významný pokrok az v souvislosti s prudkým rozvojem počítačů.

Příklad Je známo, že nemocí A trpí 1% populace. Nemoc je diagnostikována na základě vyšetření,
jehož spolehlivost je 95 %, jestliže vyšetřovaná osoba trpí chorobou A, a je 70 %, pokud chorobou A netrpí.
Je vyšetřena náhodně zvolená osoba.
Určete pravděpodobnost správné diagnózy

(i) byl-li výsledek vyšetření pozitivní,

(ii) byl-li negativní.

Řešení. Jev A – vyšetřovaná osoba trpí chorobou A a jev B – výsledek vyšetření byl pozitivní.

$$P(A) = 0,01; \quad P(A^c) = 0,99;$$

$$P(B|A) = 0,95; \quad P(B^c|A) = 0,05; \quad P(B^c|A^c) = 0,70; \quad P(B|A^c) = 0,30.$$

$$P(B) = 0,95 \times 0,01 + 0,30 \times 0,99 = 0,3065$$

$$P(A|B) = \frac{0,0095}{0,3065} = 0,031, \quad P(A^c|B^c) = \frac{0,693}{0,6935} = 0,9993$$

$P(A)$ – pravděpodobnost, že náhodně vybraná osoba trpí chorobou A

$P(A|B)$ – pravděpodobnost, že náhodně vybraná osoba ze ´ skupiny osob s pozitivním vysledkem vyšetření trpí chorobou A

$P(A), P(A^c)$– apriorní pravděpodobnosti (pravděpodobnosti sledovanych jevů před pokusem – před vyšetřením)

$P(A|B), P(A^c|B^c)$ – aposteriorní pravděpodobnosti (pravděpodobnosti sledovaných jevů po pokuse – bere se v úvahu výsledek vyšetření).

Zakladní formulace a princip tzv. bayesovského přístupu:

$X_1, \ldots, X_n$ – nezávislé stejně rozdělené náhodné veličiny s distribuční funkcí $F(x;\boldsymbol{\theta}), x \in \mathcal{R}, \boldsymbol{\theta} \in \Theta$
$\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)^T$ – parametr

klasický (frekventistický) přístup – $\theta_j, j = 1, \ldots, k$ jsou neznámé konstanty

bayesovský přístup – $\theta_j, j = 1, \ldots, k$ jsou náhodné veličiny

Základní úlohy při obou přístupech stejné:

bodové odhady, intervalové odhady, testování hypotéz a další

Příklad 1. Problém určení IQ sledovaného dítěte.
Výsledek testu X ma rozdělení $N(\theta, 100)$, θ je IQ sledovaného dítěte.
Dlouholeté výzkumy ukazují, že θ má rozd. $N(100, 225)$ – to vyjadřuje informaci o možných hodnotách θ před provedením testu.

Příklad. Kontrola jakosti, θ - podíl vadnych výrobků, pozorujme X - počet vadných v dané várce

Bayesovské metody I
Marie Hušková
Outline

Bayesova věta

Značení

$\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)^\top$ – náhodný vektor s hustotou $q(\boldsymbol{\theta})$ vzhl. k σ - konečné míře λ na $(\Theta, \mathcal{B}(\Theta))$, Θ - neprázdná borelovská podmnožina $\mathcal{R}^k$, $\mathcal{B}(\Theta)$ – borelovské podmnožiny Θ

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

$$P(\boldsymbol{\theta} \in B, \boldsymbol{X} \in C) = \int_B \Big(\int_C r(\boldsymbol{x}|\boldsymbol{\theta}) d\nu_n(\boldsymbol{x}) \Big) q(\boldsymbol{\theta}) d\lambda(\boldsymbol{\theta})$$

pro B, C lib. měř. množiny

ν_n, λ – typicky Lesqueova míra nebo čítací míra

Bayesova věta Pro podmíněnou hustotu $\pi(\boldsymbol{\theta}|\boldsymbol{x})$ náhodného vektoru $\boldsymbol{\theta}$ při daném $\boldsymbol{x}$ platí

$$\pi(\boldsymbol{\theta}|\boldsymbol{x}) = \frac{q(\boldsymbol{\theta})r(\boldsymbol{x}|\boldsymbol{\theta})}{\int_\Theta q(\boldsymbol{\theta})r(\boldsymbol{x}|\boldsymbol{\theta})d\lambda(\boldsymbol{\theta})}, \quad \int_\Theta q(\boldsymbol{\theta})r(\boldsymbol{x}|\boldsymbol{\theta})d\lambda(\boldsymbol{\theta}) > 0$$
$$= 0, \qquad \textit{jinak}$$

$\pi(\boldsymbol{\theta}|\boldsymbol{x}) \propto q(\boldsymbol{\theta})r(\boldsymbol{x}|\boldsymbol{\theta})$ – zkrácené značení (pozor- proměnná je $\boldsymbol{\theta}$, nikoli $\boldsymbol{x}$)

$r(\boldsymbol{x}) = \int_\Theta q(\boldsymbol{\theta})r(\boldsymbol{x}|\boldsymbol{\theta})d\lambda(\boldsymbol{\theta})$ -marginální hustota $\boldsymbol{x}$

$q(\boldsymbol{\theta})$ – apriorní hustota

$\pi(\boldsymbol{\theta}|\boldsymbol{x})$ – aposteriorní hustota

typicky λ - Lebesgueova míra

rozmyslet verzi pro spojitá a diskrétní rozděleni X

Příklad 1, pokračování. Problém určení IQ sledovaného dítěte.
Výsledek testu X má rozdělení $N(\theta, 100)$, θ je IQ sledovaného dítěte.

Dlouholeté výzkumy ukazují, že θ má rozd. $N(100, 225)$ – to vyjadřuje informaci o možných hodnotách θ před provedením testu.

Aposteriorní rozdělení θ při daném X je normální

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

Při klasickém přístupu použijeme k závěrům jen rozdělení X – odhad θ je X.

Při bayesovkém přístupu použijeme k závěrům aposteriorní rozdělení θ – odhad θ je $X\frac{225}{225+100}+100\frac{100}{225+100}$

Základní problémy

• Volba apriorního rozdělení

• konstrukce odhadů, testů, konfidenčních množin a dalších statistických procedur založených na aposteriorních rozděleních

Apriorní rozdělení

Apriorní rozdělení vyjadřuje (apriorní) informaci o možných hodnotách sledovaného parametru či jeho funkci před provedenim experimentu. Tato informace může mít objektivní charater, např. v příkladu o IQ sledovaného dítěte, podobně u příkladu z kontroly jakosti. Někdy však apriorní rozdělení vyjadřuje subjektivní pravděpodobnost – víru nějakého subjektu a různé subjekty mohou pro stejnou situaci mít odlišnou apriorní pravděpodobnost. Ta poslední zmíněná situace může souviset s filosofií.

V současné době se ukazuje je bayesovský přístup je užitečný i v jiných případech, např., když máme více řešení sledovaného problému, popř.více vhodných modelů.