Bayesova formule a teorie Bayesovských sítí

Sergii Babichev

Univerzita Jana Evangelisty Purkyně v Ústí nad Labem

sergii.babichev@ujep.cz

Motivace a pojetí pravděpodobnosti

- Pojem „pravděpodobnost“ má různé interpretace: frekventistický vs. subjektivní.
- Klasický přístup vychází z opakovatelných experimentů (např. høžení mincí).
- Subjektivní přístup je vhodný pro jedinečné události a je matematicky zdůvodnitelný.
- Bayesovské sítě umožňují modelovat znalosti i při nekompletních datech.

Podmíněná pravděpodobnost

Nechť E a H jsou dvě události (nebo náhodné proměnné). Podmíněná pravděpodobnost události E za podmínky, že nastala událost H:

$$P(E|H) = \frac{P(E \cap H)}{P(H)} \tag{1}$$

kde $P(E \cap H)$ je pravděpodobnost současného výskytu E i H.

Úplná disjunktní množina událostí

Mějme úplnou a disjunktní množinu hypotéz $H_1, H_2, \ldots, H_n$, tak že:

$$\bigcup_{i=1}^{n} H_i = \Omega, \quad H_i \cap H_j = \emptyset, \; i \neq j \tag{2}$$

Pak lze pravděpodobnost E vyjádřit pomocí podmíněných pravděpodobností:

$$P(E) = \sum_{i=1}^{n} P(H_i) P(E|H_i) \tag{3}$$

Toto je tzv. věta o úplné pravděpodobnosti.

Přechod k Bayesově větě (krok 1)

Z definice podmíněné pravděpodobnosti máme:

$$P(H_k|E) = \frac{P(H_k \cap E)}{P(E)} \tag{4}$$

Zároveň platí:

$$P(H_k \cap E) = P(H_k)P(E|H_k) \tag{5}$$

Dosaďme do (1):

$$P(H_k|E) = \frac{P(H_k)P(E|H_k)}{P(E)} \tag{6}$$

Přechod k Bayesově větě (krok 2)

Dosadíme za $P(E)$ z věty o úplné pravděpodobnosti:

$$P(E) = \sum_{i=1}^{n} P(H_i)P(E|H_i) \tag{7}$$

Tím dostaneme Bayesovu formuli:

$$P(H_k|E) = \frac{P(H_k)P(E|H_k)}{\sum_{i=1}^{n} P(H_i)P(E|H_i)} \tag{8}$$

Shrnutí a význam

- Bayesova formule vyjadřuje, jak aktualizovat naši důvěru v platnost hypotézy H_k po pozorování E.
- Apriorní pravděpodobnosti $P(H_k)$ mohou být odhadnuty expertně.
- Bayesovský přístup je základem mnoha moderních metod učení a rozhodování.

Bayesova věta jako pomoc při rozhodování

Bayesovu větu lze použít v reálných situacích jako nástroj podpory rozhodování. Příklad:

- Mladý muž se pokouší posoudit postoj ženy na základě jejího chování.
- Předpokládá 5 možných stavů (hypotéz):
  - H_1: Zcela se nelíbí
  - H_2: Nelíbí se
  - H_3: Neutrální pocity
  - H_4: Líbí se
  - H_5: Zamilovanost na první pohled
- Apriorní pravděpodobnosti:

$$P(H) = [0{,}1,\ 0{,}15,\ 0{,}5,\ 0{,}2,\ 0{,}05] \tag{9}$$

Pozorování a expertní znalosti

Pravděpodobnosti pozorování při daném stavu ženy:

Pozorování	H_1	H_2	H_3	H_4	H_5
Kontakt pohledem	0,1	0,2	0,4	0,2	0,1
Úsměv	0,1	0,1	0,2	0,4	0,2
Požádala o pití	0,52	0,45	0,029	0,0009	0,0001
Dívá se na jiné muže	0,3	0,4	0,2	0,08	0,02
Krátké fráze	0,1	0,4	0,3	0,1	0,1
Toaleta + make-up	0,05	0,05	0,1	0,2	0,6

Výpočet pro první pozorování: kontakt pohledem

Použijeme Bayesovu větu pro každou hypotézu:

$$P(H_i|E) = \frac{P(H_i)P(E|H_i)}{\sum_{j=1}^{5} P(H_j)P(E|H_j)} \tag{10}$$

Vektory:

$$P(H) = [0{,}1,\ 0{,}15,\ 0{,}5,\ 0{,}2,\ 0{,}05]$$
$$P(E|H) = [0{,}1,\ 0{,}2,\ 0{,}4,\ 0{,}2,\ 0{,}1]$$

Součinové členy:

$$P(H_i)P(E|H_i) = [0{,}01,\ 0{,}03,\ 0{,}20,\ 0{,}04,\ 0{,}005]$$

Normační člen:

$$\sum = 0{,}01 + 0{,}03 + 0{,}20 + 0{,}04 + 0{,}005 = 0{,}285 \tag{11}$$

Výsledné posteriorní pravděpodobnosti

- $P(H_1|E) = \frac{0,01}{0,285} \approx 0,035$
- $P(H_2|E) = \frac{0,03}{0,285} \approx 0,105$
- $P(H_3|E) = \frac{0,20}{0,285} \approx 0,702$
- $P(H_4|E) = \frac{0,04}{0,285} \approx 0,140$
- $P(H_5|E) = \frac{0,005}{0,285} \approx 0,018$

Závěr: Nejpravděpodobnější je stav H_3: neutrální pocity (70 %).
Tuto proceduru lze opakovat pro další pozorování (např. „úsměv“, „toaleta“...). Konečná důvěra v hypotézu je dána vynásobením příslušných členů vektoru pro všechna pozorování.

Shrnutí příkladu

- Bayesova věta umožňuje na základě pozorování aktualizovat náš názor na hypotézu (stav).
- V praxi: rozpoznávání objektů, klasifikace, rozhodování.
- V Bayesovských sítích každý uzel (proměnná) může být vyhodnocen pomocí podmíněných pravděpodobností.
- Kombinace více pozorování umožňuje robustní rozhodování.

Proč uvažovat násobná pozorování?

- V reálných situacích máme často více nezávislých pozorování.
- Každé nové pozorování může upravit naši důvěru v platnost hypotézy.
- Iterativní Bayesovská aktualizace umožňuje akumulaci důkazů.
- Výsledná posteriorní pravděpodobnost závisí na všech dostupných údajích.

Iterativní Bayesovská aktualizace (postup)

- Mějme apriorní rozdělení pravděpodobností $P(H_k)$ a první pozorování E_1.
- Provedeme první aktualizaci:

$$P(H_k|E_1) = \frac{P(H_k) \cdot P(E_1|H_k)}{\sum_j P(H_j) \cdot P(E_1|H_j)}$$

- Tento posterior se stává novým apriorním rozdělením pro další pozorování E_2:

$$P(H_k|E_1, E_2) = \frac{P(H_k|E_1) \cdot P(E_2|H_k)}{\sum_j P(H_j|E_1) \cdot P(E_2|H_j)}$$

- Tento proces lze iterovat pro $E_3, E_4, \ldots, E_n$.

Alternativa: souhrnná aktualizace všech pozorování

- Pokud všechna pozorování $E_1, E_2, \ldots, E_n$ jsou podmíněně nezávislá vzhledem k H_k:

$$P(E_1, E_2, \ldots, E_n | H_k) = \prod_{i=1}^{n} P(E_i | H_k)$$

- Pak celková aktualizace se provede jedním krokem:

$$P(H_k | E_1, E_2, \ldots, E_n) = \frac{P(H_k) \cdot \prod_{i=1}^{n} P(E_i | H_k)}{\sum_j P(H_j) \cdot \prod_{i=1}^{n} P(E_i | H_j)}$$

- Výsledek je identický s iterativní aktualizací (za podmíněné nezávislosti).

Shrnutí

- Iterativní Bayesovská aktualizace umožňuje postupně zapracovávat nové důkazy.
- Při podmíněné nezávislosti pozorování je možné sloučit důkazy do jednoho výpočtu.
- Tato metoda je využitelná v sekvenčním rozhodování, strojovém učení, medicínské diagnostice atd.

Probléma Montyho Halla

Tento problém spočívá v soutěži, ve které soutěžící musí vybrat jednu ze tří dveří, za jednou z nichž je skrytá cena. Moderátor show (Monty) odemkne prázdné dveře a zeptá se soutěžícího, zda chce po tom, co si soutěžící vybral jedny dveře, změnit svůj výběr na jiné dveře.

Rozhodnutí spočívá v tom, zda si původní dveře ponechat, nebo je nahradit novými. Je výhodnější vybrat jiné dveře, protože pravděpodobnost výhry je vyšší. Abychom vyřešili tuto nejednoznačnost, vytvořme pro toto zadání model s bayesovskou sítí.

Guest	0	0	0	1	1	1	2	2	2
Price	0	1	2	0	1	2	0	1	2
Host(0)	0	0	0	0	0.5	1	0	1	0.5
Host(1)	0.5	0	1	0	0	0	1	0	0.5
Host(2)	0.5	1	0	1	0.5	0	0	0	0

Výpočet: $P(\text{Prize} \mid \text{Guest} = 0, \text{Host} = 2)$

Cíl: Vypočítat pravděpodobnost, že cena je za konkrétními dveřmi, za předpokladu:

- Hráč si zvolil dveře č. **0** ($Guest = 0$),
- Moderátor otevřel dveře č. **2** ($Host = 2$).

Použijeme Bayesův vzorec:

$$P(Prize = i \mid Guest = 0, Host = 2) =$$
$$\frac{P(Guest = 0) \cdot P(Prize = i) \cdot P(Host = 2 \mid Guest = 0, Prize = i)}{P(Host = 2 \mid Guest = 0)}$$

Apriorní pravděpodobnosti:

$$P(Guest = 0) = \frac{1}{3}, \quad P(Prize = i) = \frac{1}{3} \text{ pro } i = 0, 1, 2$$

Podmíněné pravděpodobnosti:

$$P(Host = 2 \mid Guest = 0, Prize = 0) = 0{,}5$$
$$P(Host = 2 \mid Guest = 0, Prize = 1) = 1$$
$$P(Host = 2 \mid Guest = 0, Prize = 2) = 0$$

Výpočet: $P(\text{Prize} \mid \text{Guest} = 0, \text{Host} = 2)$ (pokračování)

Čitatel pro jednotlivé varianty:

$$P_0 = \frac{1}{3} \cdot \frac{1}{3} \cdot 0{,}5 = \frac{1}{18}$$
$$P_1 = \frac{1}{3} \cdot \frac{1}{3} \cdot 1 = \frac{1}{9}$$
$$P_2 = \frac{1}{3} \cdot \frac{1}{3} \cdot 0 = 0$$

Normalizační konstanta:

$$Z = P_0 + P_1 + P_2 = \frac{1}{18} + \frac{1}{9} = \frac{1}{6}$$

Posteriorní pravděpodobnosti:

$$P(Prize = 0 \mid G = 0, H = 2) = \frac{\frac{1}{18}}{\frac{1}{6}} = \frac{1}{3}$$
$$P(Prize = 1 \mid G = 0, H = 2) = \frac{\frac{1}{9}}{\frac{1}{6}} = \frac{2}{3}$$
$$P(Prize = 2 \mid G = 0, H = 2) = 0$$

Závěr: Vyplatí se změnit volbu — pravděpodobnost výhry po změně je **2/3**.