13. cvičení z PaSti – 2021-06-02

Bayesovská statistika

1. Kvído zase píše test, tentokrát s deseti otázkami, každá má tři volby. Pro každou z otázek nastane (nezávisle na ostatních) jedna ze dvou stejně pravděpodobných možností: buď se Kvído tuto látku učil, otázku „umí“ a odpoví určitě správně, nebo se neučil a tipne uniformně náhodnou volbu.

(a) Pokud první otázku odpověděl Kvído správně, jaká je pravděpodobnost, že tuto otázku uměl?

(b) Jaká je apriorní distribuce (pravděpodobnostní funkce) počtu otázek, které Kvído umí?

(c) Kvído odpověděl správně na šest otázek z deseti. Jaká je posteriorní distribuce počtu otázek, které Kvído uměl?

2. Stejný test píše několik studentů. Každý z nich patří (uniformně náhodně) do jedné ze tří skupin – všichni v první skupině umí otázku s pravděpodobností $\vartheta_1 = 0.3$, ve druhé s pravděpodobností $\vartheta_2 = 0.7$, a ve třetí s pravděpodobností $\vartheta_3 = 0.95$. Náhodně vybraný student odpoví správně k otázek (z deseti).

(a) Metodou MAP rozhodněte, do které skupiny student patří.

(b) U budiž počet otázek, které student umí. Student odpoví správně na 5 otázek. Odvoďte posteriorní distribuci, MAP odhad U a odhad U pomocí střední hodnoty.

3. Máme dvě krabičky, v první jsou dva bílé a jeden černý míček, ve druhé dva černé a jeden bílý. Vybereme náhodně jednu z krabiček (pravděpodobnost volby první krabičky je p) a pak vytáhneme míček.

(a) Popište, jak metodou MAP rozhodnout podle barvy taženého míčku, zda jsme vybrali první nebo druhou krabičku.

(b) Pro $p = 1/2$ spočtete pravděpodobnost chybného rozhodnutí. Srovnejte s chybou chybného rozhodnutí bez tahání míčku.

4. Házíme cinknutou mincí s pravděpodobností líce Θ. Naše apriorní rozdělení je dáno funkcí $f = f_\Theta$ takovou, že $f(0) = f(1) = 0$, $f(0.5) = 2$, f je lineární na intervalech $[0, 0.5]$ a $[0.5, 1]$ a f je nulová mimo $[0, 1]$.

Najděte MAP odhad Θ, pokud z n nezávislých hodů padl líc k-krát.

5. Předpokládejme, že policejní radar naměří rychlost auta vyšší o $U \sim U(0, 5)$. Předpokládejme, že skutečná rychlost auta je $V \sim U(45, 65)$. Metodou podmíněné pravděpodobnosti odhadněte rychlost auta na základě naměřené rychlosti.

6. Počet nákupních vozíků v obchodě je Θ – uniformně rozdělená náhodná veličina s hodnotami $\{1, 2, \ldots, 100\}$. Každý vozík má číslo $(1, \ldots, \Theta)$. Na našem vozíku (předpokládáme, že je uniformně náhodný) je číslo X. Odhadněte Θ

(a) metodou MAP.

(b) pomocí podmíněné pravděpodobnosti.

(Úloha je podobná příkladu o Romeovi a Julii z přednášky, ale tam šlo o spojité náhodné veličiny.)

7. Počet minut mezi příjezdy autobusů na zastávku u kolejí má exponenciální rozdělení s parametrem Θ, použijeme apriorní distribuci s hustotou $f_\Theta(\vartheta) = 10\vartheta$ pro $\vartheta \in [0, 1/\sqrt{5}]$ (a nula jinde – ověřte, že se jedná o hustotu).

(a) Jdeme na zastávku a musíme čekat 30 minut. Jaká je posteriorní hustota a odhad Θ (metodou MAP nebo podmíněnou střední hodnotou)?

(b) Provedeme pět měření (jdeme na autobus pětkrát a zapíšeme si dobu čekání). Musíme čekat 30, 25, 15, 40 a 20 minut, předpokládáme, že jednotlivé dny jsou nezávislá měření se stejným rozdělením. Jaká je posteriorní hustota a odhad Θ (metodou MAP nebo podmíněnou střední hodnotou)?

8. Varianta problému z přednášky: Máme nezávislé náhodné veličiny $X_1, \ldots, X_n$. Všechny jsou uniformní na intervalu $[0, \vartheta]$, kde ϑ je hodnota n.v. Θ. Apriorní distribuce Θ je uniformní na $[0, 1]$. Na přednášce jsme měli $n = 1$, teď budeme předpokládat $n > 3$.

(a) Najděte odhad Θ pomocí pomíněné pravděpodobnosti (při daných hodnotách $x_1, \ldots, x_n$ veličiny $X_1, \ldots, X_n$).

(b) Nakreslete graf podmíněné střední kvadratické chyby (MSE) pro odhady MAP a podmíněné střední hodnoty, jako funkce $\bar{x} = \max\{x_1, \ldots, x_n\}$, pro hodnotu $n = 5$.

(c) Pokud $\bar{x} = 0.5$, jak se vyvíjejí odhad MAP a podmíněná střední hodnota, a odpovídající střední kvadratická chyba, v závislosti na $n \to \infty$?

The first rule of the Bayesian Conspiracy is that you talk about it exactly as much as a non-member would.