3. cvičení z PSI

12. - 16. října 2015

3.1 (Vylepšování generátoru náhody) Alice a Bob chtějí spravedlivě vybrat jednoho z nich. Mohou si hodit mincí, ale ta je zdeformovaná. Dohodnou se tedy, že hodí mincí dvakrát. Alice vyhrává, pokud padnou stejné výsledky, Bob při různých výsledcích. Kdo z nich má větší šancí vyhrát?

Zkuste zlepšit tento postup tak, aby pravděpodobnost výhry byla ještě bližší k $\frac{1}{2}$.

Řešení:
Předpokládejme, že jedna ze stran (např. líc) padá s pravděpodobností $\frac{1}{2} + \varepsilon$, a druhá (rub) s pravděpodobností $\frac{1}{2} - \varepsilon$, kde $|\varepsilon| < \frac{1}{2}$). Definujme si jevy:

L_i = "při i-tém hodu padl líc",
R_i = "při i-tém hodu padl rub",
A = "vyhraje Alice",
B = "vyhraje Bob".

Máme $\overline{L_i} = R_i$, $\overline{A} = B$ a $P(L_i) = \frac{1}{2} + \varepsilon$. Přitom jednotlivé hody považujeme za nezávislé, tedy např. L_i a L_j jsou nezávislé pro $i \neq j$. Ostatní nezávislé vztahy jsou díky doplňkům už určeny. Máme

$$A = \Big(L_1 \cap L_2\Big) \cup \Big(R_1 \cap R_2\Big)$$

$$P(A) = P(L_1 \cap L_2) + P(R_1 \cap R_2) = P(L_1) \cdot P(L_2) + P(R_1) \cdot P(R_2) =$$
$$= \Big(\frac{1}{2} + \varepsilon\Big)^2 + \Big(\frac{1}{2} - \varepsilon\Big)^2 = \frac{1}{2} + 2\varepsilon^2$$
$$P(B) = 1 - P(A) = \frac{1}{2} - 2\varepsilon^2$$

Větší šanci má tedy Alice. Z generátoru náhody s chybou $a_1 = |\varepsilon|$ (hod jednou minci jedenkrát) jsme se dostali ke generátoru náhody s chybou $a_2 = 2\varepsilon^2$ (hod touto minci dvakrát). Protože $|\varepsilon| < \frac{1}{2}$, je nová chyba menší $a_2 < a_1$. Tento postup můžeme opakovaně iterovat a z rekurentního vzorce $a_{n+1} = 2(a_n)^2$ dostáváme $a_n = \frac{1}{2}\Big(2\varepsilon\Big)^{2^n}$. Protože je $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} \frac{1}{2}\Big(2\varepsilon\Big)^{2^n} = 0$ můžeme opakovaným iterováním vyrobit generátor s libovolně malou odchylkou, která se velmi rychle blíží k nule. Současně ale exponenciálně vzrůstá počet kroku, který bude tento generátor potřebovat, takže bude velmi pomalý.

3.2 ((Ne)závislost jevů) Vybereme bod (x, y) ze čtverce $I = \langle -1, 1\rangle \times \langle -1, 1\rangle$ s rovnoměrným rozdělením. Ukažte, že jevy

$A = \{(x, y) \in I \mid x > 0\}$,
$B = \{(x, y) \in I \mid y > 0\}$,
$C = \{(x, y) \in I \mid x \cdot y > 0\}$,

nejsou nezávislé, ale pouze po dvou nezávislé.

Řešení:

Máme zde obvyklou geometrickou pravděpodobnost, tj. $P(X) = \frac{\mathrm{vol}(X)}{\mathrm{vol}(I)}$ pro jev X. Snadno dostaneme, ze

$$P(A) = P(B) = P(C) = \frac{1}{2}$$
$$A \cap B = A \cap C = B \cap C = A \cap B \cap C$$

a

$$P(A \cap B) = P(A \cap C) = P(B \cap C) = P(A \cap B \cap C) = \frac{1}{4}.$$

Tedy

$$P(A \cap B) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(A) \cdot P(B)$$

a podobně je to pro ostatní případy, zatímco

$$P(A \cap B \cap C) = \frac{1}{4} \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = P(A) \cdot P(B) \cdot P(C).$$

Jevy jsou tak po dvou nezávislé, ale ne celkově nezávislé.

Je dobré poznamenat, ze pokud máme nezávislé jevy s pravděpodobnostmi $p_1, \ldots, p_n$, které nejsou ani jisté ani nemožné, tak je to vždy ekvivalentní případu, kdy házíme n nezávislými mincemi, kde jedna strana i-té mince padá s pravděpodobností p_i. Přesněji se to dá říci také tak, že:

- σ-algebra generovaná n nezávislými jevy (které nejsou ani jisté ani nemožné) je izomorfní (tj. chová se jako)
$$\exp\Big(\exp\big(\{1, \ldots, n\}\big)\Big)$$
což je tzv. *volná* Booleova algebra. Její velikost je 2^{2^n}.

Jisté a nemožné jevy z toho vylučujeme kvůli tomu, že

- jestliže $A_1, \ldots, A_n$ jsou nezávislé jevy a $P(A) \in \{0, 1\}$ pro nějaký jev A, pak $A_1, \ldots, A_n, A$ jsou opět nezávislé jevy a také $A_1 \cap A, \ldots, A_n \cap A$ jsou nezávislé jevy. Tedy pokud přidáme nebo sebereme takovýto jev, tak sice získáme opět nezávislé jevy, ale je to nezajímavá informace.

Příklad po dvou nezávislých objektů, co ale nejsou celkově nezávislé známe i z lineární algebry: tři vektory mohou být závislé i když jsou po dvou nezávislé. Obě situace a oba pojmy nezávislosti jsou si podobné v následujícím smyslu:

Nechť $v_1, \ldots, v_n$ jsou nezávislé vektory a W nějaký vektorový prostor. Pak libovolné zobrazení $f : \{v_1, \ldots, v_n\} \to V$ se dá rozšířit na lineární zobrazení $\overline{f} : V \to W$, kde V je lineární prostor generovaný vektory $v_1, \ldots, v_n$. Tedy lineárně nezávislé vektory si můžeme zobrazovat, jak chceme (díky nezávislosti nemáme žádná omezení).

A podobně: Nechť $A_1, \ldots, A_n$ jsou nezávislé jevy (kde $0 < P(A_i) < 1$ pro $i = 1, \ldots, n$) a $\mathcal{B}$ nějaká σ-algebra. Pak libovolné zobrazení $f : \{A_1, \ldots, A_n\} \to \mathcal{B}$ se dá rozšířit na homomorfismus σ-algeber $\tilde{f} : \mathcal{A} \to \mathcal{B}$, kde $\mathcal{A}$ je σ-algebra generovaná jevy $A_1, \ldots, A_n$ a je to právě výše zmiňovaná volná Booleova algebra. Tedy nezávislé jevy (které nejsou ani jisté ani nemožné) si zase můžeme zobrazovat, jak chceme.

Ještě doplňme, že

- zobrazení $g : \mathcal{A} \to \mathcal{B}$ je homomorfismus σ-algeber když přenáší spočetná sjednocení a doplňky, tedy
$$g\Big(\bigcup_{n \in \mathbb{N}} X_n\Big) = \bigcup_{n \in \mathbb{N}} g(X_n)$$
$$g\Big(\overline{X}\Big) = \overline{g(X)}$$
pro jevy $X, X_n \in \mathcal{A}$.

3.3 ((Ne)závislost jevů) Pro hod dvěma mincemi uvažujme jevy:

$A =$ "na první minci padl líc",
$B =$ "na druhé minci padl líc",
$C =$ "na mincích padly různé výsledky".

Jak je to s nezávislostí jevů A, B, C?

Řešení:
Bude to podobné jako v předchozím příkladu. Prostor elementárních jevu bude $\Omega = \{\text{líc, rub}\} \times \{\text{líc, rub}\}$ a každý elementární jev bude stejně pravděpodobný. Pak máme

$$P(A) = P(B) = P(C) = \frac{1}{2}$$

$$P(A \cap B) = P(A \cap C) = P(B \cap C) = \frac{1}{4}$$

a

$$P(A \cap B \cap C) = 0$$

protože $A \cap B \cap C = \emptyset$. Tedy jevy jsou po dvou nezávislé, ale ne celkově nezávislé.

3.4 (Bayesovská pravděpodobnost v informačním kanálu se šumem) Na vstupu informačního kanálu jsou posílány znaky "0"a "1", přitom znak "1"je posílán s pravděpodobností p. Na výstupu je daný znak přečten s pravděpodobností chyby $r = 0.1$, která nezávisí na frekvenci s jakou znak chodí (tj. na hodnotě p). Určete podmíněné pravděpodobnosti vstupu při známém výstupu, je-li

(a) $p = 0.4$,

(b) $p = 0.1$.

Řešení:
Máme jevy

$V_i =$ "vyšleme znak i",
$Z_i =$ "zachytíme znak i",

kde i je nula nebo jednička. Víme, že

$$\overline{V_0} = V_1 \quad \overline{Z_0} = Z_1 \quad \text{a} \quad P(V_1) = p.$$

Pravděpodobnost r chyby znaku "1"na výstupu je dána procentem zachycených znaku "0"v množině odeslaných znaku "1", tj. $r = \frac{P(Z_0 \cap V_1)}{P(V_1)} = P(Z_0|V_1)$. Podobně $r = P(Z_1|V_0)$. Pro zjednodušení si uvědomíme, že funkce $\tilde{P}(A) := P(A|B)$ je pravděpodobnost v proměnné A, speciálně tedy

$$P(Z_0|V_0) = 1 - P(Z_1|V_0) = 1 - r$$

a

$$P(Z_1|V_1) = 1 - P(Z_0|V_1) = 1 - r.$$

Zajímají nás podmíněné pravděpodobností $P(V_i|Z_j)$ pro $i, j \in \{0, 1\}$. Opět stačí spočítat jen některé z nich (pro zbylé máme vztahy jako např. $P(V_0|Z_1) = 1 - P(V_1|Z_1)$.) Dále pro snadnější zápis ještě použijeme, že znak $1 - i$ je odlišný od znaku i.

Podle Bayesových vět teď máme

$$P(V_i|Z_j) = \frac{P(Z_j|V_i)P(V_i)}{P(Z_j)} = \frac{P(Z_j|V_i)P(V_i)}{P(Z_j|V_i)P(V_i) + P(Z_j|V_{1-i})P(V_{1-i})} =$$

$$= \frac{1}{1 + \frac{P(Z_j|V_{1-i})P(V_{1-i})}{P(Z_j|V_i)P(V_i)}}$$

takže

$$P(V_0|Z_0) = \frac{1}{1 + \frac{P(Z_0|V_1)P(V_1)}{P(Z_0|V_0)P(V_0)}} = \frac{1}{1 + \frac{r \cdot p}{(1-r)\cdot(1-p)}}$$

a

$$P(V_1|Z_1) = \frac{1}{1 + \frac{P(Z_1|V_0)P(V_0)}{P(Z_1|V_1)P(V_1)}} = \frac{1}{1 + \frac{r\cdot(1-p)}{(1-r)\cdot p}} \ .$$

Vzorce uvádíme v tomto výsledném tvaru, aby se zvýraznila závislost na jednotlivých parametrech. Při praktickém počítání je ale vhodnější to nechat v původním zápisu a nepřevádět na tvar $\frac{1}{1+\text{něco}}$.

(a) Pro $p = 0.4$ tak máme

$$P(V_0|Z_0) = \frac{1}{1 + \frac{0.1\cdot 0.4}{0.9\cdot 0.6}} = \frac{27}{29} \doteq 0.93$$

$$P(V_1|Z_0) = 1 - \frac{27}{29} = \frac{2}{29} \doteq 0.07$$

$$P(V_1|Z_1) = \frac{1}{1 + \frac{0.1\cdot 0.6}{0.9\cdot 0.4}} = \frac{6}{7} \doteq 0.86$$

$$P(V_0|Z_1) = 1 - \frac{6}{7} = \frac{1}{7} \doteq 0.14$$

Tedy poměrně vysoká spolehlivost pro oba znaky.

(b) Pro $p = 0.1$ bude situace podstatně jiná:

$$P(V_0|Z_0) = \frac{1}{1 + \frac{0.1\cdot 0.1}{0.9\cdot 0.9}} = \frac{81}{82} \doteq 0.99$$

$$P(V_1|Z_0) = 1 - \frac{81}{82} = \frac{1}{82} \doteq 0.01$$

$$P(V_1|Z_1) = \frac{1}{1 + \frac{0.1\cdot 0.9}{0.9\cdot 0.1}} = \frac{1}{2} = 0.5$$

$$P(V_0|Z_1) = 1 - \frac{1}{2} = \frac{1}{2} = 0.5$$

Pokud procento vysílaných znaku "1"dosáhne hladiny šumu, tj. $p = r$, nedá se pak při zachyceni znaku "1"určit, jestli pochází z vyslaného signálu (tj. znaku "1") nebo naopak ze šumu (tj. chyby při vyslání znaku "0").

Ještě je dobré si uvědomit, že z Bayesových vět máme následující vztahy (a trochu jiný způsob řešení):

$$\begin{pmatrix} P(Z_0) \\ P(Z_1) \end{pmatrix} = \begin{pmatrix} P(Z_0|V_0) & P(Z_0|V_1) \\ P(Z_1|V_0) & P(Z_1|V_1) \end{pmatrix} \begin{pmatrix} P(V_0) \\ P(V_1) \end{pmatrix} =$$

$$= \begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix} \begin{pmatrix} 1-p \\ p \end{pmatrix} = \begin{pmatrix} 0.9 - 0.8p \\ 0.1 + 0.8p \end{pmatrix}$$

$$\begin{pmatrix} P(V_0|Z_0) & P(V_0|Z_1) \\ P(V_1|Z_0) & P(V_1|Z_1) \end{pmatrix} = \begin{pmatrix} P(V_0) & 0 \\ 0 & P(V_1) \end{pmatrix} \begin{pmatrix} P(Z_0|V_0) & P(Z_0|V_1) \\ P(Z_1|V_0) & P(Z_1|V_1) \end{pmatrix}^T \begin{pmatrix} P(Z_0)^{-1} & 0 \\ 0 & P(Z_1)^{-1} \end{pmatrix} =$$

$$= \begin{pmatrix} 1-p & 0 \\ 0 & p \end{pmatrix} \begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix} \begin{pmatrix} \frac{1}{0.9-0.8p} & 0 \\ 0 & \frac{1}{0.1+0.8p} \end{pmatrix} = \begin{pmatrix} \frac{0.9-0.9p}{0.9-0.8p} & \frac{0.1-0.1p}{0.1+0.8p} \\ \frac{0.1p}{0.9-0.8p} & \frac{0.9p}{0.1+0.8p} \end{pmatrix}.$$

3.5 (Bayesovská pravděpodobnost) Dveřní rám v obchodě se rozezvučí, pokud se někdo pokusí projít se zbožím, které nemá deaktivovaný čip. Systém spustí alarm v 95% případů, kdy prochází někdo s kradeným zbožím. V 5% se deaktivace čipu z nějakého důvodu nepovede a poplach bude spuštěn i při projití s legálně zakoupeným zbožím. Statisticky jsou 3% návštěvníků zloději a ostatní chtějí zboží normálně zakoupit. Jaká je pravděpodobnost, že alarm správně upozorní na kradené zboží?

Řešení:
Určíme si jevy:

$A =$ "rozezní se alarm",
$K =$ "zboží je kradené",

I když zadání může na první pohled svádět k jiné interpretaci, pravděpodobností budou tyto:

$$P(A|K) = 0.95$$

$$P(A|\overline{K}) = 0.05$$

$$P(K) = 0.03$$

Přestože součet prvních dvou podmíněných pravděpodobností dává 1, jde obecně o dvě nezávislé hodnoty a ne doplňkové pravděpodobnosti!

Nás teď zajímá, jaká bude pravděpodobnost $P(K|A)$. Podobně jako v předchozím příkladě dostaneme

$$P(K|A) = \frac{P(A|K)P(K)}{P(A)} = \frac{P(A|K)P(K)}{P(A|K)P(K) + P(A|\overline{K})P(\overline{K})} =$$

$$= \frac{1}{1 + \frac{P(A|\overline{K})(1-P(K))}{P(A|K)P(K)}} = \frac{1}{1 + \frac{0.05 \cdot (1-0.03)}{0.95 \cdot 0.03}} = \frac{57}{154} \doteq 0.37.$$

To se zdá být poměrně málo ve srovnání se zadanými hodnotami. Ve skutečnosti je to ale dáno velmi nízkým počtem zlodějů a rám tak vlastně spíše detekuje to, že se nepovede deaktivace čipu při obvyklém nákupu (viz doplňková pravděpodobnost $P(\overline{K}|A) \doteq 1 - 0.37 = 0.63$). Je také vidět, že čím nižší bude počet zlodějů, tím nespolehlivější bude v tomto směru rám - a naopak rám bude vysoce spolehlivý, pokud bude krást skoro každý...

Také bychom si mohli uvědomit, co vlastně říkají podmíněné pravděpodobnosti v zadání a jak se asi prakticky zjistí jejich hodnota:

$P(A|K)$: V obchodě si zaznamenávají počet úspěšně odhalených krádeží s pomoci rámu. Aby ale věděli, kolik zboží jim celkově chybí, musí udělat inventuru. Pak teprve mohou zjistit, jak účinný je v tomto směru rám.

$P(A|\overline{K})$: Zde zase zaznamenávají, kolikrát rám zazvonil "zbytečně", tj. kolik nastalo chyb při deaktivaci čipu u pokladny, a porovnají to s tím, kolik zboží prodali.

3.6 (Hypergeometrické rozdělení) Mezi M výrobky je K vadných. Jaká je pravděpodobnost, že mezi m náhodně vybranými výrobky je právě k vadných?

Určete, k čemu se blíží hodnota pravděpodobnosti pro pevné k a m pokud $M \to \infty$ a $K/M \to q$, kde $0 \leq q \leq 1$. Jak byste tento výsledek interpretovali?

Řešení:
Pravděpodobnost bude dána podílem příznivých možností ku všem. Příznivé jsou dány počtem způsobů jak vybrat k výrobků z K vadných násobeno počtem způsobů jak vybrat zbytek, tj. $m-k$ výrobků z $M-K$ bezvadných. Celkem tedy

$$p_{K,M}(k,m) = \frac{\binom{K}{k} \cdot \binom{M-K}{m-k}}{\binom{M}{m}}$$

Rozdělení náhodné veličiny X, která označuje počet vadných výrobků ve vybraném vzorku m výrobků (z množiny M obsahující K vadných výrobků), se nazývá hypergeometrické. Tedy

$$P(X = k) = p_{K,M}(k,m).$$

Určíme ještě limitu ze zadání:

$$\lim_{M\to\infty} p_{K,M}(k,m) = \lim_{M\to\infty} \frac{m!}{k!(m-k)!} \frac{\prod_{i=0}^{k-1}(K-i) \cdot \prod_{j=0}^{m-k-1}(M-K-j)}{\prod_{i=0}^{m-1}(M-i)} =$$

$$= \binom{m}{k} \cdot \lim_{M\to\infty} \prod_{i=0}^{k-1} \frac{K-i}{M-i} \cdot \prod_{j=0}^{m-k-1} \frac{M-K-j}{M-k-j} = \binom{m}{k} q^k (1-q)^{m-k}$$

protože uvažujeme, ze k a m jsou pevné a předpokládáme, ze $K/M \to q$ pro $M \to \infty$.

Dostáváme tedy známe binomické rozdělení, které v tomto případě odpovídá limitní situaci, kdy taháme výrobky z nekonečného množství s určeným podílem q vadných. Při n pokusech jsme předpokládali, ze výrobky NEVRACÍME (anebo je prostě vytáhneme všechny naraz). Jak ale vidíme, i přesto jsme dostali rozdělení pravděpodobnosti, které používáme, když opakovaně uskutečňujeme tentýž pokus vždy za STEJNÝCH podmínek (např. opakovaně vytahujeme z osudí koule a VRACÍME je vždy zpátky). To je tím, ze v obrovském množství výrobků M se už v limite ztratí to, jestli tam těch pár výrobků m vrátíme nebo ne (protože $m << M$).