3. cvičení z PST

8. října 2020

Připomeňme si, co reprezentují jednotlivé složky Kolmogorova modelu $(\Omega, \mathcal{A}, P)$:

* Ω je množina všech možných výsledků (tzv. *elementárních jevů*)
* $\mathcal{A}$ představuje všechny "přípustné" množiny takovýchto výsledků, tedy všechny *jevy*, se kterými můžeme pracovat.
* $P : \mathcal{A} \to \langle 0, 1\rangle$ je zobrazení, které jevu $A \in \mathcal{A}$ přiřadí jeho *pravděpodobnost* $P(A)$. A právě kvůli tomu, abychom vůbec takovéto přiřazení P mohli získat, potřebujeme požadovat jisté speciální vlastnosti od systému všech jevů $\mathcal{A}$ (chceme, aby $\mathcal{A}$ tvořil tzv. *σ-algebru*).

3.1 (Kolmogorův model)

Zjistěte, a případně doplňte, $(\Omega, \mathcal{A}, P)$ na Kolmogorův model pravděpodobnosti, je-li dáno:

- $\Omega = \{1, 2, 3\}$,
- $\mathcal{A} = \big\{\emptyset, \{1, 2\}, \{1, 3\}, \{2\}, \{3\}, \{1, 2, 3\}\big\}$,
- $P(A) = \frac{1}{3}|A|$, $A \in \mathcal{A}$ (kde $|A|$ je počet prvků množiny A).

Řešení:

Pro Kolmogorův model je tedy potřeba ověřit, že $\mathcal{A}$ je σ-algebra, tj. že splňuje:

- $\emptyset \in \mathcal{A}$,
- $A \in \mathcal{A} \;\Rightarrow \overline{A} \in \mathcal{A}$,
- $\{A_n \mid n \in \mathbb{N}\} \subseteq \mathcal{A} \;\Rightarrow\; \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$,

a že $P : \mathcal{A} \to \langle 0, 1\rangle$ je pravděpodobnost, tj. že splňuje:

- $P(\emptyset) = 0$,
- $P(\overline{A}) = 1 - P(A)$ pro každé $A \in \mathcal{A}$,
- $\{A_n \mid n \in \mathbb{N}\} \subseteq \mathcal{A}$ & A_n jsou navzájem disjunktní $\;\Rightarrow\; P\big(\bigcup_{n\in\mathbb{N}} A_n\big) = \sum_{n\in\mathbb{N}} P(A_n)$,

První dvě podmínky pro σ-algebru jsou zřejmě splněny, poslední ne, protože

$$\{2\}, \{3\} \in \mathcal{A}, \quad \text{ale} \quad \{2\} \cup \{3\} \notin \mathcal{A} \;.$$

Množina $\mathcal{A}$ tedy *není* σ-algebra a $(\Omega, \mathcal{A}, P)$ proto *není* Kolmogorův model.

Tuto nedokonalost, ale můžeme spravit tak, že k $\mathcal{A}$ přidáme prvky, které chybí: tedy prvek $\{2\}\cup\{3\} = \{2, 3\}$ a $\overline{\{2, 3\}} = \{1\}$. Dostaneme tak celou potenční množinu $\mathcal{A}' = \exp(\Omega)$, která σ-algebrou určitě je.

Teď ještě ukážeme, že $P : \mathcal{A}' \to \langle 0, 1\rangle$ (stále uvažujeme stejný předpis) je v tomto případě pravděpodobnost. Pro ulehčení si všimneme, že $|\Omega| = 3$ a tedy

$$P(A) = \frac{|A|}{|\Omega|} \in \langle 0, 1\rangle$$

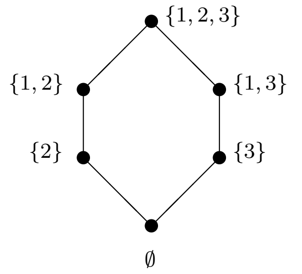
tedy jde o Laplaceovu pravděpodobnost (tj. počet příznivých případů ku počtu všech.) Pro pořádek si tedy zkontrolujeme, ze jde skutečně o pravděpodobnost

- $P(\emptyset) = \frac{|\emptyset|}{|\Omega|} = 0$
- $P(\overline{A}) = \frac{|\Omega \setminus A|}{|\Omega|} = \frac{|\Omega|-|A|}{|\Omega|} = 1 - \frac{|A|}{|\Omega|} = 1 - P(A)$
- pro $\{A_n \mid n \in \mathbb{N}\} \subseteq \exp(\Omega)$ navzájem disjunktní máme

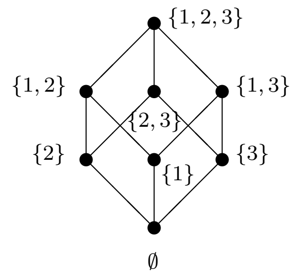
$$P\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) = \frac{\left|\bigcup_{n\in\mathbb{N}} A_n\right|}{|\Omega|} = \frac{1}{|\Omega|}\sum_{n\in\mathbb{N}} |A_n| = \sum_{n\in\mathbb{N}} P(A_n) \ .$$

Uspořádané množiny (v našem případě je uspořádání dáno inkluzí) můžeme ještě zakreslit tzv. Hasseovým diagramem (větší prvky se zakreslují nad menší a spojují se čárkou, pokud už mezi nimi žádné další prvky nejsou). Dostáváme tak:

$\mathcal{A}$:



exp(Ω) :



To, že jsme ve druhém případě dostali obrázek, který vypadá jako krychle, není náhoda. Konečné σ-algebry budou mít vždy Hasseův diagram ve tvaru vícerozměrné krychle.

Připomenutí: Jevy $A_1, \ldots, A_n$ jsou nezávislé $\overset{\text{def.}}{\iff}$ pro každou indexovou množinu $K \subseteq \{1, \ldots, n\}$ je

$$P\Big(\bigcap_{i\in K} A_i\Big) = \prod_{i\in K} P(A_i) \ .$$

Poznámka: Pokud jevy $A_1, \ldots, A_n$ jsou nezávislé, pak také jevy

- $A_1 \cup A_2, A_3, \ldots, A_n$ jsou nezávislé
- $A_1 \cap A_2, A_3, \ldots, A_n$ jsou nezávislé
- $\overline{A_1}, A_2, A_3, \ldots, A_n$ jsou nezávislé

Nezávislé jevy tedy můžeme libovolně sdružovat nebo pronikat (daný jev vždy sjednotíme nebo pronikneme vždy jen s jednou skupinou jevů), a můžeme je libovolně převracet na jejich doplňky. Výsledek jsou opět nezávislé jevy.

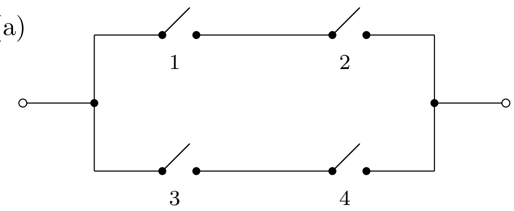
3.2 (operace s nezávislými jevy)
Čtyři spínače v zabezpečovacím zařízení pracují nezávisle, každý s pravděpodobností $p \in (0, 1)$. Jsou zapojeny (viz obrázek)

(a) po dvou sériově a pak paralelně

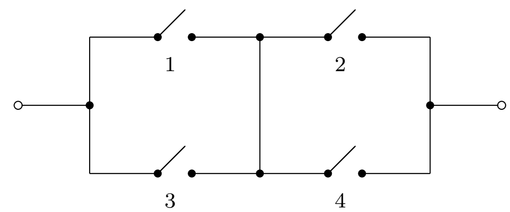
(b) po dvou paralelně a pak sériově.

S jakou pravděpodobností bude zařízení propouštět proud v jednotlivých případech? Pro které zapojení je tato pravděpodobnost větší?

(a)



(b)



Řešení:
Pro $i = 1, 2, 3, 4$ si označme jevy

A_i= "*i-tý spínač je zapnutý*"

B= "*zařízením prochází proud*"

Víme, že jevy $A_1, \ldots, A_4$ jsou nezávislé a $P(A_i) = p$.

(a) Aby proud procházel zařízením, musí jít buď horní větví nebo spodní větví:

$$B = (A_1 \cap A_2) \cup (A_3 \cap A_4)\ .$$

Pro pravděpodobnost pak (díky nezávislosti) máme

$$P(B) = P\Big((A_1 \cap A_2) \cup (A_3 \cap A_4)\Big) = P(A_1 \cap A_2) + P(A_3 \cap A_4) - P(A_1 \cap A_2 \cap A_3 \cap A_4) =$$

$$= p^2 + p^2 - p^4 = p^2(2 - p^2)\ .$$

(b) Aby proud procházel zařízením, musí projít levou částí a současně pravou částí:

$$B = (A_1 \cup A_3) \cap (A_2 \cup A_4)\ .$$

Z nezávislosti jevů A_i vyplývá, že jevy $A_1 \cup A_3$ a $A_2 \cup A_4$ jsou také nezávislé. Můžeme tak psát

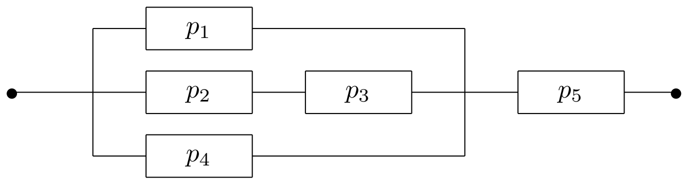
$$P(B) = P\Big((A_1 \cup A_3) \cap (A_2 \cup A_4)\Big) = P(A_1 \cup A_3) \cdot P(A_2 \cup A_4) =$$

$$= \Big(P(A_1) + P(A_3) - P(A_1) \cdot P(A_3)\Big) \cdot \Big(\cdots\Big) = (2p - p^2)^2 = p^2(2 - p)^2\ .$$

Už ze schématu zapojení je jasné, že obecně větší pravděpodobnost průchodu proudu zařízením je v případě (b), kde je jeden spoj navíc. To lze potvrdit i z vypočtené pravděpodobnosti:

$$p^2(2 - p^2) < p^2(2 - p)^2 \quad \Leftrightarrow \quad 0 < 2p^2(p - 1)^2\ .$$

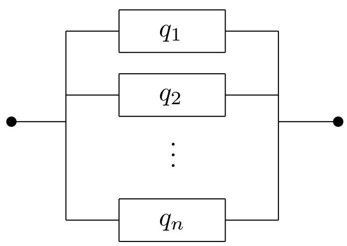
3.3 (operace s nezávislými jevy)
Zařízení na obrázku je tvořeno zapojením bloků, které pracují nezávisle na sobě a pravděpodobností výskytu poruch jsou zadány. Vypočtěte pravděpodobnost poruchy funkce celého zařízení.



Pravděpodobnosti vyčíslete pro $p_1 = 0.2$, $p_2 = p_3 = 0.4$, $p_4 = 0.3$ a $p_5 = 0.1$.

Řešení:
Úlohu si zjednodušíme tím, že budeme postupně nahrazovat více bloků jedním blokem, který bude mít stejnou pravděpodobnost poruchy.

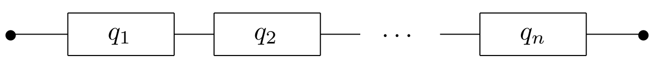
- Pro paralelní zapojení



a jevy $A_i =$"*i-tý blok (seshora) má poruchu*" je pravděpodobnost poruchy tohoto zapojení rovna

$$P(\text{\textit{"porucha paralelního zapojení"}}) = P(A_1 \cap \cdots \cap A_n) = P(A_1) \cdots P(A_n) = q_1 \cdots q_n \ .$$

- Pro sériové zapojení



a jevy $B_i =$"*i-tý blok (zleva) má poruchu*" je pravděpodobnost poruchy tohoto zapojení rovna

$$P(\text{\textit{"porucha sériového zapojení"}}) = P(B_1 \cup \cdots \cup B_n) = 1 - P\big(\overline{B_1 \cup \cdots \cup B_n}\big) =$$

$$= 1 - P(\overline{B_1} \cap \cdots \cap \overline{B_n}) = 1 - P(\overline{B_1}) \cdots P(\overline{B_n}) = 1 - (1 - q_1) \cdots (1 - q_n) \ .$$

Pro vyřešení původního zadání teď

(a) nejdříve nahradíme sériové zapojení dvou bloků s pravděpodobnostmi poruch $p_2 = 0.4$ a $p_3 = 0.4$ jediným blokem s pravděpodobností poruchy

$$p_{2,3} = 1 - (1 - p_2)(1 - p_3) = 1 - 0.6 \cdot 0.6 = 0.64 \ .$$

(b) dále nahradíme paralelní zapojení tří bloků s pravděpodobnostmi poruch $p_1 = 0.2$, $p_{2,3} = 0.64$ a $p_4 = 0.3$ jediným blokem s pravděpodobností poruchy

$$p_{1,2,3,4} = p_1 \cdot p_{2,3} \cdot p_4 = 0.2 \cdot 0.64 \cdot 0.3 = 0.0384 \ .$$

(c) a nakonec nahradíme sériové zapojení dvou bloků s pravděpodobnostmi poruch $p_{1,2,3,4} = 0.0384$ a $p_5 = 0.1$ jediným blokem, který odpovídá celému zařízení a má pravděpodobnost poruchy

$$p_{1,2,3,4,5} = 1 - (1 - p_{1,2,3,4})(1 - p_5) = 1 - 0.9616 \cdot 0.9 = 1 - 0.86544 = 0.13456 \ .$$

Důležitá poznámka: Pro jev $A \subseteq \Omega$, kde $P(A) \neq 0$, má funkce

$$\widetilde{P}(\cdot) := P(\,\cdot\,|A)$$

všechny vlastnosti pravděpodobnosti (pro množinu výsledků Ω a σ-algebru $\mathcal{A}$). Tedy podmíněná pravděpodobnost se chová v prvním argumentu jako obyčejná pravděpodobnost. Pozor, pro druhý argument už podobné chování neplatí!

V následujících větách používáme tento základní vztah pro jevy A a B:

$$P(B \cap A) = P(B|A) \cdot P(A)$$

pokud je podmíněná pravděpodobnost $P(B|A)$ definovaná, tj. pokud je $P(A) \neq 0$.

- -

Věta o úplné pravděpodobnosti: Nechť $A_1, \ldots, A_n$ je úplný disjunktní systém jevů na prostoru všech výsledků Ω (tedy jejich sjednocením je celé Ω a jevy jsou pro dvou disjunktní).
Nechť $P(A_i) \neq 0$ pro všechna i. Pak pro každý jev $B \subseteq \Omega$ platí

$$P(B) = \sum_{k=1}^{n} P(B \cap A_k) = \sum_{k=1}^{n} P(B|A_k) \cdot P(A_k)\ .$$

Bayesova věta: Pro jevy A a B platí

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

pokud $P(A) \neq 0$ a $P(B) \neq 0$.

- -

A ve spojení s větou o úplné pravděpodobnosti máme

$$P(A_i|B) = \frac{P(B|A_i) \cdot P(A_i)}{\sum\limits_{k=1}^{n} P(B|A_k) \cdot P(A_k)}\ .$$

(Všimněte si, že výraz v čitateli je jedním ze sčítanců ve jmenovateli.)

3.4 (bayesovská pravděpodobnost)

Máme 3 krabice stejného vzhledu. V první jsou 3 bílé a 2 černé koule, ve druhé jsou 2 bílé a 2 černé koule, ve třetí je 1 bílá a 4 černé koule.

(a) Určete pravděpodobnost, že z náhodně vybrané krabice náhodně vytáhneme bílou kouli. (Jaká by byla pravděpodobnost vytažení bílé koule, pokud bychom všechny koule z krabic vysypali do urny a z té bychom vybírali?)

(b) Pokud nám někdo řekl, že náhodně vybral jednu z krabic a vytáhl 1 kouli, která byla bílá, s jakou pravděpodobností můžeme usuzovat, že v téže krabici se nachází alespoň 3 černé koule?

Řešení:
Označme jevy:

$A_i =$*"byla vybrána i-tá krabice"*,
$B =$*"koule vytažená z vybrané krabice je bílá"*.

pro $i = 1, 2, 3$. Víme, že A_1, A_2 a A_3 je úplný disjunktní systém jevů, o kterých předpokládáme, že jsou stejně pravděpodobné. Tedy

$$P(A_1) = P(A_2) = P(A_3) = \tfrac{1}{3}$$

$$P(B|A_1) = \tfrac{3}{5} \quad P(B|A_2) = \tfrac{2}{4} \quad P(B|A_3) = \tfrac{1}{5} \ .$$

(a) Z věty o úplné pravděpodobnosti máme

$$P(B) = \sum_{i=1}^{3} P(B|A_i) \cdot P(A_i) = \tfrac{1}{3} \cdot \left(\tfrac{3}{5} + \tfrac{2}{4} + \tfrac{1}{5} \right) = \tfrac{13}{30} \doteq 0.4333 \ .$$

Jestliže sesypeme koule do jedné nádoby budeme v ní mít $3 + 2 + 2 + 2 + 1 + 4 = 15$ koulí a z toho $3+2+1 = 6$ bílých. Pravděpodobnost vytažení bílé koule je pak $\frac{6}{15} = 0.4$ (což je jiné číslo než spočítané $P(B) \doteq 0.4333$). Uvědomme si, že zde vybíráme koule jinak. Předtím jsem měli rovnocenné krabice, ale nerovnocenné koule v nich (protože v každé krabici je obecně různý počet koulí, viz Kolmogorův model níže) a v případě sesypanání máme zase rovnocenné koule.

(b) Protože jediná krabice obsahující alespoň 3 černé koule je třetí krabice, zajímá nás $P(A_3|B)$. Z Bayesovy věty máme:

$$P(A_3|B) = \frac{P(B|A_3) \cdot P(A_3)}{P(B)} = \frac{\frac{1}{5} \cdot \frac{1}{3}}{\frac{13}{30}} = \frac{2}{13} = 0.1538 \ .$$

Ještě si můžeme pro zajímavost sestrojit příslušný Kolmogorův model:

- elementární jev ω bude dvojice "(výběr dané krabice, vytažení koule z této krabice)"a Ω bude tedy množina všech takových ω,
- $A_i = \{\omega \in \Omega \mid \omega = (\text{výběr } i\text{-té krabice, vytažení koule z této krabice})\}$,
- pro počet krabic $k = 3$ a počet koulí k_i v i-té krabici pak pro $\omega \in A_i$ máme $P(\{\omega\}) = \frac{1}{k \cdot k_i}$,
- pravděpodobnost jevu A_i, tedy výběr i-té krabice, pak je $P(A_i) = \sum_{\omega \in A_i} \frac{1}{k \cdot k_i} = \frac{k_i}{k \cdot k_i} = \frac{1}{k}$.

Můžeme si tedy všimnout, že pravděpodobnosti vytažení koule z dané vybrané krabice nejsou všechny stejné, zatímco pravděpodobnosti výběru dané krabice ano.

3.5 (bayesovská pravděpodobnost)
Do obchodu dodávají čipy tři výrobci, po řadě 50 %, 30 % a 20 % zásoby obchodu. Pravděpodobnosti výroby funkčního čipu od jednotlivých výrobců jsou po řadě $p_1 = 0.98$, $p_2 = 0.95$ a $p_3 = 0.99$.

(a) Určete pravděpodobnosti, že zakoupený náhodně vybraný čip je vadný.

(b) Určete pravděpodobnost, že čip je od 2. výrobce, za předpokladu, že je funkční?

Řešení:
Označme si jevy:

$A_i =$ *"zakoupený čip byl od i-tého výrobce"*,
$B =$ *"zakoupený čip byl funkční"*.

Ze zadání plyne, že

$$P(A_1) = 0.5 \quad P(A_2) = 0.3 \quad P(A_3) = 0.2$$
$$P(B|A_1) = 0.98 \quad P(B|A_2) = 0.95 \quad P(B|A_3) = 0.99$$

(a) Počítáme pravděpodobnost jevu $\overline{B}$. Z věty o úplné pravděpodobnosti máme:

$$P(B) = \sum_{i=1}^{3} P(B|A_i) \cdot P(A_i) = 0.98 \cdot 0.5 + 0.95 \cdot 0.3 + 0.99 \cdot 0.2 = 0.973$$

a

$$P(\overline{B}) = 1 - P(B) = 1 - 0.973 = 0.027 \ .$$

Všimněme si ještě, že díky výpočtu $P(B)$ platí, že

$$0.95 = \min\{0.98, 0.95, 0.99\} \leq \underbrace{P(B)}_{0.973} \leq \max\{0.98, 0.95, 0.99\} = 0.99 \ .$$

Pokud nám tedy takovéto omezení nebude vycházet, někde jsme museli udělat chybu.

$P(\overline{B})$ jsme mohli spočítat i přímo jako

$$P(\overline{B}) = \sum_{i=1}^{3} \underbrace{P(\overline{B}|A_i)}_{1-P(B|A_i)} \cdot P(A_i) = 0.02 \cdot 0.5 + 0.05 \cdot 0.3 + 0.01 \cdot 0.2 = 0.027 \ .$$

(b) Zajímá nás $P(A_2|B)$. Z Bayesovy věty máme:

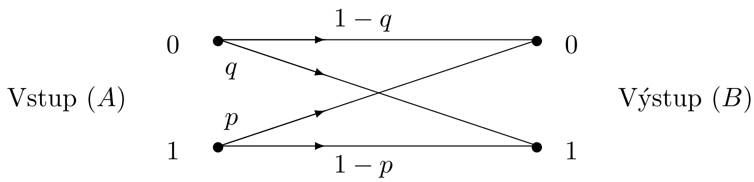
$$P(A_2|B) = \frac{P(B|A_2) \cdot P(A_2)}{P(B)} = \frac{0.95 \cdot 0.3}{0.973} = 0.293 \ .$$

I tady se podíváme, jak by vypadal ten nejjednodušší Kolmogorův model:

- elementární jev ω si můžeme definovat jako dvojici "(čip je od daného výrobce, čip je/není funkční)"a pro jednoduchost si je označme jako
$$\omega_i = (\text{čip je od } i\text{-tého výrobce, čip je funkční})$$
$$\omega_i' = (\text{čip je od } i\text{-tého výrobce, čip není funkční})$$
pro $i = 1, 2, 3$. Jevové pole pak bude $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_1', \omega_2', \omega_3'\}$. Má tedy 6 prvků.
- $A_i = \{\omega_i, \omega_i'\}$
- $B = \{\omega_1, \omega_2, \omega_3\}$
- pokud máme skutečně obdržet Kolmogorův model, musí pravděpodobnosti jednotlivých elementárních jevů nutně ze zadání být tyto:
$$P\Big(\{\omega_i\}\Big) = P(A_i \cap B) = P(B|A_i) \cdot P(A_i)$$
$$P\Big(\{\omega_i'\}\Big) = P(A_i \cap \overline{B}) = \Big(1 - P(B|A_i)\Big) \cdot P(A_i)$$
Takto definované pravděpodobnosti elementárních jevů nám pak zpětně poskytnou hodnoty pravděpodobností požadovaných v zadání:
$$P(A_i) = P\Big(\{\omega_i\}\Big) + P\Big(\{\omega_i'\}\Big)$$
$$P(B|A_i) = \frac{P(B \cap A_i)}{P(A_i)} = \frac{P\Big(\{\omega_i\}\Big)}{P\Big(\{\omega_i\}\Big) + P\Big(\{\omega_i'\}\Big)} \ .$$

3.6 (bayesovská pravděpodobnost v informačním kanálu se šumem)

Binární komunikační kanál přenáší symboly 0 a 1 ze vstupu (jevy A_0 a A_1) na výstup (jevy B_0 a B_1) podle uvedeného schématu, kdy je vlivem šumu s určitou pravděpodobností zaměněn symbol 1 za symbol 0 nebo naopak.



Nechť $1-q = P(B_0|A_0) = 0.9,\ \ 1-p = P(B_1|A_1) = 0.8.$

(a) Jestliže pravděpodobnosti na *vstupu* jsou $P(A_0) = 0.7,\ \ P(A_1) = 0.3$, určete

 (1) pravděpodobnosti $P(B_0)$ a $P(B_1)$ výskytu symbolů na výstupu

 (2) pravděpodobnost toho, že vyslaný symbol bude přenesen správně.

(b) Jestliže pravděpodobnosti na *výstupu* jsou $P(B_0) = 0.6,\ \ P(B_1) = 0.4$, určete pravděpodobnosti $P(A_0)$ a $P(A_1)$ výskytu symbolů na vstupu.

Řešení:

Pro $i = 0, 1$ máme jevy:

$A_i =$ *"byl vyslán znak i,"*

$B_i =$ *"byl přijat znak i."*

(a1) Pravděpodobnosti výskytu symbolů na výstupu určíme pomocí vzorce pro úplnou pravděpodobnost (tj. pomocí stochastických matic). Je tedy

$$\left[P\left(B_0\right), \quad P\left(B_1\right)\right] = \left[P\left(A_0\right), \quad P\left(A_1\right)\right] \cdot \begin{bmatrix} P\left(B_0|A_0\right) & P\left(B_1|A_0\right) \\ P\left(B_0|A_1\right) & P\left(B_1|A_1\right) \end{bmatrix} =$$

$$= \left[0.7, \quad 0.3\right] \cdot \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} = \left[0.69, \quad 0.31\right]$$

Stochastická matice $n \times k$ je taková, jejíž vstupy jsou nezáporné prvky a součty v rámci každého řádku se rovnají 1.

(a2) Zajímá nás jev

$$C = \text{"vyslaný znak bude přečten správně,"}$$

který je vyjádřen jako

$$C = (A_0 \cap B_0) \cup (A_1 \cap B_1)$$

Pravděpodobnost správného přenosu symbolů 0 a 1 tedy je

$$P(C) = P(A_0 \cap B_0) + P(A_1 \cap B_1) =$$
$$= P(B_0|A_0) \cdot P(A_0) + P(B_1|A_1) \cdot P(A_1) = 0.9 \cdot 0.7 + 0.8 \cdot 0.3 = 0.87\,.$$

(b) Podobně jako v (a1) máme

$$\left[P\left(B_0\right), \quad P\left(B_1\right)\right] = \left[P\left(A_0\right), \quad P\left(A_1\right)\right] \cdot \begin{bmatrix} P\left(B_0|A_0\right) & P\left(B_1|A_0\right) \\ P\left(B_0|A_1\right) & P\left(B_1|A_1\right) \end{bmatrix}$$

$$\left[0.6, \quad 0.4\right] = \left[P\left(A_0\right), \quad P\left(A_1\right)\right] \cdot \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}$$

tedy

$$\left[P\left(A_0\right), \quad P\left(A_1\right)\right] = \left[0.6, \quad 0.4\right] \cdot \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}^{-1} =$$

$$= \left[0.6, \quad 0.4\right] \cdot \frac{1}{0.72 - 0.02} \begin{bmatrix} 0.8 & -0.1 \\ -0.2 & 0.9 \end{bmatrix} = \left[\tfrac{4}{7}, \quad \tfrac{3}{7}\right]$$

Poznámka: Aby nedošlo k omylu - inverzní matice (ke stochastické matici typu "$P(B|A)$"), kterou jsem právě použili, NENÍ stochastická matice typu "$P(A|B)$"!

Stochastická matice typu "$P(B|A)$" se počítá pomocí Bayesovy věty, tedy speciálně závisí i na hodnotách $P(A_i)$ a nikoliv jen na hodnotách $P(B_i|A_j)$:

$$\begin{bmatrix} P\left(A_0|B_0\right) & P\left(A_1|B_0\right) \\ P\left(A_0|B_1\right) & P\left(A_1|B_1\right) \end{bmatrix} = \begin{bmatrix} \frac{1}{P(B_0)} & 0 \\ 0 & \frac{1}{P(B_1)}) \end{bmatrix} \cdot \underbrace{\begin{bmatrix} P\left(B_0|A_0\right) & P\left(B_0|A_1\right) \\ P\left(B_1|A_0\right) & P\left(B_1|A_1\right) \end{bmatrix}}_{\mathbb{A}} \cdot \begin{bmatrix} P\left(A_0\right) & 0 \\ 0 & P\left(A_1\right) \end{bmatrix}$$

Zde konkrétně vyjde jako:

$$\begin{bmatrix} P\left(A_0|B_0\right) & P\left(A_1|B_0\right) \\ P\left(A_0|B_1\right) & P\left(A_1|B_1\right) \end{bmatrix} = \begin{bmatrix} \frac{1}{0.6} & 0 \\ 0 & \frac{1}{0.4} \end{bmatrix} \cdot \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix} \cdot \begin{bmatrix} \frac{4}{7} & 0 \\ 0 & \frac{3}{7} \end{bmatrix} = \begin{bmatrix} \frac{6}{7} & \frac{1}{7} \\ \frac{1}{7} & \frac{6}{7} \end{bmatrix}$$

Zatímco inverze k typu "$P(B|A)$"$=\mathbb{A}^T$ (pokud vůbec má smysl - tj. pokud je $\mathbb{A}^T$ čtvercová matice a je regulární) je

$$\begin{bmatrix} P\left(B_0|A_0\right) & P\left(B_1|A_0\right) \\ P\left(B_0|A_1\right) & P\left(B_1|A_1\right) \end{bmatrix}^{-1} = \left(\mathbb{A}^T\right)^{-1} = \frac{1}{\det(\mathbb{A})} \cdot \begin{bmatrix} P\left(B_1|A_1\right) & -P\left(B_1|A_0\right) \\ -P\left(B_0|A_1\right) & P\left(B_0|A_0\right) \end{bmatrix}$$

a (obecně) to není vůbec stochastická matice (může obsahovat i záporné hodnoty)!

Uvědomme si ještě, že pokud stochastická matice typu "$P(A|B)$" má rozměr $n \times k$, pak analogická stochastická matice typu "$P(B|A)$" má rozměr $k \times n$!

3.7 (bayesovská pravděpodobnost v informačním kanálu se šumem)

Na vstupu informačního kanálu jsou posílány znaky "0"a "1", přitom znak "1"je posílán s pravděpodobností r. Na výstupu je daný znak přečten s pravděpodobností chyby $p = 0.1$, která nezávisí na frekvenci s jakou znak chodí (tj. na hodnotě r).

(a) Jaké jsou pravděpodobnosti výstupu při $r = 0.4$?

(b) Určete podmíněné pravděpodobnosti vstupu při známém výstupu, je-li $r = 0.4$ a je-li $r = 0.1$.

(c) Jestliže je pravděpodobnost znaku "0"na výstupu 0.8, jaké jsou pak pravděpodobnosti vstupů?

Řešení:
Máme jevy

$A_i =$ "*vyšleme znak i*",
$B_i =$ "*zachytíme znak i*",

kde i je nula nebo jednička. Víme, že

$$\overline{A_0} = A_1 \quad \overline{B_0} = B_1 \quad \text{a} \quad P(A_1) = r \qquad \text{(Toto je vlastnost zprávy.)}$$

Pravděpodobnost p chyby znaku "1"na výstupu je dána procentem zachycených znaku "0"v množině odeslaných znaku "1", tj.

$$p = \frac{P(B_0 \cap A_1)}{P(A_1)} = P(B_0|A_1) \qquad \text{(Toto je vlastnost přijímacího zařízení.)}$$

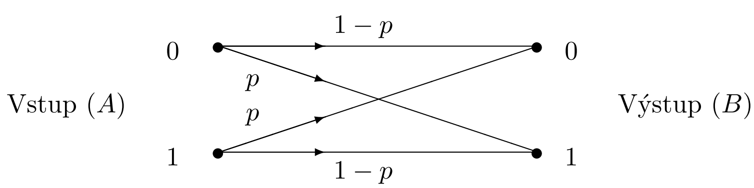
Podobně $p = P(B_1|A_0)$. Pro zjednodušení si uvědomíme, že funkce $\tilde{P}(A) := P(A|B)$ je pravděpodobnost v proměnné A, speciálně tedy

$$P(B_0|A_0) = 1 - P(B_1|A_0) = 1 - p$$

a

$$P(B_1|A_1) = 1 - P(B_0|A_1) = 1 - p.$$

Informační kanál je tedy popsán schématem:



(a) Pravděpodobnosti výskytu symbolů na výstupu určíme pomocí vzorce pro úplnou pravděpodobnost (tj. pomocí stochastických matic). Je tedy

$$\begin{bmatrix} P(B_0), & P(B_1) \end{bmatrix} = \begin{bmatrix} P(A_0), & P(A_1) \end{bmatrix} \cdot \begin{bmatrix} P(B_0|A_0) & P(B_1|A_0) \\ P(B_0|A_1) & P(B_1|A_1) \end{bmatrix} =$$
$$= \begin{bmatrix} 0.6, & 0.4 \end{bmatrix} \cdot \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix} = \begin{bmatrix} 0.58, & 0.42 \end{bmatrix}$$

Pamatujme na to, že ve stochastické matici jsou vždy nezáporná čísla a ŘÁDKY v součtu musí dávat vždy 1!

(b) Zajímají nás podmíněné pravděpodobnosti

$$P(A_i|B_j) \quad \text{pro} \quad i, j \in \{0, 1\} \qquad \text{(Toto zajímá toho, kdo zprávy přijímá.)}$$

Podle Bayesovy věty a věty o úplné pravděpodobnosti teď máme

$$P(A_0|B_0) = \frac{P(B_0|A_0)P(A_0)}{P(B_0|A_0)P(A_0) + P(B_0|A_1)P(A_1)} = \frac{(1-p)\cdot(1-r)}{(1-p)\cdot(1-r) + p \cdot r} = \frac{1}{1 + \frac{p \cdot r}{(1-p)\cdot(1-r)}}$$

a

$$P(A_1|B_1) = \frac{P(B_1|A_1)P(A_1)}{P(B_1|A_1)P(A_1) + P(B_1|A_0)P(A_0)} = \frac{(1-r)\cdot r}{(1-p)\cdot r + p \cdot (1-r)} = \frac{1}{1 + \frac{p\cdot(1-r)}{(1-p)\cdot r}} .$$

Pro zbylé podmíněné pravděpodobnosti máme opět vztahy $P(A_0|B_1) = 1 - P(A_1|B_1)$ a $P(A_1|B_0) = 1 - P(A_0|B_0)$.

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 $\frac{1}{1+\text{něco}}$.

- Pro $r = 0.4$ tak máme

$$P(A_0|B_0) = \frac{1}{1 + \frac{0.1 \cdot 0.4}{0.9 \cdot 0.6}} = \frac{27}{29} \doteq 0.93$$

$$P(A_1|B_0) = 1 - \frac{27}{29} = \frac{2}{29} \doteq 0.07$$

$$P(A_1|B_1) = \frac{1}{1 + \frac{0.1 \cdot 0.6}{0.9 \cdot 0.4}} = \frac{6}{7} \doteq 0.86$$

$$P(A_0|B_1) = 1 - \frac{6}{7} = \frac{1}{7} \doteq 0.14$$

  Tedy je to poměrně vysoká spolehlivost pro oba znaky.

- Pro $r = 0.1$ bude situace podstatně jiná:

$$P(A_0|B_0) = \frac{1}{1 + \frac{0.1 \cdot 0.1}{0.9 \cdot 0.9}} = \frac{81}{82} \doteq 0.99$$

$$P(A_1|B_0) = 1 - \frac{81}{82} = \frac{1}{82} \doteq 0.01$$

$$P(A_1|B_1) = \frac{1}{1 + \frac{0.1 \cdot 0.9}{0.9 \cdot 0.1}} = \frac{1}{2} = 0.5$$

$$P(A_0|B_1) = 1 - \frac{1}{2} = \frac{1}{2} = 0.5$$

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

(c) Podobně jako v (a) máme

$$\left[P\left(B_0\right), \quad P\left(B_1\right)\right] = \left[P\left(A_0\right), \quad P\left(A_1\right)\right] \cdot \begin{bmatrix} P\left(B_0|A_0\right) & P\left(B_1|A_0\right) \\ P\left(B_0|A_1\right) & P\left(B_1|A_1\right) \end{bmatrix}$$

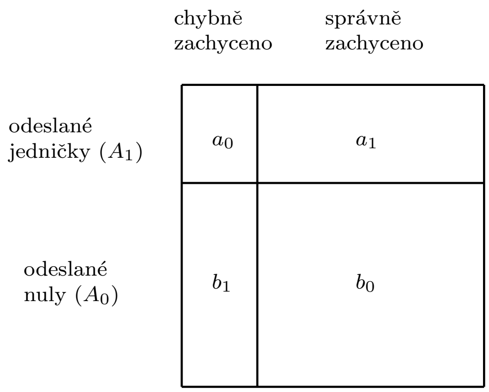
$$\left[0.8, \quad 0.2\right] = \left[P\left(A_0\right), \quad P\left(A_1\right)\right] \cdot \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix}$$

tedy

$$\left[P\left(A_0\right), \quad P\left(A_1\right)\right] = \left[0.8, \quad 0.2\right] \cdot \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix}^{-1} =$$

$$= \left[0.8, \quad 0.2\right] \cdot \frac{1}{0.81 - 0.01} \begin{bmatrix} 0.9 & -0.1 \\ -0.1 & 0.9 \end{bmatrix} = \left[\tfrac{7}{8}, \quad \tfrac{1}{8}\right]$$

Poznámka: Situaci z (b) si ještě můžeme znázornit obrázkem:



Zde a_i a b_j znamenají počty daných znaků v rámci daného jevu (např. b_1 je počet odeslaných znaků 0, které byly zachyceny jako znaky 1, neboli $|A_0 \cap B_1| = b_1$). Dále je např. $|A_0| = b_1 + b_0$ a $|B_1| = a_1 + b_1$. Speciálně máme

$$P(A_0|B_1) = \frac{|A_0 \cap B_1|}{|B_1|} = \frac{b_1}{a_1 + b_1}$$

a

$$P(A_1|B_1) = \frac{|A_1 \cap B_1|}{|B_1|} = \frac{a_1}{a_1 + b_1} \ .$$

Problém s rozpoznáním znaku 1 nastane právě když

$$\frac{a_1}{a_1 + b_1} = P(A_0|B_1) = P(A_1|B_1) = \frac{b_1}{a_1 + b_1}$$

tedy když

$$a_1 = b_1 \ .$$