# 3. cvičení z PST

18. října 2017

**3.1** (výběr s proměnnými podmínkami, Stirlingův vzorec)

V urně jsou dvě koule, bílá a černá. Provádí se výběr po jedné kouli do doby, než se vytáhne černá koule. Kdykoliv se vytáhne bílá koule, vrátí se do urny a přidají se ještě dvě bílé koule. Určete pravděpodobnost toho, že se při prvních 50 tazích černá koule nevytáhne.

**Řešení:**
Použijeme následující vztah, který si obecně hodí, pokud provádíme sérii pokusů:

Mějme posloupnost jevů $A_1 \supseteq A_2 \cdots \supseteq A_n$ kde $P(A_n) \neq 0$. Protože máme rovnost $A_i = A_i \cap A_{i-1}$, tak dostáváme

$$P(A_i) = \frac{P(A_i \cap A_{i-1})}{P(A_{i-1})} \cdot P(A_{i-1}) = P(A_i|A_{i-1}) \cdot P(A_{i-1}) \ .$$

Postupně tak iterací dostaneme

$$P(A_n) = P(A_1) \cdot P(A_2|A_1) \cdots P(A_n|A_{n-1}) \ .$$

Při praktickém použití pak jev $A_i$ znamená výsledky série prvních $i$ pokusů. Podmíněné pravděpodobnosti $P(A_i|A_{i-1})$ pak znamenají, jaká je pravděpodobnost výsledků $i$-tého pokusu za předpokladu, že už nastaly dané výsledky prvních $i-1$ pokusů.

V našem případě budeme mít

$A_i =$ "*v prvních i tazích se vytáhne bílá koule*",

podmíněné pravděpodobnosti pak budou

$$P(A_i|A_{i-1}) = \frac{2i-1}{2i}$$

protože po prvních $i-1$ pokusech v urně je $2i-1$ bílých koulí a 1 černá koule. A samozřejmě $P(A_1) = \frac{1}{2}$. Pro $n = 50$ tedy máme

$$P(A_n) = \frac{1}{2} \cdot \frac{3}{4} \cdot \frac{5}{6} \cdots \frac{2n-3}{2n-2} \cdot \frac{2n-1}{2n} = \frac{1 \cdot 2 \cdots (2n-1) \cdot 2n}{\big(2 \cdot 4 \cdots (2n-2) \cdot 2n\big)^2} = \frac{(2n)!}{(n!)^2 \cdot 2^{2n}} = \binom{2n}{n} \cdot \left(\frac{1}{2}\right)^{2n} \ .$$

Použitím Stirlingova vzorce $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$ dostaneme přibližnou hodnotu

$$P(A_n) = \frac{(2n)!}{(n!)^2 \cdot 2^{2n}} \doteq \frac{\sqrt{4\pi n}\left(\frac{2n}{e}\right)^{2n}}{2\pi n\left(\frac{n}{e}\right)^{2n} \cdot 2^{2n}} = \frac{1}{\sqrt{\pi n}} \ ,$$

tedy pro $n = 50$ to je

$$P(A_{50}) = \frac{100!}{(50!)^2 \cdot 2^{100}} \doteq \frac{1}{\sqrt{50\pi}} \doteq 0.0798 \ .$$

Současně je vidět, že pro $n \to \infty$ se $P(A_n)$ blíží k nule (přestože se podmíněná pravděpodobnost $P(A_i|A_{i-1}) = \frac{2i-1}{2i}$ vytažení bílé koule v dalším tahu postupně blíží k jedné pro $i \to \infty$.)

**3.2** ((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:**
Jevové pole 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

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

a podobně je to pro ostatní případy, zatímco

$$P(A \cap B \cap C) = 0 \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = P(A) \cdot P(B) \cdot P(C).$$

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

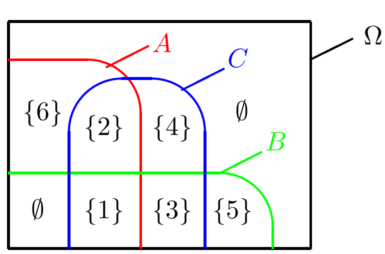
**3.3** (podmíněná nezávislost)

Nechť $\Omega = \{1, \ldots, 6\}$ je pravděpodobnostní prostor s Laplaceovou pravděpodobností.

(i) Položme $A = \{1, 2, 6\}$, $B = \{1, 3, 5\}$ a $C = \{1, 2, 3, 4\}$. Ukažte, že jevy $A$ a $B$ jsou *závislé* a přitom *podmíněně nezávislé* za podmínky $C$.

(ii) Položme $A' = \{1, 4\}$, $B' = \{1, 2, 3\}$ a $C' = A \cup B$. Ukažte, že jevy $A'$ a $B'$ jsou *nezávislé* a přitom *podmíněně závislé* za podmínky $C'$.

**Řešení:**
(i)



Máme $A \cap B = \{1\}$ a $|\Omega| = 6$ takže

$$P(A \cap B) = \frac{1}{6} \neq \frac{3}{6} \cdot \frac{3}{6} = P(A) \cdot P(B)$$

a jevy $A$ a $B$ tudíž jsou *závislé*.

Pro podmíněnou pravděpodobnost v Laplaceově modelu platí, že

$$P(X|C) = \frac{P(X \cap C)}{P(C)} = \frac{|X \cap C|}{|C|} \ .$$

Protože máme

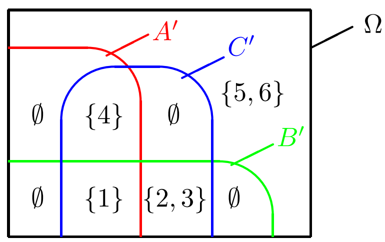
$$A \cap B \cap C = \{1\}, \quad A \cap C = \{1, 2\} \quad \text{a} \quad B \cap C = \{1, 3\}$$

dostaneme

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

Jevy $A$ a $B$ jsou tedy *podmíněně nezávislé* za podmínky $C$.

(ii)




Opět máme $A' \cap B' = \{1\}$ takže

$$P(A' \cap B') = \frac{1}{6} = \frac{2}{6} \cdot \frac{3}{6} = P(A') \cdot P(B')$$

a jevy $A'$ a $B'$ tudíž jsou *nezávislé*.

A dále máme $|C| = 4$

$$A \cap B \cap C = A \cap B = \{1\}, \quad A \cap C = A \quad \text{a} \quad B \cap C = B$$

dostaneme

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

Jevy $A$ a $B$ jsou tedy *podmíněně závislé* za podmínky $C$.

**3.4** (operace s nezávislými jevy)

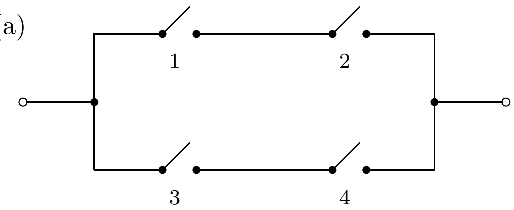
Čtyři spínače v zabezpečovacím zařízení pracují nezávisle, každý s pravděpodobností $p \in (0, 1)$. Jsou zapojeny (viz obrázek)

(a) po dvou sériově a pak paralelně

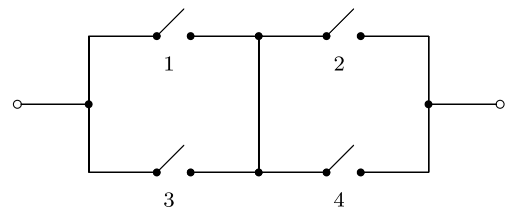
(b) po dvou paralelně a pak sériově.

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 $A_1, \dots, A_4$ 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

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

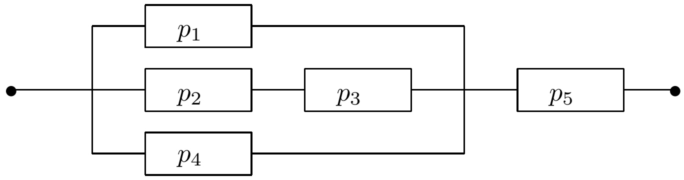
$$= \Big(P(A_1) + P(A_3) - P(A_1) \cdot P(A_3)\Big) \cdot \Big(\cdots\Big) = (2p - p^2)^2 = p^2(2 - p)^2 \ .$$

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:

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

Použili jsme to, že pokud jevy $A_1, \dots, A_n$ jsou nezávislé, pak také jevy $A_1 \cup A_2, A_3, \dots, A_n$ jsou nezávislé. Nezávislé jevy tedy můžeme libovolně sdružovat (daný jev vždy sjednotíme vždy jen s jednou skupinou jevu) a výsledek jsou opět nezávislé jevy.

**3.5** (operace s nezávislými jevy)
Zařízení na obrázku se tvořeno zapojením bloků, které pracují nezávisle na sobě a pravděpodobností výskytu poruch jsou zadány. Vypočtěte pravděpodobnost $P(A)$ funkčnosti systému a pravděpodobnost $P(\overline{A})$ poruchy funkce zařízení.




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

**Řešení:**
Příklad 19: `https://math.feld.cvut.cz/prucha/psi/wcv2.pdf`

**3.6** (geometrické rozdělení)

Bob a Alice házejí střídavě na koš, dokud se jeden z nich netrefí. Začíná Alice. Pravděpodobnost, že se při hodu trefí Alice je $a$, pravděpodobnost, že se při hodu trefí Bob je $b$.

Jaká je pravděpodobnost, že vyhraje Alice a jaká je pravděpodobnost, že vyhraje Bob?

**Řešení:**
Pořadí, jak jednotliví hráči hází míč, je toto (samozřejmě, pokud k hodu vůbec dojde):

1.hod: Alice, 2.hod: Bob, 3.hod: Alice, 4.hod: Bob, atd.

Označme si jevy

- $A$= "*vyhraje Alice*"
- $B$= "*vyhraje Bob*"
- $T_i$= "*proběhlo i hodů a při i-tém hodu se hráč, který zrovna hází, **trefí***"
- $N_i$= "*proběhlo i hodů a při i-tém hodu se hráč, který zrovna hází, **netrefí***"

Současně víme, že

$$A = \bigcup_{k=1}^{\infty} T_{2k-1}, \qquad B = \bigcup_{k=1}^{\infty} T_{2k}$$

a $\{T_i \mid i \in \mathbb{N}\}$ je disjunktní systém jevů.

Pozor, $T_i$ *nejsou* v důsledku toho nezávislé (protože $P(T_i \cap T_j) = P(\emptyset) = 0$ pro $i \neq j$)!

K výpočtu pravděpodobnosti $P(T_i)$ využijeme už známý trik, tj. klesající posloupnosti jevů:

$$N_1 \supseteq N_2 \supseteq \cdots \supseteq N_{i-1} \supseteq T_i$$

a toho, že známe podmíněné pravděpodobnosti mezi jednotlivými členy této posloupnosti (to jsou přesně ty pravděpodobnosti, že se při konkrétním hodu daný hráč netrefí, případně trefí):

$$P(T_i) = P(N_1) \cdot P(N_2|N_1) \cdots P(N_{i-1}|N_{i-2}) \cdot P(T_i|N_{i-1}) \ .$$

Dostaneme tak

$$P(T_{2k-1}) = \underbrace{(1-a)(1-b)\cdots(1-a)(1-b)}_{2(k-1)\ \text{členů}} \cdot a$$

$$P(T_{2k}) = \underbrace{(1-a)(1-b)\cdots(1-a)(1-b)}_{2(k-1)\ \text{členů}} \cdot (1-a)b$$

Výsledné pravděpodobnosti výher jsou tedy

$$P(A) = P\left(\bigcup_{k=1}^{\infty} T_{2k-1}\right) = \sum_{k=1}^{\infty} P(T_{2k-1}) = \sum_{k=1}^{\infty} \Big((1-a)(1-b)\Big)^{k-1} \cdot a =$$

$$= \frac{a}{1-(1-a)(1-b)} = \frac{a}{a+b-ab} \ .$$

a

$$P(B) = P\left(\bigcup_{k=1}^{\infty} T_{2k}\right) = \sum_{k=1}^{\infty} P(T_{2k}) = \sum_{k=1}^{\infty} \Big((1-a)(1-b)\Big)^{k-1} \cdot (1-a)b =$$

$$= \frac{(1-a)b}{1-(1-a)(1-b)} = \frac{b-ab}{a+b-ab} \ .$$

Současné vidíme, že $P(A) + P(B) = 1$ a tedy pravděpodobnost toho, že hra neskončí a bude se házet nekonečněkrát, je nulová.

A kde se objevuje to *geometrické rozdělení*? Uvažujme Alici a Boba jako jednu skupinu a hru rozdělme na kola, kde kolo bude znamenat, že vždy seskupíme dohromady nejdříve hod Alice a pak hod Boba. Pravděpodobnost $q$, že (pokud se kolo uskuteční) tak se alespoň jeden ze skupiny v daném kole trefí, je zřejmé $q = a + (1-a)b = a + b - ab$ (buď se trefí Alice anebo ne a pak se musí trefit Bob).

Uvažujme teď náhodnou veličinu $X : \Omega \to \mathbb{R}$ definovanou jako

$X =$"*počet neúspěšných kol před prvním úspěšným kolem*".

Pak je

$$P(X = k) = (1-q)^k q$$

což znamená, že veličina $X$ má geometrické rozdělení s kvocientem $q$.

(Ve skutečnosti jsme při definici veličiny $X$ vynechali případ, kdy hra neskončí. Tento případ má ale nulovou pravděpodobnost, takže na definici veličiny zde nezáleží - na jakékoliv výpočty s $X$ to nebude mít vliv.)

**Úloha se dá ještě řešit následujícím způsobem:**

Pokud se uskuteční první dva neúspěšné hody (tj. nastane jev $N_2$), můžeme si představit, že hra vlastně začíná znovu. To vyjádříme takto:

$$P(A|N_2) = P(A) \quad \text{a} \quad P(B|N_2) = P(B) \ .$$

Díky disjunktnosti jevů $T_i$ máme

$$P(A) = P(T_1) + P\left(\bigcup_{k=2}^{\infty} T_{2k-1}\right)$$

a zřejmě platí

$$\bigcup_{k=2}^{\infty} T_{2k-1} = \text{“\textit{Alice vyhraje při jiném než 1. hodu}”} = A \cap N_2$$

Takže můžeme psát

$$P\left(\bigcup_{k=2}^{\infty} T_{2k-1}\right) = P(A \cap N_2) = P(A|N_2) \cdot P(N_2) =$$

$$= P(A) \cdot (1-a)(1-b) \ .$$

Dosazením tak dostáváme poměrně intuitivní vztah

$$P(A) = a + P(A) \cdot (1-a)(1-b) \ ,$$

ze kterého vypočítáme

$$P(A) = \frac{a}{a+b-ab}\ .$$

Podobně určíme druhou pravděpodobnost z rovnice

$$P(B) = (1-a)b + P(B)\cdot(1-a)(1-b)$$

jako

$$P(B) = \frac{b-ab}{a+b-ab}\ .$$

**3.7** (bayesovská pravděpodobnost)

Máme 4 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, ve čtvrté 5 bílých a 1 černá koule. 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 $i = 1,2,3,4$. Víme, že $A_1$, $A_2$, $A_3$ a $A_4$ je úplný disjunktní systém jevů, o kterých předpokládáme, že jsou stejně pravděpodobné. Tedy

$$P(A_1) = P(A_2) = P(A_3) = P(A_4) = \frac{1}{4}$$

$$P(B|A_1) = \frac{3}{5} \quad P(B|A_2) = \frac{2}{4} \quad P(B|A_3) = \frac{1}{5} \quad P(B|A_4) = \frac{5}{6}\ .$$

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

$$P(A_3|B) = \frac{P(B|A_3)\cdot P(A_3)}{P(B)}$$

$$P(B) = \sum_{i=1}^{4} P(B|A_i)\cdot P(A_i) = \frac{1}{4}\cdot\left(\frac{3}{5}+\frac{2}{4}+\frac{1}{5}+\frac{5}{6}\right) = \frac{8}{15} \doteq 0.5333$$

a tedy

$$P(A_3|B) = \frac{1/5\cdot 1/4}{8/15} = \frac{3}{32} = 0.09375\ .$$

Ještě je dobré uvědomit si, jak vypadá Kolmogorův model, speciálně jaké jsou pravděpodobnosti elementárních jevů:

- elementární jev $\omega$ bude dvojice "(výběr krabice, vytažení koule z této krabice)" a $\Omega$ bude tedy množina všech takových $\omega$,
- $A_i = \{\omega \in \Omega \mid \omega = (\text{výběr } i\text{-té krabice, vytažení koule z této krabice})\}$,
- pro počet krabic $k = 4$ 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 pravděpodobnost vytažení koule z dané vybrané krabice nejsou všechny stejné, zatímco pravděpodobnosti výběru dané krabice ano.

Jak by situace vypadala, kdybychom v zadání předpokládali, že koule nevybíráme z krabic, ale sesypeme je do jedné nádoby, přičemž si na každou z nich napíšeme z jaké krabice pochází? Pravděpodobností se pak změní:

- elementární jev $\omega'$ bude dvojice "(číslo krabice, koule pocházející z krabice s tímto číslem)"a $\Omega'$ bude tedy množina všech takových $\omega'$,
- jev $A'_i = \{\omega' \in \Omega' \mid \omega' = (i, \text{koule pocházející z } i\text{-té krabice})\}$ představuje všechny koule, které pocházejí z $i$-té krabice,
- protože nyní taháme koule z nádoby, budeme považovat pravděpodobnost jejich vytažení za stejnou pro všechny koule, tj. $P(\{\omega\}) = \frac{1}{K}$, kde $K = \sum_j k_j$,
- pravděpodobnost jevu $A'_i$, tedy podíl počtu koulí z $i$-té krabice, pak bude $P(A'_i) = \sum_{\omega \in A'_i} \frac{1}{K} = \frac{k_i}{K}$.

Sesypáním koulí jsme tak způsobili to, že šance na vytažení kterékoliv koule se teď vyrovnaly a naopak pravděpodobnosti "příslušející krabicím"jsou různé. Pokud bychom tedy řešili takto pozměněné zadání, tak pro jev

$B' =$ "koule vytažená z nádoby je bílá"

a úplný disjunktní systém jevů $A'_1$, $A'_2$, $A'_3$ a $A'_4$ dostáváme (s celkovým počtem koulí $5 + 4 + 5 + 6 = 20$), že

$$P(A'_1) = \frac{5}{20} \qquad P(A'_2) = \frac{4}{20} \qquad P(A'_3) = \frac{5}{20} \qquad P(A'_4) = \frac{6}{20}$$

$$P(B'|A'_1) = \frac{3}{5} \qquad P(B'|A'_2) = \frac{2}{4} \qquad P(B'|A'_3) = \frac{1}{5} \qquad P(B'|A'_4) = \frac{5}{6}\ .$$

Opět jediná krabice obsahující alespoň 3 černé koule je třetí krabice, takže nás bude zajímat $P(A'_3|B')$. Z Bayesovy věty máme:

$$P(A'_3|B') = \frac{P(B'|A'_3) \cdot P(A'_3)}{P(B')}$$

$$P(B') = \sum_{i=1}^{4} P(B'|A'_i) \cdot P(A'_i) = \frac{3}{5} \cdot \frac{5}{20} + \frac{2}{4} \cdot \frac{4}{20} + \frac{1}{5} \cdot \frac{5}{20} + \frac{5}{6} \cdot \frac{6}{20} = \frac{11}{20} = 0.55$$

(což se snadněji dalo spočítat prostě podílem bílých koulí v nádobě jako $P(B') = \frac{3+2+1+5}{20} = \frac{11}{20}$) a tedy

$$P(A'_3|B') = \frac{1/5 \cdot 5/20}{11/20} = \frac{1}{11} \doteq 0.0909$$

(což se opět snadněji dalo spočítat prostě podílem bílých koulí pocházejících ze 3. krabice v rámci všech bílých koulí, tj. $P(A'_3|B') = \frac{1}{3+2+1+5} = \frac{1}{11}$).

**3.8** (bayesovská pravděpodobnost)

Po skončení aktivní služby odchází do důchodu 60 námořních kapitánů. Z této skupiny jich 5 zažilo ztroskotání. Podle statistiky při ztroskotání zahyne třetina kapitánů. Odhadněte pravděpodobnost, že kapitán během své aktivní služby zažije ztroskotání. (Možnost opakovaného ztroskotání a úmrtí z jiné příčiny během aktivní služby zanedbáváme.)

**Řešení:**

Uvažujme jevy:

$A =$ "kapitán se dožije důchodu",
$B =$ "kapitán zažije ztroskotání" .

Ze zadání máme vztahy

$$P(B|A) = \frac{5}{60} = \frac{1}{12}\ , \qquad P(\overline{A}|B) = \frac{1}{3} \quad \text{a} \quad \overline{A} \subseteq B$$

kde poslední vztah odpovídá tomu, že během aktivní služby nemůže nastat úmrtí z jiné příčiny než kvůli ztroskotání. Z posledního vztahu plyne také $\overline{B} \subseteq A$ a tudíž dostáváme tyto podmíněné pravděpodobnosti

$$P(B|\overline{A}) = \frac{P(B \cap \overline{A})}{P(\overline{A})} = 1 \quad \text{a} \quad P(A|\overline{B}) = \frac{P(A \cap \overline{B})}{P(\overline{B})} = 1 \ .$$

Nás zajímá $P(B)$. Z Bayesovy věty máme:

$$P(B) = \frac{P(B|A)}{P(A|B)} \cdot P(A) = \frac{\frac{1}{12}}{1 - \frac{1}{3}} \cdot P(A) = \frac{1}{8} \cdot P(A) \ .$$

Pomocí věty o úplné pravděpodobnosti můžeme teď zase $P(A)$ vyjádřit pomocí $P(B)$:

$$P(A) = P(A|B) \cdot P(B) + P(A|\overline{B}) \cdot P(\overline{B}) = \left(1 - \frac{1}{3}\right) \cdot P(B) + 1 \cdot \Big(1 - P(B)\Big) =$$

$$= 1 - \frac{1}{3} P(B) \ .$$

Celkem tedy

$$P(B) = \frac{1}{8}\Big(1 - \frac{1}{3}P(B)\Big) \quad \text{a} \quad P(B) = \frac{3}{25} = 0.12 \ .$$

Použitý vzorec je obecně:

$$P(B) = \frac{P(B|A)}{P(A|B)} \cdot P(A) = \frac{P(B|A)}{P(A|B)} \cdot \Big[P(A|B) \cdot P(B) + P(A|\overline{B}) \cdot \big(1 - P(B)\big)\Big]$$

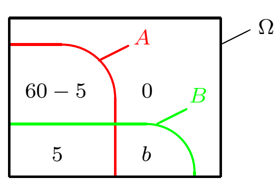
neboli

$$\color{red}{P(B)} = \frac{P(B|A) \cdot P(A|\overline{B})}{P(B|A) \cdot P(A|\overline{B}) + \big(1 - P(B|A)\big) \cdot P(A|B)} =$$

$$= \color{red}{\frac{P(B|A) \cdot P(A|\overline{B})}{P(B|A) \cdot P(A|\overline{B}) + P(\overline{B}|A) \cdot P(A|B)}} =$$

$$= \frac{\frac{1}{12} \cdot 1}{\frac{1}{12} \cdot 1 + \frac{11}{12} \cdot \frac{2}{3}} = \frac{3}{25} \ .$$

Můžeme také použít intuitivnější přístup:




kde čísla znamenají velikost dané množiny ve smyslu geometrické pravděpodobnosti (např. $\mathrm{vol}(A \cap B) = 5$ apod.). Přitom víme ještě, že

$$\frac{1}{3} = P(\overline{A}|B) = \frac{b}{5 + b}$$

takže

$$\mathrm{vol}(B \setminus A) = b = 2.5 \ .$$

Proto máme

$$P(B) = \frac{5 + b}{60 + b} = \frac{5 + 2.5}{60 + 2.5} = \frac{3}{25} \ .$$

**3.9** (bayesovská pravděpodobnost)

U 10% řidičů, kteří způsobili dopravní nehodu, bylo prokázáno požití alkoholu. Rozsáhlý průzkum ukázal, že riziko nehody se požitím alkoholu zvyšuje 7×. Odhadněte, kolik procent řidičů požilo alkohol.

**Řešení:**
Označme jevy

$A =$ "požil alkohol,"

$H =$ "způsobil nehodu."

Pak máme $P(A|H) = 0.1$ a $P(H|A) = 7 \cdot P\left(H|\overline{A}\right)$. Tudíž

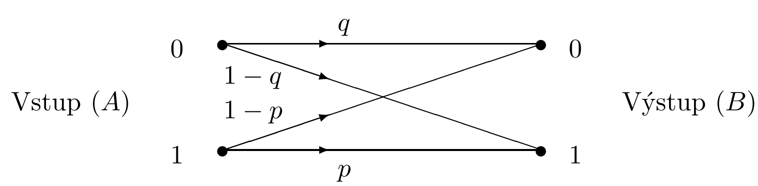
$$0.1 = P(A|H) = \frac{P(H|A) \cdot P(A)}{P(H|A) \cdot P(A) + P\left(H|\overline{A}\right) \cdot P\left(\overline{A}\right)} = \frac{7 \cdot P(A)}{7 \cdot P(A) + (1 - P(A))} \ .$$

Výsledek je

$$P(A) = \frac{1}{64} \ .$$

**3.10** (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.




Jestliže je $q = P(B_0|A_0) = 0.9$, $p = P(B_1|A_1) = 0.8$ a $P(A_0) = 0.7$, $P(A_1) = 0.3$, určete:

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

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

**Řešení:**
Pro $i = 0, 1$ máme jevy:

$A_i =$ "byl vyslán znak $i$,"

$B_i =$ "byl přijat znak $i$."

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

$$\begin{bmatrix} P(B_0) & P(B_1) \end{bmatrix} = \begin{bmatrix} P(A_0) & P(A_1) \end{bmatrix} \cdot \begin{bmatrix} P(B_0|A_0) & P(B_1|A_0) \\ P(B_0|A_1) & P(B_1|A_1) \end{bmatrix} =$$

$$= \begin{bmatrix} 0.7 & 0.3 \end{bmatrix} \cdot \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} = \begin{bmatrix} 0.69 & 0.31 \end{bmatrix}$$

(b) Pravděpodobnost $P$ správného přenosu je rovna součtu pravděpodobností správného přenosu symbolů 0 a 1, tedy

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

**3.11** (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

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

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

$$r = \frac{P(Z_0 \cap V_1)}{P(V_1)} = P(Z_0|V_1)$$ (Toto je vlastnost přijímacího zařízení.)

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 pravděpodobnosti

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

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:

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

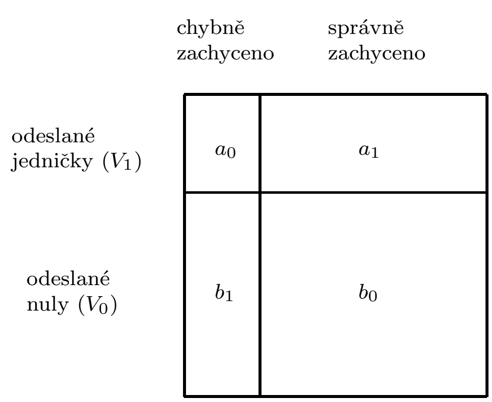
$$P(V_1|Z_0) = 1 - \frac{81}{82} = \frac{1}{82} \doteq 0.01$$

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

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

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").

Situaci 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 $|V_0 \cap Z_1| = b_1$). Dale je např. $|V_0| = b_1 + b_0$ a $|Z_1| = a_1 + b_1$. Speciálně máme

$$P(V_0|Z_1) = \frac{|V_0 \cap Z_1|}{|Z_1|} = \frac{b_1}{a_1 + b_1}$$

a

$$P(V_1|Z_1) = \frac{|V_1 \cap Z_1|}{|Z_1|} = \frac{a_1}{a_1 + b_1} \ .$$

Problém s rozpoznáním znaku 1 nastane právě když

$$\frac{a_1}{a_1 + b_1} = P(V_0|Z_1) = P(V_1|Z_1) = \frac{b_1}{a_1 + b_1}$$

tedy když

$$a_1 = b_1 \ .$$

**3.12** (stochastické matice)

Při korektuře dat se opraví 80% chybných položek. Při následném přepisování dat se do 1% položek dostanou nové chyby. Pro jaké počáteční procento chyb se po přepisu jejich počet sníží?

**Řešení:**

Příklad 21: `https://math.feld.cvut.cz/prucha/psi/wcv2.pdf`