# 4. cvičení z STP

9. - 13. březen 2020

**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 $\Omega$ a $\sigma$-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 $\Omega$ (tedy jejich sjednocením je celé $\Omega$ 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.)

**Příklad 4.1** *Na fakultě je 50% studujících na informatice, 30% na matematice a 20% na fyzice. Z těch, co studují na informatice je 10% žen, na matematice 30% a na fyzice 20%.*

*(a) Jaká je pravděpodobnost, že náhodně vybraný studující je žena?*

*(b) Jaká je pravděpodobnost, že náhodně vybraná studentka studuje matematiku?*

**Řešení:**
Označme si jevy

$A_1 =$ *"náhodně vybraný studující je informatik"*

$A_2 =$ *"náhodně vybraný studující je matematik"*

$A_3 =$ *"náhodně vybraný studující je fyzik"*

$B =$ *"náhodně vybraný studující je žena"*

Jevy $A_1, A_2, A_3$ jsou navzájem disjunktní a jejich sjednocením je celý pravděpodobnostní prostor $\Omega$ (tvořený všemi studujícími). Tedy máme úplný systém disjunktních jevů na $\Omega$. Dále známe

$$P(A_1) = 0.5 \qquad P(A_2) = 0.3 \qquad P(A_3) = 0.2$$
$$P(B|A_1) = 0.1 \quad P(B|A_2) = 0.3 \quad P(B|A_3) = 0.2$$

(a) Chceme znát $P(B)$. Podle věty o úplné pravděpodobnosti máme

$$P(B) = \sum_{j=1}^{3} P(B \cap A_j) = \sum_{j=1}^{3} P(B|A_j) \cdot P(A_j) =$$

$$= 0.1 \cdot 0.5 + 0.3 \cdot 0.3 + 0.2 \cdot 0.2 = 0.18\ .$$

(b) Chceme znát $P(A_2|B)$. Podle Bayesovy věty máme

$$P(A_2|B) = \frac{P(B|A_2) \cdot P(A_2)}{P(B)} = \frac{0.3 \cdot 0.3}{0.18} = 0.5\ .$$

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

Příklad můžeme řešit i bez explicitního použití výše uvedených vět a to pomocí velikosti množin. Tato velikost bude vyjadřovat "počty"studentů v dané množině ovšem s tím, že tento počet může být i desetinné číslo (používáme tak vlastně geometrický pravděpodobnostní model). Tato desetinná čísla pochopitelně nesmíme zaokrouhlovat!

Prostoru $\Omega$ přiřadíme nějakou velikost např. $vol(\Omega) = 100$ (obvykle je dobré si volit takovou velikost, kterou můžeme snadno dělit hodnotami ve jmenovatelích zlomků v zadaných pravděpodobnostech).

Teď si postupně můžeme začít vyplňovat tabulku níže:

- *počet studujících* $= vol(\Omega) = 100$
- *počet studujících na informatice* $= vol(A_1) = P(A_1) \cdot vol(\Omega) = 0.5 \cdot 100 = 50$ a podobně pro matematiku a fyziku
- *počet žen na informatice* $= vol(B \cap A_1) = P(B|A_1) \cdot vol(A_1) = 0.1 \cdot 50 = 5$ a podobně pro matematiku a fyziku
- *počet mužů na informatice* $= vol(B^c \cap A_1) = vol(A_1) - vol(B \cap A_1) = 50 - 5 = 45$ a podobně pro matematiku a fyziku

| | infor. | matem. | fyz. | |
|---|---|---|---|---|
| muži | 45 | 21 | 16 | 82 |
| ženy | 5 | 9 | 4 | 18 |
| | 50 | 30 | 20 | 100 |

Odsud pak ihned máme např. že

$$\text{\textit{počet žen}} = vol(B) = \sum_{i=1}^{3} vol(B \cap A_i) = 5 + 9 + 4 = 18$$

a tudíž

$$P(B) = \frac{\text{\textit{počet žen}}}{\text{\textit{počet studujících}}} = \frac{vol(B)}{vol(\Omega)} = \frac{18}{100} = 0.18$$

a dále

$$P(A_2|B) = \frac{\text{\textit{počet žen na matematice}}}{\text{\textit{počet žen}}} = \frac{vol(B \cap A_2)}{vol(B)} = \frac{9}{18} = 0.5\ .$$

**Příklad 4.2** *Na fakultě je 50% studujících na informatice, 30% na matematice a 20% na fyzice. Z těch, co studují na informatice je 10% žen a (podobně) na matematice 30% je žen. Mezi studujícími je celkově 80% mužů.*

*(a) Jaké procento z mužů studuje na matematice?*

*(b) Jaká je pravděpodobnost, že náhodně vybraný studující fyziky je muž?*

**Řešení:**
Opět si označme (jako u příkladu **4.1**) jevy

$A_1$ = *"náhodně vybraný student je informatik"*

$A_2$ = *"náhodně vybraný student je matematik"*

$A_3$ = *"náhodně vybraný student je fyzik"*

$B$ = *"náhodně vybraný student je žena"*

Opět máme úplný systém disjunktních $A_1, A_2, A_3$ jevů na $\Omega =$ *"všichni studující"*. Tentokrát známe tyto pravděpodobnosti

$$P(A_1) = 0.5 \qquad P(A_2) = 0.3 \quad P(A_3) = 0.2$$
$$P(B|A_1) = 0.1 \quad P(B|A_2) = 0.3$$
$$P(B^c) = 0.8$$

(a) Chceme znát $P(A_2|B^c)$. Podle Bayesovy věty máme

$$P(A_2|B^c) = \frac{P(B^c|A_2) \cdot P(A_2)}{P(B^c)} = \frac{(1 - 0.3) \cdot 0.3}{0.8} = \frac{21}{80} = 0.2625\ .$$

(b) Chceme znát $P(B^c|A_3)$. Protože neznáme $P(A_3|B^c)$, využijeme vztah

$$P(B^c|A_3) = \frac{P(B^c \cap A_3)}{P(A_3)}\ .$$

Z věty o úplné pravděpodobnosti máme

$$P(B^c \cap A_3) = P(B^c) - \sum_{j=1}^{2} P(B^c \cap A_j) = P(B^c) - \sum_{j=1}^{2} P(B^c|A_j) \cdot P(A_j) =$$

$$= 0.8 - \Big((1-0.1)\cdot 0.5 + (1-0.3)\cdot 0.3\Big) = 0.8 - 0.66 = 0.14$$

takže

$$P(B^c|A_3) = \frac{P(B^c \cap A_3)}{P(A_3)} = \frac{0.14}{0.2} = 0.7\ .$$

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

Příklad opět můžeme řešit i bez explicitního použití výše uvedených vět (viz příklad **4.1**). Prostoru $\Omega$ přiřadíme opět velikost např. $vol(\Omega) = 100$.

Teď postupně můžeme začít vyplňovat tabulku níže:

- *počet studujících* $= vol(\Omega) = 100$
- *počet studujících na informatice* $= vol(A_1) = P(A_1)\cdot vol(\Omega) = 0.5\cdot 100 = 50$ a podobně pro matematiku
- *počet žen na informatice* $= vol(B\cap A_1) = P(B|A_1)\cdot vol(A_1) = 0.1\cdot 50 = 5$ a podobně pro matematiku
- *počet mužů na informatice* $= vol(B^c\cap A_1) = vol(A_1) - vol(B\cap A_1) = 50 - 5 = 45$ a podobně pro matematiku
- *počet mužů* $= vol(B^c) = P(B^c)\cdot vol(\Omega) = 0.8\cdot 100 = 80$ a podobně pro ženy
- *počet mužů na fyzice* $= vol(B^c\cap A_3) = vol(B^c) - vol(B^c\cap A_1) - vol(B^c\cap A_2) = 80 - 45 - 21 = 14$ a podobně pro ženy na fyzice

| | infor. ($A_1$) | matem. ($A_2$) | fyz. ($A_3$) | |
|---|---|---|---|---|
| muži ($B^c$) | 45 | 21 | 14 | 80 |
| ženy ($B$) | $0.1\cdot 50 = 5$ | $0.3\cdot 30 = 9$ | 6 | 20 |
| | 50 | 30 | 20 | 100 |

Odsud pak ihned máme že

$$P(A_2|B^c) = \frac{\textit{počet mužů na matematice}}{\textit{počet mužů}} = \frac{vol(B^c\cap A_2)}{vol(B^c)} = \frac{21}{80} = 0.2625$$

a

$$P(B^c|A_3) = \frac{\textit{počet mužů na fyzice}}{\textit{počet studujících na fyzice}} = \frac{vol(B\cap A_3)}{vol(A_3)} = \frac{14}{14+6} = 0.7\ .$$

Často se Bayesova věta používá pro úplný disjunktní systém jevů $A$ a $A^c$, kde $0 < P(A) < 1$. Pak pro $B$ máme

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

pokud $P(B) \neq 0$.

**Příklad 4.3** *V populaci je infikována 1/4 jedinců, ale jen u 2/3 infikovaných se nákaza projevuje (a u žádných neinfikovaných).*

(a) *Jaká je pravděpodobnost, že jedinec bez příznaků není infikovaný?*

(b) *Jaká je pravděpodobnost, že jedinec je bez příznaků a současně je infikovaný? (Neboli: Jaký podíl v populaci tvoří ti, co jsou infikovaní a nemají příznaky?)*

**Řešení:**
Označme jevy

$A =$ "člověk je infikovaný,"

$B =$ "člověk má příznaky."

Ze zadání máme $P(A) = 1/4$, $P(B|A) = 2/3$ a $A^c \subseteq B^c$. Vztah $A^c \subseteq B^c$ (který, je ekvivalentní se vztahem $B \subseteq A$) znamená, že $A^c \cap B^c = A^c$. Odsud plyne, že

$$P(B^c|A^c) = \frac{P(A^c \cap B^c)}{P(A^c)} = \frac{P(A^c)}{P(A^c)} = 1 \ .$$

.

(a) Zde nás zajímá podmíněná pravděpodobnost $P(A^c|B^c)$. Z Baeysovy věty máme

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

$$= \frac{1 \cdot 3/4}{1 \cdot 3/4 + 1/3 \cdot 1/4} = 0.9 \ .$$

(b) Zajímá nás pravděpodobnost $P(B^c \cap A)$. Máme

$$P(B^c \cap A) = P(B^c|A) \cdot P(A) = \frac{1}{3} \cdot \frac{1}{4} = \frac{1}{12} \ .$$

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

Příklad opět můžeme řešit i bez explicitního použití výše uvedených vět (viz příklad **4.1**). Prostoru $\Omega$ přiřadíme opět velikost např. $vol(\Omega) = 100$.

Teď postupně můžeme začít vyplňovat tabulku níže:

- $\textit{velikost populace} = vol(\Omega) = 100$
- $vol(A) = P(A) \cdot vol(\Omega) = \frac{1}{4} \cdot 100 = 25$
- $vol(A^c) = vol(\Omega) - vol(A) = 100 - 25 = 75$
- $B \subseteq A \Rightarrow vol(B \cap A^c) = 0$
- $vol(B^c \cap A^c) = vol(A^c) - vol(A^c \cap B) = 75 - 0 = 75$
- $vol(B \cap A) = P(B|A) \cdot vol(A) = \frac{2}{3} \cdot 25$
- $vol(B^c \cap A) = P(B^c|A) \cdot vol(A) = \frac{1}{3} \cdot 25$

| | $A$ | $A^c$ | |
|---|---|---|---|
| $B^c$ | $\frac{1}{3} \cdot 25$ | 75 | |
| $B$ | $\frac{2}{3} \cdot 25$ | 0 | |
| | $\frac{1}{4} \cdot 100 = 25$ | 75 | 100 |

Odsud pak ihned máme že

$$P(A^c|B^c) = \frac{vol(A^c \cap B^c)}{vol(B^c)} = \frac{75}{\frac{1}{3} \cdot 25 + 75} = 0.9$$

a

$$P(B^c \cap A) = \frac{vol(B^c \cap A)}{vol(\Omega)} = \frac{\frac{1}{3} \cdot 25}{100} = \frac{1}{12} \ .$$

**Příklad 4.4** *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 druhého 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 $B^c$. 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

$$P(B^c) = 1 - P(B) = 1 - 0.973 = 0.027\ .$$

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

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

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

$P(B^c)$ jsme mohli spočítat i přímo jako

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

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

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

**Příklad 4.5** *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

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

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

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

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

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

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

$$P(A) = P(A|B) \cdot P(B) + P(A|B^c) \cdot P(B^c) = \left(1 - \tfrac{1}{3}\right) \cdot P(B) + 1 \cdot \left(1 - P(B)\right) =$$
$$= 1 - \tfrac{1}{3} P(B) \ .$$

Celkem tedy

$$P(B) = \tfrac{1}{8}\left(1 - \tfrac{1}{3}P(B)\right) \quad \text{a} \quad P(B) = \tfrac{3}{25} = 0.12 \ .$$

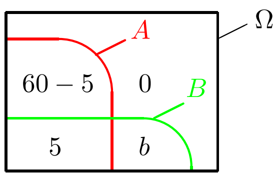
Použitý vzorec je obecně:

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

neboli

$$P(B) = \frac{P(B|A) \cdot P(A|B^c)}{P(B|A) \cdot P(A|B^c) + \big(1 - P(B|A)\big) \cdot P(A|B)} =$$
$$= \frac{P(B|A) \cdot P(A|B^c)}{P(B|A) \cdot P(A|B^c) + P(B^c|A) \cdot P(A|B)} =$$
$$= \frac{\frac{1}{12} \cdot 1}{\frac{1}{12} \cdot 1 + \frac{11}{12} \cdot \frac{2}{3}} = \frac{3}{25} \ .$$

Můžeme také použít intuitivnější přístup:



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

$$\frac{1}{3} = P(A^c|B) = \frac{P(A^c \cap B)}{P(B)} = \frac{\frac{vol(A^c \cap B)}{vol(\Omega)}}{\frac{vol(B)}{vol(\Omega)}} = \frac{vol(A^c \cap B)}{vol(B)} = \frac{b}{5+b}$$

takže

$$vol(B \setminus A) = vol(B \cap A^c) = b = 2.5 \ .$$

Proto máme

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

**Příklad 4.6** *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 $P(A|H) = 0.1$.

Dále je potřeba správně interpretovat údaj ve druhé větě "riziko nehody se požitím alkoholu zvyšuje". Při prvním (méně pozorném) pohledu to vypadá, že bychom mohli použít buď podmíněné pravděpodobnosti

"pravděpodobnost nehody, *jestliže* řidič požije/nepožije alkohol" (tedy hodnotu poměru $\frac{P(H|A)}{P(H|A^c)}$)

ale možná také přímé pravděpodobnosti

"pravděpodobnost, že řidič bude mít nehodu *a* současně požije/nepožije alkohol" (tedy hodnotu poměru $\frac{P(H \cap A)}{P(H \cap A^c)}$).

První možnost posuzuje pravděpodobnost z hlediska daného řidiče, který pochopitelně ví, jestli si dá/nedá alkohol, a tím i rozhoduje o dalších následcích. Tedy jev $A$ pro něj není daný náhodně (tím, že by řidič nevěděl, zda se uskuteční) ale naopak je daný volbou.

Z toho vyplývá i to, že druhou možnost nemůžeme popisovat z hlediska daného řidiče a musíme se na ni dívat "z venku". Druhá možnost (tedy průniky jevů) proto popisuje, jaké je procento všech jízd opilých/střízlivých řidičů, které skončily nehodou, v rámci všech možných jízd všech řidičů (třeba za nějaký rok).

Smyslem průzkumu ale jistě bylo spíš varovat před následky pití alkoholu před jízdou. To, že se jedná o podmíněnou pravděpodobnost nakonec dokládá i původní vyjádření "riziko nehody požitím alkoholu", které znamená "riziko nehody *za předpokladu* požití alkoholu".

Správná interpretace tak je, že

$$P(H|A) = 7 \cdot P(H|A^c) \ .$$

Nás nyní zajímá $P(A)$. Z Bayesovy věty a z věty o úplné pravděpodobnosti tak postupně máme

$$P(A) = \frac{P(A|H)}{P(H|A)} \cdot P(H) = \frac{P(A|H)}{P(H|A)} \cdot \Big[ P(H|A) \cdot P(A) + P(H|A^c) \cdot P(A^c) \Big] =$$

$$= P(A|H) \cdot \Big[P(A) + \frac{P(H|A^c)}{P(H|A)} \cdot \big(1 - P(A)\big)\Big] = 0.1 \cdot \big[P(A) + \tfrac{1}{7} \cdot \big(1 - P(A)\big)\big]$$

a tedy

$$10 \cdot P(A) = \tfrac{1}{7} + \tfrac{6}{7} \cdot P(A)$$

Výsledek je

$$P(A) = \frac{\frac{1}{7}}{10 - \frac{6}{7}} = \tfrac{1}{70-6} = \tfrac{1}{64} .$$

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

Také na to můžeme jít následujícím způsobem:

$$0.1 = P(A|H) = \frac{P(H|A) \cdot P(A)}{P(H|A) \cdot P(A) + P(H|A^c) \cdot P(A^c)} = \frac{7 \cdot P(A)}{7 \cdot P(A) + (1 - P(A))} .$$

opět s výsledkem

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

**Důležitá poznámka:** Náhodná veličina je funkce, která náhodnému výsledku (tj. elementárnímu jevu) přiřadí konkrétní hodnotu. Např. vybranému člověku přiřadí jeho tělesnou výšku. Jak je vidět, náhodná veličina vůbec není náhodná, co do hodnoty, kterou přiřazuje (ta je určena naprosto jasně). Náhodnost výstupu veličiny je dána náhodností jejího vstupu!

Abychom mohli s náhodnou veličinou $X$ vůbec pracovat, potřebujeme umět určovat pravděpodobnosti toho, že hodnoty $X$ budou např. v intervalu $\langle 1.5,\ 7.2) \subseteq \mathbb{R}$, což znamená, že množině $X^{-1}\langle 1.5,\ 7.2) = \{\omega \in \Omega \mid 1.5 \leq X(\omega) < 7.2\}$ musíme umět přiřadit její pravděpodobnost. To půjde jen tehdy, jestliže to bude množina měřitelná v daném systému jevů. Proto následující definice:

**Definice:** Náhodná veličina $X$ v pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ je takové zobrazení

$$X : \Omega \to \mathbb{R}$$

že vzor každého intervalu $I$ v $\mathbb{R}$ je "přípustná" množina (neboli jev), tj. $X^{-1}(I) \in \mathcal{A}$. Tuto vlastnost stačí ověřit jen pro určité typy intervalů v $\mathbb{R}$:

$$X \text{ je náhodná veličina} \quad \Leftrightarrow \quad (\forall t \in \mathbb{R}) \ \ X^{-1}\Big((-\infty, t\rangle\Big) \in \mathcal{A}$$

**Důležitá poznámka (možná i nejdůležitější vůbec....):** Co umožňuje náhodná veličina:

Náhodná veličina $X : \Omega \to \mathbb{R}$ umožňuje jednu podstatnou věc - zapomenout na původní prostor $\Omega$ všech výsledků (i s celou jeho strukturou jevů a jejich pravděpodobnostmi) a přitom stále umět vyčíslovat pravděpodobnosti pro $X$, které potřebujeme znát.

Veličina $X$ totiž umí přenést a vytvořit rozdělení pravděpodobnosti tam, kde má své hodnoty - konkrétně vytvoří nový pravděpodobnostní prostor na reálné přímce $\mathbb{R}$. A to tak, že každý interval $I \subseteq \mathbb{R}$ bude mít prostě pravděpodobnost $P_X(I) := P\Big(X^{-1}(I)\Big)$.

Pravděpodobnosti $P_X$ se říká rozdělení pravděpodobnosti veličiny $X$ (na $\mathbb{R}$). Toto rozdělení můžeme úplně popsat pokud opět známe pravděpodobnosti jen některých speciálních intervalů: opět jsou to intervaly $(-\infty, t\rangle$ pro $t \in \mathbb{R}$. Jejich pravděpodobnosti pak přirozeně definují tzv. *distribuční funkci* $F_X : \mathbb{R} \to \langle 0, 1\rangle$ jako

$$F_X(t) := P(X \leq t) \qquad \Big( = P\big(X^{-1}(-\infty, t\rangle\big)\Big)$$

*Toto je jeden z nejdůležitějších pojmů v celém kurzu! Proto se jej snažte pochopit.*

**Důležité vlastnosti distribuční funkce:** Pro distribuční funkci $F_X$ veličiny $X$ platí, že

- má hodnoty v intervalu $\langle 0, 1\rangle$,
- je neklesající,
- je zprava spojitá,
- $\lim_{t\to -\infty} F_X(t) = 0$ a $\lim_{t\to +\infty} F_X(t) = 1$.

(Tyto vlastnosti dokonce už distribuční funkce úplně charakterizují v tom smyslu, že každá funkce splňující výše uvedené vlastnosti je distribuční funkcí nějaké náhodné veličiny.)

**Příklad 4.7** *Uvažujme hod mincí s následujícími výsledky*

- $\omega_1 =$ *"padl rub" (s pravděpodobností* 0.49*)*
- $\omega_2 =$ *"padl líc" (s pravděpodobností* 0.49*)*
- $\omega_3 =$ *"nastala výjimečná situace" (hrana, zakutálení mince apod.)(s pravděpodobností* 0.02*).*

*Sestrojte dvě různé náhodné veličiny a nakreslete jejich distribuční funkce.*

**Řešení:**
Množina všech možných výsledků je tedy $\Omega = \{\omega_1, \omega_2, \omega_3\}$. Každá náhodná veličina na tomto prostoru $\Omega$ má maximálně tři různé hodnoty, takže určitě bude *diskrétní*

(tj. má nejvýše spočetně mnoho hodnot $A = \{x_1, x_2, \dots\} \subseteq \mathbb{R}$ takových, že $\sum_{u\in A} P(X = u) = 1$.)

Pro distribuční funkci diskrétní veličiny $X$ pak platí

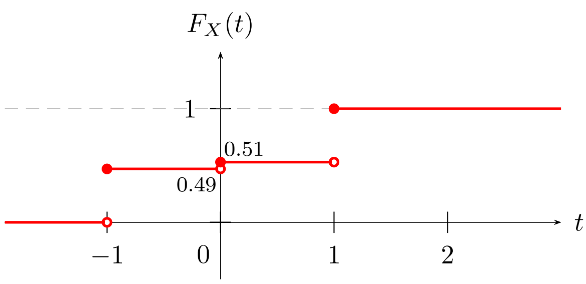
$$F_X(t) = P(X \le t) = \sum_{u\le t} P(X = u) \ .$$

Veličiny mohou být např.

- $X : \begin{array}{ccc} \omega_1 & \mapsto & 1 \\ \omega_2 & \mapsto & -1 \\ \omega_3 & \mapsto & 0 \end{array}$

Distribuční funkce je skokovitá se skoky v bodech $-1$, 0 a 1 o velikostech 0.49, 0.02 a 0.49, tj.

$$F_X(t) = \sum_{u\le t} P(X = u) = \begin{cases} & 0 \quad \text{pro} \quad t \in (-\infty, -1) \\ P(X=-1) = & 0.49 \quad \text{pro} \quad t \in \langle -1, 0) \\ P(X=-1) + P(X=0) = & 0.51 \quad \text{pro} \quad t \in \langle 0, 1) \\ P(X=-1) + P(X=0) + P(X=1) = & 1 \quad \text{pro} \quad t \in \langle 1, \infty) \end{cases}$$



- $Y: \begin{array}{ccc} \omega_1 & \mapsto & 1 \\ \omega_2 & \mapsto & 1 \\ \omega_3 & \mapsto & 3 \end{array}$

  Distribuční funkce je skokovitá se skoky v bodech 1 a 3 o velikostech 0.98 a 0.02 (protože obraz roven 1 mají dva elementární jevy $\omega_1$ a $\omega_2$, jejichž souhrnná pravděpodobnost je $0.49{+}0.49 = 0.98$), tj.

$$F_Y(t) = \sum_{u \leq t} P(Y = u) = \left\{ \begin{array}{llll} & 0 & \text{pro} & t \in (\infty, 1) \\ P(Y=1) = P(\{\omega_1, \omega_2\}) = & 0.98 & \text{pro} & t \in \langle 1, 3) \\ P(Y=1) + P(Y=3) = & 1 & \text{pro} & t \in \langle 3, \infty) \ . \end{array} \right.$$