# Řešená cvičení: NMAI059 Pravděpodobnost a statistika 1

Karel Král, TODO

3. března 2021

Tento text není určen k šíření. Všechny chyby v tomto textu jsou samozřejmě záměrné. Reportujte je prosím na adresu `kralka@iuuk.mff.cuni`....

# Obsah

**1 Zadání** **5**

1.1 1. Cvičení . . . . . . . . . . 5

**2 Tahák** **7**

2.1 Pravděpodobnostní prostor . . . . . . . . . . 7

2.2 Podmíněná pravděpodobnost . . . . . . . . . . 7

2.3 Bayesova věta . . . . . . . . . . 8

2.4 Nezávislé jevy . . . . . . . . . . 8

**3 Řešení** **9**

3.1 1. Cvičení . . . . . . . . . . 9

# Kapitola 1

# Zadání

## 1.1 Cvičení

1. Úvodní informace:
   (a) Slyšíte mě všichni dobře?
   (b) Literatura.
   (c) Pravidla zápočtu (domácí úkoly).

   Řešení: 1

2. Jak se generuje náhoda programem.
   - Python3
   - C++
   - R

   Řešení: 2

3. Připomeňte si definici pravděpodobnostního prostoru (Definice 2.1). Určete, co je
   - *množina elementárních jevů (sample space)*, tedy množina $\Omega$,
   - *prostor jevů (event space)*, tedy množina $\mathcal{F} \subseteq \mathcal{P}(\Omega)$,
   - *pravděpodobnost (probability)*, tedy funkce $\Pr\colon \mathcal{F} \to [0,1]$

   pro následující příklady:

   (a) Hod spravedlivou alkoholovou trojhrannou tužkou (není to kostka kvůli popisu prostoru jevů):

   ```
   import random
   drink = random.choice([
                   "světlé pivo",
                   "tmavé pivo",
                   "slivovice",
                   ])
   ```

   (b) Uniformně náhodné číslo z intervalu $[0,1)$.

   ```
   import random
   print(random.random())
   ```

Řešení: 3

4. Dokažte, že $\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B]$.

   Řešení: 4

5. Zopakujte si základní kombinatoriku:

   - Kolik je různých permutací množiny $\{A, B, C\}$?
   - Kolik různých slov skládajících se z písmen $\{A, B\}$ má délku 3?
   - Kolik různých podmnožin množiny $\{A, B, C, D, E\}$ má velikost 3?
   - Kolik různých kombinací s opakováním z množiny $\{A, B, C, D, E\}$ velikosti 3?

   Řešení: 5

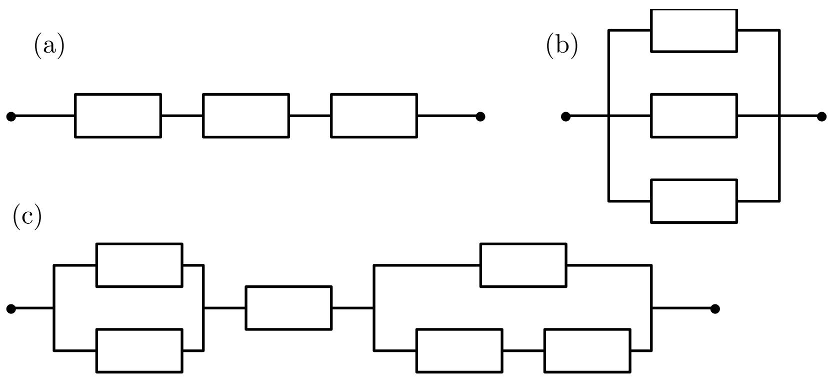
6. Jaká je pravděpodobnost, že při hodu šesti rozlišitelných spravedlivých šestistěnných kostek padnou aspoň na třech kostkách aspoň tři? Jaký je množina elementárních jevů, prostor jevů a pravděpodobnost?

   Řešení: 6

7. Nechť $\Omega$ jsou všechny permutace prvních 100 přirozených čísel, prostor jevů jsou všechny podmnožiny $\Omega$ a každý elementární jev je stejně pravděpodobný. Označme jev $A_j$ že náhodně zvolená permutace $\pi \in \Omega$ splňuje $\pi(j) = j$ (pro $1 \leq j \leq 100$). Jsou $A_1, A_2$ nezávislé jevy?

   Řešení: 7

8. Každý obdélník na obrázku je součástka, která se může porouchat s pravděpodobností $p$. Přesněji řečeno porucha znamená, že skrz ní neteče proud. Poruchy součástek jsou na sobě nezávislé. Jaká je pravděpodobnost, že stále poteče proud mezi dvěma puntíky.




   Řešení: 8

# Kapitola 2

# Tahák

## 2.1 Pravděpodobnostní prostor

**Definice.** Pravděpodobnostní prostor *je trojice* $(\Omega, \mathcal{F}, \Pr)$, *kde*

1. $\Omega$ *je* množina elementárních jevů *(sample space) (je to množina)*
2. $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ *je* prostor jevů *(event space) (je to množina podmnožin* $\Omega$*, jednotlivé prvky* $\mathcal{F}$ *nazýváme jevy)* $\mathcal{F}$ *je* $\sigma$*-algebra, tedy musí platit:*

   $\mathcal{F} = \{\emptyset, \{SU.P., TM.P., S.\}\}$

   Padlo sudé číslo $\in \mathcal{F}$

   jev $\to \{2, 4, 6\}$

   elem. jevy

   (a) $\emptyset \in \mathcal{F}$ *a zároveň* $\Omega \in \mathcal{F}$ *(celá množina elementárních jevů je jev)*

   (b) $A \in \mathcal{F} \Rightarrow (\Omega \setminus A) \in \mathcal{F}$ *(prostor jevů je uzavřený na doplňky)*

   (c) $(\forall n \in \mathbb{N} \colon A_n \in \mathcal{F}) \Rightarrow (\cup_{n \in \mathbb{N}} A_n) \in \mathcal{F}$ *(prostor jevů je uzavřený na spočetná sjednocení, poznámka může se stát, že* $A_1 \neq A_2 = A_3 = \cdots$*, specielně tedy je uzavřený i na všechna konečná sjednocení)*
3. $\Pr$ *je funkce* $\Pr \colon \mathcal{F} \to [0, 1]$ *je* pravděpodobnost *jevu z prostoru jevů, musí splňovat:*

   (a) $\Pr[\emptyset] = 0$ *a zároveň* $\Pr[\Omega] = 1$

   (b) $\Pr$ *je spočetně aditivní, tedy pro každou* $I \subseteq \mathbb{N}$ *a každou posloupnost jevů* $(A_j)_{j \in I}$*, které jsou po dvou disjunktní (tedy* $\forall i, j \in I \colon i \neq j \Rightarrow A_i \neq A_j$*) platí:*

$$\boxed{\Pr\left[\cup_{j \in I} A_j\right] = \sum_{j \in I} \Pr[A_j]}$$

! $A_i \cap A_j = \emptyset$ $i \neq j$

*(tedy* $\Pr$ *je pravděpodobnostní míra)*

## 2.2 Podmíněná pravděpodobnost

**Definice.** *Nechť* $(\Omega, \mathcal{F}, \Pr)$ *je pravděpodobnostní prostor. Nechť* $A, B \in \mathcal{F}$ *a navíc* $\Pr[B] > 0$*. Pak definujeme podmíněnou pravděpodobnost*

$$\Pr[A \mid B] = \frac{\Pr[A \cap B]}{\Pr[B]}$$

## 2.3 Bayesova věta

**Věta 1.** *Nechť* $(\Omega, \mathcal{F}, \Pr)$ *je pravděpodobnostní prostor. Nechť* $B_1, B_2, \ldots \in \mathcal{F}$ *je rozklad* $\Omega$ *(na spočetně mnoho množin). Nech* $A \in \mathcal{F}$ *a navíc platí* $\Pr[A] > 0$ *a navíc* $\Pr[B_i] > 0$ *pro každé* $i$*. Pak*

$$\Pr\left[B_j \mid A\right] = \frac{\Pr[A \mid B_j] \Pr[B_j]}{\sum_i \Pr[A \mid B_i] \Pr[B_i]}$$

## 2.4 Nezávislé jevy

**Definice.** *Nechť* $(\Omega, \mathcal{F}, \Pr)$ *je pravděpodobnostní prostor. Nechť* $A, B \in \mathcal{F}$ *jsou dva jevy. Pak řekneme, že* $A, B$ *jsou* nezávislé jevy*, pokud platí:*

$$\Pr[A \cap B] = \Pr[A] \Pr[B]$$

*(také se dá říct* $\Pr[A \mid B] = \Pr[A]$ *pokud* $\Pr[B] > 0$*).*

# Kapitola 3

# Řešení

## 3.1 Cvičení

1. **Úvodní informace:**
   (a) **Slyšíte mě všichni dobře?**
   (b) **Literatura.**
   ***Řešení:*** TODO viz přednáška
   (c) **Pravidla zápočtu (domácí úkoly).**
   ***Řešení:*** TODO ? DÚ, docházka, písemka (moodle), statistický program

2. **Jak se generuje náhoda programem.**

- **Python3** ***Řešení:***

```python
# https://docs.python.org/3/library/random.html
# Nepoužívejte pro šifrování!
import random

# https://docs.python.org/3/library/itertools.html
import itertools

# Generuj náhodné celé číslo 1 <= x <= 10 (tedy x in range(1, 11))
x = random.randint(1, 10)
print(x)

# Generuj náhodný prvek dané neprázdné posloupnosti.
drink = random.choice([
    "světlé pivo",
    "tmavé pivo",
    "slivovice",
    "bílé víno",
    "červené víno",
    "čaj",
    ])
print(drink)

# Výsledek 'B' je třikrát pravděpodobnější než 'A'.
print(random.choice(['A', 'B', 'B', 'B']))

# Vrátí list k náhodných prvků z dané sekvence s danými váhami
# (cum_weights jsou trochu rychlejší).
# https://docs.python.org/3/library/random.html#random.choices
print(random.choices(['A', 'B', 'C'], weights=[1/6, 1/6, 2/3], k=5))

# Náhodná permutace (mění přímo daný list).
my_list = ['a', 'b', 'c', 'd']
random.shuffle(my_list)
print(my_list)

# Vybere dva prvky dané posloupnosti, ve výsledku se nebudou opakovat
# pozice. Pokud se prvky opakují, tak se mohou ve výsledku opakovat.
print(random.sample(['w', 'x', 'y', 'z'], k=2))  # two distinct letters
print(random.sample(['w', 'x', 'y', 'x'], k=2))  # 'x' can repeat

# Pro přesné počítání (vrací iterable):
print(list(itertools.permutations([1, 2, 3])))
print(list(itertools.combinations('ABCD', 2)))
print(list(itertools.combinations_with_replacement('ABCD', 2)))
```

AA, AB, ~~BA~~ ...

```python
# Může se hodit scipy: sudo apt-get install python3-scipy
scipy.special.comb(n, k, exact=True)  # vrátí n nad k
```

- **C++** ***Řešení:***

```cpp
std::default_random_engine generator;
std::uniform_int_distribution<int> distribution(0, 9);
```

- **R** ***Řešení:***

```r
x <- "hello"
```

$(\Omega, \mathcal{F}, Pr)$

$\mathcal{F} \subseteq \mathcal{P}(\Omega)$ ← na přednášce P



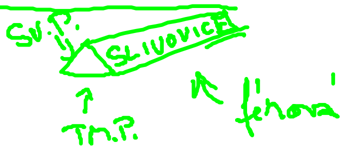
3. **Připomeňte si definici pravděpodobnostního prostoru (Definice 2.1). Určete, co je**
   - ***množina elementárních jevů (sample space)*, tedy množina $\Omega$,**
   - ***prostor jevů (event space)*, tedy množina $\mathcal{F} \subseteq \mathcal{P}(\Omega)$,**
   - ***pravděpodobnost (probability)*, tedy funkce $\Pr\colon \mathcal{F} \to [0,1]$**

   **pro následující příklady:**

   (a) **Hod spravedlivou alkoholovou trojhrannou tužkou (není to kostka kvůli popisu prostoru jevů):**

uniformní

```
import random
drink = random.choice([
                    "světlé pivo",
                    "tmavé pivo",
                    "slivovice",
                    ])
```




***Řešení:*** V počítači nám stačí předchozí kód (pseudonáhodné číslo vs náhodné číslo). Fyzicky můžeme generovat pomocí hodu trojhrannou tužkou, která má značky na stěnách (a zaručeně nemůže padnout nastojato).

- Množina elementárních jevů $\Omega = \{$světlé pivo, tmavé pivo, slivovice$\}$
- Prostor jevů $\mathcal{F} = \mathcal{P}(\Omega)$, tedy $|\mathcal{F}| = 2^{|\Omega|} = 2^3$:

$$\begin{aligned}
\mathcal{F} = \{ & \\
& \emptyset, \\
& \{\text{světlé pivo}\}, \\
& \{\text{tmavé pivo}\}, \\
& \{\text{slivovice}\}, \\
& \{\text{světlé pivo, tmavé pivo}\}, \\
& \{\text{světlé pivo, slivovice}\}, \\
& \{\text{tmavé pivo, slivovice}\}, \\
& \{\text{světlé pivo, tmavé pivo, slivovice}\} \\
\} &
\end{aligned}$$

- Pravděpodobnost

$$\begin{aligned}
\Pr[\emptyset] &= 0 \quad \text{viz def.} \\
\Pr[\{\text{světlé pivo}\}] &= 1/3 \\
\Pr[\{\text{tmavé pivo}\}] &= 1/3 \\
\Pr[\{\text{slivovice}\}] &= 1/3 \\
\Pr[\{\text{světlé pivo, tmavé pivo}\}] &= 2/3 \\
\Pr[\{\text{světlé pivo, slivovice}\}] &= 2/3 \\
\Pr[\{\text{tmavé pivo, slivovice}\}] &= 2/3 \\
\Pr[\{\text{světlé pivo, tmavé pivo, slivovice}\}] &= 1 \quad \text{viz def.}
\end{aligned}$$

$\Omega$ spočetná

$\mathcal{F} = \mathcal{P}(\Omega)$ je OK

ale nemám uniformní pravděpodobnost

$\Omega = \mathbb{N}$, $Pr[\{1\}] = \frac{1}{2}, \dots Pr[\{k\}] = \frac{1}{2^k}$

$Pr[\Omega] = 1$

$Pr[\{1,3\}] = \frac{1}{2} + \frac{1}{8}$

Většinou ale nepopisujeme jev jako podmnožinu elementárních jevů, ale nějak lidsky:

$$\Pr[\text{nějaké pivo}] = \Pr[\text{světlé pivo} \vee \text{tmavé pivo}] = \Pr[\{\text{světlé pivo, tmavé pivo}\}] = 2/3$$

„$\mathcal{F}$ jsou jevy jejichž pravděpodobnost uvažuju“

$$\Pr[\text{ne pivo}] = \Pr[\{\text{slivovice}\}] = 1 - \Pr[\{\text{světlé pivo, tmavé pivo}\}] = 1/3$$

(b) **Uniformně náhodné číslo z intervalu** $[0,1)$**.**

```python
import random
print(random.random())
```

disk $= \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 \leq 1\}$

***Řešení:*** V počítači nám funkce `random.random` vrátí float $x$, který je přesně reprezentovatelný, platí $0.0 \leq x < 1.0$ a zároveň $x$ je celočíselný násobek $2^{-53}$. To ale znamená, že nikdy nedostaneme 0.05954861408025609 i když to je číslo přesně reprezentovatelné jako python float. Můžeme generovat i vícebitová čísla, ale vždy to bude číslo! A to je dobře, většina reálných čísel nejde reprezentovat konečnou posloupností symbolů (pozor, $\sqrt{2}$ jde reprezentovat jako kořen polynomu, ale i třeba $1/\pi$ jde reprezentovat například algoritmem, který ho počítá).

$[0,1)$

$\frac{\varphi}{2\pi}$

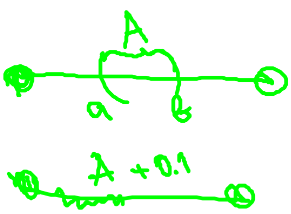
Jak bychom fyzikálně generovali uniformně náhodné číslo z intervalu $[0,1)$? Uniformně náhodné znamená, že každé číslo bude stejně pravděpodobné. Hod šipkou na interval má nevýhodu, že buď budou konce intervalu méně pravděpodobné než prostředek nebo se nám může stát že hodíme šipku mimo interval (nejspíš obojí). Můžeme ale udělat terč s jedním vyznačeným poloměrem, který bude kruh, připevnit jeho střed k vrtačce, roztočit a hodit šipku (tak abychom netrefili střed, ale určitě trefili kruh). Pak náhodné číslo $x \in [0,1)$ bude úhel který svírá vyznačný poloměr a naše šipka dělený $2\pi$.

- Toto je jen myšlenkový experiment, fyzická implementace je nejspíš poměrně nebezpečná. Doma to nezkoušejte!
- Můžete namítnout, že šipka nevybere přesně bod, že terč je tvořen atomy a tedy je také z nějakého pohledu diskrétní. Asi ano, ale já neříkal, že reálná čísla existují. Pro představu atom může mít poloměr okolo $10^{-10}m$, tedy na délce $1m$ jich vedle sebe vyskládáme zhruba $10^{10}$. Srovnejte s přesností $2^{-53} \approx 10^{-16}$, tedy pokud bychom výsledek `random.random()` brali jako pozici v metrech, pak máme přesnost zhruba na dvě miliontiny atomu.

Python random.random()

- *množina elementárních jevů (sample space)*, tedy množina $\Omega = [0,1)$
- *prostor jevů (event space)*, tedy množina $\mathcal{F} \subseteq \mathcal{P}(\Omega)$

$\Pr[A] = b - a$



Napřed se zeptejme, jestli by nemohlo platit, že $\mathcal{F} = \mathcal{P}(\Omega)$. Asi bychom chtěli následující vlastnosti:

  - pokud $A \in \mathcal{F}$ je interval, pak $\Pr[A]$ je rovna délce $A$
  - pokud $A \in \mathcal{F}$ a navíc $x + A = \{x + a \mid a \in A\} \subseteq \Omega$ pro nějaké reálné číslo $x$, pak $\Pr[A] = \Pr[x + A]$.
  - každá podmnožina $\Omega$ má přiřazenou nějakou pravděpodobnost

Banach Tarski 5 kusů

Ale to nejde, protože ne každá množina má míru (tady Pr odpovídá takzvané pravděpodobnostní míře – míra celého prostoru je rovna jedné). Nejznámějším příkladem je https://en.wikipedia.org/wiki/Banach%E2%80%93Tarski_paradox. Hezké video: https://www.youtube.com/watch?v=s86-Z-CbaHA.

Naše řešení: $\mathcal{F}$ je množina Lebesgueovsky měřitelných podmnožin $\Omega$.

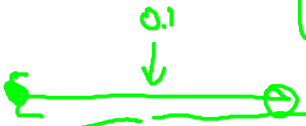
- *pravděpodobnost (probability)*, tedy funkce $\Pr\colon \mathcal{F} \to [0,1]$ je Lebesgueova míra.

Pokud jste neslyšeli o tom, co je to míra, tak se nelekejte. Dobré k zapamatování:

- Pravděpodobnost je číslo mezi nulou a jedničkou (obecná míra celého prostoru nemusí být rovna jedné, ale nás zajímá pravděpodobnostní míra).

$\Omega = \{1, 2 \dots 6, 7\}$
$\mathcal{F} = \mathcal{P}(\Omega)$
$\Pr[A] = 0$ pokud $7 \in A$

- "Nic se nestane" má pravděpodobnost nula – $\Pr[\emptyset] = 0$ (tedy specielně $\emptyset$ je měřitelná).
- "Něco se stane" má pravděpodobnost jedna – $\Pr[\Omega] = 1$ (tedy specielně $\Omega$ je měřitelná).
- "Stane se $A$" má pravděpodobnost 1-"Nestane se $A$" – $\Pr[A] = 1 - \Pr[\Omega \setminus A]$ (tedy pokud je $A$ měřitelná, pak je i její doplněk měřitelný).
- Pr[Na kostce padne jedna tečka nebo dvě tečky] = Pr[padne jedna tečka]+Pr[padnou dvě tečky] – pravděpodobnost sjednocení disjunktních množin je rovna součtu jejich pravděpodobností (platí také pro *spočetně* mnoho disjunktních množin a navíc sjednocení spočetně mnoha disjunktních měřitelných množin je taky měřitelné)

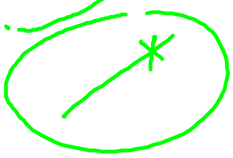



- Lebesgueova míra jednoho bodu je nulová (tedy ze spočetného disjunktního sjednocení máme že pravděpodobnost že uniformně náhodné reálné číslo z intervalu $[0, 1)$ je racionální je nulová).

  kontrast na definici nemožného jevu

  !

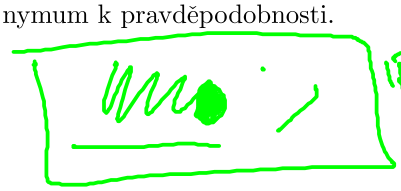
  Pozor na to, že sice platí $\Pr[\{0.1\}] = 0$, ale to neznamená, že jev že padne 0.1 je nemožný (akorát velice velice nepravděpodobný). Naopak pokud je jev nemožný Pr[na šestistěnné kostce padne sedm], pak je jeho pravděpodobnost nulová. $= 0$
- Úsečka v $[0, 1] \times [0, 1] \subseteq \mathbb{R}^2$ má Lebesgueovu míru nula.



Takže ty definice, které vám přijdou divné jsou tam kvůli teorii míry (případně později kvůli teorii integrálu).

Pozor na to, že ne každá pravděpodobnost je Lebesgueova míra, můžeme uvažovat například $\Omega = [0, 1)$, $\mathcal{F}$ jsou Lebesgueovsky měřitelné, Pr která $\Pr[\{0.1\}] = 1/2$ a $\Pr[A] = \lambda(A)/2 + 1/2$ pokud $0.1 \in A$ a $\Pr[A] = \lambda(A)/2$ jinak (kde $\lambda(A)$ značí Lebesgueovu míru množiny $A$).

Někdy také mluvíme o pravděpodobnostní distribuci, zatím to můžete brát jako synonymum k pravděpodobnosti.

$\mathbb{R}^2$ … L. míra je plocha obsah

$\mathbb{R}^3$ L. míra je objem


disj. aditivita



4. **Dokažte, že** $\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B]$.

***Řešení:*** Předpokládejme, že platí $A \cap B, A \cup B, A, B \in \mathcal{F}$ (jinak by výraz nahoře neměl smysl). Z definice pravděpodobnostního prostoru (Definice 2.1) máme spočetnou aditivitu pro disjunktní jevy. Rozdělíme tedy $A \cup B$ na disjunktní množiny:

$$C_1 = A \cap B$$
$$C_2 = A \setminus B$$
$$C_3 = B \setminus A$$

tedy máme $A \cup B = C_1 \cup C_2 \cup C_3$ a zároveň $C_i \cap C_j = \emptyset$ pro každé dvě $1 \leq i < j \leq 3$. Dle spočetné aditivity můžeme psát:

$$\begin{aligned}\Pr[A \cup B] &= \Pr[C_1 \cup C_2 \cup C_3] \\ &= \Pr[C_1] + \Pr[C_2] + \Pr[C_3]\end{aligned}$$

$C_i$ jsou disjunktní!

Máme také:

$$A = C_2 \cup C_1 = (A \setminus B) \cup (A \cap B)$$
$$B = C_3 \cup C_1 = (B \setminus A) \cup (A \cap B)$$

takže můžeme psát:

$$\begin{aligned}\Pr[A \cup B] &= \Pr[C_1] + \Pr[C_2] + \Pr[C_3] \\ &= \Pr[C_1] + \Pr[C_2] + 2\Pr[C_3] - \Pr[C_3] \\ &= (\Pr[C_1] + \Pr[C_3]) + (\Pr[C_2] + \Pr[C_3]) - \Pr[C_3] \\ &= (\Pr[A]) + (\Pr[B]) - \Pr[C_3] \\ &= \Pr[A] + \Pr[B] - \Pr[A \cap B]\end{aligned}$$

což jsme chtěli dokázat.

PIE

$\Pr[A \cup B] \leq \Pr[A] + \Pr[B]$ $(\Pr[A \cap B] \geq 0)$

UNION BOUND

Subaditivita, Booleho nerovnost

5. **Zopakujte si základní kombinatoriku:**

- **Kolik je různých permutací množiny $\{A, B, C\}$?**

  ***Řešení:*** $3! = 3 \cdot 2 \cdot 1$, obecně $n!$ pokud máme $n$ rozlišitelných prvků

```
import itertools

permutations = list(itertools.permutations(['A', 'B', 'C']))
print(f'There are {len(permutations)} permutations: {permutations}')

# There are 6 permutations:
#    [('A', 'B', 'C'),
#     ('A', 'C', 'B'),
#     ('B', 'A', 'C'),
#     ('B', 'C', 'A'),
#     ('C', 'A', 'B'),
#     ('C', 'B', 'A')]
```

- **Kolik různých slov skládajících se z písmen $\{A, B\}$ má délku 3?**

  ***Řešení:*** $2^3 = 8$, obecně $n^r$ kde $n$ je počet písmen, $r$ délka slova

```
import itertools

all_words = list(itertools.product('AB', repeat=3))
print(f'There are {len(all_words)} words: {all_words}')

# There are 8 words:
#    [('A', 'A', 'A'),
#     ('A', 'A', 'B'),
#     ('A', 'B', 'A'),
#     ('A', 'B', 'B'),
#     ('B', 'A', 'A'),
#     ('B', 'A', 'B'),
#     ('B', 'B', 'A'),
#     ('B', 'B', 'B')]
```

- **Kolik různých podmnožin množiny $\{A, B, C, D, E\}$ má velikost 3?**

  ***Řešení:*** $\binom{n}{r} = \frac{n!}{r!(n-r)!} = \binom{5}{3} = 10$

```
import itertools

all_subsets = list(itertools.combinations('ABCDE', 3))
print(f'There are {len(all_subsets)} subsets: {all_subsets}')

# There are 10 subsets:
#     [('A', 'B', 'C'),
#      ('A', 'B', 'D'),
#      ('A', 'B', 'E'),
#      ('A', 'C', 'D'),
#      ('A', 'C', 'E'),
#      ('A', 'D', 'E'),
#      ('B', 'C', 'D'),
#      ('B', 'C', 'E'),
#      ('B', 'D', 'E'),
#      ('C', 'D', 'E')]
```

AAA, AAB, ~~ABA~~, ...



- **Kolik různých kombinací s opakováním z množiny $\{A, B, C, D, E\}$ velikosti 3?**

  ***Řešení:*** $\binom{n+r-1}{r}$ kde $n = 5$, $r = 3$. Protože máme $n-1$ svislítek a $r$ hvězdiček a kódujeme takto:

  $$AAD = \underbrace{**}_{AA}|||\underbrace{*}_{D}|$$

  tedy

  $$\boxed{\text{počet A}|\text{počet B}|\text{počet C}|\text{počet D}|\text{počet E}}$$

  kde každý počet je reprezentován počtem hvězdiček a vybíráme které z $n+r-1$ symbolů budou hvězdičky.

  n hvězdiček

```
import itertools

sorted_sequences = list(itertools.combinations_with_replacement('ABCDE', r=3))
print(f'There are {len(sorted_sequences)} sorted sequences: {sorted_sequences}')

# There are 35 sorted sequences:
#    [('A', 'A', 'A'), ('A', 'A', 'B'), ('A', 'A', 'C'),
#     ('A', 'A', 'D'), ('A', 'A', 'E'), ('A', 'B', 'B'),
#     ('A', 'B', 'C'), ('A', 'B', 'D'), ('A', 'B', 'E'),
#     ('A', 'C', 'C'), ('A', 'C', 'D'), ('A', 'C', 'E'),
#     ('A', 'D', 'D'), ('A', 'D', 'E'), ('A', 'E', 'E'),
#     ('B', 'B', 'B'), ('B', 'B', 'C'), ('B', 'B', 'D'),
#     ('B', 'B', 'E'), ('B', 'C', 'C'), ('B', 'C', 'D'),
#     ('B', 'C', 'E'), ('B', 'D', 'D'), ('B', 'D', 'E'),
#     ('B', 'E', 'E'), ('C', 'C', 'C'), ('C', 'C', 'D'),
#     ('C', 'C', 'E'), ('C', 'D', 'D'), ('C', 'D', 'E'),
#     ('C', 'E', 'E'), ('D', 'D', 'D'), ('D', 'D', 'E'),
#     ('D', 'E', 'E'), ('E', 'E', 'E')]
```

6×  1 6 1111 ✗

666666 ✓



6. **Jaká je pravděpodobnost, že při hodu šesti rozlišitelných spravedlivých šestistěnných kostek padnou aspoň na třech kostkách aspoň tři? Jaký je množina elementárních jevů, prostor jevů a pravděpodobnost?**

***Řešení:***

- Množina elementárních jevů je $\Omega = \{1,2,3,4,5,6\}^6 = \{111111, 111112, \ldots, 666666\}$, tedy $|\Omega| = 6^6 = 46656$.
- Prostor jevů je $\mathcal{F} = \mathcal{P}(\Omega)$, tedy $|\mathcal{F}| = 2^{46656}$
- Pravděpodobnost $\Pr[\{abcdef\}] = 1/6^6$ pro libovolná $a,b,c,d,e,f \in \{1,2,3,4,5,6\}$.

Jak takový příklad řešit? Uvědomit si přesně o čem mluvíme, pak zkusit přemýšlet o jednodušších jevech.

- Na jedné kostce padnou aspoň tři s pravděpodobností $4/6 = 2/3$ (musí padnout $3,4,5,6$, tedy nepadne $1,2$).
- Pokud vybereme $k$ kostek, pak pravděpodobnost, že přesně na těchto kostkách padne aspoň tři je přesně $(2/3)^k(1/3)^{6-k}$ (kostky jsou nezávislé). Kupříkladu pokud vybereme první čtyři kostky, pak nás zajímá:

$k=4$

$$\Pr\left[\{ABCDef \mid A,B,C,D \in \{3,4,5,6\}, e,f \in \{1,2\}\}\right] = \frac{4^4 \cdot 2^2}{6^6} = (2/3)^4(1/3)^{6-4}$$

- Přesně $k$ kostek vybereme $\binom{6}{k}$ způsoby.
- Rozdělíme pravděpodobnostní prostor na jevy, kde přesně na $k$ kostkách padne číslo aspoň tři (tedy máme disjunktní rozklad) a $k \geq 3$, tedy použijeme definici pravděpodobnosti a sečteme předchozí pro $k \in \{3,4,5,6\}$:


$$\begin{aligned}
\Pr[\text{aspoň na třech kostkách aspoň tři}] =& \binom{6}{3}(2/3)^3(1/3)^{6-3} && k=3\\
&+\binom{6}{4}(2/3)^4(1/3)^{6-4} && k=4\\
&+\binom{6}{5}(2/3)^5(1/3)^{6-5} && k=5\\
&+\binom{6}{6}(2/3)^6(1/3)^{6-6} && k=6\\
=&0.8998628257887514
\end{aligned}$$

Tady jsme vlastně použili větu z přednášky, že pokud $B_0, B_1, \ldots, B_{2^6-1}$ jsou rozklad $\Omega$ (tedy $B_i \neq B_j$ pro $i \neq j$ a zároveň $\cup B_i = \Omega$), pak $\boxed{\Pr[A] = \sum_i \Pr[A \mid B_i] \Pr[B_i]}$. Kde jev $A$ je že na aspoň třech kostkách padne aspoň tři. Jev $B_i$ je že na přesně určených kostkách padne aspoň tři (tedy jevy $B_i, B_j$ jsou opravdu disjunktní). Konkrétně 22 zapsané binárně je 010110, pak jev $B_{22}$ je jev že na druhé, čtvrté a páté kostce padlo číslo aspoň tři. Pak $\Pr[A \mid B_x] = 1$ pokud $x$ má v binárním zápisu aspoň tři jedničky a $\Pr[A \mid B_x] = 0$ jinak. Už jsme spočítali, že $\Pr[B_x] = (2/3)^k(1/3)^{6-k}$ pokud $x$ má v binárním zápisu $k$ jedniček.

Pomocí programu:

```
import itertools
import scipy.special
import random
```

```
# Přesný výsledek pomocí kombinatoriky:
def p(k):
    """ Probability that there are exactly k out of 6 dice with at least 3. """
    return scipy.special.comb(6, k, exact=True) * ((2/3)**k) * ((1/3)**(6-k))

exact_computed = sum(p(k) for k in range(3, 7))
print(f'Přesný výsledek: {exact_computed}')
```

$\binom{6}{k} \cdot (2/3)^k \cdot (1/3)^{6-k}$

$p(3) + p(4) + p(5) + p(6)$

```
# Přesný výsledek spočítaný hrubou silou:
def indicator(dice):
    """ Return 1 if at least three dice have at least 3, otherwise return 0. """
    if sum(1 for x in dice if x >= 3) >= 3:
        return 1
    else:
        return 0

all_outcomes = itertools.product(range(1, 7), repeat=6)
exact_bruteforce = sum(indicator(d) for d in all_outcomes) / (6**6)
print(f'Hrubá síla: {exact_bruteforce}')
```

$\Omega = \{111111, \dots, 666666\}$

```
# Simulace:
N = 1000  # Number of tries
simulated = sum(indicator(random.choices(range(1, 7), k=6)) for _ in range(N)) / N
print(f'Simulace: {simulated}')
```

hod 6 kostkama

N hodů

```
# Možný výsledek:
# Přesný výsledek: 0.8998628257887514
# Hrubá síla: 0.8998628257887518
# Simulace: 0.903
```

Porovnejme naše metody:

- Výpočet vzorcem:
  - přesný výsledek
  - potřebovali jsme kombinatoriku a přemýšlet
  - pokud se změní zadání, tak řešení se změní celkem dost (kostky jsou nerozlišitelné)
  - velice rychlý výpočet
- Procházení všech možností:
  - jednodušší vymýšlení
  - potřebujeme programovat
  - přesný výsledek (liší se v posledních místech floatu, dáno nepřesnostmi floatové reprezentace, `exact=True` nevrací nativní float)
  - pokud se změní zadání, tak se řešení skoro nezmění
  - pokud je množina elementárních jevů velká, tak se tento postup nepoužitelný $\Omega$
- Simulace:
  - jednoduché vymýšlení (skoro jako předchozí případ)

- – pokud se změní zadání, tak se řešení skoro nezmění
- – časová složitost není lineární ve velikosti množiny elementárních jevů ($N$ krát vybíráme náhodný prvek $\Omega$, což často zvládáme v $\mathcal{O}(\log|\Omega|)$ krocích)
- – nepřesný výsledek – závisí na náhodných bitech počítače a počtu pokusů
- – budeme potřebovat trochu teorie abychom odhadli jak jistí si jsme výsledkem

Def $A, B$ jsou nezávislé $\iff \Pr[A \cap B] = \Pr[A]\Pr[B]$



7. **Nechť $\Omega$ jsou všechny permutace prvních 100 přirozených čísel, prostor jevů jsou všechny podmnožiny $\Omega$ a každý elementární jev je stejně pravděpodobný. Označme jev $A_j$ že náhodně zvolená permutace $\pi \in \Omega$ splňuje $\pi(j) = j$ (pro $1 \le j \le 100$). Jsou $A_1, A_2$ nezávislé jevy?**

***Řešení:***

- Počet permutací v $A_j$ je přesně 99! (jeden prvek je fixní, zbytek permutujeme), tedy

$$\Pr[A_j] = \frac{99!}{100!} = \frac{1}{100}$$

- Počet permutací v $A_1 \cap A_2$ je přesně 98! (dva prvky jsou fixní, zbytek permutujeme), tedy

$$\Pr[A_1 \cap A_2] = \frac{98!}{100!} = \frac{1}{9900}$$

- Dle definice nezávislých jevů bychom potřebovali $\Pr[A_1]\Pr[A_2] = \Pr[A_1 \cap A_2]$ (Definice 2.4), ale to neplatí:

$$\Pr[A_1 \cap A_2] = \frac{1}{9900} \neq \frac{1}{10000} = \Pr[A_1]\Pr[A_2]$$

$\Rightarrow$ nejsou nezávislé

Zamysleme se nad počítačovým řešením:

```
import random

# indexujeme od nuly
def fixed(my_list, j):
    return my_list[j] == j

my_list = list(range(100))
N = 100000

A1 = 0
for _ in range(N):
    random.shuffle(my_list)
    A1 += 1 if fixed(my_list, 0) else 0
A1 = A1 / N
print(f'Pr[A_1] = Pr[A_2] = {A1} (= {1/100})')

A1A2 = 0
for _ in range(N):
    random.shuffle(my_list)
    A1A2 += 1 if fixed(my_list, 0) and fixed(my_list, 1) else 0
A1A2 = A1A2 / N
print(f'Pr[A_1 and A_2] = {A1A2} (= {1/9900})')

# Možný výsledek:
# Pr[A_1] = Pr[A_2] = 0.00993 (= 0.01)
# Pr[A_1 and A_2] = 0.00012 (= 0.00010101010101010101)
```

\# pokusů

náhodná permutace

$N = 1 \rightsquigarrow A1 = 0.0$

- jednoduchá simulace
- potřebujeme více pokusů, protože potřebujeme odhadnout s větši přesností (menší pravděpodobnost, tak abychom nedostali nulu)

:(

- vůbec nemůžeme použít hrubou sílu, neboť $100! \approx 9.33 \cdot 10^{157}$, pro představu:
  - počítač vykoná zhruba $10^9$ instrukcí za sekundu
  - lineární algoritmus (další permutaci najdeme v jednotkovém čase) by trval zhruba $10^{148}$ sekund
  - stáří vesmíru se odhaduje na $13.787 \cdot 10^9$ let
  - jeden rok trvá zhruba $\pi \cdot 10^7$ sekund
  - stáří vesmíru je tedy zhruba $4.34 \cdot 10^{17}$ sekund
  - tedy lineární algoritmus který by prošel všechny permutace 100 prvkové množiny by běžel zhruba $10^{131}$ stáří vesmíru

8. **Každý obdélník na obrázku je součástka, která se může porouchat s pravděpodobností $p$. Přesněji řečeno porucha znamená, že skrz ní neteče proud. Poruchy součástek jsou na sobě nezávislé. Jaká je pravděpodobnost, že stále poteče proud mezi dvěma puntíky.**

Pr[ všechny OK] $(1-p)^3$

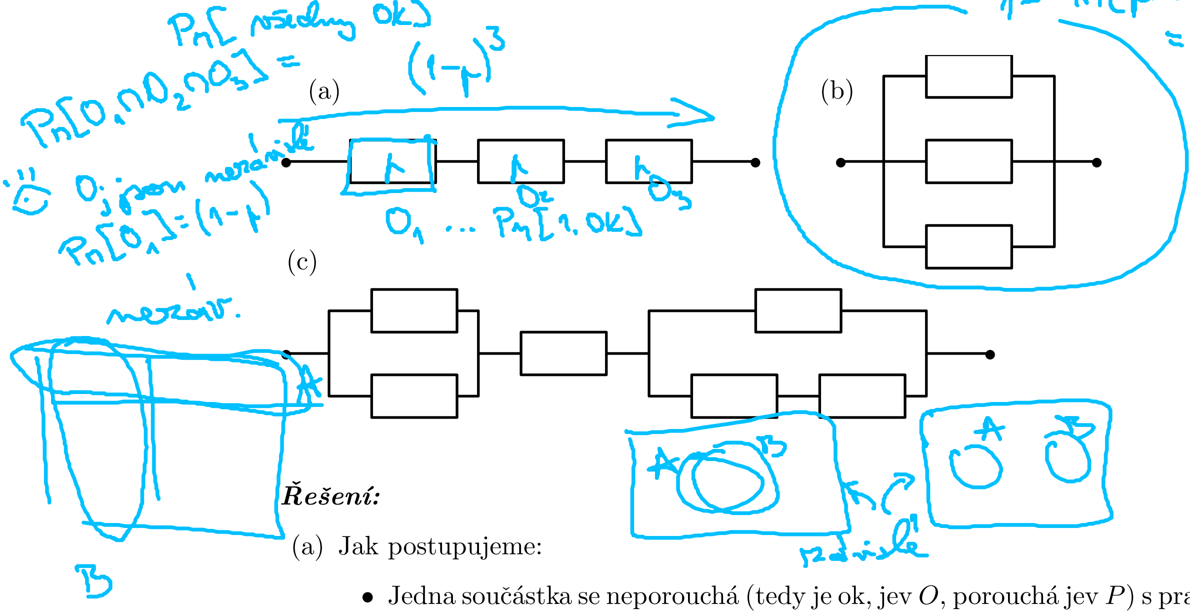
$\Pr[O_1 \cap O_2 \cap O_3] =$

$O_i$ jsou nezávislé

$\Pr[O_1] = (1-p)$

nezáv.

$1 - \Pr[\text{porucha všech}] = 1 - p^3$




***Řešení:***

(a) Jak postupujeme:

- Jedna součástka se neporouchá (tedy je ok, jev $O$, porouchá jev $P$) s pravděpodobností $\Pr[O] = 1 - \Pr[P] = 1 - p$.
- Aby proud tekl, tak všechny součástky musí být ok. Tedy nás zajímá

$$\Pr[O_1 \cap O_2 \cap O_3]$$

- Z nezávislosti jevů $P_1, P_2$ máme i nezávislost jevů $O_1, O_2$ (tedy jejich doplňků):
  - Chceme: $\Pr[A \cap B] = \Pr[A]\Pr[B]$ právě tehdy když $\Pr[\overline{A} \cap \overline{B}] = \Pr[\overline{A}]\Pr[\overline{B}]$ kde $\overline{A} = \Omega \setminus A$ je doplněk
  - Z minulého příkladu přeuspořádáním dostaneme (pro libovolné jevy)

$$\Pr[A \cap B] = \Pr[A] + \Pr[B] - \Pr[A \cup B]$$

  - Tedy:

$$\Pr[\overline{A} \cap \overline{B}] = \Pr[\overline{A}] + \Pr[\overline{B}] - \Pr[\overline{A} \cup \overline{B}]$$

  - Použijeme že $\overline{A} \cup \overline{B} = \overline{A \cap B}$
  - Tedy píšeme:

$$\begin{aligned}
\Pr[\overline{A} \cap \overline{B}] &= \Pr[\overline{A}] + \Pr[\overline{B}] - \Pr[\overline{A} \cup \overline{B}] \\
&= \Pr[\overline{A}] + \Pr[\overline{B}] - \Pr[\overline{A \cap B}] \\
&= (1 - \Pr[A]) + (1 - \Pr[B]) - (1 - \Pr[A \cap B]) \\
&= 1 - \Pr[A] - \Pr[B] + \Pr[A]\Pr[B] \qquad \text{(z nezávislosti } A, B)
\end{aligned}$$

  - Chtěli jsme $\Pr[\overline{A} \cap \overline{B}] = \Pr[\overline{A}]\Pr[\overline{B}] = (1 - \Pr[A])(1 - \Pr[B])$, což je akorát jinak napsaný předchozí řádek.

- Z nezávislosti jevů $O_1, O_2, O_3$ máme rovnou

$$\begin{aligned}\Pr[O_1 \cap O_2 \cap O_3] &= \Pr[O_1]\Pr[O_2]\Pr[O_3] \\ &= (1 - \Pr[P_1])(1 - \Pr[P_2])(1 - \Pr[P_3]) \\ &= (1 - p)^3\end{aligned}$$

(b) Druhý příklad je podobný, ale potřebujeme aby aspoň jedna součástka fungovala, tedy chceme

$$\Pr[O_1 \cup O_2 \cup O_3]$$

Jako nápovědu použijte pozorování že $\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B]$.

Jednodušší postup je, uvědomit si, že se nemůžou porouchat všechny a použít nezávislost, tedy

$$\begin{aligned}\Pr[O_1 \cup O_2 \cup O_3] &= 1 - \Pr[P_1 \cap P_2 \cap P_3] \\ &= 1 - \Pr[P_1]\Pr[P_2]\Pr[P_3] \\ &= 1 - p^3\end{aligned}$$

(c) Kombinace myšlenek předchozích.