

*If the mathematics is not yet useful, it should at least be beautiful.*

*Albrecht Pietsch, Jena*

*Mathematics has advanced a bit since the time of Euclid.*

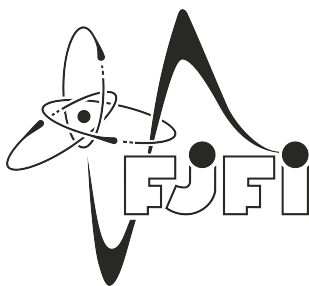
*Terence Tao, UCLA*

# Úvod do pravděpodobnosti 1

## Diskrétní rozdělení

**Jan Vybíral**

Department of Mathematics  
Faculty of Nuclear Sciences and Physical Engineering  
Czech Technical University



# 1 Elementární pravděpodobnost

## 1.1 Náhoda, pravděpodobnost a statistika

Neumíme-li dostatečně dobře vysvětlit či předpovědět výsledek nějakého procesu či experimentu, jeví se nám tyto výsledky *náhodné*. Např. házíme-li opakovaně (na první pohled prakticky totožným způsobem) mincí či kostkou, padají nakonec pokaždé jiná čísla či strany mince. Pokud bychom znali přesně polohu kostky i síly, které na ni během hodu a po dopadu působí, tak bychom sice (alespoň teoreticky) měli být schopni vysvětlit, které číslo padne a proč, ale obvykle nám chybí jak přesná znalost počátečních a okrajových podmínek, tak i schopnost (či ochota) tyto simulace provést.

Náhodný experiment je tedy dán seznamem možných výsledků a pravděpodobnostmi, že k nim dojde. V této přednášce budeme uvažovat pouze experimenty, které mají jen konečný (případně nekonečný, ale spočetný) počet všech možných výsledků. Tak například hod (férovou, šestistěnnou) kostkou je dán množinou všech možných výsledků  $\Omega = \{1, 2, \dots, 6\}$  a vektorem  $p = (1/6, 1/6, \dots, 1/6)$ . Tento vektor musí splňovat dva jednoduché předpoklady:

1.  $p_\omega \geq 0$  pro všechna  $\omega \in \Omega$ ,

2.  $\sum_{\omega \in \Omega} p_\omega = 1$ .

Vektor  $p = (p_\omega)_{\omega \in \Omega}$ , který splňuje tato dvě pravidla, nazveme *rozdělení pravděpodobnosti* na  $\Omega$ .

**Volba modelu:** Pro daný experiment je často možné volit různé modely, které vedou na různé množiny  $\Omega$ , případně vektory  $p = (p_\omega)_{\omega \in \Omega}$ . Tak například pro šestistěnnou kostku je možné volit i  $\Omega = \{1, 2, \dots, 7\}$  s  $p = (1/6, 1/6, \dots, 1/6, 0)$ . Zde  $p_7 = 0$  říká jen prostý fakt, že číslo 7 padne s pravděpodobností 0. Protože se obvykle snažíme volit náš model kompletní, ale co nejjednodušší<sup>1</sup>, tak samozřejmě preferujeme model s šestiprvkovým  $\Omega$ .

Někdy je však situace složitější. Uvažujme hod dvěma (férovými, šestistěnnými) kostkami. Výsledek takového hodu je možné popsat dvojicí padlých čísel seřazených podle velikosti, tedy

$$\Omega = \{(1, 1), (1, 2), \dots, (1, 6), (2, 2), \dots, (2, 6), \dots, (6, 6)\}$$

a

$$p = (1/36, 1/18, \dots, 1/18, 1/36, \dots, 1/18, \dots, 1/36).$$

---

<sup>1</sup>[https://cs.wikipedia.org/wiki/Occamova\\_břítva](https://cs.wikipedia.org/wiki/Occamova_břítva)

Skutečně, pravděpodobnost, že padne jednička a dvojka je dvakrát větší než že padnou dvě jedničky a

$$\sum_{j=1}^6 p_{(j,j)} + \sum_{1 \leq j < k \leq 6} p_{(j,k)} = 6 \cdot \frac{1}{36} + 15 \cdot \frac{1}{18} = 1.$$

Počet prvků  $\Omega$  je  $\#\Omega = 21$ .

Stejně tak si ale můžeme představit, že kostky jsou očíslovány (a víme tedy, která je první a která druhá) a volit

$$\Omega = \{1, \dots, 6\} \times \{1, \dots, 6\} = \{(1, 1), (1, 2), \dots, (1, 6), (2, 1), \dots, (2, 6), \dots, (6, 6)\}$$

a

$$p = (1/36, 1/36, \dots, 1/36, \dots, 1/36).$$

Výsledné  $\Omega$  je sice větší ( $\#\Omega = 36$ ), ale vektor  $p$  je nyní jednodušší, všechny jeho komponenty jsou stejně veliké. Z toho důvodu bývá obvyklé volit pro tento pokus druhý model.

Kromě označení  $\Omega$  a  $p = (p_\omega)_{\omega \in \Omega}$  bývá zvykem používat ještě jiný způsob značení. Označíme-li  $X$  výsledek hodu kostkou, pak

$$p_\omega = \mathbb{P}(X = \omega) = 1/6$$

pro všechna  $\omega \in \{1, \dots, 6\}$  označuje (znovu) onen fakt, že pravděpodobnost hodu každého čísla od 1 do 6 je stejná, a sice  $1/6$ . V tomto slova smyslu je  $X$  náhodná veličina. Je-li  $A \subset \Omega$ , tak definujeme

$$\mathbb{P}(X \in A) = \sum_{\omega \in A} p_\omega. \quad (1.1)$$

Funkce  $A \rightarrow \mathbb{P}(X \in A)$  splňuje mnoho zajímavých vlastností.

- Předpis (1.1) definuje  $\mathbb{P}(X \in A)$  pro *všechny* množiny  $A \subset \Omega$  - řada v (1.1) obsahuje pouze nezáporné členy, nemůže tedy oscilovat nebo nekonvergovat. Všechny množiny  $A \subset \Omega$  jsou tedy *měřitelné*.
- $\mathbb{P}(X \in A) \geq 0$  pro každou množinu  $A \subset \Omega$ .
- Jsou-li  $A_1, A_2, \dots, A_N$  disjunktní podmnožiny  $\Omega$ , pak platí, že

$$\mathbb{P}\left(X \in \bigcup_{j=1}^N A_j\right) = \sum_{j=1}^N \mathbb{P}(X \in A_j).$$

- Pokud je  $\Omega$  nekonečná, ale spočetná, a  $A_1, A_2, \dots$  je nekonečná posloupnost disjunktních podmnožin  $\Omega$ , platí opět

$$\mathbb{P}\left(X \in \bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mathbb{P}(X \in A_j).$$

- $\mathbb{P}(X \in \Omega) = 1$ .

Funkci definované na všech podmnožinách  $\Omega$ , která splňuje výše uvedené požadavky, říkáme *pravděpodobnostní míra*. Při tomto postupu jsme tedy měli na mysli nejprve nějaký náhodný jev (hod kostkou), jeho (proměnlivý) výsledek jsme označili  $X$ , obor hodnot  $X$  (tedy všechny možné výsledky) jsme nazvali  $\Omega$  a teprve nakonec zkonstruovali pravděpodobnostní míru  $A \rightarrow \mathbb{P}(X \in A)$ .

V principu je možné postupovat i opačně. Formálně je to naprosto v pořádku, celý postup je jen mnohem méně intuitivní. Nejprve předpokládáme zadanou (konečnou, nebo nekonečnou spočetnou) množinu  $\Omega$  a poté i funkci  $P : A \rightarrow P(A)$ , která je definována pro všechny  $A \subset \Omega$  - a která splňuje výše zavedené vlastnosti. (Reálná) náhodná veličina  $X$  je pak libovolné zobrazení  $X : \Omega \rightarrow \mathbb{R}$ . Hod kostkou pak modelujeme následujícím způsobem:  $\Omega = \{1, 2, \dots, 6\}$ ,  $P(A) = \#A/6$  a  $X(\omega) = \omega$  pro všechna  $\omega \in \Omega$ .

V dalším textu už nebudeme mezi těmito dvěma přístupy rozlišovat (v obou případech máme nakonec jak množinu  $\Omega$ , tak i pravděpodobnostní míru na ní) a budeme používat  $\mathbb{P}(A)$  a  $\mathbb{P}(X \in A)$  jako záměnné.

**Uniformní rozdělení:** *Uniformní* (nebo *rovnoměrné*) rozdělení  $p = (p_\omega)_\Omega$  je charakterizováno tím, že všechna  $p_\omega$  jsou stejně veliká, tedy  $p_\omega = \frac{1}{\#\Omega}$ , kde  $\#\Omega$  označuje počet prvků množiny  $\Omega$ . V tomto případě požadujeme, aby byl počet prvků množiny  $\Omega$  konečný, jinak totiž  $p_\omega = 0$  a  $\sum_\omega p_\omega = 0$  a  $p$  není rozdělení. V případě uniformního rozdělení lze vyjádřit  $\mathbb{P}(X \in A)$  pomocí

$$\mathbb{P}(X \in A) = \sum_{\omega \in A} p_\omega = \sum_{\omega \in A} \frac{1}{\#\Omega} = \frac{\#A}{\#\Omega}.$$

Pro výpočet pravděpodobnosti, že náhodná veličina  $X$  splňuje nějakou podmínku (vyjádřenou množinou  $A$ ) tedy stačí pouze umět spočítat počet všech možností (tedy  $\#\Omega$ ) a počet všech “pozitivních případů”, tedy  $\#A$ . To bohužel často nebývá tak jednoduché. Nejčastěji se pro tento účel používají elementární pojmy z kombinatoriky, které teď připomeneme.

## 1.2 Opakování kombinatorických pojmů

### Faktoriál:

Pro přirozené číslo  $n \in \mathbb{N}$  je  $n! := 1 \cdot 2 \cdot \dots \cdot n$ .<sup>2</sup> Toto číslo označuje počet permutací  $n$ -prvkové množiny, tedy počet prostých zobrazení  $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ .

Uvažujme například, kolika způsoby může  $n$  žáků vystoupit z autobusu vystupující-li po jednom předními dveřmi. Pokud si žáky očíslováme  $\{1, \dots, n\}$  a první vystoupivší bude  $\sigma(1)$ , druhý  $\sigma(2)$  atd., tak každý způsob vysednutí odpovídá právě jedné permutaci a celkem je tedy  $n!$  možností, jak mohou žáci vystoupit.

<sup>2</sup><https://cs.wikipedia.org/wiki/Faktoriál>

**Variace:**<sup>3</sup> Mějme  $n$  objektů, z nichž chceme postupně  $k$  vybrat. Pokud je vybráním každého objektu tento objekt odstraněn a při dalším výběru již nemůže být znovu vybrán, dostáváme tzv. *variace bez opakování*. Jejich počet je

$$V(k, n) = n(n-1) \dots (n-k+1) = \frac{n!}{(n-k)!}, \quad 1 \leq k \leq n.$$

Pokud je po výběru daný objekt vrácen nazpět do množiny všech objektů a při dalším výběru může být znovu vybrán, jde o *variace s opakováním*, jejíž počet je

$$V'(k, n) = n^k, \quad 1 \leq k.$$

**Kombinace:**<sup>4</sup> Pokud u  $k$  prvků vybraných z  $n$  prvkové množiny bez opakování nezáleží na jejich pořadí, jde o *kombinace* a jejich počet je

$$C(k, n) = \frac{n!}{k!(n-k)!} = \binom{n}{k}, \quad 0 \leq k \leq n.$$

### 1.3 Typické úlohy

1. Kolik různých 4-místných PIN kódů můžeme vytvořit z čísel 2,4,6,8, pokud se žádná číslice nemá v kódu opakovat?
2. Kolik různých 4-místných PIN kódů můžeme vytvořit z čísel 1,3,5,7,9 pokud se žádná číslice nemá v kódu opakovat?
3. Anglická abeceda má 26 písmen. Potřebujeme na její zakódování pomocí Morseovy abecedy s tečkami a čárkami i pětiznakové kódy?
4. Čínská abeceda má 10 000 znaků. Kdybychom chtěli pro tyto znaky vytvořit Morseovu abecedu, kolikaznakové kódy budeme potřebovat?
5. a) Jak pravděpodobné je, že při hodech 7 mincemi padne právě třikrát hlava?  
b) Jak pravděpodobné je, že při 5 hodech kostkou padnou právě dvě šestky?  
c) Jaká je pravděpodobnost, že při dvou hodech kostkou bude číslo v druhém hodu větší než v prvním?
6. Házíme čtyřmi šestistěnnými hracími kostkami. Určete, jaká je pravděpodobnost, že (a) padnou čtyři různá čísla, (b) padnou pouze lichá čísla, (c) součet čísel na všech kostkách dohromady bude roven 6, (d) součet čísel bude větší než 5, (e) padne alespoň jedna šestka.

---

<sup>3</sup>[https://cs.wikipedia.org/wiki/Variace\\_\(kombinatorika\)](https://cs.wikipedia.org/wiki/Variace_(kombinatorika))

<sup>4</sup><https://cs.wikipedia.org/wiki/Kombinace>

7. a) Jaká je pravděpodobnost, že při hodu třemi kostkami hodíme součet 9?  
b) Jaká je pravděpodobnost, že při hodu třemi kostkami hodíme součet 10?  
c) Jaký součet je při hodu třemi kostkami nejpravděpodobnější?
8. Který z následujících jevů má největší pravděpodobnost:  
a) při hodu 6 kostkami padne aspoň jedna šestka,  
b) při hodu 12 kostkami padnou aspoň dvě šestky,  
c) při hodu 18 kostkami padnou aspoň tři šestky?
9. Jaká je pravděpodobnost, že při čtyřech po sobě následujících hodech šestistěnnou kostkou padne aspoň jednou šestka?  
Jaká je pravděpodobnost, že ve 24 hodech dvěma kostkami padnou aspoň jednou dvě šestky?  
Na kterou z obou událostí je výhodnější si vsadit?

## 1.4 Složitější příklady

## 2 Základní pojmy

Cílem této kapitoly je zavést základní pojmy nutné pro studium náhodných událostí.

### 2.1 Nezávislé jevy

Mějme opět (konečnou či spočetnou)  $\Omega$  a na ní definovanou pravděpodobnostní míru  $\mathbb{P}$ . Prvky  $\omega \in \Omega$  se nazývají *elementární jevy*.<sup>5</sup> Podmnožiny  $A \subset \Omega$  se nazývají *jevy*. Na  $\Omega$  předpokládáme zdefinovanou pravděpodobnostní míru  $\mathbb{P}$ , která je tedy definována na všech jevech. Její hodnotu na množině  $A$  označujeme  $\mathbb{P}(A)$  a nazýváme *pravděpodobností*  $A$ .

- Je-li  $A \subset \Omega$  nějaký jev, pak  $A^c = \Omega \setminus A$  nazýváme *jev opačný*.
- Jsou-li  $A, B \subset \Omega$  disjunktní, pak jim říkáme i *neslučitelné*.
- Jsou-li  $A_1, \dots, A_m$  (případně  $A_1, A_2, \dots, A_m, \dots$ ) vzájemně neslučitelné (tedy  $A_i \cap A_j = \emptyset$  pro  $i \neq j$ ), ale vždy musí právě jeden z nich nastat (tedy  $\cup_j A_j = \Omega$ ), pak  $A_1, A_2, \dots$  nazveme *úplný systém jevů* (nebo také *disjunktní rozklad*  $\Omega$ ).

U jednoduchých náhodných pokusů se můžeme přesný výsledek dozvědět velmi jednoduše - podíváme-li se na kostku, hned víme, co padlo. Tuto informaci už pak nelze (v rámci našeho modelu) nijak upřesnit. U složitějších pokusů je třeba provést několik kontrol či měření, abychom se dopátrali přesného výsledku. Po každém měření pak můžeme některé možnosti vyloučit, jiné se stávají naopak pravděpodobnější.

Jako jednoduchý příklad zkusme lékařskou diagnózu. Náhodně vybraný člověk z celé populace má pravděpodobnost  $p_0$  na to, že je zdravý, a pravděpodobnosti  $p_1, \dots, p_N$  na to, že trpí jednou z  $N$  nemocí. Označme  $A$  jev, kdy je člověk zdravý, a  $B_1, \dots, B_N$  jevy, kdy trpí nemocí 1, resp. 2, 3, atd. Pro jednoduchost předpokládejme, že  $B_1, \dots, B_N$  jsou disjunktní a položme  $B = A^c = \cup_{j=1}^N B_j$ . Celkem je tedy

$$\Omega = A \cup \bigcup_{j=1}^N B_j, \quad p_0 + \sum_{j=1}^N p_j = 1.$$

Přišel-li ovšem někdo do nemocnice, můžeme vyloučit jev  $A$  - zdraví lidé se na vstupu do nemocnice (prakticky) neobjevují. To nám umožní zjednodušit celý model. Položme

$$\Omega' = \bigcup_{j=1}^N B_j.$$

---

<sup>5</sup>Množina  $\Omega$  se pak někdy nazývá *množina elementárních jevů*, což se ale jeví jako příliš krkolomné označení.

Na vchodu do nemocnice tedy platí  $\mathbb{P}(\Omega') = 1$  a náš model opravdu obsahuje všechny možné výsledky. Jaká je nyní pravděpodobnost, že náhodně vybraný člověk na vrátnici nemocnice má nemoc 1, 2, 3, atd.? Je to

$$p'_1 = \frac{p_1}{p_1 + \dots + p_N}, \quad p'_2 = \frac{p_2}{p_1 + \dots + p_N}, \quad \dots, \quad p'_N = \frac{p_N}{p_1 + \dots + p_N}.$$

Nyní opět platí  $\sum_{j=1}^N p'_j = 1$ . Pravděpodobnost, že náhodný člověk má nemoc 1 za podmínky, že vyhledal nemocinici, je tedy

$$\mathbb{P}(B_1|B) = p'_1 = \frac{\mathbb{P}(B_1)}{\mathbb{P}(B)}.$$

Získaná dodatečná informace - nebereme náhodné lidi z ulice, ale jen ty, kteří vyhledali lékaře - nám umožňuje zúžit prostor  $\Omega$  a přepočítat/upřesnit pravděpodobnost jednotlivých událostí.

**Definice 2.1.** Nechtě  $A, B \subset \Omega$  jsou dva jevy a nechtě  $\mathbb{P}(B) > 0$ . Pak

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \tag{2.1}$$

nazýváme *podmíněná pravděpodobnost* jevu  $A$  za podmínky  $B$ .

Pokud dodatečná informace (tedy například, že došlo k jevu  $B$ ) nemá vliv na to, jestli došlo k jevu  $A$  či nikoliv, tedy pokud

$$\mathbb{P}(A) = \mathbb{P}(A|B), \quad \text{tj. pokud} \quad \mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B),$$

tak výskyt jevu  $A$  nezávisí na  $B$ . Pokud ale i  $\mathbb{P}(A) > 0$ , pak platí i

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

a výskyt jevu  $B$  nezávisí na  $A$ . Celkově v takovém případě řekneme, že jevy  $A$  a  $B$  jsou *nezávislé*.

**Definice 2.2.** Dva jevy  $A, B$  jsou nezávislé, pokud

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B).$$

Pojem nezávislosti lze snadno rozšířit na větší počet jevů.

**Definice 2.3.** Nechtě  $A_1, A_2, \dots, A_k$  jsou jevy. Řekneme, že

1. Jevy  $A_1, \dots, A_k$  jsou “po dvou nezávislé”,<sup>6</sup> pokud

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i) \cdot \mathbb{P}(A_j)$$

pro všechna  $i \neq j$ .

---

<sup>6</sup>pairwise independent

2. Jevy  $A_1, \dots, A_k$  jsou “nezávislé”,<sup>7</sup> pokud pro každou podmnožinu  $\{i_1, \dots, i_j\} \subset \{1, \dots, k\}$  platí

$$\mathbb{P}(A_{i_1} \cap \dots \cap A_{i_j}) = \mathbb{P}(A_{i_1}) \cdot \dots \cdot \mathbb{P}(A_{i_k}).$$

Mezi těmito dvěma definicemi je zásadní rozdíl. Zatímco první představuje  $\binom{k}{2}$  podmínek, druhá vyžaduje  $2^k$  rovností. Druhá definice zahrnuje jako speciální případ i první (pro  $j = 2$ ) a nezávislé jevy jsou tedy i po dvou nezávislé. Opačná implikace obecně neplatí<sup>8</sup>, jak ukazuje následující jednoduchý (a klasický) příklad.

Házíme dvěma mincemi. Událost, kdy na první minci padne hlava, označme  $A$ . Událost, kdy na druhé minci padne hlava, označme  $B$ . A konečně  $C$  bude událost, kdy na obou mincích padne totéž - tedy buď na obou hlava nebo na obou orel. Pak  $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = 1/2$ . Dále platí (ověřte!)

$$\mathbb{P}(A \cap B) = \mathbb{P}(A \cap C) = \mathbb{P}(B \cap C) = 1/4,$$

ale

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

Události  $A, B, C$  jsou tedy po dvou nezávislé, ale nejsou nezávislé.

Počítání s podmíněnými pravděpodobnostmi nebývá složité a typicky je založeno přímo na definici (2.1). Uveďme dva příklady takového postupu.

**Věta 2.4.** *Nechť  $A$  je náhodný jev a nechť  $B_1, \dots, B_n$  je úplný systém jevů (a.k.a. disjunkttní rozklad  $\Omega$ ) a nechť  $\mathbb{P}(B_1), \dots, \mathbb{P}(B_n) > 0$ .*

1. *Pak platí*

$$\mathbb{P}(A) = \sum_{j=1}^n \mathbb{P}(A|B_j) \cdot \mathbb{P}(B_j). \quad (2.2)$$

2. *Pokud je navíc i  $\mathbb{P}(A) > 0$ , pak platí*

$$\mathbb{P}(B_k|A) = \frac{\mathbb{P}(A|B_k) \cdot \mathbb{P}(B_k)}{\sum_{j=1}^n \mathbb{P}(A|B_j) \cdot \mathbb{P}(B_j)}.$$

*Důkaz.* Množiny  $A \cap B_j$  tvoří disjunkttní rozklad  $A$  a proto platí

$$\mathbb{P}(A) = \sum_{j=1}^n \mathbb{P}(A \cap B_j).$$

---

<sup>7</sup>mutually independent

<sup>8</sup>[https://en.wikipedia.org/wiki/Independence\\_\(probability\\_theory\)#Pairwise\\_and\\_mutual\\_independence](https://en.wikipedia.org/wiki/Independence_(probability_theory)#Pairwise_and_mutual_independence)

Protože navíc  $\mathbb{P}(A \cap B_j) = \mathbb{P}(A|B_j) \cdot \mathbb{P}(B_j)$ , je první bod dokázán.

I druhý bod plyne rychle

$$\mathbb{P}(B_k|A) = \frac{\mathbb{P}(A \cap B_k)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A \cap B_k)}{\mathbb{P}(B_k)} \cdot \frac{\mathbb{P}(B_k)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B_k) \cdot \mathbb{P}(B_k)}{\mathbb{P}(A)}$$

a následně dosadíme z předchozího vzorce (2.2). □

## 2.2 Náhodné veličiny

Připomeňme, že předpokládáme, že  $\Omega$  je konečná, nebo nekonečná spočetná a že pro každou její podmnožinu  $A \subset \Omega$  je definována její míra  $\mathbb{P}(A)$ . Pak náhodná veličina  $X$  je libovolné zobrazení z  $\Omega$  do  $\mathbb{R}$ , tedy  $X : \Omega \rightarrow \mathbb{R}$ .

Jako příklad náhodné veličiny uvažujme výšku dětí ve třídě. Nechť  $\Omega = \{d_1, \dots, d_N\}$ , kde  $d_1, \dots, d_N$  jsou (například podle abecedy seřazené) jednotlivé děti ve třídě. Nechť  $X(d_1)$  je výška prvního dítěte (v centimetrech) a stejně definujeme i  $X(d_2), \dots, X(d_N)$ . Děti budeme vybírat náhodně, tedy  $\mathbb{P}(\{d_j\}) = 1/N$  pro všechna  $1 \leq j \leq N$ . Pravděpodobnost, že náhodně vybrané dítě má výšku mezi 130 a 140 centimetry, pak můžeme zapsat jako

$$\mathbb{P}(\{\omega \in \Omega : 130 \leq X(\omega) \leq 140\}).$$

Skutečně,  $\{\omega \in \Omega : 130 \leq X(\omega) \leq 140\}$  označuje množinu všech dětí, které mají požadovanou výšku, a pomocí  $\mathbb{P}$  pak změříme pravděpodobnost, že náhodně vybrané dítě padne do této množiny. Celkově je ale tento zápis příliš dlouhý, obvykle jej zkracujeme na

$$\mathbb{P}(130 \leq X \leq 140).$$

Pro náhodnou veličinu  $X$  lze definovat její *distribuční funkci*  $F : \mathbb{R} \rightarrow \mathbb{R}$  jako  $F(x) := \mathbb{P}(X \leq x)$ . Pro spojitě rozdělené náhodné veličiny je to velice užitečný pojem, pro diskrétní náhodné veličiny se lze bez něj docela dobře obejít a nahradit jej případně pomocí

$$F(x) = \mathbb{P}(X \leq x) = \sum_{x_j \leq x} \mathbb{P}(X = x_j).$$

## 2.3 Střední hodnota

**Definice 2.5.** Střední hodnota náhodné veličiny  $X$  definované na diskrétním pravděpodobnostním prostoru  $\Omega$  a nabývajících hodnot  $x_1, x_2, \dots$  je

$$\mathbb{E} X = \sum_j x_j \mathbb{P}(X = x_j). \quad (2.3)$$

**Poznámka 2.6.** 1. Pokud je  $\Omega$  konečná, tak je  $\mathbb{E} X$  dobře definováno pro každou náhodnou veličinu  $X$  a je to vlastně vážený průměr všech možných výsledků  $x_1, \dots, x_n$ , kde váhy jsou pravděpodobnosti toho, že při pozorování  $X$  opravdu uvidíme  $x_1, x_2, \dots$ .

2. Je-li  $\Omega$  nekonečná, může dojít k problémům s konvergencí v (2.3). Vzhledem k tomu, že pořadí prvků  $\Omega$  by nemělo mít vliv na zavedené pojmy, tak vyžadujeme, aby (2.3) konvergovala ke stejné hodnotě při libovolném přeuspořádání, tedy (podle Riemannovy věty) aby konvergovala absolutně. Pokud tomu tak není, řekneme, že střední hodnota  $X$  není definována.

3. Pokud například  $\Omega = \{1, 2, 3, \dots\}$ ,  $\mathbb{P}(\{j\}) = \frac{1}{2^j}$  a  $X(j) = 2^j$ , tak

$$\sum_{j=1}^{\infty} x_j \cdot \mathbb{P}(X = x_j) = \sum_{j=1}^{\infty} 2^j \cdot \frac{1}{2^j} = \sum_{j=1}^{\infty} 1.$$

Náhodná veličina  $X$  nabývá tedy jen konečných hodnot, ale i přesto je její střední hodnota nekonečná.

Máme-li dány dvě náhodné veličiny  $X$  a  $Y$  a znalost hodnoty jedné z nich nám nedává žádnou výhodu ve znalosti hodnoty druhé z nich, pak řekneme, že  $X$  a  $Y$  jsou nezávislé.

**Definice 2.7.** Řekneme, že náhodné veličiny  $X$  a  $Y$ , definované na témže prostoru  $\Omega$ , jsou nezávislé, pokud

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y)$$

pro všechny možné  $x$  z oboru hodnot  $X$  a všechny  $y$  z oboru hodnot  $Y$ .<sup>9</sup>

V našem případě s třídou, kde  $X$  označovalo výšku dítěte, by například  $Y$  mohlo označovat váhu dítěte. Veličiny  $X$  a  $Y$  ale nejspíše nebudou nezávislé - víme-li, že náhodně vybrané dítě má velkou výšku, bude mít spíše i vyšší váhu.

Příkladem nezávislých náhodných veličin je například hod dvěma kostkami - předpokládáme, že výsledný počet ok na jedné kostce nijak neovlivňuje počet ok na druhé kostce. Proto pro jejich modelování s výhodou použijeme techniku (tenzorového či kartezského) součinu. Jsou-li  $(\Omega, \mathbb{P})$  a  $(\Omega', \mathbb{P}')$  dva (diskrétní) pravděpodobnostní prostory, pak  $\Omega \times \Omega'$  vybavený mírou splňující  $\mathbb{P}(A \times B) = \mathbb{P}(A) \cdot \mathbb{P}'(B)$  je opět pravděpodobnostní prostor.

---

<sup>9</sup>V řeči distribučních funkcí lze toto ekvivalentně přeformulovat jako

$$F_{(X,Y)}(x, y) := \mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(X \leq x) \cdot \mathbb{P}(Y \leq y) = F_X(x) \cdot F_Y(y).$$

Předpoklad nezávislosti je často velmi praktický, protože umožňuje rozdělit složitější problémy na nezávislé komponenty. Často je ale používán i přesto, že víme, že platí jen přibližně. Například předpoklad, že zpoždění jednotlivých autobusů na stejné lince MHD, jsou na sobě nezávislá, je více či méně zavádějící.

*Varování:*<sup>10</sup> Stejně jako pro jevy bylo třeba rozlišovat mezi “po dvou nezávislými” a “nezávislými” jevy, je toto třeba rozlišovat i pro náhodné veličiny. Náhodné veličiny  $X_1, \dots, X_n$  jsou po dvou nezávislé, je-li každá dvojice  $X_i, X_j$  nezávislá pro všechna  $i \neq j$ . Veličiny  $X_1, \dots, X_n$  jsou (vzájemně) nezávislé, pokud

$$\mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \mathbb{P}(X_1 = x_1) \cdot \dots \cdot \mathbb{P}(X_n = x_n)$$

pro všechny možné hodnoty  $x_1, \dots, x_n$ .<sup>11</sup> Na rozdíl od náhodných jevů zde není nutné uvažovat i všechny podmnožiny  $\{1, \dots, n\}$ . Je-li např.  $n = 3$ , tak

$$\begin{aligned} \mathbb{P}(X_1 = x_1, X_3 = x_3) &= \sum_{x_2} \mathbb{P}(X_1 = x_1, X_2 = x_2, X_3 = x_3) \\ &= \sum_{x_2} \mathbb{P}(X_1 = x_1) \cdot \mathbb{P}(X_2 = x_2) \cdot \mathbb{P}(X_3 = x_3) \\ &= \mathbb{P}(X_1 = x_1) \cdot \mathbb{P}(X_3 = x_3) \cdot \sum_{x_2} \mathbb{P}(X_2 = x_2) \\ &= \mathbb{P}(X_1 = x_1) \cdot \mathbb{P}(X_3 = x_3). \end{aligned}$$

Příklad tří náhodných veličin, které jsou po dvou nezávislé, ale nejsou nezávislé, je podobný příkladu, který jsme uvažovali pro analogický efekt u jevů. A sice předpokládejme, že házíme dvěma mincemi. Veličina  $X$  nabývá hodnoty 1, pokud na první minci padla hlava, a hodnoty 0, pokud padl orel. Stejně tak  $Y$  je 1, pokud na druhé minci padne hlava, jinak je  $Y = 0$ . A konečně  $Z$  je 1, pokud padla právě jedna hlava, jinak je  $Z = 0$ . Snadno se ověří, že  $X, Y, Z$  jsou po dvou nezávislé, ale nejsou nezávislé, protože

$$0 = \mathbb{P}(X = 1, Y = 1, Z = 1) \neq \mathbb{P}(X = 1) \cdot \mathbb{P}(Y = 1) \cdot \mathbb{P}(Z = 1) = 1/8.$$

Všimněte si, že v tomto případě je vždy  $XYZ = 0$  (pokud  $X$  a  $Y$  nejsou rovny nule, tak  $Z = 0$ ) a tedy i  $0 = \mathbb{E}[XYZ] \neq \mathbb{E}[X] \cdot \mathbb{E}[Y] \cdot \mathbb{E}[Z] = 1/8$ .

**Věta 2.8** (Vlastnosti střední hodnoty). 1. *Nechť  $X$  je náhodná veličina s konečnou střední hodnotou. Nechť  $a, b \in \mathbb{R}$ . Pak  $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$ .*

2. *Nechť  $X_1, \dots, X_n$  jsou náhodné veličiny s konečnými středními hodnotami. Pak  $\mathbb{E}[X_1 + \dots + X_n] = \mathbb{E}[X_1] + \dots + \mathbb{E}[X_n]$ .*<sup>12</sup>

<sup>10</sup>[https://en.wikipedia.org/wiki/Pairwise\\_independence](https://en.wikipedia.org/wiki/Pairwise_independence)

<sup>11</sup>Ekvivalentně lze opět vše přeformulovat pomocí distribučních funkcí.

<sup>12</sup>Varianta pro nekonečně mnoho náhodných veličin obecně neplatí!

3. Jsou-li  $X, Y$  **nezávislé** náhodné veličiny s konečnými středními hodnotami. Pak  $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$ .

4. Jsou-li  $X_1, \dots, X_n$  **nezávislé** náhodné veličiny s konečnými středními hodnotami, pak

$$\mathbb{E} \left[ \prod_{j=1}^n X_j \right] = \prod_{j=1}^n \mathbb{E}[X_j].$$

*Důkaz.* 1. Jsou-li  $x_1, x_2, \dots$  všechny možné hodnoty náhodné veličiny  $X$ , pak  $ax_1 + b, ax_2 + b, \dots$  jsou všechny možné hodnoty  $aX + b$  a platí

$$\begin{aligned} \mathbb{E}[aX + b] &= \sum_j [ax_j + b] \mathbb{P}(aX + b = ax_j + b) = \sum_j [ax_j + b] \mathbb{P}(X = x_j) \\ &= a \sum_j x_j \mathbb{P}(X = x_j) + b \sum_j \mathbb{P}(X = x_j) = a \mathbb{E}[X] + b. \end{aligned}$$

2. Tvzení dokážeme pro  $n = 2$  (tedy dvě náhodné veličiny  $X$  a  $Y$ ), dále pak plyne indukcí. Jsou-li  $x_1, x_2, \dots$  možné hodnoty  $X$  a  $y_1, y_2, \dots$  hodnoty  $Y$ , pak  $X + Y$  nabývá hodnot  $x_i + y_j$  s pravděpodobnostmi  $p_{i,j} = \mathbb{P}(X = x_i, Y = y_j)$ . Poznamenejme, že může nastat i  $p_{i,j} = 0$ . Pak platí

$$\begin{aligned} \mathbb{E}[X + Y] &= \sum_{i,j} (x_i + y_j) p_{i,j} = \sum_{i,j} x_i p_{i,j} + \sum_{i,j} y_j p_{i,j} \\ &= \sum_i x_i \sum_j p_{i,j} + \sum_j y_j \sum_i p_{i,j} \\ &= \sum_i x_i \sum_j \mathbb{P}(X = x_i, Y = y_j) + \sum_j y_j \sum_i \mathbb{P}(X = x_i, Y = y_j) \\ &= \sum_i x_i \mathbb{P}(X = x_i) + \sum_j y_j \mathbb{P}(Y = y_j) = \mathbb{E}[X] + \mathbb{E}[Y]. \end{aligned}$$

3. Opět označme  $x_1, x_2, \dots$  možné hodnoty  $X$  a  $p_i = \mathbb{P}(X = x_i)$  a  $y_1, y_2, \dots$  hodnoty  $Y$  s  $q_j = \mathbb{P}(Y = y_j)$ . Podobně jako dříve obdržíme (s využitím nezávislosti  $X$  a  $Y$ )

$$\begin{aligned} \mathbb{E}[X \cdot Y] &= \sum_{i,j} (x_i \cdot y_j) \mathbb{P}(X = x_i, Y = y_j) = \sum_{i,j} (x_i \cdot y_j) \cdot p_i \cdot q_j \\ &= \left( \sum_i p_i \cdot x_i \right) \cdot \left( \sum_j q_j y_j \right) = \mathbb{E}[X] \cdot \mathbb{E}[Y]. \end{aligned}$$

4. Konečně obecný případ plyne opět indukcí, stačí si rozmyslet, že  $X_1 \cdot \dots \cdot X_{n-1}$  a  $X_n$  jsou nezávislé, jsou-li  $X_1, \dots, X_n$  nezávislé.

□

## 2.4 Rozptyl

Rozptyl náhodné veličiny vyjadřuje, jak moc se daná veličina obvykle liší od svého průměru.

**Definice 2.9.** Nechť  $X$  je náhodná veličina a necht' má dobře definovanou střední hodnotu  $\mu := \mathbb{E} X$ . Pak rozptyl  $X$  definujeme jako

$$\text{Var}(X) = \mathbb{E}(X - \mu)^2,$$

pokud je i tato střední hodnota dobře definována.

Rozptyl  $X$  lze vyjádřit i jako

$$\begin{aligned}\text{Var}(X) &= \mathbb{E}(X - \mu)^2 = \mathbb{E}(X^2 - 2\mu X + \mu^2) = \mathbb{E}[X^2] - 2\mu\mathbb{E}[X] + \mu^2 \\ &= \mathbb{E}[X^2] - [\mathbb{E} X]^2.\end{aligned}$$

**Věta 2.10** (Rozptyl součtu nezávislých náhodných veličin). *Nechť  $X_1, \dots, X_n$  jsou nezávislé náhodné veličiny s dobře definovaným rozptylem. Pak*

$$\text{Var}(X_1 + \dots + X_n) = \text{Var}(X_1) + \dots + \text{Var}(X_n).$$

*Důkaz.* Označme  $X = X_1 + \dots + X_n$ ,  $\mu_j = \mathbb{E} X_j$  a  $\mu = \mu_1 + \dots + \mu_n = \mathbb{E}[X]$ . Pak

$$\begin{aligned}\text{Var}(X) &= \mathbb{E}(X - \mu)^2 = \mathbb{E} \left[ \sum_{j=1}^n (X_j - \mu_j) \right]^2 = \mathbb{E} \left[ \sum_{j,k=1}^n (X_j - \mu_j)(X_k - \mu_k) \right] \\ &= \sum_{j,k=1}^n \mathbb{E}(X_j - \mu_j)(X_k - \mu_k) \\ &= \sum_{j=1}^n \mathbb{E}(X_j - \mu_j)^2 + \sum_{j \neq k} \mathbb{E}(X_j - \mu_j) \cdot \mathbb{E}(X_k - \mu_k) \\ &= \sum_{j=1}^n \text{Var}(X_j).\end{aligned}$$

□

## Reference

- [1] J. Anděl, Matematika náhody, Matfyzpress, 4. vydání, 2020
- [2] J. Anděl, Statistické úlohy, historiky a paradoxy, 2018