Pravděpodobnost, proč je to tak těžké?

Tomáš Svoboda and Petr Pošík

Vision for Robots and Autonomous Systems, Center for Machine Perception
Department of Cybernetics
Faculty of Electrical Engineering, Czech Technical University in Prague

15. března 2023

Outline

- Mathematics of uncertainty
- Random Experiment, Outcomes, Sample Space, Events, ...
- Probability, Conditional Probability, Independence
- Random Variable, Expectation

Pravděpodobnost (nejistota) je všude

- Pravděpodobnost srážek zítra je 70%.
- Jakou mám šanci vyhrát v loterii?
- Mám pozitivní test na nemoc X, jsem opravdu nemocný?
- Svědectví jsou X, Y, a Z, je obviněný vinen?
- Nezaměstnanost se změnila o X, jaká bude inflace?
- Jak se bude vyvíjet cena akcií?
- Vybraná akce je X, o kolik se robot pohne?
- Jaká je pravděpodobnost, že na fotce je osoba XY?
- Jak dlouho mi bude trvat cesta, když pojedu tramvají?
- ...

Potřebujeme matematický popis ...

(Náhodný) pokus

Experiment :

- Vaguely: the act of observing certain feature of the world
- A procedure that
  - can be repeated many times under the same conditions and
  - has a well-defined set of possible outcomes.
- Deterministic experiment has only a single possible outcome.
- Random experiment has more than one possible outcomes.
- Before executing random experiment, we do not know the actual outcome. After execution this uncertainty vanishes.

Příklad 1: Házíme třikrát mincí (Head/Tails)

Jaká je pravděpodobnost, že padne třikrát panna (Head)?

Prostor (množina!) S všech elementárních jevů . Jak je velká?

A 3^2

B 2^3

C $2 \cdot 3$

D ∞

Jevy :

- A - $3\times$ panna (hlava), $P(A) = ?$
- B - $3\times$ stejný symbol, $P(B) = ?$
- C - alespoň jeden orel (tail), $P(C) = ?$ ...

Příklad 1: Házíme třikrát mincí (Head/Tails)

Jaká je pravděpodobnost, že padne třikrát panna (Head)?

Prostor (množina!) $\mathcal{S}$ všech elementárních jevů . Jak je velká?

A 3^2

B 2^3

C $2 \cdot 3$

D ∞

Jevy :

- A - $3\times$ panna (hlava), $P(A) = ?$
- B - $3\times$ stejný symbol, $P(B) = ?$
- C - alespoň jeden orel (tail), $P(C) = ?$ ...

Příklad 1: Házíme třikrát mincí (Head/Tails)

Jaká je pravděpodobnost, že padne třikrát panna (Head)?

Prostor (množina!) $\mathcal{S}$ všech elementárních jevů . Jak je velká?

A 3^2

B 2^3

C $2 \cdot 3$

D ∞

Jevy :

- A - $3\times$ panna (hlava), $P(A) = ?$
- B - $3\times$ stejný symbol, $P(B) = ?$
- C - alespoň jeden orel (tail), $P(C) = ? \ldots$

Příklad 1: Házíme třikrát mincí (Head/Tails)

Jaká je pravděpodobnost, že padne třikrát panna (Head)?

Prostor (množina!) $\mathcal{S}$ všech elementárních jevů . Jak je velká?

A 3^2

B 2^3

C $2 \cdot 3$

D ∞

Jevy :

- A - $3\times$ panna (hlava), $P(A) = ?$
- B - $3\times$ stejný symbol, $P(B) = ?$
- C - alespoň jeden orel (tail), $P(C) = ?$ ...

Příklad 1: Házíme třikrát mincí (Head/Tails)

Jaká je pravděpodobnost, že padne třikrát panna (Head)?

Prostor (množina!) $\mathcal{S}$ všech elementárních jevů . Jak je velká?

A 3^2

B 2^3

C $2 \cdot 3$

D ∞

Jevy :

- A - 3$\times$ panna (hlava), $P(A) = ?$
- B - 3$\times$ stejný symbol, $P(B) = ?$
- C - alespoň jeden orel (tail), $P(C) = ? \ldots$

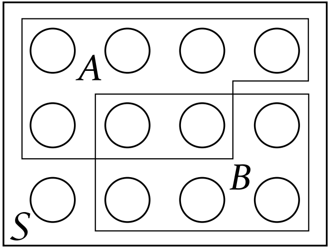
(Random) Events /(Náhodné) jevy/

Elementary events /elementární jevy/ are *all possible, mutually exclusive outcomes* of certain experiment.

The set of elementary events is called a sample space /množina elementárních jevů/ , denoted as $\mathcal{S}$.

An event /jev/ is any subset of the sample space, $A \subseteq \mathcal{S}$.

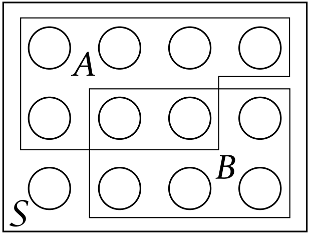
- Event A occured if the experiment outcome belongs to A.
- An event is any statement about the experiment outcome for which we can decide if it occured or not.



Naive probability (Bernoulli/Laplace)

$$P(A) = \frac{|A|}{|\mathcal{S}|} = \frac{\text{number of outcomes favorable to } A}{\text{total number of outcomes in } \mathcal{S}}$$

- Limited to *equally likely* outcomes/elementary events. (*Equally likely?*)
- It does not allow for infinite sample spaces, geometric probability, ...
- *Combinatorics!* Counting (variations, permutations, combinations, ...)



Events and their combinations

Important events:

- Certain event : $\mathcal{S}$, 1
- Impossible event : $\emptyset$, 0

Event combinations:

- Conjunction $(A$ and $B)$: $A \cap B$
- Disjuction $(A$ or $B)$: $A \cup B$
- Complementary event /jev opačný/ to A: $A^c = \mathcal{S} \setminus A$
- $A \Rightarrow B$: $A \subseteq B$
- Disjoint events /jevy neslučitelné/ : $A_1, \ldots, A_n : \bigcap_{i \leq n} A_i = \emptyset$
- Mutually exclusive events /Jevy po dvou neslučitelné = vzájemně se vylučující/ : $A_1, \ldots, A_n : \forall i,j \in \{1, \ldots, n\},\ i \neq j : A_i \cap A_j = \emptyset$

Events and their combinations

Important events:

- Certain event : $\mathcal{S}$, 1
- Impossible event : $\emptyset$, 0

Event combinations:

- Conjunction (A and B): $A \cap B$
- Disjuction (A or B): $A \cup B$
- Complementary event /jev opačný/ to A: $A^c = \mathcal{S} \setminus A$
- $A \Rightarrow B$: $A \subseteq B$
- Disjoint events /jevy neslučitelné/ : $A_1, \ldots, A_n : \bigcap_{i \leq n} A_i = \emptyset$
- Mutually exclusive events /Jevy po dvou neslučitelné = vzájemně se vylučující/ : $A_1, \ldots, A_n : \forall i, j \in \{1, \ldots, n\},\ i \neq j : A_i \cap A_j = \emptyset$

Partition of sample space /Úplný systém jevů/

Partition of sample space $\mathcal{S}$ /Úplný systém jevů/ is composed of events $B_1, \ldots, B_n$ if they are *mutually exclusive* and $\bigcup\limits_{i=1}^{n} B_i = \mathcal{S}$.

- The sample space $\mathcal{S}$ is its own partition by definition.
- Events $\{C, C^c\}$ form a partition: $C \cap C^c = \emptyset$ and $C \cup C^c = \mathcal{S}$.

Proč je úplný systém jevů důležitý koncept?

Partition of sample space /Úplný systém jevů/

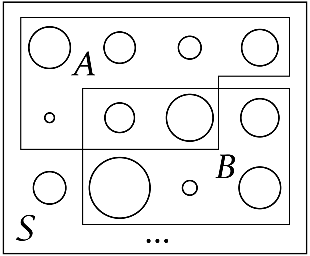
Partition of sample space $\mathcal{S}$ /Úplný systém jevů/ is composed of events $B_1, \ldots, B_n$ if they are *mutually exclusive* and $\bigcup_{i=1}^{n} B_i = \mathcal{S}$.

- The sample space $\mathcal{S}$ is its own partition by definition.
- Events $\{C, C^c\}$ form a partition: $C \cap C^c = \emptyset$ and $C \cup C^c = \mathcal{S}$.

Proč je úplný systém jevů důležitý koncept?

Axiomatic probability (Kolmogorov)

- Sample space $\mathcal{S}$ may be infinite.
- Elementary events do not have to be equally likely.
- Axiomatic:
  1. state a set of constraints the probability function must obey
  2. find a function that fulfills them (next slides)



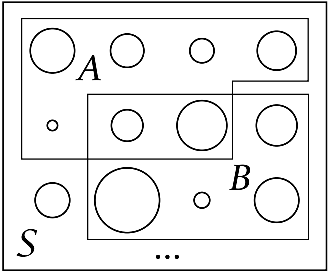
Definition of probability

- Probability function /pravděpodobnostní funkce/ P is a function that assigns a real number between 0 and 1 to each event $A \subseteq \mathcal{S}$.
- P must satisfy the following axioms:

  1. $P(\emptyset) = 0$, $P(\mathcal{S}) = 1$
  2. For any mutually exclusive events $A_1, A_2, \ldots, A_n$:

  $$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

  (n may be infinite)



Interpretations of probability

Frequentist :

- Relative frequency of an event after many repetitions of random experiment.

Bayesian :

- Degree of belief that an event occurs.
- This allows us to assign probabilities to statements like "candidate A wins elections" or "suspect X is guilty", although we cannot repeat the same elections or the same crime over and over.

Příklad 2: Vlastnosti P, hod kostkou

$$\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$$

Uvažujme jevy:

- A - padne 6
- B - padne sudé číslo

Množinově: $A \subset B$

Pravděpodobnost: $P(A) < P(B)$

Další jev:

- C - padne 2 nebo 4

Množinově: $C = B \setminus A$

Pravděpodobnost: $P(C) = P(B) - P(A)$

Příklad 2: Vlastnosti P, hod kostkou

$$\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$$

Uvažujme jevy:

- A - padne 6
- B - padne sudé číslo

Množinově: $A \subset B$

Pravděpodobnost: $P(A) < P(B)$

Další jev:

- C - padne 2 nebo 4

Množinově: $C = B \setminus A$

Pravděpodobnost: $P(C) = P(B) - P(A)$

Příklad 2: Vlastnosti P, hod kostkou

$$\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$$

Uvažujme jevy:

- A - padne 6
- B - padne sudé číslo

Množinově: $A \subset B$

Pravděpodobnost: $P(A) < P(B)$

Další jev:

- C - padne 2 nebo 4

Množinově: $C = B \setminus A$

Pravděpodobnost: $P(C) = P(B) - P(A)$

Příklad 2: Vlastnosti P, hod kostkou

$$\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$$

Uvažujme jevy:

- A - padne 6
- B - padne sudé číslo

Množinově: $A \subset B$

Pravděpodobnost: $P(A) < P(B)$

Další jev:

- C - padne 2 nebo 4

Množinově: $C = B \setminus A$

Pravděpodobnost: $P(C) = P(B) - P(A)$

Příklad 2: Vlastnosti P, hod kostkou

$$\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$$

Uvažujme jevy:

- A - padne 6
- B - padne sudé číslo

Množinově: $A \subset B$

Pravděpodobnost: $P(A) < P(B)$

Další jev:

- C - padne 2 nebo 4

Množinově: $C = B \setminus A$

Pravděpodobnost: $P(C) = P(B) - P(A)$

Příklad 2: Vlastnosti P, hod kostkou

$$\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$$

Uvažujme jevy:

- A - padne 6
- B - padne sudé číslo

Množinově: $A \subset B$

Pravděpodobnost: $P(A) < P(B)$

Další jev:

- C - padne 2 nebo 4

Množinově: $C = B \setminus A$

Pravděpodobnost: $P(C) = P(B) - P(A)$

Příklad 2: Vlastnosti P, hod kostkou

$$\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$$

Uvažujme jevy:

- A - padne 6
- B - padne sudé číslo

Množinově: $A \subset B$

Pravděpodobnost: $P(A) < P(B)$

Další jev:

- C - padne 2 nebo 4

Množinově: $C = B \setminus A$

Pravděpodobnost: $P(C) = P(B) - P(A)$

Příklad 2: Vlastnosti P, hod kostkou

$$\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$$

Uvažujme jevy:

- A - padne 6
- B - padne sudé číslo

Množinově: $A \subset B$

Pravděpodobnost: $P(A) < P(B)$

Další jev:

- C - padne 2 nebo 4

Množinově: $C = B \setminus A$

Pravděpodobnost: $P(C) = P(B) - P(A)$

Příklad 2: Vlastnosti P, hod kostkou

$$\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$$

Uvažujme jevy:

- A - padne 6
- B - padne sudé číslo

Množinově: $A \subset B$

Pravděpodobnost: $P(A) < P(B)$

Další jev:

- C - padne 2 nebo 4

Množinově: $C = B \setminus A$

Pravděpodobnost: $P(C) = P(B) - P(A)$

Příklad 2: Vlastnosti P (pokr.)

Hod kostkou, $\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$, A - padne 6, B - padne liché číslo.

Zjevně $A \cap B = \emptyset$,

$$P(A \cup B) = P(A) + P(B)$$

Čerpadlo v elektrárně je zálohováno pomocí druhého identického čerpadla. Jev A_i znamená, že čerpadlo i funguje. Jaká je pravděpodonost, že funguje alespoň jedno?

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

Obě čerpadla fungují:

$$P(A_1 \cap A_2) = ?$$

Příklad 2: Vlastnosti P (pokr.)

Hod kostkou, $\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$, A - padne 6, B - padne liché číslo.
Zjevně $A \cap B = \emptyset$,

$$P(A \cup B) = P(A) + P(B)$$

Čerpadlo v elektrárně je zálohováno pomocí druhého identického čerpadla. Jev A_i znamená, že čerpadlo i funguje. Jaká je pravděpodonost, že funguje alespoň jedno?

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

Obě čerpadla fungují:

$$P(A_1 \cap A_2) = ?$$

Příklad 2: Vlastnosti P (pokr.)

Hod kostkou, $\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$, A - padne 6, B - padne liché číslo.
Zjevně $A \cap B = \emptyset$,

$$P(A \cup B) = P(A) + P(B)$$

Čerpadlo v elektrárně je zálohováno pomocí druhého identického čerpadla. Jev A_i znamená, že čerpadlo i funguje. Jaká je pravděpodonost, že funguje alespoň jedno?

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

Obě čerpadla fungují:

$$P(A_1 \cap A_2) = ?$$

Příklad 2: Vlastnosti P (pokr.)

Hod kostkou, $\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$, A - padne 6, B - padne liché číslo.
Zjevně $A \cap B = \emptyset$,

$$P(A \cup B) = P(A) + P(B)$$

Čerpadlo v elektrárně je zálohováno pomocí druhého identického čerpadla. Jev A_i znamená, že čerpadlo i funguje. Jaká je pravděpodonost, že funguje alespoň jedno?

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

Obě čerpadla fungují:

$$P(A_1 \cap A_2) = ?$$

Příklad 2: Vlastnosti P (pokr.)

Hod kostkou, $\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$, A - padne 6, B - padne liché číslo.
Zjevně $A \cap B = \emptyset$,

$$P(A \cup B) = P(A) + P(B)$$

Čerpadlo v elektrárně je zálohováno pomocí druhého identického čerpadla. Jev A_i znamená, že čerpadlo i funguje. Jaká je pravděpodonost, že funguje alespoň jedno?

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

Obě čerpadla fungují:

$$P(A_1 \cap A_2) = ?$$

Příklad 2: Vlastnosti P (pokr.)

Hod kostkou, $\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$, A - padne 6, B - padne liché číslo.
Zjevně $A \cap B = \emptyset$,

$$P(A \cup B) = P(A) + P(B)$$

Čerpadlo v elektrárně je zálohováno pomocí druhého identického čerpadla. Jev A_i znamená, že čerpadlo i funguje. Jaká je pravděpodonost, že funguje alespoň jedno?

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

Obě čerpadla fungují:

$$P(A_1 \cap A_2) = ?$$

Příklad 2: Vlastnosti P (pokr.)

Hod kostkou, $\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$, A - padne 6, B - padne liché číslo.
Zjevně $A \cap B = \emptyset$,

$$P(A \cup B) = P(A) + P(B)$$

Čerpadlo v elektrárně je zálohováno pomocí druhého identického čerpadla. Jev A_i znamená, že čerpadlo i funguje. Jaká je pravděpodonost, že funguje alespoň jedno?

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

Obě čerpadla fungují:

$$P(A_1 \cap A_2) = ?$$

Properties of probability

For any valid probability function:

- $P(A) \in \langle 0, 1 \rangle$ (definition)
- $P(\emptyset) = 0, \qquad P(\mathcal{S}) = 1$ (axioms)
- $P(A^c) = 1 - P(A)$
- If $A \subseteq B$, then $P(A) \leq P(B)$
- If $A \subseteq B$, then $P(B \setminus A) = P(B) - P(A)$
- If $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$ (*aditivity*)
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Příklad 3: Pravděpodobnost částí

Vybereme-li z populace náhodně jednoho člověka, pak pro něj platí:

- Trpí nemocí X a je mladší 18 let s pravděpodobností 0.01.
- Trpí nemocí X a je mu/jí mezi 18 a 65 lety s pravděpodobností 0.05.
- Trpí nemocí X a je starší 65 let s pravděpodobností 0.09.

Jaká je pravděpodobnost, že náhodně vybraný člověk má nemoc X?

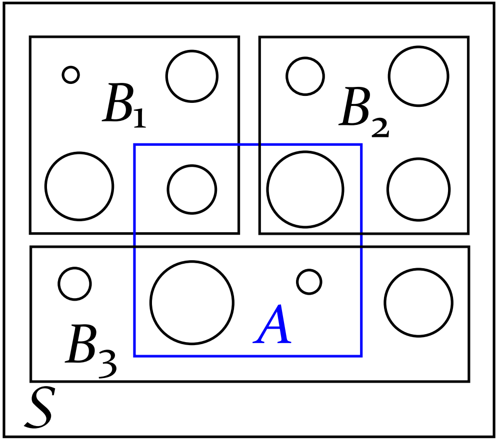
Properties of probability (cont.)

If $\{B_1, \ldots, B_n\}$ is a partition of sample space then for any event $A \subseteq \mathcal{S}$

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) \ .$$

In particular, for partition $\{C, C^c\}$

$$P(A) = P(A \cap C) + P(A \cap C^c) \, .$$



Independent events /Nezávislé jevy/

Čerpadla u elektrárny, chtěli bychom, aby byla nezávislá Events A and B are independent if and only iff

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

If A, B are independent, then

- $P(A \cup B) = P(A) + P(B) - P(A) \cdot P(B)$,
- and pairs A, B^c and A^c, B and A^c, B^c are independent too.

Independent events /Nezávislé jevy/

Čerpadla u elektrárny, chtěli bychom, aby byla nezávislá Events A and B are independent if and only iff

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

If A, B are independent, then

- $P(A \cup B) = P(A) + P(B) - P(A) \cdot P(B)$,
- and pairs A, B^c and A^c, B and A^c, B^c are independent too.

Nezávislost jevů: hod dvěma mincemi

- A - na první minci padla hlava
- B - na druhé minci padla hlava
- C - na mincích padly různé symboly

Které skupiny jevů jsou nezávislé?

A žádná skupina jevů

B dvojice (A, B), (B, C), (A, C)

C dvojice (A, B), (B, C), (A, C) i trojice (A, B, C)

D pouze trojice (A, B, C)

Nezávislost jevů: hod dvěma mincemi

- A - na první minci padla hlava
- B - na druhé minci padla hlava
- C - na mincích padly různé symboly

Které skupiny jevů jsou nezávislé?

A žádná skupina jevů

B dvojice (A, B), (B, C), (A, C)

C dvojice (A, B), (B, C), (A, C) i trojice (A, B, C)

D pouze trojice (A, B, C)

Příklad: Voják nebo technik?

Tomáš je pořádkumilovný, rozhodný a má smysl pro spravedlnost. Jako dítě rád hrál na počítači strategické hry a střílečky. Vždy se zajímal o zbraně a vojenskou techniku.

Čím je podle vás Tomáš v současnosti?

A Voják

B Technik

Příklad: Voják nebo technik?

Tomáš je pořádkumilovný, rozhodný a má smysl pro spravedlnost. Jako dítě rád hrál na počítači strategické hry a střílečky. Vždy se zajímal o zbraně a vojenskou techniku.

Čím je podle vás Tomáš v současnosti?

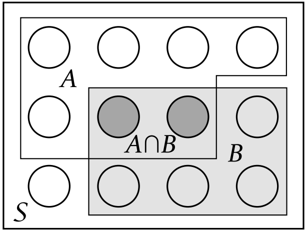
A Voják

B Technik

Conditional probability

Conditional probability of event A given event B is defined as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \qquad P(B) \neq 0.$$



- All probabilities are conditional: $P(A) = P(A|\mathcal{S})$.
- Interpretation:
  1. $P(A)$ is our current belief that event A occurs.
  2. We get a new information that a different event B occured.
  3. $P(A|B)$ is now our updated belief about A.
- Conditional probability is still a probability: it maps any event $A \subseteq \mathcal{S}$ to $\langle 0, 1 \rangle$.

Properties of Conditional Probability

- $P(\mathcal{S}|B) = 1$, $P(\emptyset|B) = 0$.
- $P(A|A) = 1$, $P(A^c|A) = 0$.
- If $B \subseteq A$, then $P(A|B) = 1$.
- If $P(A \cap B) = 0$, then $P(A|B) = 0$.
- If $A_1, \ldots, A_n$ are mutually exclusive events, then $P\left(\bigcup_{i=1}^{n} A_i \middle| B\right) = \sum_{i=1}^{n} P(A_i|B)$.
- Events A, B are independent iff $P(A|B) = P(A)$ (if $P(A|B)$ is defined).

Úprava/aktualizace pravděpodobnosti

1. Pravděpodobnost $P(A)$ je naše úvodní (apriorní) důvěra, že nastane jev A.
2. Dozvíme se, že nastal jev B.
3. Pravděpodobnost $P(A|B)$ je naše aktualizovaná (aposteriorní) důvěra, že nastane jev A.

O jevech A a B nevíme nic dalšího. Která z následujících možností je správná?

A $P(A|B) < P(A)$

B $P(A|B) = P(A)$

C $P(A|B) > P(A)$

D Může nastat kterákoli z uvedených možností.

Úprava/aktualizace pravděpodobnosti

1. Pravděpodobnost $P(A)$ je naše úvodní (apriorní) důvěra, že nastane jev A.
2. Dozvíme se, že nastal jev B.
3. Pravděpodobnost $P(A|B)$ je naše aktualizovaná (aposteriorní) důvěra, že nastane jev A.

O jevech A a B nevíme nic dalšího. Která z následujících možností je správná?

A $P(A|B) < P(A)$

B $P(A|B) = P(A)$

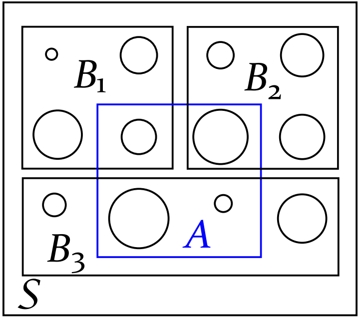
C $P(A|B) > P(A)$

D Může nastat kterákoli z uvedených možností.

The Law of Total Probability

Let $B_1, \ldots, B_n$ be a partition of the sample space $\mathcal{S}$ (i.e., the B_i are disjoint events and their union is $\mathcal{S}$), with $P(B_i) > 0$ for all i.
Then

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)$$



Bayes rule

Probability of the intersection of two events A and B, $P(A \cap B)$, can be expressed in 2 ways:

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

From that it follows Bayes rule

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

Applying the law of total probability from previous slide:

$$P(B_i|A) = \frac{P(A|B_i) \cdot P(B_i)}{P(A)} = \frac{P(A|B_i) \cdot P(B_i)}{\sum_{j \in I} P(A|B_j) \cdot P(B_j)}.$$

Bayes rule

Probability of the intersection of two events A and B, $P(A \cap B)$, can be expressed in 2 ways:

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

From that it follows **Bayes rule**

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

Applying the law of total probability from previous slide:

$$P(B_i|A) = \frac{P(A|B_i) \cdot P(B_i)}{P(A)} = \frac{P(A|B_i) \cdot P(B_i)}{\sum\limits_{j \in I} P(A|B_j) \cdot P(B_j)}.$$

Bayes rule

Probability of the intersection of two events A and B, $P(A \cap B)$, can be expressed in 2 ways:

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

From that it follows **Bayes rule**

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

Applying the law of total probability from previous slide:

$$P(B_i|A) = \frac{P(A|B_i) \cdot P(B_i)}{P(A)} = \frac{P(A|B_i) \cdot P(B_i)}{\sum\limits_{j \in I} P(A|B_j) \cdot P(B_j)}.$$

Pracovat s náhodnými jevy je těžkopádné . . .

Pokus: 3 hody mincí. Elementární jevy $s \in \mathcal{S}$. Jevy $A_i \subseteq \mathcal{S}$:

- tři hlavy – $X(s) = 3$
- alespoň jedna hlava – $X(s) \geq 1$
- tři stejné symboly – $X(s) \in \{3, 0\}$
- . . .

Můžeme definovat každý jev jako množinu (leckdy dost obsáhlou) elementárních jevů s.
Nebo si zavedeme náhodnou proměnnou/veličinu X:

$$X(s) = \text{počet hlav v } s$$

Náhodná *proměnná* je ve skutečnosti *funkce*!

Pracovat s náhodnými jevy je těžkopádné ...

Pokus: 3 hody mincí. Elementární jevy $s \in \mathcal{S}$. Jevy $A_i \subseteq \mathcal{S}$:

- tři hlavy – $X(s) = 3$
- alespoň jedna hlava – $X(s) \geq 1$
- tři stejné symboly – $X(s) \in \{3, 0\}$
- ...

Můžeme definovat každý jev jako množinu (leckdy dost obsáhlou) elementárních jevů s.

Nebo si zavedeme náhodnou proměnnou/veličinu X:

$$X(s) = \text{počet hlav v } s$$

Náhodná *proměnná* je ve skutečnosti *funkce*!

Pracovat s náhodnými jevy je těžkopádné . . .

Pokus: 3 hody mincí. Elementární jevy $s \in \mathcal{S}$. Jevy $A_i \subseteq \mathcal{S}$:

- tři hlavy – $X(s) = 3$
- alespoň jedna hlava – $X(s) \geq 1$
- tři stejné symboly – $X(s) \in \{3, 0\}$
- . . .

Můžeme definovat každý jev jako množinu (leckdy dost obsáhlou) elementárních jevů s.

Nebo si zavedeme náhodnou proměnnou/veličinu X:

$$X(s) = \text{počet hlav v } s$$

Náhodná *proměnná* je ve skutečnosti *funkce*!

Pracovat s náhodnými jevy je těžkopádné . . .

Pokus: 3 hody mincí. Elementární jevy $s \in \mathcal{S}$. Jevy $A_i \subseteq \mathcal{S}$:

- tři hlavy – $X(s) = 3$
- alespoň jedna hlava – $X(s) \geq 1$
- tři stejné symboly – $X(s) \in \{3, 0\}$
- . . .

Můžeme definovat každý jev jako množinu (leckdy dost obsáhlou) elementárních jevů s.

Nebo si zavedeme náhodnou proměnnou/veličinu X:

$$X(s) = \text{počet hlav v } s$$

Náhodná *proměnná* je ve skutečnosti *funkce*!

Pracovat s náhodnými jevy je těžkopádné …

Pokus: 3 hody mincí. Elementární jevy $s \in \mathcal{S}$. Jevy $A_i \subseteq \mathcal{S}$:

- tři hlavy – $X(s) = 3$
- alespoň jedna hlava – $X(s) \geq 1$
- tři stejné symboly – $X(s) \in \{3, 0\}$
- …

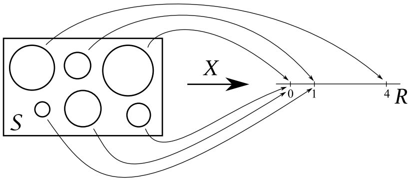
Můžeme definovat každý jev jako množinu (leckdy dost obsáhlou) elementárních jevů s.
Nebo si zavedeme náhodnou proměnnou/veličinu X:

$$X(s) = \text{počet hlav v } s$$

Náhodná *proměnná* je ve skutečnosti *funkce*!

Random Variable

Random variable (náhodná proměnná/veličina) on a probability space $(\mathcal{S}, P)$ is a function X mapping elementary events $s \in \mathcal{S}$ to real numbers $\mathbb{R}$, i.e., $X : \mathcal{S} \to \mathbb{R}$.

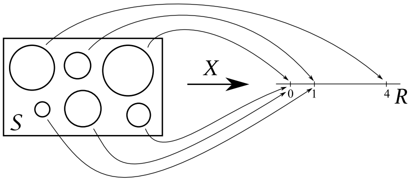


"Random variable is a numerical 'summary' of an aspect of the experiment."

- R.v. X assigns a numerical value $X(s)$ to each possible outcome $s \in \mathcal{S}$.
- The mapping is *deterministic*; the randomness comes from outcomes of random experiment (with outcome probabilities described by probability function P).
- Before the experiment, we know neither the value of s, nor the value of $X(s)$. But we can compute the probability that X will take on a given value, or a range of values.
- After the experiment, s was realized, and the r.v. crystalizes into value $X(s)$.

Random Variable

Random variable (náhodná proměnná/veličina) on a probability space $(\mathcal{S}, P)$ is a function X mapping elementary events $s \in \mathcal{S}$ to real numbers $\mathbb{R}$, i.e., $X : \mathcal{S} \to \mathbb{R}$.



"Random variable is a numerical 'summary' of an aspect of the experiment."

- R.v. X assigns a numerical value $X(s)$ to each possible outcome $s \in \mathcal{S}$.
- The mapping is *deterministic*; the randomness comes from outcomes of random experiment (with outcome probabilities described by probability function P).
- Before the experiment, we know neither the value of s, nor the value of $X(s)$. But we can compute the probability that X will take on a given value, or a range of values.
- After the experiment, s was realized, and the r.v. crystalizes into value $X(s)$.

Náhodné jevy vs hodnoty náhodné proměnné

Let X be a random variable, i.e., $X : \mathcal{S} \to \mathbb{R}$.

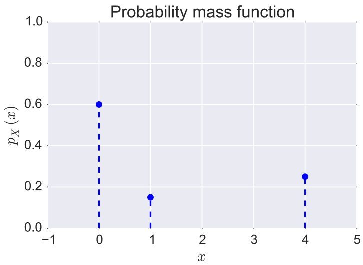
- $X = x$ denotes the event $\{s \in \mathcal{S} : X(s) = x\}$, i.e., the event consisting of all outcomes s such that $X(s) = x$.
- $X \in \langle a, b)$ denotes the event $\{s \in \mathcal{S} : a \leq X(s) < b\}$, i.e., the event consisting of all outcomes s such that $a \leq X(s) < b$.

Discrete Random Variable

Random variable X is called discrete if the values of $X(s)$ for all $s \in \mathcal{S}$ form either

- a finite set of values $a_1, a_2, \ldots, a_n$, or
- an infinite set of countably many values $a_1, a_2, \ldots$

Support of X:
$\mathcal{S}_X = \{x \in \mathbb{R} : P(X = x) > 0\} = \{a_1, a_2, \ldots\}$



Probability Mass Function (PMF) /pstní fce/ of a discrete r.v. X is the function p_X given by

$$p_X(x) = P(X = x) = P(\{s \in \mathcal{S} : X(s) = x\}).$$

Cumulative Distribution Function (CDF) /distribuční fce/ of a discrete r.v. X is the function F_X defined as

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

Discrete Random Variable

Random variable X is called discrete if the values of $X(s)$ for all $s \in \mathcal{S}$ form either

- a finite set of values $a_1, a_2, \ldots, a_n$, or
- an infinite set of countably many values $a_1, a_2, \ldots$

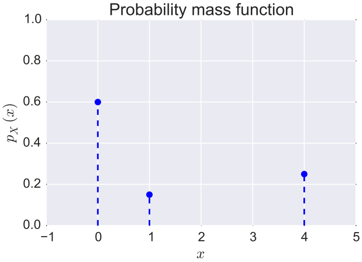
Support of X:
$\mathcal{S}_X = \{x \in \mathbb{R} : P(X = x) > 0\} = \{a_1, a_2, \ldots\}$

Probability Mass Function (PMF) /pstní fce/ of a discrete r.v. X is the function p_X given by

$$p_X(x) = P(X = x) = P(\{s \in \mathcal{S} : X(s) = x\}).$$

Cumulative Distribution Function (CDF) /distribuční fce/ of a discrete r.v. X is the function F_X defined as

$$F_X(x) = P(X \leq x) = \sum_{t \leq x} p_X(t).$$



Discrete Random Variable

Random variable X is called discrete if the values of $X(s)$ for all $s \in \mathcal{S}$ form either

- a finite set of values $a_1, a_2, \dots, a_n$, or
- an infinite set of countably many values $a_1, a_2, \dots$

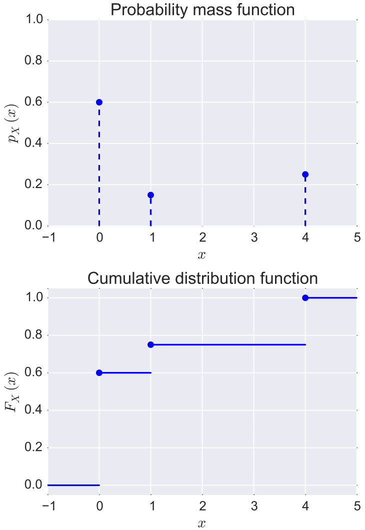
Support of X:
$\mathcal{S}_X = \{x \in \mathbb{R} : P(X = x) > 0\} = \{a_1, a_2, \dots\}$

Probability Mass Function (PMF) /pstní fce/ of a discrete r.v. X is the function p_X given by

$$p_X(x) = P(X = x) = P(\{s \in \mathcal{S} : X(s) = x\}).$$

Cumulative Distribution Function (CDF) /distribuční fce/ of a discrete r.v. X is the function F_X defined as

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



Expected value

Před začátkem experimentu, kolik **očekávám**, že padne hlav?

Expected value (střední hodnota) of a discrete r.v. X is denoted as $\mathrm{E}\,X$ and is defined as

$$\mathrm{E}\,X = \sum_{t \in \mathbb{R}} t \cdot p_X(t) = \sum_{t \in S_X} t \cdot p_X(t).$$

For equally probable outcomes $s \in S$ also $\mathrm{E}\,X = \frac{1}{|S|} \sum_{s \in S} X(s)$.

Characteristics of $\mathrm{E}\,X$:

- $\mathrm{E}\,r = r$, $\mathrm{E}(\mathrm{E}\,X) = \mathrm{E}\,X$
- $\mathrm{E}(X + Y) = \mathrm{E}\,X + \mathrm{E}\,Y$, $\mathrm{E}(X + r) = \mathrm{E}\,X + r$, $\mathrm{E}(X - Y) = \mathrm{E}\,X - \mathrm{E}\,Y$
- $\mathrm{E}(rX + sY) = r\,\mathrm{E}\,X + s\,\mathrm{E}\,Y$
- For *independent* r.v.s: $\mathrm{E}(X \cdot Y) = \mathrm{E}\,X \cdot \mathrm{E}\,Y$.

Expected value

Před začátkem experimentu, kolik očekávám , že padne hlav?

Expected value (střední hodnota) of a discrete r.v. X is denoted as $\mathrm{E}\,X$ and is defined as

$$\mathrm{E}\,X = \sum_{t \in \mathbb{R}} t \cdot p_X(t) = \sum_{t \in \mathcal{S}_X} t \cdot p_X(t).$$

For equally probable outcomes $s \in S$ also $\mathrm{E}\,X = \frac{1}{|S|} \sum_{s \in S} X(s)$.

Characteristics of $\mathrm{E}\,X$:

- $\mathrm{E}\,r = r$, $\mathrm{E}(\mathrm{E}\,X) = \mathrm{E}\,X$
- $\mathrm{E}(X+Y) = \mathrm{E}\,X + \mathrm{E}\,Y$, $\mathrm{E}(X+r) = \mathrm{E}\,X + r$, $\mathrm{E}(X-Y) = \mathrm{E}\,X - \mathrm{E}\,Y$
- $\mathrm{E}(rX+sY) = r\,\mathrm{E}\,X + s\,\mathrm{E}\,Y$
- For *independent* r.v.s: $\mathrm{E}(X \cdot Y) = \mathrm{E}\,X \cdot \mathrm{E}\,Y$.

Expected value

Před začátkem experimentu, kolik očekávám, že padne hlav?

Expected value (střední hodnota) of a discrete r.v. X is denoted as $\mathrm{E}\,X$ and is defined as

$$\mathrm{E}\,X = \sum_{t \in \mathbb{R}} t \cdot p_X(t) = \sum_{t \in \mathcal{S}_X} t \cdot p_X(t).$$

For equally probable outcomes $s \in \mathcal{S}$ also $\mathrm{E}\,X = \dfrac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} X(s)$.

Characteristics of $\mathrm{E}\,X$:

- $\mathrm{E}\,r = r$, $\mathrm{E}(\mathrm{E}\,X) = \mathrm{E}\,X$
- $\mathrm{E}(X + Y) = \mathrm{E}\,X + \mathrm{E}\,Y$, $\mathrm{E}(X + r) = \mathrm{E}\,X + r$, $\mathrm{E}(X - Y) = \mathrm{E}\,X - \mathrm{E}\,Y$
- $\mathrm{E}(rX + sY) = r\,\mathrm{E}\,X + s\,\mathrm{E}\,Y$
- For *independent* r.v.s: $\mathrm{E}(X \cdot Y) = \mathrm{E}\,X \cdot \mathrm{E}\,Y$.

Expected value

Před začátkem experimentu, kolik očekávám, že padne hlav?

Expected value (střední hodnota) of a discrete r.v. X is denoted as $\mathrm{E}\,X$ and is defined as

$$\mathrm{E}\,X = \sum_{t \in \mathbb{R}} t \cdot p_X(t) = \sum_{t \in \mathcal{S}_X} t \cdot p_X(t).$$

For equally probable outcomes $s \in \mathcal{S}$ also $\mathrm{E}\,X = \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} X(s)$.

Characteristics of $\mathrm{E}\,X$:

- $\mathrm{E}\,r = r$, $\mathrm{E}(\mathrm{E}\,X) = \mathrm{E}\,X$
- $\mathrm{E}(X+Y) = \mathrm{E}\,X + \mathrm{E}\,Y$, $\mathrm{E}(X+r) = \mathrm{E}\,X + r$, $\mathrm{E}(X-Y) = \mathrm{E}\,X - \mathrm{E}\,Y$
- $\mathrm{E}(rX + sY) = r\,\mathrm{E}\,X + s\,\mathrm{E}\,Y$
- For *independent* r.v.s: $\mathrm{E}(X \cdot Y) = \mathrm{E}\,X \cdot \mathrm{E}\,Y$.

References, further reading

Klasické české čtení [2], na domácí stránce https://cmp.felk.cvut.cz/~navara/stat/ mnoho dalších zajímavých studijních materiálů.

[1] Joseph K. Blitzstein and Jessica Hwang.
Introduction to Probability.
CRC Press, 2nd edition, 2019.
http://probabilitybook.net/.

[2] Mirko Navara.
Pravděpodobnost a matematická statistika.
ČVUT v Praze, Praha, 2007.
https://cmp.felk.cvut.cz/~navara/stat/.

[3] N. Silver.
Signál a šum.
Paseka, 2014.