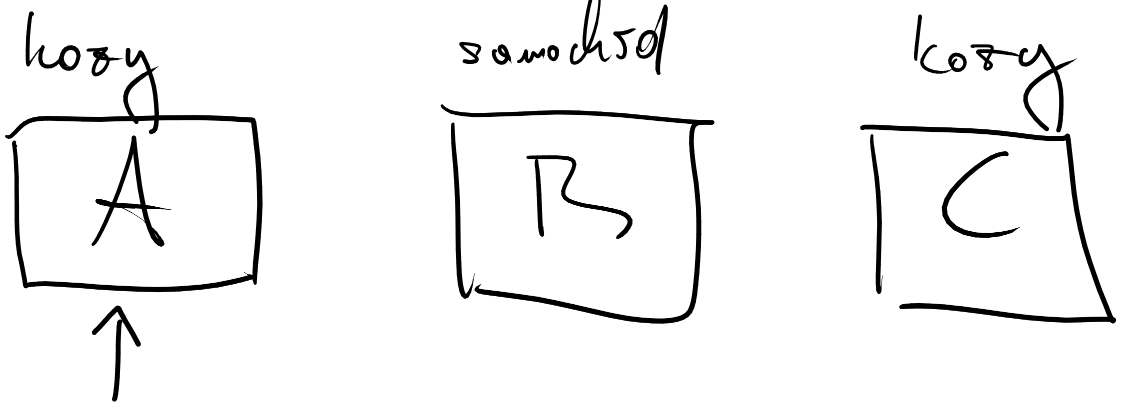
Metody Statystyczne 1

kacper.topolnicki@uj.edu.pl

- problem Montiego Halla
- aksjomaty prawdopodobieństwa (Kołmogorow 1933)
- własności prawdopodobieństwa
- twierdzenie Bayesa
- prawdopodobieństwo Bayesowskie
- dyskretne zmienne losowe
- ciągłe zmienne losowe
- funkcje gęstości prawdopodobieństwa
- własności rozkładów

MH



○ MH

○ uczestnik

- czy warto zmienić wybór

monty-hall.nb

Kolmogorow, 1933

Ω – zbiór zdarzeń elementarnych

$\mathcal{F}$ – przestrzeń zdarzeń losowych

P – funkcja przypisująca prawdopodobieństwo

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

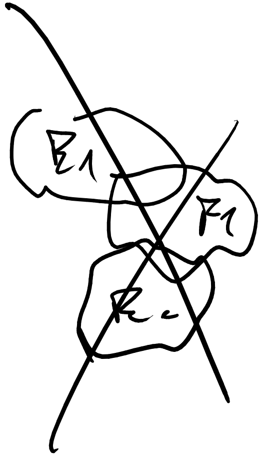
(Ω, F, P)

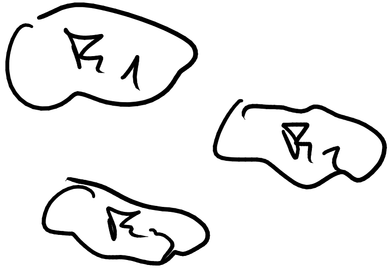
1) $\forall_{E \in F} \; P(E) \in \mathbb{R} \;,\; P(E) \geq 0$

$\Omega \in F$

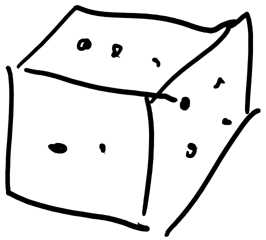
podzbiory Ω

2) $P(\Omega) = 1$





3) $\underbrace{E_1, E_2, \ldots}_{\text{nie mają części wspólnych}}$ $P\left(\bigcup_{i=1}^{N} E_i\right) = \sum_{i=1}^{N} P(E_i)$



$$\Omega = \{\text{⚀}, \text{⚁}, \text{⚂}, \text{⚃}, \text{⚄}, \text{⚅}\}$$

$$F = \{\emptyset, \underbrace{\{\text{⚀}, \text{⚂}, \text{⚄}\}}, \ldots, \Omega\}$$

$$\left.\begin{aligned}
P(\{\text{⚀}\}) &= \tfrac{1}{6}\\
P(\{\text{⚁}\}) &= \tfrac{1}{6}\\
P(\{\text{⚂}\}) &= \tfrac{1}{6}\\
P(\{\text{⚃}\}) &= \tfrac{1}{6}\\
P(\{\text{⚄}\}) &= \tfrac{1}{6}\\
P(\{\text{⚅}\}) &= \tfrac{1}{6}
\end{aligned}\right\} \xrightarrow[\text{prawdopodobieństwa}]{\text{aksjomaty}\ (\Omega, F, P)}$$

$$P(\{\text{⚀}, \text{⚂}, \text{⚄}\}) = ?$$

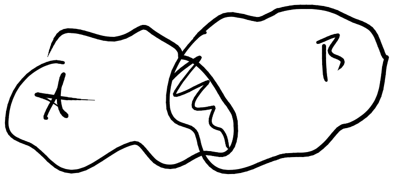
$$P(\{\text{⚁}, \text{⚃}, \text{⚅}\}) = ?$$

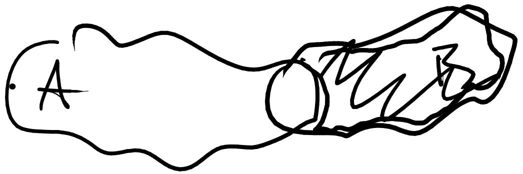
- $P(\emptyset) = 0$
- $P(\underbrace{\Omega \setminus A}_{A^*}) = 1 - P(A)$
- $A \subseteq B \Rightarrow P(A) < P(B)$
- $\forall E \in F: \quad 0 \leq P(E) \leq 1$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

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

$$P(A \cup B) = P(\underbrace{A \cup (B \setminus A)}) = P(A) + P(B \setminus A) =$$

$$= P(A) + P(B \setminus (A \cap B))$$



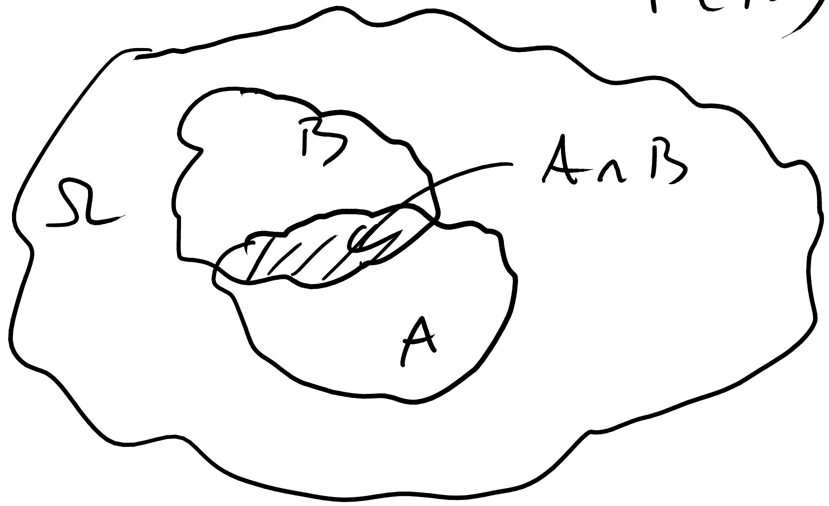


$$\begin{cases} P(A \cup B) = P(A) + P(B \setminus (A \cap B)) \\ P(B) = P(B \setminus (A \cap B)) + P(A \cap B) \end{cases}$$

$$\underbrace{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$$

Tw. Bayesa

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



$$P(A\cap B) = P(A|B)\,P(B)$$

$$P(B\cap A) = P(B|A)\,P(A)$$

$$P(A\cap B) = P(A|B)\,P(B) = P(B|A)\,P(A)$$

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

prawo Bayesa

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

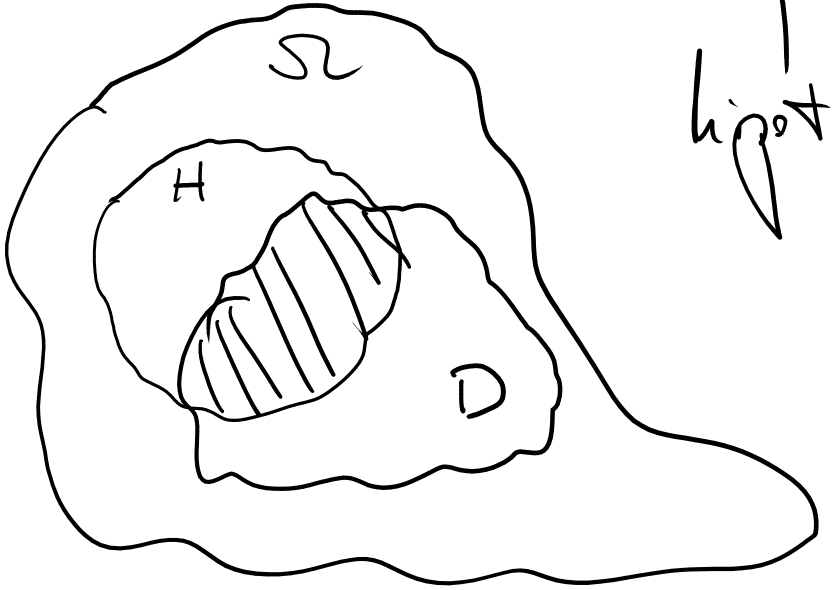
$A \to H$

$B \to D$

$$P(H|D) = P(H) \cdot \frac{P(D|H)}{P(D)}$$

hipoteze

date



$$P(A \wedge B \mid X) = \frac{P(A \wedge B \wedge X)}{P(X)} = \frac{P(B \wedge A \wedge X)}{P(X)} = P(B \wedge A \mid X)$$

$$P(A \wedge B \mid X) = \frac{P(A \wedge B \wedge X)}{P(B \wedge X)} \quad \frac{P(B \wedge X)}{P(X)} =$$

$$= P(A \mid B \wedge X) \cdot P(B \mid X) =$$

$$= P(B \mid A \wedge X) \cdot P(A \mid X)$$

$$P(A \wedge B \mid X) = P(A \mid B \wedge X) \cdot P(B \mid X) = P(B \mid A \wedge X) \cdot P(A \mid X)$$

$$P(A \mid B \wedge X) = P(A \mid X) \cdot \frac{P(B \mid A \wedge X)}{P(B \mid X)}$$

$$P(A \mid B \wedge X) = P(A \mid X) \cdot \frac{P(B \mid A \wedge X)}{P(B \mid X)}$$

$$\downarrow$$

$$P(H \mid D \wedge X) = P(H \mid X) \cdot \frac{P(D \mid A \wedge X)}{P(D \mid X)}$$

$$P(\quad) \rightarrow P(\quad \mid \quad)$$

$$P(A \wedge B \mid X) = P(A \mid B \wedge X) \cdot P(B \mid X) = P(B \mid A \wedge X) \cdot P(A \mid X)$$

$$P(A \vee B \mid X) = P(A \mid X) + P(B \mid X) - P(A \wedge B \mid X)$$

$$\begin{cases} P(A \wedge B \mid X) = P(A \mid B \wedge X) \cdot P(B \mid X) = P(B \mid A \wedge X) \cdot P(A \mid X) \\ P(A \vee B \mid X) = P(A \mid X) + P(B \mid X) - P(A \wedge B \mid X) \end{cases}$$

$\cap \to \wedge$

$\cup \to \vee$

$\underline{A,\ B,\ X}$ → propozycje logiczne

$$\underbrace{P(H|D_2 D_1 X)}_{\text{posterior}} = P(H|X) \cdot \frac{P(D_2 D_1|HX)}{P(D_2 D_1|X)} =$$

$$= P(H|X) \frac{P(D_2|D_1 HX)\, P(D_1|HX)}{P(D_2 D_1|X)} =$$

$$= P(H|X) \frac{P(D_2|D_1 HX)\, P(D_1|HX)}{P(D_2|D_1 X)\, P(D_1|X)} =$$

$$= \underbrace{P(H|D_1 X)}_{\text{prior}} \frac{\overbrace{P(D_2|H D_1 X)}^{\text{likelihood}}}{P(D_2|D_1 X)}$$

P

częstotliwość

$\frac{1}{3}$

prędkość światła

c

$c \approx 300\,000\,000 \ \frac{m}{s}$

$P(c)$

$P(c \mid X)$

$$P(A|X) = \frac{1}{3}$$

$$P(B|X) = \frac{1}{3}$$

$$P(C|X) = \frac{1}{3}$$

$$P(A|w(B)X) = \overbrace{P(A|X)}^{\frac{1}{3}} \cdot \frac{\overbrace{P(w(B)|AX)}^{\frac{1}{3}}}{\frac{1}{3}\,P(w(B)|X) = \frac{1}{3}} = \frac{1}{3}$$

$$P(B|w(B)X) = \overbrace{P(B|X)}^{\frac{1}{3}} \cdot \frac{\overbrace{P(w(B)|BX)}^{\frac{1}{3}}}{P(w(B)|X) = \frac{1}{3}} = \frac{1}{3}$$

$$P(C|w(B)X) = \overbrace{P(C|X)}^{\frac{1}{3}} \cdot \frac{\overbrace{P(w(B)|CX)}^{\frac{1}{3}}}{P(w(B)|X) = \frac{1}{3}} = \frac{1}{3}$$

$$P(A|o(A)w(B)X) = \underbrace{P(A|w(B)X)}_{\frac{1}{3}} \underbrace{P(o(A)|Aw(B)X)}_{0} / P(o(A)|\ldots) = 0$$

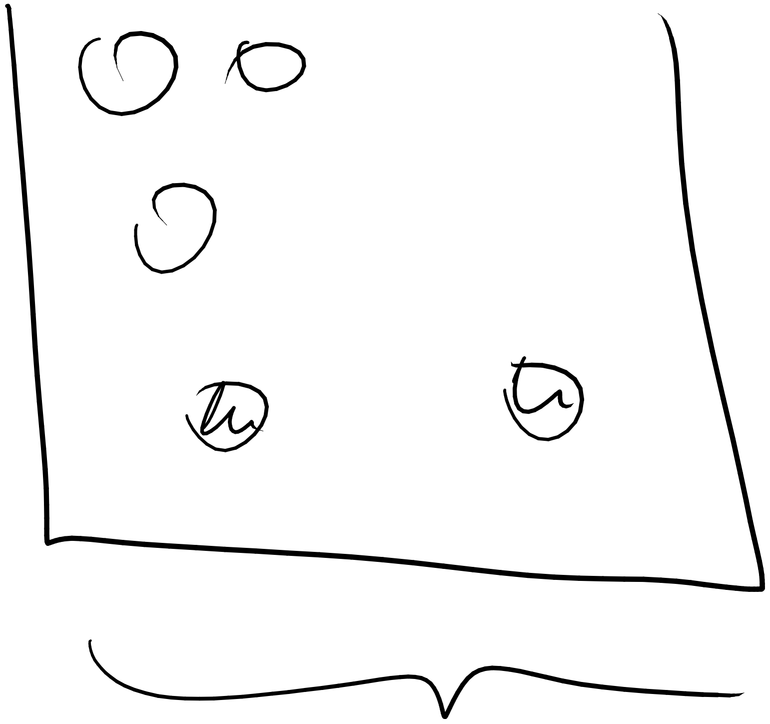
$$P(B|o(A)w(B)X) = \underbrace{P(B|w(B)X)}_{\frac{1}{3}} \underbrace{P(o(A)|Bw(B)X)}_{\frac{1}{2}} / P(o(A)|\ldots) = \frac{1}{3}$$

$$P(C|o(A)w(B)X) = \underbrace{P(C|w(C)X)}_{\frac{1}{3}} \underbrace{P(o(A)|Cw(B)X)}_{1} / P(o(A)|\ldots) = \frac{2}{3}$$

$$P(A \mid X) \longrightarrow P(A \mid w(B)\,X) \longrightarrow P(A \mid \theta(A)\,w(B)\,X)$$

$$P(B \mid X) \longrightarrow P(B \mid w(B)\,X) \longrightarrow P(B \mid \theta(A)\,w(B)\,X)$$

$$P(C \mid X) \longrightarrow P(C \mid w(B)\,X) \longrightarrow P(C \mid \theta(A)\,w(B)\,X)$$



50

$P(\theta \mid D_2 P_1 \ldots)$

$P_u(\theta)$

0 u

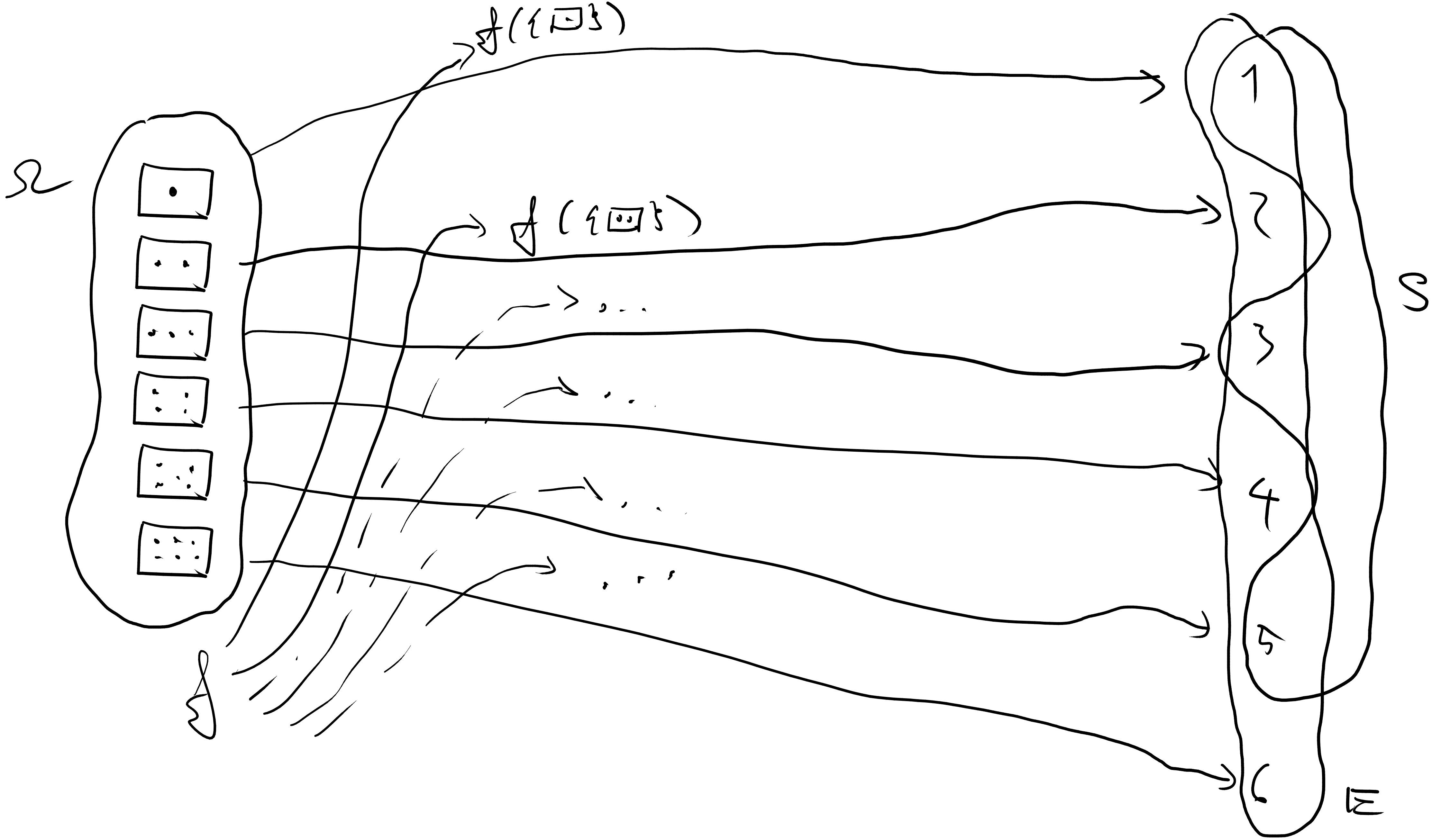
Zmienne losowe

- wartość zależy od wyniku procesu losowego
- formalnie

$$\xi : \Omega \longrightarrow \mathbb{R} \quad (E)$$

- prawdopodobieństwo

$$P(\xi \in \underset{\substack{\uparrow \\ \mathbb{R}}}{S}) = P\left(\{\omega \in \Omega \mid \xi(\omega) \in S\}\right)$$

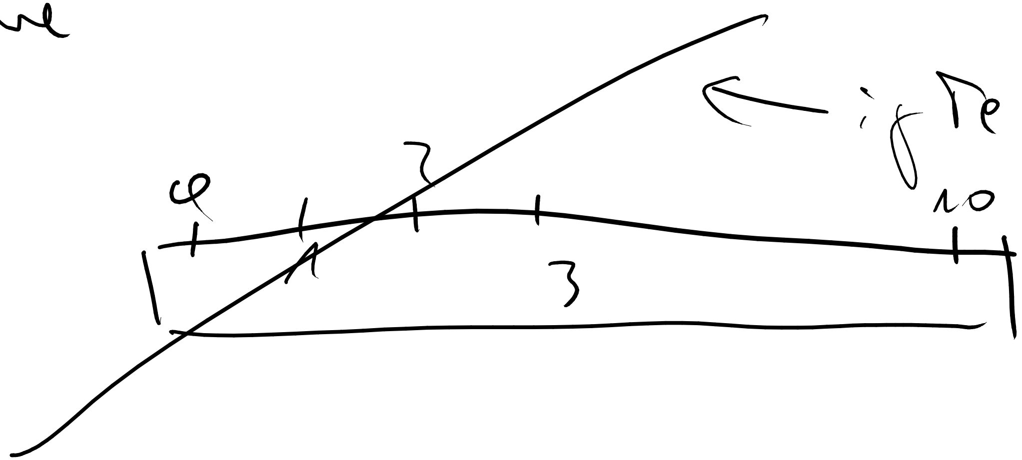

Ω
f
f({⚀})
f({⚁})
1
2
3
4
5
6
S
E

$$P(\underbrace{y \in \{1,3,5\}}_{S}) = P(\{⚀, ⚂, ⚄\}) = \frac{1}{2}$$

$$f(\{⚀\}) = 1$$

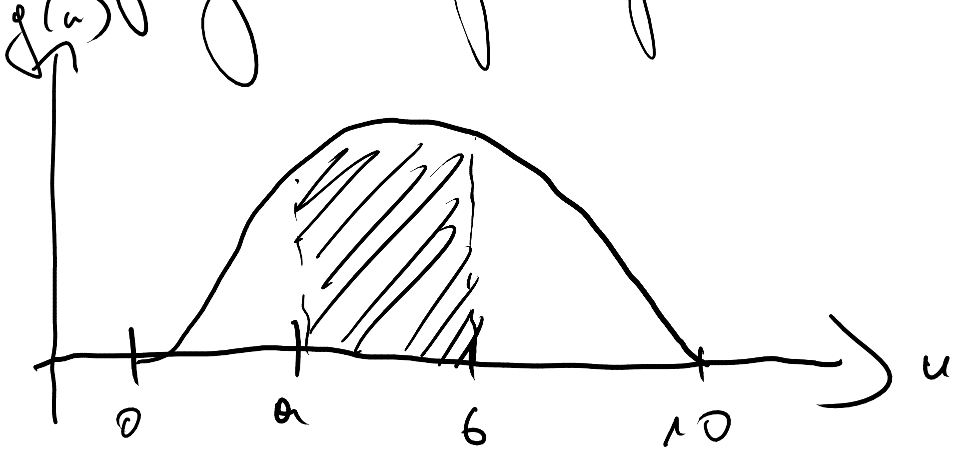
$$f(\{⚂\}) = 3$$

ciągłe zmienne losowe



$$P(y \in \{3\}) = ? = 0$$

funkcja gęstości prawdopodobieństwa



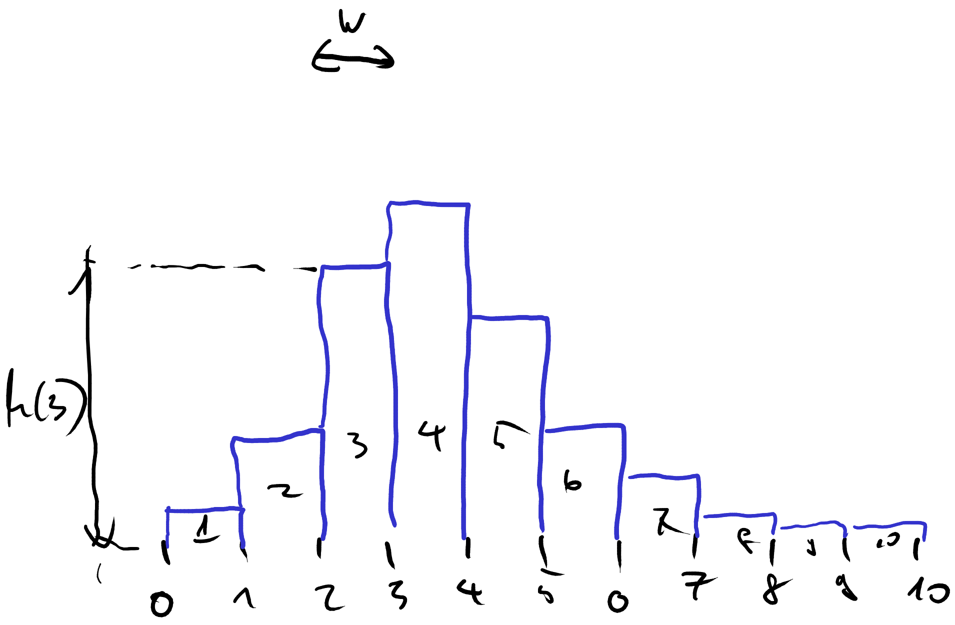
$$P(a \le y \le 6) = \int_a^6 g(x)\,dx$$

$$P(0 < y < 10) = 1 = \int_0^{10} g(x)\,dx$$

- $$\int_{u_{min}}^{u_{max}} g(u)\,du = \underline{1}$$

$$P\left(z \leq u \leq z + dz\right) = g(z)\,dz$$

- $$\forall_z\, P(z \leq u \leq z + dz) > 0 \Rightarrow \forall_u\, g(u) > 0$$



$$\frac{\text{count}(i)}{N \quad w}$$

$$P((i-1)\cdot w \le y \le i\cdot w) = \int_{(i-1)w}^{iw} g(u)\,du \approx w\,\boxed{g\left(\tfrac{1}{2}\big((i-1)w + iw\big)\right)}$$

$$P((i-1)w \le y < iw) = \text{count}(i) / N$$