metody statystyczne 1

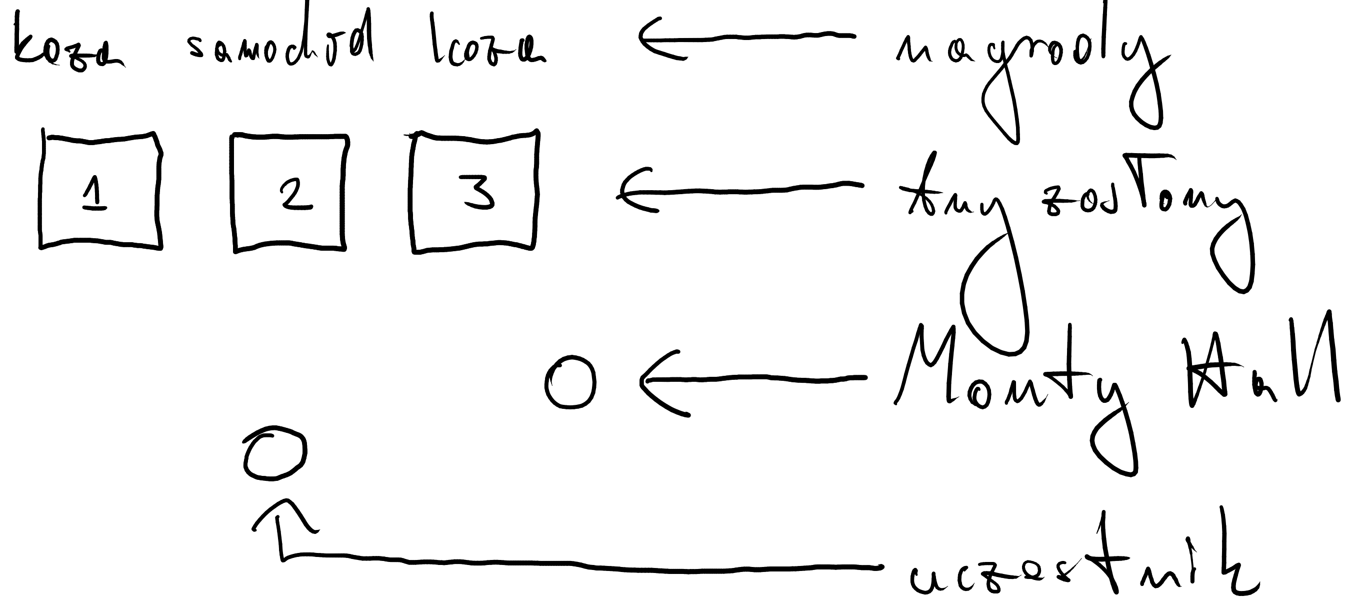
kacper.topolnicki@uj.edu.pl

plan

- problem Montiego Halla
- aksjomaty prawdopodobieństwa
- własności prawdopodobieństwa
- twierdzenie Bayesa
- interpretacje prawdopodobieństwa
- dyskretne zmienne losowe
- ciągłe zmienne losowe
- funkcja gęstości prawdopodobieństwa
- własności rozkładów prawdopodobieństwa

paradoks Montiego Halla

- teleturniej:



- przebieg:

1) uczestnik wybiera zasłonę
2) Monty Hall odsłania kozę
3) uczestnik może zmienić zasłonę
4) wszystkie zasłony idą w górę

- czy uczestnik powinien zmienić wybór

monty_hall.nb

link dostępny na stronie

aksjomaty prawdopodobieństwa

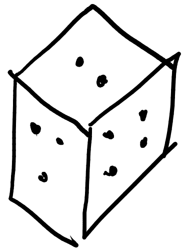
- Kołmogorow 1933

$$(\Omega, F, P)$$

przestrzeń wszystkich „zdarzeń losowych"

zbiór wszystkich „zdarzeń elementarnych"

funkcja przypisująca „zdarzeniu losowemu" „prawdopodobieństwo" $\in [0,1]$



Ω = {⚀, ⚁, ⚂, ⚃, ⚄, ⚅}

zdarzenie elementarne

zbiór pusty

nieparzysta liczba oczek

F = {∅, {⚀, ⚂, ⚄}, ...}

podzbiory Ω

zdarzenie losowe

$P(\{⚀\}) = \frac{1}{6}$

$P(\{⚁\}) = \frac{1}{6}$

$P(\{⚂\}) = \frac{1}{6}$

$P(\{⚃\}) = \frac{1}{6}$

$P(\{⚄\}) = \frac{1}{6}$

$P(\{⚅\}) = \frac{1}{6}$

aksjomaty (Ω, F, P)

dwa oczka lub trzy oczka lub cztery oczka

P({⚁, ⚂, ⚃})

P({⚀, ⚂, ⚄})

...

- $\forall E \in F : P(E) \in \mathbb{R} \land P(E) \geq 0$

np. {⚀, ⚂, ⚄} np. {∅, {⚀, ⚁, ⚄}, ...} l. rzeczywiste

- $P(\Omega) = 1$

prawdopodobieństwo zajścia przynajmniej jednego zdarzenia elementarnego

- $E_i \in F \quad \forall_{i \neq j} \; E_i \cap E_j = \emptyset \; : \; P(E_1 \cup E_2 \cup \ldots \cup E_N) = \sum_{i=1}^{N} P(E_i)$

zbiory rozłączne

liczenie prawdopodobieństwa zdarzeń rozłącznych

może to wystarczy aby policzyć:

$P(\{⚀, ⚁, ⚂\})$

$P(\{⚁, ⚃, ⚅\})$?

własności prawdopodobieństwa

- $P(\emptyset) = 0$

  zbiór pusty

- $A \subseteq B \Rightarrow P(A) < P(B)$

- $P(\Omega \setminus A) = 1 - P(A)$

  A^*

- $\forall E \in F: \; 0 \leq P(E) \leq 1$

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

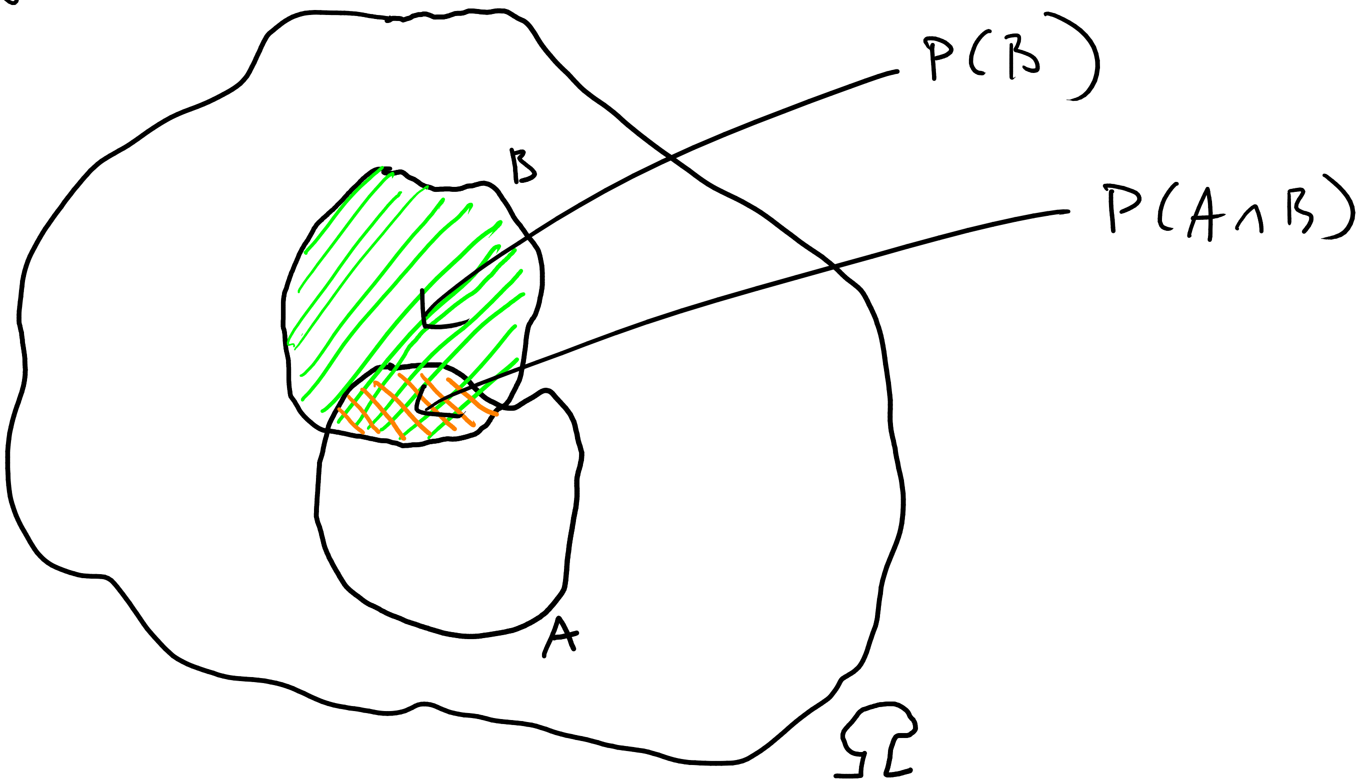
  gdy rozłączne

  $P(A \cap B) = P(\emptyset) = 0$

Twierdzenie Bayesa

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

prawdopodobieństwo warunkowe



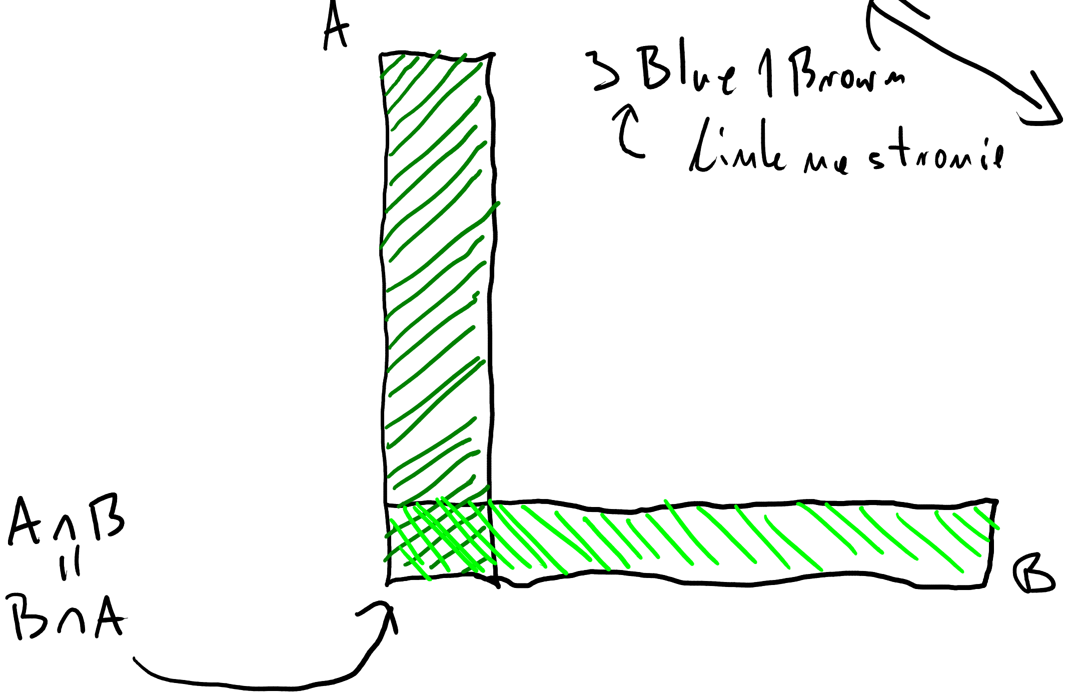
$$\left.\begin{aligned} P(A \cap B) &= P(A \mid B)\,P(B) \\ P(\underset{\underset{A \cap B}{\parallel}}{B \cap A}) &= P(B \mid A)\,P(A) \end{aligned}\right\} \Rightarrow P(A \mid B)\,P(B) = P(B \mid A)\,P(A)$$

$$P(A \mid B) = \frac{P(B \mid A)\,P(A)}{P(B)} \leftarrow \text{twierdzenie Bayesa}$$

3Blue1Brown

link na stronie

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



– Monty Hall bez zmiany:

• $P(A) = \frac{1}{3}$

• $P(B|A) = \frac{1}{2}$

odsłonięcie kozy za B pod warunkiem samochodu za A

$\boxed{1}$ $\boxed{2}$ $\boxed{3}$

wybór (1)

odsłonięto koza (2)

A – samochód za 1

B – koza odsłonięta za 2

• $P(B) = \underbrace{\frac{1}{3}}_{\text{samochód za 1}} \cdot \underbrace{\frac{1}{2}}_{\text{odsłaniamy 2 lub 3}} + \underbrace{\frac{1}{3}}_{\text{samochód za 2}} \cdot \underbrace{0}_{\text{uwaga! nie możemy odsłonić 2!}} + \underbrace{\frac{1}{3}}_{\text{samochód za 3}} \cdot \underbrace{1}_{\text{nie możemy odsłonić 1, 3}} = \frac{3}{6} = \frac{1}{2}$

tw. Bayesa

• $P(A|B) = \frac{\overset{\frac{1}{2}}{P(B|A)} \cdot \overset{\frac{1}{3}}{P(A)}}{\underset{=\frac{1}{2}}{P(B)}} = \frac{1}{3}$

samochód za 1 pod warunkiem odsłonięcia kozy za 2, prawdopodobieństwo wygranej

- Monty Hall ze zmianą

C_i – samochód za zasłoną i

X_i – pierwszy wybór gracza to i

H_i – odsłonięcie zasłony i

$$\bullet\; P(\underset{\substack{\uparrow\\ \text{samochód za zasłoną 2}}}{C_2} \mid \underbrace{(H_3 \wedge X_1)}_{\substack{\text{gracz wybiera 1}\\ \text{odsłonięto zasłonę 3}}}) = \frac{P(C_2 \wedge H_3 \wedge X_1)}{P(H_3 \wedge X_1)} = \frac{P(H_3 \mid (C_2 \wedge X_1))\, P(C_2 \wedge X_1)}{P(H_3 \wedge X_1)} =$$

samochód za zasłoną 2 gracz zmienia wybór zasłony na 2

prawdopodobieństwo zwycięstwa

1,2,3 → 2,1,3

1,2,3 → 1,3,2

...

$$= \frac{\boxed{P(H_3 \mid (C_2 \wedge X_1))}\,\boxed{P(C_2 \wedge X_1)}}{\boxed{P(H_3 \mid X_1)}\,\boxed{P(X_1)}} = \dots$$

$$P(H_3 \mid (C_1 \wedge X_1)) = \frac{1}{2} \qquad \boxed{P(H_3 \mid X_1) = \frac{1}{2}}$$

$$\boxed{P(H_3 \mid (C_2 \wedge X_1)) = 1}$$

$$P(H_3 \mid (C_3 \wedge X_1)) = 0$$

$$\boxed{P(C_i \wedge X_j) = \underbrace{P(C_i)\,P(X_j)}_{\text{niezależne}} = \frac{1}{3}\cdot\frac{1}{3} = \frac{1}{9}}$$

$$\boxed{P(X_i) = \frac{1}{3}}$$

$$\dots = \frac{1 \cdot \frac{1}{9}}{\frac{1}{2}\cdot\frac{1}{3}} = \frac{6}{9} = \frac{2}{3}$$

interpretacje prawdopodobieństwa

- prawdopodobieństwo obiektywne (frequentist probability)

$\underbrace{\text{⚀ ⚁ ⚀ ⚅ } \dots}_{\text{powtarzający się proces}}$

⚀ występuje z częstością ≈ 0.166

⚁ występuje z częstością ≈ 0.166

⚂ występuje z częstością ≈ 0.166

...

częstość kojarzymy z prawdopodobieństwem

• prawdopodobieństwo subiektywne (Bayesian probability)

$$p(h|e) = \frac{p(e|h)\; p(h)}{p(e)}$$

h – hipoteza

e – dowód

posterior probability

prior probability

prawdopodobieństwo odzwierciedla posiadaną wiedzę

na początku: $p(h)$ $\xrightarrow{\text{prawo Bayesa}}$ na końcu: $p(h|e)$

przykład:

$$p\left(h=(n, N-n) \mid e=m_i \right) =$$

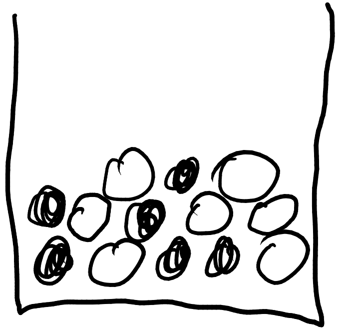
l. czarnych, l. białych

możemy wyliczyć

$$= \frac{p(e=m_i \mid h=(n, N-n)) \cdot p(h=(n, N-n))}{p(e=m_i)}$$

odzwierciedla wiedzę przed wylosowaniem

możemy wyliczyć z warunku normalizacji $\sum_h p(h|e) = 1$



N - kul

n - czarnych

N - n białych

- losujemy kulę $m_1 = c, b$
- na podstawie wylosowanych $(m_1, m_2, \ldots)$ stawiamy hipotezy dotyczące n
- znamy N

bayes.nb

link dostępny na stronie

- Dlaczego na końcu nie mamy ostrego piku dla wtórnej liczby ?
    - informacja o bieżącej liczbie kul w pudle

można spróbować dodać

zmienne losowe

- zmienna, której wartości zależy od wyniku procesu losowego
- formalnie:

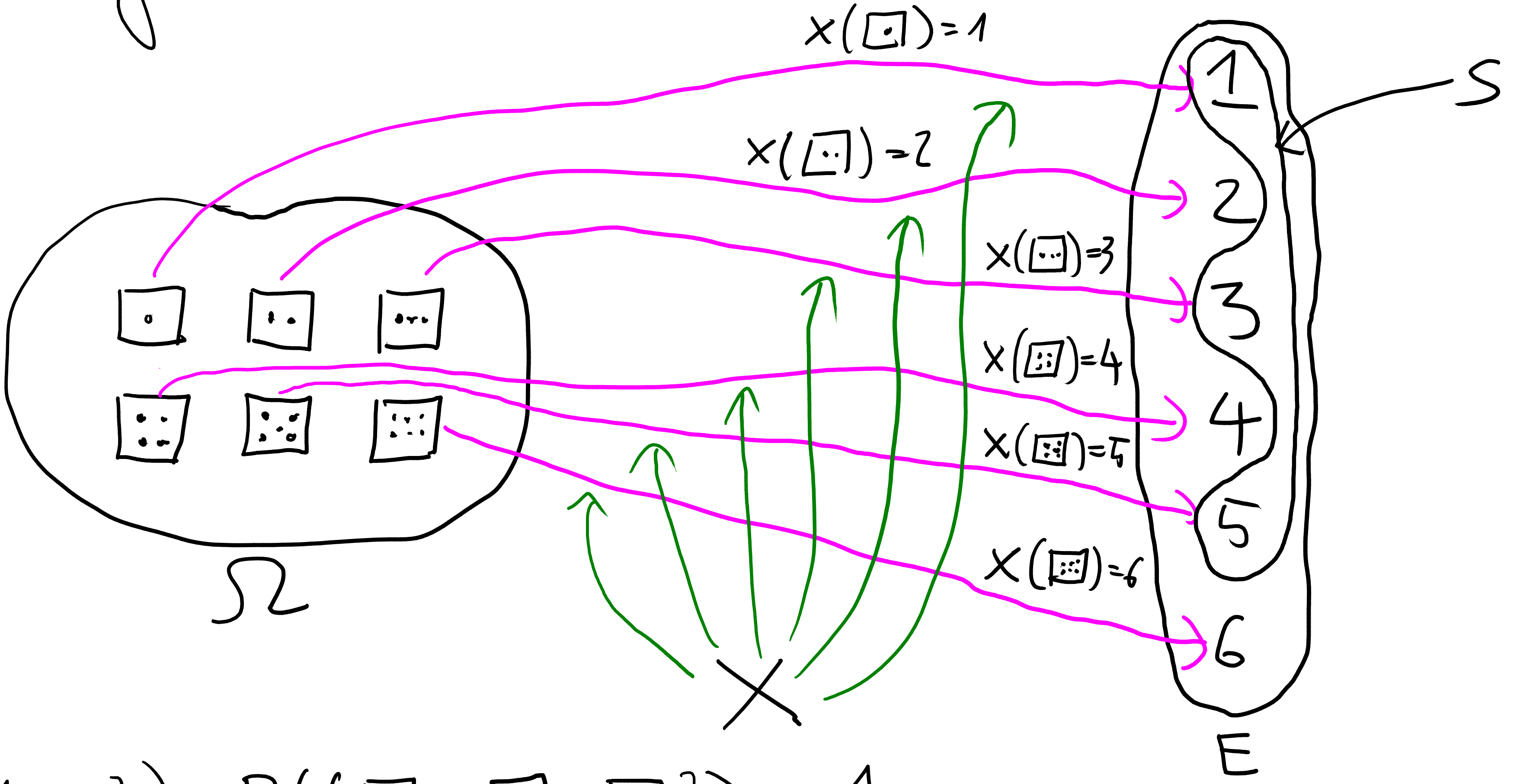
  funkcja $X : \Omega \to E$

  E – przestrzeń mierzalna np. $E = \mathbb{R}$
- prawdopodobieństwo zmiennej losowej:

$$P(y \in S) = P(\{\omega \in \Omega \mid X(\omega) \in S\})$$

$S \subset E$

w praktyce:

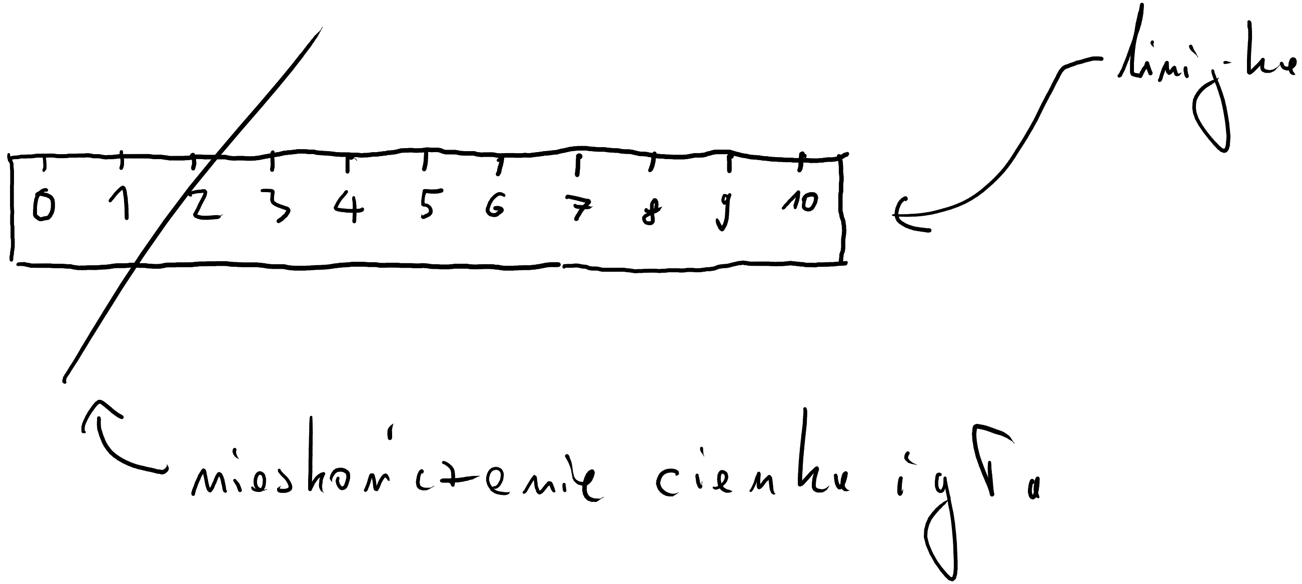


$P(y \in \{1,2,3\}) = P(\{⚀, ⚂, ⚄\}) = \frac{1}{2}$

S

$X(⚀) = 1 \in S$ $\quad X(⚂) = 3 \in S$ $\quad X(⚄) = 5 \in S$

ciągłe zmienne losowe

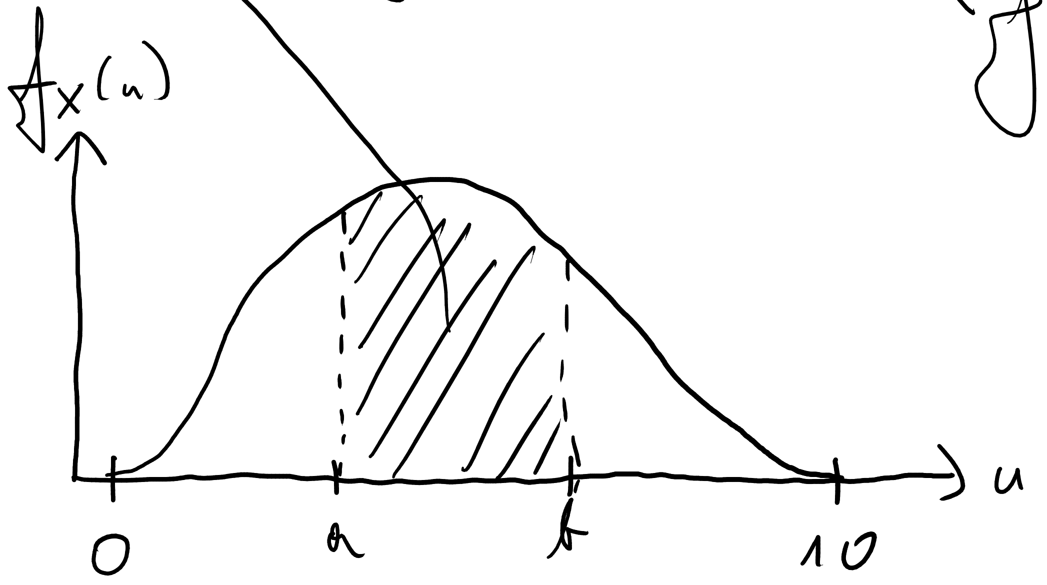


$$P(y \in \{3\}) = ? = 0$$

$$P(a \le y \le b) = \int_a^b f_X(u)\,du$$

$(y \in S = [a, b]$

funkcja gęstości prawdopodobieństwa dla zmiennej losowej X



$$P(0 \le y < 10) = 1 = \underbrace{\int_0^{10} f_X(u)\,du}_{\text{normalizacja}}$$

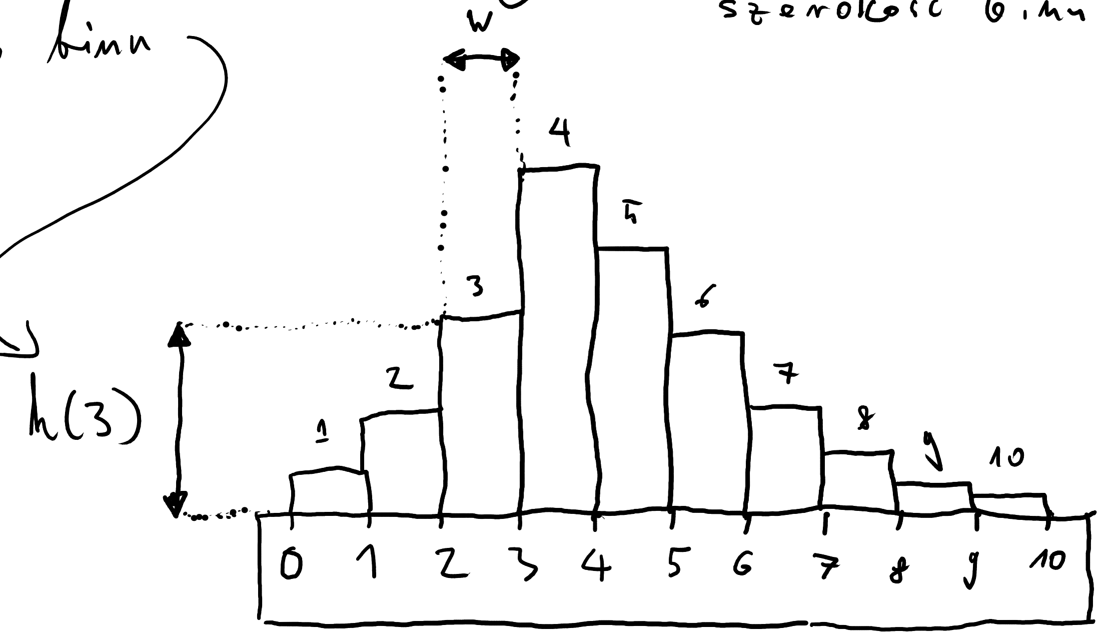
właściwości funkcji gęstości prawdopodobieństwa:

- $\int_{u_{min}}^{u_{max}} f(u)\,du = 1$
- $P(z \leq y \leq z + dz) = f(z)\,dz$
- $\forall_{z}\ P(z \leq y \leq z + dz) > 0 \Rightarrow \forall_{z}\ f(z) > 0$

histogramy

wysokość trzeciego binu

szerokość binu



jak dobrać wysokość binów?

wysokość słupka

często można wybrać w bibliotekach do tworzenia wykresu

- prawdopodobieństwo, że y wpadnie w bin o numerze i:

$$P((i-1)w \le y \le i \cdot w) = \int_{(i-1)w}^{i\cdot w} f(u)\,du$$

- ile będzie zliczeń w binie i:

$$N \cdot P((i-1)w \le y \le i \cdot w)$$

N ← całkowita liczba przypadków

- przybliżenie funkcji gęstości prawdopodobieństwa

$$P((i-1)w \le y \le i \cdot w) \approx w \cdot f\left(\frac{1}{2}((i-1)w + iw)\right) \qquad f(\ldots) \approx \frac{P(\ldots)}{w}$$

opis funkcji gęstości prawdopodobieństwa wartość oczekiwana

$$E(X) = \int_{-\infty}^{+\infty} u \, f_X(u) \, du$$

$$E(g(X)) = \int_{-\infty}^{+\infty} g(u) \, f_X(u) \, du$$

zmienna losowa $\mathbb{R} \to \mathbb{R}$

σταθε

$$g(x) = c \qquad E(c) = \int_{-\infty}^{+\infty} c\, f_x(u)\,du = c \underbrace{\int_{-\infty}^{+\infty} f_x(u)\,du}_{= 1} = c$$

$$g(x) = ax + b \qquad E(ax+b) = \int_{-\infty}^{+\infty} (au+b) f_x(u)\,du =$$

$$= \int_{-\infty}^{+\infty} a u\, f_x(u)\,du + \underbrace{\int_{-\infty}^{+\infty} b\, f_x(u)\,du}_{b} =$$

$$= a \int_{-\infty}^{+\infty} u\, f_x(u)\,du + b =$$

$$= a E(x) + b$$

Agp.nb

link na stronie

dystrybuanta

$$F_x(y) = \int_{-\infty}^{y} \underset{\geq 0}{du}\; \underset{\geq 0}{f_x(u)} = P(x \leq y)$$

zawsze rośnie

$$\frac{d F_x(y)}{dy} = f_x(y)$$

wariancja

$$\mathrm{Var}(X) = \int_{-\infty}^{+\infty} du \, (u - E(X))^2 f_X(u) =$$

$$= \int_{-\infty}^{+\infty} du \left(u^2 - 2uE(X) + E^2(X)\right) f_X(u) =$$

$$= E(X^2) - 2E^2(X) + E^2(X) =$$

$$= E(X^2) - E(X)$$

odchylenie standardowe

$$\sigma(X) = \sqrt{var(X)}$$

$$\sigma(g(X)) = \sqrt{var(g(X))}$$

momenty rozkładu

$$\bar{\mu}_k = E(X^k)$$ względem 0

$$\mu_k = E((X - \underset{\underset{E(X)}{\parallel}}{\mu_1})^k)$$

skośność

$$\beta_1 = \frac{\mu_3}{\sigma(X)^3}$$

kurtoza

$$\beta_2 = \frac{\mu_4}{\sigma(X)^4} - 3$$

kwantyl

x_p jest kwantylem rzędu p gdy $F_X(x_p) = p$

mod