

# Rozkłady prawdopodobieństwa

Wojciech Kotłowski

Statystyka i analiza danych 2019/2020

24.03.2020

## Rozkłady prawdopodobieństwa

- **Dwupunktowy**  $B(p)$ , dla zmiennej  $X \in \{0, 1\}$ .

$$P(X = 1) = 1 - P(X = 0) = p, \quad E[X] = p, \quad D^2[X] = p(1-p).$$

- **Dwumianowy**  $B(n; p)$ :

$$X = \sum_{k=1}^n X_k, \quad \text{gdzie } X_k \sim B(p).$$

Suma  $n$  niezależnych zmiennych dwupunktowych. Zachodzi:

$$E[X] = np, \quad D^2[X] = np(1-p).$$

Dlaczego?

## Użyteczne własności

Niech  $X_1, \dots, X_n$  – **i.i.d.**<sup>1</sup> oraz  $E[X_k] = \mu$ ,  $D^2[X_k] = \sigma^2$ .

- Dla sumy  $X = \sum_{k=1}^n X_k$ :

$$\begin{aligned} E[X] &= E[X_1] + \dots + E[X_n] = n\mu, \\ D^2[X] &= D^2[X_1] + \dots + D^2[X_n] = n\sigma^2. \end{aligned}$$

- Dla średniej  $\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k = \frac{1}{n} X$ :

$$\begin{aligned} E[\bar{X}] &= \frac{1}{n} E[X] = \mu, \\ D^2[\bar{X}] &= \frac{1}{n^2} D^2[X] = \frac{\sigma^2}{n}, \quad \text{czyli } D[\bar{X}] = \frac{\sigma}{\sqrt{n}}. \end{aligned}$$

---

<sup>1</sup>i.i.d.: *independent and identically distributed* – niezależne zmienne losowe o tym samym rozkładzie

# Rozkład normalny

$X \sim N(\mu, \sigma^2)$  ma gęstość prawdopodobieństwa:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Rodzina rozkładów normalnych ma wyjątkową własność:

- Dowolna kombinacja liniowa zmiennych o rozkładzie normalnym ma rozkład normalny!

## Prawo wielkich liczb

$$X_1, \dots, X_n - \text{i.i.d.}, E[X_k] = \mu, D^2[X_k] = \sigma^2.$$

$$E[\bar{X}] = \mu, \quad D[\bar{X}] = \frac{\sigma}{\sqrt{n}}.$$

Rozkład średnia z próby **koncentruje się** wokół wartości oczekiwanej  $\mu$ .

# Centralne Twierdzenie Graniczne

$X_1, \dots, X_n$  – i.i.d.,  $E[X_k] = \mu$ ,  $D^2[X_k] = \sigma^2$ .

Zdefiniujmy **ustandaryzowaną** średnią  $Z = \frac{\bar{X} - \mu}{\sigma} \sqrt{n}$ .

$$E[Z] = 0, \quad D^2[Z] = 1.$$

Zachodzi:

$$Z \xrightarrow{D} N(0, 1).$$

Przez CTW rozkład normalny występuje **powszechnie!**

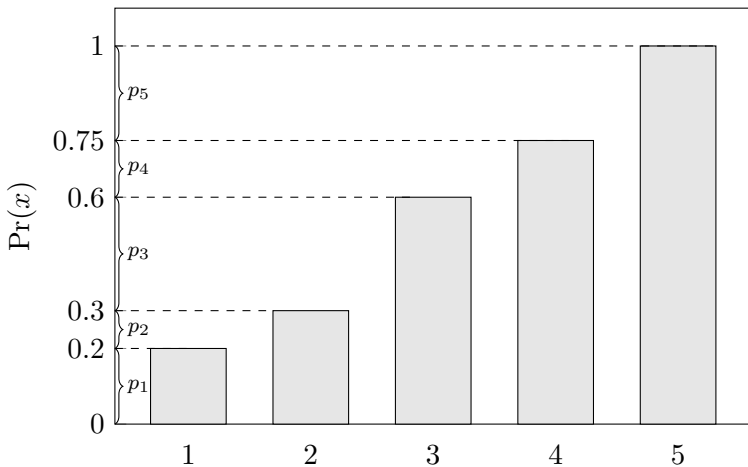
- Cechy fizyczne (wzrost, waga, BMI), wyniki testów (IQ), odstępstwa od normy w produkcji, wszelkie uśrednianie.

## Losowanie liczb

Jak wylosować liczbę z danego rozkładu  $P$  mając do dyspozycji tylko generator liczb z zakresu  $[0, 1]$  (jednostajnie)?

## Losowanie liczb

Jak wylosować liczbę z danego rozkładu  $P$  mając do dyspozycji tylko generator liczb z zakresu  $[0, 1]$  (jednostajnie)?





## Losowanie liczb

Jak wylosować liczbę z danego rozkładu  $P$  mając do dyspozycji tylko generator liczb z zakresu  $[0, 1]$  (jednostajnie)?

