Metody Statystyczne

wykład 3

17 XI 2024

kacper.topolnicki@uj.edu.pl

- rozkład normalny
- estymacja punktowa
- estymacja przedziałowa
- procesy stochastyczne
  - łańcuchy Markova

rozkład normalny

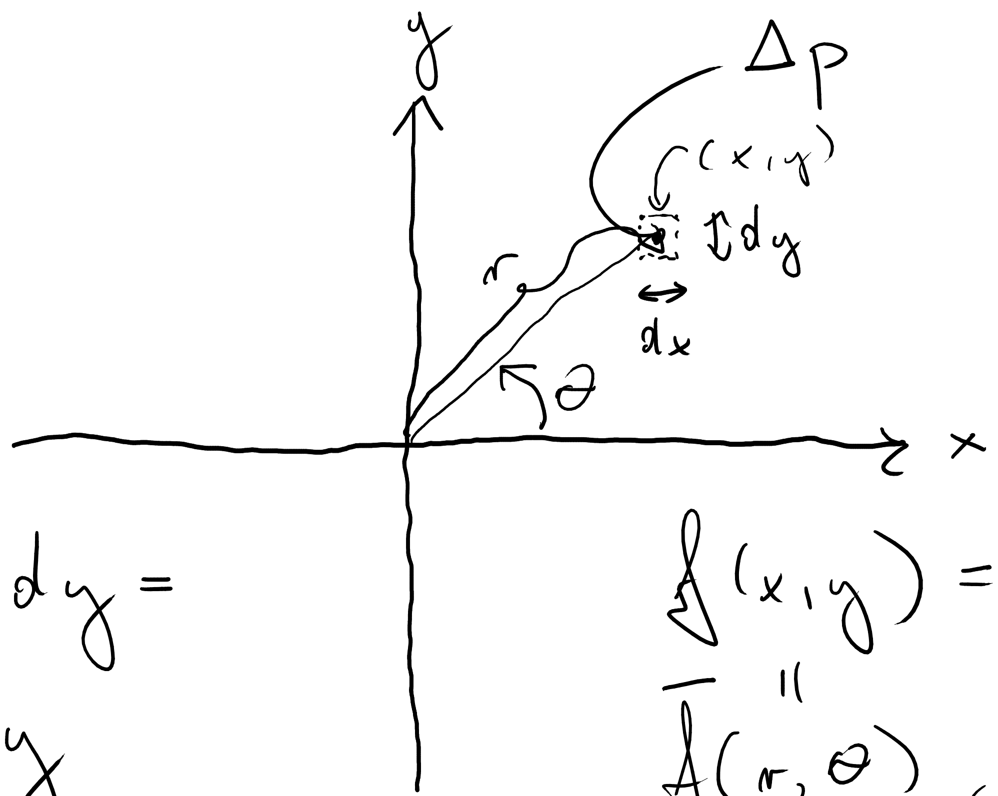
X, Y - niezależne zmienne losowe

FGP jest taka sama dla X oraz Y

f

$$f(x) \to 0 \text{ gdy } x \to \pm\infty$$

$$f(-x) = f(x)$$



$$\Delta p = f(x)\,dx\; f(y)\,dy =$$

$$= \underbrace{f(x,y)}_{\substack{\text{FGP 2D} \\ \text{dla } x,\,y}}\,dx\,dy$$

$$f(x,y) = f(x)\,f(y)$$

$$\| $$

$$\bar{f}(r,\theta)$$

$y = r\sin\theta$

$x = r\cos\theta$

$$\underbrace{0 = \frac{\partial \bar{f}}{\partial \theta}(r,\theta)}_{\text{założenie}} = f(x)\,\frac{\partial f}{\partial \theta}(y) + f(y)\,\frac{\partial f}{\partial \theta}(x)$$

$$0 = f(x)\,\frac{\partial f}{\partial \theta}(r\sin\theta) + f(y)\,\frac{\partial f}{\partial \theta}(r\cos\theta)$$

$$0 = f(\overbrace{r\cos\theta}^{x})\, f'(\overbrace{r\sin\theta}^{y})\,\overbrace{r\cos\theta}^{x} + f(\underbrace{r\sin\theta}_{y})\, f'(\underbrace{r\cos\theta}_{x})\,(-\underbrace{r\sin\theta}_{y})$$

$$0 = f(x)\, f'(y)\, x - f(y)\, f'(x)\, y$$

$$f(y)\, f'(x)\, y = f(x)\, f'(y)\, x$$

$$\underbrace{\frac{f'(x)}{x\, f(x)} = \frac{f'(y)}{y\, f(y)}}_{\text{speTmiome dllo kosolego x,y}}$$

$$\frac{f'(x)}{x\,f(x)} = c \quad , \quad \frac{f'(y)}{y\,f(y)} = c$$

$$f'(x) = cx \cdot f(x)$$

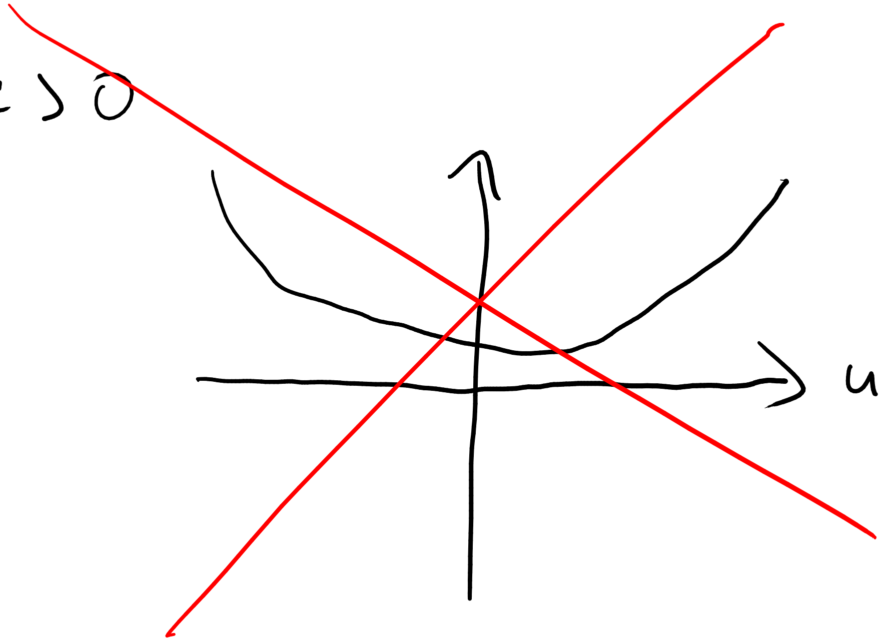
$$f(\underset{\uparrow y}{x}) = A \cdot e^{\frac{c}{2}\underset{\uparrow y}{x^2}}$$

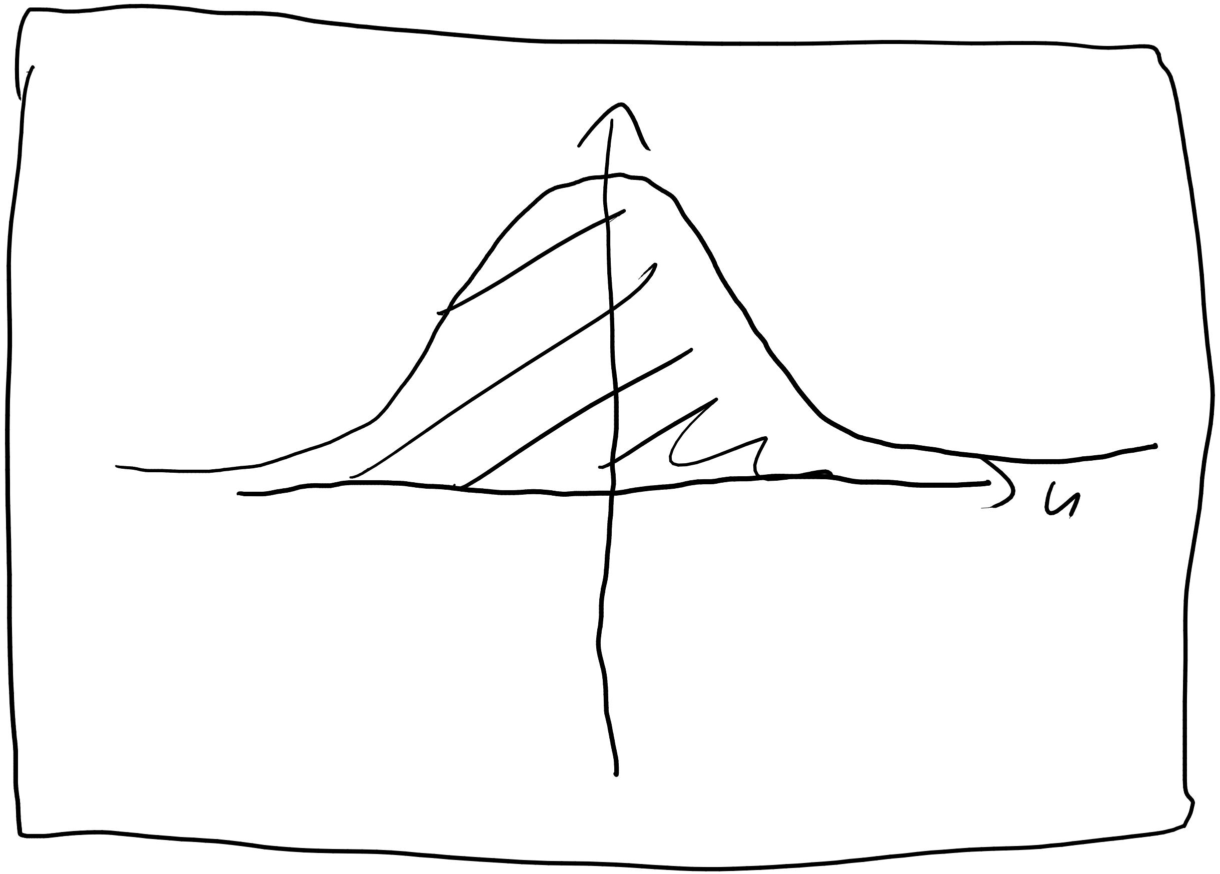
$$f'(x) = A \cdot e^{\frac{c}{2}x^2} \cdot \frac{c}{\cancel{2}} \cdot \cancel{2}x =$$

$$= c \cdot x \cdot \underbrace{A e^{\frac{c}{2}x^2}}_{f(x)}$$

$$\int_{-\infty}^{+\infty} du\, f(u) = 1$$

$c > 0$





$$\int_{-\infty}^{+\infty} f(u)\,du = 1$$

$$f(u) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(u-\mu)^2}{\sigma^2}\right)$$

wartość oczekiwana

$A, C \rightarrow \sigma, \mu$ $\quad\quad \sigma > 0$

odchylenie standardowe

rozkład Cauchiego

rozkład normalny

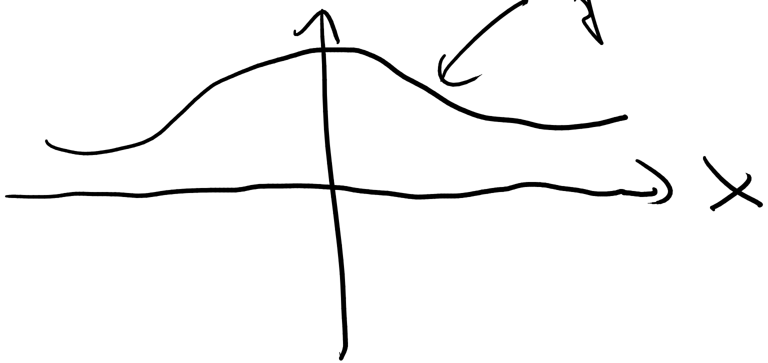
$$Z = X + Y$$

rozkład normalny

rozkład cauchiego

estymacja punktowa

FGP



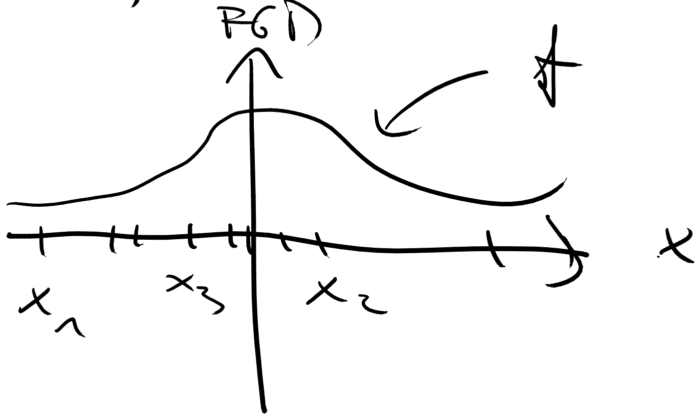
$$E(X) = \int_{-\infty}^{+\infty} dx \; x \, f(x)$$

$$Var(X) = \int_{-\infty}^{+\infty} dx \, f(x) \, (x - E(X))^2$$

$\rho_{x,x}$

w praktyce: zamiast f mamy

$\{x_1, x_2, \dots, x_N\}$ losowane z f



jak oszacować $E(X)$, $Var(X)$, $\rho_{X,Y}$

korzystamy z estymatorów

$$\overbrace{T_N(\theta)}^{\text{jest zmienną losową}} = \hat{\theta}$$

E, Var

liczba danych

$$B(\hat{\theta}) = E(\hat{\theta}) - \theta$$ ← "mele"

błąd

$$\underbrace{B(\hat{\theta}) = 0}_{\text{estymatorem nieobciążony}}$$

$$\hat{\theta} = \overline{X} = \frac{X_1 + X_2 + \ldots + X_n}{N}$$

θ

dowolny

$E(X_i)$

$$B(\hat{\theta}) = E(\hat{\theta}) - \theta = E(\overline{X}) - \theta = E(X_i) - \theta =$$

$$= E(X_i) - E(X_i) =$$

$$= 0$$

$$\hat{\theta}_1 = X_1$$

wartości oczekiwan

$$\hat{\theta} = \frac{X_1 + X_2 + \dots + X_n}{N}$$

$$B(\hat{\theta}_1) = E(\hat{\theta}_1) - \theta =$$

$$= E(X_1) - \theta = \underline{0}$$

mean $\quad$ error

squared

$$MSE(\hat{\theta}) = E\left((\hat{\theta} - \theta)^2\right)$$

$$\hat{\theta}_1 = X_1$$

$$\hat{\theta}_2 = \frac{X_1 + X_2 + \dots + X_n}{N} = \overline{X}$$

$$MSE(\hat{\theta}_1) = E\left(\left(\hat{\theta}_1 - \theta\right)^2\right) = E\left(\left(X_1 - E(X_1)\right)^2\right) =$$

$$= Var(X_1) = \sigma^2$$

$$MSE(\hat{\theta}_2) = E\left(\left(\hat{\theta}_2 - \theta\right)^2\right) = E\left(\left(\overline{X} - \theta\right)^2\right) =$$

$$= \underbrace{Var\left(\overline{X} - \theta\right)}_{Var(\overline{X})} + E\left(\overline{X} - \theta\right)^2 = \frac{\sigma^2}{N}$$

$$\underbrace{\underbrace{\lim_{N \to \infty}} P\left(|\hat{\theta}_N - \theta| \geq \epsilon \right) = 0 \quad \text{dla dowolnego } \epsilon}_{\text{estymator konsystentny}}$$

l. eksperymentów

$$\sigma^2 = E\left((X - \cancel{\mu})^2 \right)$$

w. oczekiwana

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{k=1}^{N} (X_k - \cancel{\mu})^2$$

$$\bar{X}$$

$$\bar{S}^2 = \frac{1}{N} \sum_{k=1}^{N} (X_k - \bar{X})^2 =$$

$$= \frac{1}{N} \left(\sum_{k=1}^{N} X_k^2 - N\bar{X}^2 \right)$$

$\mathbb{E}(X_i) = \underline{\mu}$ $\qquad$ $Var(X_i) = \underline{\sigma^2}$

$$B(\bar{S}^2) = \mathbb{E}(\bar{S}^2) - \sigma^2$$

$$\mathbb{E}(\bar{X}^2) = \mathbb{E}(\bar{X})^2 + Var(\bar{X}) = \mu^2 + \frac{\sigma^2}{n}$$

$$\mathbb{E}(\bar{S}^2) = \frac{1}{N}\left(\sum_{k=1}^{N} \mathbb{E}(X_k^2) - n\,\mathbb{E}(\bar{X}^2)\right) =$$

$$= \frac{1}{N}\left(N(\mu^2 + \sigma^2) - n\left(\mu^2 + \frac{\sigma^2}{n}\right)\right) =$$

$$= \frac{N-1}{N}\sigma^2$$

$$B(\bar{S}^2) = \mathbb{E}(\bar{S}^2) - \sigma^2 = -\frac{\sigma^2}{N}$$

$$B\left(\underbrace{\frac{1}{N-1}\sum_{k=1}^{n}(X_k - \bar{X})^2}_{\text{estymator wariancji}}\right) = ?$$

estymacja przedziałowa

$$T_W(\theta) = \hat{\theta} \in$$

$$\downarrow$$

$$\underbrace{\left[T^L(\theta) \, , \, T^R(\theta) \right]}_{\text{przedział ufności dla } \theta}$$

$$P\left(T_N^L(\theta) \leq \theta \leq T_N^R(\theta)\right) = \gamma$$

poziom ufności

wartość oczekiwana

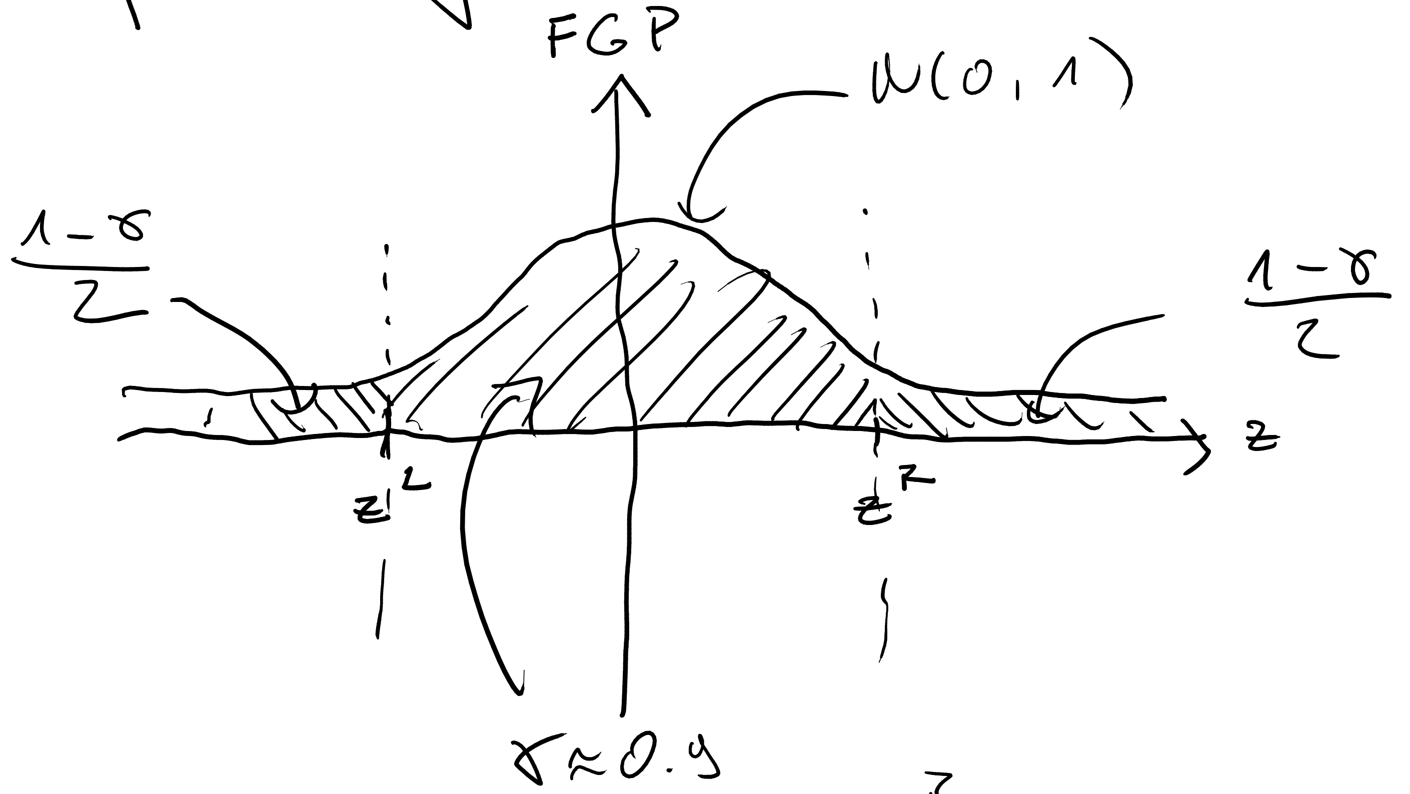
$$\sigma^2 \leftarrow \text{znana} \rightarrow \sigma(X)$$

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$z = \frac{\bar{X} - E(X)}{\sigma(\bar{X})} = \frac{(\bar{X} - E(X))\sqrt{N}}{\sigma(X)}$$

$FGP(z) \leftarrow$ rozkład normalny $N(0,1)$

$\gamma \approx 0.9$ ← poziom ufności



$$z^L = z_{\frac{1-\gamma}{2}} = - z_{\frac{\gamma+1}{2}}$$

$$z^R = z_{\frac{1-\gamma}{2}+\gamma} = z_{\frac{\gamma+1}{2}}$$

$$P(z^L \leq z \leq z^R) = \gamma =$$

$$= P\left(z^L \leq \frac{(\overline{X} - E(x))\sqrt{N}}{\sigma(x)} \leq z^R\right) =$$

$$= P\left(\frac{z^L \sigma(x)}{\sqrt{N}} \leq \overline{X} - E(x) \leq \frac{z^R \sigma(x)}{\sqrt{N}}\right) =$$

$$= P\left(\frac{z^L \sigma(x)}{\sqrt{N}} - \overline{X} \leq -E(x) \leq \frac{z^R \sigma(x)}{\sqrt{N}} - \overline{X}\right) =$$

$$= P\left(\overline{X} - \frac{z^L \sigma(x)}{\sqrt{N}} \geq E(x) \geq \underbrace{\overline{X} - \frac{z^R \sigma(x)}{\sqrt{N}}}_{T_N^R(E(x))}\right) =$$

$$= P\left(\underbrace{\overline{X} - \frac{z_{\frac{1+\gamma}{2}} \sigma(x)}{\sqrt{N}}}_{T_N^L(E(x))} \leq E(x) \leq \overline{X} + \frac{z_{\frac{1+\gamma}{2}} \sigma(x)}{\sqrt{N}}\right)$$

nie znamy $\sigma(x)$ $\left(\sigma^2 \right)$

N - pomiarów

$$t = \frac{(\overline{X} - \mathbb{E}(x))\sqrt{N}}{s(x)}$$

$$s(x) = \frac{1}{N-1} \sum_{i=1}^{N} (X_i - \overline{X})^2$$

rozkład studenta o $N-1$ stopniach swobody

$$P\left(\overline{X} - \frac{1}{\sqrt{N}} S(x)\, t_{\frac{1+\gamma}{2}} \leq E(x) \leq \overline{X} + \frac{1}{\sqrt{N}} S(x)\, t_{\frac{1+\gamma}{2}}\right) = \gamma$$

$$\left[\, \overline{T}_N^{\,L}(E(x)) \,,\; \overline{T}_N^{\,R}(E(x)) \,\right]$$

$$\chi = \frac{(N-1)\, S(x)^2}{\sigma^2(x)} \qquad S = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})^2$$

ma FGD $\quad \chi^2_{\underbrace{N-1}}$

stopniach swobody

$T^L \qquad\qquad T^R$

$$P(\ldots < \chi < \ldots) = \gamma$$

$$P\left(\sqrt{\frac{(N-1)\, S^2}{\left(\chi^2_{N-1}\right)_{\frac{1+\gamma}{2}}}} \leq \sigma(x) \leq \sqrt{\frac{(N-1)\, S^2}{\left(\chi^2_{N-1}\right)_{\frac{1-\gamma}{2}}}} \right) = \gamma$$

#

X

$X : \Omega \longrightarrow \mathbb{R}$

$\mathbb{R}$ — liczby rzeczywiste

Ω — zbiór zdarzeń elementarnych

$X(⚁) = 2$

$\omega \in \Omega$

$X(t, \omega)$

t — czas