Metody Statystyczne

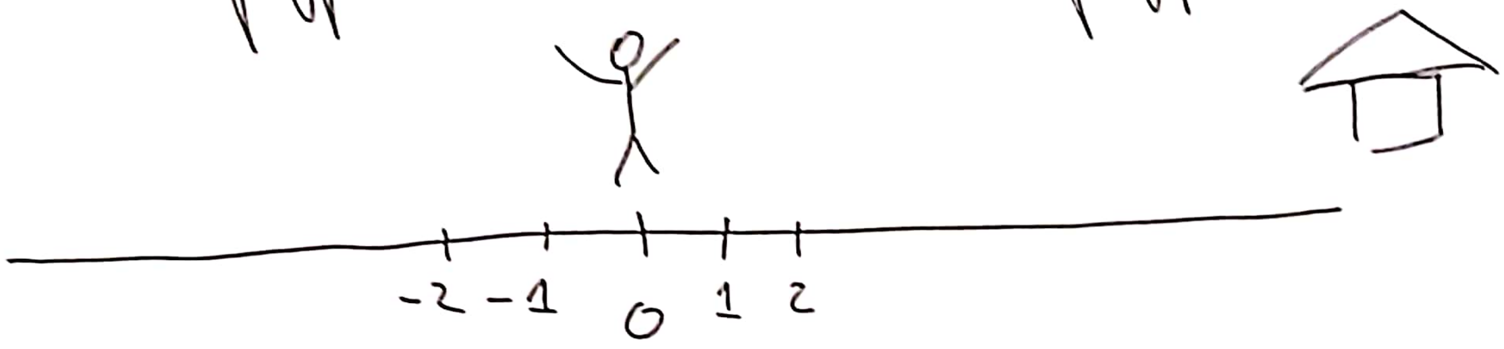
wykład 5

10 I 2021

Plan:

- przypomnienie / uzupełnienie
- procesy liczące
  - Bernoulli
  - Poisson

Błądzenie przypadkowe - Bernoulli - przypomnienie



prawdopodobieństwo sukcesu (krok w prawą stronę)

p

prawdopodobieństwo porażki (krok w lewą stronę)

$q = 1 - p$

funkcja autokorelacji:

$$\underbrace{R_{XX}}_{\text{pozycja „pijaczka”}}(\overbrace{m,n}^{\text{czas dyskretny}}) \equiv E[X_m \cdot X_n] =$$

$$= E[(X_m - X_n + X_n)\, X_n] =$$

$$= E[\underbrace{(X_m - X_n)\, X_n}_{\text{niezależne zmienne losowe}}] + E[X_n^2] = \dots$$

$$\ldots = \underbrace{E[X_m - X_n]\,E[X_n] - E[X_n^2]}_{\text{policzone na poprzednim wykładzie}} =$$

$$= \begin{cases} (p-q)^2\, m n + 4 m p q & m < n \\[2ex] (p-q)^2\, m n + 4 p q \min(m,n) & \text{ogólnie} \end{cases}$$

wartość najmniejsza

auto covariance function

$$K_{xx}(m,n) = R_{xx}(m,n) - \underbrace{\overset{E}{\mu_x(m)}\,\mu_x(n)}_{\text{wartości oczekiwane}}$$

współczynnik auto-korelacji

$$\rho_{xx}(m,n) = \frac{K_{xx}(m,n)}{\sigma[X_m]\,\sigma[X_n]}$$

procesy liczące

- proces gdzie zmienne losowe

$N(t)$

oznacza liczbę zliczeń

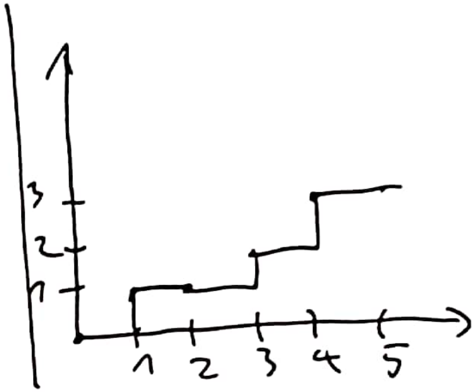
w czasie t

$N(t) \in \{0, 1, 2, 3, \dots\} = \mathbb{Z}^+$

– przykład – rzut monetą

H – reszka ← sukces

T – orzeł



$t =$	0	1	2	3	4	5	...
	T	H	T	H	H	T	
$N(t) =$	0	1	1	2	3	3	

sukcesy w próbie Bernulliego

n_1 – ~~su~~ l. sukcesów ← – wylosowanie H

sukces

n_2 – porażka – l. wylosowanych T

$q = 1 - p$

p

$\underbrace{n = n_1 + n_2}_{\text{całkowita l. rzutów monetą}}$

$\underbrace{k = n_1 - n_2}_{\text{pozycja}}$

$$P[N(n) = n_1] = \binom{n}{n_1} p^{n_1} q^{n_2}$$

$$P[N(h)] = \binom{h}{N(h)} p^{N(h)} (1-p)^{h-N(h)}$$

n. [illegible]

n. successi

1 rzut monetą co Δ sekund

$$t = n \cdot \Delta \qquad n = \frac{t}{\Delta}$$

„frame”

liczba prób (kroków, rzutów)

$$E\left[N\left(\frac{t}{\Delta}\right)\right] = \underbrace{\frac{t}{\Delta}}\, p$$

$$\lambda \equiv \frac{E\left[N\left(\frac{t}{\Delta}\right)\right]}{t} = \underbrace{\frac{p}{\Delta}}$$

„arrival rate”

T – czas pomiędzy kolejnymi zliczeniami

„interarival rate"

T H H T T T H

$T = \quad 1 \quad 4$

$E[T]$?

$$E(T) = E(j\Delta)$$

liczba krokow pomiedzy sukcesami

czas — " —

$j=1$

p

T	H	T	T	H	H	T	...
0	1	2	3	4	5	6	...

$P[j=1] = p$

$j=2$

q $\quad$ $j=1$ $\quad$ p

T	H	T	H	T	H	H	...
0	1	2	3	4	5	6	...

$j=2$

$P[j=2] = q^{2-1}p$

$$P[j] = q^{j-1}\, p = q^{j-1}(1-q)$$

$$E[j] = \sum_{j=0}^{\infty} j\, P[j] = \sum_{j} j\, q^{j-1}(1-q) =$$

$$= (1-q)\sum_{j=1}^{\infty} j\, q^{j-1} = (1-q)\sum_{j=1}^{\infty} \frac{d}{dq} q^{j} =$$

$$= (1-q)\frac{d}{dq}\overbrace{\sum_{j=1}^{\infty} q^{j}}^{\text{szereg geometryczny}} = \dots$$

$$\dots = (1-q)\underbrace{\frac{d}{dq}\,\frac{1}{1-q}} =$$

$$= (1-q)\,(1-q)^{-2}\,(-1)\,(-1) =$$

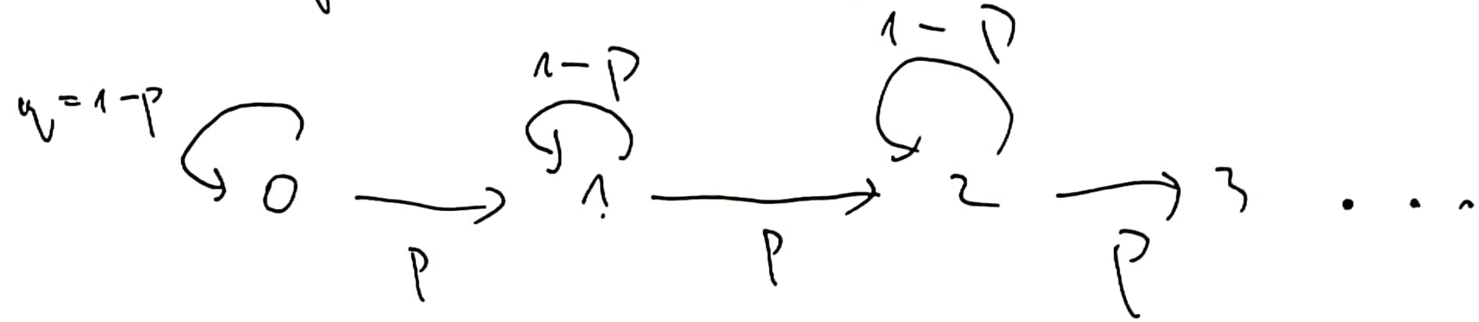
$$= (1-q)\,(1-q)^{-2} = \frac{1}{1-q} = \frac{1}{p} - E(j)$$

$$E[T] = E[j\,\Delta] = \frac{1}{p}\,\Delta = \frac{1}{\lambda}$$

$$\lambda = \frac{1}{E[T]}$$

$$\sigma^2(T) = \Delta^2 \, \frac{1-p}{p^2} = \frac{1-p}{\lambda^2}$$

graf dla procesu liczącego Bernoulliego



- nie periodyczny.

- „homogeneous" — macierz ~~przej~~ ~~przejś~~ prawdopodobieństw taka sama po każdym kroku czasowym.

macierz prawdopodobieństw

$$P = \begin{pmatrix} 1-p & p & 0 & 0 & \cdots \\ 0 & 1-p & p & 0 & \cdots \\ 0 & 0 & 1-p & p & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

$$(p_0^{(t)}, p_1^{(t)}, p_2^{(t)}, p_3^{(t)}, \ldots)$$

$$\uparrow \quad \uparrow$$

$w(t)$ $n(t)$ - - -

0

$P_{ij}^{(h)}$ – prawdopodobieństwo $i-j$ sukcesów w h-próbach

$$P_{ij}^{(n)} = \begin{cases} \binom{n}{j-i} p^{j-i} q^{n-j+i} & 0 \le j-i \le n \\ 0 & \text{pozostałe przypadki} \end{cases}$$

- proces Poissona

Bernulli

$$
\begin{tikzcd}
0 \arrow[loop above, "1-p"] \arrow[r, "p"] & 1 \arrow[loop above, "1-p"] \arrow[r, "p"] & 2 & \cdots
\end{tikzcd}
$$

$p \leftrightarrow$ prawdopodobieństwo sukcesu w czasie Δ

$q = 1 - p \leftrightarrow$ prawdopodobieństwo porażki w czasie Δ

ciągły czas

$\Delta \to 0$

proces Poissona

„małe" czasy h

$\to \lambda h \leftrightarrow$ prawdopodobieństwo sukcesu w h

$\to 1 - \lambda h \leftrightarrow$ prawdopodobieństwo porażki w h

proces Poissona jest granicznym przypadkiem procesu Bernulliego

P. P. z "arrival rate" λ to proces liczący $N(t)$ który: ...

$$\underbrace{\{0, 1, 2, \ldots\}}_{\mathbb{Z}^+}$$

a) $N(0) = 0$ $\quad N(t) \in \{0, 1, 2, \dots\} = \mathbb{Z}^+$

b) $h \to 0$

$$p_{mn}(h) = \begin{cases} 1 - \lambda h + \dots & n = m \\ \lambda h + \dots & n = m+1 \\ 0 + \dots & n \geq m+2 \end{cases}$$

c) $t > s$

$$\underbrace{N(t) - N(s) \qquad N(s)}_{\text{niezależne zmienne losowe}}$$

h – „mały" czasem

$$P[N(t)=n] = p_n(t) = ?$$

$$p_n(t+h) = p_n(t)\cdot p_{nn}(h) + p_{n-1}(t)\, p_{n-1\,n}(h) + \ldots =$$

$$= p_n(t)(1-\lambda h) + p_{n-1}(t)\,\lambda\cdot h + \ldots$$

$n=0$ $$p_0(t+h) = (1-\lambda\cdot h)\, p_0(t) + \ldots = \ldots$$

$$\boxed{p_{mn}(h) = P[N(t+h)=n \mid N(t)=m]}$$

$$p_0(t+h) - p_0(t) = (1-\lambda h)\, p_0(t) - p_0(t)$$

$$p_0(t+h) - p_0(t) = -\lambda h\, p_0(t) \quad / : h$$

$$\frac{p_0(t+h) - p_0(t)}{h} = -\lambda\, p_0(t)$$

$$\underbrace{\lim_{h \to 0} \frac{p_0(t+h) - p_0(t)}{h}}_{\text{definicja pochodnej}} = -\lambda\, p_0(t)$$

$$\boxed{\frac{d\, p_0(t)}{dt} = -\lambda\, p_0(t)}$$

$n \neq 0$

$$p_n(t+h) = p_n(t)(1-\lambda h) + p_{n-1}(t)\,\lambda h + \dots$$

$$p_n(t+h) - p_n(t) = p_n(t)(-\lambda h) + p_{n-1}(t)\,\lambda h + \dots$$

$$\underbrace{\lim_{h \to 0} \frac{p_n(t+h) - p_n(t)}{h}}_{\text{definicja pochodnej}} = \lambda\, p_{n-1}(t) - \lambda\, p_n(t)$$

$n = 1$

$$\frac{d\, p_1(t)}{dt} = \lambda \underbrace{p_0(t)}_{e^{-\lambda t}} - \lambda\, p_1(t)$$

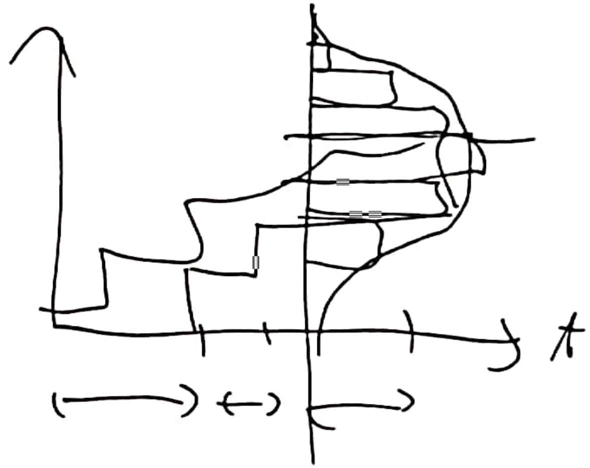
$$\frac{d\, p_1(t)}{dt} = \lambda\, e^{-\lambda t} - \lambda\, p_1(t)$$

$$p_1(t) = \lambda t\, e^{-\lambda t} + \underbrace{C_1}\, e^{-\lambda t} \qquad C_1 = 0$$

$$p_1(0) = C_1 \underbrace{e^{-\lambda \cdot 0}}_{1} = \underbrace{C_1 = 0}$$

$$p_1(t) = (\lambda t)\, e^{-\lambda t}$$

$$p_2(t) = \frac{(\lambda t)^2}{2!}\, e^{-\lambda t}$$



$$p_3(t),\ p_4(t),\ p_5(t),\ \ldots$$

$$p_n(t) = \frac{(\lambda t)^n}{n!}\, e^{-\lambda t} \qquad n = 0, 1, 2, 3, 4, \ldots$$

wprowadzić do oryginalnego równania i sprawdzić

$$\boxed{\frac{d\,p_n(t)}{dt} = \lambda\, p_{n-1}(t) - \lambda\, p_n(t)}$$

$$\boxed{\frac{d\,p_0(t)}{dt} = -\lambda\, p_0(t)}$$

$$N(0) = 0 \Rightarrow p_n(0) = \delta_{n0}$$

$$p_0(t) = \underbrace{C_0}_{(C_0 = 1)} e^{-\lambda t} \qquad p_0(0) = C_0 \underbrace{e^{-\lambda \cdot 0}}_{1} = C_0 = 1$$

Funkcja generująca (w skrócie)

$$G(z,t) \equiv \sum_{n=0}^{\infty} p_n(t)\, z^n$$

$$\frac{\partial}{\partial t} G(z,t) = \lambda (z-1)\, G(z,t)$$

$$G(z,0) = 0$$

$$G(z,t) = \underbrace{e^{\lambda (z-1) t}}$$

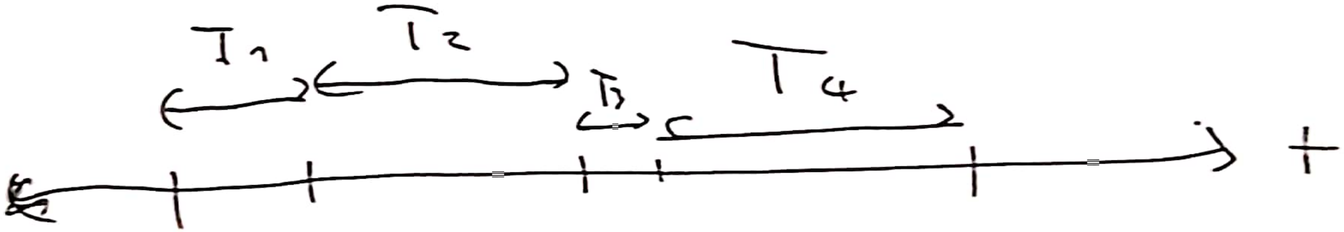
$$G(z,t) = e^{\lambda(z-1)t} = \underbrace{\sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda t} z^n}$$

$$p_n(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \qquad \sum_{n=0}^{\infty} p_n(t) z^n$$

„inter arrival times" – niezależne zmienne losowe

T

$$E[T] = \frac{1}{\lambda}$$



$T \longleftrightarrow X$ zmiana ~~oznaczeń~~ oznaczeń

$t = 0$ nie mamy zliczeń

$$P[X_1 > x] = P[N(x) = 0] =$$

$$= p_0(x) = e^{-\lambda t}$$

$t_1 = X_1$

$$P[X_2 > x \mid X_1 = t_1] =$$

$$= P[\underbrace{\text{brak zliczeń w } (t_1, t_1 + x) \mid X_1 = t_1}_{\text{niezależne zmienne losowe}}] = \ldots$$

$$\ldots = e^{-\lambda x}$$

$$\underbrace{P[X_n \le x]}_{\text{dystrybuanta}} = 1 - e^{-\lambda x}$$

$$\int_0^x f(x)\,dx$$

$$f(x) = \frac{d}{dx}\left(1 - e^{-\lambda x}\right) = \lambda e^{-\lambda x}$$

$$E[x] = \int_0^\infty x f(x) = \frac{1}{\lambda}$$