Metody Statystyczne

20 XII 2020

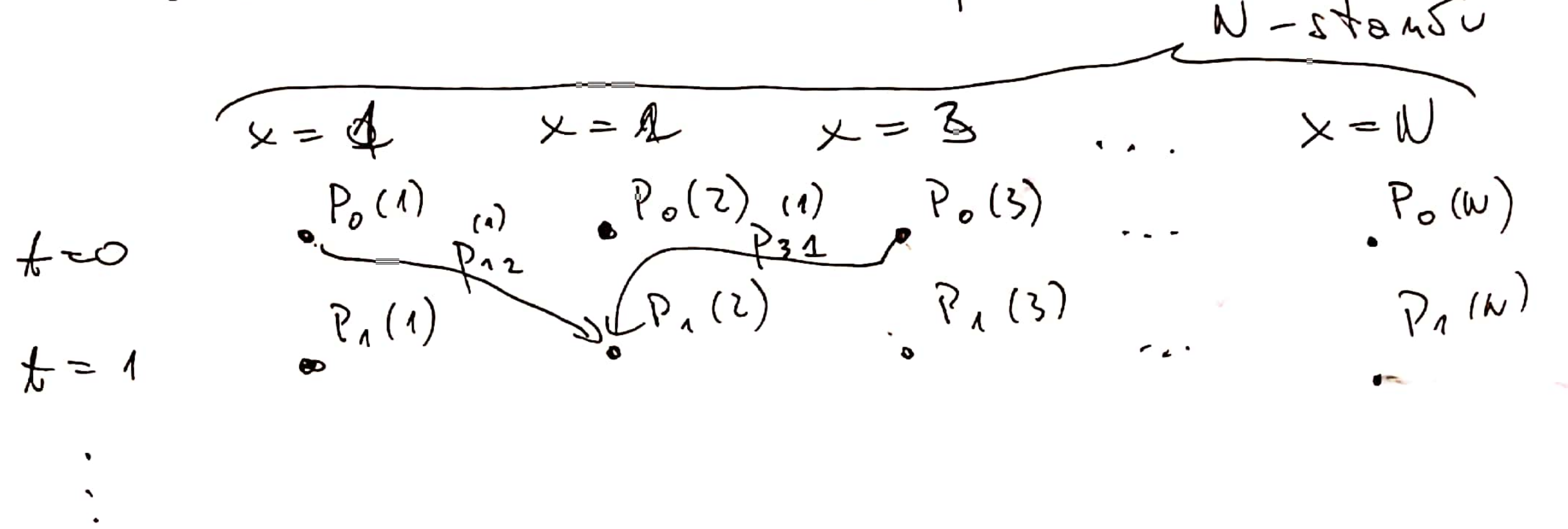
kacper.topolnicki@uj.edu.pl

Plan:

- powtórka
- Łańcuchy Markowa:
  - [illegible]
  - stany stacjonarne
- Bernulli

[stanach stacjonarnych

– na początek: Łańcuchy Markova ze skończoną liczbą stanów



- stan w chwili t:

$$\vec{S}_t = (P_t(1), P_t(2), \dots, P_t(N))$$

- macierz prawdopodobieństw

$$P^{(1)} = \begin{pmatrix} p_{11}^{(1)} & p_{12}^{(1)} & \cdots \\ p_{21}^{(1)} & p_{22}^{(1)} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

macierz $N \times N$

$$S_{t+1} = S_t \, P^{(1)}$$

– ile wynosi $P^{(2)}$

$$P^{(2)} = P^{(1)} \cdot P^{(1)} = \left(P^{(1)}\right)^2 .$$

$$P^{(n)} = \left(P^{(1)}\right)^n$$

$$\underbrace{P^{(m+n)}_{ik} :=}_{\text{definicja}} P(X_{\underbrace{m+n}_{\text{czas}}} = k \mid X_{\underset{\text{czas}}{0}} = i) =$$

zmienna losowa

wartość zmiennej losowej

$$= \sum_j P(X_{m+n} = k \wedge X_n = j \mid X_0 = i) =$$

wynika z własciwości zbioru

definicja prawdopodobieństwa warunkowego

$$= \sum_j \frac{P(X_{m+n} = k \wedge X_n = j \wedge X_0 = i)}{P(X_0 = i)} =$$

$$= \sum_j \frac{P(X_{m+n}=k \wedge (X_m=j \wedge X_0=i))}{P(X_m=j \wedge X_0=i)} \cdot \frac{P(X_m=j \wedge X_0=i)}{P(X_0=i)} =$$

$$= \sum_j P(X_{m+n}=k \mid X_m=j \wedge X_0=i)\, P(X_m=j \wedge X_0=i) =$$

$$= \sum_j P(X_{m+n}=k \mid X_m=j)\, P(X_m=j \wedge X_0=i)$$

$$p_{ik}^{(m+n)} = \sum_j p_{jk}^{(n)}\, p_{ij}^{(m)} =$$

$$= \underbrace{\sum_j p_{ij}^{(m)}\, p_{jk}^{(n)}}_{\text{mnożenie macierzy}} \qquad P^{(m+n)} = P^{(m)} \cdot P^{(n)}$$

[[stan stacjonarny

$$\overline{\pi} = (\pi(1), \pi(2), \dots, \pi(n))$$

def:

$$\pi P^{(1)} = \overline{\pi}$$

$$\pi P^{(n)} = \overline{\pi}$$

$\overbrace{\text{równanie}}^{\text{eigen equation}}$

$$A\vec{x} = \lambda\vec{x}$$

$$xA = \lambda x$$

$$\Uparrow$$

$$1$$

– kiedy istnieje?

- ile wynosi $p^{(n)} = \left(p^{(1)}\right)^n = \underbrace{p^{(1)} \cdot p^{(1)} \cdot p^{(1)} \cdot \ldots \cdot p^{(1)}}_{n}$

$$\Big\updownarrow \; n \to \infty$$

$$\left(p(1), p(2), p(3), \ldots \right) \cdot \begin{pmatrix} \pi(1) & \pi(2) & \pi(3) & \cdots \\ \pi(1) & \pi(2) & \pi(3) & \cdots \\ \pi(1) & \pi(2) & \pi(3) & \cdots \\ \vdots & \vdots & & \end{pmatrix} =$$

$$= \begin{pmatrix} p(1)\,\underline{\pi(1)} + p(2)\,\underline{\pi(1)} + p(3)\,\underline{\pi(1)} + \ldots, \\ p(1)\,\pi(2) + p(2)\,\pi(2) + p(3)\,\pi(3) + \ldots, \\ \ldots \end{pmatrix} =$$

$$= \Big(\pi(1) \underbrace{\big(p(1) + p(2) + \dots \big)}_{1},\ \pi(2) \underbrace{\big(p(1) + p(2) + \dots \big)}_{1},\ \dots \Big) =$$

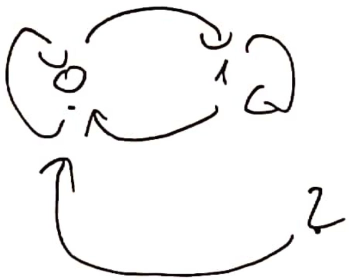
czy tak jest zawsze?

$$= \underbrace{\big(\pi(1), \pi(2), \pi(3), \dots \big)}$$

startując z dowolnego stanu łańcucha Markowa (wektora prawdopodobieństw), jeżeli n „duże" to będziemy w stanie stacjonarnym

Rodzaje stanów (wartości zmiennych losowych)

- stan i jest <u>dostępny</u> ze stanu j jeżeli



$$p_{ij}^{(n)} > 0 \text{ dla pewnego } n > 0$$

- stany i oraz j się <u>komunikują</u>

gdy i jest dostępny z j oraz j jest dostępny z i

b Klasa komunikacji

- dwa stany i, j należą do tej samej klasy komunikacji jeżeli się komunikują

{0, 1} zamknięta   {2} nie-zamknięte   {3} zamknięte

- Łańcuch Markowa jest nierozkładalny jeżeli wszystkie stany komunikują

- Klasa komunikacji jest zamknięta jeżeli nie da się jej opuścić

...

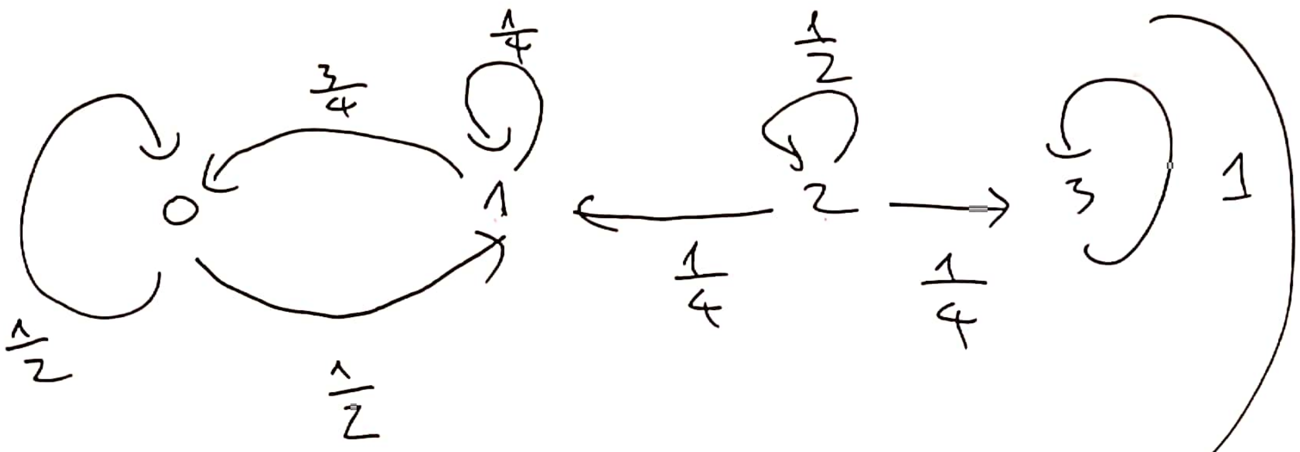
... klasa C jest zamknięta gdy

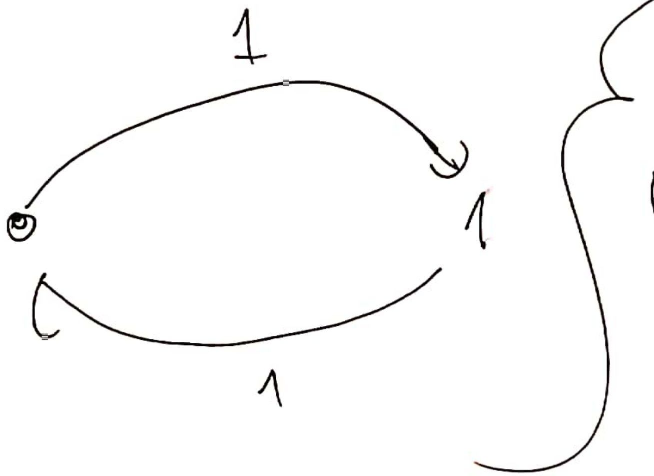
$$p_{ij}=0 \text{ jeżeli } i\in C \text{ oraz } j\notin C$$

– periodyczność stanów oraz klas

$$\begin{tikzcd} 0 \arrow[r, bend left, "1"] & 1 \arrow[l, bend left, "1"] \end{tikzcd}$$

$$P^{(1)}=\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$





dziwoląg!

prosza z.
lamchodzi!
notebook.

$$P^{(1)} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, P^{(2)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, P^{(3)} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, P^{(4)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \dots$$

$$p_{11}^{(1)} = 0 \quad , p_{11}^{(2)} = 1 \quad , p_{11}^{(3)} = 0 \quad , p_{11}^{(4)} = 1, \dots$$

<u>periodyczność stanu i:</u>

$$d_i = \gcd\{n > 0 : p_{ii}^{(n)} > 0\}$$

$i = 1 \qquad d_1 = \gcd\{2, 4, 6, 8, \dots\} = 2$

– klasa periodyczna, a-periodyczna

[] Tw. stan stacjonarny nieredukowalnego łańcucha Markova ma dodatnie prawdopodobieństwa

szkic dowodu:

niech

$$\underbrace{\pi = (\pi(1), \pi(2), \dots)}_{\text{stan stacjonarny}}$$

wybieramy taki stan a:

$$P(X_0 = a) = \pi(a) > 0$$

dla stanu b z nieredukowalności wynika, że istnieje

$$n > 0 : P(X_n = b \mid X_0 = a) > 0$$

$$\pi(b) = P(X_n = b) \geq P(X_n = b \wedge X_0 = a) =$$

$$= P(X_0 = a) \underbrace{P(X_n = b \mid X_0 = a)}_{> 0} > 0$$

$$\pi(b) > 0$$

II Tw. Nieredukowalny łańcuch Markowa ma co najwyżej jeden stan stacjonarny

szkic dowodu:

$$\pi_1 P^{(1)} = \pi_1 \qquad \pi_2 P^{(1)} = \pi_2$$

$$\text{z} \quad \alpha = a \cdot \pi_1 \qquad \beta = b \cdot \pi_2$$

$$\underline{\alpha} \cdot P^{(1)} = a\,\pi_1 \cdot P^{(1)} = a\,(\underbrace{\pi_1 \cdot P^{(1)}}_{\pi_1}) = a \cdot \pi_1 = \underline{\alpha}$$

dobieramy wartość a, b:

$$\cancel{x} \cdot \alpha - \beta \geq 0$$

$$\alpha(i) = \beta(i) \qquad \alpha_i = \beta_i$$

$\uparrow$

pewna wartość X

$\uparrow$

czy możemy tak zawsze zrobić

$$\eta = \alpha - \beta \qquad \eta_i = 0$$

$$\eta(i)$$

$$\underbrace{\eta P^{(1)} = \eta}_{\parallel}$$

$$\eta P^{(1)} = (a\pi_1 - b \cdot \pi_2) P^{(1)} = a(\pi_1 P^{(1)}) + b(\pi_2 P^{(1)}) =$$

$$= a \cdot \pi_1 + b \cdot \pi_2 = \eta$$

$$\eta_i = \eta(i) - 0 - \sum_{j \neq i} \eta_j \, p_{ji}^{(n)} = \sum_{j \neq i} \eta_j \, p_{ji}^{(1)}$$

nieredukowalny łańcuch Markowa:

istnieje n^i : $p_{ji}^{(n^i)} > 0$

$$0 = \sum_{j \neq i} \eta_j \, p_{ji}^{(n_i)}$$

$$\Downarrow$$

$$\eta_j = 0$$

unikalny

$\alpha = \beta$ - stan stacjonarny

$\pi_1 = \pi_2$

Sposoby znajdowania stanów stacjonarnych

- przeprowadzenie bezpośredniego eksperymentu
- rozwiązanie r. własnego

$$\underbrace{P^{(1)}\,\overline{\pi} = 1\,\overline{\pi}}_{\text{typowo}} \quad \longleftrightarrow \quad \overline{\pi}\, P^{(1)} = \lambda\, \overline{\pi}$$

λ: 1

$P^{(1)}$: macierz przej.

$\overline{\pi}$: stan stacjonarny (wektor)

– liczenie $\left(P^{(1)}\right)^n \quad n \to \infty$

$$\begin{pmatrix} \pi(1) & \pi(2) & \pi(3) & \cdots \\ \pi(1) & \pi(2) & \pi(3) & \cdots \\ \pi(1) & \pi(2) & \pi(3) & \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

– metoda funkcji generującej

$$f = \{ \underbrace{f_0, f_1, f_2, \dots}_{\text{wektor}} \}$$

$$F(z) = \sum_{k=0}^{\boxed{\infty}} f_k z^k$$

funkcja generująca

$$\lim_{z \to 1} (1 - z) F(z) = \lim_{k \to \infty} f_k$$

Final value theorem

$$g(z) = \sum_{n=0}^{\infty} s^{(n)} z^{n}$$

$$\left(s^{(n)}(1),\ s^{(n)}(2),\ \dots \right)$$

$$g(z)\, P = \sum_{n=0}^{\infty} s^{(n+1)} z^{n} =$$

$$= z^{-1} \sum_{n=0}^{\infty} s^{(n+1)} z^{n+1} = z^{-1} g(z) - z^{-1} s^{(0)}$$

$$g(z) = \underbrace{s^{(0)}}_{\text{wektory}} \overbrace{\left(\underbrace{\mathbb{1}}_{\text{macierz jednostkowa}} - P z \right)}^{\text{macierz}}{}^{-1} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

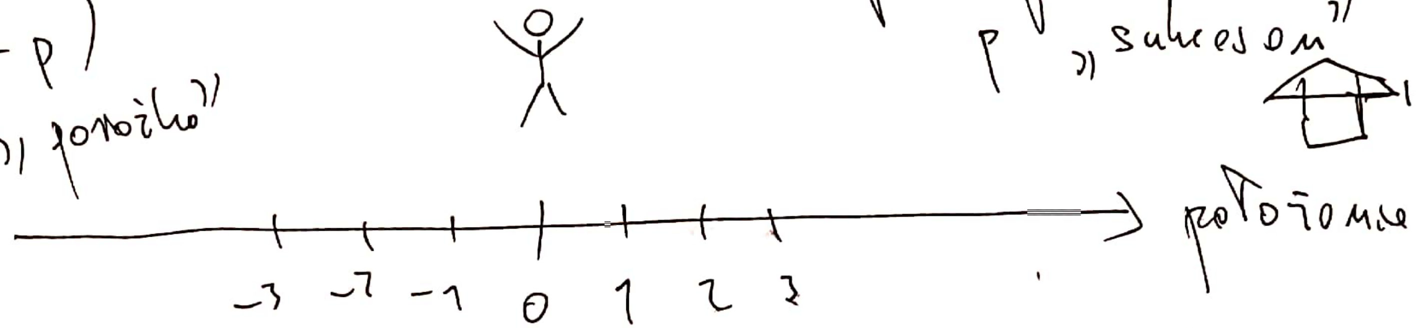
Final value theorem

$$\overbrace{\underbrace{\lim_{n \to \infty} s^{(n)}}_{\text{stato } n \text{ stazionario}} = \lim_{z \to 1} (1-z)\, \underline{g(z)}}$$

I próba Bernoulliego i błądzenie przypadkowe

krok w lewo z prawdopodobieństwem ← $q = (1-p)$ „porażka”

→ krok w prawo z prawdopodobieństwem p „sukces”



n_1 – l. kroków w prawo

n_2 = l. kroków w lewo

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

$k = n_1 - n_2$ ← pozycja

symbol Newtona:

$$\underbrace{\binom{n}{k} = \frac{n!}{k!(n-k)!}}$$

na ile sposobów można wybrać podzbiór k-elementowy ze zbioru n-elementowego

$$\overbrace{\{1\ -1\ 1\ 1\ -1\ 1\ 1\ \dots\}}$$

$$P(X_n = k) = \underbrace{\binom{n}{n_1}}_{\{1\ 1\ 1\ 1\ 1\ 1\ \dots\}} p^{n_1} q^{n_2} =$$

$$= \left| \begin{array}{l} n_1 = \frac{n+k}{2} \\ \\ n_2 = \frac{n-k}{2} \end{array} \right| = \binom{n}{\frac{n+k}{2}} p^{\frac{n+k}{2}} (1-p)^{\frac{n-k}{2}} = \dots$$

$$= \frac{n!}{\left(\frac{n+k}{2}\right)!\left(n-\frac{n-k}{2}\right)!}\, p^{\frac{n+k}{2}} (1-p)^{\frac{n-k}{2}} =$$

$$= \left| p = \frac{1}{2} \right| = \frac{n!}{2^n} \left(\left(\frac{n+k}{2}\right)! \left(\frac{n-k}{2}\right)! \right) =$$

$$= P(X_n = k)$$

[jednowymiarowe błądzenie przypadkowe

$$S_i = \begin{cases} +1 & \text{z prawd. } p \\ -1 & \text{z prawd. } q = 1-p \end{cases}$$

$$X_n = X_0 + \underbrace{\sum_{i=1}^{n} S_i}_{S_1 + S_2 + \dots}$$

[wartość oczekiwana

$$E[X_n] = E[X_0] + \sum_{i=1}^{n} E[S_i] = \dots$$

wartość oczekiw.

$$= \left| X_0 - 0 \right| = \sum_{i=1}^{n} E[S_i] = n \cdot E[S] =$$

$$= n(p-q)$$

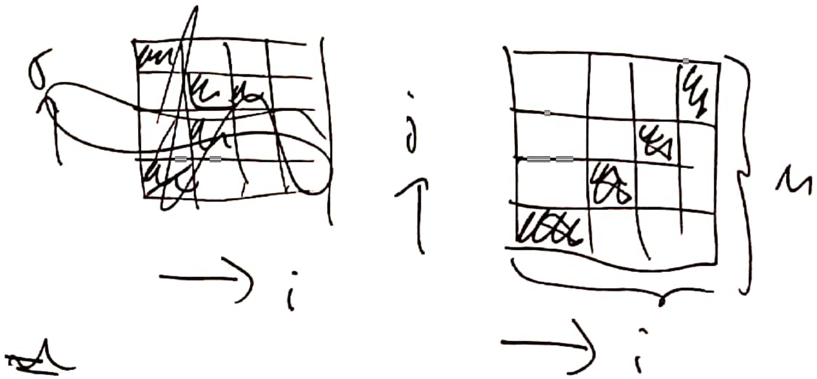
$$E[X_n] \overset{p=q=\frac{1}{2}}{=} 0$$

$$E[X_n^2] = E\left[\left(\sum_{i=1}^{n} S_i\right)\left(\sum_{i=1}^{n} S_i\right)\right] =$$

$$= E\left[\sum_{i=1}^{n} S_i^2 + \sum_{\substack{i,j \\ i \neq j}} S_i S_j\right] = \dots$$

niezależne zmienne losowe

$$\ldots = E\Big[\underbrace{\sum_{i=1}^{n} \underbrace{S_i^2}_{+1}}_{\underbrace{n\cdot(+1)}_{n}}\Big] + \underbrace{E\Big[\sum_{\substack{i,j \\ i\neq j}} S_i\, S_j\Big]}_{(n^2-n)\, E[S_i]\, E[S_j]} =$$



$$= \underbrace{n + (n^2-n)\,(p-q)^2}$$

$$\operatorname{var}[X_n] = \underbrace{\mathbb{E}[X_n^2]} - \underbrace{\mathbb{E}[X_n]^2} =$$

$$= 4 \cdot n \cdot p \cdot \underset{\substack{\parallel \\ 1-p}}{q}$$