Metody statystyczne

Zestaw zadań numer 2

Vitalii Urbanevych

vitalii.urbanevych@doctoral.uj.edu.pl

21.11.2021

- 2 użytkowników
- 1 komputer
- Do komputera może być zalogowanych
  - $x = 0$ użytkowników
  - $x = 1$ użytkownik
  - $x = 2$ użytkowniki

Niezalogowani

Prawdopodobieństwo zalogowania:

$$P_{zalogowania} = 0.2$$

Prawdopodobieństwo pozostania niezalogowanym:

$$1 - P_{zalogowania} = 0.8$$

Zalogowani

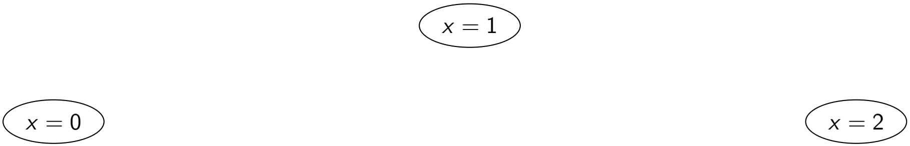
Prawdopodobieństwo wylogowania:

$$P_{wylogowania} = 0.5$$

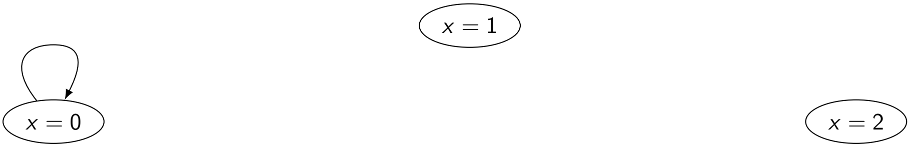
Prawdopodobieństwo pozostania zalogowanym:

$$1 - P_{wylogowania} = 0.5$$

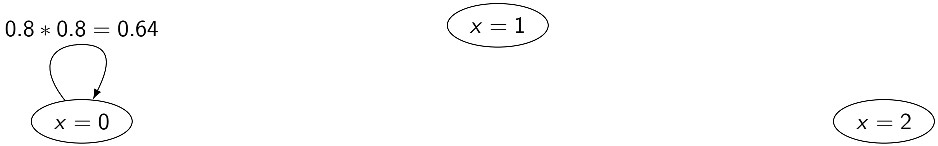
$$P_{zalogowania} = 0.2; \quad P_{wylogowania} = 0.5$$



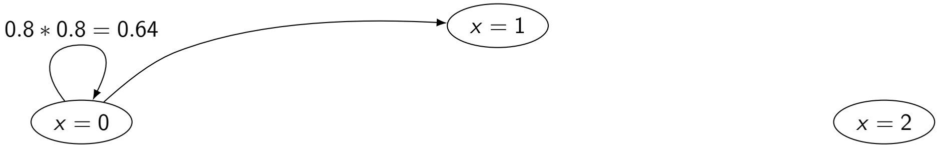
$$P_{zalogowania} = 0.2; \quad P_{wylogowania} = 0.5$$



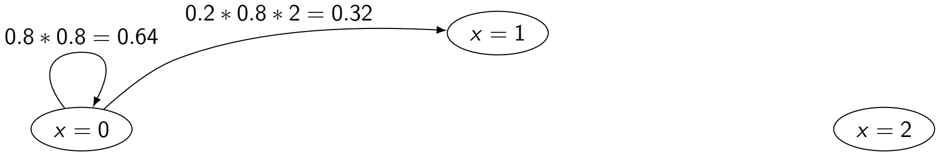
$$P_{zalogowania} = 0.2; \quad P_{wylogowania} = 0.5$$



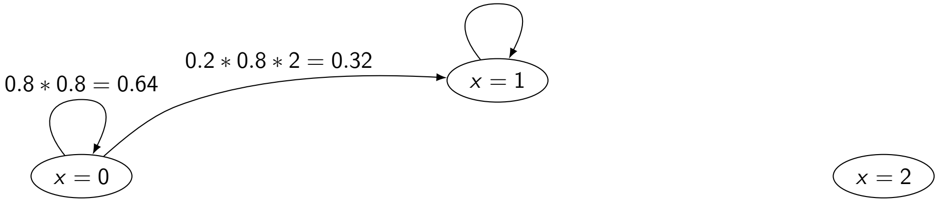
$$P_{zalogowania} = 0.2; \quad P_{wylogowania} = 0.5$$



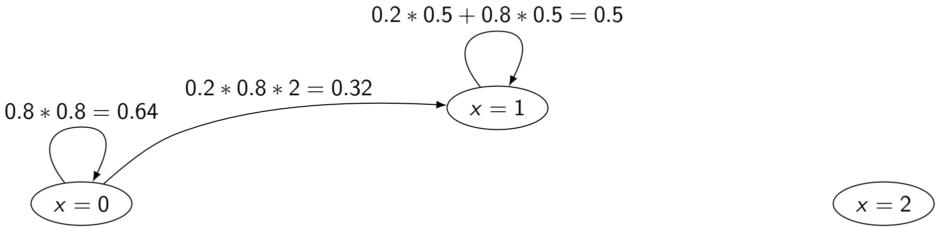
$$P_{zalogowania} = 0.2; \quad P_{wylogowania} = 0.5$$



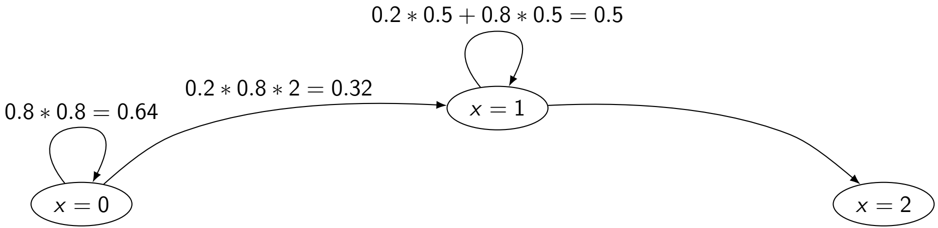
$$P_{zalogowania} = 0.2; \quad P_{wylogowania} = 0.5$$



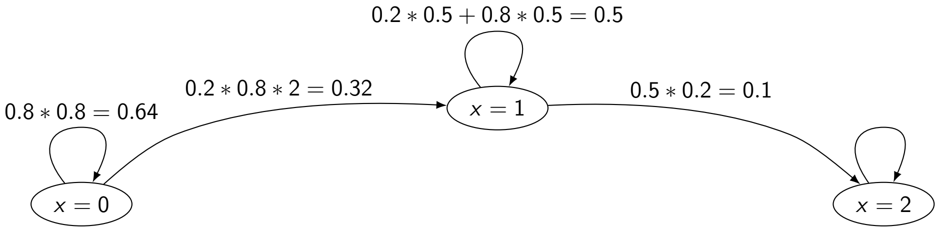
$$P_{zalogowania} = 0.2; \quad P_{wylogowania} = 0.5$$



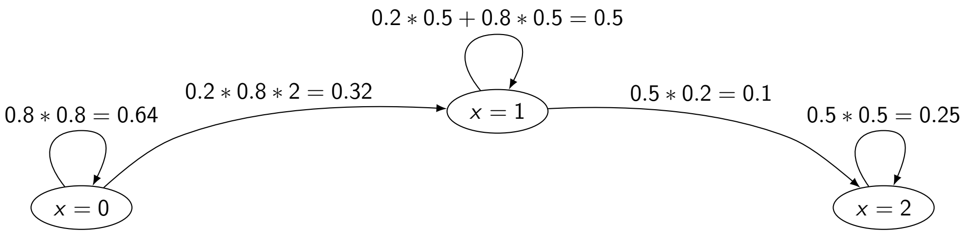
$$P_{zalogowania} = 0.2; \quad P_{wylogowania} = 0.5$$



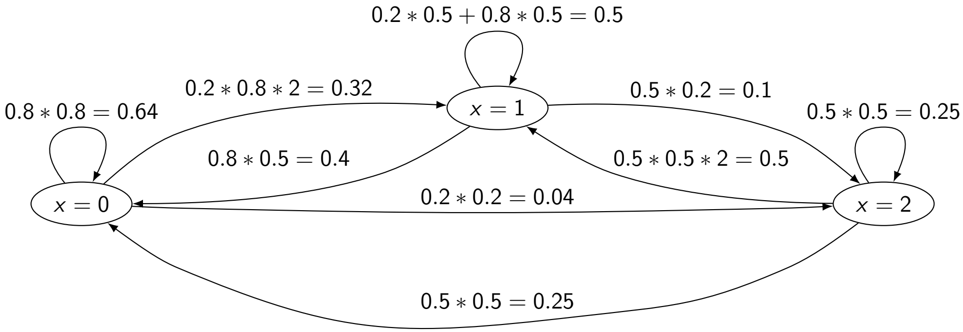
$$P_{zalogowania} = 0.2; \quad P_{wylogowania} = 0.5$$

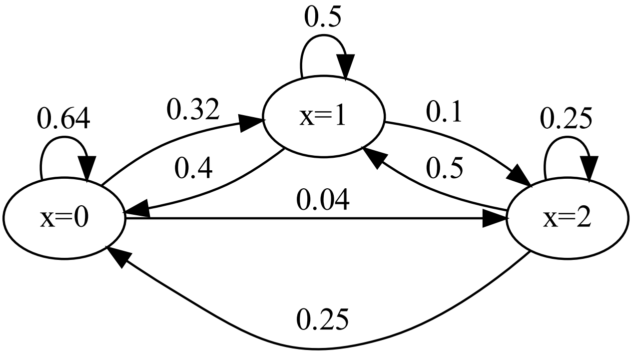


$$P_{zalogowania} = 0.2; \quad P_{wylogowania} = 0.5$$



$P_{zalogowania} = 0.2; \quad P_{wylogowania} = 0.5$





$P_{zalogowania} = 0.2$
$P_{wylogowania} = 0.5$

Macierz przejść

$$P = \begin{array}{c c} & \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \end{array} & \begin{pmatrix} 0.64 & 0.32 & 0.04 \\ 0.4 & 0.5 & 0.1 \\ 0.25 & 0.5 & 0.25 \end{pmatrix} \end{array}$$

prawdopodobieństwo ucieczki z $x = 0 \rightarrow 0.64+0.32+0.04=1$

Macierz przejść

$$P = \begin{array}{c c} & \begin{array}{ccc} \color{red}{0} & \color{red}{1} & \color{red}{2} \end{array} \\ \begin{array}{c} \color{blue}{0} \\ \color{blue}{1} \\ \color{blue}{2} \end{array} & \begin{pmatrix} 0.64 & 0.32 & 0.04 \\ 0.4 & 0.5 & 0.1 \\ 0.25 & 0.5 & 0.25 \end{pmatrix} \end{array}$$

prawdopodobieństwo przejścia z $x = \{0, 1, 2\}$ do $x = 2$

"Stanem" nazywamy taki vector $p = \begin{pmatrix} p_0 & p_1 & p_2 \end{pmatrix}$ w którym układ jest z prawdopodobieństwem:

$$\begin{matrix} p_0 & w & x = 0 \\ p_1 & w & x = 1 \\ p_2 & w & x = 2 \end{matrix}$$

Jak policzyć odpowiednie prawdopodobieństwa dla stanu po 1 iteracji?

$$\begin{pmatrix} p_0 & p_1 & p_2 \end{pmatrix} \begin{pmatrix} 0.64 & 0.32 & 0.04 \\ 0.4 & 0.5 & 0.1 \\ 0.25 & 0.5 & 0.25 \end{pmatrix} = pP = p\prime$$

$$p\prime = \begin{pmatrix} p_0\prime & p_1\prime & p_2\prime \end{pmatrix} \text{ - nowy stan}$$

Macirz prawdopodobieństw po dwoch iteracjach

Po dwóch iteracjach:

$$p'' = p' \cdot P = (p \cdot P) \cdot P = p \cdot (P \cdot P)$$

$$P \cdot P = \begin{pmatrix} 0.64 & 0.32 & 0.04 \\ 0.4 & 0.5 & 0.1 \\ 0.25 & 0.5 & 0.25 \end{pmatrix} \begin{pmatrix} 0.64 & 0.32 & 0.04 \\ 0.4 & 0.5 & 0.1 \\ 0.25 & 0.5 & 0.25 \end{pmatrix} =$$

Macirz prawdopodobieństw po dwoch iteracjach

Po dwóch iteracjach:

$$p\prime\prime = p\prime \cdot P = (p \cdot P) \cdot P = p \cdot (P \cdot P)$$

$$P \cdot P = \begin{pmatrix} 0.64 & 0.32 & 0.04 \\ 0.4 & 0.5 & 0.1 \\ 0.25 & 0.5 & 0.25 \end{pmatrix} \begin{pmatrix} 0.64 & 0.32 & 0.04 \\ 0.4 & 0.5 & 0.1 \\ 0.25 & 0.5 & 0.25 \end{pmatrix} =$$

Macirz prawdopodobieństw po dwoch iteracjach

Po dwóch iteracjach:

$$p\prime\prime = p\prime \cdot P = (p \cdot P) \cdot P = p \cdot (P \cdot P)$$

$$P \cdot P = \begin{pmatrix} 0.64 & 0.32 & 0.04 \\ 0.4 & 0.5 & 0.1 \\ 0.25 & 0.5 & 0.25 \end{pmatrix} \begin{pmatrix} 0.64 & 0.32 & 0.04 \\ 0.4 & 0.5 & 0.1 \\ 0.25 & 0.5 & 0.25 \end{pmatrix} =$$

$$= \begin{pmatrix} 0.64 \times 0.64 + 0.32 \times 0.4 + 0.04 \times 0.25 = 0.5476 \\ \\ \\ \end{pmatrix}$$

Macirz prawdopodobieństw po dwoch iteracjach

Po dwóch iteracjach:

$$p\prime\prime = p\prime \cdot P = (p \cdot P) \cdot P = p \cdot (P \cdot P)$$

$$P \cdot P = \begin{pmatrix} 0.64 & 0.32 & 0.04 \\ 0.4 & 0.5 & 0.1 \\ 0.25 & 0.5 & 0.25 \end{pmatrix} \begin{pmatrix} 0.64 & 0.32 & 0.04 \\ 0.4 & 0.5 & 0.1 \\ 0.25 & 0.5 & 0.25 \end{pmatrix} =$$

$$= \begin{pmatrix} 0.5476 & 0.3848 \\ & \\ & \end{pmatrix}$$

Macirz prawdopodobieństw po dwoch iteracjach

Po dwóch iteracjach:

$$p\prime\prime = p\prime \cdot P = (p \cdot P) \cdot P = p \cdot (P \cdot P)$$

$$P \cdot P = \begin{pmatrix} 0.64 & 0.32 & 0.04 \\ 0.4 & 0.5 & 0.1 \\ 0.25 & 0.5 & 0.25 \end{pmatrix} \begin{pmatrix} 0.64 & 0.32 & 0.04 \\ 0.4 & 0.5 & 0.1 \\ 0.25 & 0.5 & 0.25 \end{pmatrix} =$$

$$= \begin{pmatrix} 0.5476 & 0.3848 & 0.0676 \\ & & \\ & & \end{pmatrix}$$

Macirz prawdopodobieństw po dwoch iteracjach

Po dwóch iteracjach:

$$p\prime\prime = p\prime \cdot P = (p \cdot P) \cdot P = p \cdot (P \cdot P)$$

$$P \cdot P = \begin{pmatrix} 0.64 & 0.32 & 0.04 \\ 0.4 & 0.5 & 0.1 \\ 0.25 & 0.5 & 0.25 \end{pmatrix} \begin{pmatrix} 0.64 & 0.32 & 0.04 \\ 0.4 & 0.5 & 0.1 \\ 0.25 & 0.5 & 0.25 \end{pmatrix} =$$

$$= \begin{pmatrix} 0.5476 & 0.3848 & 0.0676 \\ 0.481 & 0.428 & 0.091 \\ 0.4225 & 0.455 & 0.1225 \end{pmatrix}$$

Stan stacjonarny

- Co się stanie po $N \to \infty$ iterations?
- Spodziewamy się osiągnąć stan stacjonarny
- Jak w takim przypadku będzie wyglądała macierz prawdopodobieństw?

$$P^N = \underbrace{P \cdot P \cdots P}_{N}$$

$$N \to \infty : P^N \to \begin{bmatrix} \pi_0 & \pi_1 & \pi_2 \\ \pi_0 & \pi_1 & \pi_2 \\ \pi_0 & \pi_1 & \pi_2 \end{bmatrix}$$

- Właściwość: $P^{N+1} = P^N$, przy $N \to \infty$ (kolejna iteracja nie zmienia stany)

Nie zawsze istnieje taka macierz P^N

Stan stacjonarny

$$\begin{bmatrix} p_0 & p_1 & p_2 \end{bmatrix} \begin{bmatrix} \pi_0 & \pi_1 & \pi_2 \\ \pi_0 & \pi_1 & \pi_2 \\ \pi_0 & \pi_1 & \pi_2 \end{bmatrix} = \begin{bmatrix} \pi_0(p_0 + p_1 + p_2) & \pi_1(p_0 + p_1 + p_2) & \pi_2(p_0 + p_1 + p_2) \end{bmatrix} =$$

$$= \begin{bmatrix} \pi_0 & \pi_1 & \pi_2 \end{bmatrix} \leftarrow \text{stan stacjonarny}$$

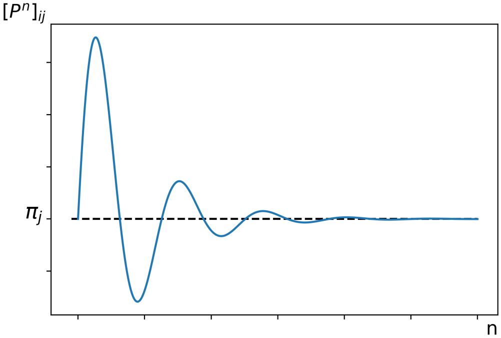
Stan stacjonarny

Taki stan Π, kiedy kolejna iteracja nie zmienia stanu:

$$\Pi \cdot P = \Pi$$

Problem A

- Policzyć P^N dla dużych N
- Kryterium zbieżności: $|P^n - P^{n-1}| < \epsilon$
- $\epsilon = 10^{-5}$ lub inne wartości
- Narysować wykres $[P^n]_{ij}$ w zależności od n
- Porównanie z wartościami π_j na wykresie



Problem B

- Symulacja
  1. Zaczynamy z wybranego węzła $x = \{0, 1, 2\}$
  2. Losowanie kolejnego węzła zgonie z P
  3. Przejście do losowanego węzła
  4. Wracamy do 2
- Losujemy $N = 10^4$ przejść
- Obliczenie $\pi_i^{exp} = \frac{N_i}{N}$, N_i - ile razy odwiędzone $x = i(0, 1, 2)$
- Porównanie z P^N
- Start z różnych węzłów

Problem C

- 100 użytkowników
- $x = 0, 1, 2, \cdots, 100$
- $P_{logowania} = 0.2$, $P_{wylogowania} = 0.5$
- Trudno jest obliczyć macierz prawdopodobieństw
- Wykonujemy symulację trajektorii
- Ile wynosi π_i^{exp} dla $i = 0, 1, \cdots, 100$?
- Wykresy zbieżności $\pi_i^{exp} = \frac{N_i}{N}$ jako zależność od N dla kilku i (np dla pięciu największych wartości π)
- Wykres końcowych wartości dla wszystkich π

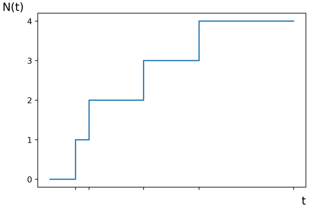
Problem D

Podobna symulacja jak w C, tylko

- $P_{logowania} = 0.2$
- $P_{wylogowania} = 1 - (0.008 \cdot x + 0.1)$

Process Poissona

- t_i - czas pomiędzy zdarzeniami $(i-1)$ i i,
  $t_0 = 0$
  $i = 1, 2, \ldots$
- t_i jest losowane z rozkładu wykładniczego, $f(t) = \lambda e^{-\lambda t}$
  - losujemy t_i metodą odwróconej dystrybuanty
  - $t_i = \frac{-\ln(n_i)}{\lambda}$
  - n_i losowane z rozkładu jednorodnego na przedziale (0,1)
- $N(t)$ - ilość zdarzeń, które wystąpiły do chwili t
  $N(0) = 0$
- Taki proces nazywa się procesem Poissona o intensywności λ
- $N(t)$ ma rozkład Poissona $P(N(t) = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$, z parametrem λt



Problem E

Symulacja procesu Poissona

- $\lambda = 1$
- $t = 1, 10, 20, 90$
- Zaimplementować symulację pojawienia zdarzeń
- Dla każdej wartości t:
  - Otrzymać rozkład prawdopodobieństwa ilości zdarzeń
  - Porównać z rozkłądem Poissona z parametrem λt
- Sprawdzić że wartość średnia jest λt

Problem F

- Mamy symulację zdarzeń jak w E
- $\lambda = 1,\ t = 1, 10, 20, 90$
- Każde zdarzenie może należeć do jednej z trzech grup: 1,2,3
- Należność do jednej z grup jest losowane i prawdopodobieństwa należności do grup: $p_1 = 0.2,\ p_2 = 0.5,\ p_3 = 0.3\ (p_1 + p_2 + p_3 = 1)$
- Sprawdzić że rozkład prawdopodobieństwa zdarzeń grupy i jest rozkładem Poissona z parametrem $\lambda t p_i$
- Sprawdzić że wartość średnia dla takiego rozkładu jest $\lambda t p_i$